Measurement of the Branching Fraction of the Decay $B^0 \to D^{*\pm}D^{\mp}$ at the Belle Experiment

Daniel Stemmer

Diplomarbeit

Am Institut für Experimentelle Kernphysik
der Fakultät für Physik des
Karlsruher Institut für Technologie (KIT)

Referent: Prof. Dr. M. Feindt
Korreferent: Prof. Dr. T. Müller

31. Mai 2011

Deutsche Zusammenfassung

In dieser Arbeit habe ich das Verzweigunsverhältnis von $B^0 \to D^{*\pm}D^{\mp}$ auf dem vollen Datensatz der Belle-Kollaboration bestimmt. Verzweigungsverhältnisse können mit Hilfe des Standardmodells der Teilchenphysik ausgerechnet werden, weshalb sich ihre Messung als Test des Standardmodells eignet. Erste Messungen des Verzweigungsverhältnisses des Zerfalls $B^0 \to D^{*\pm}D^{\mp}$ wurde bereits am BABAR [1] und am Belle Experiment [2] durchgeführt. Zwischen diesen beiden Messungen besteht eine leichte Diskrepanz, die den Fehler auf den Weltmittelwert, der von der PDG ermittelt wird, vergrößert. Der größte Beitrag zur Unsicherheit stammt von der Belle Messung aus dem Jahre 2001, die lediglich 4.3% des heutigen endgültigen Datensatzes nutzte. Der inzwischen verfügbare Gesamtdatensatz ermöglicht eine genauere Beschreibung von Signal und Untergrund, was auch zur Reduktion von systematischen Unsicherheiten führt.

Im Rahmen des Standardmodells erwartet man im Zerfall $B^0 \to D^{*\pm}D^{\mp}$ mischungsinduzierte CP-Verletzung. Die vorherigen Messungen von BABAR [3],[4] und Belle [5] konnten mischungsinduzierte CP-Verletzung auch mit einer Konfidenz von weniger als drei Standardabweichungen bestätigen. Darum ist es wichtig, die CP-verletzenden Parameter mit dem vollen Datensatz von Belle zu messen.

Das Belle-Experiment befindet sich am KEK, einem Forschungszentrum im japanischen Ort Tsukuba nördlich von Tokio. Es besteht aus dem KEKb-Teilchenbeschleuniger und dem Belle-Detektor. Der KEKb-Beschleuniger [6] ist ein Elektron-Positron Beschleuniger mit einer Schwerpunktsenergie von 10.58 GeV, was gerade der Energie zur Erzeugung des Y(4S) entspricht. Dabei hat der Elektronstrahl eine Energie von 8.0 GeV und der Positronenstrahl eine Energie von 3.5 GeV. Diese Asymmetrie führt zu einem Lorentzboost des produzierten Y(4S) im Laborsystem. Beim Zerfall des Y(4S) entsteht entweder ein $B^0\bar{B}^0$- oder ein B^+B^--Paar. Durch den Boost im Laborsystem ist eine Messung der Zerfallszeitdifferenz der beiden B-Mesonen möglich, was essentiell für die Messung von zeitabhängiger CP-Verletzung ist.

Der Belle Detektor [7] ist ein für Hochenergieexperimente typischer Vielzweckdetektor. Zur Spurrekonstruktion wird ein Silizium-Vertex-Detektor und eine gasgefüllte Driftkammer verwendet. Die Teilchenidentifikation wird zusätzlich mit dem *Time of Flight* System und dem *Aerogel Cherenkov Counter* unterstützt. Zur Energiemessung besitzt der Detektor ein Kalorimetersystem. Ein supraleitender Magnet sorgt für das zur Impulsmessung nötige Magnetfeld. Zur Identifikation von Myonen und Detektion von K_L^0 Mesonen befindet sich das *KLM* als äußerste Lage auf dem Detektor.

Die Datennahme bei Belle erfolgte bis Juni 2010. Dabei wurden 771 fb^{-1} auf der Y(4S) Resonanz aufgezeichnet. Um Signal und Untergrund besser untersuchen zu können,

werden simulierte Daten, sogenanntes Monte Carlo, verwendet. Auf Monte Carlo ist der Ursprung und die Identität jedes Teilchens bekannt. Von der Belle-Kollaboration wird ein *generisches* Monte Carlo zur Verfügung gestellt, welches alle bekannten B-Meson Zerfälle mit den entsprechenden Verzweigungsverhältnissen enthält. Diese Art von Monte Carlo ermöglichte es, die vorliegende Analyse *blind*, also ohne Betrachtung der Daten, zu entwickeln. Der Vorteil einer blinden Analyse ist die Reduktion der unbeabsichtigten Beeinflussung des Messergebnisses durch den Beobachter.

Zur Messung des Verzweigungsverhältnisses von $B^0 \rightarrow D^{*\pm} D^{\mp}$ werden nur bestimmte[1] hadronische Zerfälle rekonstruiert. Das D-Meson, welches direkt beim B-Meson-Zerfall entsteht, wird für die beiden Zerfälle $D^{\mp} \rightarrow K^{\pm} \pi^{\mp} \pi^{\mp}$ und $D^{\mp} \rightarrow K_S^0 \pi^{\mp}$ rekonstruiert. Das $D^{*\pm}$ wird in den hadronischen Zerfällen $D^{*\pm} \rightarrow D^0 \pi^{\pm}$ und $D^{*\pm} \rightarrow D^{\pm} \pi^0$ rekonstruiert. Folgende D-Mesonen aus D* werden rekonstruiert:

- $D^0 \rightarrow K^- \pi^+$
- $D^0 \rightarrow K^- \pi^+ \pi^- \pi^+$
- $D^0 \rightarrow K^- \pi^+ \pi^0$

- $D^0 \rightarrow K_S^0 \pi^+ \pi^-$
- $D^0 \rightarrow K^- K^+$
- $D^+ \rightarrow K^- \pi^+ \pi^+$

Insgesamt werden 12 Zerfallskanäle rekonstruiert, die 2.46% der Zerfälle von $B^0 \rightarrow D^{*\pm} D^{\mp}$ ausmachen.

Zur Selektion des besten Kandidaten eines Ereignisses wird der Chi-Quadrat-Wert aus den D-Mesonmassen und der D*D- Massendifferenz gebildet. Im Fall multipler B-Meson Kandidaten in einem Ereignis wird der Kandidat ausgewählt, der den kleinsten Chi-Quadrat-Wert[2] hat.

Zur Signalselektion wurde ein schnittbasierter Ansatz gewählt. Wichtige Variablen sind dabei die Massen der D-Mesonen, die Massendifferenz zwischen D^* und D-Meson und, sofern neutrale Pionen im Zerfall vorkommen, zusätzlich noch die Pionmasse und die Energie des niederenergetischeren Photons, aus dem das Pion rekonstruiert wurde. Um die bestmögliche Ereignisselektion zu erzielen, wurden die Schnittwerte aller Variablen gleichzeitig optimiert. Als Kennwert der optimiert wurde, diente die Signifikanz, die die Zahl der Signalereignisse dividiert durch die Wurzel der Gesamtereignisse in der Signalregion[3] ist.

$$\Sigma_{signi} = \frac{N_{sig}}{\sqrt{N_{sig} + N_{bkg}}}$$

Die Schnitte wurden für die unterschiedlichen D-Meson-Kanäle einzeln optimiert, da die Massenauflösung der D-Mesonen verschiedener Kanäle im allgemeinen unterschiedlich ist. Durch die Optimierung konnte die Reinheit[4] in der Signalregion von 2.4% nach der

---

[1]Eine Übersicht über alle untersuchten Zerfallskanäle bietet Abbildung 3.1.

[2]Für die genaue Berechnung des Chi-Quadrat-Wertes siehe Kapitel 3.2.2.

[3]Die Signalregion dieser Analyse wird definiert als: $|\Delta E| < 30\text{MeV}$ und $M_{bc} > 5.27\text{GeV}/c^2$. Die Definitionen von ΔE and M_{bc} sind im nächsten Abschnitt zu finden.

[4]Mit Reinheit wird der Anteil an Signalereignissen von einem Datensatz bezeichnet.

Rekonstruktion auf 60.3% erhöht werden.

Das Verzweigungsverhältnis $\mathcal{BR}(B^0 \to D^{*\pm}D^{\mp})$ errechnet sich als Quotient aus der Anzahl der gemessenen Signalereignisse durch das Produkt aus der Zahl der produzierten B-Mesonpaaren und den Rekonstruktionseffizienzen,[5] sowie den Verzweigungsverhältnissen[6] der D und D^* Mesonen.

Während die Zahl der Signalereignisse sowie die Effizienz bestimmt werden, werden die Zahl der B-Mesonpaare, sowie die D und D^* Verzweigungsverhältnisse von der Belle-Kollaboration beziehungsweise der Particle Data Group übernommen. Das Signal wird durch zweidimensionale Parameteranpassung der beiden Variablen M_{bc} und ΔE aus den Daten extrahiert. M_{bc} wird aus den Impulsen der Tochterteilchen und der Strahlenergie errechnet und ist das Äquivalent zur B-Mesonmasse:

$$M_{bc} = \sqrt{(\frac{E_{beam}}{2})^2 - p_{B^0}^2}$$

ΔE ist die Differenz zwischen gemessener Energie des B-Mesons und der halben Strahlenergie:

$$\Delta E = \frac{E_{beam}}{2} - E_{B^0}$$

Ereignisse mit echten B-Mesonen treten gehäuft bei $\Delta E = 0$ und $M_{bc} = M_{B^0} = 5.28\text{GeV}/c^2$ auf. Zur Parameterschätzung wird die *Maximum-Likelihood*-Methode verwendet. Die Parametrisierung des funktionalen Verlaufs von ΔE und M_{bc} muss nicht nur die Signalregion, sondern auch das Seitenband korrekt beschreiben. Kontinuumuntergrund wird in ΔE mit einer Geraden und in M_{bc} mit einer ARGUS-Funktion parametrisiert. Zudem sind noch weitere nicht kontinuierlich verteilte Untergrundquellen vorhanden. Durch irrtümliche Identifikation von Kaonen als Pionen aus den Zerfällen $B^0 \to D^{*\pm}D_S^{\mp}(KK\pi)$ entsteht eine Anhäufung von Ereignissen im negativen ΔE Seitenband. Der umgekehrte Fall, bei dem Pionen fälschlicherweise als Kaonen identifiziert werden, ist auf $B^0 \to D^{*\pm}D^{\mp}(\pi\pi\pi)$ Zerfälle zurückzuführen. Da die Ereignisse in beiden Fällen auch aus einem B-Meson-Zerfall kommen, verhalten sie sich in M_{bc} wie Signalereignisse. Darum wird die korrekte Beschreibung des Seitenbandes in ΔE benötigt, um die richtige Zahl von Signalereignissen zu erhalten.

Auf Monte Carlo, das ausschließlich aus simulierten Signalereignissen besteht, wurde ein nicht kontinuierlicher Untergrund entdeckt, der im Folgenden als *Crossfeed* bezeichnet wird. Dieser Untergrund kommt von teilweise falsch rekonstruierten Ereignissen, wobei neutrale Pionen mit einem falsch rekonstruierten Photon den Hauptteil des Crossfeed ausmachen. Bei der Parameteranpassung können Signal- und Crossfeedanteil nicht getrennt werden, und werden daher gemeinsam behandelt. Um die korrekte Zahl an Signalereignissen auf Daten ermitteln zu können, wird der Crossfeedanteil auf Monte Carlo bestimmt.

Eine weitere Untergrundquelle sind die nichtresonaten Zerfälle $B^0 \to D^{*\pm}K_S^0\pi^{\mp}$. Diese

---

[5]Die Effizienz ist definiert als die Zahl der rekonstruierten Signalereignisse dividiert durch die Zahl der produzierten Signalereignisse. $\epsilon = \frac{N_{rek}}{N_{all}}$

[6]Für die D und D^* Verzweigungsverhältnisse wurden die Weltmittelwerte der PDG verewendet. [8]

sind in ΔE und M_{bc} nicht von Signal zu unterscheiden. Ihr Beitrag wird im D-Massen-Seitenband durch Parameteranpassung bestimmt.

Sowohl die Signal- als auch die Untergrundbeschreibung des oben skizzierten Anpassungsmodells konnten auf Monte Carlo bestätigt werden. Zur Messung der Effizienzen, die für das Verzweigungsverhältnis notwendig sind, wurde auf Monte Carlo zurückgegriffen. Eine zur Konsistenzkontrolle durgeführte Messung des Verzweigungsverhältnisses auf Monte Carlo konnte das Verfahren bestätigen.

Die Messung wurde mit dem vollen Datensatz der Belle-Kollaboration von 772 Millionen B-Meson-Paaren durchgeführt. Dabei wurde die oben skizzierte auf Monte Carlo entwickelte Signalextraktion verwendet. Die Daten werden dabei sehr gut durch das Modell beschrieben. Es wurden 818±39 Signalereignisse gemessen. Zwischen Daten und Monte Carlo besteht ein Unterschied in der Rekonstruktionseffizienz von Kaonen und Pionen bezüglich der Teilchenidentifikation. Die auf Monte Carlo bestimmten Effizienzen wurden dementsprechend korrigiert. Die Messung des Verzweigungsverhältnisse ergibt:

$$\mathcal{BR}\left(B^0 \to D^{*\pm}D^{\mp}\right) = (5.90 \pm 0.28 \pm 0.63) \cdot 10^{-4}$$

Wie üblich entspricht der zweite Term dabei dem statistischen und der dritte Term dem systematischen Fehler des Verzweigungsverhältnisses. Der systematische Fehler errechnet sich aus der Unsicherheit der Zahl der B-Mesonpaare, den Unterschieden der Rekonstruktionseffizienz zwischen Daten und Monte Carlo für neutrale beziehungsweise geladene Pionen oder Kaonen, den Unsicherheiten der D und D^* Verzweigungsverhältnisse und den Grenzen des Anpassungsmodells.

Die Messung von $\mathcal{BR}\left(B^0 \to D^{*\pm}D^{\mp}\right)$ in dieser Arbeit ist die bis heute präziseste Einzelmessung dieses Verzweigungsverhältnisses. Sie ist sehr gut mit dem bisherigen Weltmittelwert [8] von $(6.1 \pm 1.5) \cdot 10^{-4}$ und der letzten Messung von BABAR [1] von $(5.7 \pm 0.7 \pm 0.7) \cdot 10^{-4}$ verträglich. Die Fehler auf das Verzweigungsverhältnis konnten gegenüber der letzten Messung von Belle [2] von $\left(11.7 \pm 2.6^{+2.2}_{-2.5}\right) \cdot 10^{-4}$ deutlich gesenkt werden.

Die in dieser Arbeit entwickelte Ereignisselektion des Zerfalls $B^0 \to D^{*\pm}D^{\mp}$ kann in Zukunft als Ausgangspunkt für eine CP-Verletzungsmessung in diesem Kanal dienen.

Contents

Introduction

> "Each piece, or part, of the whole nature is always an approximation to the complete truth, or the complete truth so far as we know it. In fact, everything we know is only some kind of approximation, because we know that we do not know all the laws as yet. Therefore, things must be learned only to be unlearned again or, more likely, to be corrected.......The test of all knowledge is experiment. Experiment is the sole judge of scientific truth."   Richard P. Feynman

In particle physics, the state of the art *approximation*, to say it with Feynman's words, is the Standard Model of Particle Physics. It is a very powerful theoretical model, which resolves many but not all mysteries of particle physics. The task of an experimentalist is finding appropriate ways to measure the parameters of the standard model. Precise measurements are necessary to qualify oneself in confirming or rejecting theory predictions. Thereby, it is important not to bias the results by expectations of any kind.

The Belle experiment is a high precision experiment dedicated to the investigation of flavor physics in B-meson decays, especially CP-violation studies. The well-understood production mechanism of the B-mesons at the Y(4S) resonance allows for unique experimental techniques.

The branching ratio is the ratio of a specific decay chain compared to all possible decays of the initial state. Though the standard model predicts branching fractions they are observables suitable in comparing theory and experiment. In this thesis, the measurement of the branching fraction of $B^0 \rightarrow D^{*\pm}D^{\mp}$, based on the full data set of the Belle collaboration, will be presented. This branching fraction has been measured before by the BABAR Collaboration [1] and the Belle Collaboration [2]. There is a small tension between those two measurements, which increases the error on the world average of this branching fraction [8]. The biggest contribution to the uncertainty about the world average comes from the previous Belle measurement from 2001, which only used 4.3% of the data set that is available now. The full data set allows for a better understanding of background sources, which should decrease systematic uncertainties about the branching fraction. This thesis is the first work which includes non-combinatorial background in the fit model for the signal extraction in $B^0 \rightarrow D^{*\pm}D^{\mp}$ decays in order to achieve a better separation between signal and background. To minimize biasing from the experimenters expectations, the whole analysis was performed *"blindly"* on Monte Carlo. This means that all selection criteria and fit parameters were determined and optimized on Monte Carlo, without looking at data. In this way the optimization of statistical fluctuations was excluded.

In the channel $B^0 \to D^{*\pm}D^{\mp}$, CP-violation was observed with less than 3σ confidence by previous measurements ([5],[4]), which makes it important for further studies. For a measurement of the parameters describing direct and mixing induced CP-violation, a reliable signal selection is necessary. The signal selection, which was developed for the branching fraction measurement in this analysis, can easily be used for a CP-violation measurement in the future.

1. Theoretical Basics

1.1. Standard Model of Particle Physics

Nowadays four fundamental interactions are known in nature. They are gravity, electromagnetism, weak and strong force. They are important at different length scales, as they differ in relative strength and range. Gravity, which describes the mass attraction, responsible for apples to fall, orbiting moons and the formation of galaxies, has long range. Its potential is proportional to the inverse distance of two interacting objects. Electromagnetism has in principle the same range as gravity, as the potential for two oppositely charged objects is also proportional to the inverse distance between them. But due to the fact that in universe oppositely charged objects are bound together to a neutral composite, like protons and electrons form atoms, the electromagnetic force is not important on a galactic scale. All atomic physics and chemistry can be explained only with electromagnetic interaction. The strong force is responsible for the formation of protons and neutrons, the components of the atomic nucleus. It is restricted to a short range of about one femto meter, the size of a proton. The weak interaction knows no bound states and is restricted to a range of about a few atto meters. Details of the fusion in the sun or radioactive β-decays can only be explained by the weak interaction. The Standard Model of particle physics (SM), which describes objects on a tiny length scale, includes three of the four interactions in quantum field theories. Electromagnetism is described with quantum electrodynamics (QED). Quantum chromo dynamics (QCD) describes the strong interaction and quantum flavor dynamics the weak interaction.

1.1.1. Quantum Numbers and Symmetries

In quantum physics, a quantum system, if not degenerated, is defined by its quantum numbers. Because a particle is a quantum system its characteristics are described by quantum numbers. But in contrast to atomic physics, where a different angular momentum $\vec{L}$ just means another quantum state of the same atom, in particle physics these different states themselves are all distinct particles. Due to the high number of possibilities this leads to a very rich zoo of particles. The Standard Model can explain all these particles as composites of a few elementary particles. The fundamental particles, known as the particle family of the SM, are shown in table 1.1 and 1.2. All particles shown there have been observed, except for the Higgs-boson. The fundamental particles can be ordered further. Classified by spin, we can distinguish fermions with half-integral and bosons with integral spin. Fermions are the building blocks of matter in the SM. They are divided into leptons and quarks, each occurring in three generations

of pairs. Of a lepton pair one has electric charge equal one and a finite mass; the other one, called neutrino, is a neutral particle, which was long believed to be massless. In the last years neutrino experiments established mixing from neutrinos of one generation to neutrinos of another generation. This mixing is explained with a non-zero mass of neutrinos (see [8] chapter 11). But until today only upper limits of the neutrino masses could be established. Leptons only interact by weak interactions, and when electrically charged, also by electromagnetic interactions. Quarks occur in six *flavors* ordered in doublets of an up-type quark with charge $+\frac{2}{3}$ and a down-type quark with charge $-\frac{1}{3}$. They take part in all three interactions. The bosons act as mediators of the three elementary interactions. They couple to particles carrying the charge corresponding to the interaction they are mediating.

It is common to visualize physical processes in particle physics by drawing Feynman diagrams. A schematic picture is shown in figure 1.1. I use a time axis from left to right in all diagrams for convenience. A Feynman graph is characterized by internal and external lines, which each represent a particle[1], and vertices where those lines meet and the interaction takes place. External lines, connected to only one vertex, represent the real particles we can observe in our detector if technically possible. Internal lines are called virtual particles. They are intermediate quantum states, which are not directly observable. They can have an off shell mass, which means that the mass of virtual particles is not restricted by the energy-momentum equation[2]:

$$E^2 - \vec{p}^2 = m^2$$

Although virtual particles with off shell mass are allowed by the Heisenberg uncertainty principle, the amplitude of these transitions are highly suppressed by $\sim \frac{1}{q^2-m^2}$. At every vertex conservation laws hold, which means that the sum of ingoing lines must have the same quantum numbers as the sum of the outgoing lines. The strength of the interaction is described by the coupling constant[3] α, which is different for the three interaction types. To calculate the transition probability from a known initial state to a certain final state, one has to take into account all processes, expressed through Feynman diagrams. Every vertex adds factor of α, the coupling constant of the interaction, to the transition probability.

