

Chapter 4

Event selection

Contents

4.1	Data samples and Simulations	40
4.1.1	Data samples	40
4.1.2	Simulation samples	40
4.2	Online selection	41
4.3	Offline selection	44
4.3.1	Selection requirements	45
4.3.2	Particle identification requirements	54
4.3.3	Charmless and single-charm backgrounds	54
4.3.4	Misidentified D and A_c^+ hadrons	56
4.3.5	Invariant mass vetoes	58
4.3.6	Normalisation mode veto	60
4.3.7	Multivariate analysis	62
4.3.8	Impact parameter requirements	78
4.4	Multiple candidates	81
4.5	Refitting the decay chain	82
4.6	Invariant mass distributions	83

In this chapter the procedures developed to reconstruct and select $B^+ \rightarrow D_s^+ \phi$ and $B^+ \rightarrow D_s^+ K^+ K^-$ candidates are described. In both cases the branching fractions are measured relative to the normalisation channel $B^+ \rightarrow D_s^+ \bar{D}^0$. The corresponding selection for the normalisation channel $B^+ \rightarrow D_s^+ \bar{D}^0$ is also described.

4.1 Data samples and Simulations

4.1.1 Data samples

The searches for $B^+ \rightarrow D_s^+ \phi$ and $B^+ \rightarrow D_s^+ K^+ K^-$ decays are performed using a combined data sample comprising of the entire Run I data set and a part of the Run II data set, namely the years 2015 and 2016. The total integrated luminosity obtained for each year is listed in Table 4.1, along with the corresponding centre-of-mass energies.

Year	Integrated luminosity (fb ⁻¹)	\sqrt{s} (TeV)
2011	1.0	7
2012	2.0	8
2015	0.3	13
2016	1.5	13

Table 4.1: The integrated luminosities obtained during the different data taking periods used in this analysis and the corresponding centre-of-mass energies (\sqrt{s}).

4.1.2 Simulation samples

Many aspects of the searches for $B^+ \rightarrow D_s^+ \phi$ and $B^+ \rightarrow D_s^+ K^+ K^-$ decays requires input from simulation samples. This includes the relative selection efficiencies of the signal and normalisation channels and the determination of the invariant mass distributions of the candidates. The samples are generated separately for each of the running periods listed in Table 4.1 and for both polarities of the LHCb dipole magnet. The simulation samples for $B^+ \rightarrow D_s^+ \phi$ and $B^+ \rightarrow D_s^+ \bar{D}^0$ decays are generated assuming they proceed via pseudo-two-body decays. In contrast, the phase-space dependence of $B^+ \rightarrow D_s^+ K^+ K^-$ decays must be considered when generating these simulations. Samples are generated assuming a flat distribution across the phase-space. Additionally, simulation samples are required for many of the background processes considered in the analyses. These are generated using the same framework and processed using the same reconstruction as the signal modes.

In the simulations, pp collisions are generated using PYTHIA [19,20] with a specific LHCb configuration [21]. Decays of hadronic particles are described by EVTGEN [22], in which final-state radiation is generated using PHOTOS [23]. The interaction of the generated particles with the detector, and its response, are implemented using the GEANT4 toolkit [24] as described in Ref. [25].

4.2 Online selection

The online event selection is performed by the LHCb trigger, which consists of a hardware stage, based on information from the calorimeter and muon systems, followed by a software stage, which reconstructs the full event, as detailed in Sec. 3.2.8. Trigger decisions are made globally for a given event; either the event is saved or it's discarded. However, for events that are retained more detailed information about the specific interaction that initiated the trigger can be used to help isolate events containing signal candidates. At both hardware and software level there are a number of different triggers that can lead to an event being retained. In hardware this is broadly separated according to the sub-detector that contained the triggering deposit. Deposits in the muon chambers or hadronic calorimeter would fire L0Muon or L0Hadron respectively. Electromagnetic calorimeter deposits could result in L0Electron or L0Photon firing. The $B^+ \rightarrow D_s^+ K^+ K^-$ and $B^+ \rightarrow D_s^+ \phi$ decays are both reconstructed in fully hadronic final states, therefore events retained as a result of the L0Electron, L0Photon or L0Muon trigger firing are unlikely to have been caused by the signal decays.

In addition to specifying which triggers fired, it is possible to use more granular information about the specific hits initiating the trigger to improve the selection of signal candidates. The tracks constituting the reconstructed candidate are matched to the deposits in each sub-detector. The fraction of hits required to match in order for the trigger decision to be associated to that object vary depending on the sub-detector in question, as listed in Table 4.2. Reconstructed objects are classified into categories when considering their relationship to the various triggers that fired in a given event.

Sub-detector	Fraction of matching hits
HCAL	1%
ECAL	1%
Muon	0.01%
TT	0%
IT and OT	70%
VELO	70%

Table 4.2: The fraction of matching hits required for a offline reconstructed object to be associated to a online sub-detector deposit.

Triggered on Signal: if the deposited matched to a reconstructed object was sufficient to have fired a given trigger then this object is said to be *Triggered on Signal* (**TOS**) with respect to that trigger (Fig. 4.1a). As such, if the hits that this object caused were removed from the events the trigger would no longer fire.

Triggered Independently of Signal: conversely, if the specific trigger fired solely due to other interactions within the same event the object is *Triggered Independently of Signal* (**TIS**) with respect to that trigger (Fig. 4.1b). In this case, removal of hits matched to the object would not affect the trigger decision. It is possible for a candidate to be both **TIS** and **TOS** (Fig. 4.1d).

Triggered on Both: a third category is possible, *Triggered on Both* (**TOB**), where both the signal object and another object are required to reach the threshold to fire a trigger (Fig. 4.1c). In this situation neither are sufficient to individually fire the trigger, but removing either one of them would prevent the trigger firing. Events in this category are not considered in this analysis.

At the hardware trigger stage, the selected candidates are required to be **TOS** with respect to the hadronic trigger **L0Hadron**. This ensures the selected candidates were retained due to corresponding deposits in the hadronic calorimeter. Alternatively, candidates are selected if they are **TIS** with respect to the global hardware trigger **L0Global**; any of the hardware trigger subsystems can contribute to the **L0Global**

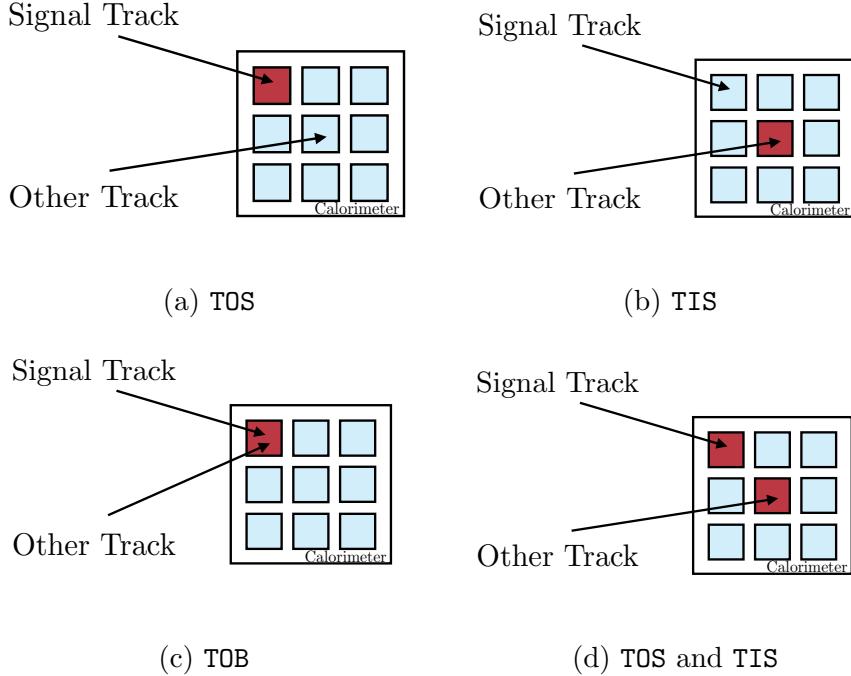


Figure 4.1: Simple schematics of the different trigger matching possibilities, illustrated with one signal track and one non-signal track. Cells with a deposit above the trigger threshold are shown in red.

decision. This allows candidates that have been retained due another highly energetic decay in the same event to contribute. This could be the decay of hadron resulting from the other b quark in a $b\bar{b}$ pair production. The relative fractions of simulated decays selected by these **L0Hadron TOS** and **L0Global TIS** for $B^+ \rightarrow D_s^+\phi$ decays decaying to various D_s^+ final states are shown in Table 4.3.

D_s^+ decay mode	TOS & !TIS	TIS&!TOS	TIS&TOS	TIS or TOS
$D_s^+ \rightarrow K^+ K^- \pi^+$	34.0%	41.2%	22.9%	96.9%
$D_s^+ \rightarrow K^+ \pi^- \pi^+$	34.9%	39.5%	23.8%	96.9%
$D_s^+ \rightarrow \pi^+ \pi^- \pi^+$	35.9%	38.7%	23.5%	96.9%

Table 4.3: The fraction of simulated $B^+ \rightarrow D_s^+\phi$ decays in each of the hardware trigger categories for the Run I conditions. Here **TOS** refers to candidates found to be **TOS** with respect to the **L0Hadron** trigger, and **TIS** to candidates found to be **TIS** with respect to the **L0Global** requirement. The percentages are given relative to the number of decays passing the reconstruction requirements.

The software trigger stage is split into two parts, HLT1 and HLT2. The first

stage, HLT1, requires that the selected candidates are associated with at least one well reconstructed track. This track must be significantly displaced from the primary interaction and has a large transverse momentum. In Run II, this trigger was reimplementing using a multivariate analysis trained to select a well reconstructed, high transverse momentum track that was significantly displaced from the primary interaction. An additional trigger was also included to select pairs of high transverse momentum tracks that form a vertex that is displaced from the primary interaction. The signal candidate is required to be TOS with respect to these trigger lines such that one of the candidate tracks passes this requirement.

At the second software stage, HLT2, two different algorithms are used to select candidates for this analysis. The first one uses a multivariate algorithm [26] to identify the presence of a secondary vertex that has two, three or four tracks and is displaced from any PV. At least one of these charged particles must have a transverse momentum $p_T > 1.7 \text{ GeV}/c$ and be inconsistent with originating from a PV. This trigger algorithm is referred to as the topological trigger. The second algorithm selects ϕ candidates decaying to two charged kaons. Each kaon must have a transverse momentum $p_T > 0.8 \text{ GeV}/c$ and be inconsistent with originating from a PV. The invariant mass of the kaon pair must be within $20 \text{ MeV}/c^2$ of the known ϕ mass [27]. This inclusive ϕ line is used to maximise the selection efficiencies as ϕ mesons can contribute directly to $B^+ \rightarrow D_s^+ \phi$ decays, as well as via the large fractions of $D_s^+ \rightarrow \phi \pi^+$ decays that contribute to the $D_s^+ \rightarrow K^+ K^- \pi^+$ final state. The relative fractions of simulated $B^+ \rightarrow D_s^+ \phi$ decays selected by these different trigger algorithms are listed in Table 4.4. It can be seen that the inclusive ϕ trigger algorithm helps additionally retain a larger fraction of decays for the $D_s^+ \rightarrow K^+ K^- \pi^+$ decay.

4.3 Offline selection

Events passing any trigger requirement are saved to tape for processing offline. The stages of the offline reconstruction are detailed in this section.

D_s^+ decay mode	Topo & !Phi	Phi&!Topo	Topo&Phi	Topo or Phi
$D_s^+ \rightarrow K^+ K^- \pi^+$	50.9%	4.6%	43.1%	98.7%
$D_s^+ \rightarrow K^+ \pi^- \pi^+$	67.6%	2.7%	28.3%	98.5%
$D_s^+ \rightarrow \pi^+ \pi^- \pi^+$	67.5%	2.6%	28.5%	98.7%

Table 4.4: The fraction of simulated $B^+ \rightarrow D_s^+ \phi$ decays in each of the HLT2 software trigger categories for the Run I conditions. Here **Topo** refers to candidates found to be **TOS** with respect to the topological algorithm, and **Phi** to candidates found to be **TOS** with respect to the inclusive ϕ algorithm. The fractions of selected decays are measured relative to the HLT1 requirements.

4.3.1 Selection requirements

The large offline data samples passing the online trigger selection are habitually processed in a procedure known within LHCb as *Stripping*. This centrally managed processing builds candidates from tracks and neutral calorimeter objects in each event according to a set of predefined *Stripping Lines*. Each line builds a specific candidate decay, applying a loose set of preselection requirements, including kinematic, geometric and invariant mass selections.

The candidate $B^+ \rightarrow D_s^+ \phi$ and $B^+ \rightarrow D_s^+ K^+ K^-$ decays are built using almost identical set of requirements due to the similar topologies of these decays. The ϕ meson has a lifetime of $(1.55 \pm 0.01) \times 10^{-22}$ s [27], therefore the kaon decay products effectively originate at the B^+ decay vertex in a similar way to $B^+ \rightarrow D_s^+ K^+ K^-$ decays. The normalisation channel, $B^+ \rightarrow D_s^+ \bar{D}^0$ is built similarly, however the difference in lifetime between the \bar{D}^0 and ϕ mesons necessitates a slightly different topology. The \bar{D}^0 meson has a lifetime of $(4.101 \pm 0.015) \times 10^{-13}$ s, allowing the meson to travel away from the B^+ decay vertex. Schematics of the signal and normalisation decay topologies are shown in Fig. 4.2.

The candidate $B^+ \rightarrow D_s^+ K^+ K^-$ and $B^+ \rightarrow D_s^+ \phi$ signal decays and $B^+ \rightarrow D_s^+ \bar{D}^0$ normalisation decay are reconstructed in fully hadronic final states. Both the \bar{D}^0 and ϕ mesons are reconstructed using pairs of oppositely charged kaons. The branching fractions are $\mathcal{B}(\bar{D}^0 \rightarrow K^+ K^-) = (3.97 \pm 0.07) \times 10^{-3}$ and $\mathcal{B}(\phi \rightarrow K^+ K^-) = (48.9 \pm 0.5)\%$ respectively [27]. Although the similar two-body hadronic decay $D^0 \rightarrow K^- \pi^+$

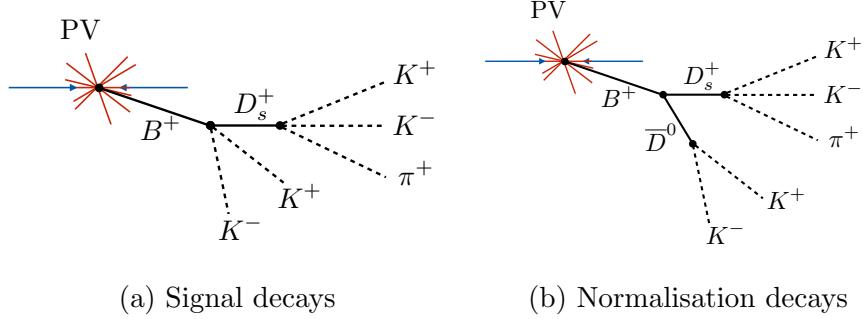


Figure 4.2: Schematics of the signal and normalisation decay topologies. The collision of the two protons (blue) results in the primary collision vertex (PV). Many promptly produced tracks (red) originate at the PV. In both cases, the long-lived B^+ meson decay vertex is displaced from the PV. The long-lived charm mesons are also displaced from the B^+ meson decay vertex.

has a larger branching fraction than mode chosen for the normalisation channel, $\mathcal{B}(D^0 \rightarrow K^-\pi^+) = (3.89 \pm 0.04)\%$, sharing the same final state helps to reduced systematic uncertainties in the ratio of selection efficiencies. Many differences in how pions and kaons interact with the detector can be neglected as they would affect the signal and normalisation channel in the same way. The D_s^+ mesons used in the search for $B^+ \rightarrow D_s^+ K^+ K^-$ decays are reconstructed using the $D_s^+ \rightarrow K^+ K^- \pi^+$ decay. The search for $B^+ \rightarrow D_s^+ \phi$ additionally includes the modes $D_s^+ \rightarrow \pi^+ \pi^- \pi^+$ and $D_s^+ \rightarrow K^+ \pi^- \pi^+$ to increase the sensitivity of the search. The branching fractions for these decays are listed in Table 4.5. In each instance the normalisation channel is reconstructed using the same D_s^+ decay mode as the signal.

