Pontifícia Universidade Católica
do Rio de Janeiro

Carolina da Silva Bolognani

Search for CP violation in the decays $D_s^+ \to K^-K^+K^+$ and $D^+ \to K^-K^+K^+$ in the LHCb experiment

Dissertação de Mestrado

Dissertation presented to the Programa de Pós–graduação em Física of PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Física.

Advisor: Prof. Carla Göbel Burlamaqui de Mello

Rio de Janeiro
August 2021

Pontifícia Universidade Católica
do Rio de Janeiro

Carolina da Silva Bolognani

Search for CP violation in the decays $D_s^+ \to K^-K^+K^+$ and $D^+ \to K^-K^+K^+$ in the LHCb experiment

Dissertation presented to the Programa de Pós–graduação em Física of PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Física. Approved by the Examination Committee:

Prof. Carla Göbel Burlamaqui de Mello
Advisor
Departamento de Física – PUC-Rio

Profa. Irina Nasteva
UFRJ

Prof. Tobias Frederico
ITA

Rio de Janeiro, August 5th, 2021

Carolina da Silva Bolognani

The author obtained the degree of Bachelor in Physics in 2019 from Universidade Federal do Rio de Janeiro (UFRJ).

Bibliographic data

Bolognani, Carolina

Search for CP violation in the decays $D_s^+ \to K^-K^+K^+$ and $D^+ \to K^-K^+K^+$ in the LHCb experiment / Carolina da Silva Bolognani; advisor: Carla Göbel Burlamaqui de Mello. – 2021.

111 f: il. color. ; 30 cm

Dissertação (mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Física, 2021.

Inclui bibliografia

1. Física – Teses. 2. Charme. 3. Violação de CP. 4. LHCb. 5. Decaimentos. 6. Física de Partículas. I. Göbel, Carla. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Física. III. Título.

CDD: 620.11

I dedicate this work to all those who
believed in me and offered their support.

Acknowledgments

This study was financed by CNPq and FAPERJ. To these institutions, my grateful appreciation. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) -Finance Code 001.

I would like to first thank my advisor Carla Göbel for guiding me through this time, introducing me to the world of research and sharing with me the same enthusiasm for physics. I would, however, also like to apologise for possibly annoying her with the dozens of sudden messages and plots I would send her out of the blue, whenever I "started thinking of things". Being your student was truly the best choice I could have made.

To the Rio Charm Group, I express my gratitude for the lengthy weekly discussions that made this work possible. To Alberto dos Reis, Sandra Amato and specially Erica Polycarpo, who has pushed me forward since I first started my undergrad studies. Also, to the many students of the group, Fernanda, Lucas, Felipe and Juan, for the occasional independent meetings where we helped each other in the most playful of tones.

I would also like to thank those that, before the pandemic, made going to PUC-Rio such a great experience. In particular, the administrative staff Gizelda Dias da Silva, Eduardo da Silva Rodrigues, Marcia Arjona and Julinho, the excellent professors, and amazing friends I made. To these friends, thanks for all the light-hearted lunches and beer-filled evenings at Garota da Gávea.

Additional thanks go to the physicist friends and mentors that supported and encouraged me through this time from outside PUC-Rio. In particular, Prof. Eduardo Fraga, and my undergrad friends, the Victors.

To my family, my utmost respect and gratefulness. My brother Henrique, who has been my best friend from the moment he was born, my parents Ana and Fábio, who have always cheered me on and been proud of my path, and my grandparents Othon and Margarida, who have inspired me to be who I am today.

Then I wish to thank my many friends who have supported me even from afar. To Marcella, my first friend ever, the sister life gave me. To my friends from "the End of the World", who have always put up with my silly takes, shared my love for fictional characters and historical figures alike, pushed me to be a great person in the weirdest way possible, and have appreciated how I genuinely found physics and mathematics to be fun and let me talk about it even during recess period in school. To the best friends life could have

unexpectedly given me, Luz and Rick, with whom I laugh, vent, gossip, do random Buzzfeed quizes, watch high-school drama movies and find comfort.

Last, but definitely not least, my gratitude for Pedro. You were my partner through all of this, and will be for much more to come.

Abstract

Bolognani, Carolina; Göbel, Carla (Advisor). **Search for CP violation in the decays $D_s^+ \to K^-K^+K^+$ and $D^+ \to K^-K^+K^+$ in the LHCb experiment**. Rio de Janeiro, 2021. 111p. Dissertação de Mestrado – Departamento de Física, Pontifícia Universidade Católica do Rio de Janeiro.

The current best theory for describing elementary particles and their interactions, the Standard Model, has charge-parity (CP) violation as a fundamental feature, which implies that particles and antiparticles behave differently. Due to CP violation, the decay of a meson into a certain final state might then be distinct from the decay of the anti-meson to that charge conjugated final state. This effect has been extensively studied and measured over the years in the decays of K and B mesons, which involve the *strange* and *beauty* quarks respectively, but only recently observed in D meson decays, involving the *charm* quark. The predicted CP asymmetries in the charm sector are very small, so, in order to make new observations, large data samples are required.

This dissertation reports the search for CP violation in the singly-Cabibbo-suppressed $D_s^+ \to K^-K^+K^+$ and doubly-Cabibbo-suppressed $D^+ \to K^-K^+K^+$ decays. The search for local asymmetries in the phase space of the decays is performed with a model-independent technique, using the data samples obtained in pp collisions with a centre-of-mass energy of 13 TeV, collected by the LHCb detector during 2016—2018 (run 2). Also included is a study of the method's sensitivity to CP violation effects in the current statistics using simulations. In this analysis, two Cabibbo-favoured decays, $D_s^+ \to K^-K^+\pi^+$ and $D^+ \to K^-\pi^+\pi^+$, are used to ensure we are not sensitive to nuisance asymmetries, from production or detection effects.

Keywords

Charm; CP Violation; LHCb; Decays; Particle Physics.

Resumo

Bolognani, Carolina; Göbel, Carla. **Busca por violação de CP nos decaimentos $D_s^+ \to K^-K^+K^+$ e $D^+ \to K^-K^+K^+$ no experimento LHCb**. Rio de Janeiro, 2021. 111p. Dissertação de Mestrado – Departamento de Física, Pontifícia Universidade Católica do Rio de Janeiro.

A melhor teoria atual para descrever partículas elementares e suas interações, o Modelo Padrão, tem como peça fundamental a violação de carga-paridade (CP), a qual implica diferença no comportamento de partículas e anti-partículas. Por conta da violação de CP, o decaimento de um méson em um certo estado final pode então ser distinto do decaimento do anti-méson nesse estado final conjugado. Esse efeito foi extensivamente estudado e medido ao longo dos anos em decaimentos de mésons K e B, que envolvem os quarks *strange* e *beauty* respectivamente, mas apenas recentemente observado em decaimentos de mésons D, envolvendo o quark *charm*. As assimetrias de CP previstas no setor charmoso são bastante pequenas, então para que novas observações sejam feitas, grandes conjuntos de dados são necessários.

Esta dissertação reporta a busca por violação de CP nos decaimentos $D_s^+ \to K^-K^+K^+$, suprimido por Cabibbo, e $D^+ \to K^-K^+K^+$, duplamente suprimido por Cabibbo. A busca por assimetrias locais nos espaços de fase dos decaimentos é realizada com uma técnica modelo-independente, usando dados obtidos em colisões *pp* de energia de centro-de-massa de 13 TeV, coletados pelo detector LHCb durante 2016—2018 (*run 2*). Também incluído está um estudo da sensibilidade do método a efeitos de violação de CP dadas as estatísticas atuais usando simulações. Nesta análise, dois decaimentos favorecidos por Cabibbo, $D_s^+ \to K^-K^+\pi^+$ e $D^+ \to K^-\pi^+\pi^+$, são utilizados para nos assegurar que não somos sensíveis a assimetrias espúrias, de efeitos de produção ou detecção.

Palavras-chave

Charme; Violação de CP; LHCb; Decaimentos; Física de Partículas.

Table of contents

List of figures

List of tables

If the Lord can lead you to it, He can lead you through it.

Trixie Mattel, *UNHhhh.*

1
Introduction

The study of symmetries and their violations in nature has always been of great importance and guided our understanding of the most elementary building blocks constituting the universe. Two important discrete symmetries in particle physics are charge conjugation and parity, which, respectively, relate particles with their antiparticles and imply invariance under spacial reflection. At first, all fundamental interactions were thought to preserve these symmetries, and although the strong and electromagnetic forces do, the weak interaction was found to violate separately both of them [1–4]. Their combination, however, was expected to be conserved. And so, since its first observation in the $K^0 - \bar{K}^0$ mixing of a neutral kaon system in 1964 [5], charge-parity (CP) violation has puzzled the particle physics community. The impact of the discovery was such that it then led Kobayashi and Maskawa, in 1973 [6], to predict the existence of new quarks in order to theoretically explain the phenomenon within the framework of the Standard Model (SM), which describes our best understanding so far of the fundamental interactions.

Since then, a lot of effort has been put into measuring new CP violating effects. In 1999, the first observation of *direct* CP violation, i.e. in the decays and not related to mixing of K mesons was made by the NA48 collaboration [7], and, since 2001, with the advent of "B-factory" experiments, many observations of CP violation in B mesons started being made [8, 9]. For the charm sector, however, only one observation of CP violation has been made so far; in 2019, a difference of CP violation asymmetries between the decays of a D^0 meson to the final states K^-K^+ and $\pi^-\pi^+$ was measured by the LHCb collaboration [10].

The SM prediction for CP violation in decays of charm hadrons is quite small, so as time goes on and more data are collected, the physics community hopes to make new observations of CP violation effects in D meson systems. This dissertation describes the analysis performed to search for these effects in two charm-meson decays, $D_s^+ \to K^-K^+K^+$ and $D^+ \to K^-K^+K^+$,[1] using data collected by the LHCb experiment in 2016—2018. These are three-body decays where we can look for the effects of local CP violation asymmetries in their phase spaces, which are potentially larger effects than the total (phase

[1] Charge conjugate decays are implicit, otherwise mentioned explicitly.

space integrated) asymmetries. These two decays are interesting for different reasons; the D_s^+ decay is a Cabibbo-suppressed decay, and, as explained later, there is a SM prediction for CP violation in it; on the other hand, the D^+ decay is doubly-Cabibbo suppressed, so the SM predicts no CP violation should be seen in it, and therefore any observation would imply New Physics beyond the SM. Given the similarities of both decays due to them having the same final state, the analysis is done in parallel. Furthermore, no CP violation analysis of either decay has been published in the literature before.

The remainder of this dissertation is organised as follows. In Chapter 2, a review of the SM is presented. The necessary fundamentals for understanding CP violation, specifically in 3-body decays, is discussed, with a focus on charm decays.

In Chapter 3, an overview of the LHCb experiment is given. Chapter 4 describes the selection criteria applied to the LHCb data used. The selection process includes requirements used to remove specific background contributions as well as a multi-variate analysis performed with machine learning to reduce combinatorial background levels.

Chapter 5 contains the analysis of the decay channels. First, the technique for performing the search for asymmetries in a model-independent way, called Mirandizing, is described. Then, given the statistics of the final data sample, studies using simulated distributions of the decays are performed to evaluate the technique sensitivity to CP violating effects. Following this, systematic checks are performed to validate the search methods as well as verify there is no sensitivity to nuisance asymmetries from the background of the signal channels as well as production or detection asymmetries, analysed with the control channels $D_s^+ \to K^-K^+\pi^+$ and $D^+ \to K^-\pi^+\pi^+$. Finally, the prospects for the analysis flow are described. This analysis is performed *blinded*, which means that the actual measurements to perform the CP violation search are not made until the analysis is approved by the internal review committee within the LHCb Collaboration.

The conclusions of this dissertation are presented in Chapter 6. Additional information pertaining to the analysis is included in the Appendices.

2
Theoretical Fundamentals

2.1
The Standard Model

The Standard Model of particle physics is the best theory so far for describing the interactions between the known elementary particles. These particles, which are divided between fermions (half-integer spins) and bosons (integer spins), have different charges to quantify how they behave according to each one of the fundamental interactions, be it Electromagnetic, Weak or Strong interactions (Gravity is not included in the SM) [11]. Amongst the fermions, which compose matter, there are six leptons and six quarks, and we refer to the different species of particles of the same type as being different flavours, which are organised in three families. For each matter particle there is its antiparticle, which has the same mass but opposite charges (and all additive quantum numbers, such as lepton or baryon number). As for the bosons, there are four types of gauge bosons, i.e. responsible for mediating interactions, as well as the Higgs boson. A summary of these particles can be found in Fig. 2.1.[1]

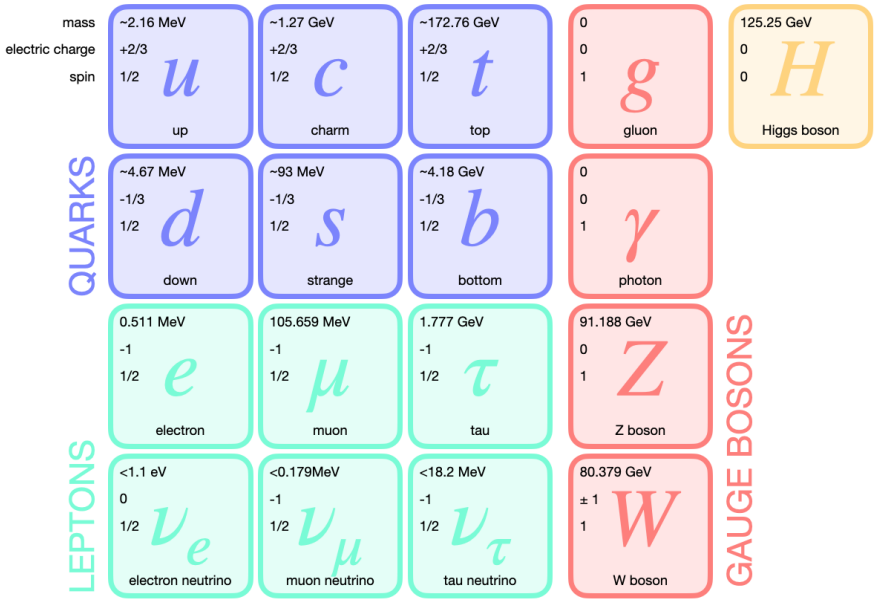


Figure 2.1: Elementary particles of the Standard Model [12].

[1]Natural units, $\hbar = c = 1$, are used throughout this work.

Formally, we say that the SM is a quantum field theory based on the gauge group $SU(3)_C \times SU(2)_L \times U(1)_Y$. This means that it is a theory described in a framework which combines quantum mechanics with special relativity, with a certain group structure that describes the local symmetries between particles. In these types of theories, we have that the fermions are the fundamental representation of the symmetry group, while the gauge bosons are the adjoint representation.

The first part of the SM, $SU(3)_C$, relates to Quantum Chromodynamics (QCD), which describes the strong interactions between colour-charged quarks and the eight massless force-carrying gluons. Due to the non-Abelian (non-commutative) structure of the $SU(3)$ symmetry, a property of QCD is asymptotic freedom, which implies that at higher energies scales (lower distances) the couplings are weaker. An observed feature of QCD is colour confinement, i.e. quarks and gluons cannot propagate freely and must be parts of composed structures, called hadrons. These hadrons are typically baryons (constituted of three quarks or three antiquarks) or mesons (constituted of a quark and an antiquark). Recently, new types of structures which are theoretically also allowed have also been observed, such as tetraquarks and pentaquarks.

The second part of the SM, $SU(2)_L \times U(1)_Y$, relates to the Glashow-Salam-Weinberg model, which takes into account the fact that the SM is chiral, i.e. distinguishes between left-hand and right-hand chirality of the particles, and introduces three massless gauge bosons which couple only to left-hand states of the fermions and one massless gauge boson with a coupling via hypercharge (Y).

We can write the gauge part of the SM Lagrangian by (see for instance [13, 14])

$$\begin{aligned}\mathcal{L}^{\text{SM}}_{\text{gauge}} = &\sum_{i=1\ldots3} \sum_{\psi=Q^i_L \ldots E^i_R} \bar{\psi} i \not{D} \psi \\ &- \frac{1}{4} \sum_{a=1\ldots8} G^a_{\mu\nu} G^{\mu\nu,a} - \frac{1}{4} \sum_{a=1\ldots3} W^a_{\mu\nu} W^{\mu\nu,a} - \frac{1}{4} B^a_{\mu\nu} B^{\mu\nu,a} \,, \end{aligned} \tag{2-1}$$

where ψ represents the fermions, $G^a_{\mu\nu}$, $W^a_{\mu\nu}$ and $B_{\mu\nu}$ denote the field strength tensors of the gluons, $SU(2)_L$ bosons and $U(1)_Y$ boson, respectively. The D_μ ($\not{D} = \gamma^\mu D_\mu$) represents the covariant derivative, which leads to both the kinetic terms for fermions as well as their interactions with the many gauge boson fields. To illustrate, for a certain gauge boson field A_μ, we may write the field strength tensor as

$$F_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f^{abc} A^b_\mu A^c_\nu \,, \tag{2-2}$$

where g is the coupling parameter and f^{abc} is the structure constant of the symmetry group with generators T^a, $[T^a, T^b] = if^{abc}T^c$ (for abelian groups we then have $f^{abc} = 0$). The last term of Eq. 2-2 implies, in case the symmetry group is non-abelian, self-interactions of the gauge boson. The covariant derivative is written as

$$D_\mu = \partial_\mu - igT^a A^a_\mu \, . \tag{2-3}$$

The last term of Eq. 2-3 indicates how the fermion-boson interaction proceeds. The fermion content is given by the left-hand doublets and right-hand singlets which, in the notation $\psi(A, B)_Y$, where A is the representation under $SU(3)_C$, B under $SU(2)_L$ and Y the hypercharge, are

$$Q^i_L(3,2)_{+1/6} \, , \; U^i_R(3,1)_{+2/3} \, , \; D^i_R(3,1)_{-1/3} \, , \; L^i_L(1,2)_{-1/2} \, , \; E^i_R(1,1)_{-1} \, , \tag{2-4}$$

where i runs over the 3 different families, $Q_L = \binom{U}{D}_L$, $L_L = \binom{\nu}{E}_L$ with U being the up-type quarks, $U^{\text{tr}} = (u, c, t)$, D the down-type quarks, $D^{\text{tr}} = (d, s, b)$, and E the charged leptons, $E^{\text{tr}} = (e, \mu, \tau)$. The hypercharge Y relates to the electric charge Q and the $SU(2)$ generator $T^3 = \sigma^3/2$ (here, σ represents the corresponding Pauli matrix) by $Q = T^3 + Y$. For singlets under $SU(2)_L$, i.e. right-handed particles, we have $T^3 = 0$, and so the electric charge and hypercharge coincide.

There is an additional dynamic part of the SM missing from $\mathcal{L}^{\text{SM}}_{\text{gauge}}$, which is the Higgs sector. When adding a $SU(2)_L$ doublet of scalar fields, ϕ, we find that the $SU(2)_L \times U(1)_Y$ symmetry is spontaneously broken when the Higgs field acquires a vacuum expectation value ($v \approx 246$ GeV), and four resulting bosons arise from the mixture of the three massless $SU(2)_L$ bosons and the massless $U(1)_Y$ boson. These resulting bosons are the known massive $W^\pm$ and Z^0 weak interaction bosons and the massless photon of electromagnetic interactions. With these new bosons, we may express the electroweak Lagrangian in terms of charged currents $J^\pm_\mu$, weak neutral currents J^0_μ and the electromagnetic currents J^{em}_μ,

$$\begin{aligned}
J^+_\mu &= \bar{U}_L \gamma_\mu D_L + \bar{E}_L \gamma_\mu \nu_L \\
J^-_\mu &= \bar{D}_L \gamma_\mu U_L + \bar{\nu}_L \gamma_\mu E_L \\
J^0_\mu &= \frac{1}{2}\left(\bar{U}_L \gamma_\mu U_L - \bar{D}_L \gamma_\mu D_L + \bar{\nu}_L \gamma_\mu \nu_L - \bar{E}_L \gamma_\mu E_L\right) \\
J^{\text{em}}_\mu &= \frac{2}{3}\bar{U}_{L,R} \gamma_\mu U_{L,R} - \frac{1}{3}\bar{D}_{L,R} \gamma_\mu D_{L,R} - \bar{E}_{L,R} \gamma_\mu E_{L,R} \, ,
\end{aligned} \tag{2-5}$$

in a charged part $\mathcal{L}_{CC}$ and a neutral part $\mathcal{L}_{NC}$,

$$\begin{aligned}\mathcal{L}_{CC} &= \frac{g}{\sqrt{2}}\left(J^{+}_{\mu}W^{-,\mu} + J^{-}_{\mu}W^{+,\mu}\right)\\ \mathcal{L}_{NC} &= eJ^{\mathrm{em}}_{\mu}A^{\mu} + \frac{g}{\cos\theta_W}\left(J^{0}_{\mu} - \sin^2\theta_W J^{\mathrm{em}}_{\mu}\right)Z^{\mu}\ ,\end{aligned} \tag{2-6}$$

where $W^{\pm,\mu}$ and Z^{μ} represent the weak boson fields, A^{μ} the photon field, g the coupling of weak interactions and e of the electromagnetic interactions, and θ_W the Weinberg angle.

Other than introducing spontaneous symmetry breaking, the Higgs also explicitly breaks the flavour symmetry via its interactions, which are of a Yukawa type, with the fermion fields. Once the Higgs acquires its vacuum expectation value, fermion masses arise:

$$-\mathcal{L}^{\mathrm{SM}}_{\mathrm{Yukawa}} = Y^{ij}_{d}\bar{Q}^{i}_{L}\phi D^{j}_{R} + Y^{ij}_{u}\bar{Q}^{i}_{L}\tilde{\phi}U^{j}_{R} + Y^{ij}_{e}\bar{L}^{i}_{L}\phi E^{j}_{R} + \mathrm{h.c.}\ \ (\tilde{\phi} = iT^2\phi^{\dagger})\ , \tag{2-7}$$

and the fermion mass matrices become $\mathcal{M} = vY$.

2.1.1 CP violation in the Standard Model

So far, CP violation has not been mentioned. This effect comes about in the weak interactions through the mass matrices $\mathcal{M}_U$ and $\mathcal{M}_D$ [15–17].[2] Defining four unitary matrices such that these mass matrices may be diagonalised,

$$V_{U,L}\mathcal{M}_U V^{\dagger}_{U,R} = \mathcal{M}^{\mathrm{diag}}_{U}\ \ ,\ \ V_{D,L}\mathcal{M}_D V^{\dagger}_{D,R} = \mathcal{M}^{\mathrm{diag}}_{D}\ , \tag{2-8}$$

we may rotate the quark fields to their mass eigenstates

$$U^{m}_{L[R]} = V_{U,L[R]}U_{L[R]}\ \ ,\ \ D^{m}_{L[R]} = V_{D,L[R]}D_{L[R]}\ . \tag{2-9}$$

In shifting the quark fields in the SM to their mass eigenstates, we see no effect in the functional form of the QCD Lagrangian or the neutral part of the electroweak Lagrangian since there is no quark flavour mixture. In the charged part $\mathcal{L}_{CC}$, however, we get for instance

$$\bar{U}_L\gamma_\mu D_L \ \rightarrow\ \bar{U}^{m}_{L}V_{U,L}\gamma_\mu V^{\dagger}_{D,L}D_L \ =\ V_{U,L}V^{\dagger}_{D,L}\bar{U}^{m}_{L}\gamma_\mu D_L\ . \tag{2-10}$$

Therefore, we see that a matrix which defines different amplitudes in the mixing of quark flavours arises in the charged sector of the weak interactions.

[2] CP violation is theoretically also possible in QCD, but there is no experimental evidence for it. This is known as the strong-CP problem [18], and will not be discussed here.

This matrix is called the Cabibbo-Kobayashi-Maskawa (CKM) matrix:

$$V_{CKM} = V_{U,L}V_{D,L}^{\dagger} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} . \tag{2-11}$$

In the SM, V_{CKM} is unitary but not necessarily real. If the matrix is complex, we may have that $V_{ij} \neq V_{ij}^{*}$, and so the amplitude for transitioning from a certain quark flavour to another differs from the amplitude for this transition for antiquarks. This is the mechanism through which CP violation can appear in the SM.

2.1.2 The CKM matrix

Before the third family of quarks was known, Cabibbo had introduced a mixing angle so that the mixing of quark flavours would preserve the universality of weak interactions [19].[3] This angle, the Cabibbo-angle θ_c, was such that it related the weak interaction eigenstates d', s' to the mass eigenstates

$$\begin{aligned} d' &= d\cos\theta_c + s\sin\theta_c \\ s' &= -d\sin\theta_c + s\cos\theta_c \, . \end{aligned} \tag{2-12}$$

In this 2×2 matrix, there are only real parameters, and no CP violation can arise. Kobayashi and Maskawa proposed, then, a third generation of quarks, seeing that 3×3 unitary matrices are minimally parametrised by four free parameters, three real angles and one complex phase,

$$\mathrm{V}_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix} , \tag{2-13}$$

where $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$ for the θ_{ij} Euler angles between the i, j families, and δ_{13} the complex phase.

From the unitarity relation $V_{CKM}V_{CKM}^{\dagger} = \mathbb{I}$ two types of conditions must

[3] Cabibbo's paper preceded the quark model, and therefore the derivation followed by comparison of rates of K's to π's decays to leptons.

be satisfied,

$$\sum_{i=1}^{3} |V_{ij}|^2 = 1 \quad j = 1 \dots 3$$
$$\sum_{i=1}^{3} V_{ji} V_{ki}^* = \sum_{i=1}^{3} V_{ij} V_{ik}^* = 0 \quad j, k = 1 \dots 3, \;\; j \neq k \; . \tag{2-14}$$

The first type of condition implies weak universality, in other words, the couplings of an up-type quark to all down-type quarks have in total a universal strength for all generations, and vice-versa. This is related to the fact that the $SU(2)$ doublets are coupled to the weak bosons with the same strength.

There are other possible parametrisations for the CKM matrix, the most used being the Wolfenstein parametrisation. In it, we write the matrix elements as expansions of $\lambda = \sin\theta_c \approx |V_{us}| \approx 0.22$, and instead of the three angles and one complex phase, the elements are given in terms of four real parameters λ, A, ρ, η.

$$\mathrm{V}_{CKM} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 - \frac{1}{8}\lambda^4 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda + \frac{1}{2}A^2\lambda^5[1 - 2(\rho + i\eta)] & 1 - \frac{1}{2}\lambda^2 - \frac{1}{8}\lambda^4(1 + 4A^2) & A\lambda^2 \\ A\lambda^3(1 - \bar{\rho} - i\bar{\eta}) & -A\lambda^2 + \frac{1}{2}A\lambda^4[1 - 2(\rho + i\eta)] & 1 - \frac{1}{2}A^2\lambda^4 \end{pmatrix} + \mathcal{O}(\lambda^6). \tag{2-15}$$

In the Wolfenstein parametrisation, it is easy to see how the magnitudes of the different matrix elements compare to each other. First, we can see that the matrix is in modulus almost symmetric and diagonal. The elements of the in-family transitions, V_{ud}, V_{cs}, V_{tb} are close to 1. As the elements get farther away from the diagonal, they get smaller. Another observation to be made is that the CP violating phases arise first at $\mathcal{O}(\lambda^3)$ in the elements V_{ub} and V_{td}, and for charm transitions the phase comes up with a further λ^2 suppression in V_{cd}. The reason for this almost diagonal and symmetric nature of the CKM matrix, which ultimately comes from understanding the quark masses, still remains a puzzle in flavour physics.

A good representation of the CP violating phases is done through what are called unitarity triangles. These triangles are defined in the complex plane from the null-relations of Eq. 2-14. The most common parametrisation is given in terms of the most relevant relation for B meson decays,

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0$$
$$\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} + 1 + \frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} = 0 \; , \tag{2-16}$$

defining a triangle, seen in Fig. 2.2, where the apex is given by

$$\bar{\rho} + i\bar{\eta} = 1 + \frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} = -\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} \,, \tag{2-17}$$

and the angles are defined as

$$\alpha \equiv \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right) \;\;,\;\; \beta \equiv \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right) \;\;,\;\; \gamma \equiv \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right) \,. \tag{2-18}$$

With these, to $\mathcal{O}(\lambda^3)$, it is a good approximation to assume real matrix elements, with the exceptions $V_{td} = |V_{td}|e^{-i\beta}$ and $V_{ub} = |V_{ub}|e^{-i\gamma}$.

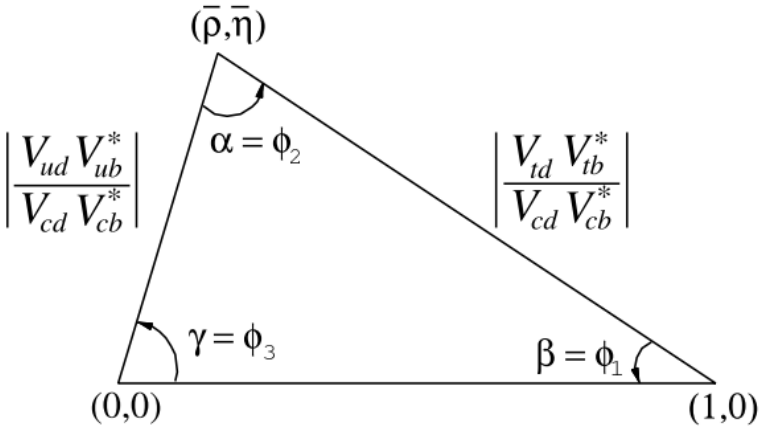


Figure 2.2: Sketch of unitarity triangle [12].

In non-leptonic weak quark transitions, the amplitude is proportional to the square of the product of the two matrix elements involved. For example, we may consider the transitions $c \to s\bar{d}u$, $c \to s\bar{s}u$ and $c \to d\bar{s}u$, depicted in Fig. 2.3 at tree-level. For charm decays, i.e. concerning the 2×2 submatrix, it is useful to name these transitions as, respectively, Cabibbo-favoured, Cabibbo-suppressed and doubly-Cabibbo-supressed based on their λ-suppression.

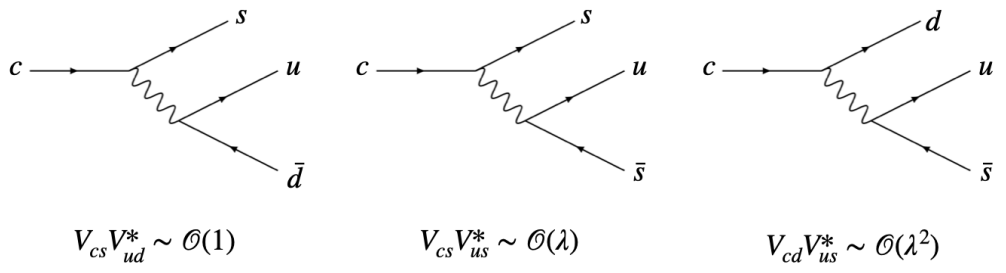


Figure 2.3: Cabibbo-favoured, Cabibbo-suppressed and doubly-Cabibbo-suppressed charm transitions.

2.2 CP violation in decays

As mentioned, quarks do not propagate freely and are confined within hadrons. Therefore, we must talk about these quark transitions in the context of hadron decays. Let a hadron P decay into a final state f, with a certain amplitude $\mathcal{A}(f) = \langle f|\mathcal{H}|P\rangle$, where $|\mathcal{A}(f)|^2$ indicates the probability amplitude of the process, and the corresponding charge-conjugates be $\bar{P}$ and $\bar{f}$ with amplitude $\bar{\mathcal{A}}(\bar{f})$. Here, $\mathcal{H}$ represents the Hamiltonian of the interaction. At scales lower than the W mass, which is where the c and b quark masses lie, we may integrate out the W and write the effective Hamiltonian in terms of the charged currents [20],

$$\mathcal{H}_{\text{eff}} = \frac{G_F}{\sqrt{2}} J^\dagger_\mu J^\mu \quad , \quad J^\mu = \begin{pmatrix} \bar{u} \\ \bar{c} \\ \bar{t} \end{pmatrix}^{\text{tr}} \gamma^\mu(1-\gamma_5) V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{2-19}$$

The hadronic degrees of freedom enter through QCD effects at the current energy scale μ. Hard (short-distance, perturbative) effects are calculable and included via Wilson coefficients C_i, while long-distance effects are left to be calculated in the matrix elements of quark current operators O,

$$\mathcal{H}_{\text{eff}} = \frac{G_F}{\sqrt{2}} \sum_i \lambda_i C_i(\mu) O_i(\mu)$$

with G_F the Fermi constant and λ_i the product of the relevant V_{CKM} elements. Methods for calculating matrix elements $\langle f|O_i|P\rangle$ will not be discussed in this dissertation. There are two different types of CP violating scenarios in the decays of hadrons:

1. Direct CP violation

   This type of CP violation implies having the probability for $P \to f$ differ from $\bar{P} \to \bar{f}$.

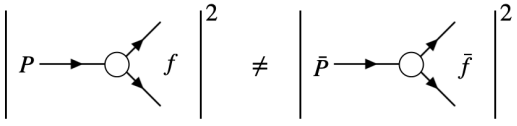


Figure 2.4: Direct CP violation.

2. Indirect CP violation

   This type of CP violation happens involving mixing $P - \bar{P}$, which can only occur for neutral hadrons.

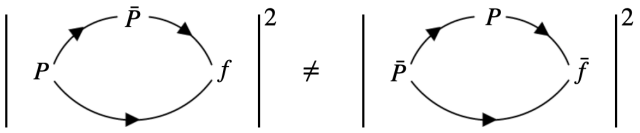


Figure 2.5: Indirect CP violation.

For neutral hadrons where both P and $\bar{P}$ decay to a common final state, we may have the interplay of both direct and indirect CP violation, i.e. in the decay and in the mixing.

In the analysis discussed in this dissertation, both $D_s^+ \to K^-K^+K^+$ and $D^+ \to K^-K^+K^+$ are charged meson decays, therefore only direct CP violation effects may appear and mixing will not be further discussed.

Since CP violation enters the SM through a complex phase, it is necessary to have an interference between at least two processes so that this phase may be observed. Therefore, let us consider the simplest case where two types of processes contribute to the amplitudes $\mathcal{A}(f)$ and $\bar{\mathcal{A}}(\bar{f})$, each with a certain magnitude M_i, a CP conserving phase, commonly refered to as strong phase, δ_i, and a CP violating weak phase ϕ_i. We have, then,

$$\begin{aligned}\mathcal{A}(P \to f) =& |M_1|e^{i\delta_1}e^{i\phi_1} + |M_2|e^{i\delta_2}e^{i\phi_2} \\ \bar{\mathcal{A}}(\bar{P} \to \bar{f}) =& |M_1|e^{i\delta_1}e^{-i\phi_1} + |M_2|e^{i\delta_2}e^{-i\phi_2} \ .\end{aligned} \tag{2-20}$$

The CP violation observable is an asymmetry defined as

$$A_{CP} = \frac{|\mathcal{A}(f)|^2 - |\bar{\mathcal{A}}(\bar{f})|^2}{|\mathcal{A}(f)|^2 + |\bar{\mathcal{A}}(\bar{f})|^2} \ , \tag{2-21}$$

and so one finds

$$A_{CP} = \frac{2|M_1||M_2|\sin(\delta_1 - \delta_2)\sin(\phi_1 - \phi_2)}{|M_1|^2 + |M_2|^2 + 2|M_1||M_2|\cos(\delta_1 - \delta_2)\cos(\phi_1 - \phi_2)} \ . \tag{2-22}$$

It is clear that for CP violation to be observed, $A_{CP} \neq 0$, we must have a weak phase difference between the two types of processes, but also a strong phase difference. These strong phases may have two different sources: either from long-distance effects from final-state interactions, or from short-distance effects. In D decays, this last type of effect can happen through the so-called penguin diagrams, in which pure QCD effects are present, which may contribute in transitions where two of the three final quarks are of the same flavour and form a $q\bar{q}$ pair, such as in Fig. 2.6.