1.1.2. Quantum Electrodynamics

The photon is the mediator of the electromagnetic interaction. It couples to all charged particles, namely quarks, charged leptons and charged W-bosons. The coupling constant in QED is known from atomic physics as fine structure constant $\alpha \approx \frac{1}{137}$. One example of a QED process is Bhabba-scattering as shown in figure 1.2. Initial and final state are the same in both figures: an electron positron pair. The right figure shows

---

[1]Convention states that particles move along and anti-particles against time in Feynman diagrams.
[2]In this theses natural units are used, where c = 1 and $\hbar$ = 1.
[3]Although it is not a constant as it varies slowly with the energy scale of the interaction.

Table 1.1.: Fermions

Quarks				
Generation	Name	Charge	Mass	couples with
I	up (u)	$+\frac{2}{3}$	1.7-3.3 MeV	γ , g , $W^{\pm}$, Z^0
	down (d)	$-\frac{1}{3}$	4.1-5.8 MeV	γ , g , $W^{\pm}$, Z^0
II	charm (c)	$+\frac{2}{3}$	$1.27^{+0.07}_{-0.09}$ GeV	γ , g , $W^{\pm}$, Z^0
	strange (s)	$-\frac{1}{3}$	101^{+29}_{-21} MeV	γ , g , $W^{\pm}$, Z^0
III	top (t)	$+\frac{2}{3}$	$172.0^{+0.9}_{-1.3}$ GeV	γ , g , $W^{\pm}$, Z^0
	bottom (b)	$-\frac{1}{3}$	$4.19^{+0.18}_{-0.06}$ MeV	γ , g , $W^{\pm}$, Z^0

Leptons				
Generation	Name	Charge	Mass	couples with
I	e	-1	0.511 MeV	γ , $W^{\pm}$, Z^0
	ν_e	0	< 2 eV	$W^{\pm}$, Z^0
II	μ	-1	105.7 MeV	γ , $W^{\pm}$, Z^0
	ν_μ	0	< 2 eV	$W^{\pm}$, Z^0
III	τ	-1	1.78 ± 0.0016 GeV	γ , $W^{\pm}$, Z^0
	ν_τ	0	< 2 eV	$W^{\pm}$, Z^0

Table 1.2.: Bosons

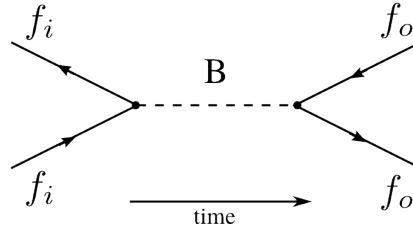| Interaction	Name	Charge	Colour	weak Isospin	Mass
electromagnetic	γ	-1	-	-	0
strong	g	0	r,g,b	-	0
weak	Z^0	0	-	0	91.18 ± 0.002 GeV
	W^+	$+1$	-	$+1$	80.40 ± 0.02 GeV
	W^-	-1	-	-1	80.40 ± 0.02 GeV

Figure 1.1.: An exemplary Feynman diagram with ingoing and outgoing fermion lines
           a boson propagator and two vertices

elastic scattering of electron and positron with a photon exchange. In the left figure
electron and positron annihilate into a photon, which then creates an electron positron
pair out of the vacuum. Both diagrams contribute to the $e^+e^- \to e^+e^-$ transition.
As both diagrams have the same final state, they interfere and the total transition
amplitude $|A_{tot}|^2$ includes also an interference term.

$$|A_{tot}|^2 = |A_{inel}|^2 + A_{inel}A_{el}^* + A_{inel}^*A_{el} + |A_{el}|^2$$

Although the two diagrams are the main contribution of the process, other processes
with more vertices are also possible. In Figure 1.3 some other possible transitions
are shown. They have more vertices for which an additional factor α comes into
the amplitude of the transition probability. Since $\alpha \ll 1$ it is reasonable to take
the additional terms as perturbation in order of α of the first diagrams. Therefore,
perturbative calculations are possible in the QED. They have achieved high precision;
for example, the anomalous magnetic moment of the electron is theoretically calculated
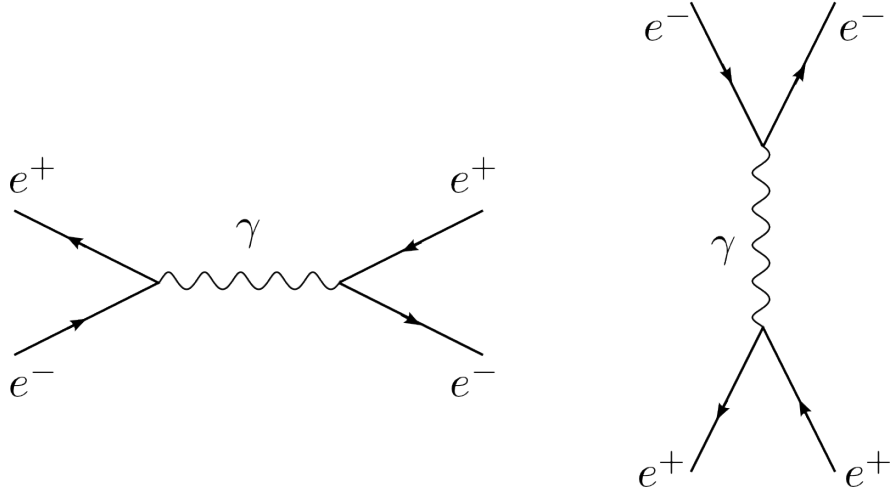very precisely [9].

Figure 1.2.: Bhabba Scattering - The left diagram shows inelastic scattering with annhi-
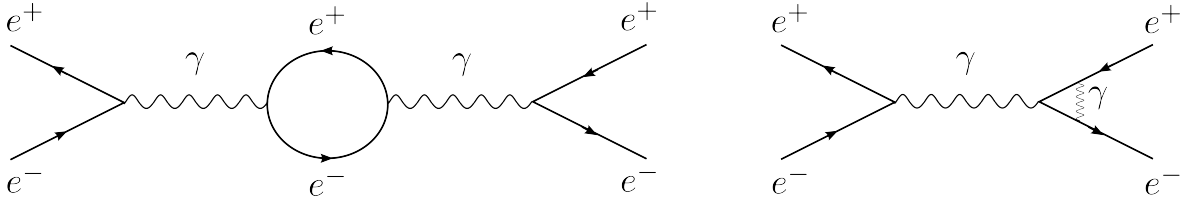   lation and the right diagram shows elastic scattering

Figure 1.3.: Radiation Corrections - Some higher order diagrams for $e^+e^- \to e^+e^-$

1.1.3. The Strong Interaction

The strong interaction is restricted to the quark and gluon sector. The quantum chromo dynamics (QCD) describes the strong interaction, which occurs between quarks via the exchange of gluons. Strong interaction is sensitive to a quantum number, called color, which only quarks and gluons possess. There are three types of color, called red, green and blue and their anti-colors (anti-red, anti-green and anti-blue). In nature only colorless objects are observed, which comes from quark *confinement*. It can be explained on the basis of a $q\bar{q}$-pair, also illustrated in figure 1.4. If quark and anti-quark, which form a $\bar{q}$-pair, are in close vicinity to each other, they interchange only a few gluons. As gluons carry color themselves, they can interact with each other. With increasing separation of the $q\bar{q}$-pair, the gluon-gluon-interaction increases, which draws more energy into the color-field, until it is enough to create a new $q\bar{q}$-pair out of the vacuum. That is why quarks cannot be isolated. Therefore, quarks[4], which all have a certain color themselves, add up to color-less objects called hadrons. Two types of hadrons, mesons and baryons, are known. A meson is a composed state of a quark and an anti-quark with opposite color. A B^0-meson, for example, is a composite of a d-quark and $\bar{b}$-quark. Baryons consist out of three quarks, who all have different color, so that their combination is colorless. The principle of the color addition can be easily visualized by the color wheel shown in figure 1.5. It can be seen, that not only color and anti-color cancel each other out, but if added vectorial also the combination of the three colors red, green and blue (or the three anti-colors).

With three different colors there are eight linear independent gluons, all carrying color and anti-color, possible. Using a specific convention [9], this gluon-octet can be written in the following way:

$$r\bar{g} \quad r\bar{b} \quad g\bar{b} \quad g\bar{r} \quad b\bar{r} \quad b\bar{g} \quad \frac{r\bar{r} - g\bar{g}}{\sqrt{2}} \quad \frac{r\bar{r} + g\bar{g} - 2b\bar{b}}{\sqrt{6}}$$

Theoretically there is also a ninth combination possible, but this gluon would be a completely symmetric color singlet that would not interact.

The gluon-gluon-coupling is responsible for the curious behavior of the potential of strong interactions. The Yukawa potential $V_S(r)$ of the strong force between a quark

---

[4]The top-quark is one exception, as its lifetime is so short that it decays before it can form hadrons.

and an anti-quark pair with distance r to each other can be written as:

$$V_S(r) = -\frac{4}{3}\frac{\alpha_S(r)}{r} + kr$$

For large r the last term, which describes the confinement, is dominating. The parameter k describes the length scale of strong interactions. For short distances, which is equal to high energies in quantum field theory, the first term which is similar to the Coulomb potential is dominating. This means that for very short distances the quarks are *quasi-free*, which is called *asymptotic freedom*. The coupling constant $\alpha_S(r)$ itself is dependent on the length (or energy) scale. The coupling constant decreases with increasing energy. This comes from an *anti-screening* [9] effect of the color-field in close vicinity of a color charge.
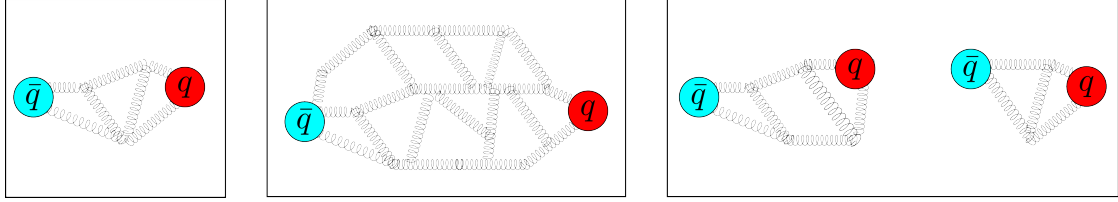

Figure 1.4.: Confinement - In the left figure a $q\bar{q}$-pair exchanging a few gluons is shown. As soon as they are separated, shown in middle figure, more and more gluons occur drawing energy into the color-field. If they are separated further, a $q\bar{q}$-pair is created, which is why quarks cannot be isolated.
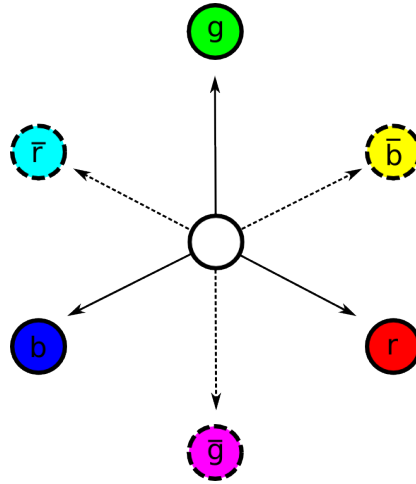

Figure 1.5.: The colors of quarks can be illustrated by the color wheel. In nature only "white" or "colorless" objects occur as free particles. The color of an object is the vectorial sum of the colors of its constituent particles.

1.1.4. The Weak Interaction

The weak interaction is mediated by the charged W-bosons and the neutral Z^0-boson. In contrast to the gluon or the photon, they have a mass of almost hundred GeV. The W- and Z-bosons interact with every elementary fermion, as can be seen in the fermion summary table 1.1. Weak transitions are highly suppressed due to the non-zero mass of the mediating bosons. They are therefore only important, if no strong or electromagnetic interaction is possible or the interaction energy is above the W (or Z) mass. But there are still many processes which are only possible through the exchange of a virtual W or Z-boson. The decay of B^0-mesons for example is only possible through weak interactions. A B^0 is the lightest meson combination with a b-quark in it. If the b-flavor would be conserved, this meson could not decay and would be stable. But W-bosons mediate flavor changing transitions, from up-type to down-type quarks and vice versa. Thereby transitions within one generation are most likely, but also transitions from one generation to another are allowed. The Cabibbo-Kobayashi-Maskawa matrix (CKM) describes the strength of the flavor changing transitions. It is a unitary 3×3 matrix, shown below (numerical values obtained from [8]). The absolute square of a matrix element V_{ij} is proportional to the transition probability[5] from q_i to q_j.

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \approx \begin{pmatrix} 0.974 \pm 0.0002 & 0.225 \pm 0.0009 & 0.004 \pm 0.0004 \\ 0.230 \pm 0.011 & 1.023 \pm 0.036 & 0.040 \pm 0.001 \\ 0.008 \pm 0.0006 & 0.039 \pm 0.002 & 0.88 \pm 0.07 \end{pmatrix}$$

$$\begin{pmatrix} |u\rangle \\ |c\rangle \\ |s\rangle \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} |d\rangle \\ |s\rangle \\ |b\rangle \end{pmatrix}$$

$$V_{CKM} V_{CKM}^{-1} = V_{CKM} V_{CKM}^{\dagger} = 1$$

As an down-type quark the b-quark of the B^0 meson can only decay into up-type quarks. The top quark is not possible, since it is heavier than the bottom, so only up-quark and charm-quark remain. Looking at the CKM-matrix, one sees that the $b \to c$ transition is favored by CKM over $b \to u$. The tree level contribution of the decay of B^0 to $D^{*\mp} D^{\pm}$ investigated in this thesis is shown in figure 1.6. The $\bar{b}$-quark with charge $+\frac{1}{3}$ is transformed into a $\bar{c}$ with charge $-\frac{2}{3}$ by emitting a W^+-boson. The W^+ decays into $c\bar{d}$, which is CKM-suppressed with $|V_{cd}|^2$.

   W-bosons can also mediate transitions from leptons to neutrinos (and vice versa) within a lepton generation[6]. The neutral Z^0-boson couples to fermion anti-fermion pairs of the same type, like $q\bar{q}$, l^+l^- or $\nu\bar{\nu}$. Weinberg, Glashow and Salam proposed

---

[5]In general the squared matrix element is multiplied with a phase space factor. The phase space includes all possible states of a certain system and is obtained by integrating over all possible momenta of the final state. Therefore, if the four-momentum difference between initial and final state in one process is larger than in another process, the phase space factor in this process is also larger.

[6]The results of neutrino appearance and disappearance experiments suggest also mixing in the lepton sector.

$$\bar{d}$$
$$D^+$$
$$W$$
$$c$$
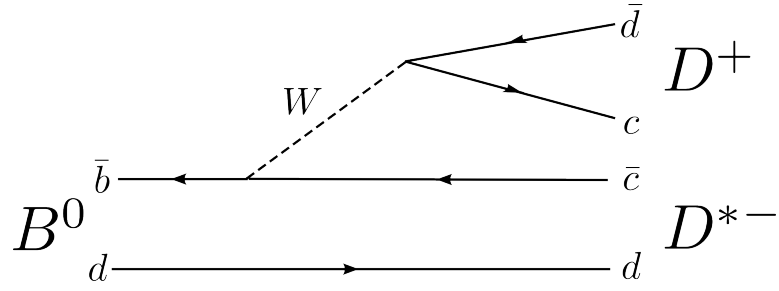$$\bar{b}$$
$$B^0$$
$$\bar{c}$$
$$D^{*-}$$
$$d$$
$$d$$

Figure 1.6.: Tree level diagram $B^0 \rightarrow D^{*-}D^+$

a unification of electromagnetic and weak interaction, which is known as electroweak interaction. The electroweak interaction is explained e.g. in reference [10].

1.2. CP-Violation

1.2.1. Discrete Symmetries

CP is the product of charge conjugation C and parity P. Therefore, I will first explain parity and charge conjugation.

Parity

A parity transformation flips the sign of all space coordinates. Therefore, if done twice the system is in the same state as before.

$$P|\vec{x}\rangle = |-\vec{x}\rangle \quad P^2|x\rangle = |x\rangle$$

Naively one would think that physical laws are independent of parity transformation, since the coordinate system is arbitrary. An observable sensitive to parity transformation is the helicity[7], which is standardized product of momentum and spin of a particle.

$$h = \frac{\vec{p}\vec{s}}{|\vec{p}||\vec{s}|}$$

For fermions, with spin $\frac{1}{2}$ there are two possible orientations of the spin to a certain axis. With a parity transformation, the momentum flips its sign, but the spin stays as it is, which is why the helicity of these particles also changes its sign under parity transformation. Particles with negative helicity are called left-handed and particles with positive helicity are called right-handed. Wu showed in 1957 [10] that parity is violated in weak decays. In fact, W-bosons only couple to left-handed fermions and right-handed anti-fermions. For neutrinos, which only interact weakly, this means that there are only left-handed neutrinos and right-handed anti-neutrinos observable.

Charge Conjugation

Charge conjugation is another discrete symmetry. A charge conjugation transforms all particles into their anti-particles. Therefore, only neutral[8] particles or neutral composites, like the $\phi = (s\bar{s})$, can be an eigenstate of charge conjugation.

$$C|f\rangle = |\bar{f}\rangle \quad C^2|f\rangle = |f\rangle$$

CP and CP-Violation

CP is the combined transformation of C and P. It transforms a left-handed neutrino ν_L in a right-handed anti-neutrino ν_R for example. After discovery of P violation many physicists hoped that in nature at least CP is not violated. They were proved wrong

---

[7]The helicity itself is not Lorentz-invariant. In quantum field theory, the *chirality*, which is Lorentz invariant, is used. For $\beta \to 1$ the helicity becomes equal to the chirality. The helicity is used here for simplicity.

[8]Neutral means here: neutral in charge and flavor.

in 1964, when Cronin and Fitch discovered CP-violation in the kaon system.
Today there are three types of CP-violation known.

1.2.2. Neutral Meson Mixing and Indirect CP-Violation

Flavored neutral mesons with one light and one heavier quark can mix via exchange
of two W-bosons, which is described through a so-called box diagram. An example for
a box diagram is shown in figure 1.7. Another box diagram of the same order can be
obtained just by rotating all internal lines clockwise by 90 degrees. The $\bar{b}$-quark emits[9]
a W^- changing its flavor to an up-type anti-quark, while the d-quark absorbs the W^-
and is turned into an up-type quark. With a second W-boson exchange, a b-quark
and a $\bar{d}$-quark, which form a $\bar{B}^0$, are obtained. The transition from $\bar{B}^0$ to B^0 works
the same way. Due to mixing, the mass eigenstates are superpositions of the flavor
eigenstates $|B^0\rangle$ and $|\bar{B}^0\rangle$. The two mass eigenstates are called $|B_L\rangle$ (light) and $|B_H\rangle$
(heavy):

$$|B_L\rangle = \frac{1}{\sqrt{2}}\left(p|B^0\rangle + q|\bar{B}^0\rangle\right)$$

$$|B_H\rangle = \frac{1}{\sqrt{2}}\left(p|B^0\rangle - q|\bar{B}^0\rangle\right)$$

With the phase convention that $CP|B^0\rangle = |\bar{B}^0\rangle$, the B_L and B_H become CP-eigenstates
if $|\frac{q}{p}| = 1$. In contrast, if $|\frac{q}{p}| \neq 1$ there exists indirect CP-violation. Indirect CP-
violation occurs if the amplitude for mixing from the first flavor-state to the second is
different from the amplitude of mixing from the second to the first. To explain this
difference one first has learn a bit more about the CKM-matrix. The CKM-matrix has
in fact complex components in its matrix elements, as it can be seen in the Wolfenstein-
parametrization[10]. The parameters λ, A, ρ and η are real, thus ρ plus $i\eta$ is a complex
number. At third order of λ two elements, V_{td} and V_{ub}, have an imaginary part and
thus also a phase. Looking again at the box diagrams 1.7 one can see that the rel-
evant CKM-elements are different for two mixing-processes. One can show [11] that
$\frac{q}{p} = \frac{V_{td}V_{tb}^*}{V_{td}^*V_{tb}}$. This leads to indirect CP-violation, since V_{td} is complex.
Indirect CP-violation is still [8], almost 50 years later, only measured in the kaon sys-
tem, where CP-violation was first observed. The great difference in lifetime, about two
orders of magnitude, of the two neutral kaon mass eigenstates makes the kaon system
an exception. It allows to create a pure sample of the long-living mass eigenstate, which
is necessary for an indirect CP-violation measurement. In the B^0, no pure samples of
B_L and B_H can be made.