Decay	Branching fraction
$D_s^+ \rightarrow K^+ K^- \pi^+$	$5.45 \pm 0.17\%$
$D_s^+ \rightarrow \pi^+ \pi^- \pi^+$	$1.09 \pm 0.05\%$
$D_s^+ \rightarrow K^+ \pi^- \pi^+$	$0.66 \pm 0.04\%$

Table 4.5: The branching fractions for the different D_s^+ decay modes used to reconstruct the signal and normalisation modes.

The selection of candidates makes requirements on many different distinguishing characteristics. The definitions of the relevant quantities are as follows:

Mass, $m(X)$: the invariant mass of the indicated particle or combination of particles.

Momentum, p : the total momentum of the particle or combination of particles.

Transverse momentum, p_T : The projection of the total momentum, p , on the plane perpendicular to the proton beam axis.

Lifetime, τ : the time taken for the candidate to travel from its production vertex to its decay vertex.

Products p_T scalar sum, $\sum |p_T|$: the sum of the magnitudes of the transverse momentum for each of the decay products.

Vertex quality, χ^2_{Vtx} : the minimised value of χ^2/N_{DOF} as determined in the fit to the vertex location.

Track quality, χ^2_{Trk} : the minimised value of χ^2/N_{DOF} as determined in the fit track hits.

Impact parameter, IP: The shortest distance between a given track and a vertex location, as shown in Fig. 4.3.

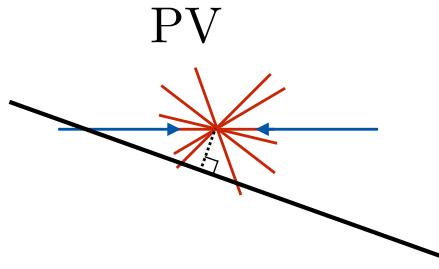


Figure 4.3: Impact parameter (dotted black line) between a track (black) and a vertex (red lines).

Impact parameter significance, χ^2_{IP} : The difference of a given vertex's χ^2/N_{DOF} with and without a specific track included in the fitting procedure.

Flight distance significance, χ^2_{FD} : A measure of how significant the flight distance of a combination of particle is. Defined as the χ^2 associated with the difference in position of the two vertices, $\vec{d} = \vec{v}_2 - \vec{v}_1$, where \vec{v}_1 and \vec{v}_2 are positions of the first and the second vertices.

Distance of closest approach, DOCA(h, h'): The shortest distance between two tracks as shown in Fig 4.4.

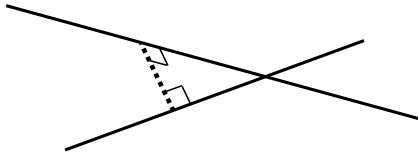


Figure 4.4: Distance of closest approach (dotted black line) between two tracks (black).

Ghost track probability, P_{Ghost} : This parameter quantifies the probability that a given track is an incorrect combination of tracking stations hits, known as a ghost track. The numerical value is the output of a Neural Network algorithm trained to separate true tracks from ghost tracks using simulated events. Various tracking parameters are inputs to the Neural Network including the number of hits in various tracking stations, the track fit quality and the number of tracks per event.

Direction angle: The angle between the particles momentum vector and the vector connecting the primary vertex and decay vertex as shown in Fig. 4.5.

Particle identification, PIDK: Something

As the final states are fully hadronic, the candidates are built from the combination of five tracks. Only *long* tracks (those with hits in the Velo and tracking stations) are used to build these mesons. The track χ^2/N_{DOF} is required to be below 4.0 to ensure these are well reconstructed. Additionally, they are required to have a total

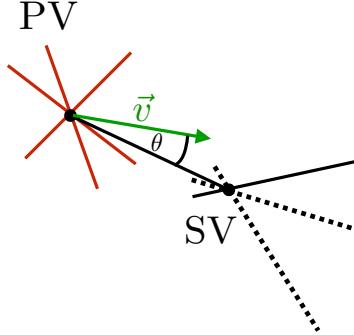


Figure 4.5: Direction angle, θ , between the line joining two vertices, primary vertex (PV) and secondary vertex (SV), and the momentum vector shown in green. The momentum vector corresponds to the momentum of the three tracks contributing to the SV.

momentum $p > 1000 \text{ MeV}/c$ and the transverse momentum is required to be $p_T > 100 \text{ MeV}/c$. Due to the relatively long lifetime of the B^+ and D mesons, the decay products originate from a vertex that is displaced from the proton-proton collision vertex. Therefore it is possible that the trajectory that the decay products followed didn't pass through the collision position as shown in Fig. 4.3. A requirement is placed on the significance of the impact parameter between the track and the proton-proton collision vertex of $\chi_{\text{IP}}^2 > 4$ to ensure all of the tracks used are inconsistent with originating at the primary interaction.

Loose requirements are placed on *Particle Identification* variables to ensure the tracks are of the required species. These are further tightened as detailed in Section 4.3.2. Incorrect track candidates created by combining unrelated VELO and tracking station hits are suppressed by requiring the ghost track probability, P_{Ghost} , to be less than 0.4.

The tracks passing these requirements are combined in pairs or threes to form the D_s^+ and ϕ or \bar{D}^0 meson candidates. To create the ϕ candidates (or K^+K^- pair), all combinations of two tracks are considered. The two tracks are assigned the mass hypothesis of kaons. Only pairs which pass within 0.5 mm of one another at their closest point are retained. A number of requirements are imposed to ensure the pair are consistent with coming from a ϕ meson (or K^+K^- pair) originating from

a B^+ meson decay. The vertex is required to be of good quality, the scalar sum of the transverse momentums must be greater than $1000 \text{ MeV}/c$ and the flight distance significance is required to be $\chi_{\text{FD}}^2 > 16$. Additionally the cosine of the direction angle, defined to be the angle between the momentum vector and flight vector, is required to be $\cos \theta > 0$, preventing the vertex being backward of the PV when the momentum is in the forward direction. In the search for $B^+ \rightarrow D_s^+ \phi$ mesons, the invariant mass of the two kaons is required to be within $150 \text{ MeV}/c^2$ of the known ϕ meson mass.

The \bar{D}^0 candidates are selecting using a similar parameters, however the exact values of the cuts are changed to reflect the differences in properties of the ϕ and \bar{D}^0 mesons. The vertex quality and flight distance significance requirements are tightened to $\chi^2/N_{\text{DOF}} < 10$ and $\chi_{\text{FD}}^2 > 36$ respectively. Additionally, the scalar p_{T} sum requirement is increased to $\sum |p_{\text{T}}| > 1800 \text{ MeV}/c$ as the decay products tend to have higher transverse momentum.

The D_s^+ mesons for the signal and normalisation modes are created using combinations of three tracks given the mass hypotheses kaons or pions, depending on the decay mode being constructed. All combinations of three tracks are considered and only those in which all three tracks are within 0.5 mm of one another are retained. Similar to the \bar{D}^0 meson the vertex quality, flight distance significance and scalar p_{T} sum requirements are $\chi^2/N_{\text{DOF}} < 10$, $\chi_{\text{FD}}^2 > 36$ and $\sum |p_{\text{T}}| > 1800 \text{ MeV}/c$ respectively.

The B^+ meson candidates are constructed by creating all possible combinations of the D_s^+ and ϕ or \bar{D}^0 meson candidates in each event. Requirements are placed on these combination to select just those consistent with B^+ mesons; the combination is required to have a lifetime $\tau_{B^+} > 0.2 \text{ ps}$ and direction angle $\cos \theta > 0.999$. The impact parameter significance is required to be $\chi_{\text{IP}}^2 < 25$ to ensure the B^+ meson originated at the primary interaction. The vertex is required to have a quality of $\chi^2/N_{\text{DOF}} < 10$. More requirements are additionally placed on the decay products that contribute to the B^+ meson. The selection requirements imposed on candidate $B^+ \rightarrow D_s^+ \phi$ and $B^+ \rightarrow D_s^+ K^+ K^-$ decays in their respective *Stripping Lines* are detailed in Table 4.6.

Two slightly different strategies are used for the normalisation channel selection in the search for $B^+ \rightarrow D_s^+ \phi$ and $B^+ \rightarrow D_s^+ K^+ K^-$ events. In the former, a dedicated $B^+ \rightarrow D_s^+ \bar{D}^0$ *Stripping Line* is used to reconstruct the normalisation channel decays. The *Stripping Line* selection for this line is listed in Table 4.7.

The *Stripping Line* used in the search for $B^+ \rightarrow D_s^+ K^+ K^-$ decays covers the full $m(K^+ K^-)$ phase-space. This includes the \bar{D}^0 mass such that this line reconstructs both the signal and normalisation channels simultaneously. Both modes are selected using this line to reduce systematic uncertainty in the ratio of selection efficiencies.

Particles	Quantity	Requirement
B^+	Mass	$4750 < m(D_s^+ \phi) < 7000 \text{ MeV}/c^2$
	Products p_T scalar sum	$\sum p_T > 5000 \text{ MeV}/c$
	Vertex quality	$\chi^2/N_{\text{DOF}} < 10$
	Lifetime	$\tau_{B^+} > 0.2 \text{ ps}$
	Impact parameter significance	$\chi^2_{\text{IP}} < 25$
	Direction angle	$\cos \theta > 0.999$
	>0 decay products with:	
	Momentum	$p > 10000 \text{ MeV}/c$
	Transverse momentum	$p_T > 1700 \text{ MeV}/c$
	Impact parameter significance	$\chi^2_{\text{IP}} > 16$
D_s^+	Impact parameter	$\text{IP} > 0.1 \text{ mm}$
	>1 decay products with:	
	Momentum	$p > 5000 \text{ MeV}/c$
	Transverse momentum	$p_T > 500 \text{ MeV}/c$
	Mass	$1770 < m(h^+ h^- h^+) < 2068 \text{ MeV}/c^2$
	Products p_T scalar sum	$\sum p_T > 1800 \text{ MeV}/c$
	Distance of closest approach	$\text{DOCA}(h^+, h^-) < 0.5 \text{ mm}$
	Distance of closest approach	$\text{DOCA}(h^-, h'^+) < 0.5 \text{ mm}$
	Distance of closest approach	$\text{DOCA}(h^+, h'^+) < 0.5 \text{ mm}$
	Direction angle	$\cos \theta > 0$
ϕ	Vertex quality	$\chi^2/N_{\text{DOF}} < 10$
	Flight distance significance	$\chi^2_{\text{FD}} > 36$
	Mass (only for $B^+ \rightarrow D_s^+ \phi$)	$ m(K^+ K^-) - m_\phi < 150 \text{ MeV}/c^2$
	Products p_T scalar sum	$\sum p_T > 1000 \text{ MeV}/c$
	Distance of closest approach	$\text{DOCA}(K^+, K^-) < 0.5 \text{ mm}$
	Direction angle	$\cos \theta > 0$
$K^\pm (\pi^\pm)$	Vertex quality	$\chi^2/N_{\text{DOF}} < 16$
	Flight distance significance	$\chi^2_{\text{FD}} > 16$
	Track quality	$\chi^2/N_{\text{DOF}} < 4.0$
	Transverse momentum	$p_T > 100 \text{ MeV}/c$
	Momentum	$p > 1000 \text{ MeV}/c$
	Impact parameter significance	$\chi^2_{\text{IP}} > 4$

Table 4.6: Selection requirements for $B^+ \rightarrow D_s^+ \phi$ and $B^+ \rightarrow D_s^+ K^+ K^-$ candidates.

Particles	Quantity	Requirement
B^+	Mass	$4750 < m(D_s^+ \bar{D}^0) < 7000 \text{ MeV}/c^2$
	Products p_T scalar sum	$\sum p_T > 5000 \text{ MeV}/c$
	Vertex quality	$\chi^2/N_{\text{DOF}} < 10$
	Lifetime	$\tau_{B^+} > 0.2 \text{ ps}$
	Impact parameter significance	$\chi_{\text{IP}}^2 < 25$
	Direction angle	$\cos \theta > 0.999$
	>0 decay products with:	
	Momentum	$p > 10000 \text{ MeV}/c$
	Transverse momentum	$p_T > 1700 \text{ MeV}/c$
	Impact parameter significance	$\chi_{\text{IP}}^2 > 16$
	Impact parameter	$\text{IP} > 0.1 \text{ mm}$
	>1 decay products with:	
	Momentum	$p > 5000 \text{ MeV}/c$
	Transverse momentum	$p_T > 500 \text{ MeV}/c$
D_s^+	Mass	$1770 < m(h^+ h^- h^+) < 2068 \text{ MeV}/c^2$
	Products p_T scalar sum	$\sum p_T > 1800 \text{ MeV}/c$
	Distance of closest approach	$\text{DOCA}(h^+, h^-) < 0.5 \text{ mm}$
	Distance of closest approach	$\text{DOCA}(h^-, h'^+) < 0.5 \text{ mm}$
	Distance of closest approach	$\text{DOCA}(h^+, h'^+) < 0.5 \text{ mm}$
	Direction angle	$\cos \theta > 0$
	Vertex quality	$\chi^2/N_{\text{DOF}} < 10$
	Flight distance significance	$\chi_{\text{FD}}^2 > 36$
\bar{D}^0	Mass	$1765 < m(h^+ h^- h^+) < 1965 \text{ MeV}/c^2$
	Products p_T scalar sum	$\sum p_T > 1800 \text{ MeV}/c$
	Distance of closest approach	$\text{DOCA}(K^+, K^-) < 0.5 \text{ mm}$
	Direction angle	$\cos \theta > 0$
	Vertex quality	$\chi^2/N_{\text{DOF}} < 10$
	Flight distance significance	$\chi_{\text{FD}}^2 > 36$
$K^\pm (\pi^\pm)$	Track quality	$\chi^2/N_{\text{DOF}} < 4.0$
	Transverse momentum	$p_T > 100 \text{ MeV}/c$
	Momentum	$p > 1000 \text{ MeV}/c$
	Impact parameter significance	$\chi_{\text{IP}}^2 > 4$
	Ghost track probability	$P_{\text{Ghost}} < 0.4$
	Particle identification	$\text{PIDK} > -10 \text{ (PIDK} < 20)$

Table 4.7: Selection requirements for $B^+ \rightarrow D_s^+ \bar{D}^0$ candidates.

4.3.2 Particle identification requirements

Particle identification variables help to determine the species of tracks passing through the LHCb detector. Using information from the RICH sub-detectors, the likelihood of different mass hypotheses are compared to the pion hypothesis. Loose requirements are made on the kaon hypothesis PID variable to reduce the contribution from other types of hadrons and background from other b -hadron decays with misidentified hadrons. The parameter is defined to be

$$\text{PIDK} = \Delta \log(K - \pi) = \log \mathcal{L}(K) - \log \mathcal{L}(\pi) \quad (4.1)$$

where $\mathcal{L}(K)$ and $\mathcal{L}(\pi)$ are the likelihoods of the kaon and pion hypotheses respectively. The requirements placed on each decay product are listed in Table 4.8.

Decay mode	Species	PID requirement
$\phi \rightarrow K^+ K^-$	K^+	$\text{PIDK} > 0$
	K^-	$\text{PIDK} > 0$
$\bar{D}^0 \rightarrow K^+ K^-$	K^+	$\text{PIDK} > 0$
	K^-	$\text{PIDK} > 0$
$D_s^+ \rightarrow K^+ K^- \pi^+$	K^+	$\text{PIDK} > -5$
	K^-	$\text{PIDK} > -5$
	π^+	$\text{PIDK} < 5$
$D_s^+ \rightarrow \pi^+ \pi^- \pi^+$	π^+	$\text{PIDK} < 5$
	π^-	$\text{PIDK} < 5$
	π^+	$\text{PIDK} < 5$
$D_s^+ \rightarrow K^+ \pi^- \pi^+$	K^+	$\text{PIDK} > -5$
	π^-	$\text{PIDK} < 5$
	π^+	$\text{PIDK} < 5$

Table 4.8: Particle identification requirements applied to kaons and pions.