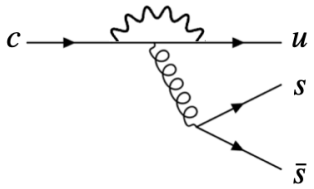


Figure 2.6: Penguin diagram topology for $c \to s\bar{s}u$ transition.

2.2.1 Three-body decays

In three-body decays, the dynamics are richer [21]. When a meson decays into two mesons, $M \to P_1P_2$, the momenta p_1, p_2 are completely determined by energy-momentum conservation, while for three particles final states where both decaying particle and products are spinless, we have two degrees of freedom, which then define a 2-dimensional phase space. For $M \to P_1P_2P_3$, the three daughters four-momenta introduce twelve parameters. From these, three degrees of freedom (d.o.f.) are removed from $E_i^2 = \mathbf{p}_i^2 + m_i^2$ relations for fixed masses and four d.o.f. are removed by energy-momenta conservation. Since for spinless particles the decay is isotropic in the centre of mass, the three Euler angle d.o.f. are also removed, leaving 2 variables to describe the decay.

Consider the decay $M \to P_1P_2P_3$, with the decaying particle having four-momenta $p = (E, \mathbf{p})$ and mass m and the product particles with p_i and m_i ($i = 1 \ldots 3$). We introduce the two-particle Lorentz invariant masses, s_{ij},

$$\begin{aligned} s_{12} &= (p_1 + p_2)^2 = (p - p_3)^2 \\ s_{13} &= (p_1 + p_3)^2 = (p - p_2)^2 \\ s_{23} &= (p_3 + p_3)^2 = (p - p_1)^2 \ , \end{aligned} \tag{2-23}$$

where it is easy to verify that they obey

$$s_{12} + s_{13} + s_{23} = m^2 + m_1^2 + m_2^2 + m_3^2 \ . \tag{2-24}$$

This way, we may use two of these invariant masses to describe the degrees of freedom of the three-body decay. The 2-D phase space of these variables is called the Dalitz plot.

The Dalitz plot includes all possible momenta configurations in the centre of mass of M that may happen in the decay, and its boundaries can be fully determined by analysing mass and three-momentum relations. For example, s_{12} is maximised when p_3 is at its minimum value of $p_3 = m_3$, so that $\mathbf{p}_3 = 0$, and P_1 and P_2 move with high equal three-momenta in opposite directions. On the

other hand, s_{12} is minimised when P_3 moves with maximum three-momentum and, using $|\mathbf{p}| = E|\mathbf{v}|$, it is easy to verify that

$$(p_1 + p_2)^2 = m_1^2 + m_2^2 + 2m_1 m_2 \left(\frac{1 - |\mathbf{v}_1||\mathbf{v}_2| \cos\theta}{\sqrt{(1 - |\mathbf{v}_1|^2)(1 - |\mathbf{v}_2|^2)}} \right) , \tag{2-25}$$

which is minimal for $|\mathbf{v}_1| = |\mathbf{v}_2|$ and $\cos\theta = 1$, i.e. same speed and direction, so P_1 and P_2 move slowly together, as a lump, such that $(p_1 + p_2)^2 = (m_1 + m_2)^2$. We then conclude

$$\begin{aligned} (m_1 + m_2)^2 &\leqslant s_{12} \leqslant (m - m_3)^2 \\ (m_1 + m_3)^2 &\leqslant s_{13} \leqslant (m - m_2)^2 \\ (m_2 + m_3)^2 &\leqslant s_{23} \leqslant (m - m_1)^2 \, . \end{aligned} \tag{2-26}$$

We can also derive the functions that define the kinematic limits of one invariant mass in terms of the other masses. For this, we have to define the reference frames R_{ij} where $\mathbf{p}_i + \mathbf{p}_j = \mathbf{0}$ for each pair of daughter particles i and j. In these frames, we may write energy and momentum expressions. For instance, in the frame R_{12}, we have $p_1 + p_2 = (E_1 + E_2, \mathbf{0})$, so $E_1 + E_2 = \sqrt{s_{12}}$, and

$$\begin{aligned} s_{12} &= m_1^2 + m_2^2 + 2p_1 \cdot p_2 \\ &= m_1^2 + m_2^2 + 2(E_1 E_2 - \mathbf{p}_1 \mathbf{p}_2) \\ &= m_1^2 + m_2^2 + 2E_1(\sqrt{s_{12}} - E_1) + 2\mathbf{p}_1^2 \\ &= m_1^2 + m_2^2 + 2E_1(\sqrt{s_{12}} - E_1) + 2(E_1^2 - m_1^2) \\ \Rightarrow E_1 &= \frac{s_{12} - m_2^2 + m_1^2}{2\sqrt{s_{12}}} \quad , \quad E_2 = \sqrt{s_{12}} - E_1 = \frac{s_{12} + m_2^2 - m_1^2}{2\sqrt{s_{12}}} \, . \end{aligned} \tag{2-27}$$

Similarly,

$$\begin{aligned} s_{12} &= m^2 + m_3^2 - 2p \cdot p_3 \\ &= m^2 + m_3^2 - 2E_3(\sqrt{s_{12}} + E_3) + 2(E_3^2 - m_3^2)^2 \\ \Rightarrow E_3 &= \frac{m^2 - s_{12} - m_3^2}{2\sqrt{s_{12}}} \, . \end{aligned} \tag{2-28}$$

By squaring the above expressions for E_k, and substituting $E_k^2 = \mathbf{p}_k^2 + m_k^2$,

we also derive the momentum expressions,

$$|\mathbf{p}_1| = |\mathbf{p}_2| = \frac{\lambda^{1/2}(s_{12}, m_2^2, m_1^2)}{2\sqrt{s_{12}}}$$
$$|\mathbf{p}_3| = \frac{\lambda^{1/2}(m^2, s_{12}, m_3^2)}{2\sqrt{s_{12}}} \ , \tag{2-29}$$

where $\lambda(x, y, z) \equiv x^2 + y^2 + z^2 - 2xy - 2yz - 2zx$ is the Källén function (also known as triangular function). In this frame, we may also want to consider the helicity angle $\theta_{13}^{R_{12}}$, as in Fig. 2.7, which is taken from $s_{13} = m_1^2 + m_3^2 + 2E_1E_3 - 2|\mathbf{p}_1||\mathbf{p}_3|\cos\theta_{13}^{R_{12}}$ using the above expressions,

$$\cos\theta_{13}^{R_{12}} = \frac{2s_{12}(m_1^2 + m_3^2 - s_{13}) + (s_{12} - m_2^2 + m_1^2)(m^2 - s_{12} - m_3^2))}{\lambda^{1/2}(s_{12}, m_2^2, m_1^2)\lambda^{1/2}(m^2, s_{12}, m_3^2)} \ . \tag{2-30}$$

Similar expressions can be derived for all reference frames and all angles within them. In this example, by taking the maximum and minimum value of the helicity angle, $\cos\theta_{13}^{R_{12}} = \pm 1$, we get the boundary expressions for $s_{13}(s_{12})$,

$$s_{13}^{\pm}(s_{12}) = m_1^2 + m_3^2 - \frac{1}{2s_{12}}\Big\{(s_{12} - m^2 + m_3^2)(s_{12} - m_2^2 + m_1^2) \mp \lambda^{1/2}(s_{12}, m_2^2, m_1^2)\lambda^{1/2}(m^2, s_{12}, m_3^2)\Big\} \ . \tag{2-31}$$

Thus, the Dalitz plot borders are defined, and can be seen in Fig. 2.8. For symmetric final states where two particles are identical, as is the case with $K^-K^+K^+$ labelled as particles 1, 2 and 3 respectively, it is common to consider s_{high} and s_{low} (instead of s_{12} and s_{13}), in which we would take, for example, K^- with the highest momentum K^+ (which means highest value amongst s_{12} and s_{13}) and with the lowest, respectively, having then a *folded* Dalitz plot, given that this equivalent to folding the Dalitz plot in half, at the diagonal.

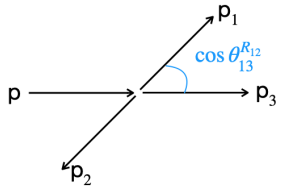


Figure 2.7: Three-body decay in the reference frame R_{12}.

The event density in the Dalitz plot is directly related to the differential decay rate, which can be written as [12]

$$d\Gamma = \frac{1}{(2\pi)^3 32m^2}|\mathcal{A}|^2 ds_{12}ds_{13} \ , \tag{2-32}$$

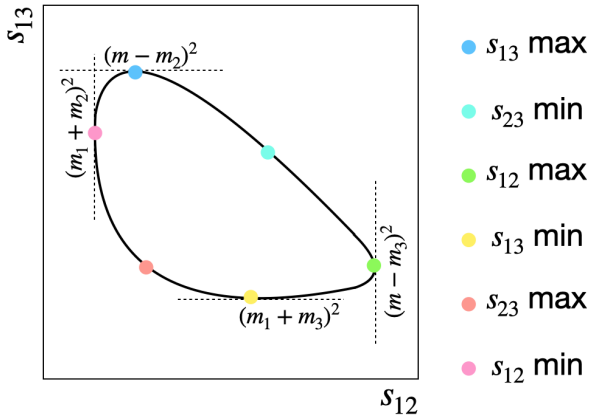


Figure 2.8: Generic Dalitz plot in $s_{12} \times s_{13}$.

so that the dynamics of the decay, given through the processes in the amplitude $\mathcal{A}2$, can be seen in the form of a non-uniform distribution across the phase space. This variation of density of events typically occurs due to the fact these decays can happen through intermediate resonances, e.g. $M \to R(\to P_1P_2)P_3$. The simplest description of this is done through the isobar model, where the total amplitude is described as a coherent sum of all resonant amplitudes (possibly including a non-resonant amplitude as well). This model supposes a $2+1$ approximation where the companion meson, i.e. that not coming from the resonance, does not interact with the other two mesons, and works well for narrow resonances and when there are non-overlapping resonances with the same spin.

The intermediate resonances have individual amplitudes usually parametrised as

$$\mathcal{A}_R = {}^{J}F_M {}^{J}F_R \times {}^{J}M_R \times BW_R \,, \tag{2-33}$$

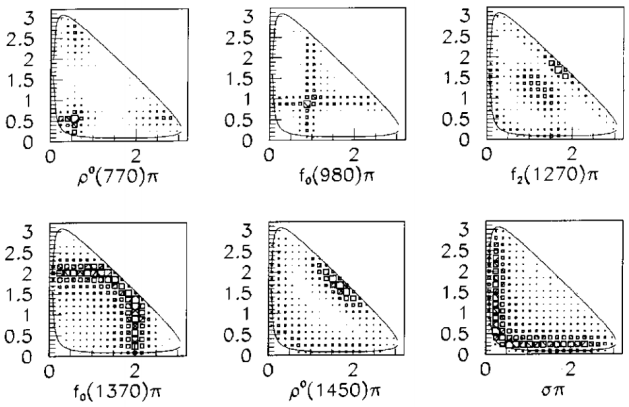
where J is the spin of the resonance, ${}^{J}F_{M,R}$ are the Blatt–Weisskopf damping factors [22] (which account for the finite sizes of M and R), ${}^{J}M_R = (-2|\mathbf{p}_k||\mathbf{p}_i|)^J P_J(\cos\theta_{ki}^{R_{ij}})$ is an angular factor (where P_L is the Legendre polynomial of order J) following the Zemach formalism in the rest frame of the resonance [23], and BW_R is the propagator which indicates the dynamical function, usually given by a shape like the relativistic Breit-Wigner,

$$BW_R = \frac{1}{m_R^2 - s_{ij} - im_R\Gamma_R(s_{ij})} \,. \tag{2-34}$$

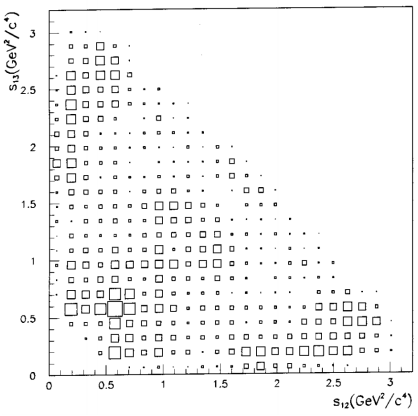
Due to the angular factor, resonances of different spins have different types of signatures across the Dalitz plot, creating as many nodes ($\cos\theta_{ki}^{R_{ij}} = 0$) as the value of their spins. When together, however, these resonances interfere and cause many different patterns to appear in the phase space, as shown in Fig. 2.9 for the $D^+ \to \pi^-\pi^+\pi^+$ decay . The final amplitude, in which a

non-resonant effect nr and N resonances may be present, each with a real magnitude a and generic phase δ where both strong and weak phases are included, is given by

$$\mathcal{A}(s_{12}, s_{13}) = a_{nr} e^{i\delta_{nr}} \mathcal{A}_{nr}(s_{12}, s_{13}) + \sum_{n}^{N} a_n e^{i\delta_n} \mathcal{A}_n(s_{12}, s_{13}) \; . \tag{2-35}$$



(a)



(b)

Figure 2.9: $D^+ \to \pi^-\pi^+\pi^+$Dalitz plot from [24, 25], (a) for simulation of possible individual resonance contributions and (b) for observed Dalitz plot.

These different resonant sub-processes may contribute differently to particle and antiparticle decays, namely to $\mathcal{A}(f)$ and $\bar{\mathcal{A}}(\bar{f})$. They might introduce different phases or have different magnitudes, generating CP violation effects. Furthermore, given their intricate interference patterns, the resonances introduce localised effects; as seen in the Breit-Wigner shape in Eq. 2-34, there is now an energy dependence in the amplitude, and since the BW is complex, there is also a non-trivial energy dependence on the strong phases.

To exemplify these local effects, let us consider that the sub-process 1 in Eq. 2-20 is resonant (consider spin zero for simplification) and 2 is non-

resonant. In this, we change $|M_1| \rightarrow \rho(s_{ij})e^{i\beta(s_{ij})}|M_1|$, where $\beta(s_{ij})$ is the energy dependent strong phase and $\rho(s_{ij}) = \sqrt{|BW_R|^2}$ the energy dependent part of the resonant amplitude, introduced by the Breit-Wigner. Extrapolating calculations presented previously, one finds

$$|\mathcal{A}(f)|^2 - |\bar{\mathcal{A}}(\bar{f})|^2 \propto \rho(s_{ij}) = \sqrt{\frac{1}{(m_R^2 - s_{ij})^2 + (m_R\Gamma)^2}} \, , \tag{2-36}$$

which, in the limit of small resonance widths, leads to $\rho \approx (m_R^2 - s_{ij})^{-1}$ [26]. From this, one concludes that the sign of the asymmetry A_{CP} should then change in crossing over a narrow range in the Dalitz plot close to a resonance's nominal mass. This and other localised effects such as a change of A_{CP} over the two regions $\cos\theta_{ki}^{R_{ij}} > 0$ and $\cos\theta_{ki}^{R_{ij}} < 0$ for spin 1 resonances, lead to the total (phase-space integrated) asymmetry to not be as large, since many opposite sign effects cancel out, but give us the opportunity to observe these potentially enhanced effects if we look separately at different Dalitz plot regions.

Experimentally, CP violation is measured in counting experiments, so by taking the yields of particle and antiparticle decays, respectively $N(P \rightarrow f)$ and $N(\bar{P} \rightarrow \bar{f})$, we define a raw asymmetry which includes CP violation effects, but also eventual initial state production and experimental final state detection asymmetries, as

$$A_{\text{raw}} = \frac{N(P \rightarrow f) - N(\bar{P} \rightarrow \bar{f})}{N(P \rightarrow f) + N(\bar{P} \rightarrow \bar{f})}. \tag{2-37}$$

This raw asymmetry can be measured both for the total phase space and for specific regions.

These effects have been observed in many B meson decays, where CP violation in the SM is large. Consider, for instance, the beauty meson decay $B^+ \rightarrow \pi^-\pi^+\pi^+$, for which we have a transition $b \rightarrow u\bar{u}d$, and so have both tree and penguin diagram contributions, but bringing a complex phase in V_{ub} already at tree level. We may then compare the measured CP violation in the full Dalitz plot with local raw asymmetries. For the example decay, these are reported in [27]. The raw local asymmetries can be seen in Fig. 2.10, while the reported total CP asymmetry is

$$A_{CP}(B^+ \rightarrow \pi^-\pi^+\pi^+) = +0.058 \pm 0.008 \pm 0.009 \pm 0.007 \, .$$

CP violation is undeniably present, but it is easy to see that localised effects are indeed enhanced. The total A_{CP} is $\sim 6\%$, where production and detection asymmetries have already been removed but are only about $\sim 1\%$, while the raw localised asymmetry in some regions of the phase space reaches

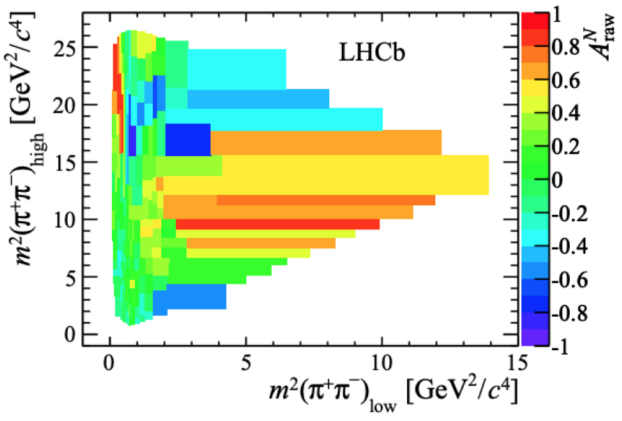


Figure 2.10: A_{raw} distribution over $B^+ \to \pi^-\pi^+\pi^+$ Dalitz plot, from [27].

$\sim$ 80%. Another feature clearly observed is the change of sign in crossing resonances; in the low $m^2(\pi^+\pi^-)_{\text{low}}$ (in our notation, s_{low}) region we have the vector $\rho(770)$ resonance [28] and can see, in the upper left, how the asymmetry changes from positive (red) to negative (blue) around $(0.77\ \text{GeV})^2 \approx 0.6\ \text{GeV}^2$.

2.3 Charm decays and CP violation

The study of CP violation in the charm sector is an interesting endeavour as it is the only sector where we are able to look at these effects for up-type quarks, being then complementary to $b-$ and $s-$hadron studies. On the bad side, the actual asymmetries predicted for the SM are much smaller, being the tree-level complex phase in V_{cd} suppressed by two powers of λ in comparison to V_{ub}, as seen in Eq. 2-15. There is a silver lining, though, seeing that the SM effect is so small, this provides an excellent probe to test New Physics (NP) effects.

Even so, the physics community has been able to design experiments capable of collecting just enough data to be sensitive to this effect. In 2019, the LHCb collaboration observed direct CP violation in D decays in the form of $\Delta A_{CP} = A_{CP}(D^0 \to K^-K^+) - A_{CP}(D^0 \to \pi^-\pi^+) = (15.4 \pm 2.9) \times 10^{-4}$ [10]. As more data are gathered, new observations are bound to come up in the following years.

In this context, analysing the decays $D_s^+ \to K^-K^+K^+$ and $D^+ \to K^-K^+K^+$ gives us the opportunity to look for two different kinds of effects, SM and NP, respectively, with the advantage of being three-body decays with potentially more sensitivity to CP violation as local effects in the phase space. To understand this, let us look at each decay individually.

2.3.1
$D_s^+ \to K^-K^+K^+$

This decay follows a $c \to s\bar{s}u$ transition, being a Cabibbo-suppressed decay, with three types of topologies contributing: annihilation, tree-level emission of W, and penguin.

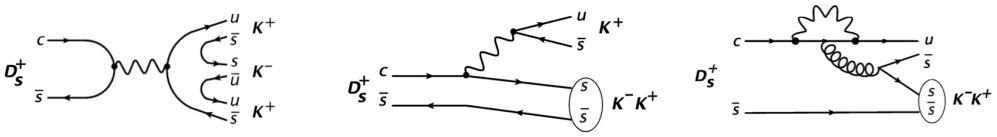


Figure 2.11: $D_s^+ \to K^-K^+K^+$topologies. For tree-level emission of W (middle) there is also a colour-suppressed amplitude when W is emitted internally and therefore does not produce a meson directly, requiring the resulting quark and antiquark to be in particular colour states so as to form neutral colour final state hadrons with the remaining components of the process.

From the diagrams, it can be seen that the annihilation and the tree-level emission diagrams carry no weak phases, and therefore no CP violation arises merely from their interference. On the other hand, in the penguin topology, a weak phase appears when the quark in the loop is b, due to V_{ub}, and as such the interference between the penguin amplitude and the tree topologies gives rise to CP violation effects. We can estimate the size of this CP asymmetry by writing in general terms [29, 30]

$$\begin{aligned} \mathcal{A}(c \to s\bar{s}u) &= \frac{G_F}{\sqrt{2}} V_{cs}V_{us}^* \mathcal{A}_T \left(1 + \frac{V_{cb}V_{ub}^*}{V_{cs}V_{us}^*}\frac{\mathcal{A}_P}{\mathcal{A}_T}\right) \\ &\Rightarrow A_{CP} \approx 2\left|\frac{V_{cb}V_{ub}^*}{V_{cs}V_{us}^*}\right| \sin\gamma \sin\delta_s \left|\frac{\mathcal{A}_P}{\mathcal{A}_T}\right| , \end{aligned} \tag{2-38}$$

where $\mathcal{A}_T$ and $\mathcal{A}_P$ refer to the leading amplitudes (tree and annihilation) and penguin amplitude, respectively. The denominator is approximated to 1 given that $\left|\frac{\mathcal{A}_P}{\mathcal{A}_T}\right|$ is small, with the penguin contribution highly suppressed in comparison to the tree, G_F is the Fermi constant, δ_s is the strong phase difference, and γ is the relevant weak phase. Using the current values [12] $|V_{us}| = 0.2245$, $|V_{cs}| = 0.987$, $|V_{cb}| = 0.041$, $|V_{ub}| = 0.00382$, $\gamma = 72.1°$, the approximation $\left|\frac{\mathcal{A}_P}{\mathcal{A}_T}\right| \approx \frac{\alpha_s}{\pi} \approx 0.1$, and the optimistic assumption $\sin\delta_s \approx 1$, Eq. 2-38 gives $A_{CP} \approx 1.3 \times 10^{-4}$ [31]. This value does not take into account long-distance effects ($\left|\frac{\mathcal{A}_P}{\mathcal{A}_T}\right|$ may be higher), so CP violation in charm can lead to $A_{CP} \simeq \mathcal{O}(10^{-3})$ [32].

Although the predicted effect is small, with enough data we will be sensitive to it. Additionally, in case there are NP effects, the effect could be

enhanced. Anyway, no search for CP asymmetry has been performed for this channel, and it represents a good way to look for a small SM effect.

2.3.2
$D^+ \to K^-K^+K^+$

Even though the final state is the same, this decay is very different from $D_s^+ \to K^-K^+K^+$. The quark level transition here is now $c \to d\bar{s}u$, being doubly-Cabibbo-suppressed, and so having no penguin diagram contribution to it. The possible topologies are displayed in Fig. 2.11.

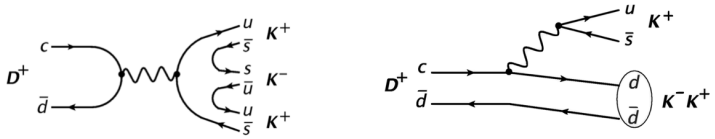


Figure 2.12: $D^+ \to K^-K^+K^+$topologies. For tree-level emission of W (right) there is also a colour-suppressed amplitude when W in emitted internally.

Since there is no penguin diagram with a weak phase, there is no SM CP violation for charm doubly-Cabibbo-suppressed decays. Because of this, studying this decay provides an interesting, and so far yet undone, way to look for NP effects, given that any CP asymmetry could only come from these. There are a few SM extensions which may lead to this effect, for example 2 Higgs-doublets models [33].

3
The LHCb experiment

3.1
The Large Hadron Collider (LHC)

In order to generate many short-lived and high-mass particles and study their properties, large particle accelerators are needed. In this context, CERN (*Conseil Européen pour la Recherche Nucléaire*) has developed machines to allow for the study of the most fundamental aspects of matter since its founding in the 1950s. Currently, the largest machine in the world resides at CERN near Geneva; called the Large Hadron Collider (LHC), it originally was designed to collide proton beams at 14 TeV centre-of-mass energy and $10^{34}\text{cm}^{-2}\text{s}^{-1}$ luminosity, being also able to collide heavy lead ions at 2.8 TeV with $10^{27}\text{cm}^{-2}\text{s}^{-1}$ peak luminosity [34]. The machine luminosity L relates the number of events per second generated in the machine, N, and the cross-section for the event under study, σ, by $N = L\sigma$.

The LHC is built in a 26.7 km tunnel between 45 m and 170 m below the surface, next to the Jura mountains in the border of Switzerland and France. It is connected by two transfer tunnels of 2.5 km to the CERN accelerator complex that injects the beams of protons. Using electromagnetic fields, the protons, which are taken from hydrogen gas by ionising it with an electric field, are led through the accelerator chain. This chain starts with the Linac2, a linear accelerator that takes these protons to 50 MeV of energy. Then, the beam proceeds to a series of synchroton accelerators; first the Proton Synchrotron Booster (PSB), which accelerates the protons to 1.4 GeV, then the Proton Synchrotron (PS), to 25 GeV, and finally the Super Proton Synchrotron (SPS), reaching 450 GeV. When the protons are finally injected in the LHC, they are separated in two beam pipes, each going in one direction, so they can be accelerated to their maximum energy, currently of 6.5 TeV. The two beams are brought into collision at one of the four experiment detectors, ATLAS, CMS, ALICE and LHCb. The accelerator complex, which also includes other experiments and parts, is shown in Fig. 3.1.

The LHC has so far had two accelerator runs, Run 1 in 2011–2012, and Run 2 in 2015–2018. Between runs, there is a long shutdown period where

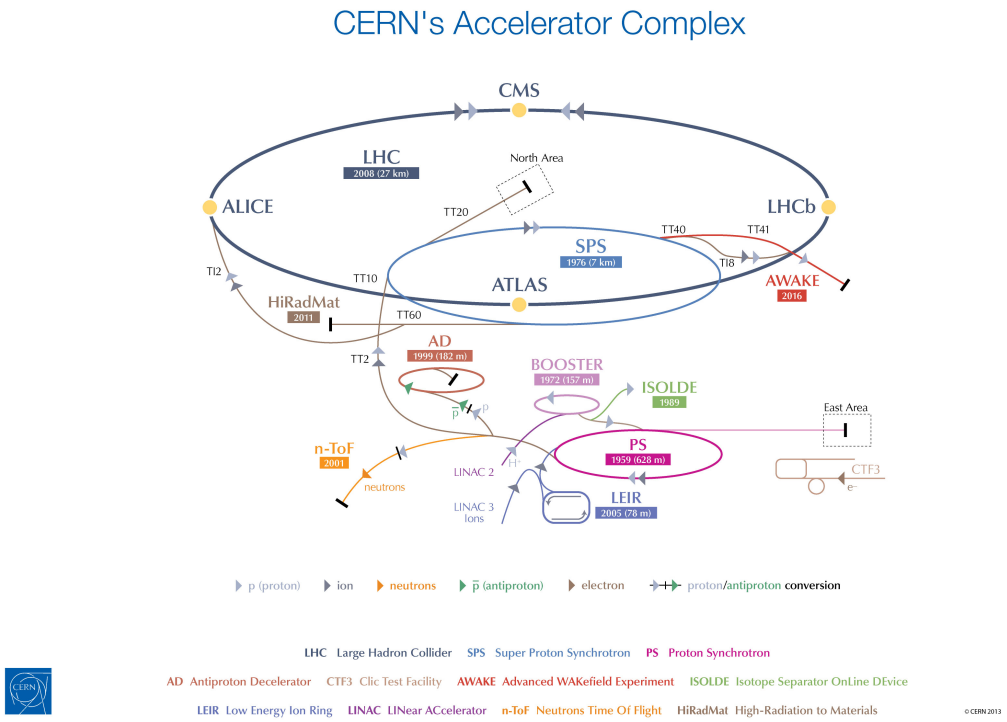


Figure 3.1: CERN accelerator complex [35].

maintenance and upgrades are performed on detectors. The LHC is scheduled to have Run 3 starting in 2022.

ALICE (A Large Ion Collider Experiment) is focused on heavy-ion physics, being mainly dedicated to studying the QCD quark-gluon plasma state of matter by colliding lead ions at high energies. The main goal of the collaboration is to better understand the phenomena of confinement and chiral-symmetry restoration [36].

ATLAS (A Toroidal LHC ApparatuS) and CMS (Compact Muon Solenoid) are both general-purpose detectors, aiming at studying a wide range of physics phenomena at the energy frontier. They cover almost the complete 4π solid angle around the collision point with very good energy precision, but using different technologies between them [37, 38]. After achieving their initial goal of discovering the Higgs boson [39, 40], both experiments have also been able to measure many different SM effects for heavy-particles, but also look for NP effects through missing energy, such as extra dimensions, dark matter or supersymmetric particles, which have not yet been observed [41].

The LHCb (Large Hadron Collider beauty) experiment is dedicated to studying flavour physics, with a focus on CP violation, rare decays and spectroscopy. Its detectors are built in the forward region so as to have high-precision tracking, and therefore make precision measurements on the production and decays of $b-$ and $c-$hadrons [42].

3.2 The LHCb detector

The LHCb detector is a single-arm spectrometer, covering a forward angular region of 10—300 mrad in the bending plane of the magnets and 10—250 mrad in the non-bending plane. This design is justified by the fact that $b-$ (and $c-$) hadrons are predominantly produced in a direction close to the beam collision direction, going either forwards or backwards. The machine is composed of many subdetector parts, with the trigger system performed in two levels (hardware and software). In dealing with the decays of these heavy hadrons, it is fundamental to have high precision in identifying the collision point (primary vertex, PV) and the decay point (secondary vertex, SV) to accurately calculate flight distances, as well as other topological quantities. Another essential characteristic of the LHCb for flavour physics studies is its excellent particle identification system. The layout of the spectrometer, with its many subdetectors, can be seen in Fig. 3.2, where de z axis goes along with the beam and y axis goes along the vertical.

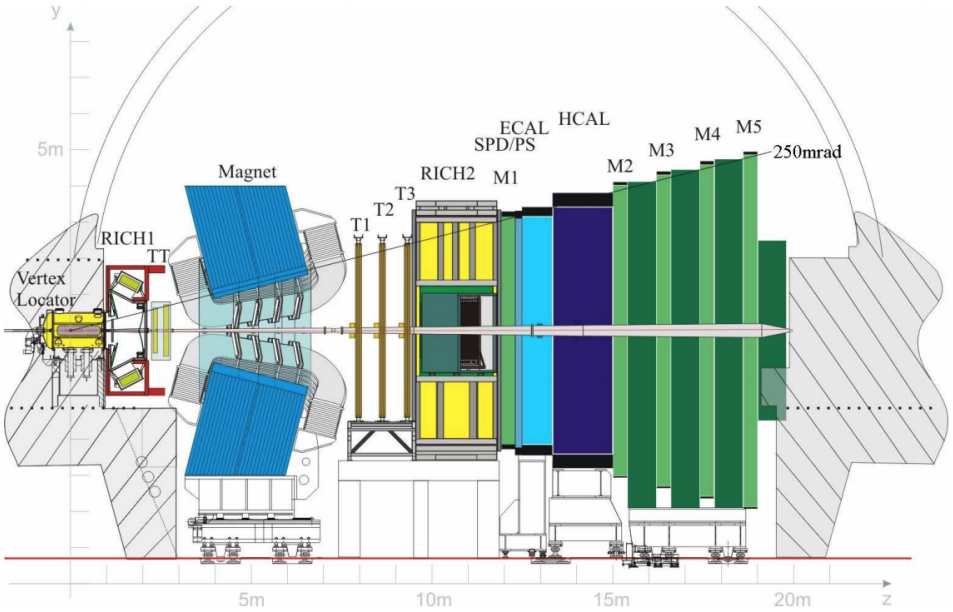


Figure 3.2: LHCb experiment layout [42].

The Vertex Locator (VELO) [43] is the first tracking station and it measures precisely the track coordinates, allowing for an identification of PV and SV and consequent measure of certain topological variables, such as the impact parameter (IP). The VELO consists of a series of silicon strip modules along the beam direction, which give a good measurement of the radial variables r and ϕ around the collision. The VELO sensors are retractable, being

open while the beams are circling the LHC but being placed very close to the beam line for the collisions. Schematics for the VELO can be seen in Fig. 3.3.

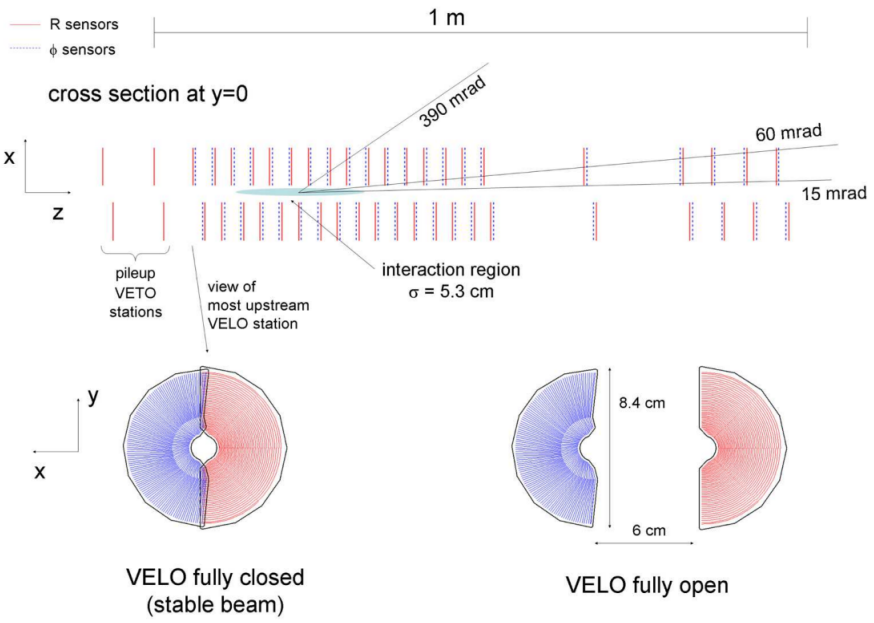


Figure 3.3: Schematics of VELO subdetector [42].

The next tracking station, Tracker Turicensis (TT), is located upstream of the magnets. TT uses silicon strips and has 4 planar layers, with the 2 inner layers having rotations of -5^o and $+5^o$ to achieve better resolution. The other tracking stations, T1, T2, and T3, consist of an inner tracker (IT) and an outer tracker (OT). The IT's are positioned in the centre of tracking stations and have the same 4 layer configuration of silicon strips as the TT, while the OT's consist of two staggered layers of straw-tube drift chambers [44, 45].

In the Magnet station, the LHCb uses a dipole magnet with a 4 Tm magnetic field that bends the trajectory of charged particles, so that information on their momenta can be obtained [46]. To reduce asymmetries due to overall detector efficiencies, the LHCb inverts the polarity of the magnetic field, taking data with either the field pointing up (MagUp) or down (MagDown).