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \approx \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ \lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4)$$

---

[9]This is equivalent to the b-quark absorbing a W^+
[10]Shown in reference [11] for example

with:[11]

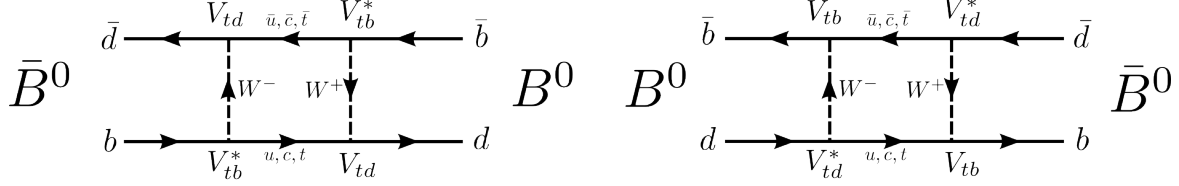$$\lambda \approx 0.225 \quad A \approx 0.808 \quad \rho \approx 0.132 \quad \eta \approx 0.341$$

Figure 1.7.: Two box diagrams responsible for $B^0\bar{B}^0$-mixing are shown. Because V_{td} is a complex number, there is a difference in the amplitude of both diagrams.

1.2.3. Direct CP-Violation

Another possibility for CP-violation is direct CP-violation. Direct CP-violation occurs when the amplitudes from an initial CP eigenstate $|i\rangle$ to a final CP eigenstate $|f\rangle$ are different for CP conjugated states.

$$|\langle i|H|f\rangle|^2 \neq |\langle \bar{i}|H|\bar{f}\rangle|^2$$

Direct CP-violation is independent of mixing and also occurs in the decays of charged mesons which do not mix. Different amplitudes for CP-conjugated decays can only arise when there are underlying processes with a complex phase difference which interfere. Figure 1.8 shows the two diagrams, which are in theory responsible for direct CP-violation in $B^0 \to D^{*\pm}D^{\mp}$. The tree diagram on the left has no complex CKM matrix elements in the transition. In the so called penguin diagram on the right, the complex matrix element V_{td}^* comes in. If only the penguin diagram would contribute to the transition, there would be no direct CP-violation expected, since the imaginary part, the phase, would be nullified by calculating the amplitude. But if combined with the amplitude of the tree diagram (A_t) the weak phase can survive the absolute quadrature.

$$|A_p|^2 = |C_p e^{i\phi}|^2 = \text{real}$$

$$|A_p + A_t|^2 = C_p^2 + \underbrace{C_t C_p e^{i\phi}}_{imaginary} + C_t^2$$

with the penguin amplitude $A_p = C_p e^{i\phi}$ and the tree amplitude $A_t = C_t$, whereby C_t, C_p and ϕ are real.
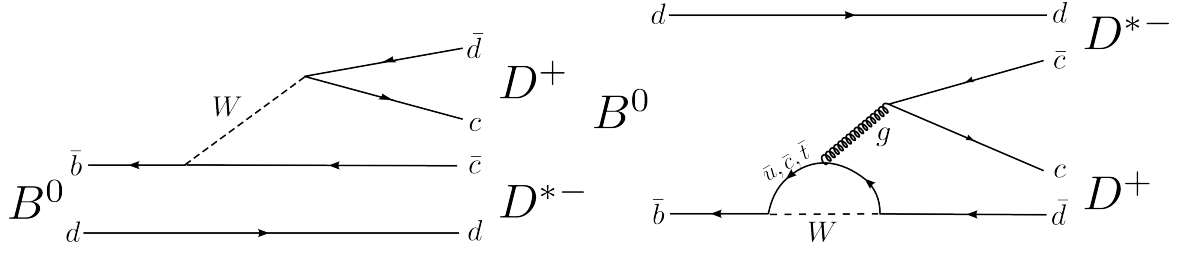
---

[11]Values obtained from [8]

Figure 1.8.: On the left side the tree level process for $B^0 \to D^{*\pm}D^{\mp}$ and on the right
        side the first order penguin diagram

1.2.4. Mixing Induced CP-Violation

The third possibility and the one expected to have the largest contribution for the
studied decay, is the mixing induced CP-violation. It can occur, if two initial CP-
eigenstates ($|i\rangle$ and $|\bar{i}\rangle$) can both decay in the same CP-eigenstate ($|f\rangle$) and the initial
states can mix into each other. There is CP-violation if:

$$A_f = |\langle i|T|f\rangle|^2 \neq |\langle \bar{i}|T|f\rangle|^2 = \bar{A}_f$$

I will explain mixing induced CP-violation on the system $B^0 \to D^{*\pm}D^{\mp}$. As we have
seen in the previous chapter about mixing, there is a weak phase difference $e^{i\phi} \sim \frac{V_{td}V_{tb}^*}{V_{td}^*V_{tb}}$
due to the mixing. The phase difference alone would yield to indirect CP-violation,
which is not measurable at present[12]. But due to the interference between decays,
which have mixed before they are decayed and those who don't, there is a measurable
CP-violation expected.
In principle there are two ways, shown in figure 1.9, for an initial state $|B^0\rangle$ to decay
into its final state $|D^*D\rangle$ with and without mixing into $\bar{B}^0$ before. The difference in
the decay amplitude of the initial $|B^0\rangle$ and the initial $|\bar{B}^0\rangle$ is measured by counting the
events coming from a certain initial state. Since the B-mesons oscillate with a certain
frequency into each other, which is given by their mass-difference, and decay into the
same final state, one cannot decide what the initial state of this decay was. But the
two B-mesons were produced as a coherent quantum state. Therefore, when the first
B-meson decays, let us say as a B^0, one can say that the other B-meson in this instance
is a $\bar{B}^0$. Later on, the $\bar{B}^0$ can oscillate, but in this instance its flavor is determined by
quantum physics. This is used to determine the flight length difference of the initial
produced B^0's and $\bar{B}^0$'s. The time-dependent asymmetry A_f is then given in the usual
([8] page 160) notation[13]:

$$A_{f\pm}(\Delta t) = \frac{\frac{d\Gamma}{dt}(|\bar{B}^0\rangle \to |D^{*\pm}D^{\mp}\rangle) - \frac{d\Gamma}{dt}(|B^0\rangle \to |D^{*\pm}D^{\mp}\rangle)}{\frac{d\Gamma}{dt}(|\bar{B}^0\rangle \to |D^{*\pm}D^{\mp}\rangle) + \frac{d\Gamma}{dt}(|B^0\rangle \to |D^{*\pm}D^{\mp}\rangle)}$$

---

[12]"The small deviation (less than one percent) of $|\frac{q}{p}|$ from 1 implies that, at the present level of
    experimental precision, CP violation in B mixing is a negligible effect." ([8] page 160)
[13]Because the time difference Δt is used, t is replaced by Δt

$$A_{f\pm}(\Delta t) = S_{f\pm}sin(\Delta M \Delta t) + C_{f\pm}cos(\Delta M \Delta t)$$

with

$$S_{f\pm} = \frac{2 \cdot \mathbf{Im}(\lambda_f)}{1 + |\lambda_f|^2}, \qquad C_{f\pm} = \frac{1 - |\lambda_f|^2}{1 + |\lambda_f|^2},$$

and

$$\lambda_f = e^{i\phi}\frac{A_f}{\bar{A}_f}.$$

The parameters $S_{f\pm}$ and $C_{f\pm}$ are called the CP-violating parameters for mixing induced CP-violation. To determine these parameters a fit to the flight-length difference of the B (which corresponds to the lifetime difference) has to be performed for initial B^0 and $\bar{B}^0$ events. As $B^0 \rightarrow D^{*+}D^-$ is distinguishable from $B^0 \rightarrow D^{*+}D^-$, the measurement of A_f has to be split into two measurements, A_{f+} and A_{f-}.
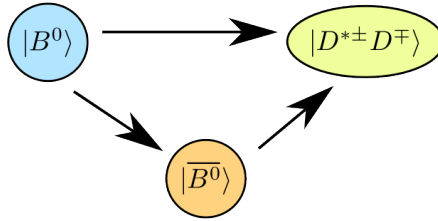
Figure 1.9.: Decay with and without mixing

1.2.5. Time Integrated Charge Asymmetry

Another quantity which can be measured in $B^0 \rightarrow D^{*\pm}D^{\mp}$ decays is the asymmetry $A_\pm$ of $D^{*+}D^-$ and $D^{*-}D^+$ decays. It is expected to be zero, according to the standard model.

$$A_\pm = \frac{N_{D^{*+}D^-} - N_{D^{*-}D^+}}{N_{D^{*+}D^-} + N_{D^{*-}D^+}}$$

2. Experimental Setup

If you want to observe natures inner fragments, you need tools which allow you to observe them. As we know from optics, one can only resolve objects, if the wavelength of the probing beam is in the order of the size of the examined object. The probing beam can be light, or a beam of electrons or protons. Quantum Mechanics explains the wave-particle dualism, which leads to the de-Broglie-relation. It states that the wavelength of a given particle is the Planck constant divided through the particles momentum.

$$\lambda = \frac{h}{|\vec{p}|}$$

In particle physics, the goal is to observe the smallest fragments of nature, therefore a probing beam with very tiny wavelength, which corresponds to high energetic beam-particles, is needed.

Due to the high momentum of the particles, a relativistic description is necessary. In special relativity, four-vectors in Minkowski-space with one time coordinate and three space coordinates are used to describe physics laws.

At B-factories, a maximal production of B-mesons is requested, therefore B-factories normally run at the Y(4S)-resonance, shown in figure 2.1, which is just above the threshold of $B\bar{B}$-production. Due to their short lifetime it is not easy to examine mesons, especially the heavy B-mesons, in nature. We can create them artificially in the lab. Therefore, we exploit Einstein's famous energy-mass relation, which states that matter can be formed out of energy and that the annihilation of a particle releases free energy proportional to the particle's mass.

$$E = m \cdot c^2.$$

In collider experiments, one uses the energy of the colliding particles to create new particles. The energy of the collision is measured in the center-of-mass system (CMS). The CMS energy of the collision can be easily expressed in four-vector notation with the four-momentum:

$$p^\mu = \left(\begin{array}{c} E/c \\ \vec{p} \end{array} \right).$$

$$\sqrt{S} = (p^\mu + p^\nu).$$

One can distinguish two designs of accelerator experiments fixed target experiments and colliding beam experiments. In fixed target experiments, one beam of particles
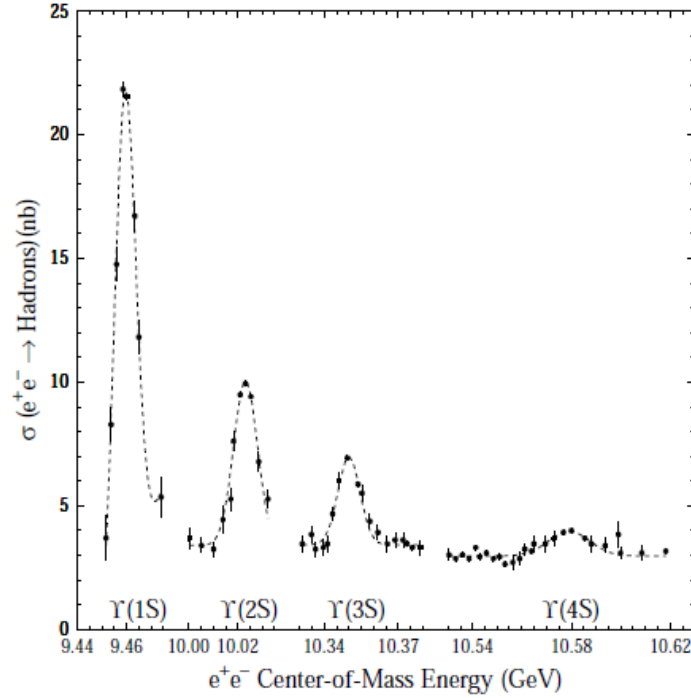
Figure 2.1.: The Y-resonances [12]

(electrons, for example) is accelerated and shot at a stationary target (a block of lead atoms for example). The CMS-energy (referred here as $\sqrt{S}$) in fixed target experiments is proportional to the square root of the energy of the colliding beam, $\sqrt{S} \propto \sqrt{E_{beam}}$. At colliding experiments in contrast to fixed target experiments, two particle beams are accelerated in opposite directions and crossed at the interaction point, where they scatter inelastically and generate new particles, which then can be measured. With the assumption of symmetric beam energies and head on head collision, the center-of-mass energy is proportional to the beam energy: $\sqrt{S} \propto E_{beam}$.

Topology of Accelerator Types

If one discriminates accelerators by their topologies, two main types: cyclic and linear accelerators can be distinguished. Both types use electric fields for acceleration. While in linear accelerators (LA) the distance of acceleration is passed only once, there are multiple acceleration cycles in circular accelerators (CA). Therefore CA can be more compact than LA and still have the same or higher distance of acceleration in total. To hold the particles in CA on their path, magnetic fields are used. The higher the energy of accelerated particle becomes, the higher the magnetic field has to be to hold it on its course. Many CA in high energy physics increase the magnetic field synchronically with the acceleration process to maintain a constant curvature of the particle beam, and are therefore called synchrotrons. Thus, the acceleration limit of synchrotrons is given by the maximal power of the magnetic field and the radius of the ring, and

by the amount of dissipative synchrotron radiation. Synchrotron radiation is a form of bremsstrahlung and occurs if charged particles are forced on a circular path. The energy loss through synchrotron radiation is proportional to inverse mass of the fourth power:

$$\Delta E_{loss} \propto \frac{1}{m^4}.$$

Hadron and Lepton Colliders

Another way to distinguish collider experiments is by their accelerated particles. Until today, all accelerators use only stable particles, namely protons and electrons and their anti-particles. Cyclic hadron colliders with proton and (anti-)proton beams can reach higher energies than lepton colliders, since light particles like electrons loose much more energy through synchrotron radiation than heavier particles like protons. Also the cross-section of B-meson production increases with the energy, which is why it is easier for hadron machines to collect a huge number of B-mesons. One drawback of hadron colliders is that the center-of-mass energy does not match the beam energy, since not all parts of the hadron interact in a collision. This restricts especially the missing energy resolution, as a crosscheck between measured energy in the calorimeters and beam energy is not possible. At hadron colliders, interesting events are produced with the two processes gluon-gluon-fusion and quark-anti-quark annihilation. The remaining parts of the (anti-)proton still can produce particles which are not wanted and therefore pollute the signal event. In comparision, an event with a lepton-lepton maschine is much more clean, which can be seen in figure 2.2. The disadvantage of cyclic lepton-maschines is the energy loss due to synchrotron radiation. Its advantage is that the center-of-mass energy is known in an event and the event itself is normally much cleaner than at hadron colliders.

There are also hadron-lepton-machines, like HERA at DESY in Hamburg, but they are not explained in detail here, as they are sensitive to other field of particle physics, which are not discussed in this thesis.

2.1. The Belle Experiment

The Belle experiment is situated at the national research center for nuclear physics KEK in Tsukuba, Japan, about 60 kilometers northeast of Tokio. About 380 scientists from 14 countries work in the Belle colaboration. The two main parts of the experiment are the KEKB accelerator explained in detail in reference [6] and the Belle detector described in detail in reference [7]. The important details of the experiment were summarized in the following.

2.1.1. The KEKB Accelerator

The B-mesons for this analysis were produced with the KekB accelerator. Hence, the KekB accelerator will be explained in more detail. The KekB accelerator is a electron-
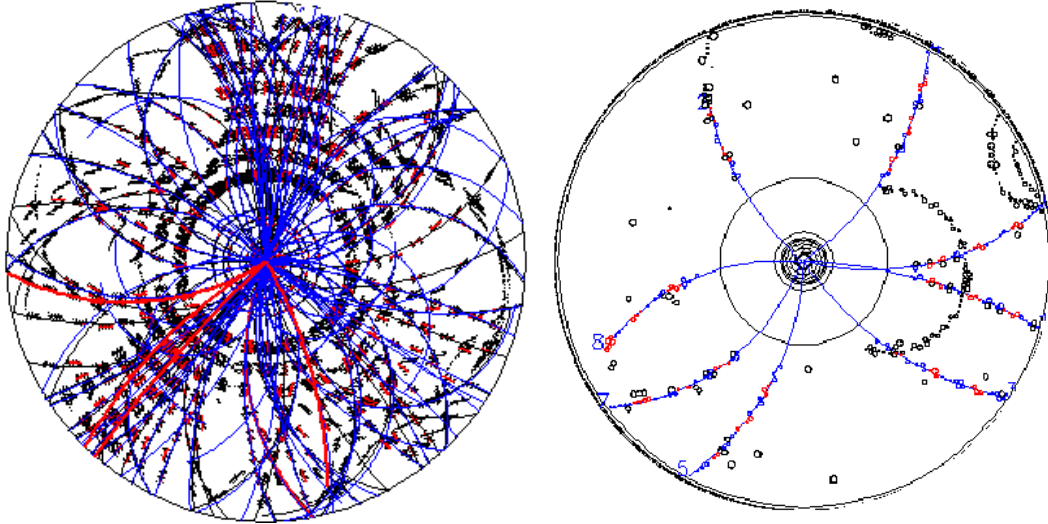
Figure 2.2.: A typical event at a hadron collider (CDF) and a lepton collider (Belle) [13]

positron synchrotron machine. The accelerator consists of two rings the high energy ring (HER) for electrons and the low energy ring (LER) for the postiron beam. Both rings are installed side-by-side in the old TRISTAN[1] tunnel. With 8 GeV for electrons and 3.5 GeV for positrons, the CMS-energy is 10.58 GeV. This is just at the Y(4S) resonance, which is shown in figure 2.1. The Y(4S) consists of a $b\bar{b}$ quark anti-quark pair and decays practically only in B-Mesons[2].

A schematic picture of the accelerator is shown in figure 2.4. Electrons and positrons are injected through the LINAC[3] complex into the two rings, already at their full energy. The rings are filled with dense packages of electrons (or positrons), the so called bunches. Though already at their intended energy when entering the ring, the bunches have to be accelerated further to compensate the energy-loss through synchrotron radiation. This is done with compressed radio frequency pulses, which require separate bunches instead of a continuous beam. The interaction point is situated at the opposite side of the injection point. At this point, where the Belle detector is located, the electron beam and the positron beam cross each other with a frequency of approximately 508,9 MHz at an angle of about 22mrad. To achieve a higher effective crossing area, special magnets called crab cavities, are used flip the bunches just before the collision so that they can collide head on head. A schematic picture of how crab cavities work is shown in figure 2.3.

---

[1]TRISTAN was a former collider experiment that also used the tunnel.

[2]In the following it is assumed that the Y(4S) decays equally into $B^0\bar{B}^0$- and B^+B^--pairs.

[3]The LINAC complex consists out of different accelerators for pre-acceleration and positron production. For more information have a look at [14].
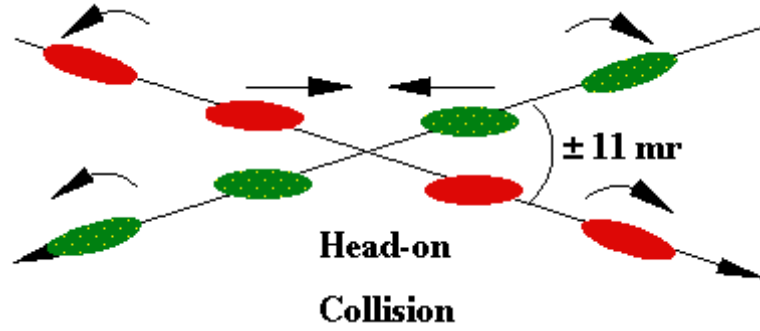
Figure 2.3.: The crab cavities [15] are special magnets, which rotate the bunches just before the collision so that they collide head-on. After the collision, the remains of the bunch are rotated back to reduce random beam scattering.
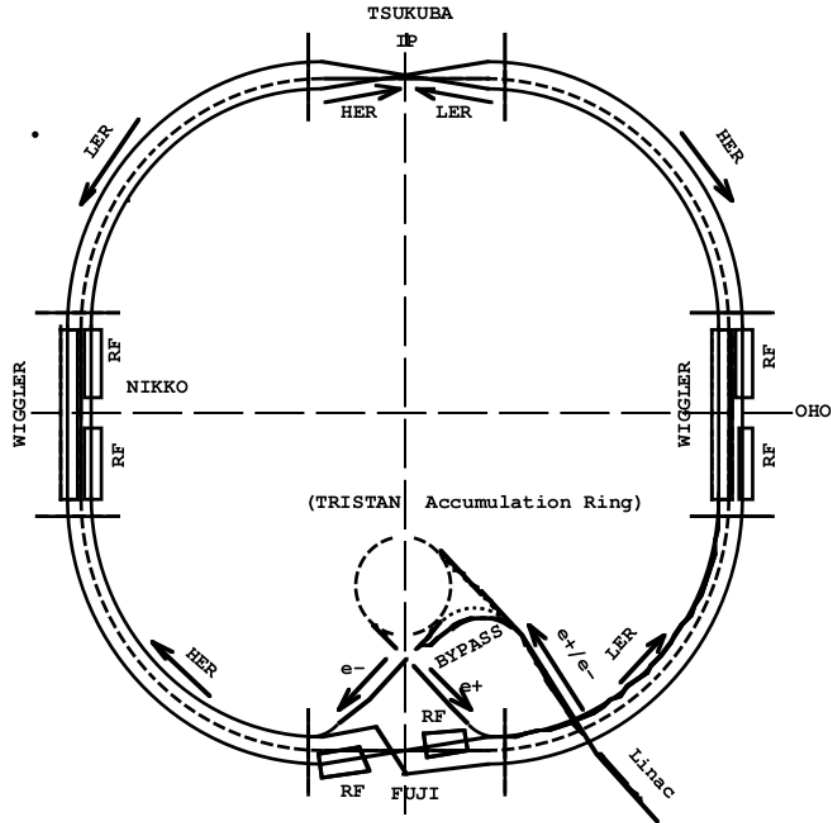

Figure 2.4.: A schematic view of the KEKb accelerator [6] is shown. The injection with the LINAC takes place at FUJI hall, which is situated opposite of the interaction point at TSUKUBA hall. Additional acceleration to maintain the energy is done with radio frequency pulses (RF) at different locations. Deflecting magnets are situated all along the curvatures.