4.3.3 Charmless and single-charm backgrounds

Decays of B^+ mesons that didn't proceed via D mesons could form a peaking background below the signal invariant mass distributions when they decay to the same final state. The signal mode could receive contributions from the decays $B^+ \rightarrow h^+ h^- h^+ \phi$ or $B^+ \rightarrow h^+ h^- h^+ h^-$, referred to as charmless backgrounds. Here h^\pm is used to

represent K^\pm or π^\pm in the specific D_s^+ , \bar{D}^0 or ϕ final state. The normalisation mode is also susceptible, however as it involves two charm mesons it could receive contributions from the decays $B^+ \rightarrow h^+ h^- h^+ \bar{D}^0$ or $B^+ \rightarrow D_s^+ h^+ h^-$, referred to as single-charm backgrounds, and $B^+ \rightarrow h^+ h^- h^+ h^-$ referred to as a charmless background. These backgrounds can be suppressed by requiring the D meson decay vertex to be displaced from the B^+ meson decay vertex. Requirements are applied to the significance of the vertex separation (χ_{FD}^2).

The residual yields of charmless backgrounds in the signal mode are estimated by performing a fit to the B^+ invariant mass for candidates with $25 < |m(h^+ h^- h^+) - m(D_s^+)| < 50 \text{ MeV}/c^2$. This background estimation is performed separately for the $B^+ \rightarrow D_s^+ \phi$ and $B^+ \rightarrow D_s^+ K^+ K^-$ searches.

For the $B^+ \rightarrow D_s^+ \bar{D}^0$ normalisation channel, a two-dimensional optimisation is performed to calculate the contribution from decays without a D_s^+ meson, \bar{D}^0 meson or both. The two-dimensional space defined by the D_s^+ and \bar{D}^0 masses is split into four types of area as shown in Fig 4.6.

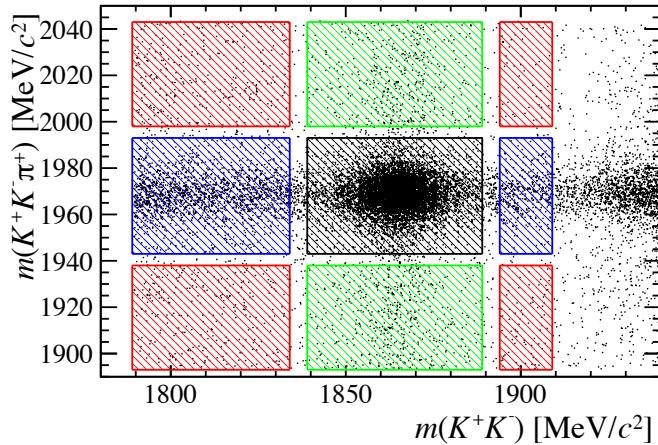


Figure 4.6: Two dimensional normalisation

1. Areas in which $B^+ \rightarrow h^+ h^- h^+ h^-$ decays contribute (red).
2. Areas in which either $B^+ \rightarrow D_s^+ h^+ h^-$ or $B^+ \rightarrow h^+ h^- h^+ h^-$ decays can contribute (blue).

3. Areas in which either $B^+ \rightarrow h^+h^-h^+\bar{D}^0$ or $B^+ \rightarrow h^+h^-h^+h^-$ decays can contribute (green).
4. The signal region in which $B^+ \rightarrow D_s^+\bar{D}^0$, $B^+ \rightarrow h^+h^-h^+\bar{D}^0$, $B^+ \rightarrow D_s^+h^+h^-$ or $B^+ \rightarrow h^+h^-h^+h^-$ decays could contribute (black).

Asymmetric \bar{D}^0 sidebands are used to prevent misidentified $B^+ \rightarrow D_s^+(\bar{D}^0 \rightarrow K^-\pi^+)$ decays from being included in the sideband sample. The optimal selection requirements are chosen such that the maximal signal efficiency is achieved for a residual charmless contribution of 2% of the normalisation yield.

The optimisation of the signal and normalisation cuts is performed separately for each different D_s^+ decay mode, and for the $B^+ \rightarrow D_s^+K^+K^-$ and $B^+ \rightarrow D_s^+\phi$ selections. The optimised requirements are listed in Table 4.9, along with the estimated residual yields of charmless and single charm yields in the signal region.

B^+ decay mode	D_s^+ decay mode	$\chi_{\text{FD}}^2(D_s^+)$	$\chi_{\text{FD}}^2(\bar{D}^0)$	Residual yields
$B^+ \rightarrow D_s^+\phi$	$D_s^+ \rightarrow K^+K^-\pi^+$	0.0	-	0.0
$B^+ \rightarrow D_s^+\phi$	$D_s^+ \rightarrow K^+\pi^-\pi^+$	25.0	-	2.6
$B^+ \rightarrow D_s^+\phi$	$D_s^+ \rightarrow \pi^+\pi^-\pi^+$	5.0	-	0.0
$B^+ \rightarrow D_s^+\bar{D}^0$	$D_s^+ \rightarrow K^+K^-\pi^+$	8.0	0.0	21.6
$B^+ \rightarrow D_s^+\bar{D}^0$	$D_s^+ \rightarrow K^+\pi^-\pi^+$	18.0	0.0	3.3
$B^+ \rightarrow D_s^+\bar{D}^0$	$D_s^+ \rightarrow \pi^+\pi^-\pi^+$	16.0	0.0	3.9
$B^+ \rightarrow D_s^+K^+K^-$	$D_s^+ \rightarrow K^+K^-\pi^+$	5.0	-	0.19
$B^+ \rightarrow D_s^+\bar{D}^0$	$D_s^+ \rightarrow K^+K^-\pi^+$	8.0	0.0	7.95

Table 4.9: Charmless and single charm minimum flight distance significance requirements applied to the D_s^+ and \bar{D}^0 candidates.

4.3.4 Misidentified D and Λ_c^+ hadrons

It is possible for the samples D_s^+ mesons to be contaminated by other misidentified decays of D^+ mesons or Λ_c^+ baryons in which one of the decay products has been incorrectly identified. The invariant mass of the D_s^+ meson is recalculated, swapping the mass hypothesis of the ambiguous track to that of the K , π or p , depending on the decay mode. The particle identification requirements are tightened within a mass

window around the D^+ or Λ_c^+ mass, effectively removing this crossfeed. For the mode $D_s^+ \rightarrow K^+ K^- \pi^+$, the vetoes are not applied to candidates for which $|m(K^- K^+) - m_\phi| < 10 \text{ MeV}/c^2$ as there are a high purity of $D_s^+ \rightarrow K^+ K^- \pi^+$ decays in this region.

The specific vetoes included in this selection are listed in Table 4.10. The invariant

Decay Mode	Misidentified decay
$D_s^+ \rightarrow \textcolor{red}{K}^+ K^- \pi^+$	$D^+ \rightarrow \pi^+ K^- \pi^+$
	$\Lambda_c^+ \rightarrow \textcolor{red}{p} K^- \pi^+$
$D_s^+ \rightarrow \textcolor{red}{K}^+ \pi^- \pi^+$	$D^+ \rightarrow \pi^+ \pi^- \pi^+$

Table 4.10: Misidentified decays targeted by vetoes. The ambiguous track is highlighted in red in each case.

mass distributions for each the misidentified $D_s^+ \rightarrow K^+ K^- \pi^+$ decays are shown with and without the MVA requirements in Figs. 4.7 and 4.8 for both the signal $B^+ \rightarrow D_s^+ \phi$ and normalisation $B^+ \rightarrow D_s^+ \bar{D}^0$ decays.

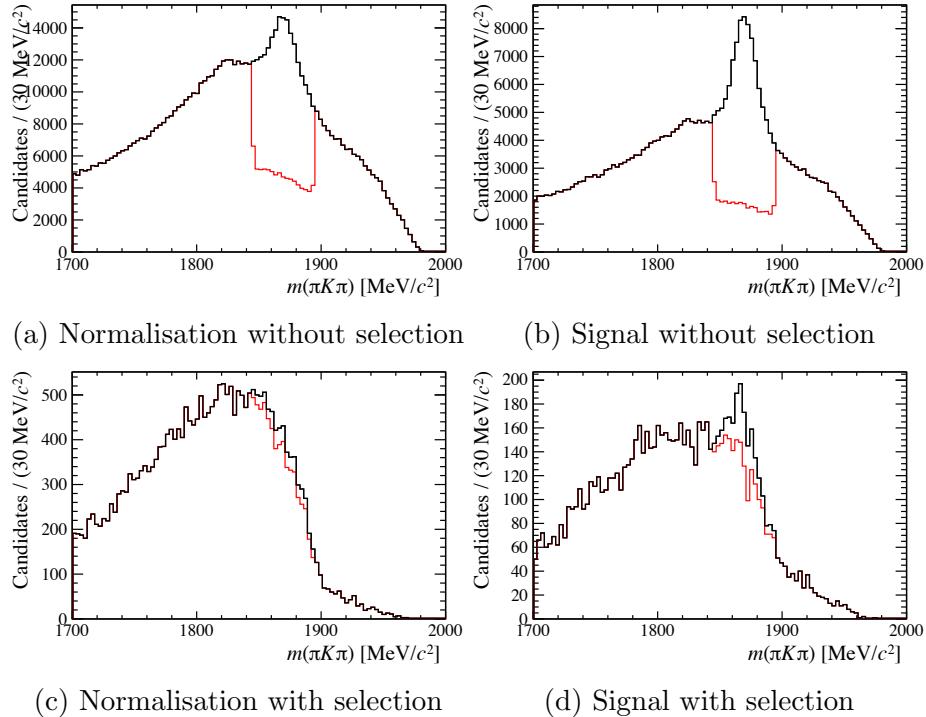


Figure 4.7: Invariant mass distributions of $D_s^+ \rightarrow K^+ K^- \pi^+$ samples reconstructed as $D^+ \rightarrow \pi^+ K^- \pi^+$ for the signal and normalisation samples. The samples are shown with (red) and without (black) the veto described in Sec. 4.3.4. The distributions are shown before (top) and after (bottom) the MVA requirements have been applied.

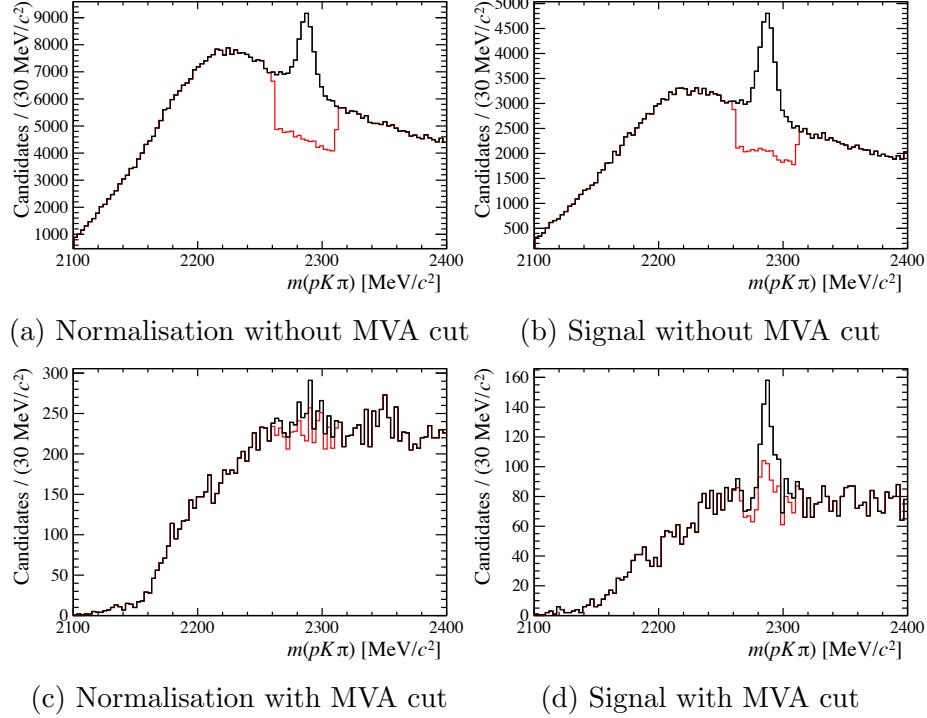


Figure 4.8: Invariant mass distributions of $D_s^+ \rightarrow K^+ K^- \pi^+$ samples reconstructed as $\Lambda_c^+ \rightarrow p K^- \pi^+$ for the signal and normalisation samples. The samples are shown with (red) and without (black) the veto described in Sec. 4.3.4. The distributions are shown before (top) and after (bottom) the MVA requirements have been applied.

4.3.5 Invariant mass vetoes

Sharp peaking structures are observed in subsets of the final state particles. These are removed with simple invariant mass cuts to remove combinatorial or partially reconstructed backgrounds that result from these incorrectly reconstructed decays. For simplicity the final state particles for each mode are labelled with a number between 1–5 as described in Table 4.11.

Decay Mode	1	2	3	4	5
$B^+ \rightarrow (D_s^+ \rightarrow K^+ K^- \pi^+) \phi$	K^+	K^-	π^+	K^+	K^-
$B^+ \rightarrow (D_s^+ \rightarrow \pi^+ \pi^- \pi^+) \phi$	π^+	π^-	π^+	K^+	K^-
$B^+ \rightarrow (D_s^+ \rightarrow K^+ \pi^- \pi^+) \phi$	K^+	π^-	π^+	K^+	K^-
$B^+ \rightarrow (D_s^+ \rightarrow K^+ K^- \pi^+) K^+ K^-$	K^+	K^-	π^+	K^+	K^-

Table 4.11: Particle labels used when studying invariant mass vetoes for $B^+ \rightarrow D_s^+ \phi$ and $B^+ \rightarrow D_s^+ K^+ K^-$ candidates.

All combinations of the final state particles that create a neutral or singly-charged candidate are investigated. Significant structures are observed for all three D_s^+ decay modes in some combination.

The following vetos are applied to remove these incorrectly reconstructed decays.

- For the mode $B^+ \rightarrow (D_s^+ \rightarrow K^+ K^- \pi^+) \phi$
 - $|m(1245) - m(B_s^0)| > 50 \text{ MeV}/c^2$
 - $|m(345) - m(D_s^+)| > 25 \text{ MeV}/c^2$ and $|m(345) - m(D^+)| > 25 \text{ MeV}/c^2$
- For the mode $B^+ \rightarrow (D_s^+ \rightarrow \pi^+ \pi^- \pi^+) \phi$
 - $|m(145) - m(D_s^+)| > 25 \text{ MeV}/c^2$ and $|m(145) - m(D^+)| > 25 \text{ MeV}/c^2$
 - $|m(245) - m(D_s^+)| > 25 \text{ MeV}/c^2$ and $|m(245) - m(D^+)| > 25 \text{ MeV}/c^2$
 - $|m(345) - m(D_s^+)| > 25 \text{ MeV}/c^2$ and $|m(345) - m(D^+)| > 25 \text{ MeV}/c^2$
- For the mode $B^+ \rightarrow (D_s^+ \rightarrow K^+ \pi^- \pi^+) \phi$
 - $|m(245) - m(D_s^+)| > 25 \text{ MeV}/c^2$ and $|m(245) - m(D^+)| > 25 \text{ MeV}/c^2$
 - $|m(345) - m(D_s^+)| > 25 \text{ MeV}/c^2$ and $|m(345) - m(D^+)| > 25 \text{ MeV}/c^2$

In the search for $B^+ \rightarrow D_s^+ K^+ K^-$ decays the increased size of the $m(K^+ K^-)$ phase-space means more of the combinations of final state particles are susceptible to sharp peaking structure from incorrectly reconstructed backgrounds. Those spectra found to have significant peaking structures are additionally vetoed, as shown in Fig. 4.10.