For particle identification (PID), three types of stations are used, the RICH stations, the calorimeters and the muon system. The RICH (Ring-Imaging Cherenkov System) has the main goal of distinguishing final state pions, kaons and protons [47]. While the RICH1 uses aerogel and C_4F_{10} radiators to cover the $\sim 1 - 60$ GeV range of charged particle momentum, RICH2 uses a CF_4 radiator and covers the high $\sim 15 - 100$ GeV. When particles pass through these detectors with speed greater than that of light in the medium, Cherenkov light is created which is focused with a combination of spherical and flat mirrors and reflected out of the spectrometer acceptance.

Hybrid Photon Detectors (HPDs) are used to detect the Cherenkov photons, which form a cone whose angle θ_c with the trajectory is related to the speed of the particle. Using information of the associated momentum from the tracking stations to infer particle mass, it is then possible to relate probabilities for identification of the particle as each possible type. A plot of Cherenkov angle versus momentum for different particles in the RICH C_4F_{10} radiators can be seen in Fig. 3.4.

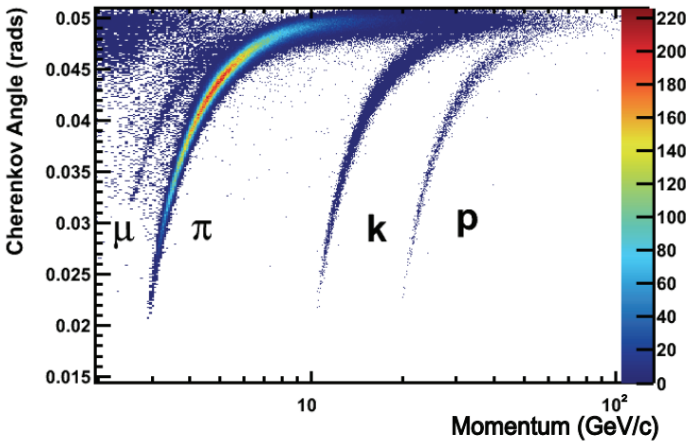


Figure 3.4: Cherenkov angle vs. particle momentum for the RICH C_4F_{10} radiators [48].

The electromagnetic (ECAL) and hadronic (HCAL) calorimeters, together with the stations SPD (Scintillator Pad Detector), PS (Preshower Detector), have the main goal of measuring the energies, but also providing information on the identification of electrons, photons and hadrons [49, 50]. One particular function of these stations is to identify and select high transverse energy hadron, electron and photon candidates for the first trigger level (L0), with the decision made $4\mu s$ after the interaction. The SPD/PS system is positioned before the calorimeters, being complementary to them, where the SPD identifies charged particles and allows electrons to be distinguished from pions, and is separated from the PS, which identifies electromagnetic particles, by a lead wall. The principle of the calorimeters is to stop the passing particles, measuring their energy loss in this process, and so LHCb uses the usual design of having the ECAL (stops lighter particles) upstream from the HCAL (stops heavier particles). The passing particles hit the detectors and produce showers of secondary particles which excite the medium and produce scintillation light, transmitted to a PhotoMultiplier. Neutral particle identification (such as photons or neutrons) is mainly obtained from calorimeters. A schematic representation of the signal deposited on the different parts of the calorimeter system can be seen in Fig. 3.5.

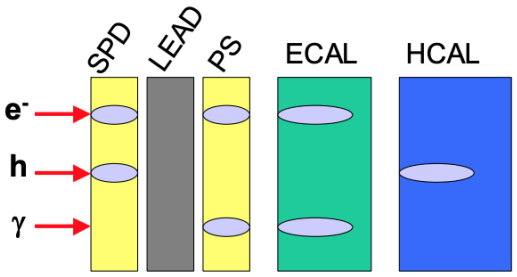


Figure 3.5: Signal deposited in the calorimeter parts by an electron, a hadron and a photon [50].

There are five muon chambers, M1, M2, M3, M4 and M5, with the goal of providing information for high transverse momentum, p_T, muon L0 trigger and also muon identification for high-level trigger (HLT) [51–53]. Station M1 is placed before the calorimeter to improve the p_T measurement in trigger, while the remaining M2—M5 are placed after the calorimeter, interleaved with iron absorbers. Stations M1—M3 have high spatial resolution along the bending plane, and are used to define the direction of the track and calculate the p_T, while stations M4 and M5, which have limited spatial resolution, have the main goal of identifying penetrating particles. The muon trigger is based on the muon track reconstruction and p_T, requiring that the hits be alligned in all 5 stations.

The trigger system acts in reducing the readout frequency from the crossing of 40 MHz to 12.5 kHz [54, 55], by selecting interesting events to be written to storage and further analysed. This is achieved in two trigger levels, a hardware Level-0 (L0) and a software High Level Trigger (HLT). The L0 takes information from the calorimeter and muon systems to select high E_T hadron, electron and photon, and high p_T muon candidates, while also rejecting high multiplicity events from the SPD system. The following HLT is executed asynchronously on a processor farm, being subdivided in two stages, HLT1 and HLT2. HLT1 has the purpose of refining the candidates found by L0, by requiring the candidates to be reconstructed in the VELO or tracking stations, which sufficiently reduces the rates for full pattern recognition of the remaining events. Then, HLT2 performs a combination of inclusive (partial reconstruction of decay) and exclusive (full reconstruction of decays) trigger algorithms to select interesting final states. For Run 2, LHCb introduced the *Turbo stream*, where the HLT2 selection of candidates is saved to disk. Trigger responses are classified as either `TIS` (trigger independent of signal) or `TOS` (trigger on signal) with respect to a defined signal, i.e. a $D^+_{(s)}$ candidate reconstructed from combining three kaon-like tracks that can be traced back to approximately the same vertex in our case.

4
Data sample and selection

4.1
Definition of variables

In order to understand how the $D^+_{(s)}$ candidates for the analysis are selected, both from trigger and offline analysis, the relevant variables are defined here. Mostly, the variables are based on topological characteristics of the decay and on particle identification, and are associated to physical quantities so that requirements on them may distinguish as much as possible signal from background events. In the topology of the three-body decay in Fig. 4.1, from the primary vertex where the $D^+_{(s)}$ is produced and the secondary vertex where it decays, the flight distance (FD), impact parameter (IP), $D^+_{(s)}$ momentum p and transverse momentum p_T and direction angle θ variables can be seen.

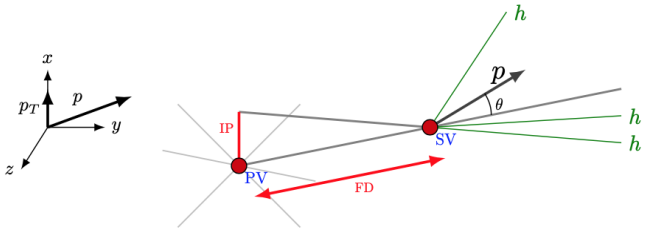


Figure 4.1: Topology of three-body decay.

The relevant variables for our analysis are:

mass$(K^-K^+K^+)$	Invariant mass combination of the three kaon candidates.
Flight distance (FD)	Distance from the PV to the SV.
FDχ^2	χ^2 of FD, that is, the square of FD over the square of its uncertainty.
Impact parameter (IP)	Minimum distance of a particle's trajectory to the PV.
IPχ^2	Difference of χ^2 of the reconstructed PV with and without the particle being considered.

Momentum (p) Modulus of a particle's momentum.

Transverse momentum (p_T) Momentum in the direction perpendicular to the beam axis.

Lifetime Time it takes for the $D^+_{(s)}$ candidate to decay, determined from momentum and FD information.

Direction angle (DIRA) Cosine of the angle between the momentum of the particle and the direction of flight from the associated PV to the SV.

Vertex χ^2 Information on the quality of the SV reconstruction from the three final state tracks.

Pointing Comparison of $D^+_{(s)}$ momentum perpendicular to flight direction to reconstructed final state particles transverse momenta,

$$\text{Pointing} = \frac{p\sin\theta}{p\sin\theta + \sum_i^3 p_{T_i}} \ ,$$

where θ is the angle between the $D^+_{(s)}$ momentum and the flight direction, and the sum runs over the transverse momenta of the three daughter particles.

Pseudorapidity (η) Describes the angle of the particle relative to the beam axis,

$$\eta = -\ln\left[\tan\frac{\theta}{2}\right].$$

Combination logIP Variable that combines IPχ^2 for mother and daughter particles,

$$\text{logIP} = \log\frac{\prod_i^3(\text{IP}\chi^2)_i}{\text{IP}\chi^2}.$$

Track χ^2/ndf χ^2 per degree of freedom of the track fit.

PIDK Information on PID from the RICH, it is the delta-log-likelihood of being a kaon with re-

spect to the pion hypothesis,

$$\text{PIDK} = \Delta \log \mathcal{L} = \log \frac{\mathcal{L}_K}{\mathcal{L}_\pi}.$$

ProbNNk Kaon probability variable, it is the output of multivariate techniques that result in a single probability value for the given particle hypothesis of being a kaon, which is created combining tracking and PID information.

Since experimentally the reconstructed mass of the $D^+_{(s)}$ candidates follow a distribution with a certain width, due to momentum resolution, the Dalitz plots of three-body decays extend over the theoretical borders as defined in Sec. 2. To account for this, a commonly used internal package for handling LHCb data is the DecayTree Fitter (DTF) package, which, in our case, forces the $D^+_{(s)}$ mass into its nominal values by changing the momentum, ensuring that the events are placed within their physical phase space. In this analysis, the Dalitz plots are always constructed using DTF invariants.

4.2 Data samples

The analysis is performed using the data collected in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ by LHCb in the years 2016–2018, which in total correspond to $5.6\,\text{fb}^{-1}$ of integrated luminosity.

The data set comes directly from the exclusive HLT2 Turbo lines `Hlt2CharmHadDspToKmKpKp, Hlt2CharmHadDpToKmKpKp` which select the $D^+_s \to K^-K^+K^+$ and $D^+ \to K^-K^+K^+$ candidates, respectively. The selection criteria are almost the same for these lines, except for the DIRA, the lifetime and the mass spectrum. These criteria are shown in Table 4.1 .

These HLT2 trigger lines require the D^+ and D^+_s candidates to be TOS on any of the two HLT1 Track Lines, `Hlt1TrackMVA` and `Hlt1TwoTrackMVA`, which in turn requires the events to pass any of the L0 hardware trigger lines. Details of the trigger specifications can be found in [56].

In order to ensure good control of the online selection, candidates are required to satisfy either a `TOS` decision at L0 of being a hadron, or a combination of `TIS` decisions of including hadrons, electrons, photons or muons, such that the events represent the inellastic collisions.

Table 4.1: HLT2 selection criteria.

Daughter requirements	
Track χ^2/ndf	< 3.0
p_T[MeV]	> 250
IP χ^2 wrt any PV	> 4
PIDK	> 5
Combination requirements	
Mass [MeV]	1779 – 1959 (D^+) , 1879 – 2059 (D_s^+)
$\sum p_T$ [MeV]	> 3000
p_T at least one track [MeV]	> 1000
p_T at least two tracks [MeV]	> 400
IP χ^2	> 10 (at least one track)
IP χ^2	> 4 (at least two tracks)
Mother requirements	
Track vertex χ^2/ndf	< 10
Mass [MeV]	1789 – 1949 (D^+) , 1889 – 2049 (D_s^+)
lifetime [ps]	> 0.4 (D^+) , > 0.2 (D_s^+)
acos(DIRA) [mrad]	< 10 (D^+) , < 14.1 (D_s^+)

4.3 Data offline selection

On top of the trigger level requirements we also apply some offline selection requirements. The 2015 data were initially considered for the analysis of selection criteria, but they were discarded due to divergences in the trigger definitions, and so, in order to define the selection criteria, we analyse further requirements using the 2016 data sample, corresponding to $1.67\,\mathrm{fb}^{-1}$.

First, we use loose pre-selection requirements, shown in Table 4.2. These requirements consist of kinematic requirements including SPD multiplicity, mass spectrum requirements and a loose IPχ^2 requirement on the mother particle to ensure it is better reconstructed. The mass distributions and Dalitz plots for the 2016 data after the pre-selection are shown in Figs. 4.2 and 4.3, and the total number of events in the 2016—2018 samples are given in Table 4.3.

Table 4.2: Additional pre-selection requirements.

nSPDHits < 1000		
$1.5 < \eta_{\text{daughters}} < 5$		
$p_{\text{daughters}} < 100$ GeV		
IP$\chi^2 < 15$		
D_s^+	:	$1905 < M < 2035$ MeV
D^+	:	$1805 < M < 1935$ MeV

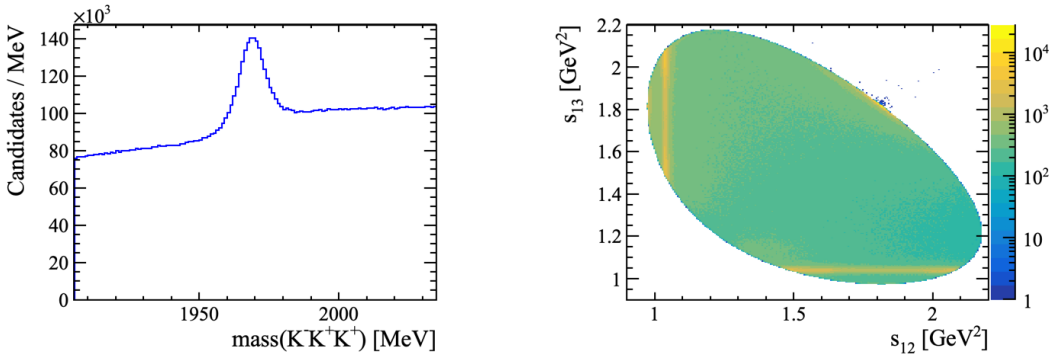


Figure 4.2: $D_s^+ \to K^-K^+K^+$ mass distribution (left) and Dalitz plot (right) where the colour scale indicated population.

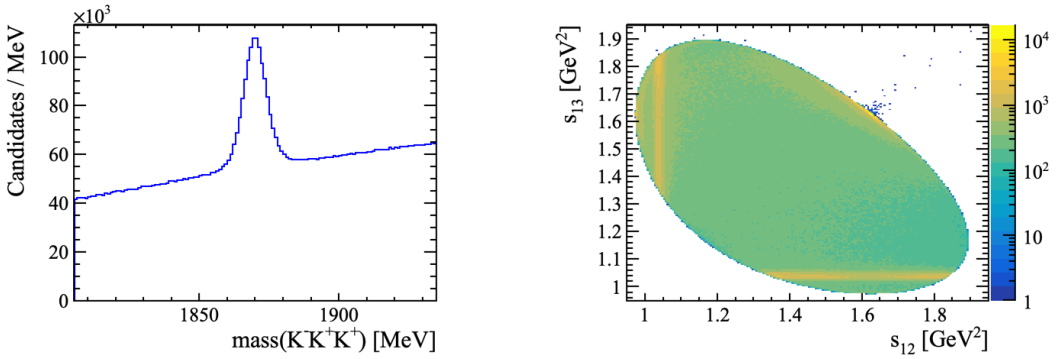


Figure 4.3: $D^+ \to K^-K^+K^+$ mass distribution (left) and Dalitz plot (right) where the colour scale indicated population.

Our data samples have high levels of background, as can be seen in the mass distributions. This background could either introduce nuisance charge asymmetries or dilute potential CP violation signals. To reduce the background, we analyse variables to reduce both combinatorial (random three-track associations) and specific backgrounds, including contamination from other channels that have one or more of their final state particles misidentified as a kaon, cloned tracks and high asymmetry regions of instrumental source.

Except for requirements needed to reduce specific backgrounds, to define the selection criteria we use the signal significance, defined as

$$\text{Significance} = \frac{S}{\sqrt{S+B}}, \tag{4-1}$$

where S and B are the estimated number of signal and background candidates using the signal and background regions as defined in Table 4.4, and where S is obtained by sideband subtraction assuming a linear background distribution over the $K^-K^+K^+$ mass spectrum. To improve the signal significance, we further use a multi-variate analysis selection, explained later in this chapter.

Table 4.3: Total number of candidates after pre-selection.

channel	year	Polarity	N. of candidates
$D_s^+ \to K^-K^+K^+$	2016	Up	5546513
		Down	7030984
	2017	Up	6631108
		Down	6859547
	2018	Up	8133104
		Down	7664855
$D^+ \to K^-K^+K^+$	2016	Up	3325296
		Down	4252244
	2017	Up	4060371
		Down	4188992
	2018	Up	4992958
		Down	4700974

Table 4.4: $D^+ \to K^-K^+K^+$ and $D_s^+ \to K^-K^+K^+$ signal and background regions used to estimate S and B.

D^+	Signal region	1860 – 1880 MeV
	Left wing	1820 – 1830 MeV
	Right wing	1910 – 1920 MeV
D_s^+	Signal	1960 – 1980 MeV
	Left wing	1920 – 1930 MeV
	Right wing	2010 – 2020 MeV

4.3.1 Charm background

A source of specific background to the data is the cross-feed from other decay channels with mis-identification (mis-ID) of a daughter particle. By examining the mass distribution of the $D_s^+ \to K^-K^+K^+$ sample in Fig. 4.2, the contamination from $D^+ \to K^-K^+\pi^+$ at the right mass sideband is clear, as the background level on the right is much higher than that in the left (it also contributes in the signal region). This and other charm background in the $K^-K^+K^+$ mass spectra for both D^+ and D_s^+ are removed by specific requirements based on PID.

To remove from $D^+ \to K^-K^+\pi^+$ in the $D_s^+ \to K^-K^+K^+$ sample, where the pion is misidentified as a kaon, both PIDK and ProbNNk variables are tested with the goal of reducing the D^+ peak (centered at mass ~ 1870 MeV) in the invariant mass spectrum where we reconstruct the events by assigning a pion mass to one daughter. Since the mis-ID can happen for either one of

the two K^+ daughters, the same requirement is applied on both of them. The effect of different PIDK requirements is shown in Fig. 4.4.

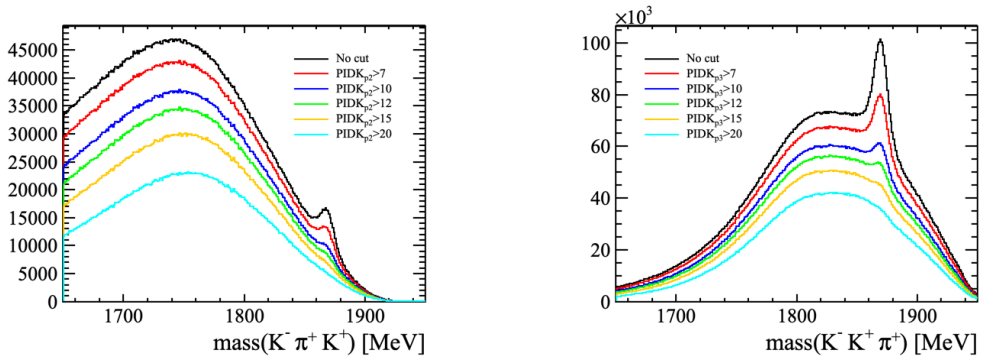


Figure 4.4: Three-body invariant mass reconstructed by assigning the pion mass to the daughter of lowest (left) and highest (right) momentum, in $D_s^+ \to K^-K^+K^+$ data, for multiple values of PIDK requirements.

We see that the $D^+ \to K^-K^+\pi^+$ contribution is reduced to a negligible level with $\text{PIDK}_{2,3} > 15$. Using the ProbNNk variable, the requirement would be at $\text{ProbNNk}_{2,3} > 0.95$. The significance of the signal yield with these requirements is 375.4 ± 1.1 and 358.0 ± 1.0, respectively. Therefore, the requirement on PIDK is chosen. No need was found to apply a PID requirement for the odd-charge kaon (on top of that from trigger): no gain in significance and no specific background to be avoided.

Other possible sources of contamination to both $D_s^+ \to K^-K^+K^+$ and $D^+ \to K^-K^+K^+$ from mis-ID are $\Lambda_c^+ \to \text{K}^-K^+p$ and $\Lambda_c^+ \to \text{K}^-\pi^+p$, which are not immediately visible in the mass distributions of the data, but are seen as peaks centered at mass ~ 2290 MeV in the invariant mass reconstruction by assigning different masses to the K^+ daughters. For $D_s^+ \to K^-K^+K^+$, we find that the PIDK requirement already applied removes this cross-feed, being negligible in the signal region, as seen in Fig. 4.5.

For the selection of $D^+ \to K^-K^+K^+$ candidates, we also look initially at PID requirements for the kaons. In this case, with no clear cross-feed seen in the mass spectrum, the study is based on the significance of the signal, using Eq. 4-1 and with the signal and background regions defined in Table 4.4. Both PIDK and ProbNNk variables are considered, and we find that the ProbNNk requirement yields both a better significance and efficiency (defined as the ratio between the estimated number of signal events after applying the requirement and that of before). The optimal requirement found is $\text{ProbNNk}_{1,2,3} > 0.6$, as shown in Fig. 4.6, with a relative efficiency of $\sim 86\%$. This requirement also reduces the Λ_c^+ contributions to a negligible level, as seen in Fig. 4.7.

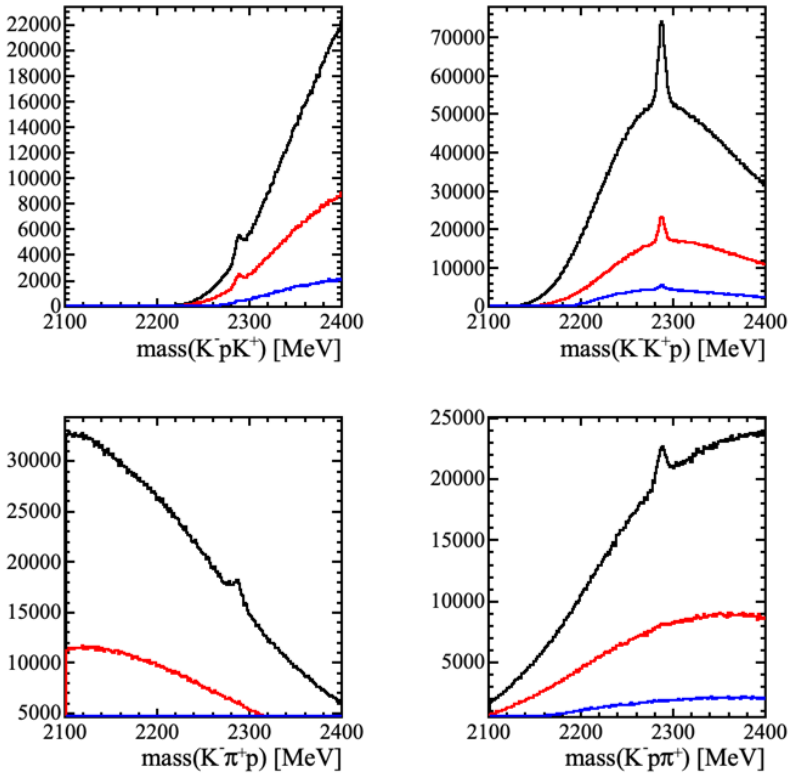


Figure 4.5: Possible misID contributions from Λ_c decays for $D_s^+ \to K^-K^+K^+$ (mass combinations indicated in the X-axis label). Plots include full data after pre-selection (black), data after applying $\mathrm{PIDK}_{2,3} > 15$ (red) in the full $K^-K^+K^+$ mass spectrum, and only in the signal region (blue).

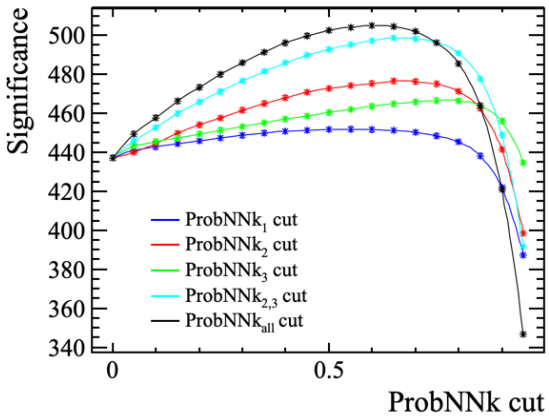


Figure 4.6: Significance of different combinations of ProbNNk requirements for $D^+ \to K^-K^+K^+$ daughters.

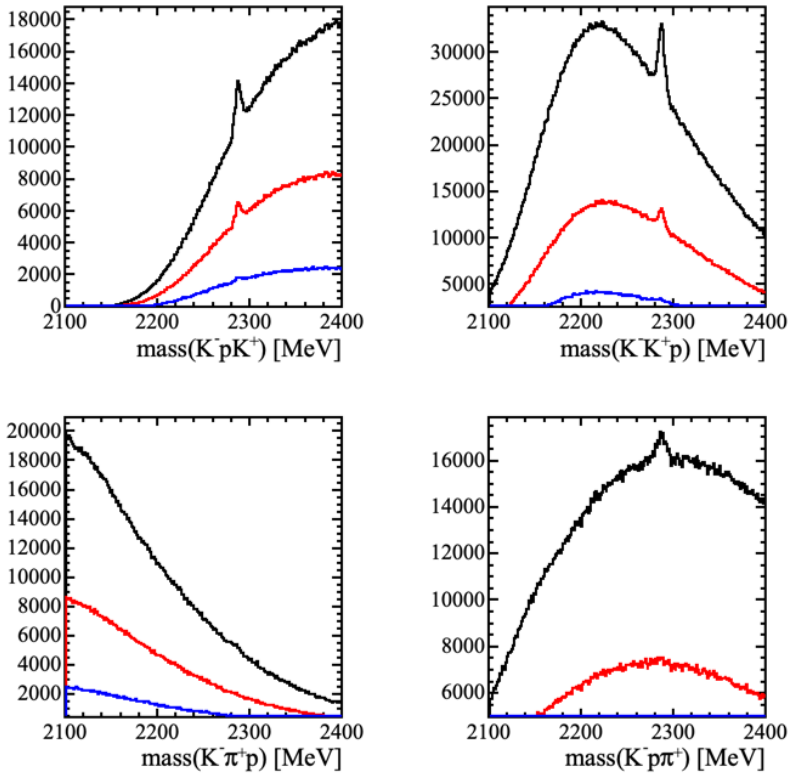


Figure 4.7: Possible misID contributions from Λ_c^+ decays for $D^+ \to K^-K^+K^+$. Plots include full data after pre-selection (black), data after applying ID requirement (red), and data after the requirement but only in the signal region (blue).

4.3.2
Clones and fiducial requirements

The next contribution eliminated was that of cloned tracks of the K^+ daughters. A single track can sometimes be duplicated in the reconstruction process, and as a result, some of the pairs of K^+ reconstructed in the final state may be due to just one track. This contribution can be seen in the Dalitz plots, Figs. 4.2 and 4.3, as a high concentration of events close to the furthermost border (of lowest K^+ momenta) of the Dalitz around the diagonal direction (in which $p_2 = p_3$). To differentiate the tracks, two slope difference variables, difTX23 and difTY23, are defined as:

$$\text{difTX23} = \left| \frac{p_{x_2}}{p_{z_2}} - \frac{p_{x_3}}{p_{z_3}} \right| \quad , \quad \text{difTY23} = \left| \frac{p_{y_2}}{p_{z_2}} - \frac{p_{y_3}}{p_{z_3}} \right| . \tag{4-2}$$

When we look at the distribution of these two variables, presented in Fig. 4.8 for $D_s^+ \to K^-K^+K^+$, the contribution of the clones is clearly seen as a peak close to zero. The requirement was chosen to be the same for $D_s^+ \to K^-K^+K^+$ and $D^+ \to K^-K^+K^+$ and in such a way as to sufficiently remove the contribution on the Dalitz plot while maintaining a good significance and

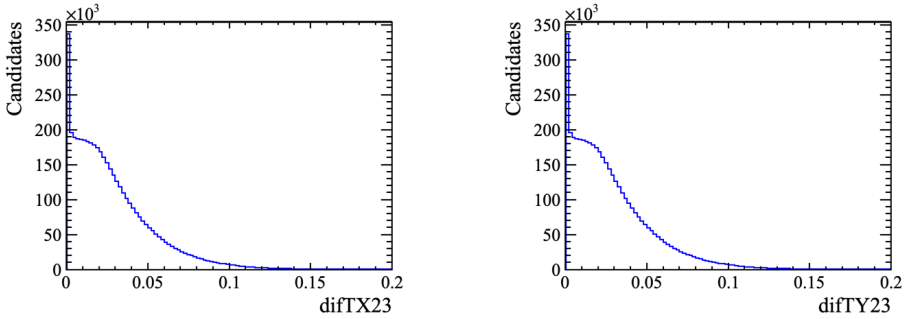


Figure 4.8: $D_s^+ \to K^-K^+K^+$ slope difference variables, on the X direction (left) and on the Y direction (right).

efficiency. The applied requirement was difTX23 $> 2.3 \times 10^{-3}$ || difTY23 $> 2.3 \times 10^{-3}$, with efficiencies of over 99% for both D_s^+ and D^+.

Next, we look at regions with high instrumental charge asymmetries, defined as $N(D_{(s)}^+) - N(D_{(s)}^-)/N(D_{(s)}^+) + N(D_{(s)}^-)$. Low momentum tracks may be driven away from the detector acceptance by the magnetic field, inducing a charge asymmetry. The asymmetries are plotted on the $p_{x,y} \times p_z$ planes for Up and Down magnet polarities separately for each daughter and the fiducial requirements are defined in such a way as to remove the high asymmetry regions; they are the same for both $D_s^+ \to K^-K^+K^+$ and $D^+ \to K^-K^+K^+$, and are defined as

$$p_{z_{1,2,3}} > 1 + 3.57 \times |p_{x_{1,2,3}}| \ ,$$

$$p_{z_{1,2,3}} > 4.25 \times |p_{y_{1,2,3}}| \ ,$$

$$p_{z_{1,2}} > 4 \ .$$

An example for the $p_x \times p_z$ plane for $D_s^+ \to K^-K^+K^+$ can be seen in Fig. 4.9. These requirements have a relative efficiency of over 96% for both D_s^+ and D^+.

4.3.3 Multi-variate analysis

The last step in the offline selection is the use of a multi-variate analysis (MVA) to select the candidates of each decay channel (separately) providing a continuous way of improving the significance and purity ($\frac{S}{S+B}$ in signal region) of the samples. In this type of analysis, the neural network is first trained on the signal and background data to make a decision by simultaneously taking into account the many discriminatory variables, and then tested on a different part of the signal and background data so as to assign a value to the new decision classifier variable for each event. This response can then be applied

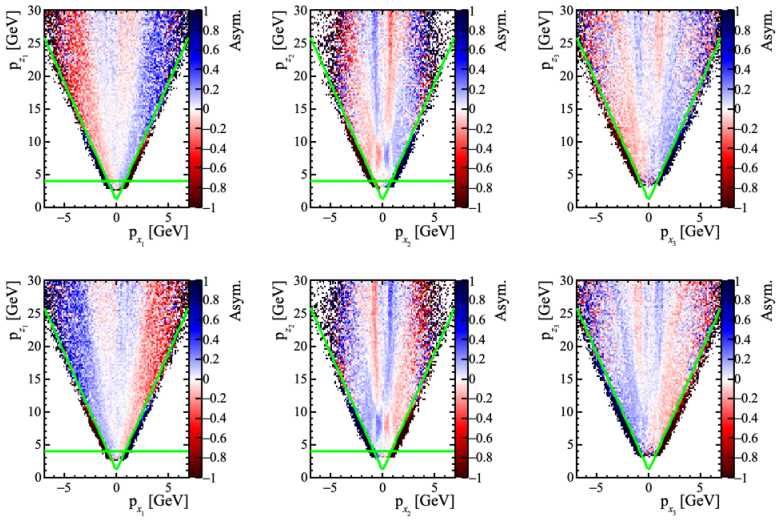


Figure 4.9: $D_s^+ \to K^-K^+K^+$ asymmetries in the $p_x \times p_z$ plane for MagUp (top plots) and MagDown (bottom plots) for each daughter particle. The green lines represent the requirements applied.

to a larger data sample, where we get the associated decision variable and can analyse different requirements on it.

The MVA was based on a Boosted Decision Tree (BDT) technique [57]. This optimisation technique consists of a binary tree that makes decisions one variable at a time, aiming to refine the classification of signal and background events.

In the training and testing phase, a random 50 thousand candidate sample (half MagUp, half MagDown) from the 2016 data set for each decay channel is used. To separate signal from background events, an sPlot fit is used [58]. The sPlot assigns a signal and a background weight to every event of the dataset, based on the signal and background components of an invariant mass fit probability density function (PDF). To obtain the signal PDF, consisting of a Gaussian and two Crystal Balls [59] to account for the tails in both sides of the distribution, full LHCb Monte Carlo simulations are used, which are explained in Appendix A. The Crystal Ball function consists of a Gaussian core with a power-law tail. This function is parametrised by its mean μ, width σ, and tail parameters N and α, being given by

$$\mathrm{CB}(m|\mu,\sigma,\alpha,N) = \mathcal{N}\cdot\begin{cases}\exp(-\frac{(m-\mu)^2}{2\sigma^2}), & \text{for } \frac{m-\mu}{\sigma} > -\alpha \\ A\cdot(B-\frac{m-\mu}{\sigma})^{-N}, & \text{for } \frac{m-\mu}{\sigma} \leqslant -\alpha\end{cases}, \tag{4-3}$$

where $\mathcal{N}$ is a normalisation factor and

$$A = \left(\frac{N}{|\alpha|}\right)^N \cdot \exp\left(-\frac{|\alpha|^2}{2}\right) \ , \ B = \frac{N}{|\alpha|} - |\alpha| \ , \ \mathcal{N} = \frac{1}{\sigma(C+D)} \ ,$$
$$C = \frac{N}{|\alpha|} \cdot \frac{1}{N-1} \cdot \exp\left(-\frac{|\alpha|^2}{2}\right) \ , \ D = \sqrt{\frac{\pi}{2}}\left(1 + \mathrm{erf}\left(\frac{|\alpha|}{\sqrt{2}}\right)\right) \ . \quad (4\text{-}4)$$

The background is parametrised by a Bernstein polynomial. The sPlot fits for both $D_s^+ \to K^-K^+K^+$ and $D^+ \to K^-K^+K^+$ can be seen in Figs. 4.10.

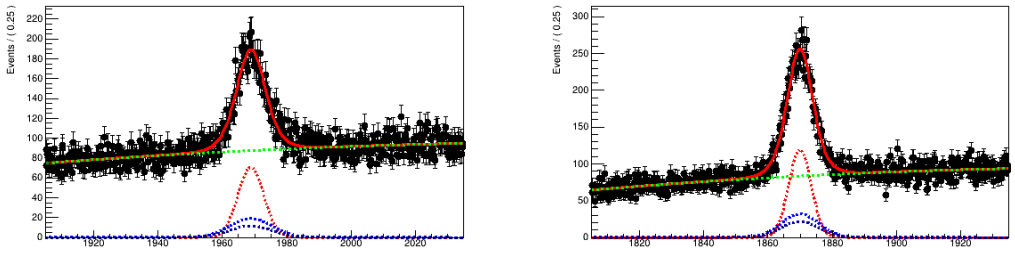


Figure 4.10: sPlot fits on the 50 thousand events samples of $D_s^+ \to K^-K^+K^+$ (left) and $D^+ \to K^-K^+K^+$ (right).

Of these 50 thousand events training samples, half is used for actual training and half for testing, using 10 discriminatory variables which pertain mostly to the mother and symmetric contributions of the daughters: IP, IPχ^2, Combination logIP, FD, log(FDχ^2), Pointing, Vertex χ^2, DIRA, p_T and p.