2.1.2. The Belle Detector

The actual measurement is done by the Belle detector, which is situated around the interaction point. The word Belle is a coinage which should paraphrase the goal of the experiment. The *B* stands for the B-mesons produced in the experiment, *el* stands for the electron and the *le* for its anti-particle, the positron which expresses that a electron-positron collider is used.

The Belle detector is constructed around the beam pipe, where the collisions take place. In consists out of different components for various purposes, such as particle identification, momentum determination and energy measurement. Due to the difference of electron and positron beam energy, the Y(4S) is produced with a boost in the lab frame. Therefore, the lifetime of the B-mesons is prolonged significantly so that it is possible to determine the vertex position, which is crucial for a time dependent measurement, necessary in mixing induced CP-violation measurements, for example. For the Belle detector, the coordinate system is chosen in a way that the z-axis points along the boost of the Y(4S), the so-called forward direction. The y-axis is set to be along the vertical and the x-axis along the horizontal. The Polar angle Θ is measured from z-axis to the xy-plane. The angle Φ is measured in the xy-plane with respect to the x-axis.

The Belle detector is shown in figure 2.5. It is a typical large-solid-angle magnetic spectrometer which consists out of different components.
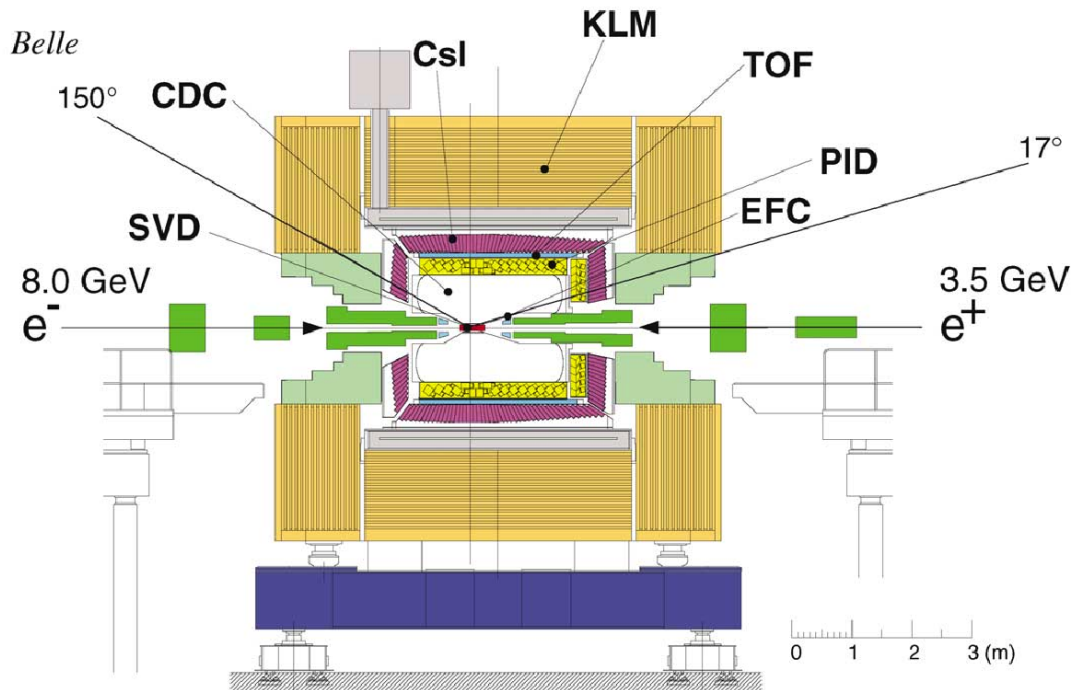

Figure 2.5.: The Belle Detector [7]

Beam Pipe

The beam pipe is situated in the innermost part of the detector. It has a total diameter of 48mm. The purpose of the beam pipe is simple: it has to provide a vacuum to minimize undesired bunch scattering and present as little resistance as possible to the produced particles traversing through it towards the actual detector. Another issue is the heat dissipation due to coulomb scattering at the beam pipe wall. Therefore, a double layer beryllium wall of 0.5mm thickness each is used. There is a 2.5mm gap between the beryllium walls, which is filled with helium for cooling.

Silicon Vertex Detector

The first layer directly put on top of the beam pipe is the silicon vertex detector (SVD). In order to measure the lifetime of short living particles, such as B-mesons, it is crucial to know the vertices where they decayed. While fast decaying particles already decay in the beam pipe, secondary vertex reconstruction (SVR) is the only way to determine their decay vertex. In SVR, the tracks of the two particles are fitted to a common point, which requires a good track resolution the SVD fortunately can provide. Actually, the z-resolution of the SVD is so good that it is possible to measure B-meson lifetimes with it.
In 2003, the Belle detector was upgraded and the old SVD (often called SVD1) had to be replaced by a new one (SVD2), because of high radiation damage. While the SVD1 consisted out of 3 layers with a total polar angular coverage of 23° to 139°, the SVD2 consists out of 4 layers with radii from 20mm to 88mm and a total angular coverage from 17° to 150°.

Central Drift Chamber

The Central Drift Chamber (CDC) with a diameter of 1.2 meters and a polar angle coverage of 17° to 150° is responsible for the main track reconstruction. It is a classical gas filled drift chamber with 50 cylindrical ordered layers of wires around the SVD. Charged particles passing through the CDC produce free electrons by ionizing gas atoms, which are then accelerated in the electric field of the wires so that they themselves can ionize electrons and create an avalanche which drifts to a wire where it can be measured. The use of stereo wires with a small angle between them allows to measure the z-position more precisely by comparing the signal yield in neighboring wires. Because the whole CDC is also in the overall magnetic field of the solenoid, one can determine the charge and the momentum out of the curvature of a track. A mixture of 50% helium and 50% ethane is used to avoid spontaneous gas discharge and reduce coulomb scattering. The specific energy deposition ($\frac{dE}{dx}$) of tracks, which is useful for particle identification, can also be measured.

Aerogel Cherenkov Counter

The Aerogel Cherenkov Counter (ACC) is placed adjacent to the CDC in barrel[4] and in forward region. When a charged particle traverses through a material with a velocity greater than the speed of light in the particular medium, Cherenkov light is produced. Cherenkov light is emitted in a cone along the track of the traversing particle. In the CDC the momentum of charged particles is already measured. But knowing the momentum is not enough. If one wants to decide if a certain particle will emit Cherenkov radiation, knowing the rest mass is necessary. This can be exploited to identify particles. For instance, a Pion starts radiating Cherenkov light with a momentum of $1.2 GeV/c$, whereas a Kaon needs $3.5 GeV/c$[5] momentum to reach the threshold velocity. Therefore, in the momentum region between 1.2 and $3.5 GeV/c$, the ACC can distinguish very well between Kaons and pions.

Time of Flight

To identify low momentum tracks, the Time of Flight detector was installed outside of the ACC. The TOF uses plastic scintiliation counters to count traversing particles. Because the TOF has a comparatively low dead time, it is used to trigger the CDC or the calorimeter. It is also used to determine the flight time of particles reaching it. This time together with the flight length and momentum, which can be calculated out of the track, is a good criterion to identify particles. With a time resolution of about 100ps, the TOF is only effective in the low momentum region ($p_t < 1.2\frac{GeV}{c^2}$) and therefore complementary to the ACC.

Electronic Calorimeters ECL and EFC

So far, all components of the detector are only sensitive to charged particles. Because we need to measure also the neutral photons, the calorimeters which are sensitive to photons are a crucial part of the detector. The calorimeter is based on the principle that all particles[6] passing through matter loose energy by interacting with it. Highly energetic photons in the vicinity, of a nucleus can make e^+e^- pair production, while charged particles, like electrons, emit photons as bremsstrahlung. The photons coming from bremsstrahlung can themselves create e^+e^- pairs which then can emit bremsstrahlung. This leads to cascades - the so-called electromagnetic showers. The size of the calorimeter has to be adapted, so that in principle particles loose all their energy within the calorimeter. Since muons emit only a very small amount of bremsstrahlung due to their high mass, they are not absorbed in the calorimeter and must therefore be identified in the muon system.

Because photons can occur in every direction, two calorimeters are used at Belle to measure them. The main Calorimeter (ECL) is built out of 8736 CsI(Tl) crystals and covers a polar angular region of 12.4˙ to 155.1˙. To distinguish two photons in close

---

[4]The barrel region is situated around the beam pipe at the interaction point.
[5]In this specific detector material.
[6]Except neutrinos, which interact so little that the calorimeter cannot measure them.

vicinity a high spatial resolution is necessary. Therefore, a huge amount of 8736 crystal cells is used. To cover a greater region, the Extreme Forward Calorimeter (EFC) is built directly upon the beam pipe in forward and backward region in addition to the ECL. Due to obvious (see figure 2.5) geometrical reasons, not the whole remaining area can be covered with the EFC, but it gives an additional angular coverage in forward and backward region. Because of its proximity to the beam pipe and hence higher exposure of radiation, a Bismuth Germanite Calorimeter is used.

Superconducting Magnet

To make momentum determination in the CDC reliable, a magnetic field of high magnitude is needed. At Belle, a superconducting solenoid of 4.4 meters length and 3.4 meters in diameter is used to create the needed magnetic field along the z-axis. A superconductor is a conductor with zero resistance, if it is operating below a characteristic temperature T_C. It has to be cooled to operate properly, which is done with liquid helium. It takes approximately 6 days to cool down the magnet to its operating temperature. Furthermore it takes 30 minutes to charge the magnet so that it reaches its working field strength of 1.5 Tesla. If fully charged, the magnet is cooled below the critical temperature so that it becomes superconducting and the field is sustained without further energy supply. The superconductor used in the Belle detector is made out of NbTi/Cu composition.

KLM

To identify muons and K_L^0, another detection system is needed. Muons, for instance, interact very weakly with the detector and are not absorbed in the calorimeter. Therefore, the calorimeter information for muons, if there is any available, cannot be used. It is important to know if a particle was a muon or a K_L^0, which can be decided with the KLM. Because no other particles than muons and K_L^0 mesons (and maybe neutrons) should be able to traverse through the calorimeter, they can be identified simply by detecting them after they left the calorimeter. To fulfill this task, an alternating system of iron absorbers and glass-resistive plate counters is used. The iron absorbers of 4.7cm thickness are necessary to increase the interaction length[7]. They are also used as iron yoke for the superconducting magnet. The glass-resistive plate counters can measure the ionizing track of muons and the decay products of the K_L^0.

2.1.3. Data and Monte Carlo at Belle

From 2000 to 2010, with a longer interruption during the SVD upgrade, data was taken at the Belle experiment (figure 2.6). The integrated luminosity of data taken in the interesting region for this analysis, the Y(4S)-resonance, is 711fb^{-1}. This corresponds to 771.6 million $B\bar{B}$-pairs. The sample itself can be splitted into two parts: The first

---

[7]The interaction length is the mean path, for a particle traveling through matter, to loose $\frac{1}{e}$th of its energy in interaction with matter.
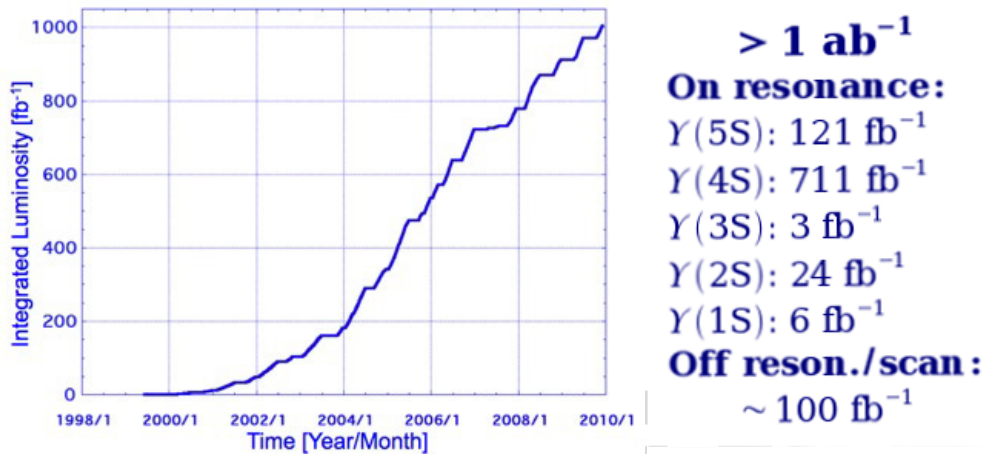
Figure 2.6.: Integrated luminosity [16]

part recorded with the original silicon vertex detector (SVD1) and the one with the new silicon vertex detector (SVD2). The track reconstruction is different for these two data sets, which leads to different reconstruction efficiencies.

For a better understanding of background sources, simulated data, called Monte Carlo, is useful. To create Monte Carlo, two steps are necessary. At first, a random generator, which creates all particles event by event, from originating particles to stable end-products, and their four-momenta, according to an input distribution, which is given as a so called decay table. A decay table (an example is given in Appendix A.5) lists the possible originating particles and their branching ratio as well as their possible daughter-particles and the branching ratios of them, up to the final particles. At this point, all simulated particles are only list entries with an exact four-momentum, which is not like the measured ones. Therefore, in the second step a detector simulation program is used to simulate detector responses, like tracks in CDC and hits in the calorimeter, for the input particles.

Based on the input decay table, one can distinguish between generic Monte Carlo and signal Monte Carlo. Generic Monte Carlo has all possible decays and known branching ratios in the decay table, so it should provide a complete image of nature, if one assumes that previous measurements that are used as input are correct. At the Belle collaboration, generic Monte Carlo is produced centrally. The luminosity of the generic Monte Carlo is scaled to data to make it comparable. The amount of generic Monte Carlo that is equivalent to data is called a stream. Currently six complete streams of generic Monte Carlo are available.

Signal Monte Carlo in contrast describes *only* certain decay chains, namely signal and can be created with high statistics to study even rare decays. Not only measured but also predicted decays, which are not included in generic MC, can be simulated.

3. Reconstruction and Selection Optimization

3.1. Reconstruction

3.1.1. Decay Channels

The Belle experiment ran predominantly at the Y(4S) resonance, where two B-mesons are produced as $B\bar{B}$-pair. In this analysis, the decays $B^0 \to D^{*\pm}D^{\mp}$ and $\bar{B}^0 \to D^{*\pm}D^{\mp}$ are reconstructed to study the branching ratio of them. Thereby, only hadronic decays of the D-mesons are studied. In figure 3.1, all decay channels used in this analysis are shown. The primary D-meson coming directly from the B-meson decay is reconstructed in two final states. In this analysis, it will be referred to as D_1. The $D^{*\pm}$ can either decay in a neutral or a charged D-meson and the corresponding pion. Five modes of neutral and one of charged D-mesons coming from D^* decays are reconstructed. Other channels which were studied were dismissed for different reasons. Channels, where a D-meson coming from a B-meson goes into a final state with a π^0, like $B^0 \to D^{*-}D^+(K^-\pi^+\pi^+\pi^0)$, are not used due to their high background level. Channels with a π^0 in them, could only be handled if the D was coming from a D^*, since then a D^*D-mass-difference cut could be applied, which is powerful to discriminate between signal and background.
The channel $B^0 \to D^{*-}D^+(\to K^-K^+\pi^+)$ was promising on Signal Monte Carlo, since it was quite clean. But on generic Monte Carlo, it was realized that the contribution from non-resonant decays $B^0 \to D^{*-}K^-K^+\pi^+$ is higher than the contribution of $B^0 \to D^{*-}D^+(\to K^-K^+\pi^+)$, which is why it was also dismissed.

3.1.2. General Principle of Reconstruction

B-mesons themselves cannot be seen in the detector, as they decay before they leave the beam pipe. The particles which can be seen in the detector are the relatively long living charged kaons, pions, protons, electrons, muons and photons. Although not all of these particles are stable, they are called "final state particles" (FSP). At Belle, short living neutral kaons, called K_S^0, are reconstructed from two charged pions. Neutral pions are reconstructed from two photons. These two neutral mesons are also called final state particles, although they are not in a narrower sense.[1]
Charged and neutral kaons and pions are combined to charged and neutral D-meson

---

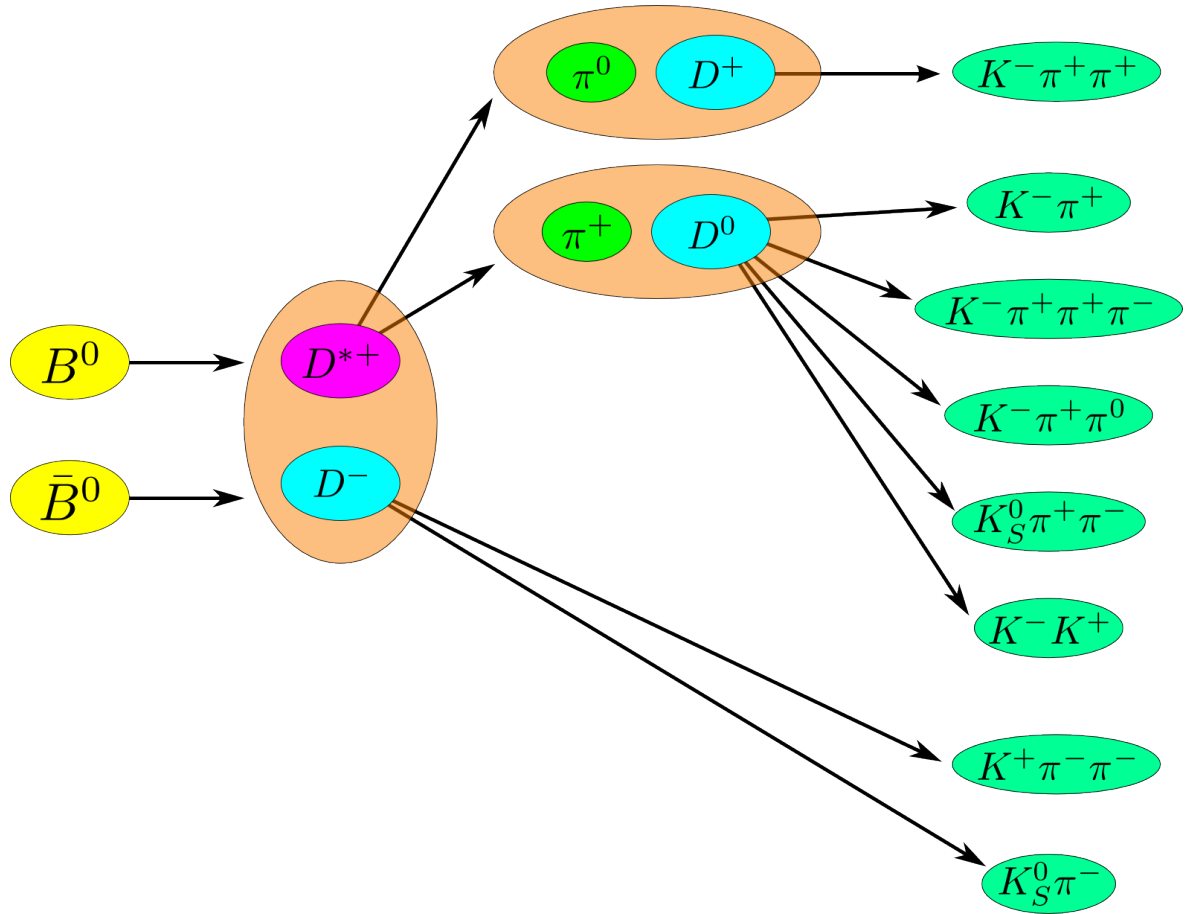[1]The π^0 decays into to photons. The K_S^0 decays almost only into to pions.