- Vetoes for the mode $B^+ \rightarrow (D_s^+ \rightarrow K^+ K^- \pi^+) K^+ K^-$:
 - $|m(1245) - m(B_s^0)| > 50 \text{ MeV}/c^2$
 - $|m(345) - m(D_s^+)| > 25 \text{ MeV}/c^2$ and $|m(345) - m(D^+)| > 25 \text{ MeV}/c^2$
 - $|m(135) - m(D_s^+)| > 25 \text{ MeV}/c^2$
 - $|m(234) - m(D_s^+)| > 25 \text{ MeV}/c^2$

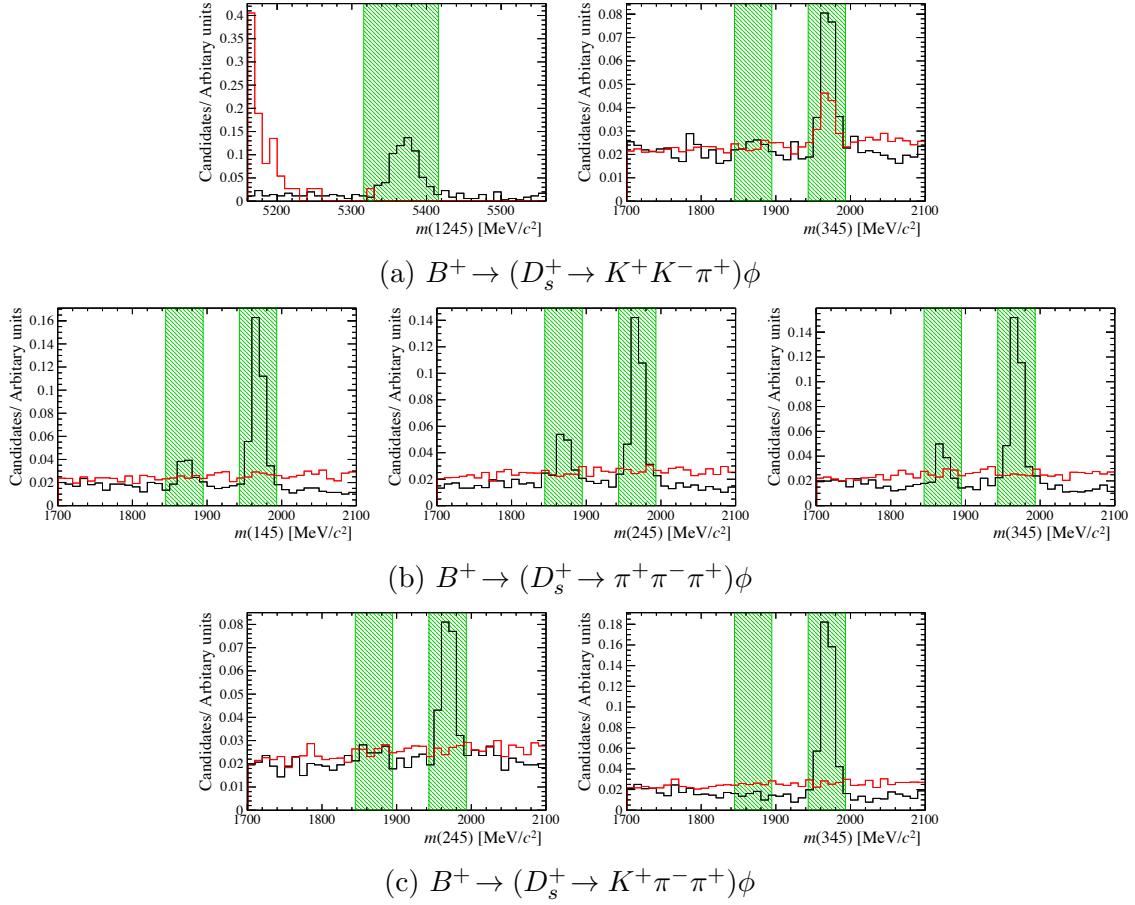


Figure 4.9: Invariant mass distributions for subsets of decay products in data (black) and simulation (red). The green region show the regions removed by the vetoes listed in Sec 4.3.5.

Another set of vetoes rejects decays where the tracks forming the D_s^+ candidate originate from an excited charged charm meson decay, for example $D^{*+} \rightarrow (D^0 \rightarrow h^+ h'^-) \pi^+$. By requiring $\Delta m = m(h^+ h'^- \pi^+) - m(h^+ h'^-) > 150 \text{ MeV}/c^2$ decays of this type are efficiently removed. These are applied to both the signal and normalisation modes for all D_s^+ decays.

4.3.6 Normalisation mode veto

In the search for $B^+ \rightarrow D_s^+ K^+ K^-$ decays, the entire $m(K^+ K^-)$ phasespace is used. This ranges from the $K^+ K^-$ mass threshold at around $990 \text{ MeV}/c^2$ to the kinematic limit at $m(B^+) - m(D_s^+) = 3300 \text{ MeV}/c^2$. This range is wide enough to include the mass of the \bar{D}^0 meson, $m(D^0) = 1864 \text{ MeV}/c^2$. Consequently, when inspecting the

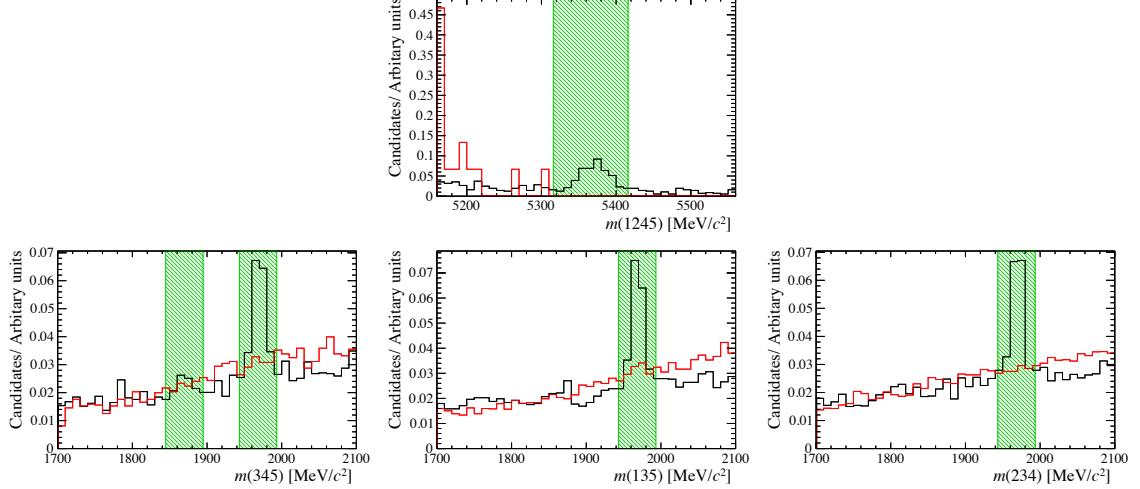


Figure 4.10: Invariant mass distributions for subsets of decay products for $B^+ \rightarrow (D_s^+ \rightarrow K^+ K^- \pi^+) K^+ K^-$ decays in data (black) and simulation (red). The green region show the regions removed by the vetoes listed in Sec 4.3.5.

$m(K^+ K^-)$ spectrum for selected signal candidates there is an excess of events at the D^0 mass. It is necessary to remove these from the signal samples as, unsurprisingly, they result in a peak at the B^+ mass in the $m(D_s^+ K^+ K^-)$ spectrum and lead to an incorrect signal yield. As described in Section 4.3.1, these removed events are used as the normalisation channel for the measurement of $B^+ \rightarrow D_s^+ K^+ K^-$ decays. The region affected by the veto $|m(K^+ K^-) - m(\bar{D}^0)| > 25 \text{ MeV}/c^2$ is shown in Fig. 4.11a.

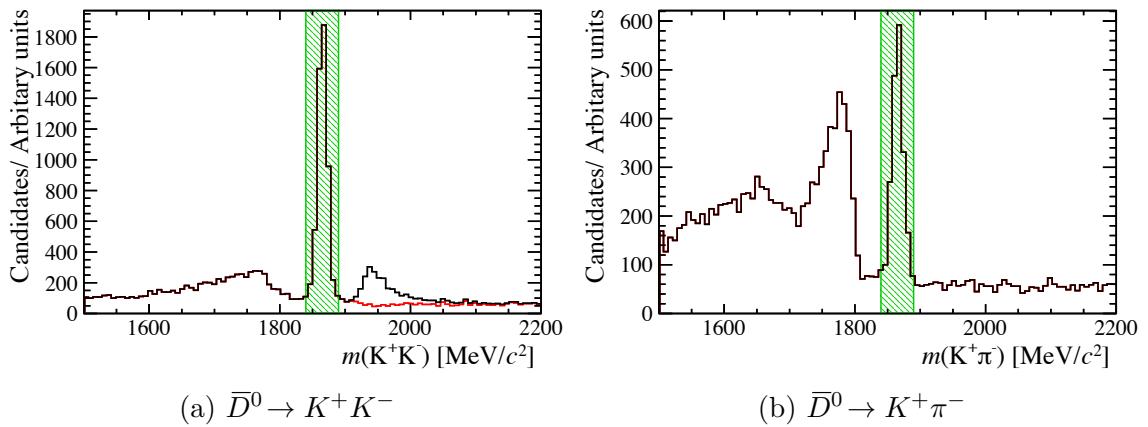


Figure 4.11: Normalisation channel vetoes.

In addition to the correctly reconstructed normalisation channel, the presence of the incorrectly reconstructed $B^+ \rightarrow D_s^+(\bar{D}^0 \rightarrow K^+ \pi^-)$ decay is observable in

Fig. 4.11a. This appears as smeared out peak to the right of the \bar{D}^0 peak. Although the probability of the π^- meson being misidentified as a K^- meson is low, the branching fraction for $\bar{D}^0 \rightarrow K^+ \pi^-$ is larger, leading to the observed excess. It is possible, and necessary, to remove this contribution. The $K^+ K^-$ candidates are reconstructed again, swapping the mass hypothesis of the K^- to π^- . The distribution of these candidates in the vicinity of the \bar{D}^0 mass is shown in Fig. 4.11b. A distinct peak is observed at the \bar{D}^0 mass. This contribution is removed by the requirement $|m(K^+ \pi^-) - m(\bar{D}^0)| > 25 \text{ MeV}/c^2$. The effect of this requirement on the $m(K^+ K^-)$ spectrum is shown by the red line in Fig. 4.11a, which represents the sample with this veto applied. This structure is effectively removed.

Another structure appears to be present to the left of the \bar{D}^0 peak in Fig. 4.11a. This is likely to be due to $\bar{D}^0 \rightarrow K^+ K^- X$ decays in which one or more particles have not been reconstructed. This partially reconstructed background would not peak at the B^+ mass in the $m(D_s^+ K^+ K^-)$ spectrum due to the missing particles, therefore no attempt is made to remove this contribution.

4.3.7 Multivariate analysis

Multivariate Analyses (MVAs) are used to help discriminate between genuine D_s^+ and ϕ meson decays and combinations of unrelated tracks. These MVAs are trained using large samples of candidates from other B mesons decays with similar topologies reconstructed in data. This data-driven approach can benefit from an expanded set of variables that are not perfectly represented in simulation, including track quality and particle identification information, in addition to the widely used kinematic and geometric properties. The method is based on the approach used in Ref. [28], however the choice of input variables has been optimised and training samples expanded to include Run II data.

The sample of D_s^+ mesons is obtained from the relatively abundant $\bar{B}_s^0 \rightarrow D_s^+ \pi^-$ decay. Similarly, the sample of ϕ mesons is obtained from $B_s^0 \rightarrow J/\psi \phi$ decays. Large, high purity samples are reconstructed using similar requirements to those applied in the selection of signal D_s^+ and ϕ mesons. A sample is selected for each of the D_s^+ and

ϕ meson decays uses in this analysis, as listed in Table 4.12. The training of separate MVAs for the different D_s^+ modes allows the use of particle identification variables that separate kaons and pions to be fully exploited. The MVA for $\phi \rightarrow K^+K^-$ decays is used to select both ϕ and \bar{D}^0 mesons. For the normalisation mode this may be suboptimal, however it ensures the selection of the signal and normalisation channels are almost identical such that the systematic uncertainty in the ratio of efficiencies is minimised. The samples of $\bar{B}_s^0 \rightarrow D_s^+ \pi^-$ and $B_s^0 \rightarrow J/\psi \phi$ decays are randomly split into two subsamples. The first is used to train the MVAs and the second is used to determine the efficiency of the selection. To prevent the difference between ϕ and \bar{D}^0 decays from affecting the ratio of efficiencies, the normalisation channel MVA efficiency is instead determined from a dedicated sample of $\bar{D}^0 \rightarrow K^+K^-$ decays as detailed in Table 4.12. The $\phi \rightarrow K^+K^-$ MVA is also used to select the K^+K^- pair in $B^+ \rightarrow D_s^+ K^+ K^-$ decays. A total of eight MVAs are trained, one for each of the D_s^+ and ϕ modes, in both Run I and Run II. Changes to the particle identification variables between the two running periods necessitates separate trainings.

Sample	Mode	Use
$\bar{B}_s^0 \rightarrow D_s^+ \pi^-$	$D_s^+ \rightarrow K^+ K^- \pi^+$	Training, Efficiency
$\bar{B}_s^0 \rightarrow D_s^+ \pi^-$	$D_s^+ \rightarrow K^+ \pi^- \pi^+$	Training, Efficiency
$\bar{B}_s^0 \rightarrow D_s^+ \pi^-$	$D_s^+ \rightarrow \pi^+ \pi^- \pi^+$	Training, Efficiency
$B_s^0 \rightarrow J/\psi \phi$	$\phi \rightarrow K^+ K^-$	Training, Efficiency
$B^+ \rightarrow \bar{D}^0 \pi^+$	$\bar{D}^0 \rightarrow K^+ K^-$	Efficiency

Table 4.12: The decay modes used to train and determine the efficiency of the various MVAs used in this analysis.

Preselection

Before the samples of D_s^+ and ϕ mesons are used to train the MVAs, precautions are taken to ensure the samples are representative of the $B^+ \rightarrow D_s^+ \phi$ and $B^+ \rightarrow D_s^+ K^+ K^-$ signal decays. The $\bar{B}_s^0 \rightarrow D_s^+ \pi^-$ and $B_s^0 \rightarrow J/\psi \phi$ decays are selected using a similar procedure to the signal modes. Firstly, *Stripping Lines* reconstruct the candidates. The $\bar{B}_s^0 \rightarrow D_s^+ \pi^-$ decays are built using the same software module as the signal and

normalisation channel. Therefore, the selection requirements for D_s^+ candidates in $\bar{B}_s^0 \rightarrow D_s^+ \pi^-$ decays are identical to those listed for the signal in Tables 4.6 and 4.7. The $B_s^0 \rightarrow J/\psi \phi$ decays are reconstructed using $J/\psi \rightarrow \mu^+ \mu^-$ decays, therefore they are built using a different software module as the final state is not fully hadronic. As such there are some differences in the selection requirements for the ϕ mesons from the two sources. In general the selection requirements for the $B_s^0 \rightarrow J/\psi \phi$ channel are looser as the presence of the muons allows an efficient triggering without the need to make tight selections on the $K^+ K^-$ pair. A direct comparison of the relevant quantities are listed in Table 4.13.

Particle	Quantity	Signal	Control
ϕ	Mass minimum	$870 \text{ MeV}/c^2$	$980 \text{ MeV}/c^2$
	Mass maximum	$1170 \text{ MeV}/c^2$	$1050 \text{ MeV}/c^2$
	Transverse Momentum	-	$p_T > 500 \text{ MeV}/c$
	Products p_T scalar sum	$\sum p_T > 1000 \text{ MeV}/c$	-
	DOCA(K^+, K^-)	$< 0.5 \text{ mm}$	-
	Direction angle	$\cos \theta > 0$	-
	Vertex quality	$\chi^2/N_{\text{DOF}} < 16$	$\chi^2/N_{\text{DOF}} < 25$
	Flight distance significance	$\chi^2_{\text{FD}} > 16$	-
K^\pm	Track quality	$\chi^2/N_{\text{DOF}} < 4.0$	$\chi^2/N_{\text{DOF}} < 5.0$
	Transverse momentum	$p_T > 100 \text{ MeV}/c$	-
	Momentum	$p > 1000 \text{ MeV}/c$	-
	Impact parameter significance	$\chi^2_{\text{IP}} > 4$	-
	Ghost track probability	$P_{\text{Ghost}} < 0.4$	-
	Particle identification	$\text{PIDK} > -10$	$\text{PIDK} > 0$

Table 4.13: The *Stripping Line* requirements for $\phi \rightarrow K^+ K^-$ candidates in the signal and MVA training (control) mode selection. All requirements are looser for the control channel with the exception of the particle identification requirements.