The BDT responses for the testing sample for both channels can be seen in Fig. 4.11. These responses are then applied in the 2016–2018 samples after all previous selection requirements, excluding the 50 thousand candidates used for training. To determine the BDT classifier requirement, we also look at the significance in combination with efficiency for different values of the BDT requirement, in Figs. 4.12 and 4.13. The requirement on the value of the BDT classifier was applied at BDT> -0.02 for $D_s^+ \to K^-K^+K^+$, with an efficiency of over 84%, and at BDT> -0.04 for $D^+ \to K^-K^+K^+$, over 88% efficient.

This last step concludes our offline selection, summarised in Table 4.5.

4.4 Final signal sample sizes

The invariant mass of the final samples are fitted to derive their yields. Using a Gaussian and two Crystal Balls as the signal PDF,

$$\mathcal{P}_{signal}(m) = f_G \times G(\mu, \sigma_G) + (1 - f_G) \times$$
$$[f_{CB} \times CB_1(\mu, R_1\sigma_G, \alpha_1, N_1) + (1 - f_{CB}) \times CB_2(\mu, R_2\sigma_G, \alpha_2, N_2)], \quad (4\text{-}5)$$

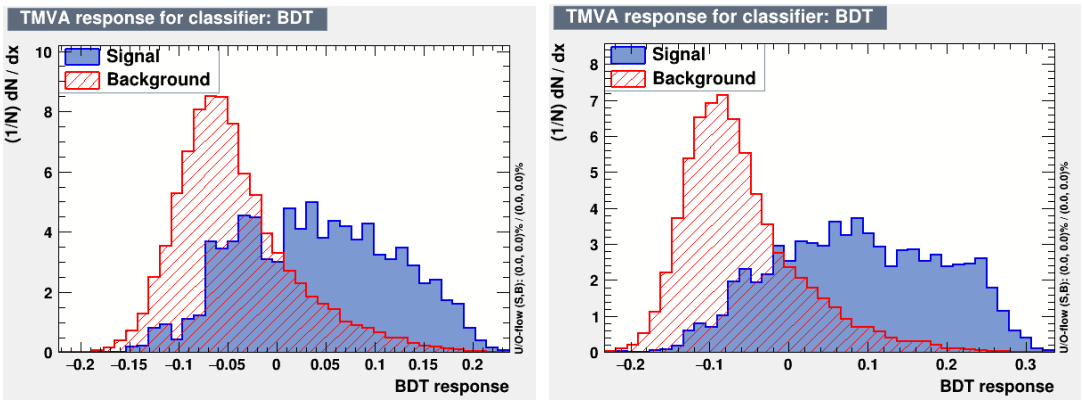


Figure 4.11: BDT response on the training sample of $D_s^+ \to K^-K^+K^+$ (left) and $D^+ \to K^-K^+K^+$ (right).

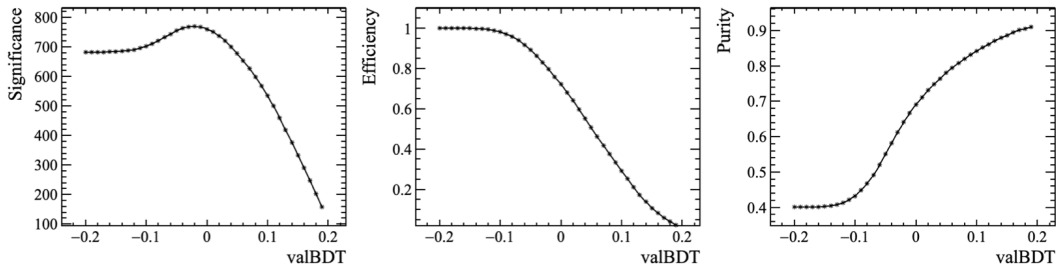


Figure 4.12: $D_s^+ \to K^-K^+K^+$ analysis of FoM (left), efficiency (middle) and purity (right) for BDT requirement.

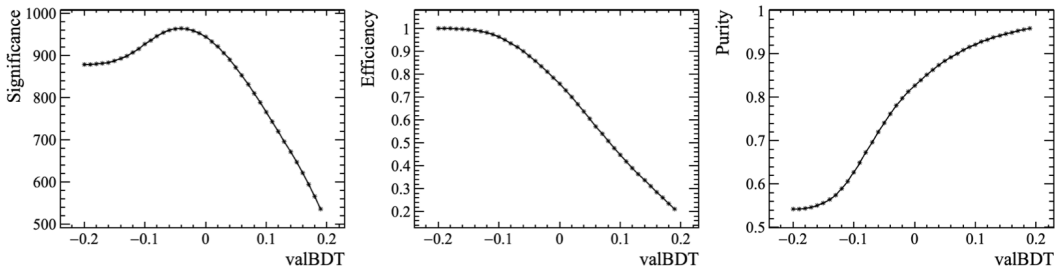


Figure 4.13: $D^+ \to K^-K^+K^+$ analysis FoM (left), efficiency (middle) and purity (right) for BDT requirement.

Table 4.5: Summary of offline requirements.

Requirements	$D_s^+ \to K^-K^+K^+$	$D^+ \to K^-K^+K^+$
Pre-selection	nSPDHits < 1000	
	$1.5 < \eta_{\text{products}} < 5$	
	$p_{\text{products}} < 100\,\text{GeV}$	
	IP$\chi^2 < 15$	
	$1905 < M < 2035$	$1805 < M < 1935$
PID	PIDK$_{2,3} > 15$	ProbNNk$_{1,2,3} > 0.6$
Clone	difTX23$> 2.3 \times 10^{-3}$|| difTY23 $> 2.3 \times 10^{-3}$	
Fiducial	$p_{z_{1,2,3}} > 1 + 3.57 \times \lvert p_{x_{1,2,3}} \rvert$ $p_{z_{1,2,3}} > 4.25 \times \lvert p_{y_{1,2,3}} \rvert$ $p_{z_{1,2}} > 4$	
BDT	BDT> -0.02	BDT> -0.04

with fixed Crystal Ball parameters from the Monte Carlo sample (see Appendix A), and a third-order Bernstein polynomial as the background PDF ($n = 3$),

$$\mathcal{P}_{bkg}(m) = \sum_{i=0}^{n} a_i \binom{n}{i} m^i \cdot (1-m)^{n-i}, \tag{4-6}$$

the fits can be seen in Figs. 4.14 and 4.15.

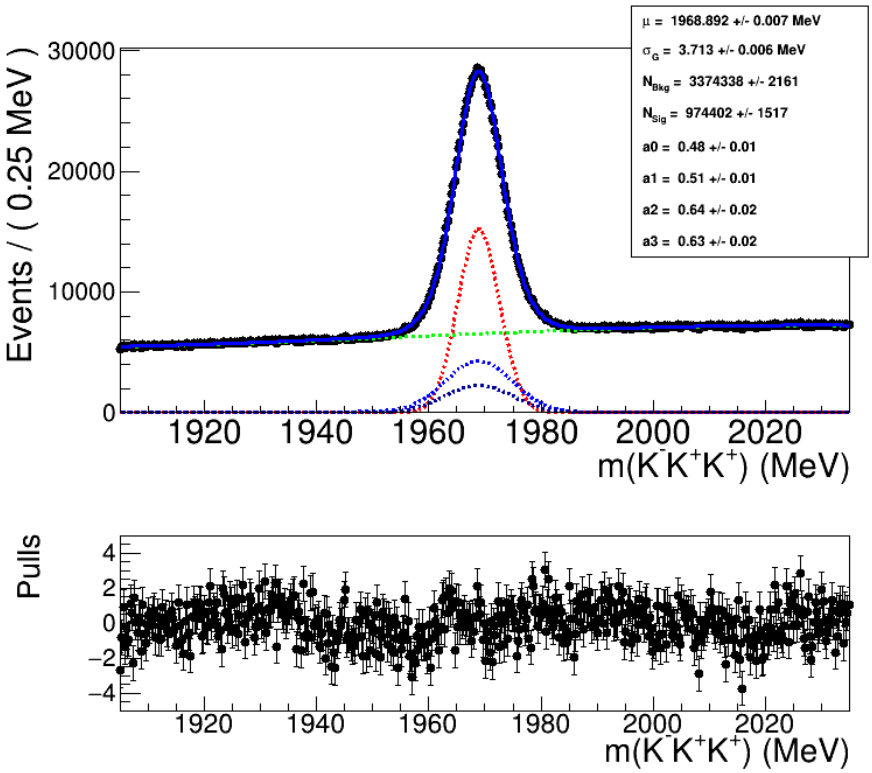


Figure 4.14: Final sample invariant mass fit of $D_s^+ \to K^-K^+K^+$.

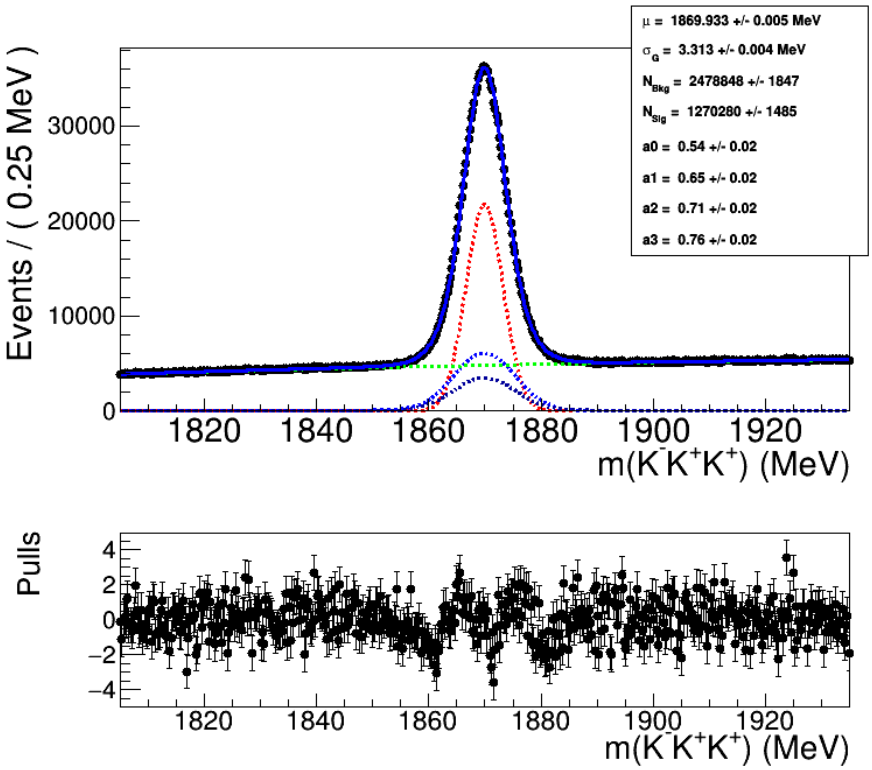


Figure 4.15: Final sample invariant mass fit of $D^+ \to K^-K^+K^+$.

From the fitted standard deviation σ_G, we can derive an effective one, σ_{eff},

$$\sigma_{eff} = \sqrt{f_G \times \sigma_G^2 + (1 - f_G) \times (R_1\sigma_G)^2 + (1 - f_G) \times (1 - f_1) \times (R_2\sigma_G)^2}\,, \tag{4-7}$$

and get the yields within a $2\sigma_{eff}$ signal region. For $D_s^+ \to K^-K^+K^+$, $\sigma_{eff} = 4.823 \pm 0.005$, defining the signal region as $(1959.258 - 1978.550)$ MeV , with a purity of $(64.44 \pm 0.04)\%$. For $D^+ \to K^-K^+K^+$, $\sigma_{eff} = 4.412 \pm 0.004$, defining the signal region as $(1861.108 - 1878.756)$ MeV , with a purity of $(77.77 \pm 0.02)\%$. The yields, given by individual fits of each charge done separately, are presented in Table 4.6.

Table 4.6: Signal and background yields of $D_s^\pm \to K^\mp K^\pm K^\pm$ and $D^\pm \to K^\mp K^\pm K^\pm$.

	Full spectrum yields		Signal region yields		
	Background	Signal	Background	Signal	Purity
D_s^+	$1\,708\,255 \pm 1\,456$	$489\,098 \pm 951$	$255\,589 \pm 218$	$458\,473 \pm 892$	$(64.21 \pm 0.05)\%$
D_s^-	$1\,665\,855 \pm 1\,438$	$485\,522 \pm 944$	$248\,673 \pm 215$	$455\,121 \pm 885$	$(64.67 \pm 0.05)\%$
D^+	$1\,253\,105 \pm 1\,249$	$638\,787 \pm 973$	$171\,782 \pm 171$	$598\,777 \pm 912$	$(77.71 \pm 0.03)\%$
D^-	$1\,225\,774 \pm 1\,236$	$631\,455 \pm 967$	$168\,487 \pm 170$	$591\,904 \pm 906$	$(77.84 \pm 0.03)\%$

4.5 Final sample Dalitz plots

Taking the total events within the $2\sigma_{eff}$ signal regions we have ~ 1.5 million candidates for both D_s^+ and D^+ (even though the D^+ decay is more suppressed, the D_s^+ selection is much tighter in order to remove the $D^+ \to K^-K^+\pi^+$ contamination, leaving similar final sample sizes). We may then look at their Dalitz plots, seen in Fig. 4.16 in terms of s_{high} and s_{low} (folded Dalitz plots).

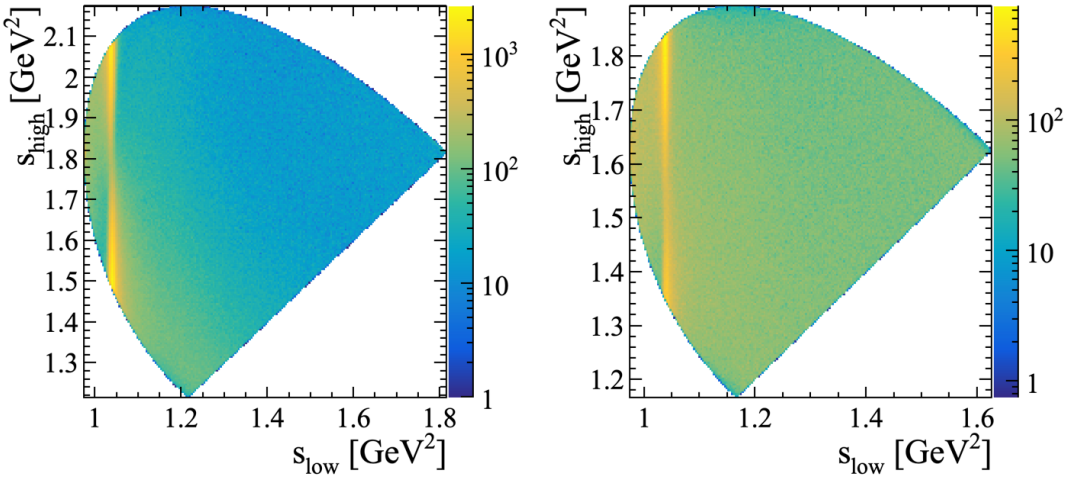


Figure 4.16: Dalitz plots of the final sample $2\sigma_{eff}$ signal regions of $D_s^+ \to K^-K^+K^+$ (left) and $D^+ \to K^-K^+K^+$ (right). The Z-axis containing the number of events is displayed in log scale for better visualisation of the interference pattern.

In the $D^+ \to K^-K^+K^+$ Dalitz plot, we can see a clear contribution from the vector resonance $\phi(1020)$, while outside this region the distribution of events is mostly uniform. Observing the $\phi(1020)$ resonance band more carefully, we note that the upper lobe has a higher density of events than the lower lobe, which indicates the interference between this vector resonance and the S-wave (scalar contributions). Furthermore, looking closely at the left side of the Dalitz, we note a slight linear concentration of events near the edge, which would indicate a resonance such as the $f_0(980)$ lying just outside the Dalitz boundaries. Indeed, the amplitude analysis for this channel, in [60], shows this decay is well parametrised by the coherent sum of $\phi(1020)K^+$ and two scalar amplitudes, $f_0(980)K^+$ and the broad $f_0(1370)K^+$.

In the $D_s^+ \to K^-K^+K^+$ Dalitz plot, we also see the $\phi(1020)$ band clearly but here the distribution of events is not as uniform, being more concentrated below 1.2 GeV^2 in s_{low}. We may also note, as in the $D^+ \to K^-K^+K^+$ Dalitz,

a contribution which could arise from the $f_0(980)$ lying just left of the $\phi(1020)$ band. From the interference pattern of the $\phi(1020)$ resonance, where we observe the left of the upper lobe to be more populated than the right, and, in contrast, the right of the lower lobe to be more populated than the left, we can also conclude that there is some broad scalar contribution. For this decay channel, however, no amplitude analysis is available, and more details on the Dalitz plot resonance model will be discussed in Sec. 5.2.1.

4.6 Candidate selection for control channels

We use the Cabibbo-favoured decays $D_s^+ \to K^-K^+\pi^+$ and $D^+ \to K^-\pi^+\pi^+$, for which no CP violation effects are expected within the SM, as our control channels. With these channels, we may check for eventual production asymmetry effects for both D_s^+ and D^+, as well as verify possible detection asymmetry effects of the kaon daughters. Production asymmetries may arise because the initial state (proton-proton collisions) is not charge symmetric; as a consequence, for example, $D^-(\bar{c}d)$ are slightly more often produced than $D^+(c\bar{d})$. Detection asymmetries occur since the detector may introduce charge asymmetries, as well as the effect of K^+/K^- different interaction with the material .

We look at both control channels passing through the selection criteria of both of our main channels, with the exception of the BDT requirements which are not needed given the high purity of the samples.

These channels have their own exclusive Turbo lines, `Hlt2CharmHadDspToKmKpPip, Hlt2CharmHadDpToKmPipPip`, the only difference in criteria in the requirements shown in Table 4.7.

Table 4.7: HLT2 selection criteria for control channels (black) different from signal channels (red).

Combination requirements	
IPχ^2	> 50 (10) (at least one tracks)
IPχ^2	> 10 (4) (at least two tracks)
Mother requirements	
Track vertex χ^2/DOF	< 6 (10)

We separate 10 samples of 10 million events out of their 2016 datasets, half MagUp polarity and half MagDown, for both control channels with the trigger level requirements and the additional pre-selection requirements of Table 4.2, with a further IP $\chi^2 < 12$ requirement for the $D_{(s)}^+$ candidate.

These 10 million event samples go through the specific requirements of the offline selection of each of the signal channels, applying the PID requirements only to their K daughters. For the $D_s^+ \to K^-K^+\pi^+$ channel, an additional requirement with almost 100% efficiency is applied to remove the specific mis-ID contribution of $D^* \to D^0(\to K^-K^+)\pi^+$, at $m_{K^-K^+} < 1.84$ GeV.

Finally, we look at the yields in the full mass spectra and in the signal regions, now of 40 MeV, as seen in Fig. 4.17, defined in the ranges $1850 - 1890$ MeV and $1950 - 1990$ MeV , for $D^+ \to K^-\pi^+\pi^+$ and $D_s^+ \to K^-K^+\pi^+$, respectively. The average number of events over the 10 samples are presented in Table 4.8.

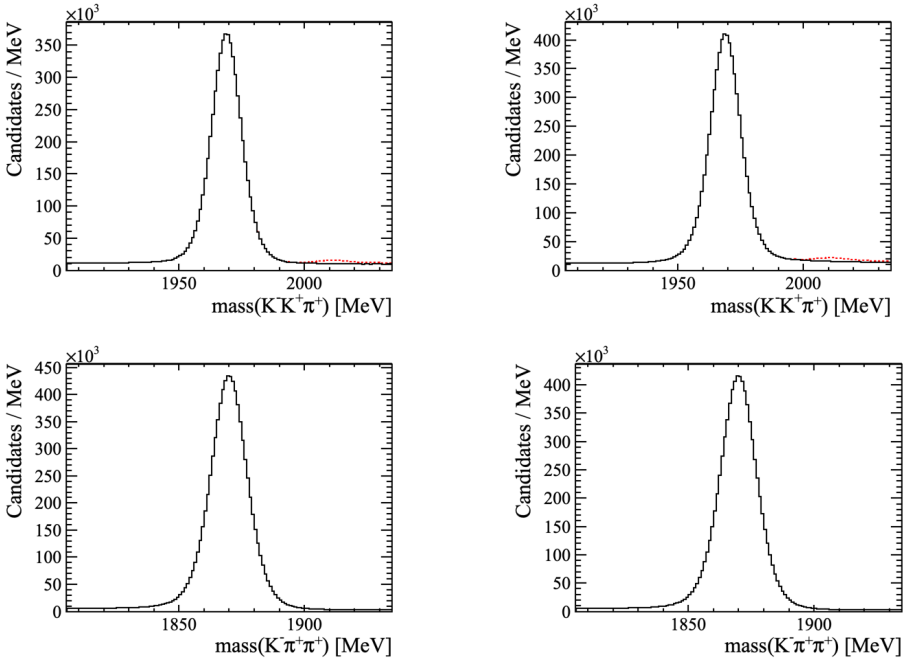


Figure 4.17: Examples of mass distributions of the control channels $D_s^+ \to K^-K^+\pi^+$ (top) and $D^+ \to K^-\pi^+\pi^+$ (bottom) with offline selection requirements of $D_s^+ \to K^-K^+K^+$ (left) and $D^+ \to K^-K^+K^+$ (right). In the $D_s^+ \to K^-K^+\pi^+$ plots, we include the distributions without the $m_{K^-K^+} < 1.84$ GeV requirement (red).

After the final selection, the yields of the control channels are approximately five times larger than those of the signal modes.

Table 4.8: Average number of events in control channels $D_s^+ \to K^-K^+\pi^+$ and $D^+ \to K^-\pi^+\pi^+$.

Selection criteria	Channel		Full spectrum candidates ($\times 10^6$)	Signal region candidates ($\times 10^6$)
$D_s^+ \to K^-K^+K^+$	$D_s^+ \to K^-K^+\pi^+$	+	3.479	2.958
		−	3.499	2.980
	$D^+ \to K^-\pi^+\pi^+$	+	4.258	3.948
		−	4.384	4.070
$D^+ \to K^-K^+K^+$	$D_s^+ \to K^-K^+\pi^+$	+	4.064	3.390
		−	4.096	3.419
	$D^+ \to K^-\pi^+\pi^+$	+	4.071	3.783
		−	4.193	3.902

5
Search for local CP violation in $D^+_{(s)} \to K^-K^+K^+$ decays

5.1
Analysis Strategy

To perform the search for CP violation, a technique, called Mirandizing [61, 62], which performs a statistical comparison between the Dalitz plots of particle and antiparticle, is applied. The Mirandizing technique is model-independent, i.e. it does not rely on parametrisations of the Dalitz plots of the decays, and with it we get information on local effects. The strategy consists of dividing the Dalitz plot in bins and, for each bin, a significance on the difference of $D^+_{(s)}$ and $D^-_{(s)}$ candidates is computed by

$$\mathcal{S}^i_{CP} = \frac{N^i(D^+_{(s)}) - \alpha N^i(D^-_{(s)})}{\sqrt{\alpha(\delta^2_{N^i(D^+_{(s)})} + \delta^2_{N^i(D^-_{(s)})})}}\ , \quad \alpha = \frac{N_{\rm tot}(D^+_{(s)})}{N_{\rm tot}(D^-_{(s)})}. \tag{5-1}$$

In the above expression $N^i(D^+_{(s)})$ and $N^i(D^-_{(s)})$ are the numbers of observed decays in the i^{th} bin of the $D^+_{(s)}$ and $D^-_{(s)}$ Dalitz plots, respectively, and $N_{\rm tot}(D^+_{(s)}), N_{\rm tot}(D^-_{(s)})$ the sum over all bins. The associated errors on the number of observed decays are given by $\delta_{N^i(D^\pm_{(s)})}$. The factor $\alpha \equiv N_{\rm tot}(D^+_{(s)})/N_{\rm tot}(D^-_{(s)})$ is introduced in the computation of $\mathcal{S}^i_{CP}$ to account for global asymmetry effects. These effects, like production asymmetry, can lead to an overall charge asymmetry, and, in most cases, is expected to be constant across the Dalitz plot.

If no statistically significant local asymmetries exist, the $\mathcal{S}^i_{CP}$ values are distributed according to a Gaussian distribution with zero mean and unit width. From the distribution of $\mathcal{S}^i_{CP}$, a χ^2 ($\chi^2 = \sum_i (\mathcal{S}^i_{CP})^2$) is obtained, and for the number of degrees of freedom being the number of bins minus one (due to the $D^+_{(s)}/D^-_{(s)}$ normalisation), a p-value can be calculated. The p-value obtained from the χ^2 test represents the probability, for a given number of degrees of freedom, and under the assumption of the null-hypothesis, i.e. no CP violation in this case, of obtaining a χ^2 that is at least as high as the value observed (this derives by comparing the obtained χ^2 value to the known χ^2-distribution for this given number of degrees of freedom) [63]. The p-value can be converted to a statistical significance in terms of the standard deviation between the

samples, σ, using Gaussian statistics. For instance, p-value≈ 0.03 implies the differences are within $\sim 3\sigma$.

This p-value allows us to measure the degree to which we are confident that the $D^+_{(s)}$ and $D^-_{(s)}$ Dalitz plots differ only by statistical fluctuations, with no significant difference that could point to CP violation. In order to imply the observation of CP violation, we require the p-value$< 3\times10^{-7}$, which represents a CP violation effect of more than 5σ. However, if there is no signal of CP violation, an upper limit for the effect cannot be set in a model-independent way. This method allows us only to access whether there is or not sensitivity to charge-asymmetry effects, and, to quantify it, i.e. calculate A_{CP}, further steps would be necessary, and are not the scope of this work.

To ensure any possible indication comes from CP violation and not from eventual nuisance asymmetries, a careful study of these effects is necessary. Nuisance local effects, such as the different production mechanisms for $D^+_{(s)}$ and $D^-_{(s)}$, differences in the reconstruction, selection or trigger efficiencies, left-right detector asymmetries, etc., are investigated through the control channels $D^+_s \to K^-K^+\pi^+$ and $D^+ \to K^-\pi^+\pi^+$. The study of nuisance asymmetries is also performed in the backgrounds of the signal channels, to ensure they do not introduce any charge-asymmetric effects to the data.

The quantities $N^i(D^+_{(s)})$ and $N^i(D^-_{(s)})$ and their uncertainties are determined in two ways in this analysis. Following the "standard" Mirandizing method, as it was first introduced and has been previously used in other works [64–66], we first consider these quantities as being the *total* number of entries in the $D^+_{(s)}$ and $D^-_{(s)}$ Dalitz plot bins within their signal regions (within $2\sigma_{eff}$), that is, including background events. In this case, the uncertainties are taken simply as the square root of the number of candidates in each bin.

Given the low purity after the selection and, on the other hand, a great loss of efficiency if the selection criteria are tightened to improve the purity, we consider a variation of the method, in which we take $N^i(D^+_{(s)})$ and $N^i(D^-_{(s)})$ and their uncertainties as the signal yields from invariant-mass fits over the full $K^-K^+K^+$ mass spectrum in each Dalitz plot bin for each charge. With this method, we are able to remove the effect of the background both locally and globally, in such a way as to account for variations of the signal and background shapes across the phase space. The signal shapes for each bin are obtained and fixed using Monte Carlo samples (see Appendix A), and the background parameters are left free for all bins.

For the standard Mirandizing method, we look at two types of binning schemes; uniform and physics motivated. The uniform binning schemes are such that the phase space is divided into equal-sized bins. The physics

motivated schemes were designed by hand as to take into account the possible local effects introduced by the spin-1 $\phi(1020)$ resonance, therefore separating the two regions around the $\phi(1020)$ resonance's nominal mass and lobes, while having roughly the same number of signal events per bin of the Dalitz plot. To estimate the number of signal candidates per bin, background subtraction is used, aiming at maintaining similar signal yields per bin. Given how narrow the distribution of the $\phi(1020)$ resonance is, bin division needs to be done by hand to be able to carefully separate the regions of interest. The binning schemes are such as to have two pairs of a uniform one and a physics motivated one with the same total number of bins, using a 5×5 grid with 21 total bins, and a 8×8 grid with 50 total bins. For the fit-per-bin method, we use only the physics motivated binning schemes, which are, as will be shown later, the most sensitive to CP violation effects, as well as the fact that for the uniform schemes some bins have negligible signal yields. These binning schemes can be seen in Figs. 5.1 and 5.2, where we display the folded Dalitz plots after all selection requirements.

This analysis is being performed blinded, i.e. performing all necessary steps and checks without actually looking at the result for the search for CP violation in the signal channels until there is absolute confidence in the method responses. Blinded analyses ensure the final result is not driven by *experimenter's bias* [67].

5.2 Sensitivity studies

In order to verify the technique sensitivity given our current sample sizes, we perform studies using toy Dalitz plots (pseudo-experiments). For these studies, the only available simulated variables are the Dalitz variables (no $K^-K^+K^+$ mass spectrum), and therefore we can only test the standard Miranda procedure. In order to generate Dalitz simulations, three components are needed: a resonance model, to describe the interference pattern of the Dalitz; an efficiency model, to include reconstruction and selection effects that reduce the acceptance across regions of the phase space; and a background model, so that the toys can properly reproduce the considered data sample.

5.2.1 Dalitz plot models

The Dalitz simulations are generated using the Laura++ software [68]. For $D^+ \to K^-K^+K^+$, an amplitude analysis, i.e. an analysis of the resonant components of the decay, has been done by the LHCb Collaboration [60]. For

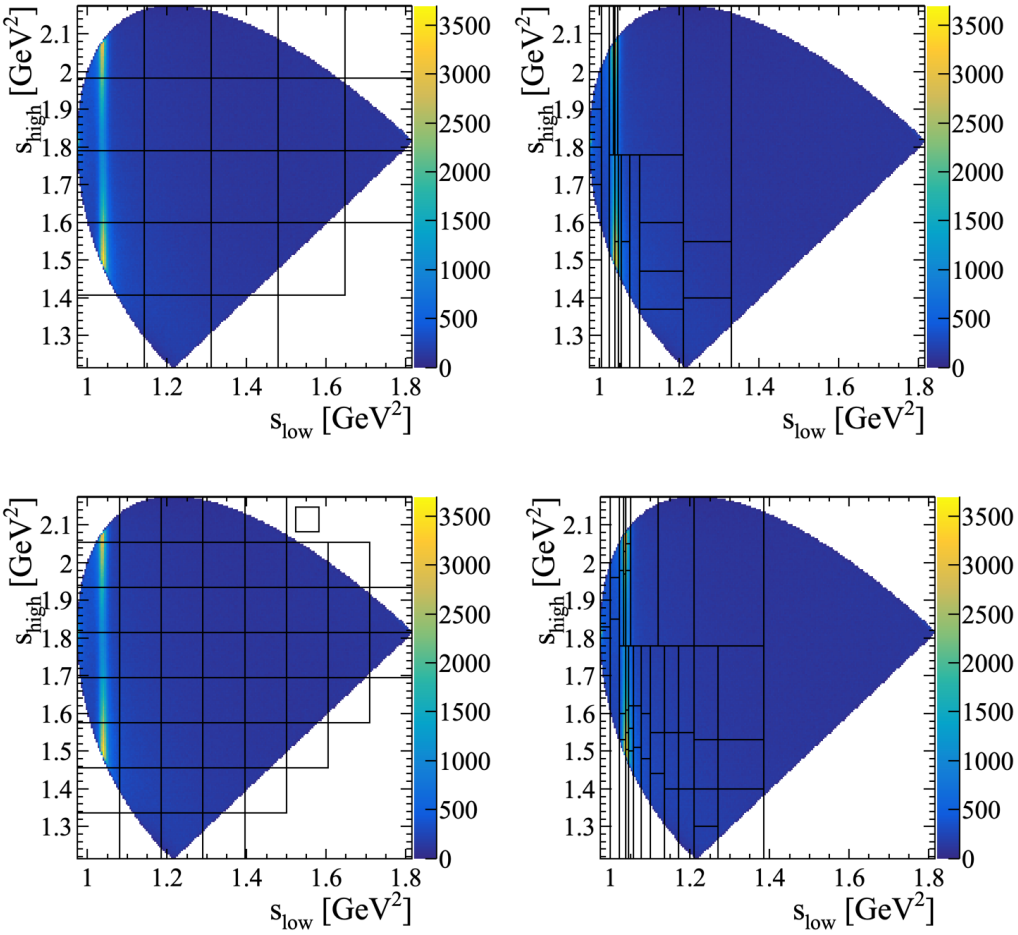


Figure 5.1: $D^+_s \to K^-K^+K^+$ binning schemes plotted over the final sample Dalitz plot. Uniform 5×5 grid (top left), physics motivated with 21 bins (top right), uniform 8×8 grid (bottom left) and physics motivated with 50 bins (bottom right).

$D^+_s \to K^-K^+K^+$, however, this analysis has not been done before, and so it is necessary to perform a Dalitz plot fit to derive an approximate resonant model.

The $D^+ \to K^-K^+K^+$ model used is that of the isobar fit (Eq. 2-35) performed in [60] (Model 3).[1] In this model, with magnitude (a_n) and phase (δ_n) parameters for each resonance shown in Table 5.1, three K^-K^+ resonances were used, the spin-1 $\phi(1020)$ and the spin-0 $f_0(1370)$ and $f_0(980)$. In addition, the $f_0(1370)$ mass and width used were fitted in the amplitude analysis, and the central values $m_{f_0(1370)} = 1.422$ GeV and $\Gamma_{f_0(1370)} = 0.324$ GeV were obtained.

[1] In the paper, the published parameters are given according to the normalisation used by the Rio++ software, which are converted to the Laura++ normalisation by maintaining the phase, although converting from degrees to radians, and using for the resonances' magnitudes $a_n = \sqrt{\frac{f_n}{f_\phi}}$, related to their fit fractions.

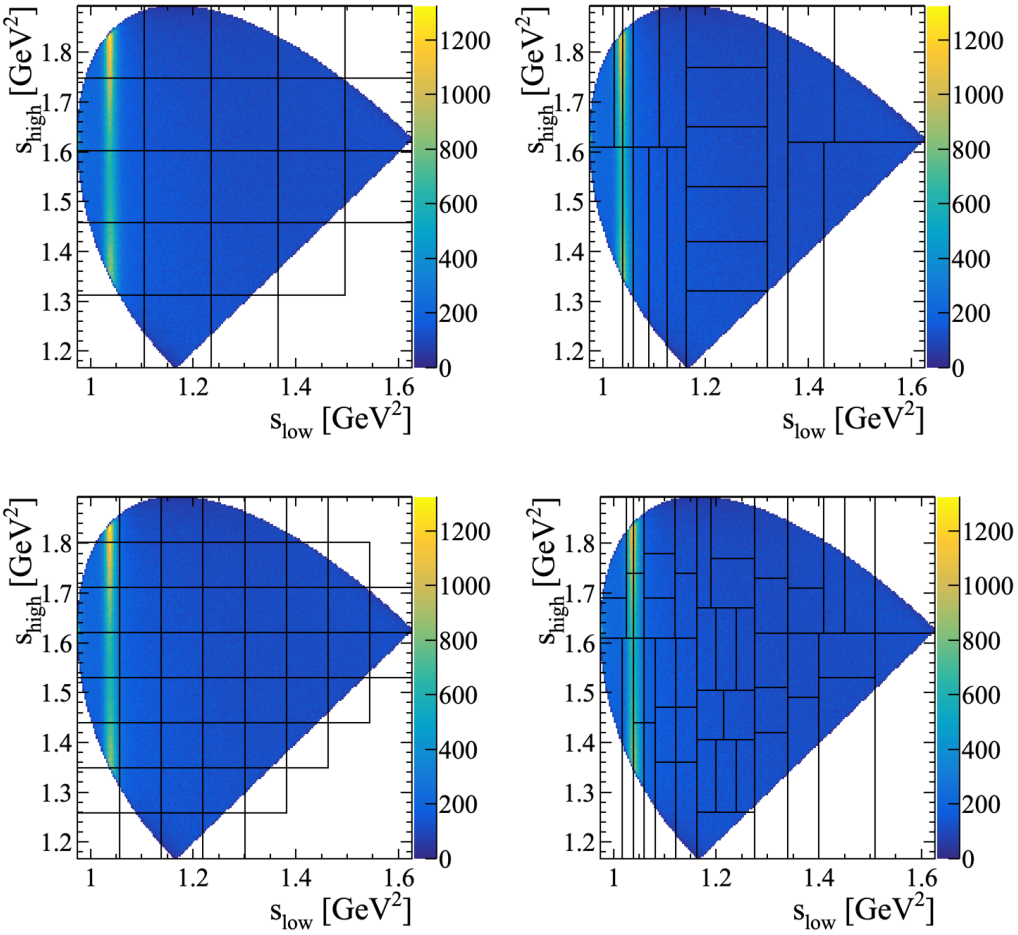


Figure 5.2: $D^+ \to K^-K^+K^+$ binning schemes plotted over the final sample Dalitz plot. Uniform 5×5 grid (top left), physics motivated with 21 bins (top right), uniform 8×8 grid (bottom left) and physics motivated with 50 bins (bottom right).