Figure 3.1.: The reconstructed decay channels of B^0 and $\bar{B}^0$ into $D^{*+}D^-$ are shown in the diagram. The charged conjugated decays $D^{*-}D^+$ are also reconstructed, but not shown here.
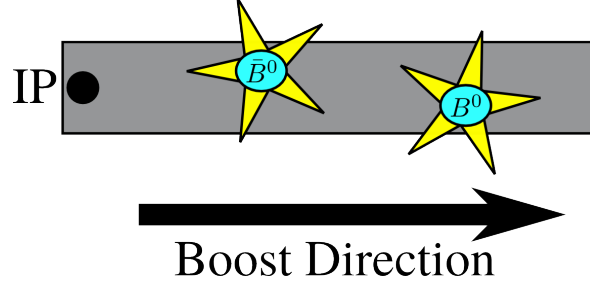
Figure 3.2.: Schematic figure of the IP-t4ube

candidates. The energy of charged particles is calculated with a mass hypothesis and the measured momentum using $E^2 = p^2 + m^2$. Therefore, the energy resolution of a particle can be improved with a more accurate track. To achieve this a Chi-square vertex fit of the FSP to a common origin is performed for each D-meson candidate. Only those D-meson candidates are taken, where no error in the fit occurs. The fit can change the track parameters of the FSP, which results in slightly different momenta that are usually more precise. The four-momentum of the D-mesons, which was originally calculated out of the measured four-momenta of its assumed daughter particles before the fit, is recalculated taken the momenta of the fitted particles into account. To form $D^{*\pm}$-mesons, neutral (charged) D-mesons are combined with charged (neutral) pions. These pions gain only low momentum in the lab frame since most of the momentum of the D^* is carried by the D-meson. These "slow" pions, especially the neutral ones, are hard to measure and therefore not used to determine the B-meson decay vertex. This vertex is obtained by fitting both (non-star) D-mesons to the IP-tube. The IP-tube is a cylinder from the interaction point (IP) along the boost direction (schematically shown in figure 3.2). The use of the IP-tube in the vertex fit takes into account the low momentum of the B-mesons in x and y direction and the boost in z-direction. Due to the fact that the two D-mesons do not carry the whole B-momentum, the obtained fit results cannot be used to recalculate the B-momentum. D^*-mesons decay via strong interactions, thus they decay at the B-meson vertex where they are produced. Therefore, the slow pions are fitted to the B-meson vertex, obtained by the fit to the IP-tube. The D^*-momentum is recalculated using the information of the changed slow pion information due to the fit before.

3.1.3. Reconstruction Criteria

Particle Identification

Data files (and also Monte Carlo files) contain different lists for different types of particles, which are filled during processing (or in case of MC by the detector simulation). There is a list of all K_S^0, a list of all photons, a list of π^0 mesons and a list of all charged particles. Charged pions and kaons can be found in the charged particle list, but also other particles like protons, electrons and muons are in the charged particle list. It is

easy to separate muons from charged pions and kaons, because only muons can pass the calorimeter. To distinguish kaons and pions from electrons, a electron veto is applied. For the separation of kaons and pions the so-called ATC[2] particle identification is used. The ATC particle identification is recommended to be used in 10 bins from zero to one:

$$\mathcal{L}_{\pi/K}^{ATC} = \frac{\mathcal{L}_{\pi}^{ATC}}{\mathcal{L}_{\pi}^{ATC} + \mathcal{L}_{K}^{ATC}}$$

It is the product of the three binned likelihood functions, which are calculated from the information of the detector system:

$$\mathcal{L}^{ATC} = \mathcal{L}^{ACC} \cdot \mathcal{L}^{TOF} \cdot \mathcal{L}^{CDC}.$$

It can happen that not every likelihood function $\mathcal{L}^{i}$ [3] is defined for every particle, since not all particles are registered in every detector part. If this happens the combined function $\mathcal{L}^{ATC}$ is set to certain values, depending on the other likelihood functions. Therefore, the ATC pid is not a likelihood function in the sense that it returns a probability, but its bins, in steps of 0.1 from zero to one, are ordered with increasing probability. One exception is the value 0.5, which should not be used as cut value, since particles with no proper ATC information are set to this value. It states that $\mathcal{L}_{\pi/K}^{ATC} = 1 - \mathcal{L}_{K/\pi}^{ATC}$ per bin. With increasing value of $\mathcal{L}_{\pi/K}^{ATC}$, it becomes more likely that a candidate is a pion.

In this analysis, loose standard ATC particle identification cuts for kaons $\mathcal{L}_{\pi/K}^{ATC} < 0.9$ and pions $\mathcal{L}_{\pi/K}^{ATC} > 0.1$ are used. This rejects a lot of background without sacrificing too much signal.

Neutral Meson Identification

At Belle, neutral short-living kaons K_S^0 are reconstructed only via their decay into a $\pi^+\pi^-$-pair, which occurs in about 69.2% of the decays according to PDG [8]. From all opposite charged pion candidates coming from a common vertex K_S^0-candidates are formed and put in a K_S^0-list. To identify K_S^0 mesons, the standard method used by Belle is the so-called *good-K_S^0-selection*. It is also used in this analysis. The *good-K_S^0-selection* is based on cuts applied on four K_S^0 variables [17]:

- The flight length fl of the K_S^0 in the x-y-plane

- The z-distance dz of the two daughter-pions at the interception point.

- The azimuthal angle $d\phi$ between the momentum and the decay vertex vector of the K_S^0

---

[2]The letters stand for different detector parts: A stands for Aerogel Cherenkov Counter, T for Time of Flight and C for Central Drift Chamber

[3]with i = ACC or TOF or CDC

- The minimal difference dr between interaction point and x-y plane of the two daughter tracks.

The cuts are applied in three bins of the K_S^0 momentum, which are shown in table 3.1. Neutral pions are reconstructed out of two photons. To suppress obvious π^0 back-

Table 3.1.: good-K_S^0-selction

Momentum K_S^0 GeV	dr (cm)	$d\phi$ (rad.)	dz (cm)	fl (cm)	ϵ on MC	ϵ on data
< 0.5	> 0.05	< 0.3	< 0.8	-	$62.0 \pm 2.7\%$	$58.3 \pm 3.8\%$
0.5–1.5	> 0.03	< 0.1	< 1.8	> 0.08	$79.0 \pm 1.0\%$	$75.3 \pm 1.2\%$
> 1.5	> 0.02	< 0.03	< 2.4	> 0.22	$83.5 \pm 1.2\%$	$83.5 \pm 1.4\%$

ground, some quality cuts were made during reconstruction . It is required that the minimal energy of each photon is larger than 30 MeV. Furthermore, the momentum of the π^0 is required to be greater than 50 MeV and the invariant mass of the π^0 has to lie in a 50 MeV/c^2 window around the nominal π^0 mass.

Summary of Reconstruction Cuts

There are other cuts made during reconstruction to keep the sample at a hand-able size. The impact parameter of a charged track is the minimal distance of the track with the interaction point. It is required that the minimal perpendicular distance dr is less than 2 cm and the minimal distance in z-direction is less than 4 cm for all charged kaons and pions. These cuts are made to ensure that the reconstructed charged tracks originate from the interaction point. All cuts made during reconstruction are summarized in table 3.2.

Table 3.2.: skim cuts

variable	cut window
$|dr|$	<2cm
$|dz|$	<4cm
$\mathcal{L}(\pi^\pm/K^\pm)$	> 0.1
$\mathcal{L}(K^\pm/\pi^\pm)$	> 0.1
D-mass	$m_{PDG} \pm 70$ MeV/c^2
$m_{D^*} - m_D$	$130 - 160$ MeV/c^2
π^0 mass	$130 - 160$ MeV/c^2
M_{bc}	5.2-5.3 GeV/c^2
ΔE	± 200 MeV

3.1.4. Skimming

After termination of data taking in June 2010, Belle had recorded 711fb^{-1} of data at the Y(4S) resonance. This corresponds to a data size of about 100TB. It is not useful to run over the whole data sample every time one has changed something in their reconstruction. Therefore, a subset of the whole dataset, the skim, is created. The requirements of the skim are minimal data size without any efficiency loss. Many computations that are done in the reconstruction program, like vertex fits for example, are not necessary for a skim and are only slowing it down. Therefore, a slim program, which does in principle only the combination of FSP, is satisfactory. Every event is taken where at least one B-candidate, of the ones that are reconstructed, can be formed. Only very loose mass-cuts several sigma away from the nominal value are performed. The skim for this analysis reduces the data size to about 0.7% of the original size. The cut windows used in the skim are listed in table 3.3.

Table 3.3.: skim cuts

variable	cut window
π^0 mass	$130 - 160$ MeV
D-mass	$m_{PDG} \pm 100$ MeV
$m_{D^*} - m_D$	$130 - 160$ MeV
M_{bc}	5.2-5.3 GeV
ΔE	± 200 MeV

3.2. Optimizaton of Selection

3.2.1. General Idea of Signal Extraction

The separation between signal and background is a typical classification problem. One wants to decide to which class, signal or background, a certain event belongs. An event is described by several parameters λ_i, the measured physical observables, which can be expressed as vector of the n-dimensional parameter space.

$$\vec{\Lambda} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ ... \\ \lambda_n \end{pmatrix}$$

On real data one can only decide according to these parameters whether an event is signal or background. Only on Monte Carlo the "truth" is known for every event. Therefore, it is important to learn more about the conditional probability $P(S|\vec{\Lambda})$ of a signal event, with a given parameter set $\vec{\Lambda}$ on MC to apply this knowledge later on

on data. A simple example with two arbitrary parameters x and y is shown in figure 3.3. The red dots symbolize signal events distributed in the x-y parameter space. Background can be suppressed by cuts on the parameters x and y. Two quantities which describe the performance of the selection are purity and efficiency. Efficiency is defined as the number of selected signal events divided by the number of all signal events in the sample:

$$\epsilon = \frac{S_{sel}}{S_{all}}.$$

The marked cuts do not cut away any signal events, and therefore conserve the efficiency at 100 percent. Purity is defined as the fraction of signal events in the selection:

$$P = \frac{S_{sel}}{S_{sel} + B_{sel}}.$$

With more narrow cuts on the parameters x and y, a higher purity could be achieved. But it would also result in lower efficiency. A better result could be achieved, if the correlation between the parameters x and y would be exploited. With a cut-based approach this is not possible. Some multivariate analysis tools, like NeuroBayes$^{®}$, an analysis tool based on a neural network, can handle correlations and thereby improve their abilities of classification. To decide which analysis method should be used, it is essential to compare the gain in separation power with the difficulties the more sophisticated technique introduces. A gain in efficiency by a few percent, with same level of purity, could be nullified if the technique would increase systematic uncertainties, for example. If the variables used for separation are not correlated, a cut based approach may be completely sufficient. But the more complex the problems get, the more sophisticated methods and tools have to be used.

The next thing to do is to find a figure of merit, which can be used to optimize the parameters in respect to the given problem.

3.2.2. Best Candidate Selection

During reconstruction, final state particles are combined to form D-mesons and onward B-meson candidates. In an event, a certain set of final state particles is observed. This set is used to combine the intermediate state D-mesons, whereby multiple combinations are possible. Multiple candidates occur, for example, if particles are mixed up during combination. This can be a problem with π^0 mesons, because low energetic γ coming from π^0 are hard to distinguish from background photons. But also events where only correctly reconstructed particles are combined can have multiple candidates. In figure 3.4 a schematic example is shown. Assuming that all final state particles are correctly reconstructed, there are still two possible D-meson combinations possible. In the left picture, kaons and pions are combined to form a D^+ and a D^-. With the additional π^0 and the D^-, a D^{*-} is created. In the right picture, the same final state particles are used to form a D^+ and a $\bar{D}^0$, where the additional π^- is used to form a D^{*-} with the $\bar{D}^0$. Both combinations are possible, but only one is correct.

To not artificially increase the statistics with multiple combinations for a single event,
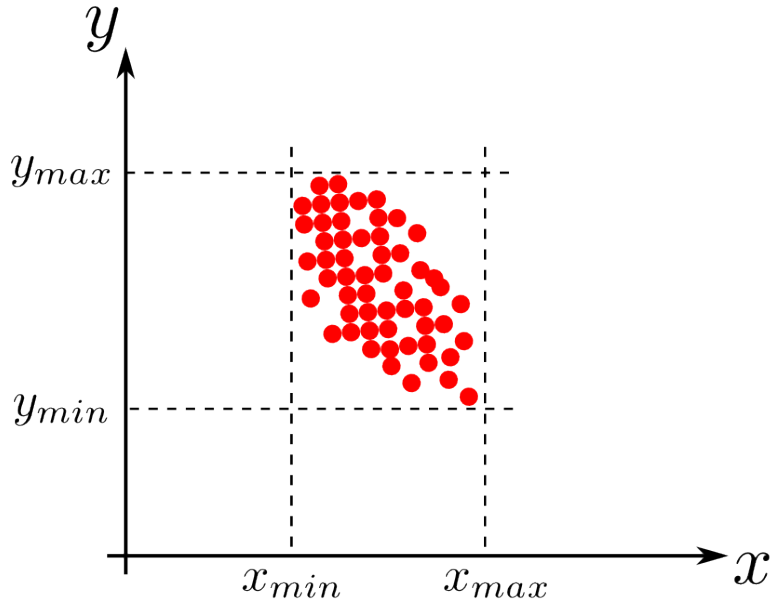
Figure 3.3.: Schematic picture of cut based signal selection. The red dots symbolize the signal which is distributed according to arbitrary parameters x and y. The correlation between x and y can't be exploited using only cuts.
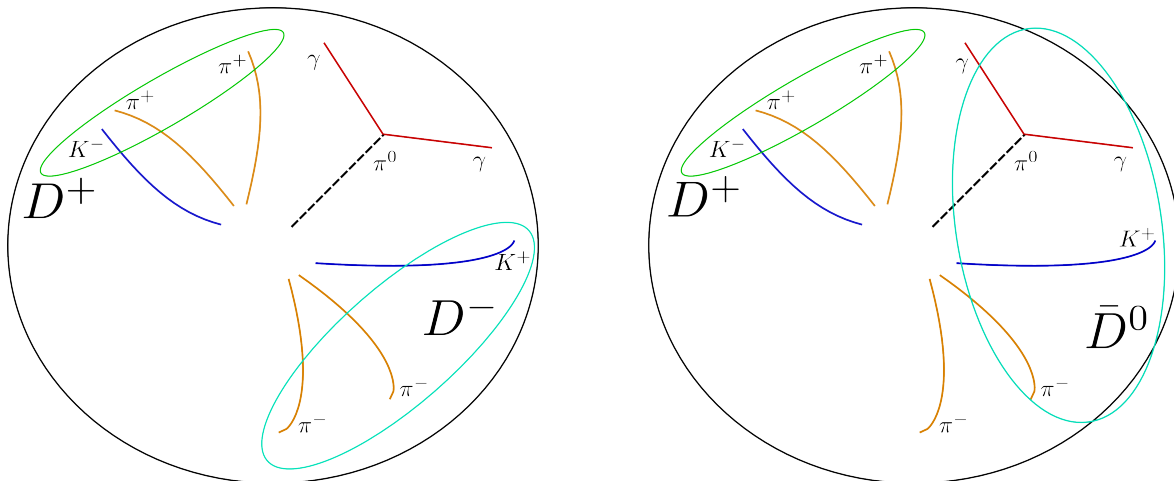

Figure 3.4.: Schematic event selection

one candidate has to be selected. Multiple candidates occur in 9.18% of the events on generic Monte Carlo after all optimized[4] cuts are applied. For each B-meson candidate, a chi-square value is calculated from both D-meson masses, m_{D_1} and m_{D_2}[5], and the D^*D mass-difference Δm_{D^*D}. The absolute value of the D-meson masses minus the nominal values obtained from PDG are divided by the width of the D-meson mass resolution, which is obtained by fitting, as explained below. The same holds for the mass-difference.

$$\chi^2 = \left| \frac{m_{D_1} - m_{D_{nom}}}{\sigma_{D_1}^{fit}} \right|^2 + \left| \frac{m_{D_2} - m_{D_{nom}}}{\sigma_{D_2}^{fit}} \right|^2 + \left| \frac{\Delta m_{D^*D} - \Delta m_{D^*D}^{nom}}{\sigma_{\Delta m}^{fit}} \right|^2$$

The candidate with the smallest χ^2 value is taken as best candidate of an event. In a signal event with multiple candidates, this selection finds the right candidate in 61 % of the cases on generic Monte Carlo.

The widths of the D-meson masses and the mass-difference are obtained by fitting signal on Monte Carlo. For each D-meson decay mode, a separate unbinned maximum likelihood fit is performed. The sum of 2 Gaussian functions[6] was used to fit the invariant mass distributions of D-meson candidates. In the case of a D-meson decaying to $K\pi\pi^0$, a bifurcated Gaussian plus a Gaussian was used instead, due to the asymmetric shape which comes from the γ-resolution. The D^*D mass-difference is fitted separately for $D^* \to D^+\pi^0$ and $D^* \to D^+\pi^+$ decays with the sum of three Gaussian functions. Figure A.1 shows the fits for the D-masses of the four D-modes $K^-\pi^+\pi^+$, $K_S^0\pi^+$, $K^-\pi^+$ and $K^-\pi^+\pi^0$ and both mass-difference fits. The fits for the other D-modes can be found in appendix A.2.

---

[4]The optimization of the cuts will be explained in the next section.

[5]Here m_{D_1} is the mass of the D-meson from the B decay and m_{D_2} is the mass of the D-meson from the D^* decay.

[6]For the description of a Gaussian function see appendix A.1.

(a) $D^+ \rightarrow K^- \pi^+ \pi^+$

(b) $D^+ \rightarrow K_S^0 \pi^+$

(c) $D^0 \rightarrow K^- \pi^+$

(d) $D^0 \rightarrow K^- \pi^+ \pi^0$

(e) $\Delta m = D^{*-} - D^0$

(f) $\Delta m = D^{*-} - D^-$
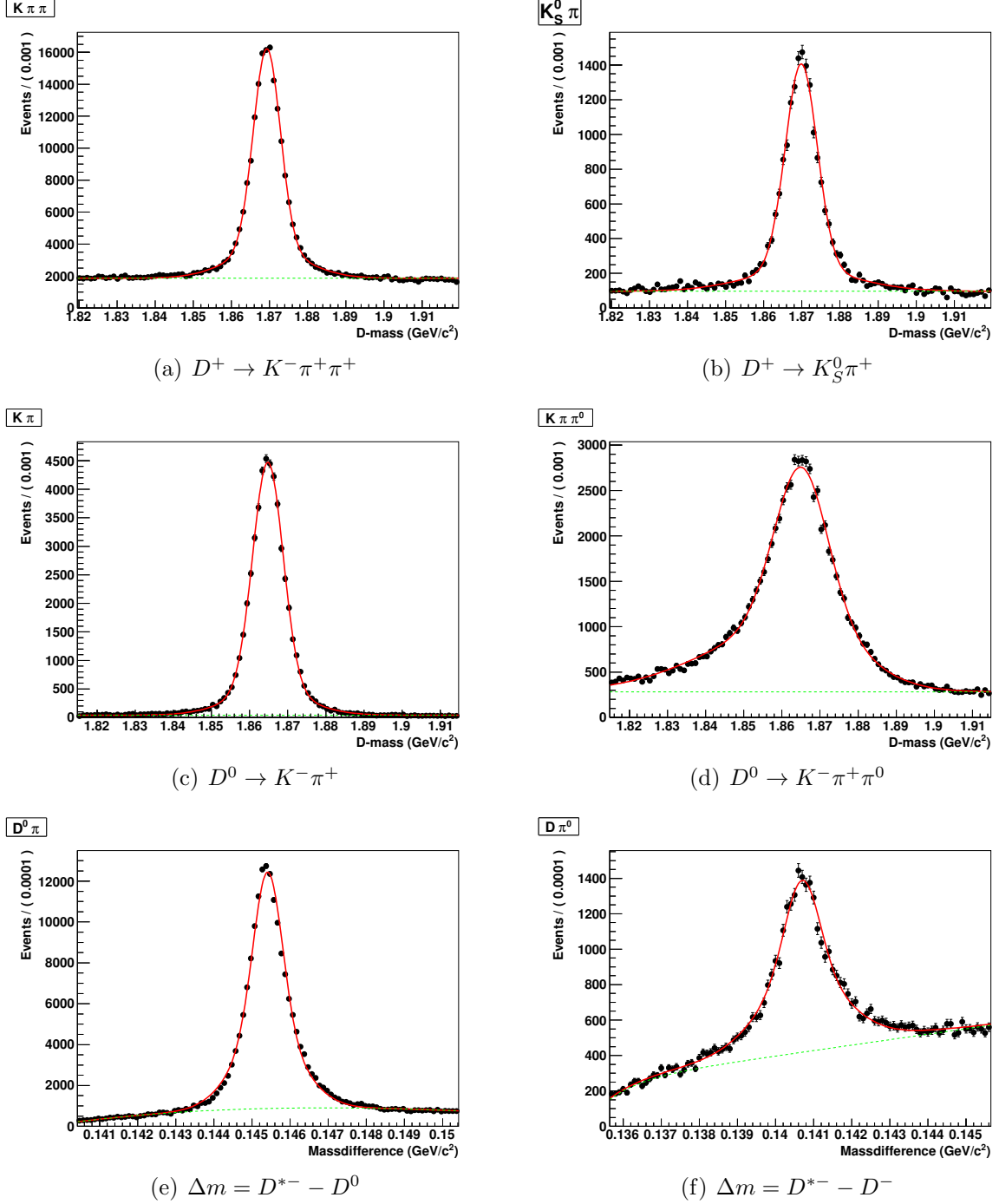
Figure 3.5.: Fit of D-Meson masses and $D^* D$ mass-difference on generic Monte Carlo

3.2.3. Multidimensional Significance Optimization

In the decay $B^0 \rightarrow D^{*\pm}D^{\mp}$, powerful variables to discriminate between signal and background exist. Therefore, it makes sense to perform a cut based selection. The question which discriminating variables can be used in the optimization is the first that arises. Two powerful variables to distinguish between signal and background are the beam constrained B-meson mass M_{bc} and the energy difference ΔE. The beam constrained mass is the mass of the B calculated from energy-momentum relation with the energy substituted by half of the beam energy:

$$= \sqrt{(\frac{E_{beam}}{2})^2 - p_{B^0}^2}$$

The energy difference is calculated between half of the beam energy and the measured energy of the B^0:

$$\Delta E = \frac{E_{beam}}{2} - E_{B^0}$$

For correctly reconstructed B-mesons, M_{bc} peaks at the invariant B-meson mass of 5.28 GeV/c^2 and ΔE peaks at zero. The beam energy is used because it can be measured much more precisely than the energy calculated from the measured energies of the daughters of the B-meson. No selection cuts can be performed on them, because they are used to extract the signal yield later on in the fit of the branching ratio measurement. In figure 3.6, a scatter plot of the the two variables is shown. The signal box in this analysis is defined as: $|\Delta E| < 30$MeV and $M_{bc} > 5.27$ GeV/c^2

In the optimization, particle masses are used as separating variables. For a better handling not the measured mass values, but the absolute of the difference between the measured mass and the nominal mass, obtained from PDG [8], are taken. The variables used for signal separation are:

- The mass of the D-meson coming from the B-meson decay

- The mass of the D-meson coming from the D^*-meson decay

- The difference between the D^*-mass and its D-meson daughter

- The mass of the neutral pion

- The minimal energy of the photons coming from π^0 decays.