The difference in the two selections could be potentially biasing when calculating the efficiency of the MVA, therefore the requirements are tightened on the half of the $B_s^0 \rightarrow J/\psi \phi$ sample used to calculate the efficiency. The looser set of requirements are still used when training the MVA methods to maximise the sample sizes to help prevent overtraining. As detailed in Section 4.3.7, when training MVA the ϕ meson invariant mass sidebands in the $B_s^0 \rightarrow J/\psi \phi$ sample are used as the background

sample. It was found that using the tighter set of requirements led to a negligible amount of background candidates, resulting in overtraining.

The second step of preselection aims to apply the same sequence of requirements to the modes used to train the MVAs as are applied to the signal and normalisation before the MVAs are applied. These are made up of the same trigger, background veto and PID requirements. For D_s^+ candidates from $\bar{B}_s^0 \rightarrow D_s^+ \pi^-$ decays, the \bar{B}_s^0 meson is required to either be **TOS** with respect to the **L0Hadron** trigger or **TIS** with respect to **L0Global** as with the signal. The \bar{B}_s^0 candidates are then required to be **TOS** with respect to the same HLT1 and HLT2 triggers as used with the signal modes.

For the $B_s^0 \rightarrow J/\psi \phi$ candidates, a large fraction are triggered due to the muons from the J/ψ decay. To ensure the sample is reflective of the signal decays the B_s^0 meson is required to be **TOS** with respect to the **L0Hadron** trigger or **TIS** with respect to **L0Global**. This effectively excludes events that have been selected solely due to the muons firing **L0Muon**. This consequently results in a large decrease in the available statistics.

As shown in Table 4.13, the MVA training mode has tighter particle identification requirements than the signal mode. Particle identification plays a crucial role in the efficacy of the MVA methods, therefore the signal PID requirements for the ϕ meson decay products are tightened accordingly as discussed in Section 4.3.2.

Finally, the same misidentified D and A_c^+ hadron vetoes as detailed in Section 4.3.4 are applied to the $\bar{B}_s^0 \rightarrow D_s^+ \pi^-$ sample. These help to ensure the samples of D_s^+ candidates are free from contamination.

Background subtraction

In order to use the $B_s^0 \rightarrow J/\psi \phi$ and $\bar{B}_s^0 \rightarrow D_s^+ \pi^-$ samples to train MVAs, background subtracted distributions must be obtained for the variables used to discriminate the signal from backgrounds. Firstly, unbinned extended maximum likelihood fits are performed to the D_s^+ and ϕ invariant mass distribution in order to determine the yield of D_s^+ and ϕ candidates respectively. Maximum likelihood fits determine the values of the signal and background yields for which the given data set was most likely.

The likelihood, \mathcal{L} , is constructed from the probability density functions (PDFs) for the fit model, $F(m, \vec{p})$, for each entry i in the data set. The PDF for each entry is evaluated at the corresponding value of the invariant mass, $m = m_i$, such that \mathcal{L} is just a function of the PDF parameters, \vec{p} ,

$$\mathcal{L}(\vec{p}) = \prod_i^N F(m = m_i, \vec{p}). \quad (4.2)$$

The maximum value of $\mathcal{L}(\vec{p})$ is achieved for the set of PDF parameters for which the data was most likely. It is computationally beneficial to instead compute the negative log-likelihood (NLL) rather than the likelihood directly, as addition is less intensive than multiplication. The NLL,

$$-\log \mathcal{L}(\vec{p}) = -\sum_i^N \log F(m = m_i, \vec{p}), \quad (4.3)$$

is then minimised with respect to the PDF parameters.

This likelihood can be *extended* to allow the total yield attributed to the PDF to be determined as a parameter as well. This is necessary to determine the yields of both signal and background contributions (n_S and n_B). The likelihood is multiplied by the Poisson probability density for measuring $n = n_S + n_B$ events,

$$\mathcal{L}(n, \vec{p}) = \frac{n^N e^{-n}}{N!} \prod_i^N F(m = m_i, \vec{p}) = \frac{e^{-n}}{N!} \prod_i^N nF(m = m_i, \vec{p}). \quad (4.4)$$

The extended NLL becomes

$$-\log \mathcal{L}(n, \vec{p}) = -\sum_i^N \log nF(m = m_i, \vec{p}) + n + \log N!. \quad (4.5)$$

The $\log N!$ term can be ignored as it is a constant. The fit model is constructed from signal and background contributions

$$nF(m|n_S, n_B, \vec{p}', \vec{p}'') = n_S f(m|\vec{p}') + n_B g(m|\vec{p}''), \quad (4.6)$$

where n_S and n_B are the signal and background yields, f and g are the PDFs for the signal and background respectively, and \vec{p}' and \vec{p}'' represent the parameters controlling the PDF shapes.

After the extended maximum likelihood fits have been performed and the yields established, the subtraction of candidates attributed to the background component is performed using the *sPlot* technique [29]. The *sPlot* method assigns weights for each species in the fit model, namely signal and background. These are assigned to each event in the data set and are only dependent on the value of the fitted variable, in this case ϕ or D_s^+ mass. When plotting other, uncorrelated variables and weighting each event accordingly, the contribution from each component in the fit can be extracted.

For both $B_s^0 \rightarrow J/\psi \phi$ and $\bar{B}_s^0 \rightarrow D_s^+ \pi^-$ decays, simple fit models are found to be sufficient when performing the *sPlot* technique. For each mode the signal components are modelled with a single Gaussian probability density function

$$f(m|\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \times e^{-\frac{(m-\mu)^2}{2\sigma^2}}, \quad (4.7)$$

where μ and σ are the mean and width of the Gaussian distribution and m is the observable invariant mass. The parameters μ and σ are allowed to vary freely. The background contributions to the $\phi \rightarrow K^+ K^-$ decays are modelled using a Chebychev polynomial with two degrees of freedom

$$g(m|a, b) = a \times (2m^2 - 1) + b \times m + 1, \quad (4.8)$$

where a and b are parameters that can vary freely and m is the observable invariant mass. The ϕ meson mass is close to the threshold for $K^+ K^-$ pair production so this parametrisation has sufficient freedom to successfully model the increasing background shape. The background contribution to all three D_s^+ decays are modelled using exponential functions with a single degree of freedom controlling the effective slope

$$g(m|c) = e^{-m \times c}, \quad (4.9)$$

where c is a freely varying parameter and m is the observable invariant mass. The fits to the ϕ and D_s^+ meson masses in the $B_s^0 \rightarrow J/\psi \phi$ and $\bar{B}_s^0 \rightarrow D_s^+ \pi^-$ samples are performed separately for each year of data taking and polarity of the LHCb dipole magnet. These distributions are shown in Fig. 4.12 for each of the ϕ and D_s^+ decay

modes. The yields of candidates in each of the subsamples are tabulated in Table 4.14, along with the totals for each mode.

Mode	Year	MagDown	MagUp
$\phi \rightarrow K^+ K^-$	2011	3190 ± 170	2120 ± 140
	2012	5800 ± 200	5540 ± 200
	2015	2240 ± 110	1640 ± 100
	2016	11200 ± 300	11100 ± 300
	Total	42830 ± 600	
$D_s^+ \rightarrow K^+ K^- \pi^+$	2011	97100 ± 500	68600 ± 400
	2012	216600 ± 800	212800 ± 800
	2015	68600 ± 400	43400 ± 300
	2016	321400 ± 900	310800 ± 900
	Total	1339300 ± 1900	
$D_s^+ \rightarrow \pi^+ \pi^- \pi^+$	2011	24300 ± 500	17100 ± 400
	2012	57700 ± 800	56500 ± 900
	2015	20400 ± 500	13100 ± 400
	2016	99200 ± 1200	92100 ± 1200
	Total	380400 ± 2300	
$D_s^+ \rightarrow K^+ \pi^- \pi^+$	2011	13200 ± 600	9500 ± 400
	2012	29300 ± 900	29300 ± 900
	2015	10000 ± 500	6200 ± 400
	2016	47800 ± 1200	44200 ± 1100
	Total	189500 ± 2300	

Table 4.14: Yields of the samples used to train data-driven MVAs. The total yields are summed over all years and both magnet polarities.

Input variables

The MVA method is trained using a large set of variables chosen to help discriminate between the signals of interest and combinations of unrelated tracks. These variables include the properties of the K^\pm or π^\pm decay products as well as those of the ϕ or D_s^+ candidate itself. Many of the quantities are the same as those previously defined in Section 4.3.1. The additional parameters relate to the track quality and particle identification information and are defined as follows:

Track matching quality, χ^2_{TRKMATCH} : this parameter is a key component of the pattern recognition software that matches the VELO and tracking station

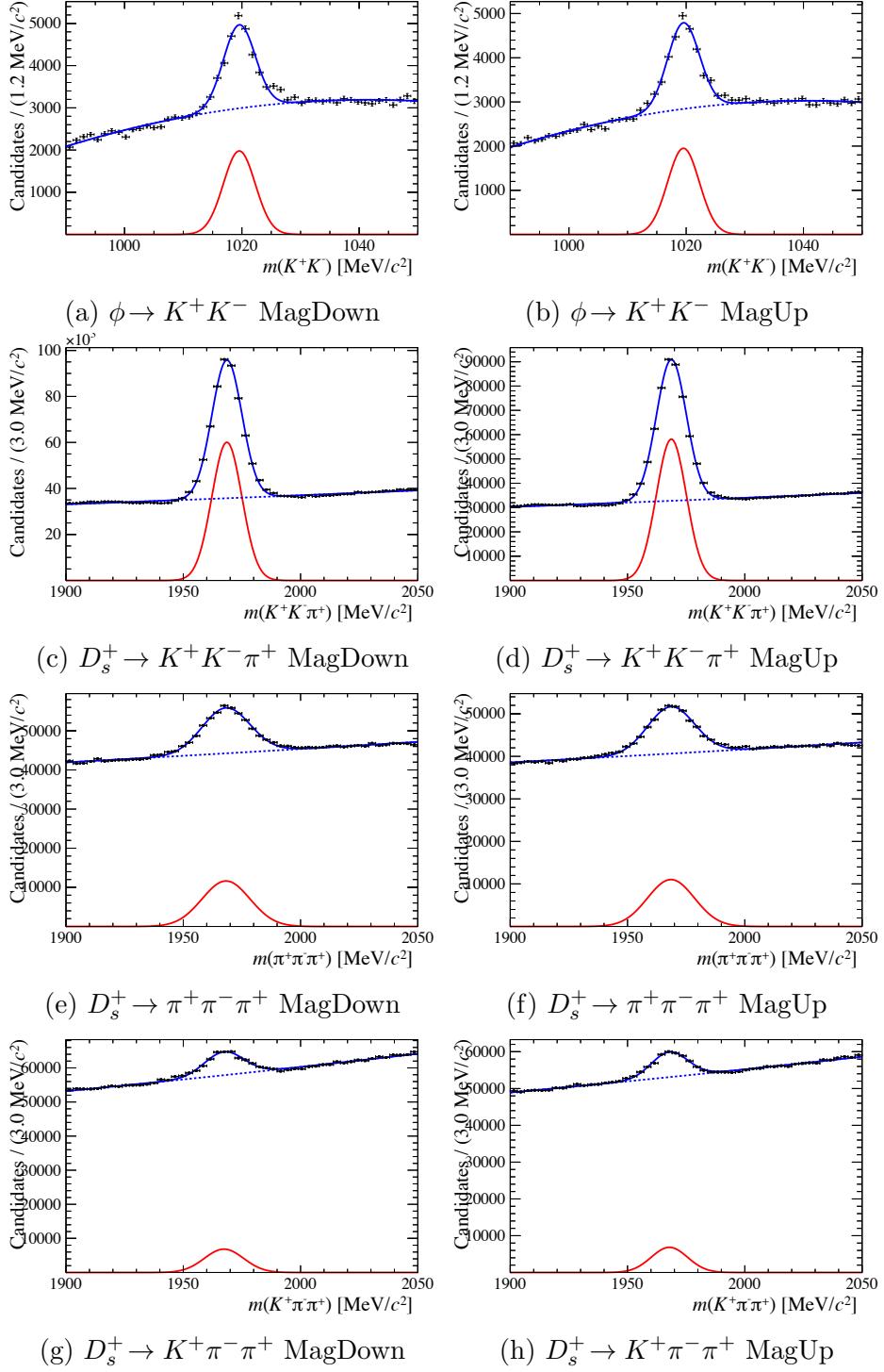


Figure 4.12: Unbinned maximum likelihood fit to $B_s^0 \rightarrow J/\psi \phi$ (top) and $\bar{B}_s^0 \rightarrow D_s^+ \pi^-$ (bottom three rows) candidates. These distributions correspond to just the 2016 data samples.

track sections together to form *long* tracks. It quantifies the quality of the track matching.

VELO track quality, χ^2_{TRKVELO} : this quantifies the quality of the track fit using just the hits in the VELO.

Kaon, proton, pion and ghost probabilities, ProbNNx: these variables are an alternative form of *Particle Identification* variables in addition to those already discussed in Section 4.3.2. Each ProbNNx variable acts as the probability for the reconstructed track to be of the species x. These variables are the responses of Artificial Neural Networks trained to differentiate between the various particle types using a large number of inputs from different sub-detectors. The trainings use simulated samples of signal and background decays and have been retuned between Run I and Run II to reflect the difference in the detector configuration. Four of the available ProbNNx tuning are used in this analysis. Three of these identify kaons, protons and pions. The last aims to identify ghost tracks, similar to the P_{Ghost} variable already discussed.

The variables used in the ϕ meson MVA are listed in Table 4.15. Some of the variables are input into the MVA training algorithm as their logarithm, rather than the variable directly. These variables tend to be rapidly increasing in at certain values, so using the logarithm aids visualisation. A total of 24 variables are used in this MVA.

When training the ϕ meson MVA, the flight distance significance is not included as an input variable. This MVA is used to select both ϕ and \bar{D}^0 mesons which are likely to have different flight distance significance distributions as a result of their different lifetimes.

The variables used in the D_s^+ MVA are listed in Table 4.16. These are largely very similar to those used for the ϕ meson MVA, however it also includes the D_s^+ meson's flight distance significance as, unlike the ϕ meson, the D_s^+ meson decays displaced from its production vertex. A total of 35 variables are used in these MVAs.

Particle	Description	Quantity
ϕ	Momentum	p
	Transverse Momentum	p_T
	Vertex quality	$\log_{10}(\chi^2_{\text{VXT}})$
	Impact parameter significance	$\log_{10}(\chi^2_{\text{IP}})$
K^\pm	Momentum	p
	Transverse momentum	p_T
	Longitudinal momentum	p_Z
	Impact parameter significance	$\log_{10}(\chi^2_{\text{IP}})$
	Track quality	χ^2_{TRK}
	Track matching quality	χ^2_{TRKMATCH}
	Kaon probability	ProbNNk
	Proton probability	ProbNNp
	Pion probability	ProbNNpi
	Ghost probability	ProbNNghost

Table 4.15: Discriminating variables used to train the ϕ MVA.

When training the MVA methods, the variables are ranked according to their discrimination power. Examples of these ranking are shown for the Run II $\phi \rightarrow K^+ K^-$ MVA and Run II $D_s^+ \rightarrow K^+ K^- \pi^+$ MVA in Table 4.17. It can be seen that for both the D_s^+ and ϕ meson MVAs that the kaon particle identification variable and various $\log_{10}(\chi^2_{\text{IP}})$ variables are the most discriminating.