Table 5.1: $D^+ \to K^-K^+K^+$ isobar model parameters, from [60].

Component	Magnitude	Phase (rad)
$\phi(1020)K^+$	1.0	0.0
$f_0(980)K^+$	1.955	-1.028
$f_0(1370)K^+$	2.024	0.229

Since for the $D_s^+ \to K^-K^+K^+$ no amplitude analysis is available, we have performed fits to the Dalitz plot of a sample of 100 thousand events from the signal region (Table 4.6) of our final dataset, using Laura++ as the fitter. The baseline model used consists of an isobar model with the three resonances obtained in the $D^+ \to K^-K^+K^+$ model, also fixing the same mass and width parameters for $f_0(1370)$. Other variations were tested, but no gain in fit quality was obtained (more details can be seen in Appendix B). The parameters of the isobar model can be seen in Table 5.2.

Table 5.2: $D^+_s \to K^-K^+K^+$ isobar model parameters, from Dalitz fit.

Component	Magnitude	Phase (rad)
$\phi(1020)K^+$	1.0	0.0
$f_0(980)K^+$	1.908	-2.252
$f_0(1370K^+)$	1.037	2.432

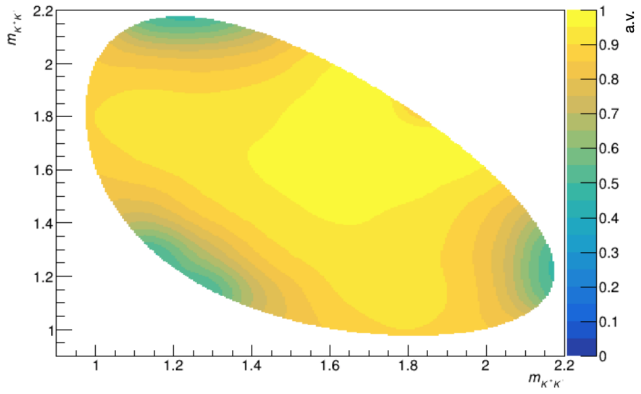


Figure 5.3: $D^+_s \to K^-K^+K^+$ smoothed efficiency map.

The efficiency maps for each channel, which are used in toy generation and for the $D^+_s \to K^-K^+K^+$ fit, are obtained from the full LHCb Monte Carlo samples (discussed in Appendix A). After applying the selection criteria used in data to the MC samples, which already include detector effects, the phase space (which does not have any resonant model) is divided by a completely flat phase space to correct for bin edge effects near the borders, and smoothed. As an example, the efficiency for $D^+_s \to K^-K^+K^+$ can be seen in Fig. 5.3.

The background models are taken from the $K^-K^+K^+$ mass sideband events from the data samples projected onto the folded Dalitz plots. For $D^+ \to K^-K^+K^+$, we use the candidates in 1820−1830 MeV and 1910−1920 MeV, and for $D^+_s \to K^-K^+K^+$, in 1920−1930 MeV and 2010−2020 MeV.

The toys are generated according to the sample sizes observed in the $2\sigma_{eff}$ signal regions of our data, given in Table 4.6. Comparisons between the s_{high} and s_{low} invariants for the final data sample and the generated toys can be seen in Figs. 5.4 and 5.5, for $D^+_s \to K^-K^+K^+$ and $D^+ \to K^-K^+K^+$ respectively, showing a good agreement.

5.2.2 Study of *p*-value bias

We perform an analysis of the bias of p-value responses due to the global correction factor α in Eq. 5-1. The background, which has a different distribution across the Dalitz plot from the signal, is considered in the calculation of this factor for the standard method, and can therefore lead

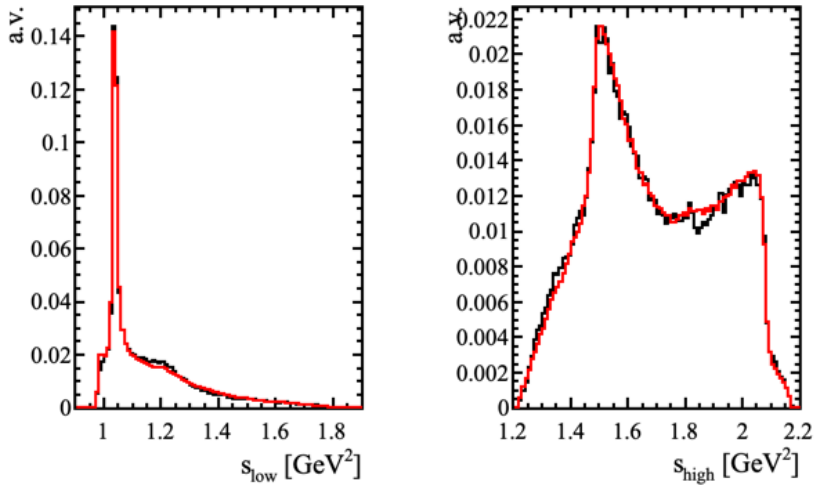


Figure 5.4: Comparison of data (black) to model (red) for $D^+_s \to K^-K^+K^+$.

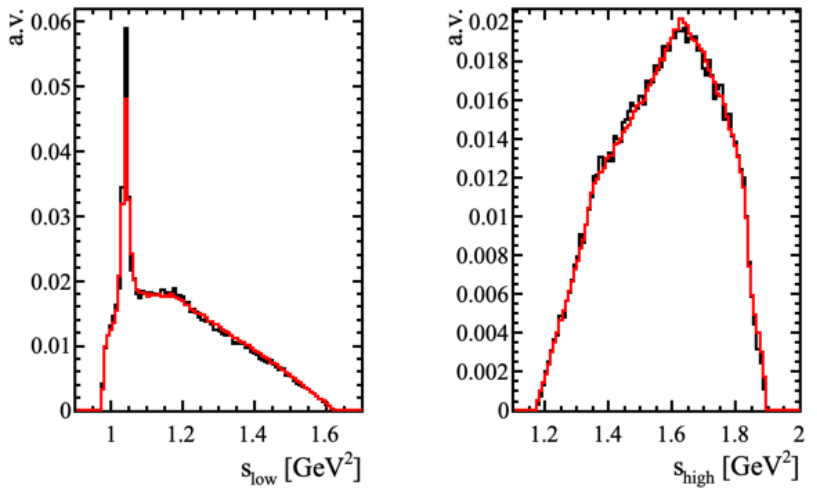


Figure 5.5: Comparison of data (black) to model (red) for $D^+ \to K^-K^+K^+$.

to a fake asymmetry, specially when the purity of the samples is low. To better illustrate this effect, we consider the following example: let there be CP-conserving $D^+_{(s)}$ and $D^-_{(s)}$ samples, with a production asymmetry such that we have more $D^+_{(s)}$ than $D^-_{(s)}$ signal events, but having both samples with the same number of background events and distributions across the Dalitz plots, and therefore having different purities among these sample. Then, when applying the standard Mirandizing method, α takes into account the total number of events, and corrects both signal and background contributions to $\mathcal{S}_{CP}$; in doing so for the background, however, it may potentially generate a fake asymmetry, as no correction was initially needed, with the effect being larger for higher background levels or lower purities.

For this test, we generate 100 pairs of $D^+_{(s)}$ and $D^-_{(s)}$ toys with the

same Dalitz plot models, the only difference being their number of signal and background events within the $2\sigma_{eff}$ regions, presented in Table 4.6. If there is no bias, the distribution of p-values is expected to come out as uniform between the values of 0 and 1, with a mean of 0.5. This test represents, therefore, a null-test for the standard Mirandizing method.

The distributions of p-values for these toys can be seen on Figs. 5.6 and 5.7, for $D^+_s \to K^-K^+K^+$ and $D^+ \to K^-K^+K^+$, respectively.

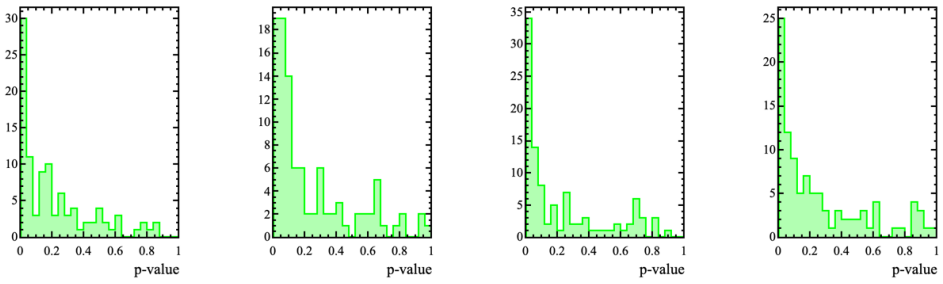


Figure 5.6: $D^+_s \to K^-K^+K^+$ p-value distributions for uniform 5×5 grid, uniform 8×8 grid, physics motivated of 21 bins, physics motivated of 50 bins binning schemes for 100 pairs of toys generated with no model difference between them, using the statistics of the data sample within $2\sigma_{eff}$.

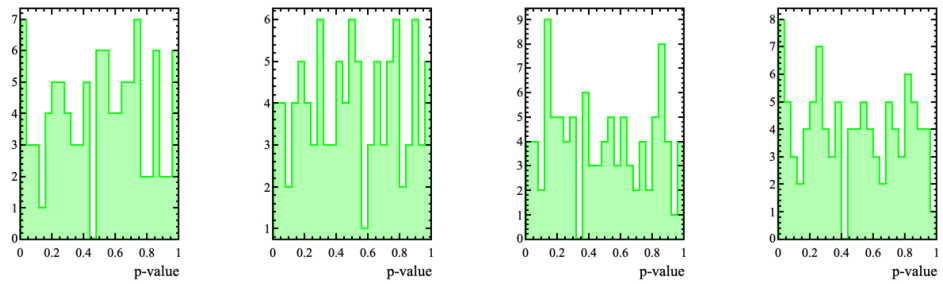


Figure 5.7: $D^+ \to K^-K^+K^+$ p-value distributions for uniform 5×5 grid, uniform 8×8 grid, physics motivated of 21 bins, physics motivated of 50 bins binning schemes for 100 pairs of toys generated with no model difference between them, using the statistics of the data sample within $2\sigma_{eff}$.

We conclude that for $D^+ \to K^-K^+K^+$ no bias is found. For $D^+_s \to K^-K^+K^+$ we see a bias; while none of the toys gave an effect of over 5σ response (p-value $< 3\times10^{-7}$), eight of them still showed very low responses (p-value $< 3\times10^{-3}$). To overcome the bias using the standard Miranda procedure, it would be necessary to reduce the background levels in the data sample, by tightening the BDT requirement. This, however, would imply a significant loss of signal candidates. For instance, to achieve a purity of $\sim 70\%$ in the signal region, more than 15% of signal candidates would be lost. For this reason, the

alternative fit method is analysed in data, in which the background is naturally removed from the $\mathcal{S}_{CP}$ calculation, as discussed in Sec. 5.1.

To perform a study of the sensitivity to CP violation while ensuring the p-value is not being driven down due to the bias, we use slightly different number of events from the data sample for $D^+_s \to K^-K^+K^+$. By generating the toys of D^+_s and D^-_s with the same purity, we find there to be no bias, as shown in Fig. 5.8. The same number of signal events from data as reported in Table 4.6 is used, but the number of background events was changed to have both D^+_s and D^-_s toys with the total data sample purity of 64.44%.

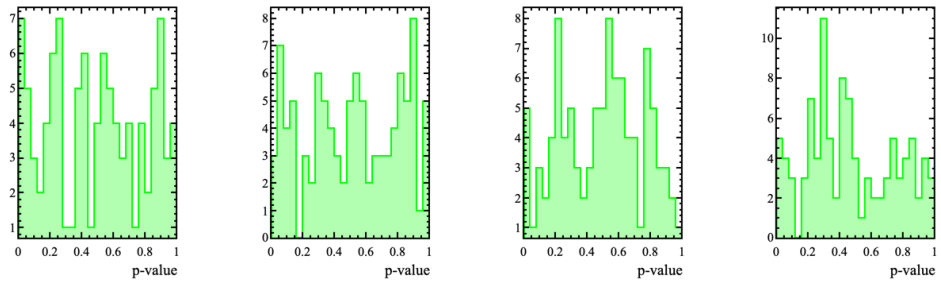


Figure 5.8: $D^+_s \to K^-K^+K^+$ p-value distributions for uniform 5×5 grid, uniform 8×8 grid, physics motivated of 21 bins, physics motivated of 50 bins binning schemes for 100 pairs of toys generated with no model difference between them, using the statistics of the signal yield in the data sample within $2\sigma_{eff}$, but with a purity of 64.44% for both D^+_s and D^-_s toys.

5.2.3 Sensitivity to CP violation

It is not known how CP violation would manifest in the decays. To verify possible CP violation scenarios, we perform sensitivity studies, in which we check what level of differences between the decay amplitudes of particle and antiparticle would lead to observation of CP asymmetries. This is done by generating a $D^+_{(s)}$ Dalitz plot according to the isobar models discussed in the previous section and a $D^-_{(s)}$ Dalitz plot with either a phase difference or a magnitude difference for one of the resonances at a time.

For $D^+ \to K^-K^+K^+$, we generate toys with the number of signal and background events of D^+ and D^- within the signal region of $2\sigma_{eff}$, again using the standard method to look for asymmetries. Given the bias observed in $D^+_s \to K^-K^+K^+$ when using the same number of events of the data sample, we perform the CP violation sensitivity studies using the data signal yields and same purity of 64.44% for D^+_s and D^-_s toys, as explained in the previous section.

Initially, we track the p-value responses to get an idea of where CP violation appears. In Tables 5.3 and 5.4, we see the p-value responses for progressively introducing a phase difference in $D_s^+ \to K^-K^+K^+$ and $D^+ \to K^-K^+K^+$, respectively. In Tables 5.5 and 5.6, we see the responses for a magnitude difference. Through these tables, we see how the p-value decreases as the differences increase, and can see around which values the CP violation signals start showing up.

Table 5.3: $D_s^+ \to K^-K^+K^+$ phase difference sensitivity.

binning	phase diff (^{o})	channels $\phi(1020)$	$f_0(980)$	$f_0(1370)$
Uniform 5×5 (21 bins)	1	5.27×10^{-1}	4.03×10^{-4}	8.55×10^{-2}
	2	3.46×10^{-4}	1.19×10^{-5}	4.01×10^{-2}
	5	5.45×10^{-8}	1.28×10^{-33}	8.34×10^{-29}
	7	5.11×10^{-12}	3.43×10^{-59}	1.88×10^{-56}
Uniform 8×8 (50 bins)	1	2.63×10^{-1}	4.84×10^{-4}	1.55×10^{-1}
	2	2.65×10^{-1}	3.21×10^{-6}	2.67×10^{-1}
	5	9.35×10^{-9}	4.41×10^{-43}	4.22×10^{-31}
	7	8.28×10^{-20}	1.88×10^{-77}	8.09×10^{-59}
Physics mot. 21 bins	1	6.03×10^{-1}	3.01×10^{-3}	5.30×10^{-2}
	2	8.04×10^{-2}	7.21×10^{-6}	3.05×10^{-3}
	5	2.29×10^{-16}	3.59×10^{-45}	2.37×10^{-19}
	7	5.48×10^{-32}	2.40×10^{-76}	6.83×10^{-38}
Physics mot. 50 bins	1	2.04×10^{-1}	1.32×10^{-2}	1.58×10^{-1}
	2	2.71×10^{-1}	4.71×10^{-6}	8.71×10^{-3}
	5	4.58×10^{-13}	8.61×10^{-50}	4.24×10^{-23}
	7	1.52×10^{-30}	2.95×10^{-80}	1.95×10^{-55}

Table 5.4: $D^+ \to K^-K^+K^+$ phase difference sensitivity.

binning	phase diff (^{o})	channels $\phi(1020)$	$f_0(980)$	$f_0(1370)$
Uniform 5×5 (21 bins)	1	6.21×10^{-1}	2.05×10^{-1}	4.35×10^{-1}
	5	6.19×10^{-13}	1.01×10^{-3}	1.07×10^{-1}
	10	3.48×10^{-35}	1.88×10^{-21}	1.99×10^{-4}
	15	5.56×10^{-79}	4.77×10^{-80}	5.35×10^{-9}
Uniform 8×8 (50 bins)	1	6.83×10^{-1}	4.92×10^{-1}	3.86×10^{-1}
	5	2.47×10^{-11}	5.78×10^{-2}	4.16×10^{-2}
	10	1.74×10^{-36}	4.21×10^{-17}	3.50×10^{-3}
	15	1.48×10^{-83}	9.83×10^{-71}	5.71×10^{-7}
Physics mot. 21 bins	1	5.20×10^{-1}	7.11×10^{-1}	7.30×10^{-1}
	5	1.43×10^{-15}	7.30×10^{-6}	2.07×10^{-3}
	10	2.32×10^{-51}	1.60×10^{-23}	5.77×10^{-12}
	15	1.71×10^{-134}	1.34×10^{-97}	2.12×10^{-20}
Physics mot. 50 bins	1	5.55×10^{-1}	5.43×10^{-1}	3.65×10^{-1}
	5	7.96×10^{-16}	2.89×10^{-3}	3.11×10^{-2}
	10	3.72×10^{-58}	5.78×10^{-24}	4.18×10^{-6}
	15	1.05×10^{-139}	1.16×10^{-103}	6.26×10^{-19}

Individual responses of these searches are affected by statistical fluctuations. Then, to be certain of the ranges in which we become sensitive to CP violation, we generate 100 toys for some key values of phase or magnitude differences and analyse their p-value distributions. An example can be seen in Fig. 5.9, and the final results, for which at least the two physics motivated binning schemes responded with over $\sim 90\%$ of p-values under 3×10^{-7}

Table 5.5: $D^+_s \to K^-K^+K^+$ magnitude difference sensitivity.

binning	mag. diff (%)	channels $\phi(1020)$	$f_0(980)$	$f_0(1370)$
Uniform 5×5 (21 bins)	1	5.96×10^{-1}	5.50×10^{-1}	1.14×10^{-1}
	2	2.98×10^{-5}	3.12×10^{-4}	8.38×10^{-4}
	5	5.48×10^{-17}	3.23×10^{-60}	4.58×10^{-21}
	7	1.37×10^{-27}	4.07×10^{-127}	4.62×10^{-36}
Uniform 8×8 (50 bins)	1	9.90×10^{-1}	2.73×10^{-1}	1.81×10^{-1}
	2	1.10×10^{-2}	2.54×10^{-6}	7.37×10^{-2}
	5	1.30×10^{-15}	1.15×10^{-87}	2.12×10^{-21}
	7	8.60×10^{-41}	1.13×10^{-160}	1.84×10^{-29}
Physics mot. 21 bins	1	6.43×10^{-1}	1.62×10^{-1}	6.51×10^{-2}
	2	5.52×10^{-4}	3.73×10^{-8}	1.62×10^{-5}
	5	3.40×10^{-24}	4.18×10^{-86}	8.46×10^{-18}
	7	2.38×10^{-66}	3.63×10^{-164}	7.72×10^{-23}
Physics mot. 50 bins	1	7.09×10^{-1}	3.28×10^{-2}	2.37×10^{-1}
	2	3.61×10^{-4}	3.46×10^{-10}	8.18×10^{-3}
	5	2.71×10^{-24}	1.05×10^{-89}	6.71×10^{-15}
	7	6.27×10^{-63}	1.34×10^{-177}	1.38×10^{-21}

Table 5.6: $D^+ \to K^-K^+K^+$ magnitude difference sensitivity.

binning	mag. diff (%)	channels $\phi(1020)$	$f_0(980)$	$f_0(1370)$
Uniform 5×5 (21 bins)	1	8.25×10^{-1}	2.34×10^{-1}	1.35×10^{-1}
	5	5.46×10^{-4}	7.19×10^{-6}	3.92×10^{-6}
	10	8.83×10^{-21}	5.45×10^{-1}	1.13×10^{-16}
	15	9.69×10^{-56}	6.26×10^{-11}	9.95×10^{-37}
Uniform 8×8 (50 bins)	1	7.94×10^{-1}	3.94×10^{-1}	5.72×10^{-1}
	5	1.33×10^{-7}	4.62×10^{-5}	1.09×10^{-4}
	10	2.38×10^{-25}	3.76×10^{-1}	2.02×10^{-17}
	15	3.08×10^{-69}	3.31×10^{-9}	2.02×10^{-32}
Physics mot. 21 bins	1	2.16×10^{-1}	2.83×10^{-1}	6.66×10^{-1}
	5	4.63×10^{-6}	7.79×10^{-4}	5.73×10^{-8}
	10	1.72×10^{-25}	2.34×10^{-3}	4.07×10^{-17}
	15	8.61×10^{-61}	9.59×10^{-13}	7.49×10^{-38}
Physics mot. 50 bins	1	9.64×10^{-1}	2.97×10^{-1}	4.60×10^{-1}
	5	2.37×10^{-4}	5.88×10^{-5}	1.46×10^{-4}
	10	4.95×10^{-24}	2.06×10^{-2}	4.51×10^{-13}
	15	1.46×10^{-69}	5.21×10^{-13}	2.64×10^{-32}

and therefore we are confident it clearly indicates CP vilation, are displayed in Table 5.7. Examples of the $\mathcal{S}_{CP}$ distributions in the Dalitz plot with either a phase or magnitude difference in the $\phi(1020)$ resonance can be seen in Figs. 5.10 and 5.11, for $D^+_s \to K^-K^+K^+$ and $D^+ \to K^-K^+K^+$, respectively.

We conclude that overall the physics motivated schemes are more sensitive than the uniform schemes and the sensitivity improves in having less bins. The values for phase and magnitude differences for the sensitivity to CP violation are, however, very high: the level of few degrees and percent, respectively, lead to asymmetries at $\mathcal{O}(10^{-2})$. For $D^+ \to K^-K^+K^+$ there was already no SM prediction of CP violation, but for $D^+_s \to K^-K^+K^+$, so far, the best predictions could lead to asymmetries at $\mathcal{O}(10^{-3})$, as discussed previously. This implies that any clear signal of CP violation in either of the two channels with our current data sample would be due to NP phenomena. Or, in other words,

our sensitivity studies show that with our sample sizes it is not possible to observe CP violation effects coming from the SM.

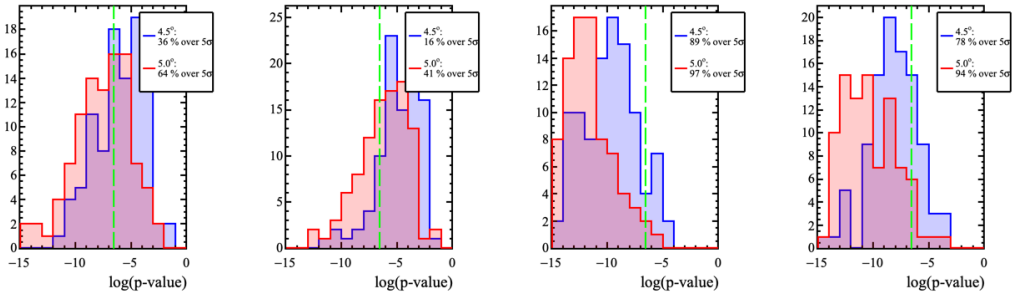


Figure 5.9: $D^+ \to K^-K^+K^+$ p-value distributions for 100 toys generated with a phase difference of either 4.5° (blue) or 5.0° (red) in the $\phi(1020)$ resonance. The green line represents the p-value corresponding to a 5σ observation of CP violation.

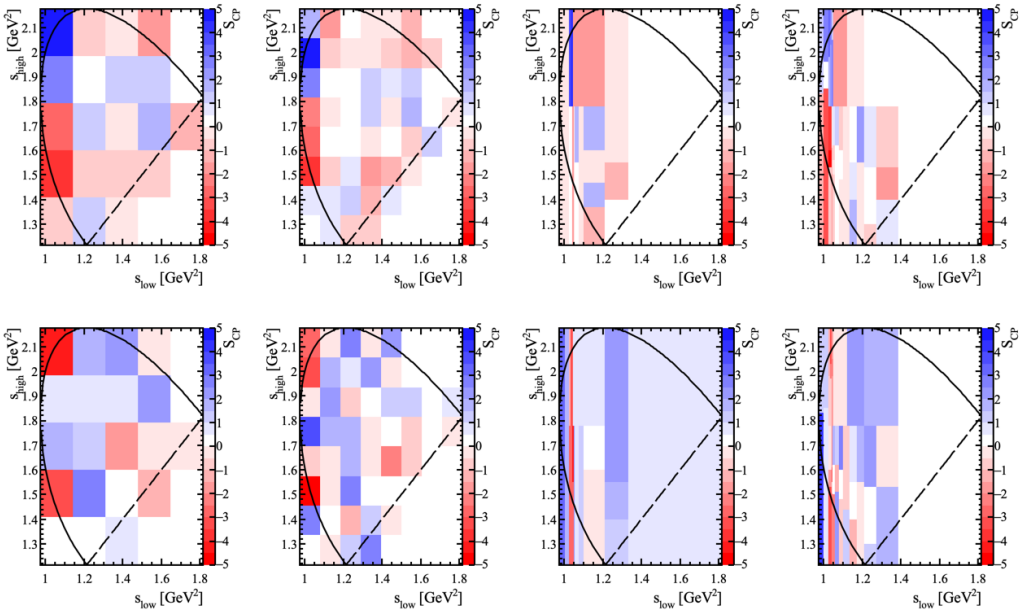


Figure 5.10: $D^+_s \to K^-K^+K^+$ examples of $\mathcal{S}_{CP}$ distributions for toys with a phase difference of 5.0^o (top) or a magnitude difference of 3.5% (bottom) in the $\phi(1020)$ resonance.

Table 5.7: Ranges of phase or magnitude differences between toys in which we are sensitive to CP violation.

decay	difference	channels		
		$\phi(1020)$	$f_0(980)$	$f_0(1370)$
$D^+_s \to K^-K^+K^+$	phase	4.5^o – 5.0^o	2.5^o – 3.0^o	4.0^o – 4.5^o
	magnitude	3.0% – 3.5%	2.0% – 2.5%	4.5% – 5.0%
$D^+ \to K^-K^+K^+$	phase	4.5^o – 5.0^o	6.5^o – 7.0^o	11.0^o – 11.5^o
	magnitude	7.0% – 7.5%	13.0% – 13.5%	9.0% – 9.5%

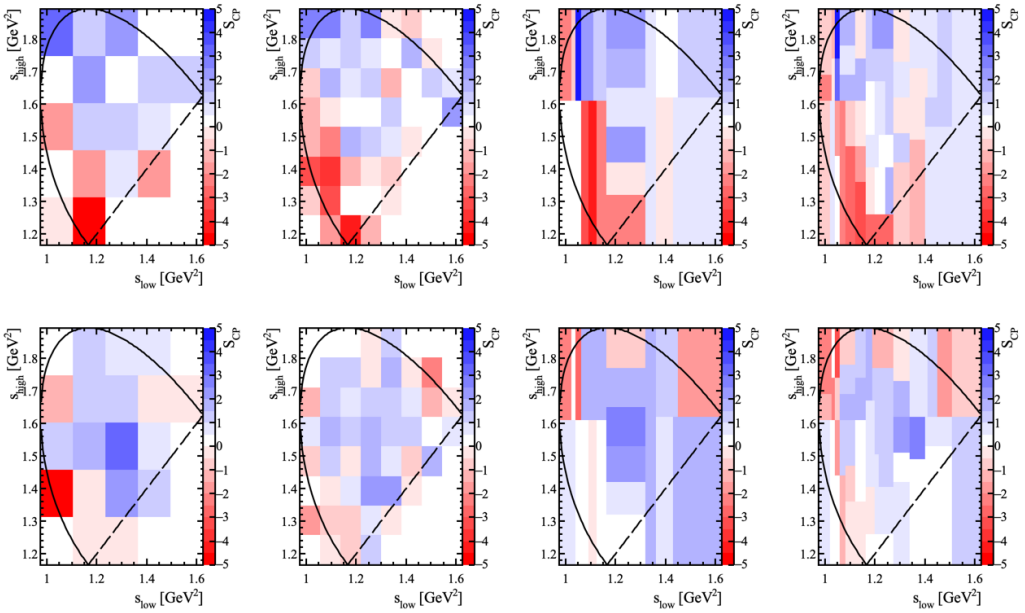


Figure 5.11: $D^+ \to K^-K^+K^+$ examples of $\mathcal{S}_{CP}$ distributions for toys with a phase difference of 5.0^o (top) or a magnitude difference of 7.5% (bottom) on the $\phi(1020)$ resonance.

5.3 Systematic checks

Systematic checks are used to verify how our analysis methods act on data while remaining blinded to CP violation results. First, we perform a null-test in data to ensure our ability to perform invariant mass fits for each bin of the physics motivated binning schemes and validate this alternative method. Next, we look at the background of our signal channels and the signal of control channels to ensure we are not sensitive to nuisance asymmetries.

5.3.1 Null-test: validation of methods

A null-test is performed in data to validate the fit-per-bin method, which is needed as an alternative to the standard Mirandizing method for removing the background effects, and consequently the bias in the p-value response. With this test, we are able to ensure we have good fit models for the different bins of the physics motivated binning schemes and compare the response to that of the standard method. For this test, we take the final data samples of both $D^+_s \to K^-K^+K^+$ and $D^+ \to K^-K^+K^+$ and divide each of them randomly in two parts. Each part has the total number of events that we have in data for $D^+_{(s)}$ and $D^-_{(s)}$, but making no charge distinction. We then perform a search for local statistical asymmetries with each method, and since there is no charge separation, this implies a null-test.

The sizes of these samples are present in Table 5.8. For the standard method, all events within the signal region are used, while for the fit-per-bin method the signal yields on the whole mass spectra are considered.

Table 5.8: Number of events of samples used for null-test.

	Sample	Total events	Signal Yield	Events in $2\sigma_{eff}$	Purity in $2\sigma_{eff}$
$D_s^+ \to K^-K^+K^+$	1	2 197 352	461 215 ± 893	716 248	(64.39 ± 0.05)%
	2	2 151 380	452 378 ± 884	701 566	(64.48 ± 0.05)%
$D^+ \to K^-K^+K^+$	1	1 891 898	601 438 ± 914	773 748	(77.80 ± 0.03)%
	2	1 857 227	589 244 ± 905	758 474	(77.75 ± 0.03)%

The results of the null-test for both methods are presented in Tables 5.9 and 5.10, for $D_s^+ \to K^-K^+K^+$ and $D^+ \to K^-K^+K^+$ respectively. The $\mathcal{S}_{CP}$ distributions for each channel can be seen in Figs. 5.12 and 5.13.

Table 5.9: Null-test for $D_s^+ \to K^-K^+K^+$.

binning		Standard χ^2/ndf	Standard p-value (%)	Fit per bin χ^2/ndf	Fit per bin p-value (%)
Uniform	5×5	27.6/20	12.0	—	—
	8×8	59.5/49	14.5	—	—
Phys. Mot.	21	13.9/20	83.4	12.1/20	91.4
	50	42.0/49	75.2	50.7/49	40.8

Table 5.10: Null-test for $D^+ \to K^-K^+K^+$.

binning		Standard χ^2/ndf	Standard p-value (%)	Fit per bin χ^2/ndf	Fit per bin p-value (%)
Uniform	5×5	28.0/20	11.0	—	—
	8×8	49.9/49	43.9	—	—
Phys. Mot.	21	15.6/20	74.0	16.6/20	67.8
	50	50.2/49	42.5	53.7/49	29.9

These responses are all compatible with the null-hypothesis. Given the similar purity of the two samples for both $D_s^+ \to K^-K^+K^+$ and $D^+ \to K^-K^+K^+$, this was expected.

Through this test we are able to confirm we have good fits for data in all regions of the Dalitz plot. From the $D^+ \to K^-K^+K^+$ tests, where purity is higher, we see that the p-value response and $\mathcal{S}_{CP}$ distribution, in Fig. 5.13, for both methods are overall similar. This gives us confidence in the fit-per-bin method and therefore verifies that it is a reliable solution to the low purity problem without having to lose signal events. Some examples of fits per bin in different regions of the physics motivated binning scheme of 21 bins can be seen

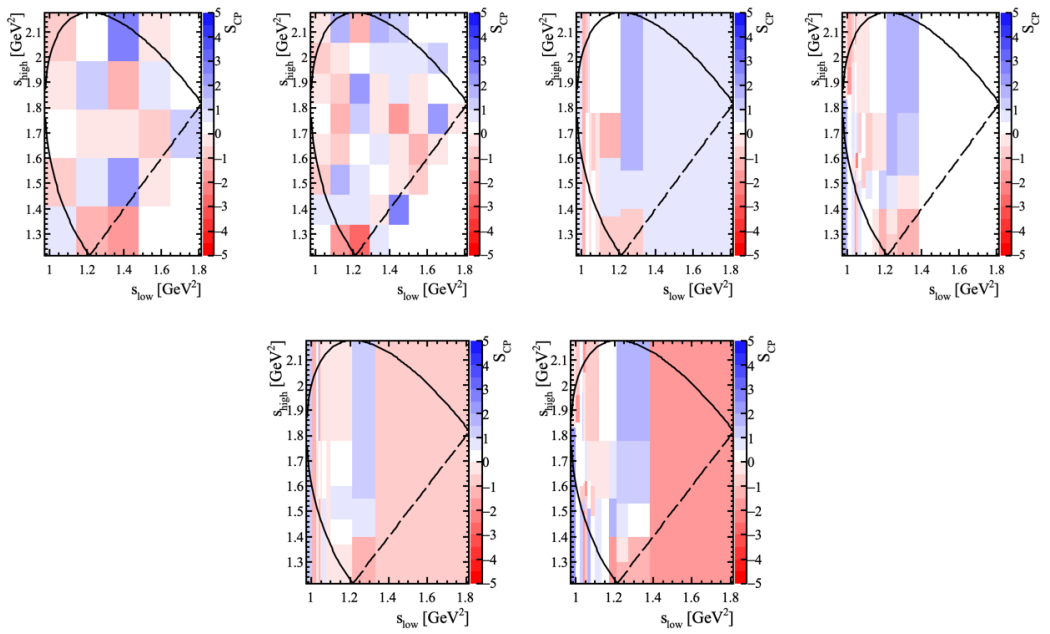


Figure 5.12: $D_s^+ \to K^-K^+K^+$ null-test using standard (top) and fit per bin (bottom) Mirandizing methods for uniform 5×5 grid, uniform 8×8 grid, physics motivated of 21 bins, physics motivated of 50 bins.