For a branching ratio measurement high efficiency and high purity are important. A figure of merit with a reasonable ratio between efficiency and purity is the statistical significance. The significance Σ_{signi} is defined as the number of signal events divided by the square root of all events in the signal box. The square root corresponds to the statistical error of the sample with the assumption of an underlying Poisson-distribution:

$$\Sigma_{signi} = \frac{N_{sig}}{\sqrt{N_{sig} + N_{bkg}}}$$
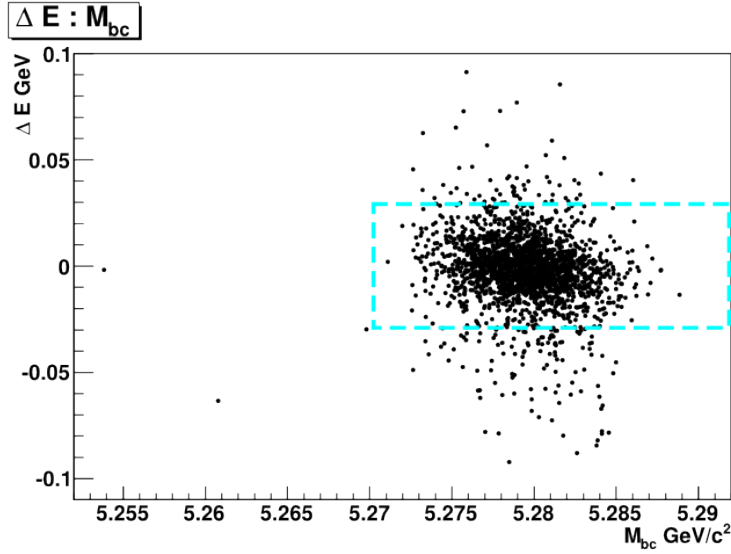
Figure 3.6.: The scatter plot shows the distribution of signal events in ΔE and M_{bc}. The blue square marks the signal area: $|\Delta E| < 30$MeV and $M_{bc} > 5.27$ GeV/c^2

In principle, one can optimize each separating variable in respect to the significance on its own. But finding the local maximum in every single variable does not automatically yield the global maximum of significance. To find the global maximum of the significance, one has to scan over each variable simultaneously. This adds an extra dimension for every variable, which is optimized, increasing the computing time exponentially. I used a program, which was written by Markus Röhrken to perform the multidimensional scan on the significance. Because it loads the whole data sample, which is optimize,, into memory, it is quite effective. The user enters the variables for which the cut range and the step size of the cuts is specified into the program. An example could be a cut on the D-meson mass in a window from 5 to 20MeV/c^2 around the nominal value in 15 steps of 1MeV/c^2. The program then calculates the significance for every point in parameter space, and returns the cut values with which the highest significance was achieved. In the optimization process, a big enough data sample is crucial to avoid training on statistical fluctuations. To account for this, an enriched sample was used for optimization. The enriched sample contains 22 times the amount of signal and two times the amount of background compared to the expected yield on data. The enrichment has to be taken into account during optimization by dividing the number of signal events and background events respectively through the enrichment factor (22 for signal and 2 for background). Thereby, it is important to make sure that the enrichment is the same for every decay channel, because every channel is optimized separately.

In the following, I will describe how the process of optimization was done. Since the separation is most important in the signal region, only events within the M_{bc} ΔE signal-

box defined above are used for optimization. Furthermore, the decay channels to be optimized have to be specified. This allows to treat different decay channels separately, which is often useful, since the resolution and signal to background ratio can differ greatly between different channels. Since the width of the D-meson mass is different for different final states, the optimization process is divided into sub-processes for every D-meson channel. It is physically plausible that the cuts on the masses of D-mesons coming directly from the B decay, in the following called *primary* D-mesons, should be the same, no matter how the D^* decays.

To achieve this, all channels are optimized separately at first with all variables mentioned above. The cut values for all variables except the mass from primary D-mesons are applied to the optimization sample to create a new sub-sample. The sub-sample is used to optimize the mass of primary D-mesons only for all D^*-decays together. In figure 3.7, the significance over the cut on the reconstructed D-mass minus the nominal D-mass is plotted. The plots are projections into the D-meson mass region, while all other cuts are set to the values at the point of the highest global significance. In both cases, $D \rightarrow K\pi\pi$ and $D \rightarrow K^0_S\pi$, the D-meson mass cut which yields the highest significance is a $9\text{MeV}/c^2$ cut window around the nominal value.

With the obtained cut values for primary D-mesons, a different sub-sample is created from the original sample, which is used for further optimization of the remaining variables. D-mesons coming from D^* decays are called *secondary* D-mesons in this thesis. The remaining variables, D^*D-mass-difference, mass of the second D-meson, mass of the π^0 and minimal photon-energy for π^0 daughters, are optimized for different *secondary* D-meson channels.[7] Because of the very low statistics of the Cabbibo-suppressed[8] decay $D^0 \rightarrow K^-K^+$, this channel is optimized together with the kinematic similar decay of $D^0 \rightarrow K^-\pi^+$. Other secondary D-meson channels are: $D^0 \rightarrow K^0_S\pi^+\pi^-$, $D^0 \rightarrow K^-\pi^+\pi^0$, $D^0 \rightarrow K^-\pi^+\pi^-\pi^+$ and $D^- \rightarrow K^+\pi^-\pi^-$. Exemplarily the results of the optimization of the channel $D^0 \rightarrow K^-\pi^+\pi^0$ are shown in figure 3.8. In plot (a) in the upper left, the significance is plotted over the cut on the D^0-mass. The significance has a plateau for a cut between $20\text{MeV}/c^2$ and $28\text{MeV}/c^2$, which is due to the fact that the D^*D-mass-difference cut already reduces much of the background. For the D^*D-mass-difference, plot (b), the maximum of the significance is clearly visible at $1.75\text{MeV}/c^2$. Also, in plot (c) of the π^0-mass and plot (d) of the minimal photon energy, the maximum of the significance is clearly visible. The results for all channels are summarized in table 3.4. The figures for the optimization of remaining channels can be found in appendix A.3. The optimization improved the purity in the signal box from 2.4% after reconstruction to 60.3% after optimization.

---

[7]A good overview of the channels is shown in figure 3.1

[8]Transitions from quarks of the second to the first generation are suppressed by the sine of the Cabbibo angle α. $sin(\alpha) \approx 0.22$
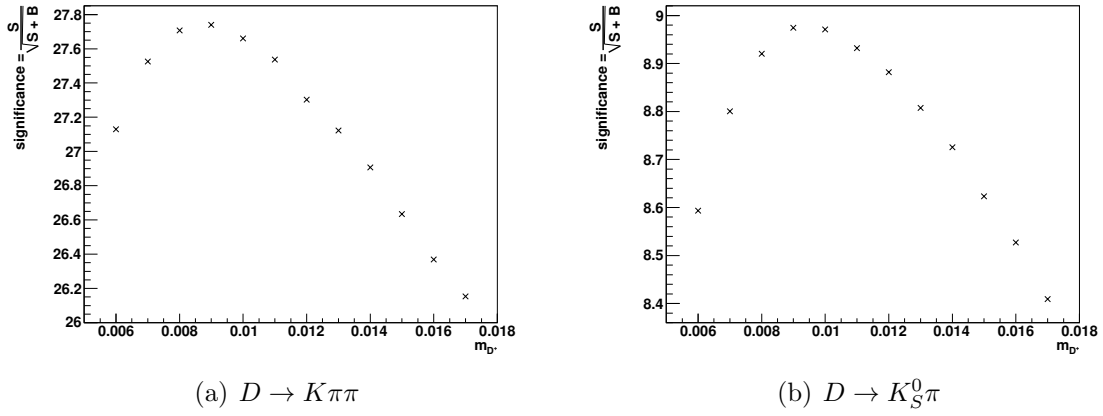
(a) $D \to K\pi\pi$

(b) $D \to K_S^0\pi$

Figure 3.7.: The plots show the significance plotted over the cut on the D-mass for the channels: $D \to K\pi\pi$ and $D \to K_S^0\pi$.

(a) Mass of D^0

(b) Mass difference Δm_{D^*D}
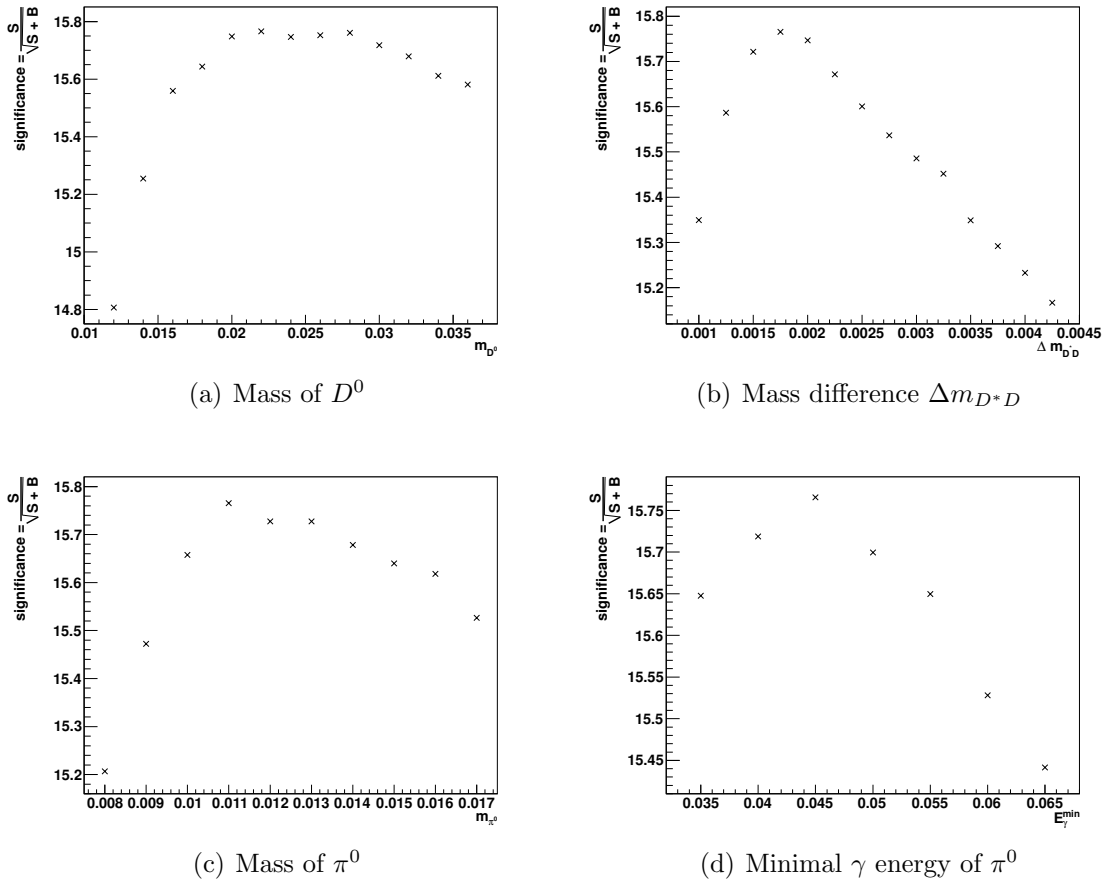
(c) Mass of π^0

(d) Minimal γ energy of π^0

Figure 3.8.: The plots of optimization are shown for channel $D^0 \to K^-\pi^+\pi^0$. The plots show the significance plotted over the cut of a certain variable, while all other variables are cut a their point of highest significance.

	Channel	$\|\Delta m_D\| <$	$\|\Delta\left(m_{D^*} - m_D\right)\| <$	$\|\Delta m_{\pi^0}\| <$	$E_\gamma^{min} >$
D^{*+}	$(K^-\pi^+/K^-K^+)\pi^+$	22 MeV/c^2	3.25 MeV/c^2	-	-
	$(K^-\pi^+\pi^-\pi^+)\pi^+$	13 MeV/c^2	1.75 MeV/c^2	-	-
	$(K^-\pi^+\pi^0)\pi^+$	22 MeV/c^2	1.75 MeV/c^2	11 MeV/c^2	45 MeV
	$(K_S^0\pi^-\pi^+)\pi^+$	11 MeV/c^2	3.0 MeV/c^2	-	-
	$(K^-\pi^+\pi^+)\pi^0$	10 MeV/c^2	1.75 MeV/c^2	13 MeV/c^2	35 MeV
D^-	$(K^+\pi^-\pi^-)$	9 MeV/c^2	-	-	-
	$(K_S^0\pi^-)$	9 MeV/c^2	-	-	-

Table 3.4.: Results of optimization - The cut values on the two *primary* D-meson masses apply for all D^* combinations. Thereby, the expression $\|\Delta m_X\|$ stands for the absolute value of the difference between the measured and the nominal value of the mass of particle X

4. Fit Procedure and Test on Monte Carlo

4.0.1. Branching Ratio

Experimentally, the branching ratio of a decay to a certain final state is determined by measuring the fraction of events into this specific final state (N_{x_1}) compared to all events (N_X).

$$\mathcal{BR}(X \to x_1) = \frac{N_{x_1}}{N_X}$$

At Belle, the number of produced $B\bar{B}$-pairs is well measured. To measure the absolute branching ratio of B^0 or $\bar{B}^0$ going to $D^{*\pm}D^{\mp}$, one needs to count all decays into this decay mode. Since not all sub-channels of D^{*-} or D^+ decays are reconstructed the branching fractions of these decays have to be taken into account as well.

Because not all produced events can finally be reconstructed, the reconstruction efficiency, which is defined as number of correctly reconstructed decays divided by the number of produced ones, is also an important factor.

In summary, the branching ratio is calculated as follows:

$$\mathcal{BR}(B^0 \to D^{*\pm}D^{\mp}) = \frac{N_{sig}}{\frac{1}{2} \cdot 2 \cdot N_{B\bar{B}} \cdot \mathcal{BR}(D^{*-} \to D\pi) \cdot \mathcal{BR}(D \to X) \cdot \epsilon}$$

The factor $\frac{1}{2}$ in the denominator comes from the fact that only 50 percent of the $B\bar{B}$-pairs are $B^0\bar{B}^0$ pairs, while the other 50 percent are B^+B^--pairs. Because a $B\bar{B}$-pair consists of B^0 and $\bar{B}^0$ and both decay in $D^{*\pm}D^{\mp}$, there is an additional factor 2 in the denominator. The branching ratio of $\mathcal{BR}\left(B^0 \to D^{*\pm}D^{\mp}\right)$ is the sum of the branching ratios $\mathcal{BR}\left(B^0 \to D^{*+}D^-\right)$ and $\mathcal{BR}\left(B^0 \to D^{*-}D^+\right)$[1]

4.1. Determination of Efficiencies

To determine the branching ratio of the reconstructed decay, the reconstruction efficiency is crucial. Efficiency is defined as number of reconstructed signal events, divided by all signal events in the data sample.

$$\epsilon = \frac{N_{rec}}{N_{all}}$$

---

[1]In principle the branching ratios $\mathcal{BR}\left(B^0 \to D^{*+}D^-\right)$ and $\mathcal{BR}\left(B^0 \to D^{*-}D^+\right)$ could be calculated seperately.

Due to that, the number of signal events in the data sample has to be determined. Of course this cannot be done on data, as I want to measure the branching ratio, which would be needed as input to estimate the number of signal events in the sample. On Monte Carlo, the number of all signal events can be counted exactly, in principle. It is more practical to use signal Monte Carlo instead of generic Monte Carlo, where one is only statistically limited by computing power. To determine the number of produced signal events I wrote a program called Monte Carlo counter.

4.1.1. Monte Carlo Counter

The purpose of the Monte Carlo counter is to run over simulated data files and count the decay channel dependent number of signal events. The Monte Carlo files contain a table, the GenHep-table, which lists all particles that were created, and all their children. For example, if one wants to find the following schematic decay, where $A \to B_1 B_2$ and $B_1 \to C_{B_1} D_{B_1}$ and $B_2 \to C_{B_2} D_{B_2}$, the program first searches in the GenHep-table for all occuring A particles. If an A is found, the Monte Carlo counter checks if the A has exactly two children and if they are B_1 and B_2. As soon as both daughters are found, it is checked for their children in the same manner. This is repeated up to the final state particles. To also take events with final state radiation into account, an additional photon is also allowed in each decay.

4.1.2. Efficiencies

In this analysis a signal Monte Carlo sample (for decay table see A.5) with approximately 500,000 signal events was used. The same reconstruction code as in the analysis with all the cuts and best candidate selection is used to reconstruct the signal MC sample. The number of signal events after reconstruction is determined using the information of the MC matching function. During reconstruction the MC matching function checks if a particle and all its children are reconstructed correctly. The use of MC matching information is necessary, since signal has to be distinguished from the *crossfeed* contribution[2]. The efficiencies are calculated for each decay channel separately, to be independent of the D- and D*-meson branching fractions. To determine the number of produced signal events per decay, the MC counter has been used. This yields to the efficiencies listed in table 4.1. As applying cuts can be seen as binomial process, the variance of the efficiency is given as:

$$V(\epsilon) = \frac{\epsilon(1 - \epsilon)}{N_{gen}}$$

---

[2]Crossfeed events are partly mis-reconstructed events with signal-like behaviour. They will be explained in section 4.2.1 in more detail.