MVA method

The MVA training and validation is implemented using the TMVA package, part of the ROOT framework [30]. During training, each method uses samples that have been defined to be ‘signal’ and ‘background’ in order to characterise the differences between the two categories. The multidimensional space defined by the input variables is effectively condensed onto a single axis given by the classifier’s response. Typically, high values of the response correspond to ‘signal’-like events and low values to ‘background’-like events. This training produces a *weights file* allowing the response to be calculated for other samples.

A second set of ‘signal’-like and ‘background’-like samples are used to blindly validate the training. The classifier response is compared in the training and validation

Particle	Description	Quantity
D_s^+	Momentum	p
	Transverse Momentum	p_T
	Vertex quality	$\log_{10}(\chi_{\text{VXT}}^2)$
	Impact parameter significance	$\log_{10}(\chi_{\text{IP}}^2)$
	Flight distance significance	$\log_{10}(\chi_{\text{FD}}^2)$
K^\pm, π^\pm	Momentum	p
	Transverse momentum	p_T
	Impact parameter significance	$\log_{10}(\chi_{\text{IP}}^2)$
	Track quality	χ_{TRK}^2
	Velo track quality	χ_{TRKVELO}^2
	Track matching quality	χ_{TRKMATCH}^2
	Kaon probability	ProbNNk
	Proton probability	ProbNNp
	Pion probability	ProbNNpi
	Ghost probability	ProbNNghost

Table 4.16: Discriminating variables used to train the D_s^+ MVAs.

samples to demonstrate if the expected separation is reproducible. This is important to determine if the trained method suffers from *overtraining*. This can happen when too few events are passed to the method, such that the statistical fluctuations in the distributions are significant. This can result in differences being identified that are actually artefacts of the low statistics and not representative of the sample as a whole. This results in the classifier being more discriminatory on the training sample than the validation sample.

The ‘signal’ and ‘background’ samples are both taken from the relevant $\bar{B}_s^0 \rightarrow D_s^+ \pi^-$ or $B_s^0 \rightarrow J/\psi \phi$ decay samples. The signal samples include all events within the fit ranges shown in Fig. 4.12. Each candidate is weighted with the appropriate Gaussian fit component weight as determined by the *sPlot* technique. This weight is passed to the TMVA framework when training the methods. The background samples are comprised of the sideband regions of the relevant ranges in Fig. 4.12. These ranges are defined by $|m(K^+ K^-) - m(\phi)| > 10 \text{ MeV}/c^2$ and $|m(h^+ h^- h^+) - m(D_s^+)| > 30 \text{ MeV}/c^2$ for the ϕ and D_s^+ MVAs respectively. No weights are used for the background samples as the regions are known to be signal-deficient.

Rank	Variable	Importance	Rank	Variable	Importance
1	K^+ ProbNNk	7.639×10^{-2}	1	K^+ ProbNNk	6.001×10^{-2}
2	K^- ProbNNk	6.043×10^{-2}	2	$\phi \log_{10}(\chi_{\text{IP}}^2)$	6.001×10^{-2}
3	$K^+ \log_{10}(\chi_{\text{IP}}^2)$	5.469×10^{-2}	3	$K^- \log_{10}(\chi_{\text{IP}}^2)$	5.997×10^{-2}
4	$\pi^+ \log_{10}(\chi_{\text{IP}}^2)$	4.984×10^{-2}	4	$\phi \log_{10}(\chi_{\text{VTX}}^2)$	5.827×10^{-2}
5	$K^+ p$	4.348×10^{-2}	5	$K^- \text{ProbNNpi}$	5.826×10^{-2}
6	$D_s^+ \log_{10}(\chi_{\text{IP}}^2)$	4.100×10^{-2}	6	$K^+ \log_{10}(\chi_{\text{IP}}^2)$	5.761×10^{-2}
7	$\pi^+ p_{\text{T}}$	3.953×10^{-2}	7	$K^+ \text{ProbNNghost}$	5.427×10^{-2}
8	$K^+ p_{\text{T}}$	3.937×10^{-2}	8	$K^- \text{ProbNNk}$	5.404×10^{-2}
9	$\pi^+ \text{ProbNNpi}$	3.881×10^{-2}	9	$K^+ \chi_{\text{TRK}}^2$	5.359×10^{-2}
10	$D_s^+ \log_{10}(\chi_{\text{FD}}^2)$	3.788×10^{-2}	10	$K^- \chi_{\text{TRK}}^2$	5.201×10^{-2}
11	$K^- \log_{10}(\chi_{\text{IP}}^2)$	3.735×10^{-2}	11	$K^+ \text{ProbNNpi}$	5.162×10^{-2}
12	$D_s^+ \log_{10}(\chi_{\text{VTX}}^2)$	3.016×10^{-2}	12	$K^- \text{ProbNNghost}$	4.985×10^{-2}
13	$K^- p$	2.770×10^{-2}	13	$K^+ \text{ProbNNp}$	4.486×10^{-2}
14	$K^+ \text{ProbNNp}$	2.758×10^{-2}	14	$K^- \text{ProbNNp}$	4.410×10^{-2}
15	$K^- \text{ProbNNp}$	2.692×10^{-2}	15	$K^+ \chi_{\text{TRKMATCH}}^2$	4.291×10^{-2}

Table 4.17: Ranking of variables for the 15 highest ranked variables used to train the $D_s^+ \rightarrow K^+ K^- \pi^+$ (left) and $\phi \rightarrow K^+ K^-$ (right) Run II MVAs.

The MVAs are trained using Gradient Boosted Decision Trees (BDTGs) [31]. This method was compared to a number of different Boosted Decision Tree (BDT) derivatives and found to be most effective.

The BDTG method categorises ‘signal’ and ‘background’ candidates by creating consecutive sets of questions called Decision Trees. Each question can have two possible answers and depends on the previous responses. Eventually the answer to a question results in a candidate being classes as ‘signal’ or ‘background’ rather than asking more questions. A definable number of decision trees are used to collectively categorise each candidate; the overall response combines the decisions of all trees. The trees are trained using *boosting*. This means candidates that are most likely to be misclassified are weighted higher, such that the later trees concentrate on these trickier areas. The boosting method utilised in the BDTG method is designed to be more robust and reproducible in noisy situations, for example with limited statistics. The BDTG method is used in training all MVAs.

The MVA classifier distributions for each of the eight MVAs trained are shown in Fig. 4.15. This figure shows both the signal and background samples, as well as the

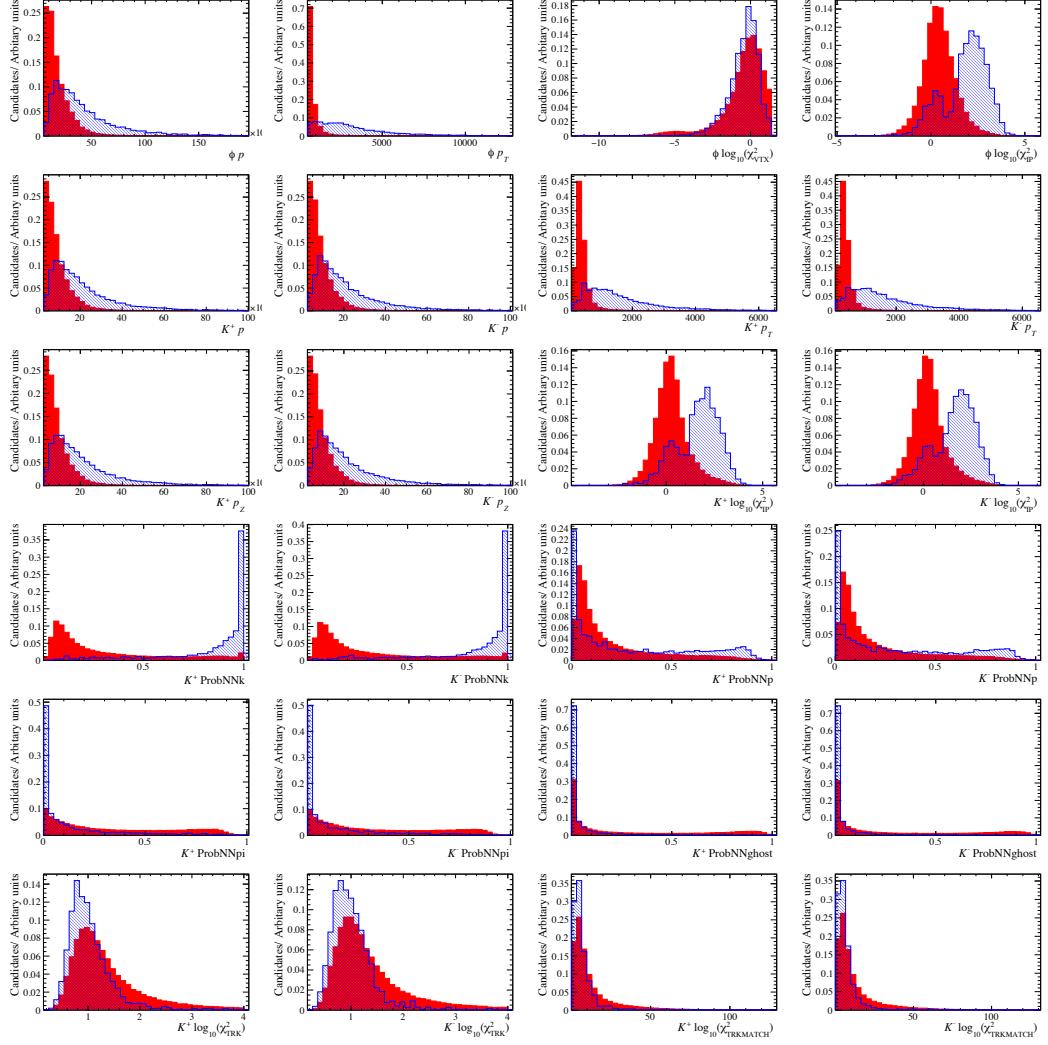


Figure 4.13: Distributions of MVA training variables in the signal (blue) and background (red) samples for the Run II $\phi \rightarrow K^+ K^-$ MVA.

training and validation subsamples. Each mode shows a good separation between the signal and background samples. For the majority of the distributions the training and testing samples have almost identical shapes, implying response is reproducible and the method has not overtrained. There are noticeable exceptions, however. The signal training and validation samples tend to show different distributions in the regions where the background distributions are maximal. This effect is likely to be a result of the use of weights in the MVA training. Ideally, the MVA classifier should remain positive across the whole range. The negative values implies that in that range the MVA classifier is correlated to the parameter used to generate the weights:

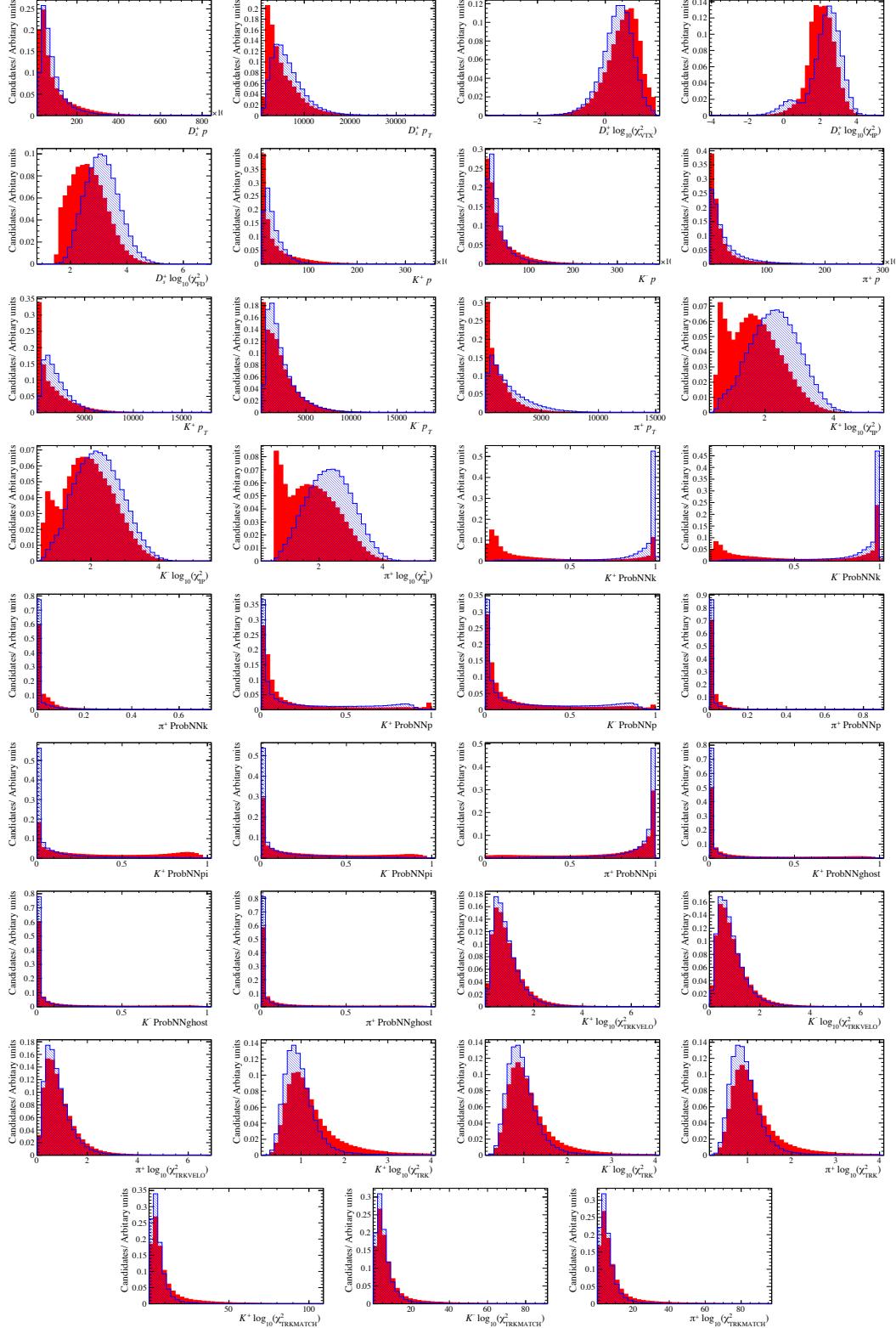


Figure 4.14: Distributions of MVA training variables in the signal (blue) and background (red) samples for the Run II $D_s^+ \rightarrow K^+ K^- \pi^+$ MVA.