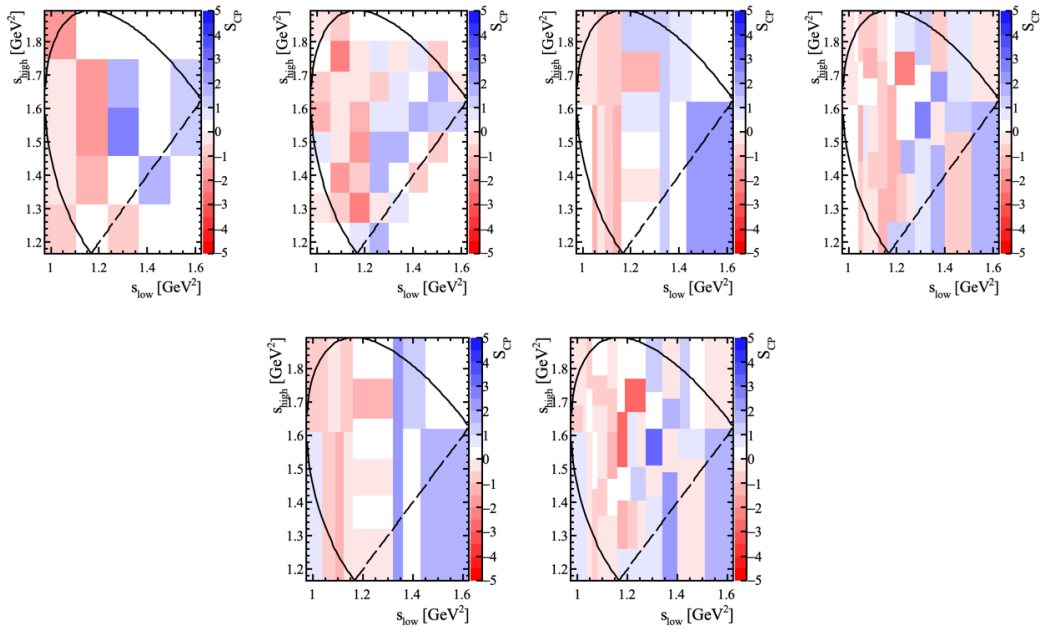


Figure 5.13: $D^+ \to K^-K^+K^+$ null-test using standard (top) and fit per bin (bottom) Mirandizing methods for uniform 5×5 grid, uniform 8×8 grid, physics motivated of 21 bins, physics motivated of 50 bins.

in Figs. 5.14 and 5.15, for $D^+_s \to K^-K^+K^+$ and $D^+ \to K^-K^+K^+$ respectively. All fits for sample 1 can be seen in Appendix C.

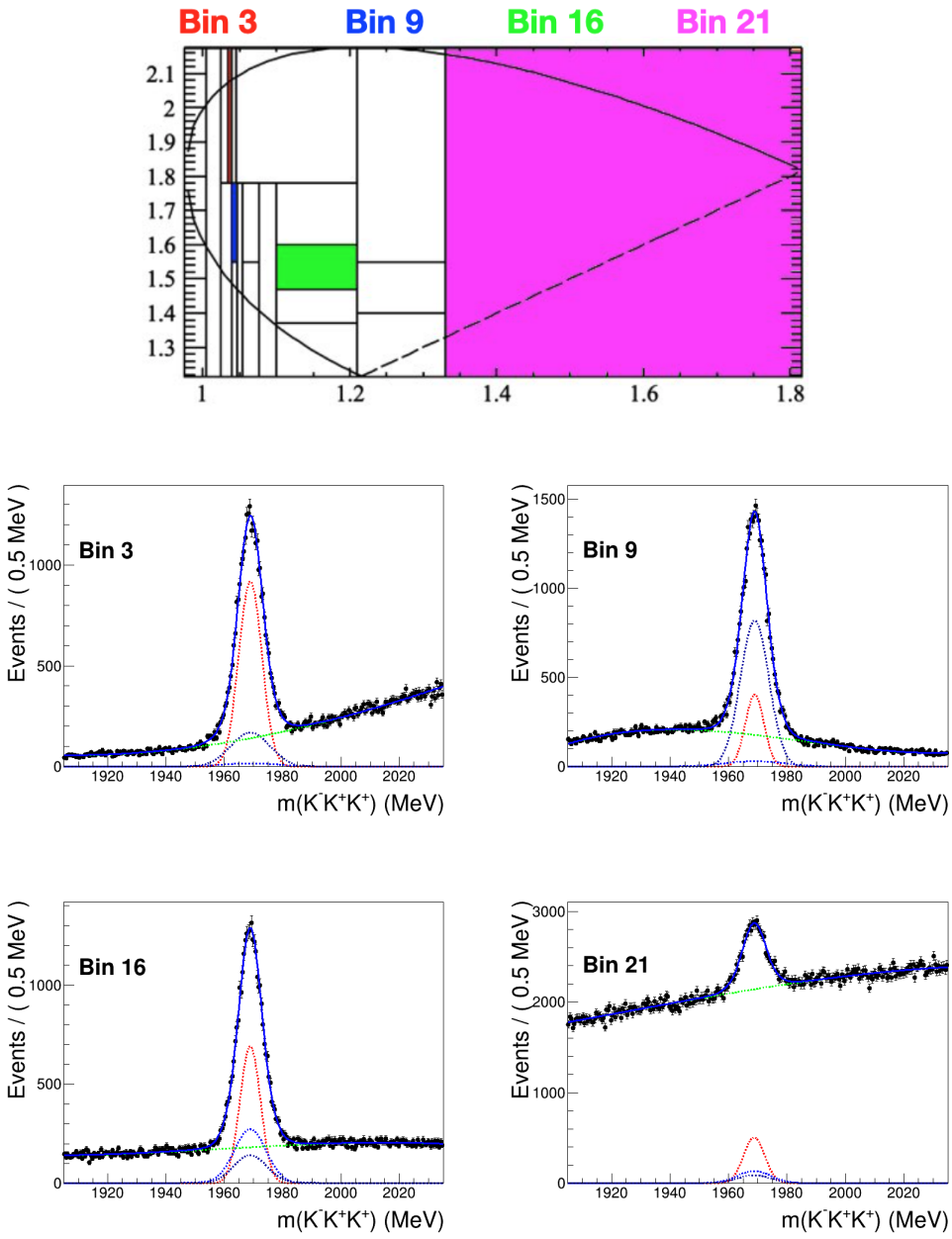


Figure 5.14: $D^+_s \to K^-K^+K^+$ examples of invariant mass fits per bin of the physics motivated binning scheme with 21 total bins in Sample 1.

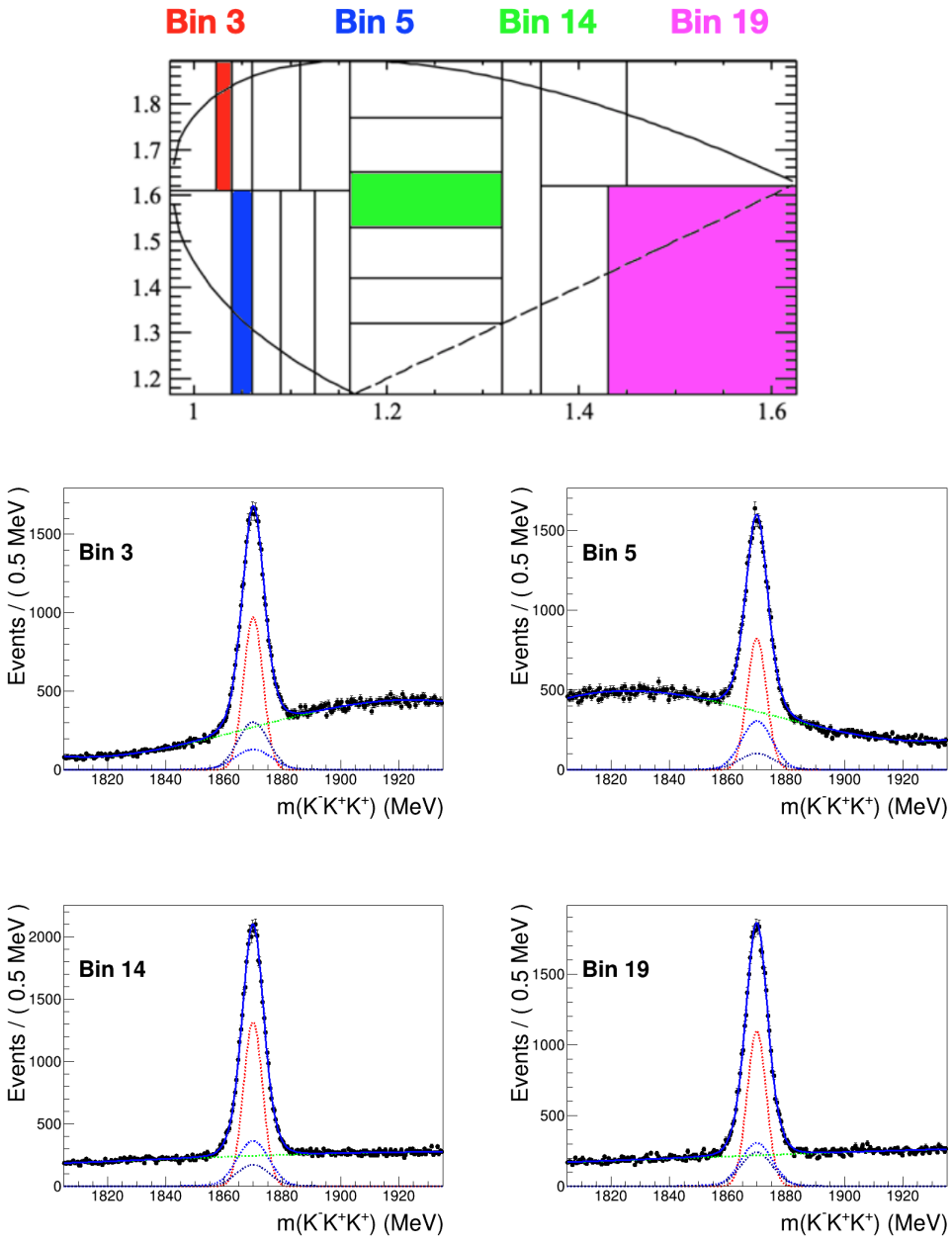


Figure 5.15: $D^+ \to K^-K^+K^+$ examples of invariant mass fits per bin of the physics motivated binning scheme with 21 total bins in Sample 1.

5.3.2 Background

We look for any possible charge asymmetries in the background of the signal channels that could cause a spurious CP violation signal in our searches using the standard method. To do this, we use both the standard Miranda and fit-per-bin methods.

First, we take the events in the same background sidebands as defined in Section 4, which have roughly the same number of background events considered in the standard Miranda method, and perform the analysis. In total, there are 517 485 background events for $D^+_s \to K^-K^+K^+$ and 381 380 for $D^+ \to K^-K^+K^+$.

Then, we consider the full samples to perform fits per bin and calculate the $\mathcal{S}_{CP}$ using the background yields. These fits are performed in such a way that we don't save the signal yield responses and the analysis remains blinded. In total, there are $\sim$ 3 376 500 background events for $D^+_s \to K^-K^+K^+$ and $\sim$ 2 478 900 for $D^+ \to K^-K^+K^+$.

The results of the background analysis are presented in Tables 5.11 and 5.12 and can be seen in Figs. 5.16 and 5.17 for $D^+_s \to K^-K^+K^+$, and Figs. 5.18 and 5.19 for $D^+ \to K^-K^+K^+$. We conclude there is no evidence of local asymmetry in the backgrounds.

Table 5.11: Background check for $D^+_s \to K^-K^+K^+$.

binning		Standard χ^2/ndf	Standard p-value (%)	Fit per bin χ^2/ndf	Fit per bin p-value (%)
Uniform	5×5	9.10/20	98.2	—	—
	8×8	61.9/49	10.3	—	—
Phys. Mot.	21	12.0/20	91.7	20.0/20	46.0
	50	40.3/49	80.6	51.6/49	37.4

Table 5.12: Background check for $D^+ \to K^-K^+K^+$.

binning		Standard χ^2/ndf	Standard p-value (%)	Fit per bin χ^2/ndf	Fit per bin p-value (%)
Uniform	5×5	21.7/20	35.7	—	—
	8×8	45.8/49	60.5	—	—
Phys. Mot.	21	20.0/20	45.7	15.3/20	75.7
	50	44.3/49	66.2	31.1/49	97.8

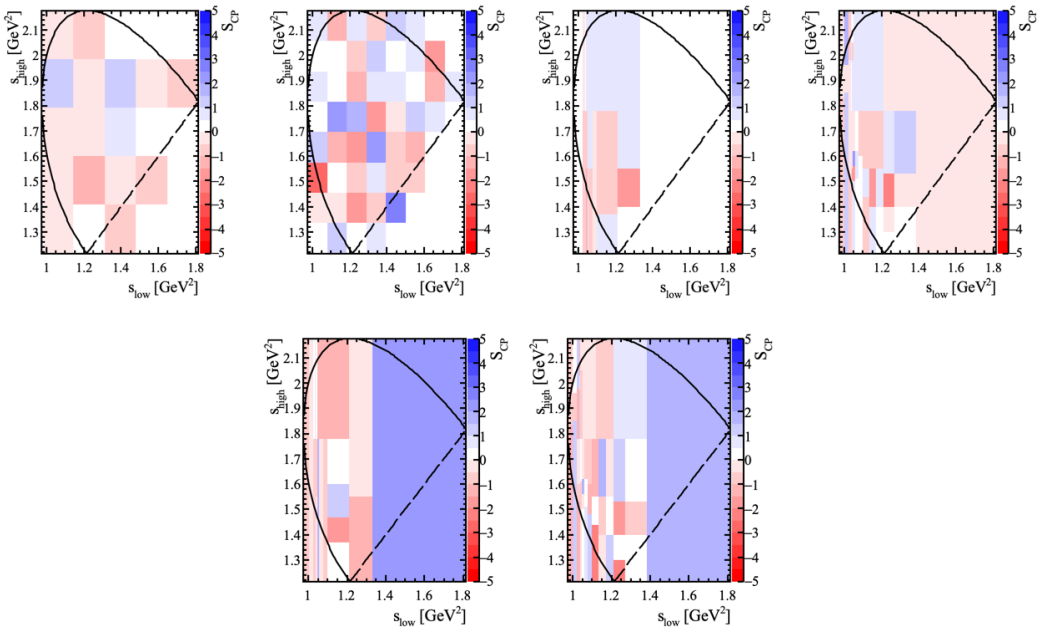


Figure 5.16: Dalitz plot of $D_s^+ \to K^-K^+K^+$ background check using standard (top) and fit per bin (bottom) Mirandizing methods for uniform 5×5 grid, uniform 8×8 grid, physics motivated of 21 bins, physics motivated of 50 bins.

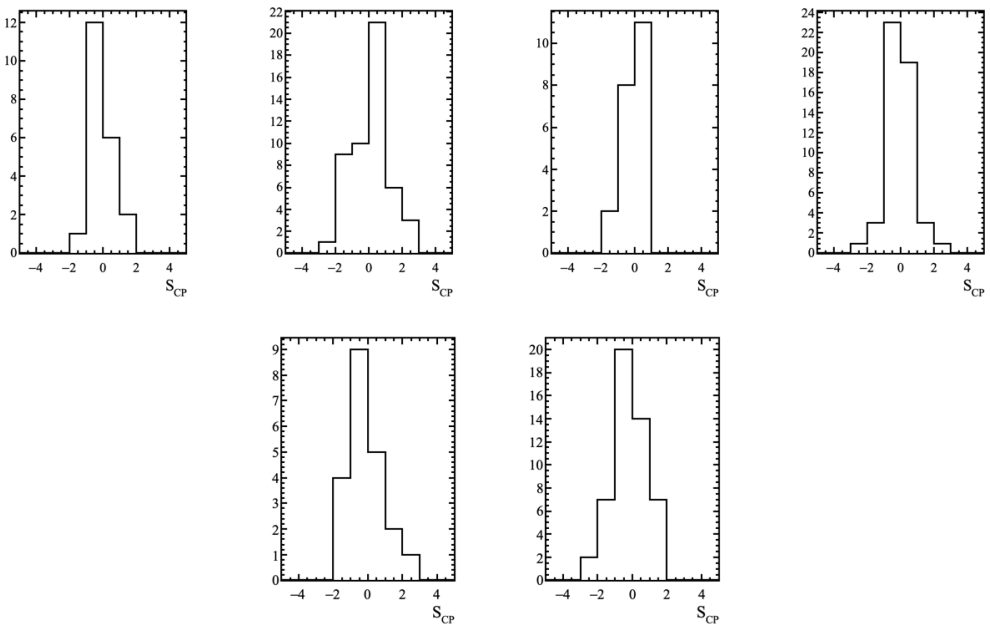


Figure 5.17: $\mathcal{S}_{CP}$ distribution of $D_s^+ \to K^-K^+K^+$ background check using standard (top) and fit per bin (bottom) Mirandizing methods for uniform 5×5 grid, uniform 8×8 grid, physics motivated of 21 bins, physics motivated of 50 bins.

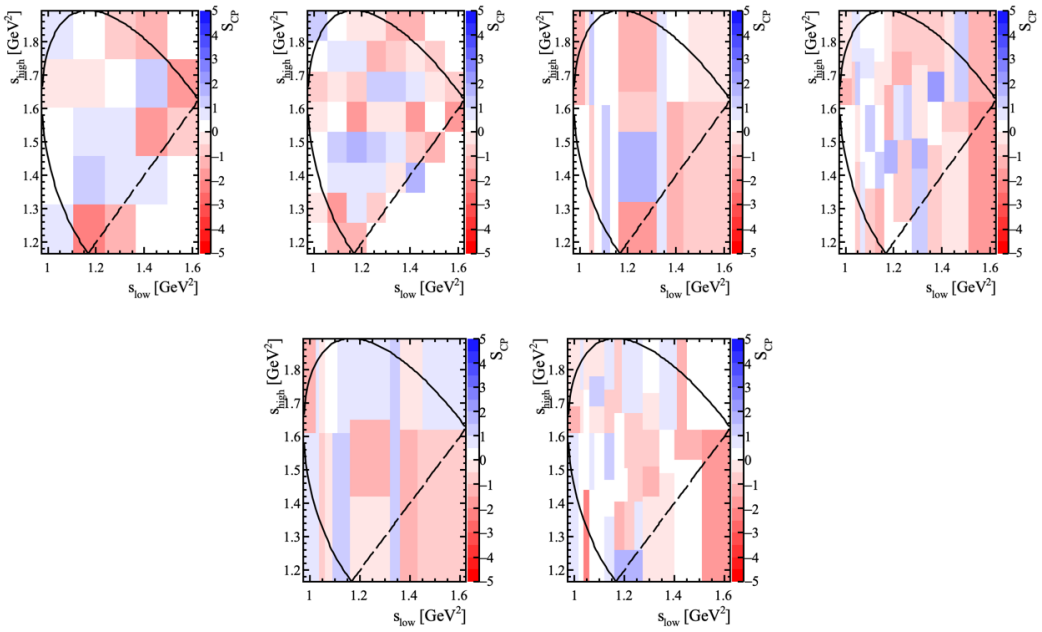


Figure 5.18: Dalitz plot of $D^+ \to K^-K^+K^+$ background check using standard (top) and fit per bin (bottom) Mirandizing methods for uniform 5×5 grid, uniform 8×8 grid, physics motivated of 21 bins, physics motivated of 50 bins.

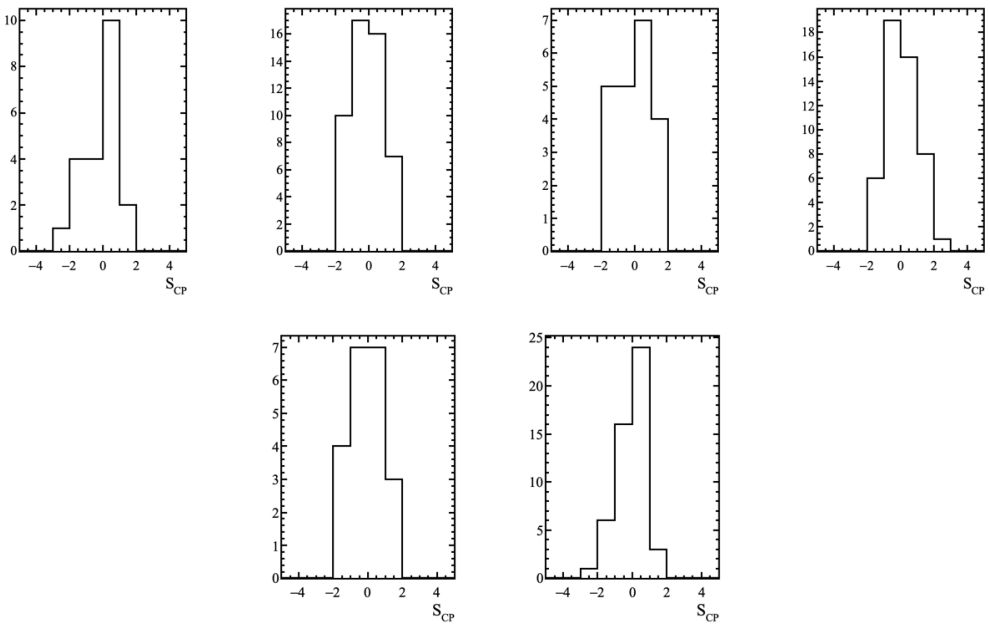


Figure 5.19: $\mathcal{S}_{CP}$ distribution of $D^+ \to K^-K^+K^+$ background check using standard (top) and fit per bin (bottom) Mirandizing methods for uniform 5×5 grid, uniform 8×8 grid, physics motivated of 21 bins, physics motivated of 50 bins.

5.3.3 Control channels

We perform a search for local asymmetries in our control channels, $D^+_s \to K^-K^+\pi^+$ and $D^+ \to K^-\pi^+\pi^+$. To do this, we use the standard method, in which we look at all events (signal+background) within the signal regions, as defined in Section 4. We perform the search in the 10 samples of each control channel with each set of selection criteria, choosing a sample size $\simeq 5\times$ larger than our signal channels, and look at the distribution of the p-value responses. Due to the high purity of these samples, of over 92%, we find there is no need to perform fits per bin in order to remove the background effects. To perform this analysis, we make use of uniform and adaptive binning schemes.

The uniform schemes are the same as the signal channels, with 5×5 and 8×8 grids, corresponding respectively to 20 bins and 46 bins for $D^+_s \to K^-K^+\pi^+$, and 21 bins and 47 bins for $D^+ \to K^-\pi^+\pi^+$. The adaptive binning schemes, where each bin has the same population, are obtained using the ROOT library `TKDTree`, having a total of 21 and 50 bins.

In Figs. 5.20 and 5.21, we present an example of the $\mathcal{S}_{CP}$ distributions for $D^+_s \to K^-K^+\pi^+$ with $D^+_s \to K^-K^+K^+$ selection and $D^+ \to K^-\pi^+\pi^+$ with $D^+ \to K^-K^+K^+$ selection, respectively. The distributions of the p-value responses of the analysis are seen in Figs. 5.22—5.25. In Table 5.13 we present the lowest values found. We conclude there is no indication of charge asymmetry in the control channels for the samples analysed, which are larger than for signal channels.

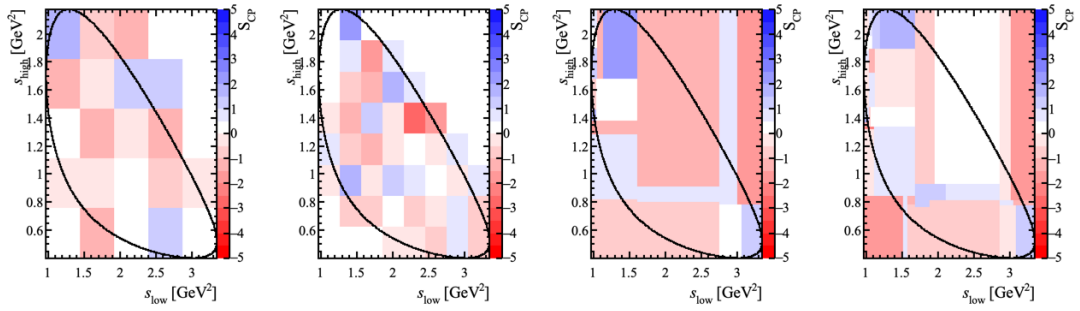


Figure 5.20: $D^+_s \to K^-K^+\pi^+$ with $D^+_s \to K^-K^+K^+$ selection; $\mathcal{S}_{CP}$ distributions in the Dalitz plots of 1 sample.

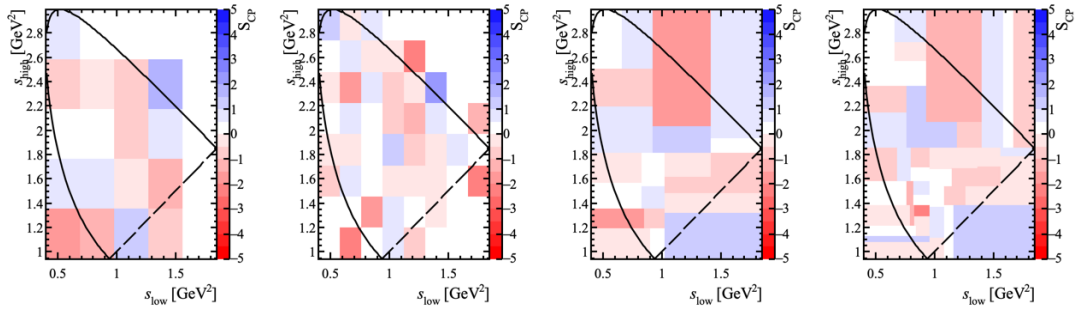


Figure 5.21: $D^+ \to K^-\pi^+\pi^+$ with $D^+ \to K^-K^+K^+$ selection; $\mathcal{S}_{CP}$ distributions in the Dalitz plots of 1 sample.

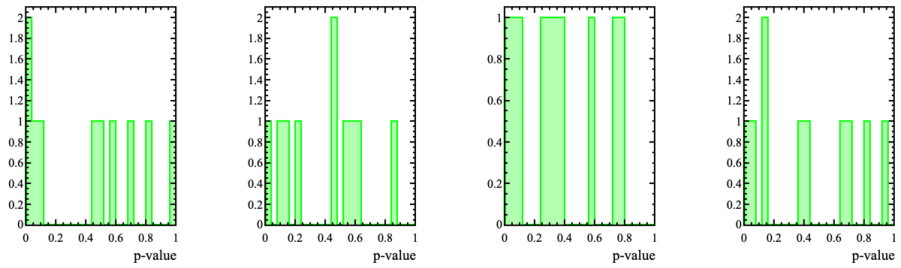


Figure 5.22: $D_s^+ \to K^-K^+\pi^+$ with $D_s^+ \to K^-K^+K^+$ selection; p-value distribution of search over the 10 samples for uniform 5×5 grid, uniform 8×8 grid, adaptive of 21 bins, adaptive of 50 bins.

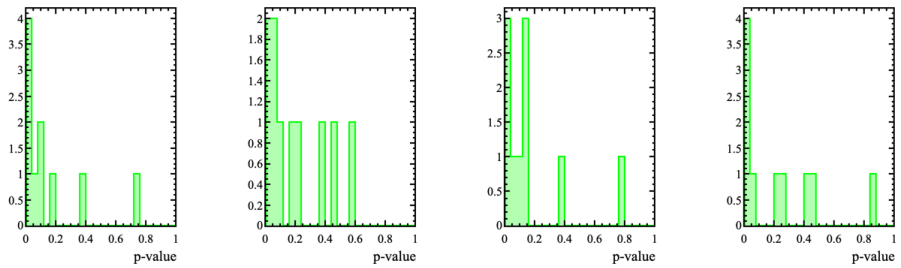


Figure 5.23: $D_s^+ \to K^-K^+\pi^+$ with $D^+ \to K^-K^+K^+$ selection; p-value distribution of search over the 10 samples for uniform 5×5 grid, uniform 8×8 grid, adaptive of 21 bins, adaptive of 50 bins.

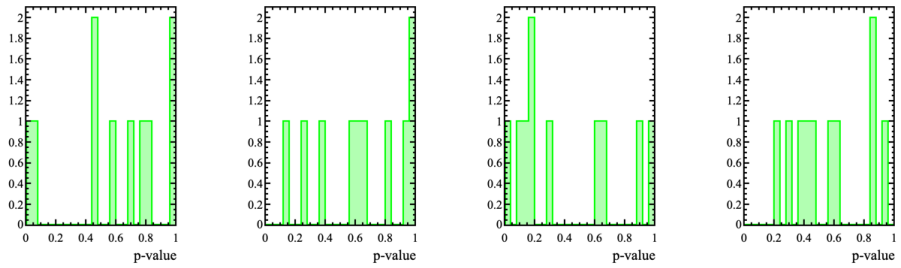


Figure 5.24: $D^+ \to K^-\pi^+\pi^+$ with $D_s^+ \to K^-K^+K^+$ selection; p-value distribution of search over the 10 samples for uniform 5×5 grid, uniform 8×8 grid, adaptive of 21 bins, adaptive of 50 bins.

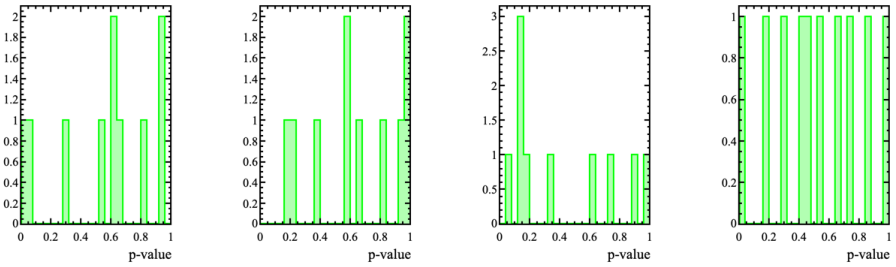


Figure 5.25: $D^+ \to K^-\pi^+\pi^+$ with $D^+ \to K^-K^+K^+$ selection; p-value distribution of search over the 10 samples for uniform 5×5 grid, uniform 8×8 grid, adaptive of 21 bins, adaptive of 50 bins.

Table 5.13: Lowest p-value responses over the search for charge asymmetries in the 10 control channel samples with each set of selection criteria.

binning		$D^+_s \to K^-K^+\pi^+$		$D^+ \to K^-\pi^+\pi^+$	
		D^+_s selection p-value	D^+ selection p-value	D^+_s selection p-value	D^+ selection p-value
Uniform	5×5	2.40×10^{-2}	4.19×10^{-4}	1.97×10^{-2}	2.70×10^{-2}
	8×8	1.14×10^{-3}	7.72×10^{-5}	1.29×10^{-1}	1.73×10^{-1}
Adaptive	21	3.08×10^{-2}	9.17×10^{-4}	2.09×10^{-2}	4.51×10^{-2}
	50	2.62×10^{-2}	1.54×10^{-2}	2.10×10^{-1}	3.85×10^{-2}

5.4
Blind analysis: foreseen strategy

As mentioned previously in this Chapter, the analysis is performed blinded. For an LHCb analysis like this, the procedure is the following: first, the blinded analysis steps are reviewed by the pertinent Working Group (WG), in this case the Charm WG; then, the analysis is approved to go to a next step: review of the full analysis by an internal Review Committee (RC) and eventually the approval to unblind the analysis. At this point, the analysis may receive the "approval to go to paper". Finally, there is a Collaboration wide review of the final paper of the analysis. At the moment, this analysis is going through the first WG review stage, and therefore no result for the actual search for CP violation in the $D^+_s \to K^-K^+K^+$ and $D^+ \to K^-K^+K^+$ decays is yet available.

For the unblinding of the result, the nominal method from which we will take the p-value is the novel fit per bin method, as it was shown the standard method introduces a bias for $D^+_s \to K^-K^+K^+$. The standard Mirandizing method will, however, also be applied as a cross-check.

The final p-values will either be greater than or lower than the 5σ threshold, at 3×10^{-7}. If the values are greater, it implies we are not yet statistically sensitive to CP violation in $D^+_s \to K^-K^+K^+$, which is

in accordance with the SM expectation, and for $D^+ \to K^-K^+K^+$ that there remains no observation of NP effects contributing to doubly-Cabibbo-suppressed decays. On the other hand, if the value is lower, it implies the sample sizes are significant enough and CP violation is being observed.

In the case of p-value$< 3 \times 10^{-7}$, our analysis technique only points towards the observation of CP violation, without being able to quantify it. To derive the A_{CP} of the decays, more complicated additional steps would be necessary, including obtaining values of the nuisance asymmetries related to production and detection effects.

6
Conclusions

In this dissertation, the analysis on the search for CP violation in the decays $D_s^+ \to K^-K^+K^+$ and $D^+ \to K^-K^+K^+$ was reported. Three-body decay studies like these, where the Dalitz plots of particle and antiparticle are statistically compared (here, by $\mathcal{S}_{CP}$), represent a direct way of verifying if the number of recorded decays are enough to be sensitive to CP violation effects, which, in charm physics, are quite small. With this analysis, we aim to look for signs of CP violation with the Cabibbo-suppressed $D_s^+ \to K^-K^+K^+$ decay, where this effect is possible in the SM at the per-mille level, while also looking for NP, with the doubly-Cabibbo-suppressed $D^+ \to K^-K^+K^+$ decay.

In this work, final samples of ~ 1.5 million events within the signal regions of the decays were analysed, being $\sim 64\%$ pure for $D_s^+ \to K^-K^+K^+$ and $\sim 78\%$ pure for $D^+ \to K^-K^+K^+$. These samples were obtained using 2016–2018 LHCb data. To perform the search for CP violation, the model-independent binned Mirandizing technique is employed according to two methods: the standard method, where both signal and background candidates within the signal regions of the decays are considered; and the new fit-per-bin method, which has not been applied to any previous analysis, where the full invariant mass spectra for every bin of the Dalitz plot are fitted in order to obtain the signal yields, and therefore remove the background effects. For these binned analyses, physics motivated binning schemes were elaborated to be able to separate relevant regions of the phase space where we expect CP violation asymmetries to change signs, and therefore increase the sensitivity to these effects.

To check the technique response, null-tests were designed. For the standard Mirandizing method, the null-test was performed using toy Dalitz plots, and we concluded that, given the overall low purity, the incompatible purities between D_s^+ and D_s^- samples and how differently the background is distributed across the bins of the phase space when compared to the signal, a bias is introduced in $D_s^+ \to K^-K^+K^+$ by inducing fake asymmetries through the global correction factor present in the calculation of $\mathcal{S}_{CP}$. For the fit-per-bin-method, the null-test is performed in data, and with it we conclude we are able to describe well all bins of the Dalitz plots by performing good invariant mass

fits, and so, given the test response was compatible with the null-hypothesis, validate this method, which is capable of removing the bias effect.

The technique sensitivity to possible CP violation effects with the current sample sizes is also tested using toys. From this, we see that our sensitivity remains low, as given by large values of either phase or magnitude differences in the resonant terms between the isobar models describing the decays of $D^+_{(s)}$ and $D^-_{(s)}$. Indeed, at the current sample sizes, the level of phase of magnitude differences to observe CP violation are at least an order of magnitude higher than those expected from the SM for $D^+_s \to K^-K^+K^+$ (asymmetries of about 10^{-4}–10^{-3}). For $D^+ \to K^-K^+K^+$, any signal would be due to NP. That means that, according to SM predictions, we expect to see no CP violation signal after unblinding. A positive result would need to be carefully investigated since it would represent NP effects.