Channel	Efficiency	Error
$(K\pi\pi)(K\pi)\pi$	18.30%	0.17%
$(K_S^0\pi)(K\pi)\pi$	11.43%	0.36%
$(K\pi\pi)(K\pi\pi\pi)\pi$	9.08%	0.09%
$(K_S^0\pi)(K\pi\pi\pi)\pi$	5.87%	0.18%
$(K\pi\pi)(K\pi\pi^0)\pi$	5.58%	0.05%
$(K_S^0\pi)(K\pi\pi^0)\pi$	3.56%	0.11%
$(K\pi\pi)(K_S^0\pi\pi)\pi$	6.16%	0.12%
$(K_S^0\pi)(K_S^0\pi\pi)\pi$	3.73%	0.25%
$(K\pi\pi)(KK)\pi$	17.28%	0.54%
$(K_S^0\pi)(KK)\pi$	9.41%	1.02%
$(K\pi\pi)(K\pi\pi)\pi^0$	5.37%	0.10%
$(K_S^0\pi)(K\pi\pi)\pi^0$	3.20%	0.19%

Table 4.1.: Efficiencies per Channel

with $\epsilon = \frac{N_{rec}}{N_{gen}}$ the standard deviation follows:

$$\sigma = \sqrt{\frac{N_{rec}(N_{gen} - N_{rec})}{N_{gen}^3}}$$

4.2. Unbinned Extended Maximum Likelihood Fit

To measure the branching ratio, one has to distinguish between signal and background events. An event, no matter if signal (S) or background (B), is characterized by a certain set of parameters $\vec{\Lambda}$, obtained as observables. The probability of a signal event S with a given parameter set $\vec{\Lambda}$ is described by the probability density function (*pdf*) $p(S, \vec{\Lambda})$ for signal. Analogically, the *pdf* for background $p(B, \vec{\Lambda})$ is defined. In general, a *pdf* is normalized to one for all parameters $\vec{\Lambda}$.

$$\int dX \cdot p(X, \vec{\Lambda}) = 1 \qquad \forall \vec{\Lambda}$$

In this example the random variable X has the realizations signal $X = S$ and background $X = B$. The *likelihood function* $\mathcal{L}$ is defined as the product of the associated *pdfs* of X. ΔE and M_{bc} [3], which were introduced in chapter 3.2.3, are used as variables in the fit, which distinguishes between signal and background. Therefore, the likelihood

---

[3]$\Delta E = \frac{E_{beam}}{2} - E_{B^0} \qquad M_{bc} = \sqrt{(\frac{E_{beam}}{2})^2 - p_{B^0}^2}$

function, which describes the data, is calculated from the *pdfs*, which describe signal and background in ΔE and M_{bc}.

$$\mathcal{L} = p(S, \Delta E) \cdot p(S, M_{bc}) + p(B, \Delta E) \cdot p(B, M_{bc})$$

For the branching ratio measurement the total number of signal events is necessary. Therefore, the likelihood function is *extended* by a term which takes the number of events N into account. Thus, it is exploited that the number of events N follows a Poisson distribution with mean η. The *extended likelihood function* $\mathcal{L}_{ext}$ is given in [18]:

$$\mathcal{L}_{ext} = \frac{1}{N!}\eta^N e^{-\eta} \cdot \mathcal{L}$$

Thus, η can be calculated from the *pdfs*:

$$\eta = \int dX \, N \cdot p(X, \vec{\Lambda})$$

The extended likelihood function is maximized using the *RooFit* package, which automatically takes care of the normalization. Internally, the negative log-likelihood function is minimized, which yields the same results, but is numerically more stable. The *pdf* of ΔE or M_{bc} is the sum from different *pdfs* each describing a part of the shape of these variables, like the shape in the signal-region. To properly describe ΔE and M_{bc} in the whole fit region background studies are necessary.
For further information about maximum likelihood fits in general, have a look at reference [19] or [20]. For specific material about *RooFit* see reference [18].
Two figures of ΔE and M_{bc} on generic Monte Carlo are shown in figure 4.1. They are plotted in the signal-region[4] . In M_{bc}, the signal peak as expected at the B-mass of $5.279\text{GeV}/c^2$ is well separated from the background. Also the signal peak in ΔE at zero is clearly visible. If only continuum background would exist, the shape of ΔE should be linearly distributed in the sidebands. But there are also bumps in the ΔE sideband visible: Two peaks in the negative sideband, and one very broad peak in the positive ΔE sideband, which is hardly visible. To correctly fit the data, all background sources have to be understood. Therefore, I will explain all important background sources in the following.

4.2.1. Signal and Signal-like Crossfeed

On signal Monte Carlo, all background sources mentioned above are not present. The only existing background should be of combinatorial nature: Thus, events where some final state particles were mixed up and the best candidate selection chose the wrong one. These events should be linearly distributed in ΔE and M_{bc} and therefore automatically

---

[4]Introduced in chapter 3.2.3 after all cuts and best candidate selection were applied. See figure 3.6
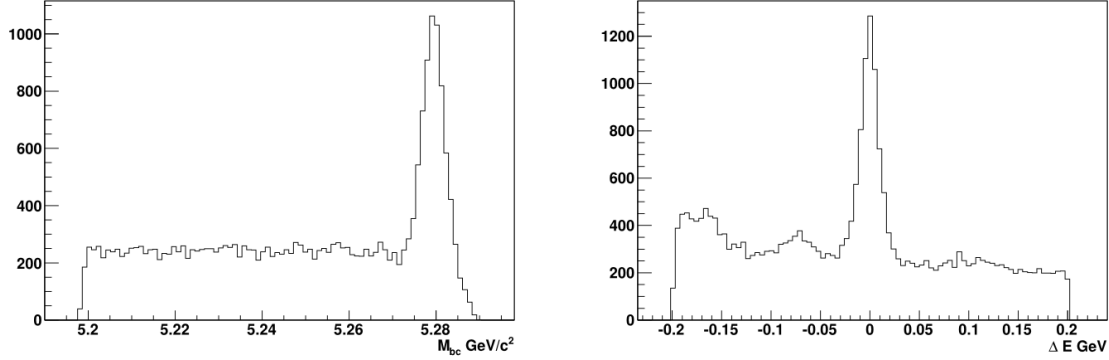
Figure 4.1.: M_{bc} with $|\Delta E| < 0.04 GeV$ and ΔE with $M_{bc} > 5.27 Gev/c^2$

be included in the background part of the fit. Due to this, signal MC is appropriate to determine the necessary parameters for the shape of the signal. But when investigated on a sample of signal Monte Carlo, where all selection cuts were applied, a peaking background contribution in ΔE and M_{bc} was found. This background source, which I will call *crossfeed*, cannot be separated from signal in the fit on data. Crossfeed events are partly wrongly reconstructed signal events: For example, both D-mesons are correctly reconstructed, but the slow pion is wrong. Or one photon of the two photons of a reconstructed π^0 is wrong.

Only a combined fit of signal and crossfeed is possible. To account for this, the ratio of signal to crossfeed has to be determined on signal MC and the crossfeed fraction has to be subtracted later on. To describe the shape of the crossfeed events, seperate fits of ΔE and M_{bc} on signal Monte Carlo have been performed. The main contribution of crossfeed were found to be in events with π^0 mesons. Because of these π^0s mesons, the crossfeed contribution has larger tails than the signal part. For the variable M_{bc}, crossfeed events are fitted with a Crystal Ball function.[5] The Crystal Ball parametrizes the shape of the crossfeed excellently in M_{bc}, which is shown in figure 4.2 (a). The signal in M_{bc} can described sufficiently with a Gaussian function. The two signal and crossfeed *pdfs* are summed up to a total *pdf*, whereby the fraction of the *pdfs* is fixed to the fraction of signal to crossfeed determined with MC-matching on signal MC. The fit for the combined function is shown in figure 4.2(b) as red line. The crossfeed contribution is indicated by the dashed blue line. The crossfeed events in ΔE also have a broader shape than the signal events. They are described with two Gaussian functions, with same mean but different widths, which is shown in 4.2 (c). The signal events in ΔE can also be described with two Gaussian functions. A combined fit of signal and crossfeed with fixed fraction is shown in 4.2 (d). The parameters of the functions describing the crossfeed are fixed in the overall fit later on. The fraction of signal to crossfeed is determined on signal Monte Carlo, and is found to be 95%:

$$\frac{N_{sig}}{N_{sig} + N_{cross}} = 0.95$$

---

[5]The Crystal Ball function is written down in the appendix A.1

(a) Crossfeed M_{bc}

(b) Signal + Crossfeed M_{bc}

(c) Crossfeed ΔE

(d) Signal + Crossfeed ΔE
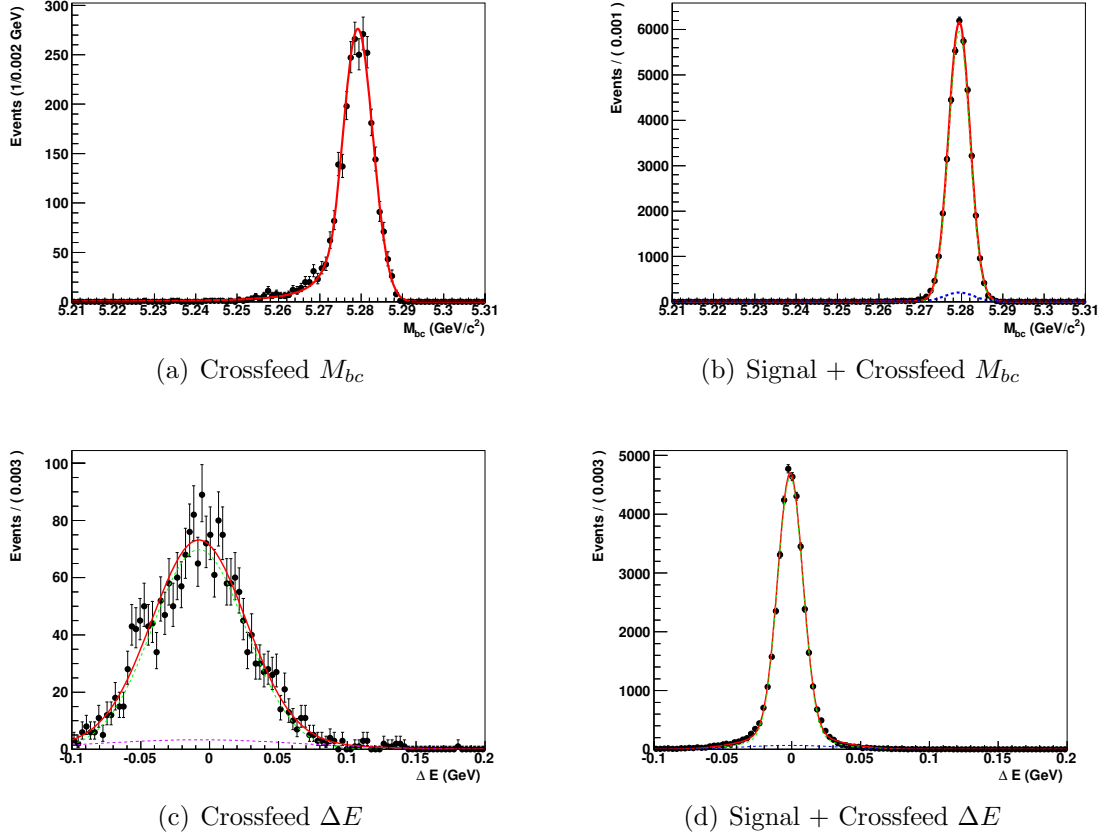
Figure 4.2.: Plot (a) shows the M_{bc} crossfeed distribution fitted with a Crystal Ball function. Plot (b) shows the signal (fitted with a Gaussian function) and the crossfeed (in blue) in M_{bc}. Plot (c) shows the ΔE crossfeed distribution fitted 2 Gaussian functions. Plot (d) shows the ΔE signal plus the crossfeed (in blue) in ΔE. Thereby the signal is also fitted with 2 Gaussian functions.

4.2.2. Misidentification

As explained in the section about particle identification (see 3.1.3), kaons and pions are both reconstructed from the charged particle list. As the separation of the particle identification is not perfect, a kaon can be identified as a pion or vice versa. The difference of the mass of kaons, with approximately 494 MeV/c^2 mass, and pions, with about 140 MeV/c^2 mass, leads to a difference in energy of these particles, because the energy is calculated from the momentum and the mass of the particle. With the wrong mass hypothesis, a wrong energy is assigned to the particle. This also effects the energy of the B-meson, which is calculated from its daughters' four-momenta. Therefore, these misidentified events are visible in ΔE. In contrast to ΔE, the mass hypothesis of the daughter particles of the B plays no role in M_{bc}, since M_{bc} is calculated only from

the momenta of the daughter particles and the beam energy. Therefore, misidentified events can appear in the M_{bc} signal peak. If massive particles are misidentified as lighter ones the energy difference is negative, if lighter ones are misidentified as heavier ones the energy difference is positive. This explains the peaks in the lower and upper ΔE sideband.

For a good fit of ΔE, it is satisfactory to cut away the negative tail and fit only the ΔE region from minus 100 to plus 200 Mev/c^2. This avoids the bump in negative ΔE around minus 175 MeV/c^2, which is broad and hard to describe. The peak in negative ΔE at 75 Mev/c^2 has to be described due to its proximity to the signal peak in order to get an accurate fit result. This peak comes from D_S decays, where a B^0 decays to $D^{*\pm}D_S^{\mp}$ and the $D_S^{\mp}$ decays to $KK\pi$. One kaon of the D_S decay is misidentified as pion and therefore reconstructed as signal decay. According to PDG [8], the branching ratio of $B^0 \to D^{*\pm}D_S^{\mp}$ is one order of magnitude higher than the signal decay $B^0 \to D^{*\pm}D^{\mp}$. This is why these misidentified decays make such a large contribution in ΔE sideband. Because the D_S originate from real B^0 mesons, the misidentified events appear as signal in M_{bc}.

With the help of the Monte Carlo matching function, these misidentified D_S events have been isolated on generic Monte Carlo. They are fitted with a single Gaussian function in ΔE, as shown in figure 4.3. For this fit, a sample three times the size of the whole data set was used. The *pdf* for misidentified kaons from D_S decays depends on three parameters. For the complete fit, later on the mean and width are fixed to the obtained values, which leaves only the number of misidentified kaons $N_{K/\pi}$ as a free parameter. Because the misidentified events look like signal events in M_{bc}, a Gaussian function with the mean and width of the signal function is taken as *pdf*.

While isolating misidentified events with MC-matching, some peaking events in positive ΔE were found, which are hardly visible by eye. These events come from B-decays into $D^{*\pm}$ and three pions, where one pion is falsely reconstructed as kaon and therefore interpreted as signal decay. As they also peak in M_{bc}, they cannot be neglected. The branching ratio for $B^0 \to D^{*\pm}\pi^+\pi^-\pi^{\mp}$ is about ten times higher than the one of $B^0 \to D^{*\pm}D^{\mp}$ according to previous measurements [8]. A Gaussian function was also chosen as *pdf* for the misidentified events in positive ΔE. The fit, which was performed on the same generic MC sample as above, is shown in 4.4. Width and mean are again fixed in the total fit. For M_{bc}, also a Gaussian function with the same mean and width as the signal function is used.

4.2.3. Continuum Background Description

Continuum background events come from various sources, like $e^+e^- \to q\bar{q}$ events,[6] for example. The characteristics of continuum events are that they are continuously distributed over the discriminating variables.

The continuum background in ΔE can be described by a linear function, which has a negative slope as it is more likely that some energy is missing than that there is

---

[6]With: q = u, d, s, or c. Often only these events are referred to as continuum events in contrast to $e^+e^- \to b\bar{b}$ events.
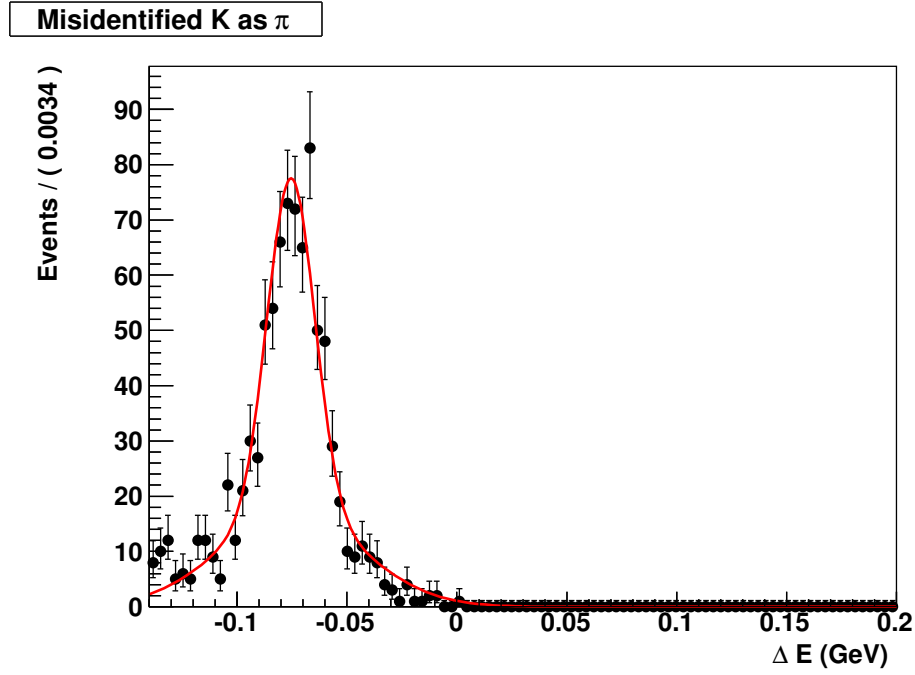
Figure 4.3.: Events where a K from a D_S decay is misidentified as a π in ΔE sideband. They are fitted with a Gaussian function
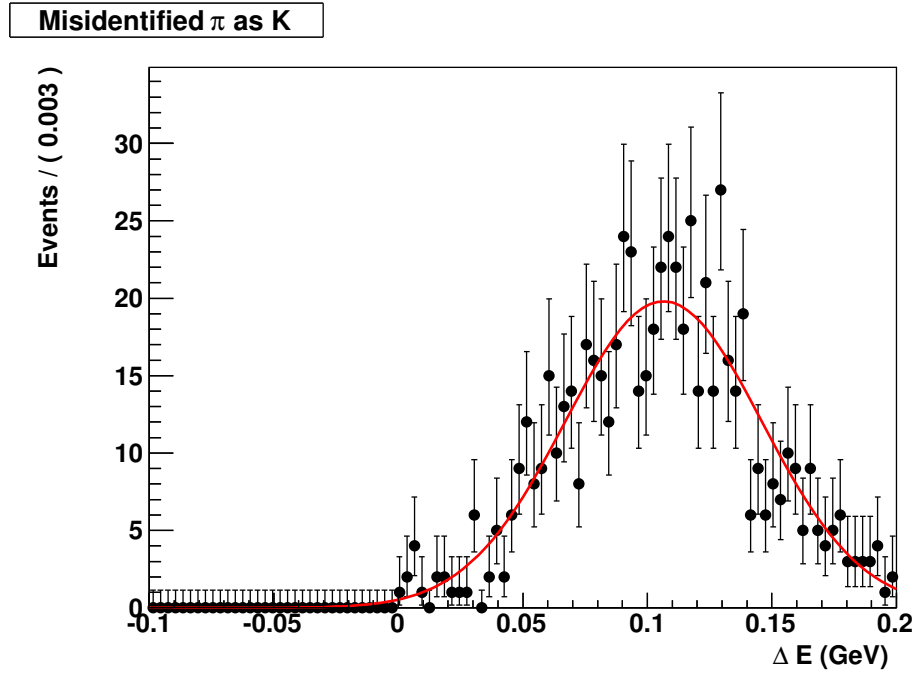
Figure 4.4.: Events where a π from is misidentified as a K in ΔE sideband. They are fitted with a Gaussian function

(a) Continuum M_{bc} (b) Continuum ΔE
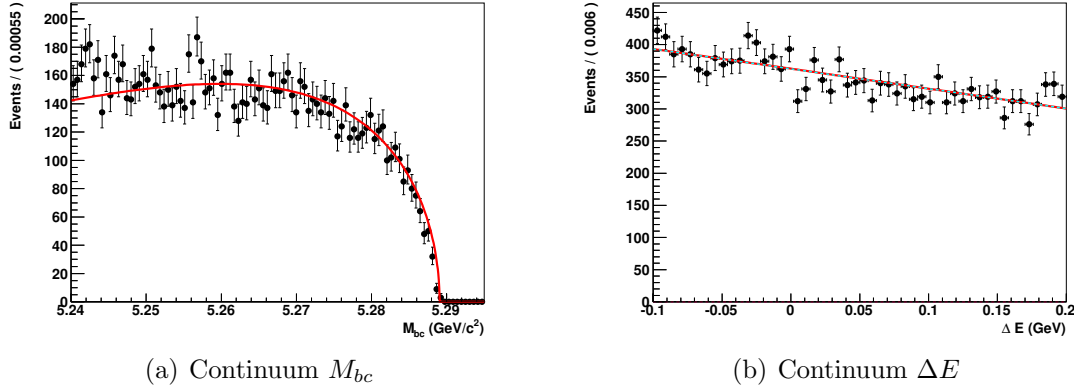
Figure 4.5.: Plot (a) shows the continuum background of M_{bc} on generic MC fitted with an ARGUS function. Plot (b) shows the continuum of ΔE on generic MC modeled with a linear function.

additional energy in an event. For M_{bc}, the continuum background is well described by the phenomenological ARGUS distribution[7], which was named after the ARGUS Collaboration. It describes the continuum background of a particle's invariant mass. In figure 4.5, two plots for the continuum of ΔE and M_{bc} are shown.