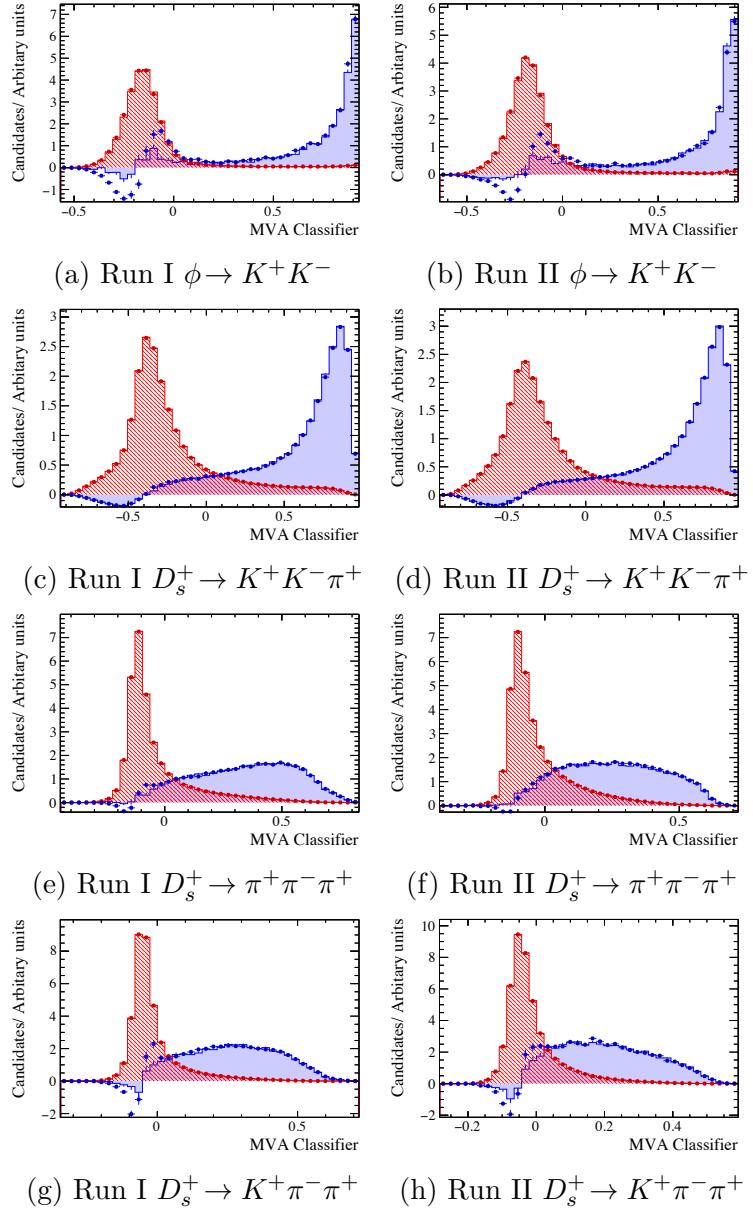


Figure 4.15: MVA classifier response for the eight MVAs trained for both signal (blue) and background (red). The training samples are represented by markers and the validation samples by filled histograms.

the ϕ or D_s^+ mass. As this only affects the range of MVA classifier values that are dominantly ‘background’-like, it likely this discrepancy is due specifically to presence of weighted background events in the signal sample passed to TMVA. This effect is further studied to determine if it results in any systematic uncertainty in the MVA efficiencies as discussed in Sections 5.7 and 6.7. In all cases the discrepancies are far

below the range of values that selection requirements are placed.

MVA efficiency

The efficiencies of the MVAs on D_s^+ and ϕ meson in the signal decays can be obtained from the validation samples of $B_s^0 \rightarrow J/\psi\phi$ and $\bar{B}_s^0 \rightarrow D_s^+\pi^-$ decays. Additionally, a sample of $B^+ \rightarrow \bar{D}^0\pi^+$ decays is used to calculate the efficiency of $\bar{D}^0 \rightarrow K^+K^-$ decays in the normalisation channel. The efficiency calculation takes into account the kinematic differences between the training and signal samples, as well as any possible correlations between the D_s^+ and ϕ kinematics, by using input from simulation samples. The signal MVA response is extracted from the validation samples in four bins of both p_T and χ_{FD}^2 . Simulation samples for $B^+ \rightarrow D_s^+\phi$ or $B^+ \rightarrow D_s^+\bar{D}^0$ decays are iterated through in turn, selecting the corresponding MVA response from the appropriate kinematic bins. The per-candidate efficiency is determined by integrating the validation sample D_s^+ and ϕ MVA responses above the required cut values

$$\varepsilon_{\text{event}} = \int_{\text{cut}_{\phi}}^1 f(x)dx \times \int_{\text{cut}_{D_s^+}}^1 g(y)dy \quad (4.10)$$

where $f(x)$ and $g(y)$ represent the weighted MVA classifier output for the ϕ and D_s^+ MVAs respectively, and cut_{ϕ} and $\text{cut}_{D_s^+}$ represent the chosen MVA cut value. The total efficiency for each mode is given by the sum of the per-candidate efficiencies within the relevant simulation sample.

The $\phi \rightarrow K^+K^-$ MVA is also used to select the K^+K^- pair in $B^+ \rightarrow D_s^+K^+K^-$ decays. Due to the similar topologies of $B^+ \rightarrow D_s^+K^+K^-$ and $B^+ \rightarrow D_s^+\phi$ decays, the K^+K^- pair share lots of similarities to the $\phi \rightarrow K^+K^-$ counterparts. One obvious difference is the invariant mass of the pair, as the $B^+ \rightarrow D_s^+K^+K^-$ decays includes a much larger phase-space. In training the MVA this could lead to the selection being non-optimal for $B^+ \rightarrow D_s^+K^+K^-$ decays at high $m(K^+K^-)$. However, when determining the efficiency of the selection using these same $B_s^0 \rightarrow J/\psi\phi$ decays these differences could lead to bias. Instead, calibration samples are used to correct for the imperfect modelling of the particle identification in the $B^+ \rightarrow D_s^+K^+K^-$

simulation samples. These corrected simulations are then used to obtain the variations in the MVA efficiencies as a function of the phase-space position, in particular of the $m(K^+K^-)$ invariant mass as discussed in Section 5.6.

MVA optimisation

The selection criteria for each of the BDTG classifiers are determined by optimising the Punzi figure of merit [32],

$$\frac{\epsilon_s}{\left(\frac{a}{2} + \sqrt{N_{\text{BKG}}}\right)} \quad (4.11)$$

with $a = 5$, where ϵ_s is the signal efficiency and N_{BKG} is the number of background candidates determined from fits to data, calculated in the signal region. The signal efficiency is calculated for the relevant mode and MVA requirement values using the data-driven technique previously described. The optimisation is performed in two dimensions, one for each of the ϕ and D_s^+ MVAs. This is performed separately for the three different D_s^+ decay modes. The results of the optimisation for each D_s^+ decay mode and running period are shown in Fig. 4.16.

The optimal values are selected separately for each of the modes and tabulated in Table 4.18.

Period	D_s^+ decay mode	D_s^+ requirement	ϕ/\bar{D}^0 requirement
Run I	$D_s^+ \rightarrow K^+K^-\pi^+$	0.5	0.8
	$D_s^+ \rightarrow \pi^+\pi^-\pi^+$	0.3	0.8
	$D_s^+ \rightarrow K^+\pi^-\pi^+$	0.2	0.8
Run II	$D_s^+ \rightarrow K^+K^-\pi^+$	0.5	0.8
	$D_s^+ \rightarrow \pi^+\pi^-\pi^+$	0.2	0.8
	$D_s^+ \rightarrow K^+\pi^-\pi^+$	0.2	0.8

Table 4.18: Optimised MVA cuts for the different D_s^+ decay modes and datasets.

4.3.8 Impact parameter requirements

The data-driven MVAs target the selection of ϕ and D_s^+ candidates, removing a significant background contribution. However it is possible to further purify the candidates by making requirements on the B^+ candidates formed from the combination of the

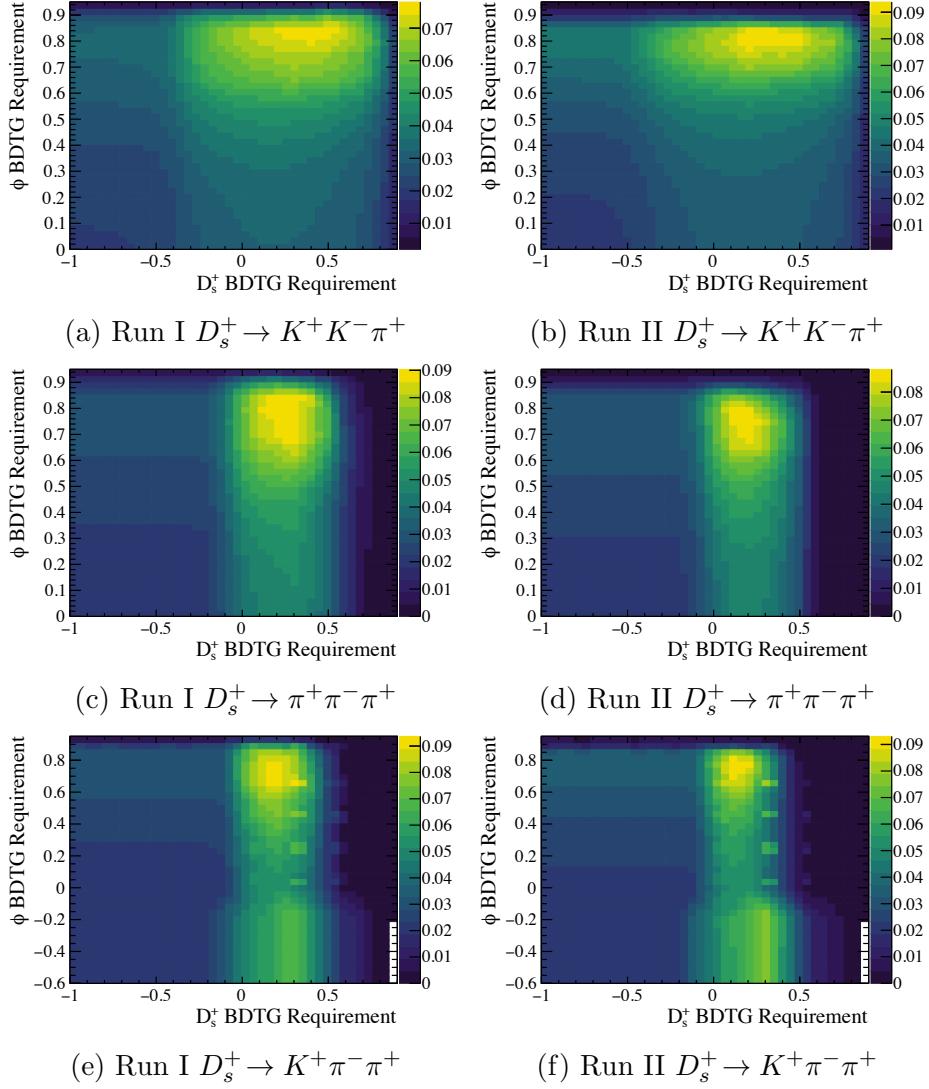


Figure 4.16: MVA optimisation

two. This helps to remove possible backgrounds coming from the incorrect combination of two unrelated mesons. Requirements are placed on the impact parameter significance χ_{IP}^2 to help ensure that the different components of the decay chain have a topology consistent with those desired.

The B^+ meson candidate is required to be consistent with originating at the PV by requiring $\chi_{\text{IP}}^2 < 10$. This helps to remove background as unrelated combinations of D_s^+ and ϕ mesons will not necessarily point back towards the PV. Additionally, the opposite requirement is made of the D_s^+ meson, $\chi_{\text{IP}}^2 > 10$. It is possible for D_s^+ mesons to be produced promptly at the PV rather than in a B^+ meson decay and

this requirement helps to remove those candidates.

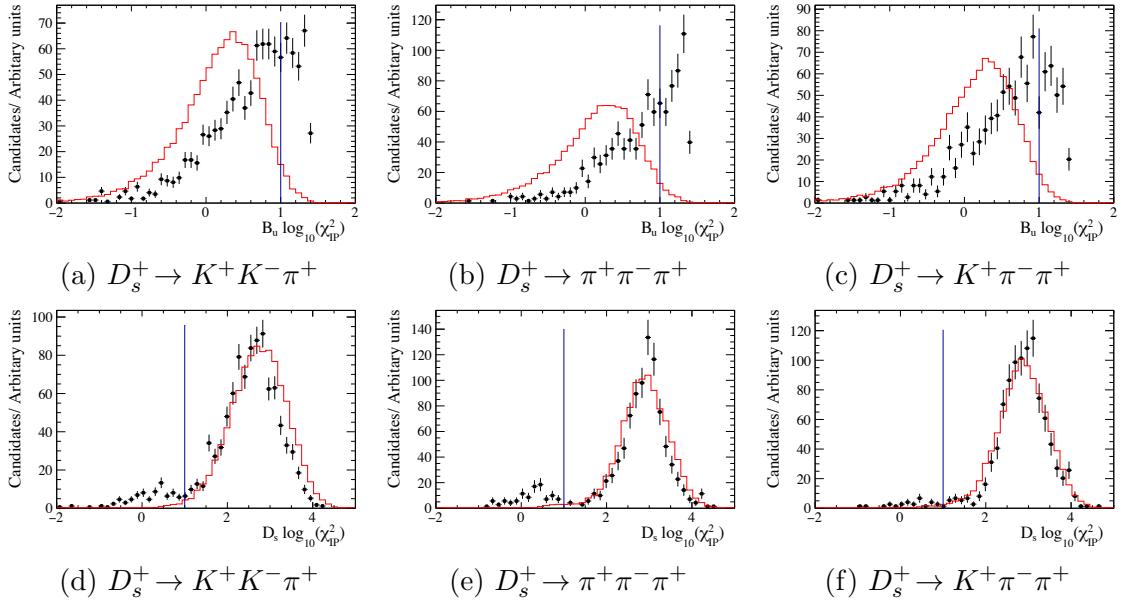


Figure 4.17: The B^+ and D_s^+ meson χ_{IP}^2 distributions for $B^+ \rightarrow D_s^+ \phi$ candidates in data (black) and simulation (red). Candidates within the range $4900 < m(D_s^+ \phi) < 5900 \text{ MeV}/c^2$ are included in these figures.

The distributions for the B^+ and D_s^+ meson impact parameter significance is shown in Fig. 4.17 for $B^+ \rightarrow D_s^+ \phi$ candidates. The requirement values are represented by vertical lines. Small peaks can be observed at low D_s^+ meson impact parameter significance, corresponding to D_s^+ mesons that originated at the PV. As a cross-check, the same distributions are produced for the normalisation channel, $B^+ \rightarrow D_s^+ \bar{D}^0$. However, instead of including all candidates in a wide $m(D_s^+ \phi)$ mass range, only those within $20 \text{ MeV}/c^2$ of the B^+ mass are included. As the normalisation sample has a very high purity this effectively isolates these decays in data. This comparison of data and simulation distributions is shown in Fig. 4.18.

Small shifts are observed in the χ_{IP}^2 distributions. These differences are studied as a source of systematic uncertainty as discussed in Sections 5.7 and 6.7.

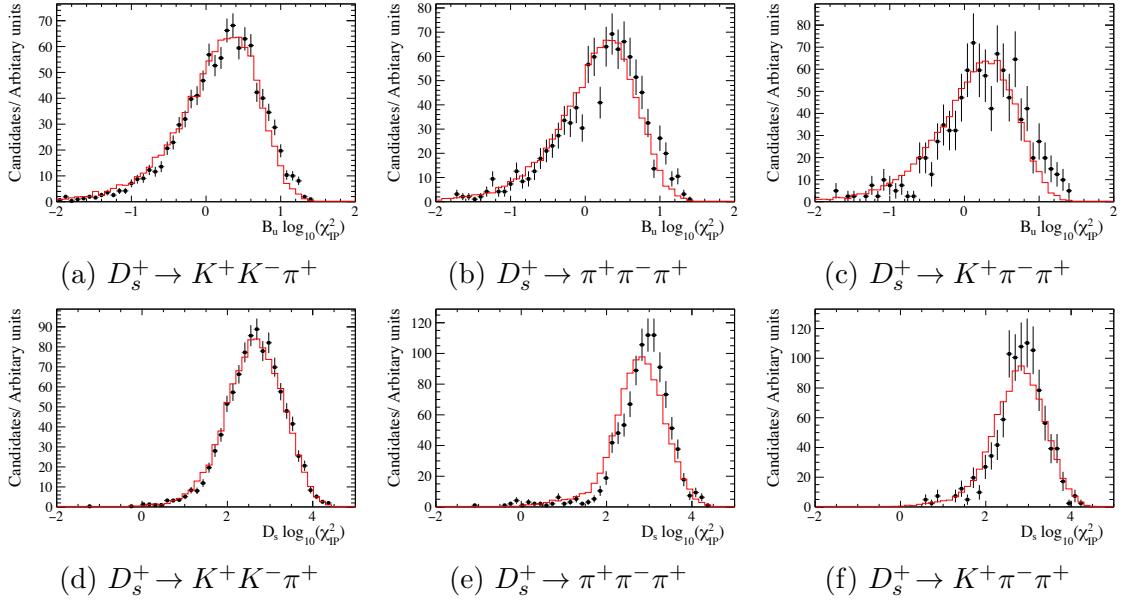


Figure 4.18: The B^+ and D_s^+ meson χ_{IP}^2 distributions for $B^+ \rightarrow D_s^+ \bar{D}^0$ candidates in data (black) and simulation (red). Candidates within the range $|m(D_s^+\phi) - m(B^+)| < 20 \text{ MeV}/c^2$ are included in these figures, effectively isolating the normalisation events in data.