Nuisance asymmetries are accounted for in checking if we are sensitive to them by analysing two Cabibbo-favoured control channels, $D^+_s \to K^-K^+\pi^+$ and $D^+ \to K^-\pi^+\pi^+$, and the backgrounds of our signal channels. We find that, even though we used 10 samples of each control channel with much higher sample sizes than for the signal channels, we are not sensitive to any spurious effect.

The final step of this analysis, i.e. the unblinding of the result for the signal channels $D^+_s \to K^-K^+K^+$ and $D^+ \to K^-K^+K^+$, has not yet been performed. The analysis is currently under scrutiny within the LHCb Collaboration when eventually the unblinding will be approved. The final results will be presented in the analysis future paper.

Bibliography

[1] WU, C. S.; AMBLER, E.; HAYWARD, R. W.; HOPPES, D. D. ; HUDSON, R. P.. **Experimental test of parity conservation in β decay**. Phys. Rev., 105:1413–1414, 1957.

[2] GARWIN, R. L.; LEDERMAN, L. M. ; WEINRICH, M.. **Observations of the failure of conservation of parity and charge conjugation in meson decays: the magnetic moment of the free muon**. Phys. Rev., 105:1415–1417, 1957.

[3] FRIEDMAN, J. I.; TELEGDI, V. L.. **Nuclear emulsion evidence for parity nonconservation in the decay chain $\pi^+ \to \mu^+ \to e^+$**. Phys. Rev., 106:1290–1293, 1957.

[4] CULLIGAN, G.; FRANK, S. G. F. ; HOLT, J. R.. **Longitudinal polarization of the electrons from the decay of unpolarized positive and negative muons**. Proceedings of the Physical Society, 73(2):169–177, Feb. 1959.

[5] CHRISTENSON, J. H.; CRONIN, J. W.; FITCH, V. L. ; TURLAY, R.. **Evidence for the 2π decay of the K_2^0 meson**. Phys. Rev. Lett., 13:138–140, Jul 1964.

[6] KOBAYASHI, M.; MASKAWA, T.. **CP violation in the renormalizable theory of weak interaction**. Prog. Theor. Phys., 49:652–657, 1973.

[7] NA48 Collaboration. **A new measurement of direct CP violation in two pion decays of the neutral kaon**. Phys. Lett. B, 465:335–348, 1999.

[8] BaBar Collaboration. **Measurement of CP violating asymmetries in b^0 decays to CP eigenstates**. Phys. Rev. Lett., 86:2515–2522, 2001.

[9] Belle Collaboration. **Observation of large CP violation in the neutral b meson system**. Phys. Rev. Lett., 87:091802, 2001.

[10] LHCb Collaboration. **Observation of CP violation in charm decays**. Phys. Rev. Lett., 122(21):211803, 2019.

[11] PESKIN, M. E.; SCHROEDER, D. V.. **An introduction to quantum field theory**. Addison-Wesley, Reading, USA, 1995.

[12] Particle Data Group. **Review of particle physics**. PTEP, 2020(8):083C01, 2020.

[13] SCHWARTZ, M. D.. **Quantum field theory and the Standard Model**. Cambridge University Press, 3 2014.

[14] HALZEN, F.; MARTIN, A. D.. **Quarks and leptons: an introductory course in modern particle physics**. 1984.

[15] BIGI, I. I.; SANDA, A. I.. **CP violation**, volumen 9. Cambridge University Press, 9 2009.

[16] ISIDORI, G.. **Flavor physics and CP violation**. In: 2012 European School of High-Energy Physics. CERN–2014–008.

[17] PICH, A.. **CP violation**. ICTP Ser. Theor. Phys., 10:14–42, 1994.

[18] KIM, J. E.. **A review on axions and the strong CP problem**. AIP Conf. Proc., 1200(1):83–92, 2010.

[19] CABIBBO, N.. **Unitary symmetry and leptonic decays**. Phys. Rev. Lett., 10:531–533, 1963.

[20] BUCHALLA, G.; BURAS, A. J. ; LAUTENBACHER, M. E.. **Weak decays beyond leading logarithms**. Rev. Mod. Phys., 68:1125–1144, 1996.

[21] BYCKLING, E.; KAJANTIE, K.. **Particle kinematics**. University of Jyvaskyla, Jyvaskyla, Finland, 1971.

[22] BLATT, J. M.; WEISSKOPF, V. F.. **Theoretical nuclear physics**. Springer, New York, 1952.

[23] ZEMACH, C.. **Three pion decays of unstable particles**. Phys. Rev., 133:B1201, 1964.

[24] E791 Collaboration. **Experimental evidence for a light and broad scalar resonance in $D^+ \to \pi^-\pi^+\pi^+$ decay**. Phys. Rev. Lett., 86:770–774, 2001.

[25] DE MIRANDA, J. M.. **Experimental evidence for a light and broad scalar resonance in $D^+ \to \pi^-\pi^+\pi^+$ decay**. Brazilian Journal of Physics, 30(2):296–303, jun 2000.

[26] BEDIAGA, I.; MIRANDA, J.; DOS REIS, A. C.; BIGI, I. I.; GOMES, A.; OTALORA GOICOCHEA, J. M. ; VEIGA, A.. **Second generation of 'Miranda procedure' for CP violation in Dalitz studies of B (\& D \& \tau) decays**. Phys. Rev. D, 86:036005, 2012.

[27] LHCb Collaboration. **Measurements of CP violation in the three-body phase space of charmless $B^{\pm}$ decays**. Phys. Rev. D, 90(11):112004, 2014.

[28] BaBar Collaboration. **Dalitz plot analysis of $B^+ \to \pi^-\pi^+\pi^+$ decays**. Phys. Rev. D, 79:072006, 2009.

[29] LENZ, A.. **What did we learn in theory from the ΔA_{CP}-saga?** In: 6th International Workshop on Charm Physics, 11 2013. arXiv:1311.6447 [hep-ph].

[30] BEDIAGA, I.; GÖBEL, C.. **Direct CP violation in beauty and charm hadron decays**. Prog. Part. Nucl. Phys., 114:103808, 2020.

[31] ISIDORI, G.; KAMENIK, J. F.; LIGETI, Z. ; PEREZ, G.. **Implications of the LHCb evidence for charm CP violation**. Phys. Lett. B, 711:46–51, 2012.

[32] GOLDEN, M.; GRINSTEIN, B.. **Enhanced CP violations in hadronic charm decays**. Phys. Lett. B, 222:501–506, 1989.

[33] BERGMANN, S.; NIR, Y.. **New physics effects in doubly Cabibbo suppressed D decays**. JHEP, 09:031, 1999.

[34] **LHC machine**. JINST, 3:S08001, 2008.

[35] MARCASTEL, F.. **CERN's accelerator complex. La chaîne des accélérateurs du CERN**. Oct 2013. General Photo OPEN-PHO-CHART-2013-001.

[36] ALICE Collaboration. **The ALICE experiment at the CERN LHC**. JINST, 3:S08002, 2008.

[37] ATLAS Collaboration. **The ATLAS experiment at the CERN Large Hadron Collider**. JINST, 3:S08003, 2008.

[38] CMS Collaboration. **The CMS experiment at the CERN LHC**. JINST, 3:S08004, 2008.

[39] ATLAS Collaboration. **Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC**. Phys. Lett. B, 716:1–29, 2012.

[40] CMS Collaboration. **Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC**. Phys. Lett. B, 716:30–61, 2012.

[41] RAPPOCCIO, S.. **The experimental status of direct searches for exotic physics beyond the Standard Model at the Large Hadron Collider**. Rev. Phys., 4:100027, 2019.

[42] ALVES, JR., A. A.; *et al.*. **The LHCb detector at the LHC**. JINST, 3:S08005, 2008.

[43] LHCb Collaboration. **LHCb VELO TDR: Vertex locator. Technical design report**. Technical Report CERN-LHCC-2001-011, 5 2001.

[44] LHCb Collaboration. **LHCb inner tracker: technical design report**. Technical Report CERN-LHCC-2002-029, Geneva, 2002.

[45] LHCb Collaboration. **LHCb outer tracker: technical design Report**. Technical Report CERN-LHCC-2001-024, Geneva, 2001.

[46] LHCb Collaboration. **LHCb magnet: technical design report**. Technical Report CERN-LHCC-2000-007, Geneva, 2000.

[47] LHCb Collaboration. **LHCb RICH: technical design report**. Technical Report CERN-LHCC-2000-037, Geneva, 2000.

[48] ADINOLFI, M.; *et al.*. **Performance of the LHCb RICH detector at the LHC**. The European Physical Journal C, 73(5), May 2013.

[49] LHCb Collaboration. **LHCb calorimeters: technical design report**. Technical Report CERN-LHCC-2000-036, Geneva, 2000.

[50] PICATOSTE OLLOQUI, E.. **LHCb preshower(PS) and scintillating pad detector (SPD): commissioning, calibration, and monitoring**. J. Phys. Conf. Ser., 160:012046, 2009.

[51] LHCb Collaboration. **LHCb muon system: technical design report**. Technical Report CERN-LHCC-2001-010, Geneva, 2001.

[52] LHCb Collaboration. **LHCb muon system: addendum to the technical design report**. Technical Report CERN-LHCC-2003-002, Geneva, 2003.

[53] LHCb Collaboration. **LHCb muon system: second addendum to the technical design report**. Technical Report CERN-LHCC-2005-012, Geneva, 2005.

[54] MATEV, R.. **The LHCb trigger in run II**. Jun 2016. LHCb-TALK-2016-214.

[55] ALFONSO ALBERO, A.; *et al.*. **Upgrade trigger selection studies**. Technical Report LHCb-PUB-2019-013, CERN-LHCb-PUB-2019-013, CERN, Geneva, Sep 2019.

[56] LHCb Collaboration. **Design and performance of the LHCb trigger and full real-time reconstruction in run 2 of the LHC**. JINST, 44:P04013–P04013, 2019.

[57] BREIMAN, L.; FRIEDMAN, J. H.; OLSHEN, R. A. ; STONE, C. J.. **Classification and regression trees**. Wadsworth international group, Belmont, California, USA, 1984.

[58] PIVK, M.; LE DIBERDER, F.. **sPlot: a statistical tool to unfold data distributions**. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 555(1):356–369, 2005.

[59] SKWARNICKI, T.. **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.

[60] LHCb Collaboration. **Dalitz plot analysis of the $D^+ \to K^-K^+K^+$ decay**. JHEP, 04:063, 2019.

[61] BEDIAGA, I.; BIGI, I. I.; GOMES, A.; GUERRER, G.; MIRANDA, J. ; REIS, A. C. D.. **On a CP anisotropy measurement in the Dalitz plot**. Phys. Rev., D80:096006, 2009.

[62] BaBar Collaboration. **Search for CP violation in neutral D meson Cabibbo-suppressed three-body decays**. Phys. Rev. D, 78:051102, 2008.

[63] LYONS, L.. **Statistics for nuclear physics**. Cambridge university press, Cambridge, England, 1986.

[64] LHCb Collaboration. **Search for CP violation in $\Xi_c^+ \to pK^-\pi^+$ decays using model-independent techniques**. Eur. Phys. J. C, 80(10):986, 2020.

[65] LHCb Collaboration. **Model-independent search for CP violation in $D^0 \to K^-K^+\pi^-\pi^+$ and $D^0 \to \pi^-\pi^+\pi^+\pi^-$ decays**. Phys. Lett. B, 726:623–633, 2013.

[66] CDF Collaboration. **Measurement of CP-violation asymmetries in $D^0 \to K_S\pi^+\pi^-$**. Phys. Rev. D, 86:032007, 2012.

[67] KLEIN, J. R.; ROODMAN, A.. **Blind analysis in nuclear and particle physics**. Ann. Rev. Nucl. Part. Sci., 55:141–163, 2005.

[68] BACK, J.; *et al.*. **LAURA^{++}: A Dalitz plot fitter**. Comput. Phys. Commun., 231:198–242, 2018.

[69] CORTI, G.; *et al.*. **How the Monte Carlo production of a wide variety of different samples is centrally handled in the LHCb experiment**. J. Phys. Conf. Ser., 664(7):072014, 2015.

[70] CLEMENCIC, M.; CORTI, G.; EASO, S.; JONES, C. R.; MIGLIORANZI, S.; PAPPAGALLO, M. ; ROBBE, P.. **The LHCb simulation application, Gauss: design, evolution and experience**. J. Phys. Conf. Ser., 331:032023, 2011.

[71] SJOSTRAND, T.; MRENNA, S. ; SKANDS, P. Z.. **A brief introduction to PYTHIA 8.1**. Comput. Phys. Commun., 178:852–867, 2008.

[72] LANGE, D. J.. **The EvtGen particle decay simulation package**. Nucl. Instrum. Meth. A, 462:152–155, 2001.

[73] AGOSTINELLI, S.; *et al.*. **GEANT4–a simulation toolkit**. Nucl. Instrum. Meth. A, 506:250–303, 2003.

[74] ALLISON, J.; *et al.*. **Geant4 developments and applications**. IEEE Trans. Nucl. Sci., 53:270, 2006.

[75] ANDERLINI, L.; *et al.*. **The PIDCalib package**. Technical Report LHCb-PUB-2016-021, CERN-LHCb-PUB-2016-021, CERN, Geneva, Jul 2016.

A
Appendix: Monte Carlo samples fits

Full LHCb simulations of the signal decays are used in this analysis [69]. These simulations, created using Monte Carlo (MC) techniques, are derived with the Gauss framework [70] using specialised programs. Events are generated using Pythia 8 [71] to reproduce generic proton-proton interactions following the LHCb operating conditions, at 13 TeV. These simulations start from the hard processes by considering the parton distributions within the high-energy incoming protons to the outgoing partons, which then produce showers of outgoing hadrons (due to confinement). The hadron decays are introduced in the MC using EvtGen [72], which then pass to the GEANT4 [73,74] in order to simulate the detector effects, so as to include the particles propagation and interaction through the detector materials. For event reconstruction, *Moore* software is used to simulate the trigger stages, while *Brunel* transforms the detector hits into objects and tracks.

In order to generate large MC samples, we provide generator level cuts, shown in Table A.1, and save candidates that pass through the trigger lines of the signal channels, `Hlt2CharmHadDspToKmKpKp, Hlt2CharmHadDpToKmKpKp`.

Table A.1: MC cuts applied at generator level.

variable	cut
each daughter $P >$	2.0 GeV
each daughter $P_T >$	0.25 GeV
$D^+_{(s)} P >$	14.0 GeV
$D^+_{(s)} P_T >$	2.1 GeV

The relevant data selection criteria are then applied to the MC, with the exception of PID requirements. In simulated samples, the response of the RICH detectors is not well modelled, and therefore, to emulate PID criteria, a data-driven method is applied. Using the PIDCalib package [75], efficiencies on the PID requirements are derived from $D^0 \to K^-\pi^+$ calibration samples.

First, the MC samples are used to derive the signal PDF parameters for $D_s^+ \to K^-K^+K^+$ and $D^+ \to K^-K^+K^+$ for the sPlot fit. Using a Gaussian and 2 Crystal Balls (Eq. 4-5), the MC samples (still with no BDT requirement)

are fitted. The fits to MC distributions can be seen in Fig. A.1, and the fit results in Table A.2. For the fit to data, the signal components fractions and Crystal Ball parameters are fixed.

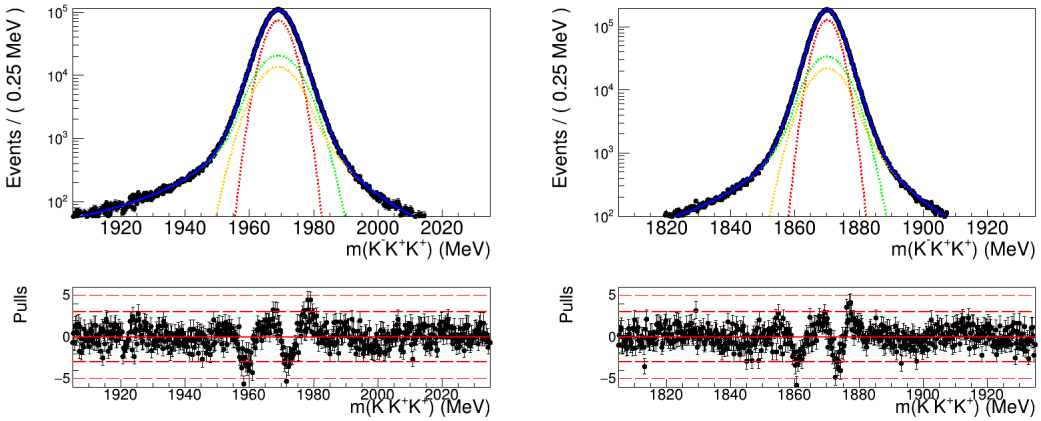


Figure A.1: Monte Carlo fits using the PDF in Eq. 4-5 for $D_s^+ \to K^-K^+K^+$ (left) and $D^+ \to K^-K^+K^+$ (right) before MVA, with Y-axis in log scale. The resulting signal shape is fixed in the sPlot fit.

Table A.2: MC fit results before MVA.

		$D_s^+ \to K^-K^+K^+$	$D^+ \to K^-K^+K^+$
	N_{sig}	3437422 ± 1854	6545644 ± 2558
Signal components fractions	f_G	0.575 ± 0.009	0.56 ± 0.01
	f_{CB}	0.62 ± 0.03	0.59 ± 0.05
Gaussian parameters	μ [MeV]	1969.039 ± 0.003	1870.100 ± 0.002
	σ_G [MeV]	3.54 ± 0.02	3.21 ± 0.02
Crystal Ball parameters	R_1	1.66 ± 0.01	1.69 ± 0.02
	R_2	1.65 ± 0.03	1.71 ± 0.04
	α_1	2.08 ± 0.03	2.05 ± 0.05
	α_2	-1.818 ± 0.04	-1.79 ± 0.06
	N_1	1.41 ± 0.03	1.68 ± 0.05
	N_2	2.21 ± 0.06	2.60 ± 0.07

The same procedure is applied to the MC samples with the BDT response applied to them, in order to derive the signal shape parameters used in the final sample invariant mass fits. These fits can be seen in Fig. A.2, and the results in Table A.3. This is repeated for every bin of the physics motivated binning schemes to derive the signal parameters used in the null-test on data, fixing different Crystal Ball parameters for each.

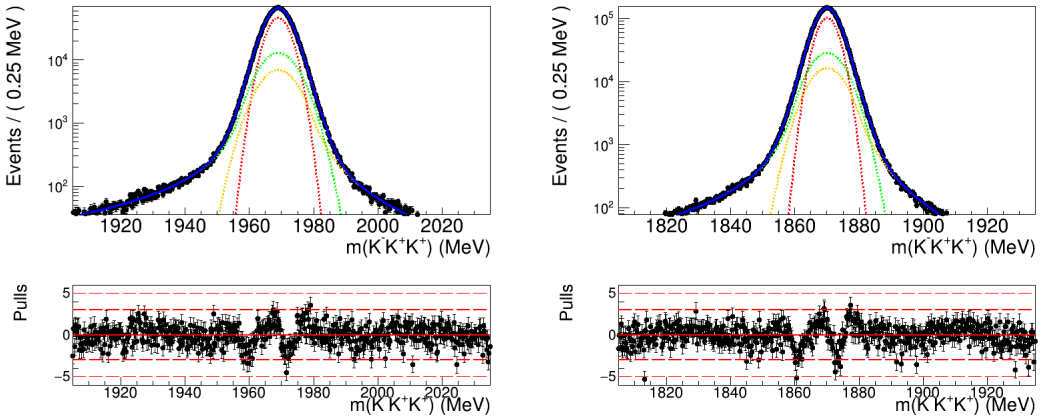


Figure A.2: Monte Carlo fits using the PDF in Eq. 4-5 for $D_s^+ \to K^-K^+K^+$ (left) and $D^+ \to K^-K^+K^+$ (right) after MVA. The resulting signal shape is fixed in the final sample fit.

Table A.3: MC fit results after MVA.

		$D_s^+ \to K^-K^+K^+$	$D^+ \to K^-K^+K^+$
	N_{sig}	2807185 ± 1675	5702413 ± 2388
Signal components fractions	f_G	0.58 ± 0.01	0.568 ± 0.006
	f_{CB}	0.65 ± 0.08	0.63 ± 0.02
Gaussian parameters	μ [MeV]	1969.020 ± 0.003	1870.096 ± 0.002
	σ_G [MeV]	3.51 ± 0.02	3.18 ± 0.01
Crystal Ball parameters	R_1	1.62 ± 0.03	1.66 ± 0.01
	R_2	1.63 ± 0.04	1.69 ± 0.02
	α_1	2.14 ± 0.06	2.13 ± 0.02
	α_2	-1.9 ± 0.1	-1.83 ± 0.03
	N_1	1.33 ± 0.03	1.58 ± 0.02
	N_2	2.0 ± 0.1	2.41 ± 0.05

B
Appendix: Dalitz Plot fit of $D_s^+ \to K^-K^+K^+$

In order to derive an approximate isobar model for $D_s^+ \to K^-K^+K^+$, a 100 thousand sample was fitted using Laura++. Our baseline model consisted of an isobar model with the three resonances obtained in the $D^+ \to K^-K^+K^+$ model, also using the same mass and width parameters for $f_0(1370)$. In addition to this model, we considered some variations:

- Model 1 (baseline): ϕ, $f_0(980)$ and $f_0(1370)$
- Model 2: ϕ, $f_0(980)$ and $f_0(1370)$, with free $M_{f_0(1370)}$ and $\Gamma_{f_0(1370)}$
- Model 3: ϕ, $f_0(980)$ and $f_0(1370)$, with free $f_0(980)$ coupling parameters
- Model 4: ϕ, $f_0(980)$, $f_0(1370)$ and $f_0(1500)$

The fit quality is measured by χ^2, defined by

$$\chi^2 = \sum_{i=1}^{n_{bins}} \chi_i^2 = \sum_{i=1}^{n_{bins}} \frac{(N_i^{obs} - N_i^{fit})^2}{\sigma_i^2} \,, \tag{B-1}$$

where N^{obs} is the number of observed events, N^{fit} the estimated number from the fit in n_{bins} of the Dalitz plot. The number of degrees of freedom (ndf) is given by $n_{bins} - n_{par} - 1$, where n_{par} is the number of free parameters in the fit, and with it we may calculate χ^2/ndf. The closer to 1 χ^2/ndf is, the better the fit quality. In addition, for method comparison, the minimised negative log-likelihood (NLL) of the fits are used. Between two models, the one with the lowest value of NLL represents a better solution.

The χ^2/ndf and NLL responses for all four models are presented in Table. B.1. We find the variations to the baseline model to not represent any significant gain in fit quality given their added complexity, and so we use this model for toy generation.

Table B.1: $D_s^+ \to K^-K^+K^+$ Dalitz plot fit responses with ΔNLL being the difference to the baseline model.

Model	χ^2/ndf	NLL	ΔNLL
1	2.39	-1136024	—
2	2.38	-1136034	-10
3	1.94	-1136183	-159
4	2.36	-1136034	-10

The Dalitz plots of data and a toy generated according to the baseline model are shown in Fig. B.1, with projections of s_{low} and s_{high} presented previously in Fig. 5.4.

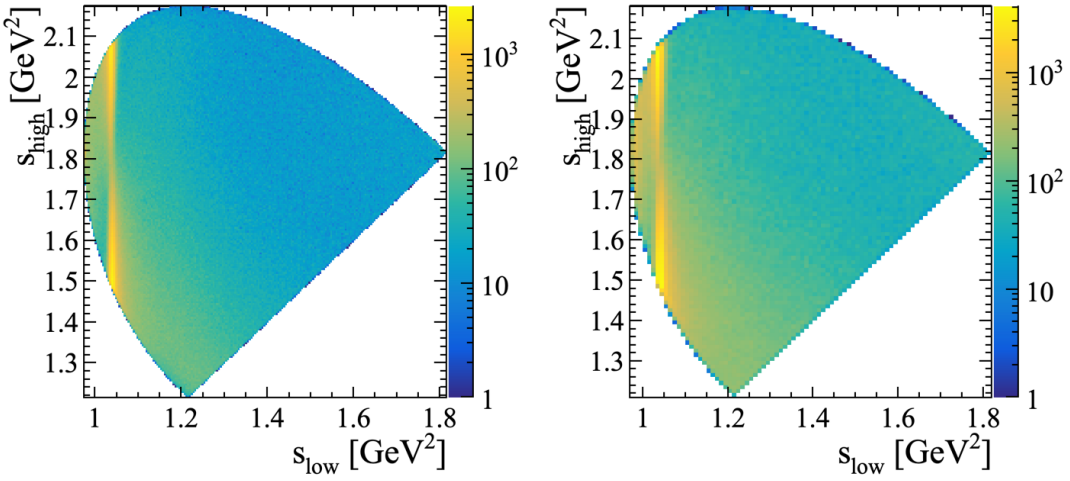


Figure B.1: Dalitz plots of the data final sample $2\sigma_{eff}$ signal region of $D_s^+ \to K^-K^+K^+$ (left) and generated from the fitted baseline model (right). The Z-axis is displayed in log scale for better visualisation of the interference pattern.

C
Appendix: Fits per bin for Sample 1 of null-test

Fits to the invariant mass distributions per bin of the Sample 1 Dalitz plots, used in the null-test on data (Sec.5.3.1), are presented here. $D_s^+ \to K^-K^+K^+$ plots for physics motivated binning scheme with 21 total bins are given in Figs. C.1 and C.2, and for 50 total bins in Figs. C.3–C.6. For $D^+ \to K^-K^+K^+$, the plots for 21 total bins are given in Figs. C.7 and C.8, and for 50 total bins in Figs. C.9–C.12.

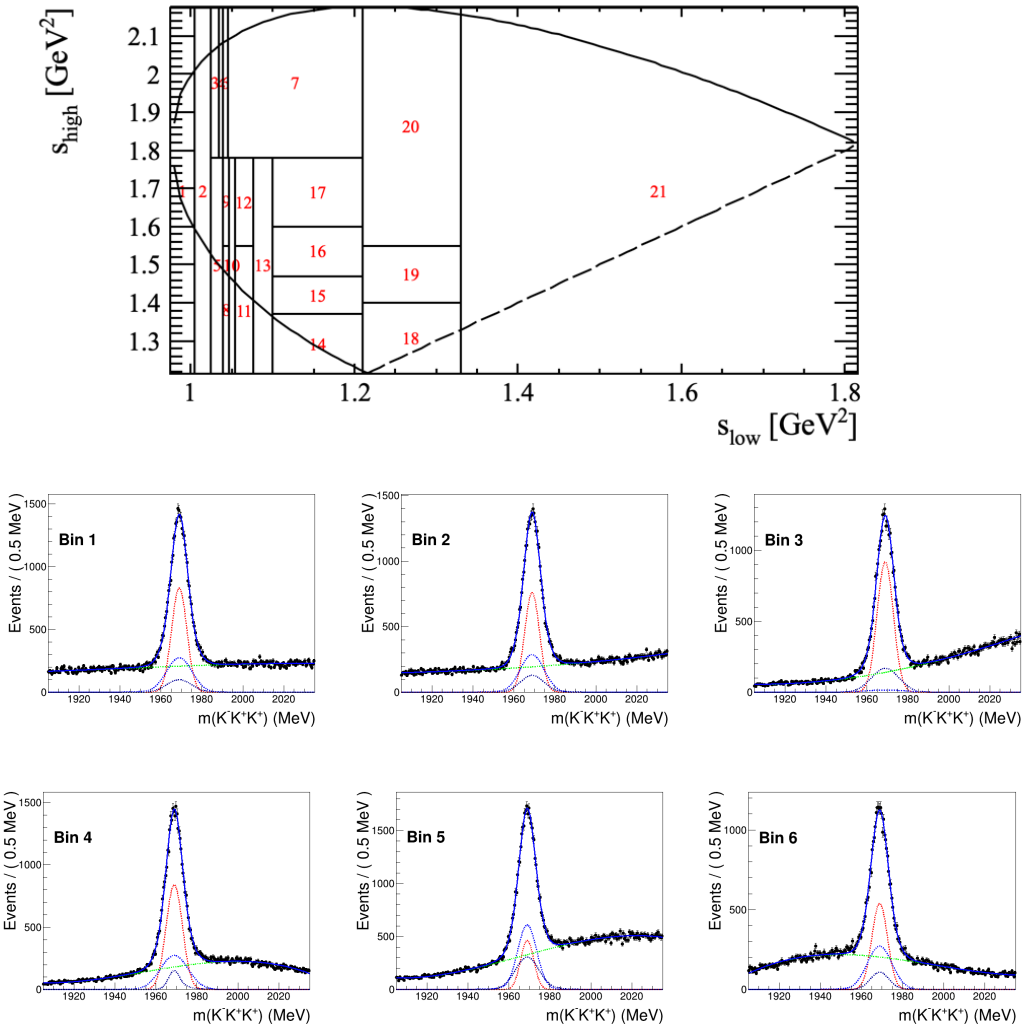


Figure C.1: $D_s^+ \to K^-K^+K^+$ invariant mass fits per bin of the physics motivated binning scheme with 21 total bins in Sample 1 (bins 1-6).

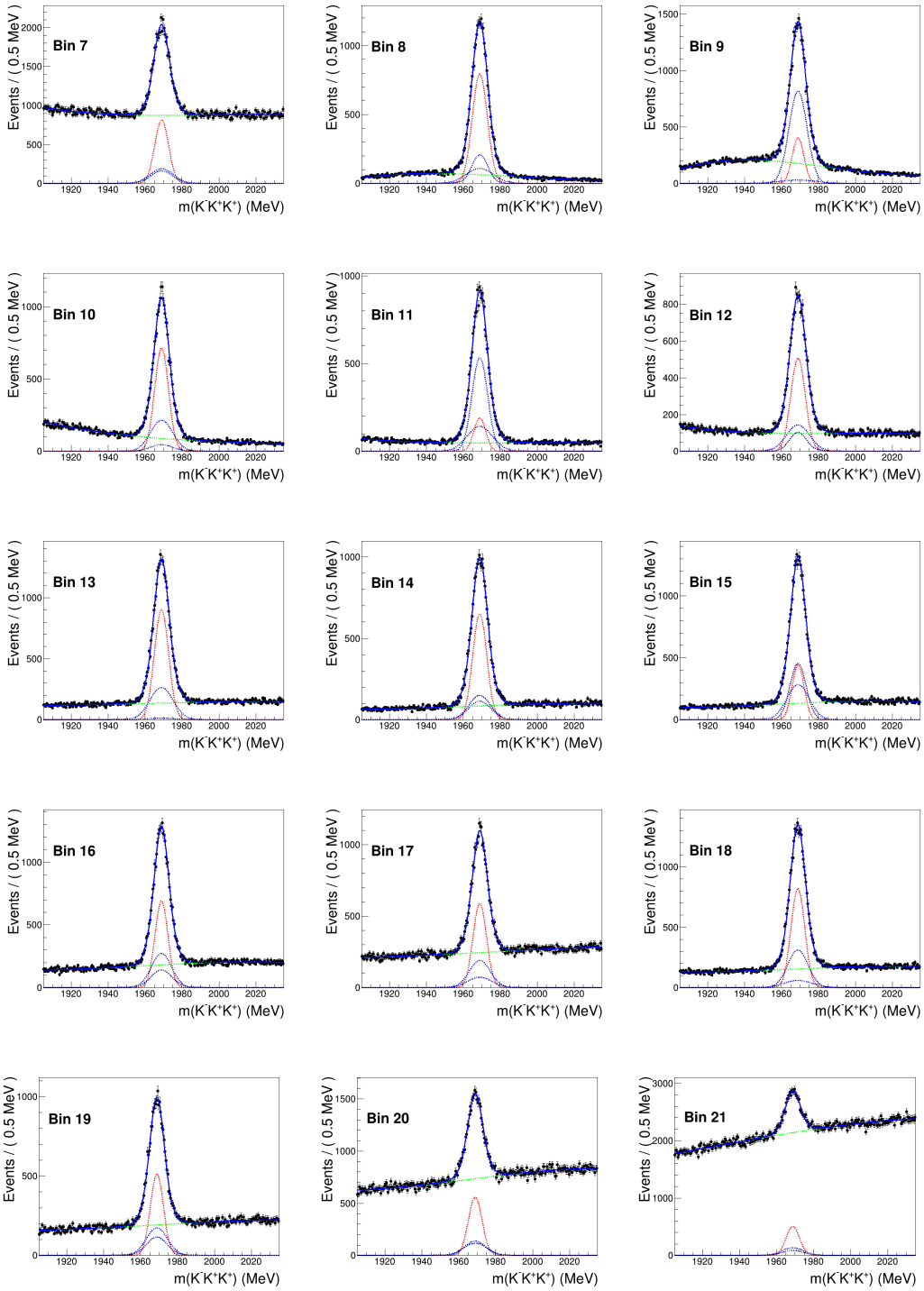


Figure C.2: $D_s^+ \to K^-K^+K^+$ invariant mass fits per bin of the physics motivated binning scheme with 21 total bins in Sample 1 (bins 7-21).

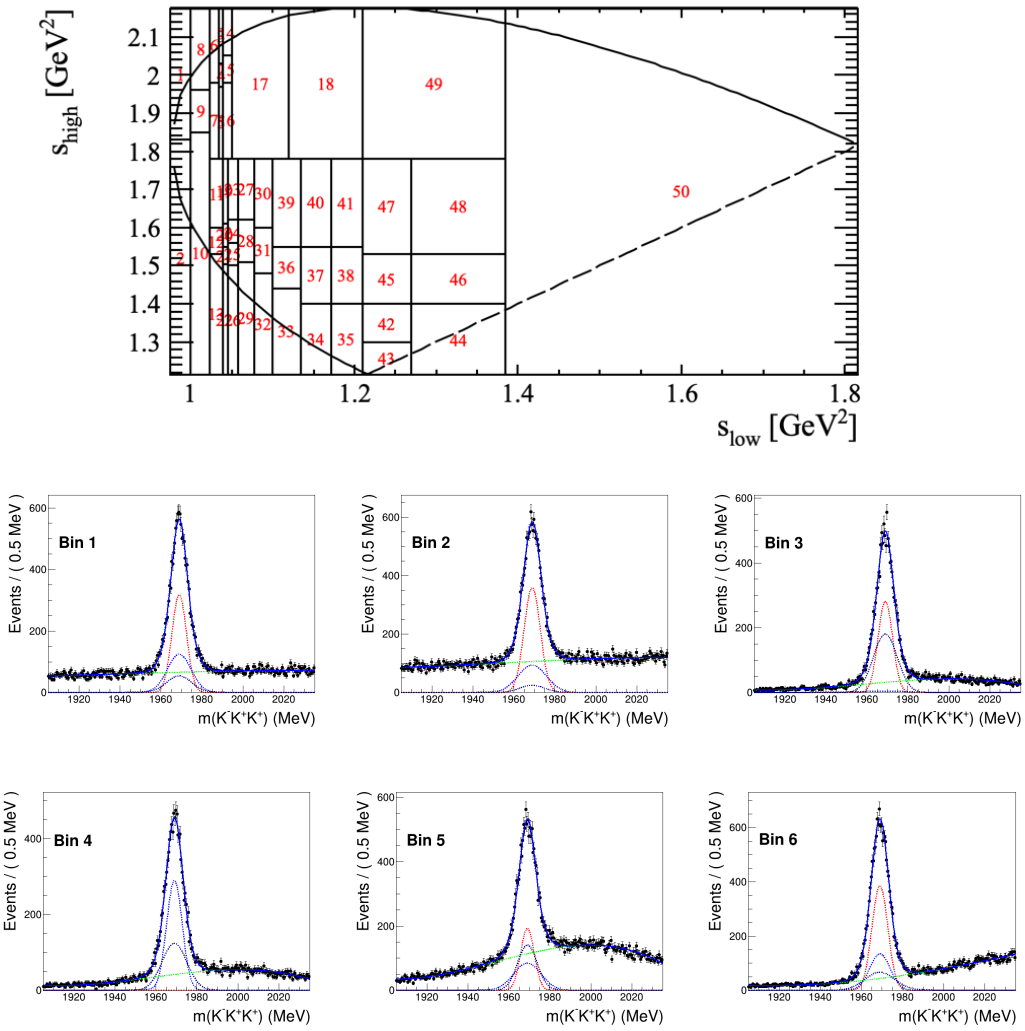


Figure C.3: $D_s^+ \to K^-K^+K^+$ invariant mass fits per bin of the physics motivated binning scheme with 50 total bins in Sample 1 (bins 1-6).

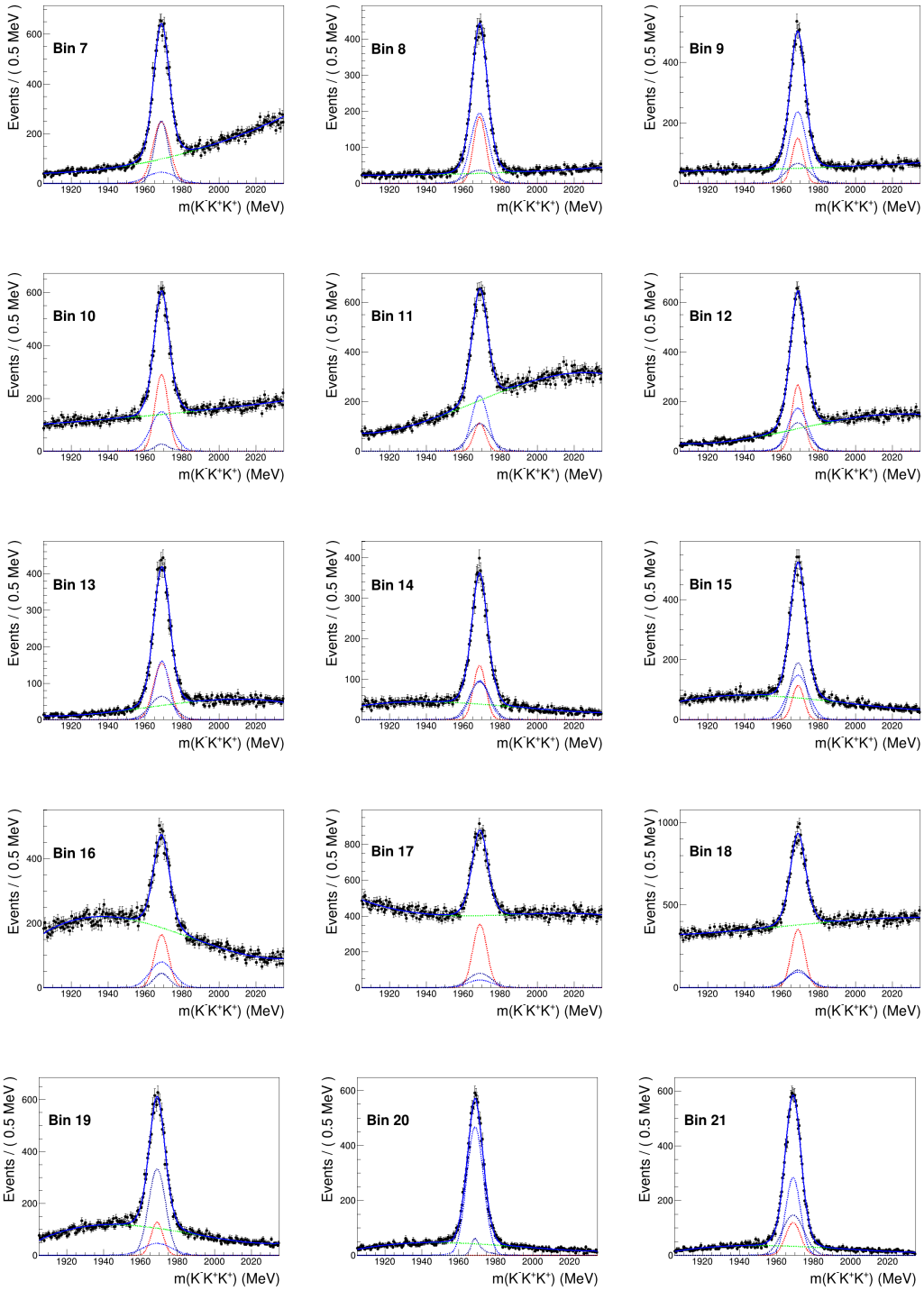


Figure C.4: $D_s^+ \to K^-K^+K^+$ invariant mass fits per bin of the physics motivated binning scheme with 50 total bins in Sample 1 (bins 7-21).

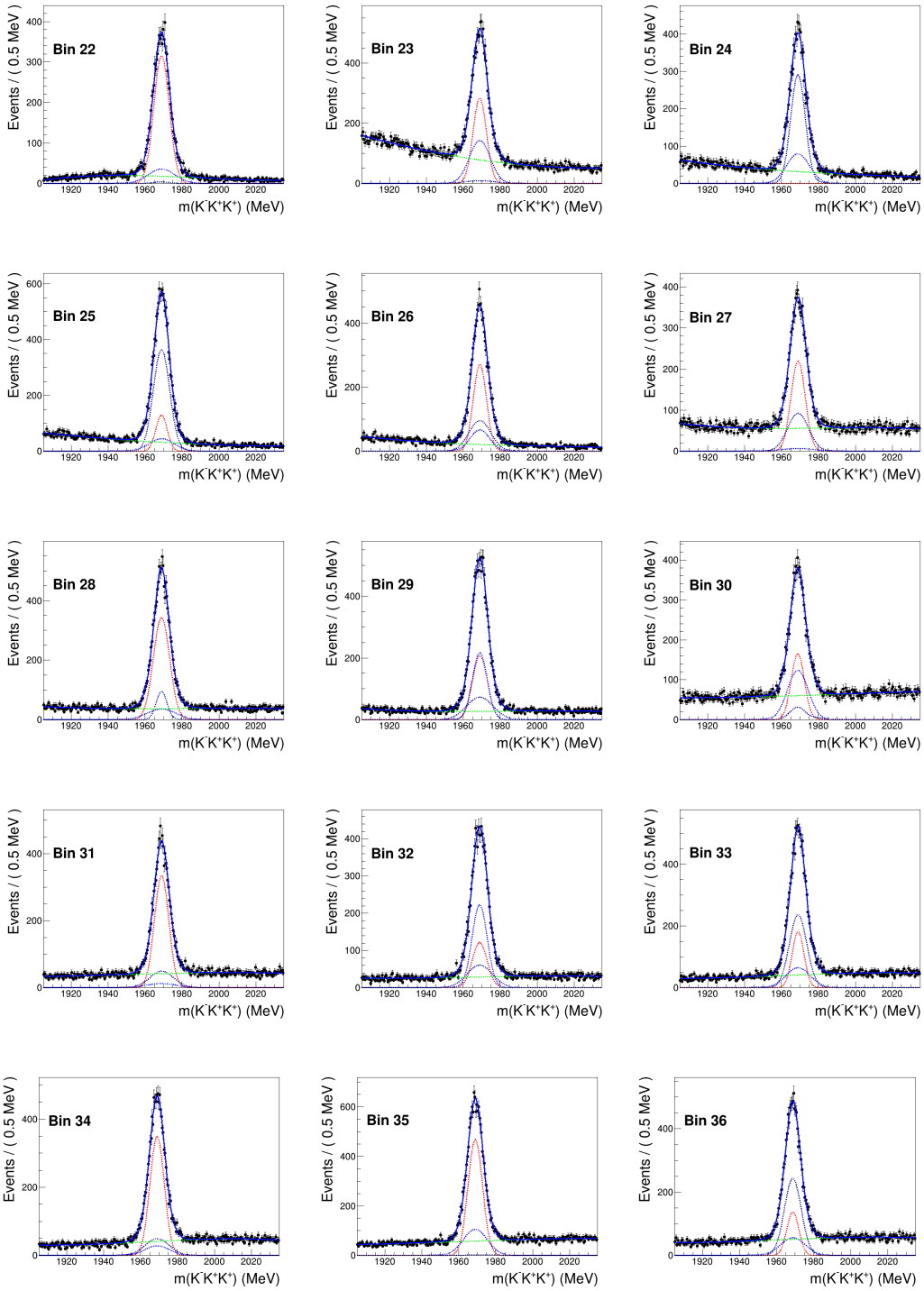


Figure C.5: $D_s^+ \to K^-K^+K^+$ invariant mass fits per bin of the physics motivated binning scheme with 50 total bins in Sample 1 (bins 22-36).

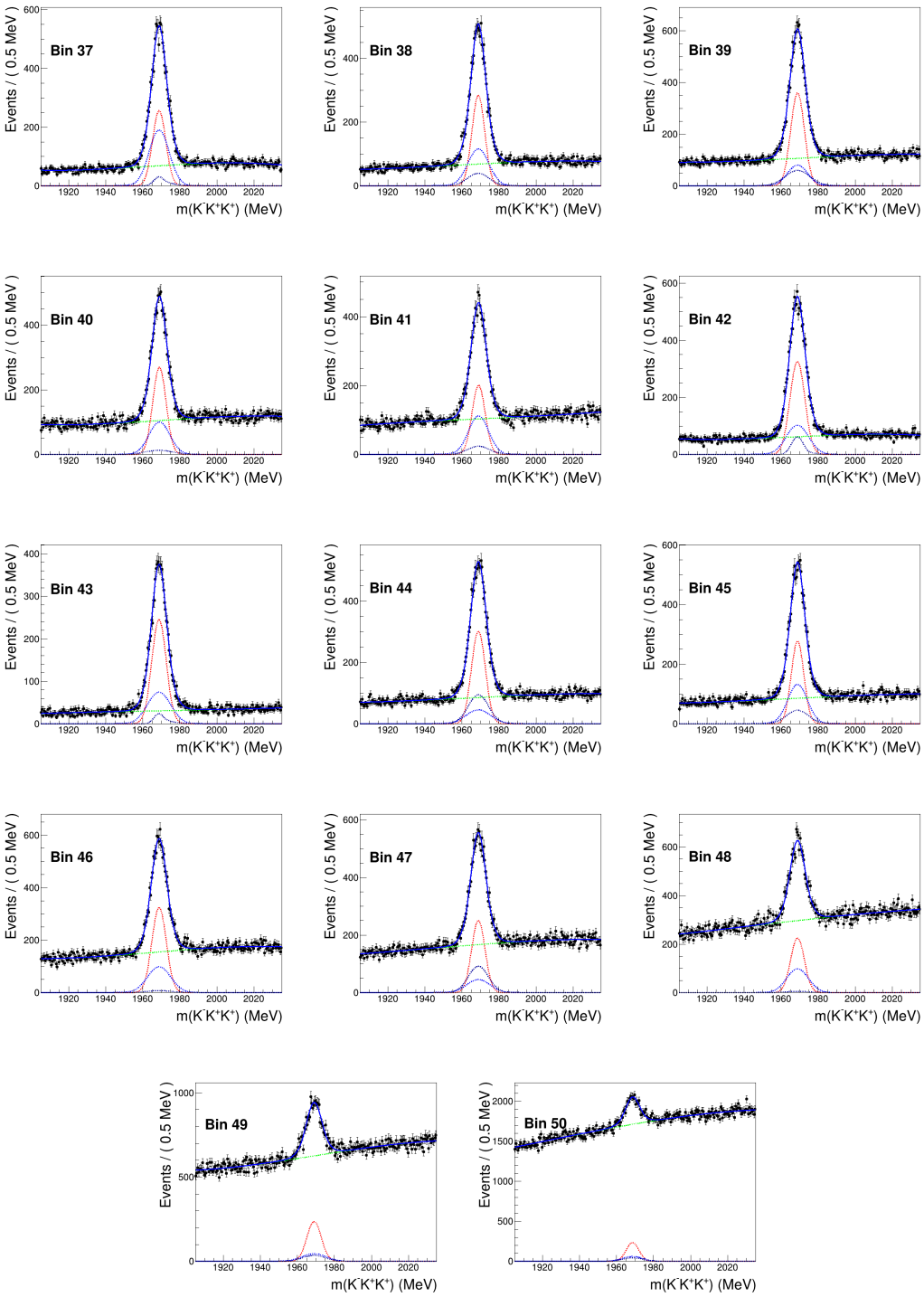


Figure C.6: $D_s^+ \to K^-K^+K^+$ invariant mass fits per bin of the physics motivated binning scheme with 50 total bins in Sample 1 (bins 37-50).

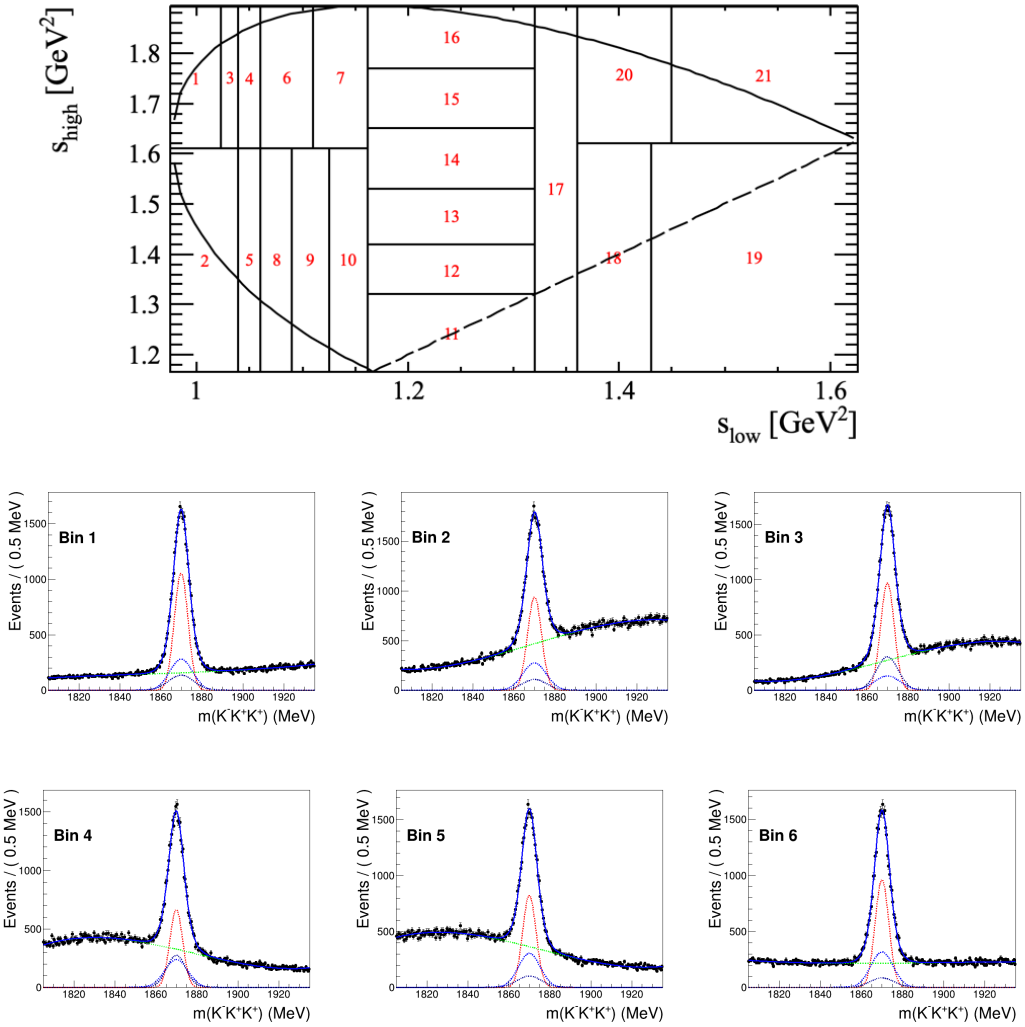


Figure C.7: $D^+ \to K^-K^+K^+$ invariant mass fits per bin of the physics motivated binning scheme with 21 total bins in Sample 1 (bins 1-6).

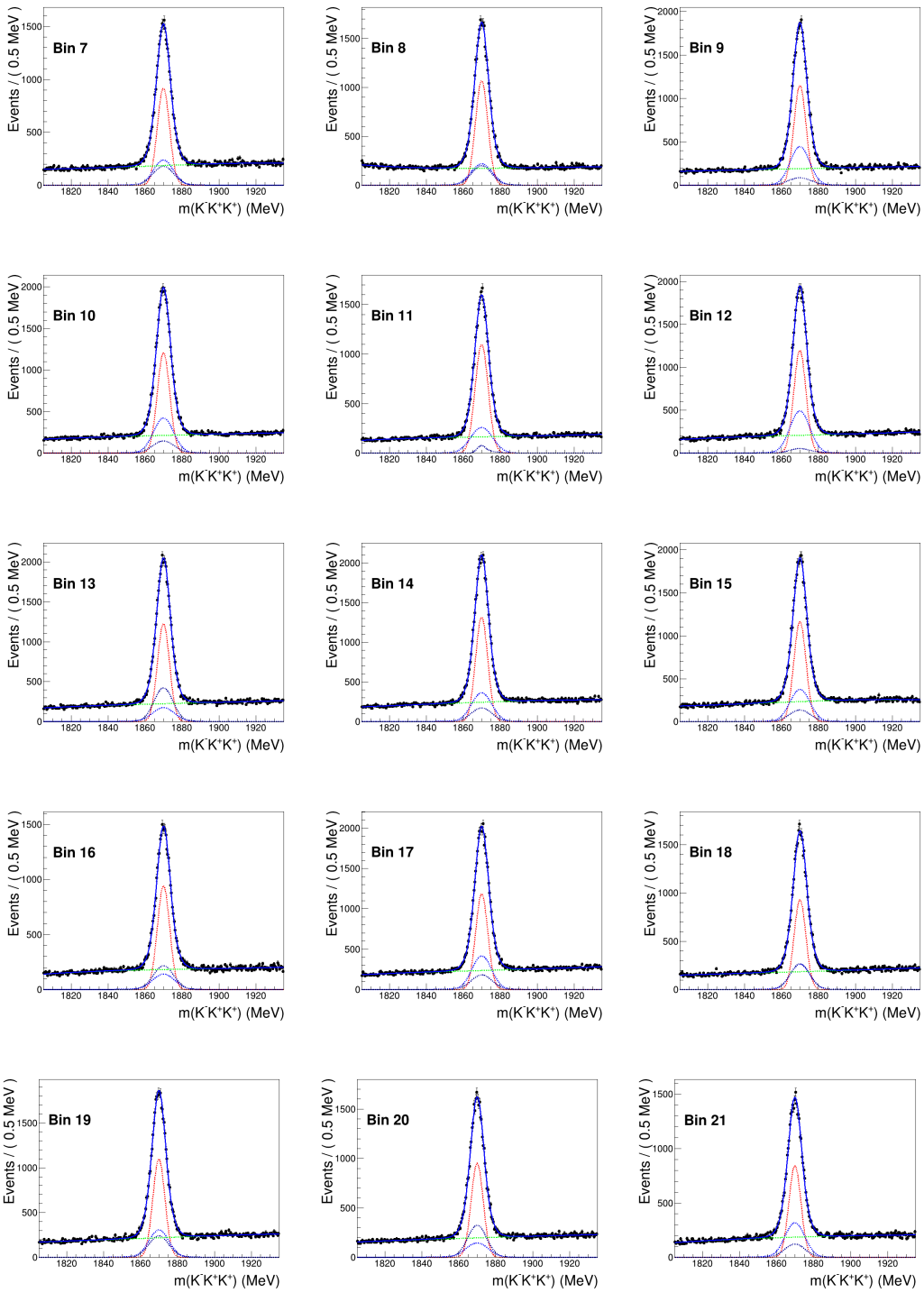


Figure C.8: $D^+ \to K^-K^+K^+$ invariant mass fits per bin of the physics motivated binning scheme with 21 total bins in Sample 1 (bins 7-21).

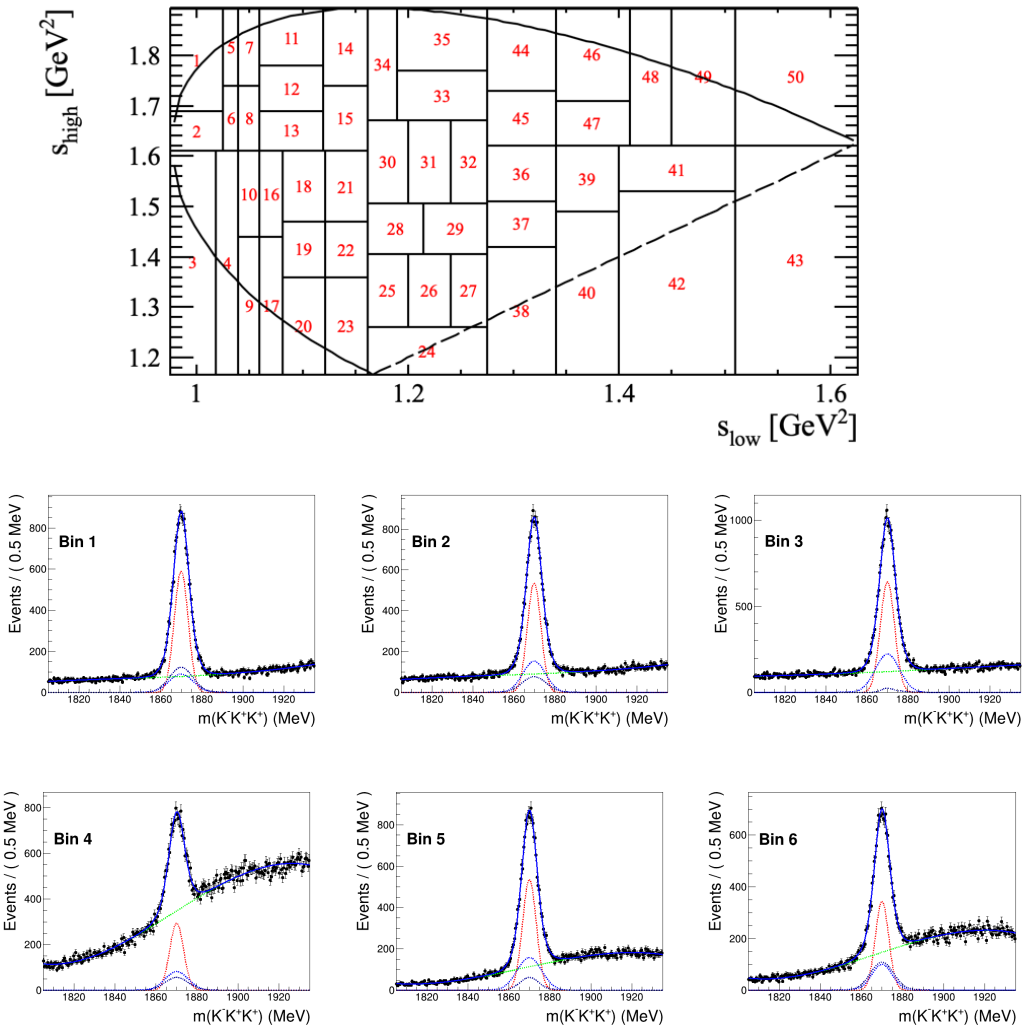


Figure C.9: $D^+ \to K^-K^+K^+$ invariant mass fits per bin of the physics motivated binning scheme with 50 total bins in Sample 1 (bins 1-6).

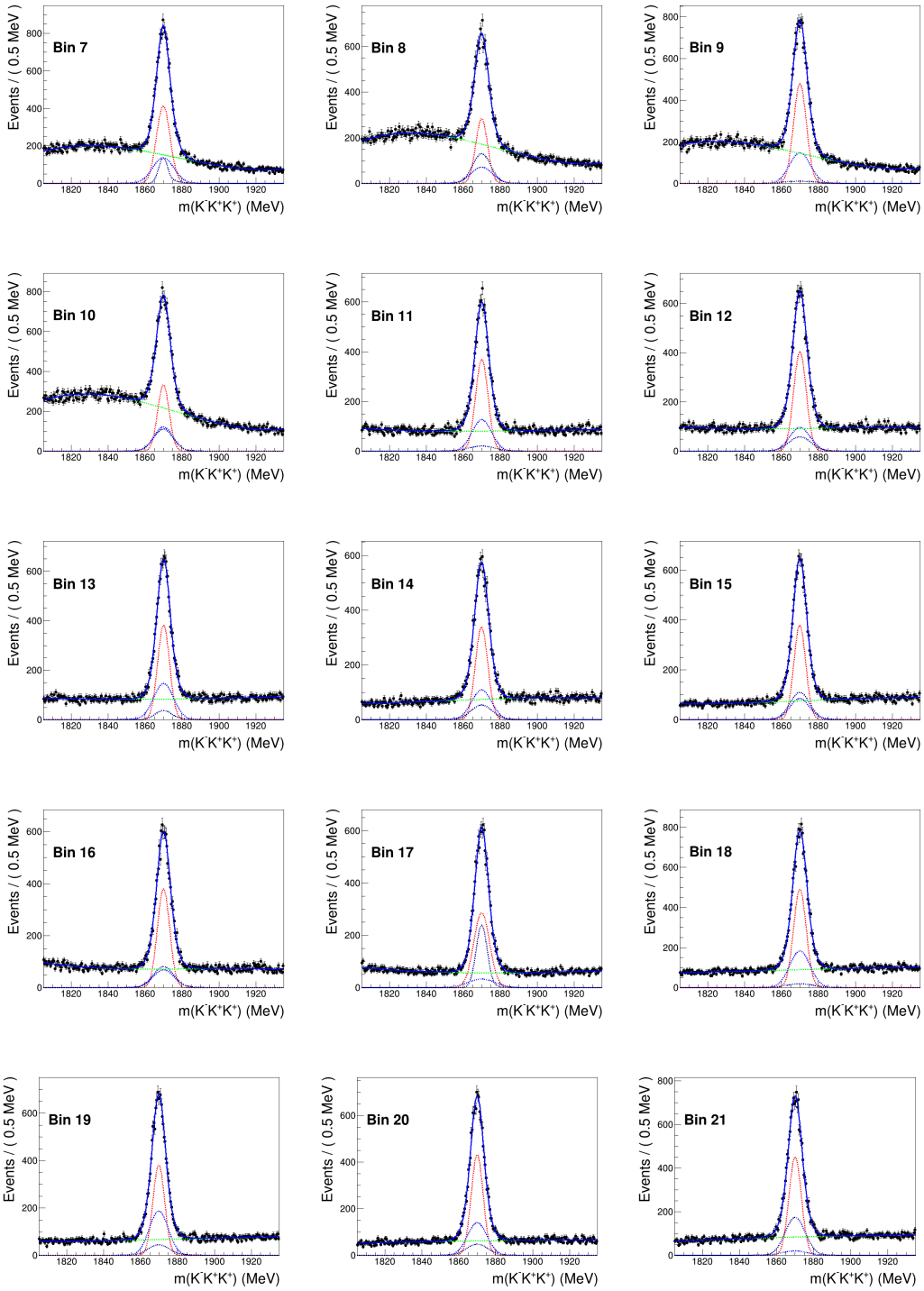


Figure C.10: $D^+ \to K^-K^+K^+$ invariant mass fits per bin of the physics motivated binning scheme with 50 total bins in Sample 1 (bins 7-21).

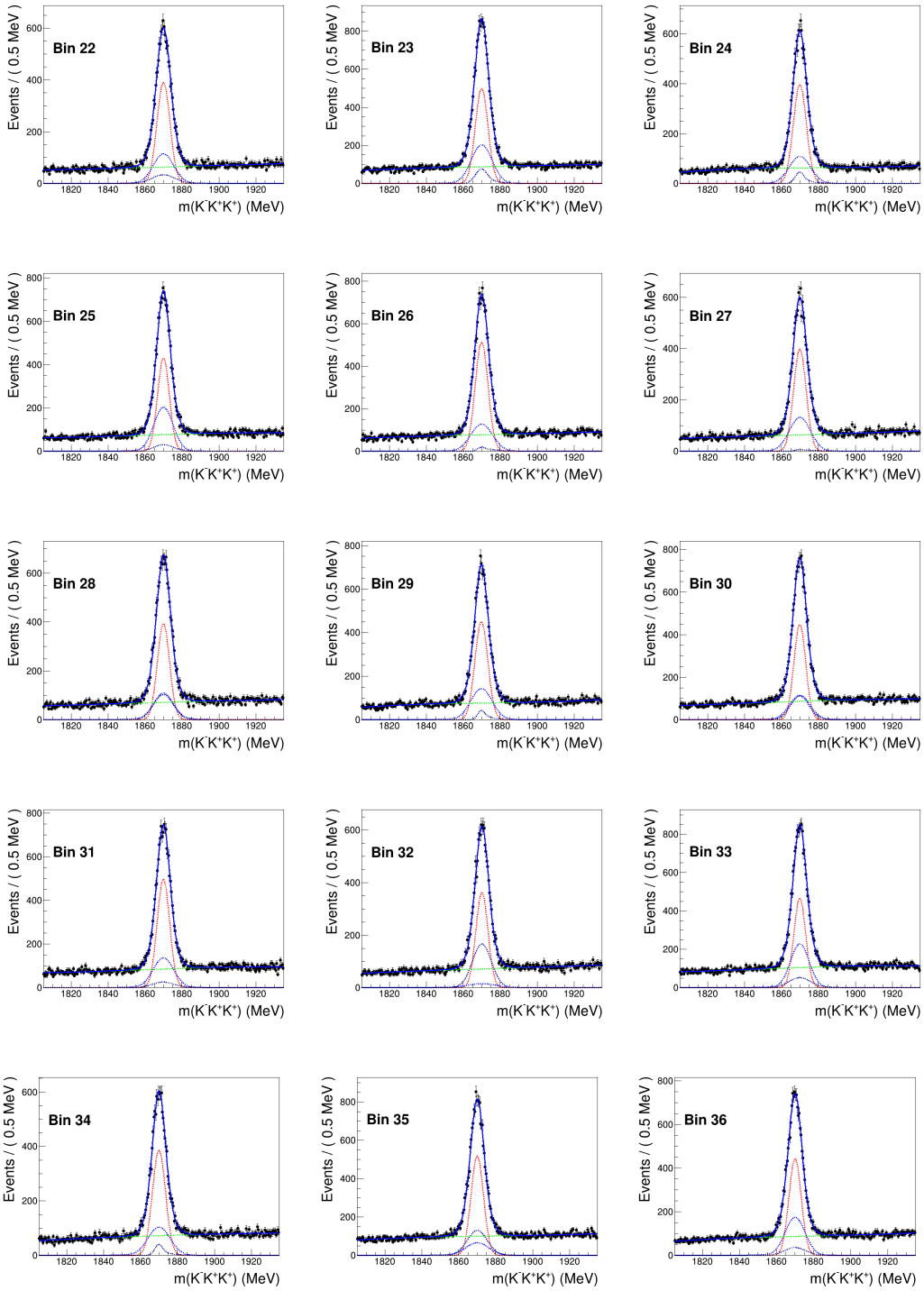


Figure C.11: $D^+ \to K^-K^+K^+$ invariant mass fits per bin of the physics motivated binning scheme with 50 total bins in Sample 1 (bins 22-36).

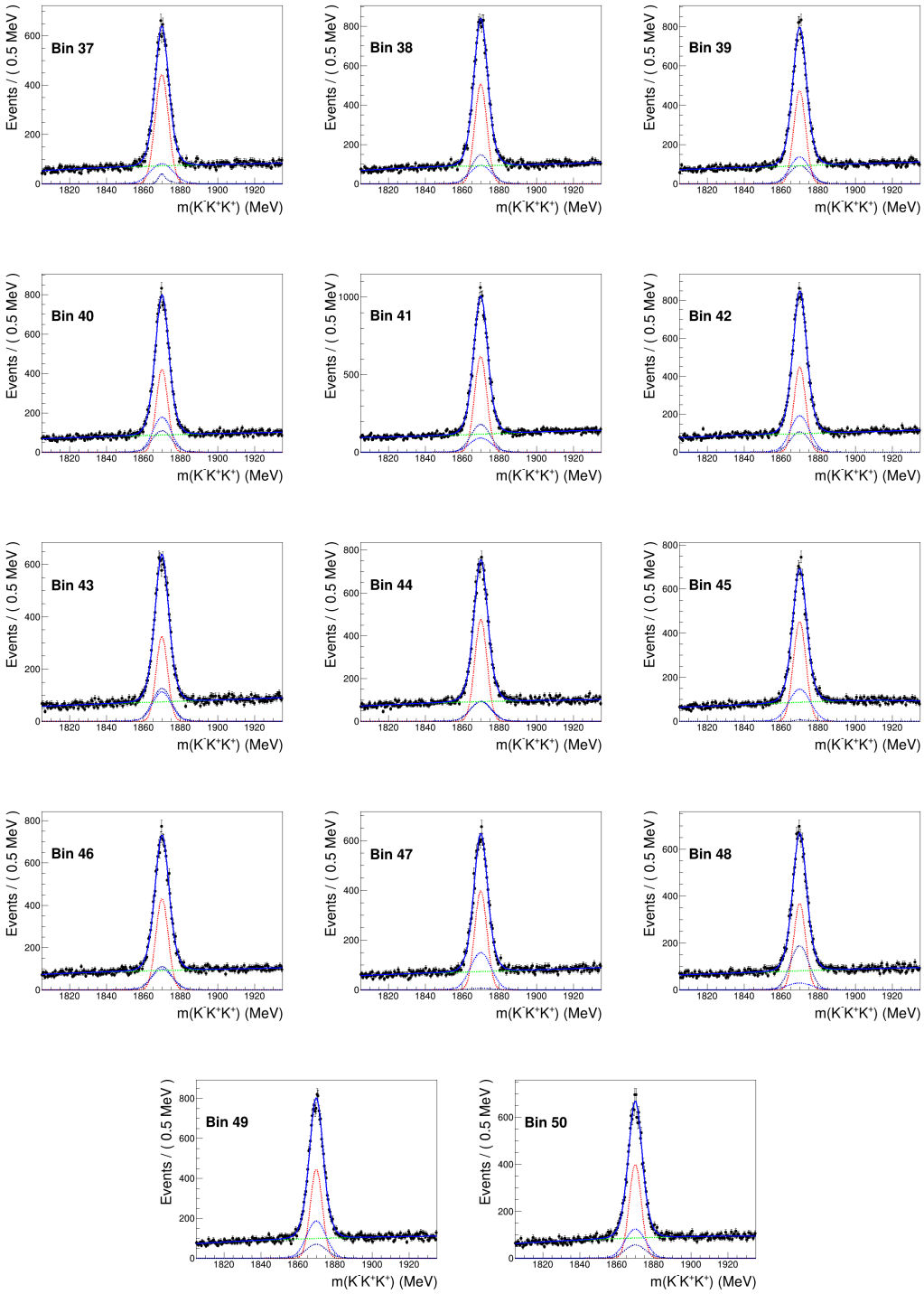


Figure C.12: $D^+ \to K^-K^+K^+$ invariant mass fits per bin of the physics motivated binning scheme with 50 total bins in Sample 1 (bins 37-50).