4.4 Multiple candidates

Signal and normalisation candidates are constructed by combining five of the large number of tracks in each event in different combinations. It is possible that after the many steps of selection have been performed, that there are still multiple candidates reconstructed in a single event. It is, of course, possible that there are more than one B meson decays in a single event. Indeed production via $b\bar{b}$ pairs would lead to two B hadrons in the same event. However both the singal and normalisation channels searched for here have relatively small branching fractions, therefore the probability for both B^+ mesons to decay the same final state is low. The rate of multiple candidates is studied for each of the samples after passing through the full selection. The run and event number of each candidate is compared, and the number of events that contain more than one candidate is divided by the total number of candidates to produce a rate. These are tabulated in Table 4.19. The rates of multiple candidates are sufficiently low that no attempt is made to remove them.

B^+ decay mode	D_s^+ decay mode	Multiple candidate rate
$B^+ \rightarrow D_s^+ \phi$	$D_s^+ \rightarrow K^+ K^- \pi^+$	0.33%
	$D_s^+ \rightarrow \pi^+ \pi^- \pi^+$	0.46%
	$D_s^+ \rightarrow K^+ \pi^- \pi^+$	0.00%
$B^+ \rightarrow D_s^+ \bar{D}^0$	$D_s^+ \rightarrow K^+ K^- \pi^+$	0.27%
	$D_s^+ \rightarrow \pi^+ \pi^- \pi^+$	0.11%
	$D_s^+ \rightarrow K^+ \pi^- \pi^+$	0.17%
$B^+ \rightarrow D_s^+ K^+ K^-$	$D_s^+ \rightarrow K^+ K^- \pi^+$	0.25%
$B^+ \rightarrow D_s^+ \bar{D}^0$	$D_s^+ \rightarrow K^+ K^- \pi^+$	0.21%

Table 4.19: Multiple candidate rates for signal and normalisation decays.

4.5 Refitting the decay chain

The yields of B^+ meson candidates for the signal and normalisation channels are determined from fits to the B^+ meson invariant mass, the exact details of which will be in Chapters 5 and 6. The *raw* B^+ meson invariant mass is calculated in the vertex fits used to constructed the B^+ candidates from combinations of D_s^+ and ϕ or \bar{D}^0 candidates. The 4-momentum vectors for the decay products trajectories are extrapolated to the vertex location and the invariant mass calculated from the sum of the decay products 4-vectors. The width of the B^+ meson distribution is dominated by the resolution of the LHCb detector, as the natural width of the meson itself is 0.4×10^{-3} eV.

It is possible to reduce the effect of the resolution by adding more constraints in the fit to the decay chain. The mass of the D_s^+ meson (and \bar{D}^0 meson for the normalisation channel) are constrained to their known values. Additionally, the momentum vector of the B^+ meson is constrained to be parallel to the line between the primary interaction and the B^+ meson decay vertex. This is implemented with the Decay Tree Fitter algorithm [33]. With these constraints applied, the whole decay chain is refitted, and the B^+ meson mass recalculated. For genuine signal and normalisation decays the resolution is improved; the smearing effects that might lead the track momentum or position to be incorrectly measured can be circumvented by including the

extra information about intermediate meson masses, or B^+ flight direction. Background candidates that have different topologies to the constrained decay chain will be randomly shifted up or down in invariant mass, leading to no overall difference in the combinatorial distribution.

It is possible to calculate all kinematic distributions for the components of the decay chain with these constraints applied, for example the momentum of the decay products. However, these variables are not used when training the MVA methods even though they may result in a greater discrimination between signal and background. The modes used to train the MVA methods have slightly different topologies, therefore the improvements in the resolution achieved in these modes may not be the same as that seen for the signal and normalisation channel.

The search for $B^+ \rightarrow D_s^+ K^+ K^-$ decays involves determining the phase-space distribution of candidates in both simulations and data using the two-body masses $m^2(D_s^+ K^-)$ and $m^2(K^+ K^-)$. These are determined using the decay product 4-vectors calculated with the B^+ meson mass constrained in addition to the D_s^+ mass and B^+ flight direction. This ensures the candidates are found within the kinematically allowed region of the two dimensional $m^2(D_s^+ K^-)$ vs. $m^2(K^+ K^-)$ space.

In the search for $B^+ \rightarrow D_s^+ \phi$ decays an angle, θ_K , is used to split candidates into two categories. This angle is calculated using the decay products 4-vectors calculated using the D_s^+ mass and B^+ flight direction constraints. This improves the separation power as opposed to using the *raw* 4-vectors.

4.6 Invariant mass distributions

The D_s^+ and ϕ meson invariant mass distributions for signal $B^+ \rightarrow D_s^+ \phi$ candidates passing the full selection requirements are shown in Fig. 4.19. The same distributions for the normalisation channel $B^+ \rightarrow D_s^+ \bar{D}^0$ are shown in Fig. 4.20. The purity of the $D_s^+ \rightarrow \pi^+ \pi^- \pi^+$ and $D_s^+ \rightarrow K^+ \pi^- \pi^+$ decays are significantly lower than the $D_s^+ \rightarrow K^+ K^- \pi^+$ mode as a result of the smaller branching fractions for these modes.

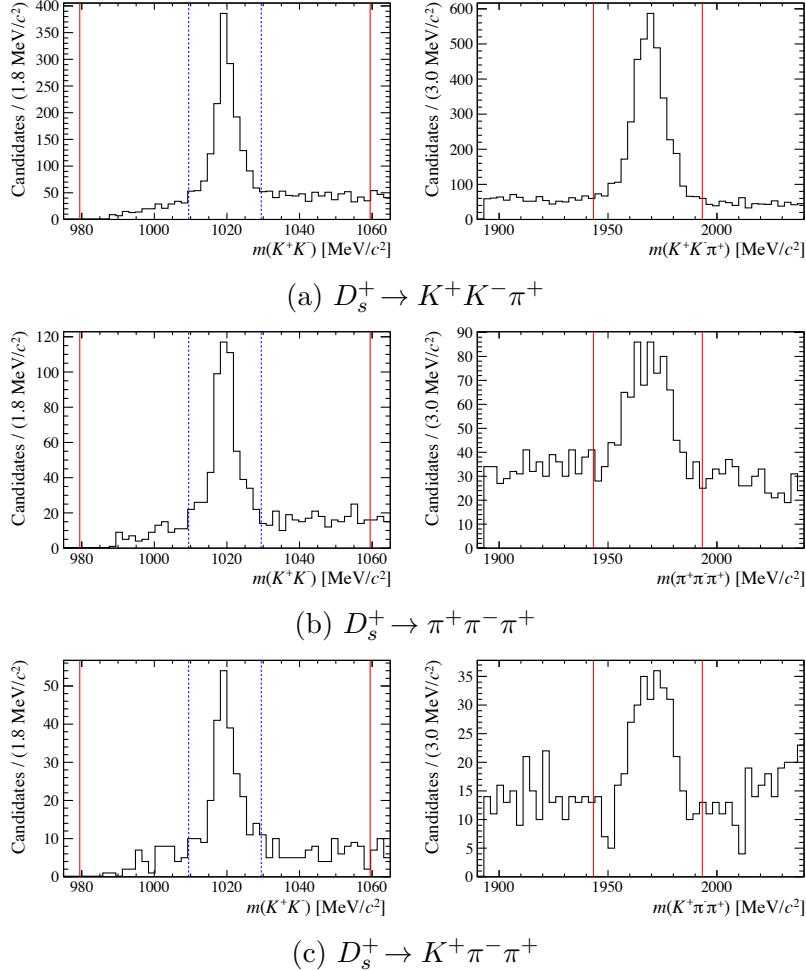


Figure 4.19: The D_s^+ and ϕ invariant mass distributions for signal $B^+ \rightarrow D_s^+ \phi$ candidates passing the selection requirements. The inner and outer $m(K^+K^-)$ ranges, as defined in Chapter 6, are represented by the blue and red vertical lines respectively.

Requirements are placed in the invariant masses of the D_s^+ , ϕ and \bar{D}^0 candidates to isolate the signal decays. The requirements are listed in Table 4.20 and represented in Figs. 4.19 and 4.20 by red vertical lines. The ϕ meson requirement is wider than those for the D_s^+ and \bar{D}^0 mesons even though the width is clearly narrower. The $B^+ \rightarrow D_s^+ \phi$ candidates are further split into two categories as detailed in Chapter 6, therefore the sidebands either side of the ϕ meson are retained.

The distribution of D_s^+ and K^+K^- invariant masses for $B^+ \rightarrow D_s^+ K^+ K^-$ and the corresponding $B^+ \rightarrow D_s^+ \bar{D}^0$ candidates are shown in Fig. 4.21.

The B^+ invariant mass distribution for the normalisation mode is shown in Fig. 4.22.

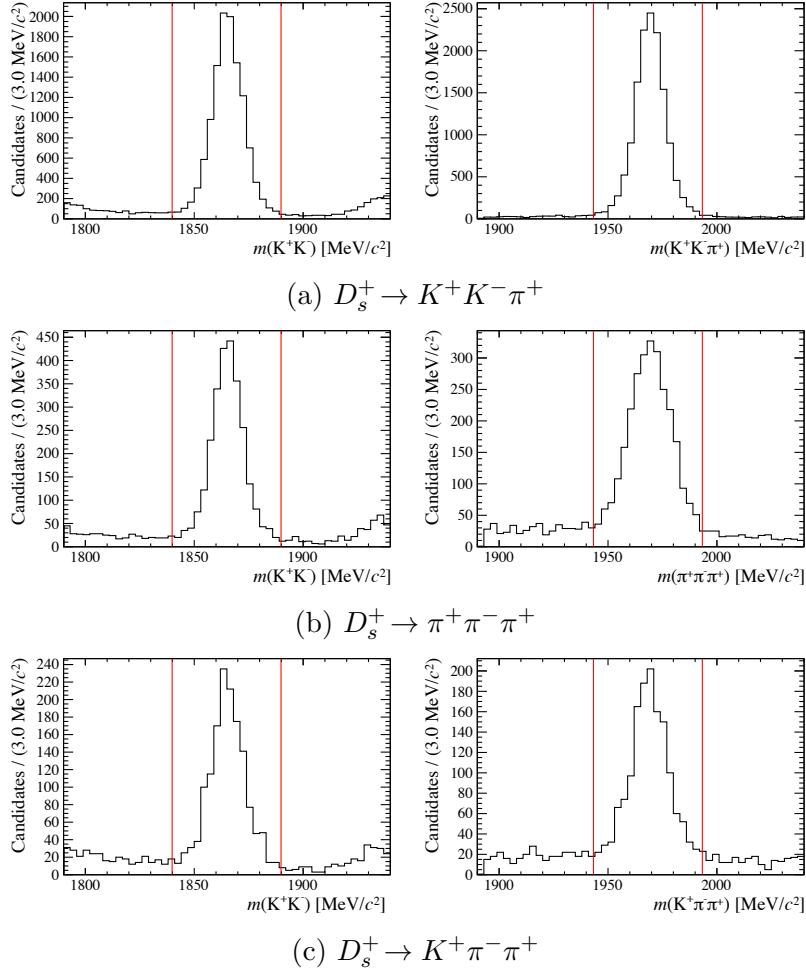


Figure 4.20: The D_s^+ and \bar{D}^0 invariant mass distributions for normalisation $B^+ \rightarrow D_s^+ \bar{D}^0$ candidates passing the selection requirements.

Particle	Requirement
D_s^+	$ m(h^+ h^- h^+) - m(D_s^+) < 25 \text{ MeV}/c^2$
\bar{D}^0	$ m(K^+ K^-) - m(\bar{D}^0) < 25 \text{ MeV}/c^2$
ϕ	$ m(K^+ K^-) - m(\phi) < 40 \text{ MeV}/c^2$

Table 4.20: Invariant mass requirements applied to all candidates.

The distribution of candidates is shown with the different aspects of the selection applied. It can be seen that the level of combinatorial background that is present across the whole mass range dramatically decreases as the different selection criteria are applied. Additionally, it is clear the normalisation channel efficiency decreases too, but at a much slower rate. The final distribution has a negligible amount of combinatorial

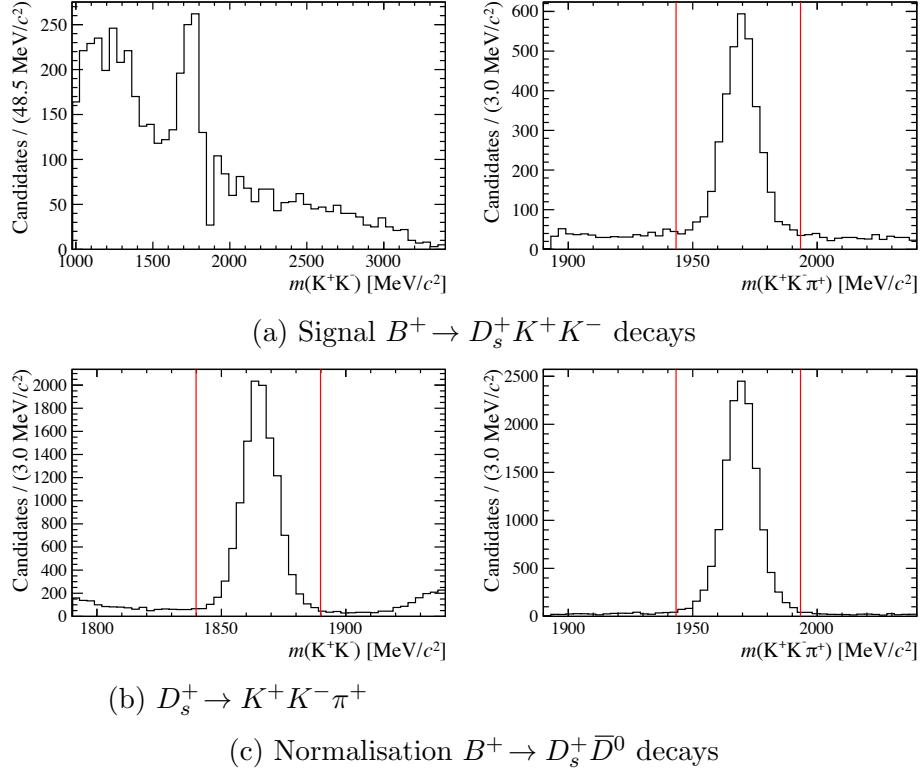


Figure 4.21: The D_s^+ and $K^+ K^-$ invariant mass distributions for $B^+ \rightarrow D_s^+ K^+ K^-$ candidates passing the selection requirements.

background with an isolated normalisation peak.

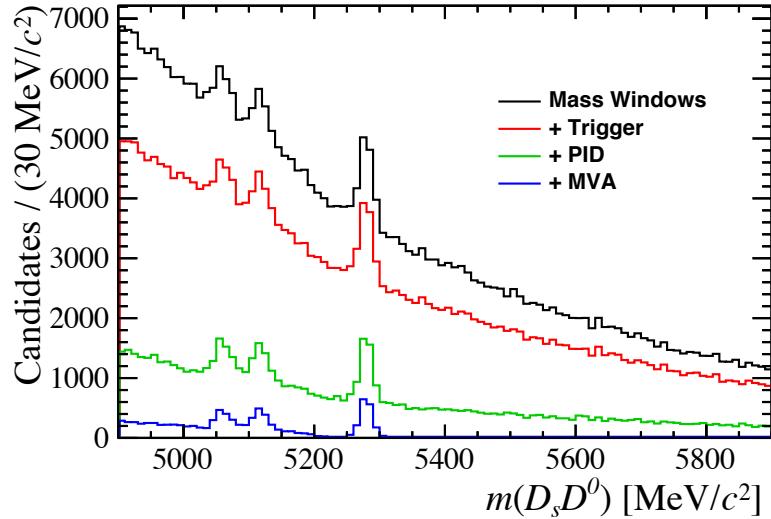


Figure 4.22: Various selection steps applied to the normalisation channel as defined in the legend.