4.2.4. Background from Non-Resonant $B^0 \to D^{*\pm} K_S^0 \pi^{\mp}$ Decays

The *non-resonant* decay of $B^0 \to D^{*\pm} K_S^0 \pi^{\mp}$ yields the same final state as the *resonant* decay $B^0 \to D^{*\pm} D^{\mp}(K_S^0 \pi^{\mp})$ which is reconstructed in this analysis. The vertex fit to the D-meson should also work for non-resonant decays, as K_S^0 and $\pi^{\pm}$ originate from the same vertex, the B-meson decay vertex. This makes it hard to reject these non-resonant decays. The non-resonant decays have a branching ratio of $(3.0 \pm 0.8) \cdot 10^{-4}$ according to [8]. This means that the branching ratio of non-resonant decays is about 30 times larger than that of resonant decays. To separate non-resonant from resonant decays, one has to look at the D-meson mass. Resonant decays only appear within a certain range of the $K_S^0 \pi^{\pm}$-meson mass, because they decay via a D-meson. To determine the number of non-resonant decays that pollute the signal region, one has to look at the sideband region of the D-meson mass. Far enough from the D-mass peak away there should be no contribution from resonant decays. The sideband region covers the region from $24 \text{ MeV}/c^2$ to $60 \text{ MeV}/c^2$ away from the nominal D-mass, which is shown in orange in figure 4.6. It is four times larger than the signal region (shown in blue), which covers a $9 \text{ MeV}/c^2$ window around the nominal value of the D-mass. The number of non-resonant decays is distributed flatly over the D-mass sideband. To obtain this number, the same fit model that describes the signal region is used to perform a fit on the sideband region.

---

[7]The analytical form of the ARGUS function is given in appendix A.1
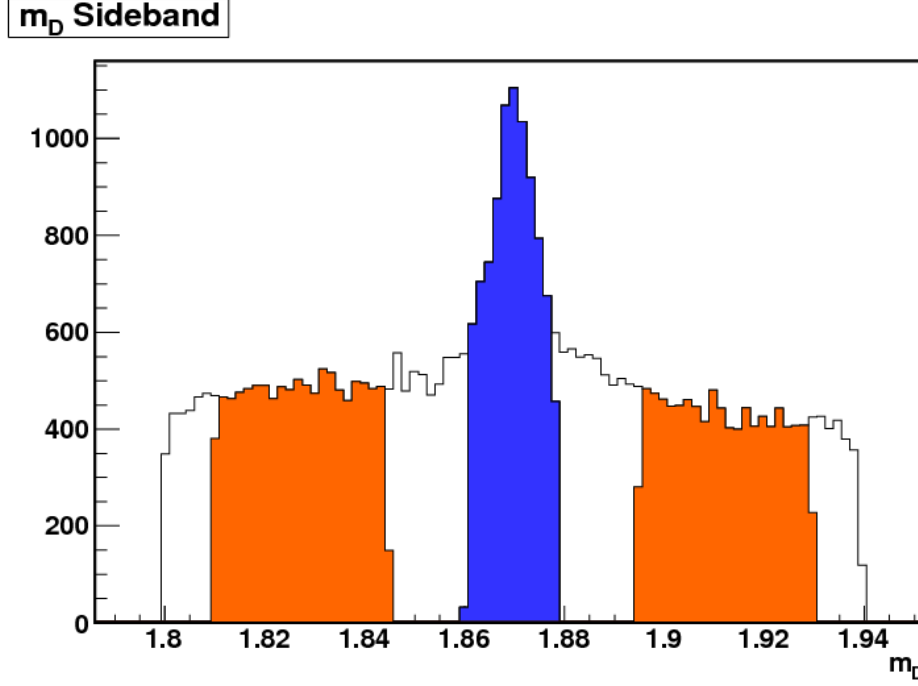
Figure 4.6.: The D-mass sideband is defined as $24\text{MeV}/c^2 < |m_{D^\pm} - m_{D^\pm}^{nom}| < 60\text{MeV}/c^2$. The signal region is defined as $|m_{D^\pm} - m_{D^\pm}^{nom}| < 9\text{MeV}/c^2$

4.2.5. Summary of Fit Parameters

As illustrated in the preceding chapters, several *pdfs* are necessary to properly describe the shape of ΔE and M_{bc}. The fit itself is two-dimensional, which means that each contribution to the likelihood function is a product of *pdfs* from M_{bc} and ΔE. The functions used in the fit are listed below. The parameters of the functions below are listed in the same order as in the description of the functions, which can be found in appendix A.1.

- Functions describing ΔE
    - Narrow signal Gaussian: $G_n^S(\Delta E; \mu_{\Delta E}, \sigma_n^S)$
    - Broad signal Gaussian: $G_b^S\left(\Delta E; \mu_{\Delta E}, \sigma_b^S := s_{n/b}^S \cdot \sigma_n^S\right)$
    - Narrow crossfeed Gaussian: $G_n^C(\Delta E; \mu_{\Delta E}, \sigma_n^C)$
    - Broad crossfeed Gaussian: $G_b^C\left(\Delta E; \mu_{\Delta E}, \sigma_b^C := s_{n/b}^C \cdot \sigma_n^C\right)$
    - Gaussian for K misidentified as π: $G^{K/\pi}(\Delta E; \mu^{K/\pi}, \sigma^{K/\pi})$
    - Gaussian for π misidentified as K: $G^{\pi/K}(\Delta E; \mu^{\pi/K}, \sigma^{\pi/K})$
    - Linear function for continuum: $f(\Delta E; a) = a \cdot \Delta E$

- Functions describing M_{bc}

- Signal Gaussian: $G\left(M_{bc}; \mu_{M_{bc}}, \sigma_{M_{bc}}\right)$
- Crossfeed Crystal Ball function: $C\left(M_{bc}; \mu_{M_{bc}}, \sigma_{M_{bc}}^{CB}, \alpha\right)$
- Gaussian for π misidentified as K: $G\left(M_{bc}; \mu_{M_{bc}}, \sigma_{M_{bc}}\right)$
- Gaussian for K misidentified as π: $G\left(M_{bc}; \mu_{M_{bc}}, \sigma_{M_{bc}}\right)$
- ARGUS distribution for continuum: $A\left(M_{bc}; c, m_0\right)$

- Functions describing the number of events
  - Poisson distribution for signal and crossfeed events: $P(N_{S/C})$
  - Poisson distribution for K misidentified as π: $P(N_{K/\pi})$
  - Poisson distribution for π misidentified as K: $P(N_{\pi/K})$
  - Poisson distribution for continuum events: $P(N_B)$

The total liklihood function is calculated using the functions defined above as:

$$\mathcal{L}_{tot} = P_{(N_{S/C})} \cdot \left[\left(\left(G_n^S{}_{(\Delta E)} + \lambda_{n/b}^S \cdot G_n^S{}_{(\Delta E)}\right) \cdot G_{(M_{bc})}\right) + \lambda_{S/C} \cdot \left(\left(G_n^C{}_{(\Delta E)}\right.\right.\right.$$

$$\left.\left.+ \lambda_{n/b}^C \cdot G_n^C{}_{(\Delta E)}\right) \cdot C_{(M_{bc})}\right)\right] + P_{(N_{K/\pi})} \cdot \left(G^{K/\pi}{}_{(\Delta E)} \cdot G_{(M_{bc})}\right)$$

$$+ P_{(N_{\pi/K})} \cdot \left(G_{(\Delta E)} \cdot G^{\pi/K}{}_{(M_{bc})}\right) + P_{(N_B)} \left(f_{(\Delta E)} \cdot A_{(M_{bc})}\right)$$

The extended liklihood function $\mathcal{L}_{tot}$ contains 13 fixed and 10 free parameters, which are shown in table 4.2.

4.2.6. Fit Results on Generic Monte Carlo

The extended likelihood function defined above was fitted on a sample of generic Monte Carlo. In figure 4.7 and 4.8, one can see the projections in ΔE and M_{bc} of the total extended maximum likelihood fit performed on one stream of Monte Carlo. The solid red line represents the total *pdf*. The dashed blue line stands for the background coming from kaons misidentified as pions, whereby the main contribution is $B^0 \rightarrow D^{*\pm}D_S^{\mp}$ decays. The fit yields to 301 ± 30 events of this type. This is in good agreement with the 278 events found by the matching function. One can see the contribution of these events in the M_{bc} signal region. Therefore, it is important to describe these events in ΔE to obtain the right number of signal events in M_{bc}. As crosscheck that the misidentified events are really peaking in M_{bc}, one can look at the projection in the ΔE sideband below -40 MeV outside the signal region. Figure 4.11 shows clearly a peak in M_{bc} coming from the misidentified events. Another peaking background in M_{bc} comes from events where a pion was misidentified as a kaon. The fit to these events is marked by the dashed red line. The number of events obtained by this fit is 215 ± 39. This is in perfect agreement with the 216 events expected from Monte Carlo matching. In figure 4.12, a M_{bc} projection of the ΔE sideband above 40 MeV shows the peaking contribution of these misidentified events. The signal contribution

Description	Parameter	Value in fit
mean of ΔE signal	$\mu_{\Delta E}$	free†
width of narrow ΔE signal Gaussian	σ_n^S	free†
width scale factor broad signal Gaussian	$s_{n/b}^S$	2.97
fraction of signal Gaussians	$\lambda_{n/b}^S$	0.88
width of narrow crossfeed Gaussian	σ_n^C	0.034 GeV
width scale factor broad crossfeed Gaussian	$s_{n/b}^C$	2.18
fraction of crossfeed Gaussians	$\lambda_{n/b}^C$	0.91
fraction of signal to crossfeed	$\lambda^{S/C}$	0.95
mean of signal M_{bc}	$\mu_{M_{bc}}$	free†
width of signal Gaussian in M_{bc}	$\sigma_{M_{bc}}$	free†
asymmetry parameter of crossfeed crystal ball	α	1.51
width of crossfeed crystal ball	$\sigma_{M_{bc}}^{CB}$	0.0036 GeV
number of signal + crossfeed events	$N_{S/C}$	free
mean negative ΔE sideband peak	$\mu^{K/\pi}$	-0.075 GeV
width negative ΔE sideband peak	$\sigma^{K/\pi}$	0.018 GeV
number of events in negative ΔE sideband peak	$N_{\pi/K}$	free
mean positve ΔE sideband peak	$\mu^{\pi/K}$	0.107 GeV
width positive ΔE sideband peak	$\sigma^{\pi/K}$	0.039 GeV
number of events in positive ΔE sideband peak	$N_{K/\pi}$	free
number of continuum events	N_B	free
shape parameter of ARGUS function	c	free
cut off parameter of ARGUS function	m_0	5.289 GeV/c^2
gradient of linear function in ΔE	a	free

Table 4.2.: Fit parameters. The values of the free parameters that are marked with a
$\dagger$, are compatible with the values obtained on signal MC.

is plotted as a dashed green line the crossfeed as a dashed orange line. The number of signal and crossfeed events combined is found to be 1778 ± 53 by the fit. The signal is described very well by the fit. Including the fraction of signal to crossfeed of 95%, which was obtained on signal Monte Carlo, the number of signal events is 1689 ± 53.[8]. As mentioned above, the data sample still includes events from non-resonant decays $B^0 \to D^{*\pm} K_S^0 \pi^{\mp}$, which have to be subtracted from the number of signal events. Therefore, the same extended maximum likelihood fit is performed on a D-mass sideband sample with $24\ \text{MeV}/c^2 < |m_D - m_D^{PDG}| < 60\ \text{MeV}/c^2$. The projections of ΔE and M_{bc} in the signal region are shown in figure 4.13 and 4.14. The fit can also describe the D-mass sideband region quite well. The number of signal events in the sideband is 213 ± 19. Since the sideband region is four times larger than the signal region, about 53 events are expected in the signal region from this fit. This is in good agreement with MC matching, which expects 58 events. If one subtracts the 53 events obtained from the sideband fit from the 1689 events obtained in the fit before, one gets 1636 ± 53 signal events, which is in good agreement with the 1613 of the MC matching. The results are summarized as "stream 04" in table 4.3, where also results of another Monte Carlo sample of same size are shown. The results of the fit on the other MC sample are also in good agreement with MC matching expectations. The branching ratio of $B^0 \to D^{*\pm} D^{\mp}$ is the sum of $\mathcal{BR}\left(B^0 \to D^{*-} D^+\right)$ and $\mathcal{BR}\left(B^0 \to D^{*-} D^+\right)$. The current world average of $\mathcal{BR}\left(B^0 \to D^{*\pm} D^{\mp}\right)$ is $(6.1 \pm 1.5) \cdot 10^{-4}$ according to PDG [8]. To calculate the branching ratio on Monte Carlo one has to know the D and D^*-meson branching fractions and the number of $B\bar{B}$-pairs on Monte Carlo. The integrated luminosity on MC is about 2.6% lower than on data. The D and D^*-meson branching fractions were estimated with a tool comparable to the Monte Carlo counter explained in section 4.1.1. They are listed in appendix A.1. One has to take all this into account to calculate the branching fractions for the two streams of generic Monte Carlo that were fitted above. The MC sample used in the fits above yields a branching ratio of $(11.8 \pm 0.4) \cdot 10^{-4}$; the other sample, called stream 01, yields to a branching fraction of $(11.9 \pm 0.4) \cdot 10^{-4}$. This is in good agreement with the expected branching ratio of $12.2 \cdot 10^{-4}$, which was used for Monte Carlo production[9].

stream	$N_{sig/cross}$	$N_{K/\pi}$	$N_{\pi/K}$	$N_{K_S^0 \pi}$	N_{sig}	
01	1790 ± 53	297 ± 30	300 ± 39	$\frac{204 \pm 20}{4} = 51$	1650	FIT
	-	270	240	54	1649	MC-MATCHING
04	1778 ± 53	301 ± 30	215 ± 39	$\frac{213 \pm 19}{4} = 53$	1636	FIT
	-	278	216	58	1613	MC-MATCHING

Table 4.3.: Fit results on generic MC

---

[8]To be conservative, the error is not scaled down with the signal to crossfeed fraction.

[9]In the generic Monte Carlo decay table the branching ratio of $B^0 \to D^{*\pm} D^{\mp}$ is exactly twice as large as the world average given by PDG.
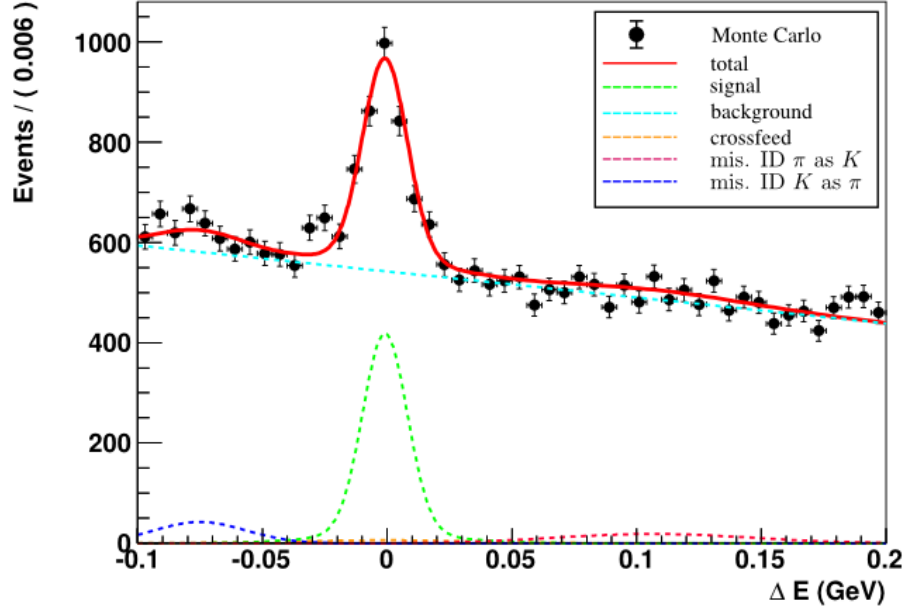
Figure 4.7.: Projection of ΔE on one stream of generic Monte Carlo
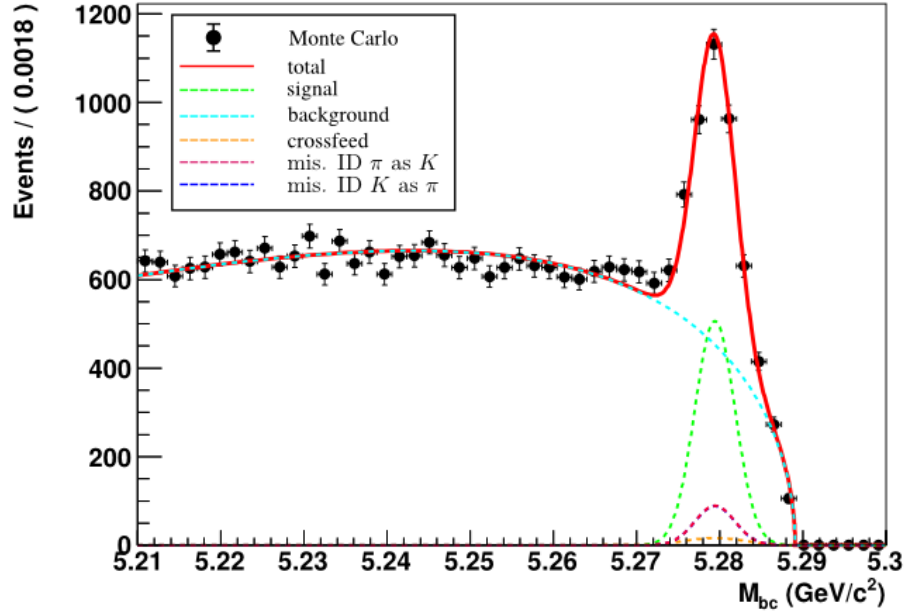
Figure 4.8.: Projection of M_{bc} on one stream of generic Monte Carlo
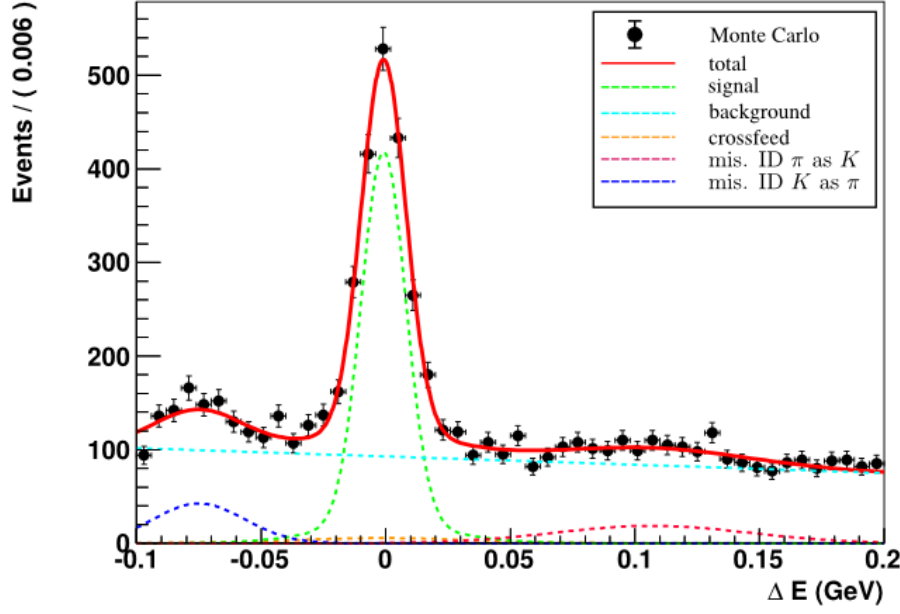
Figure 4.9.: Projection of ΔE in M_{bc} signalbox on one stream of generic Monte Carlo
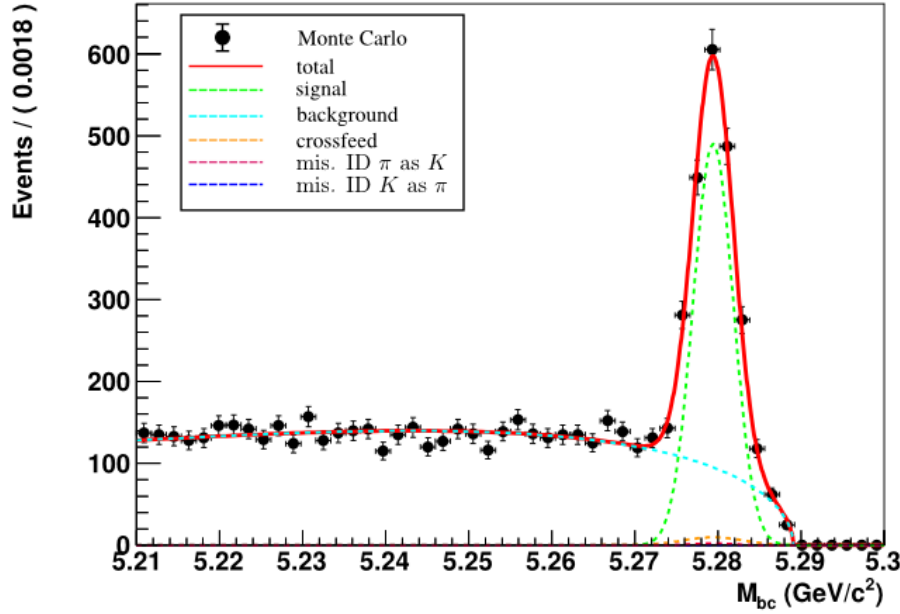
Figure 4.10.: Projection of M_{bc} in ΔE signalbox on one stream of generic Monte Carlo
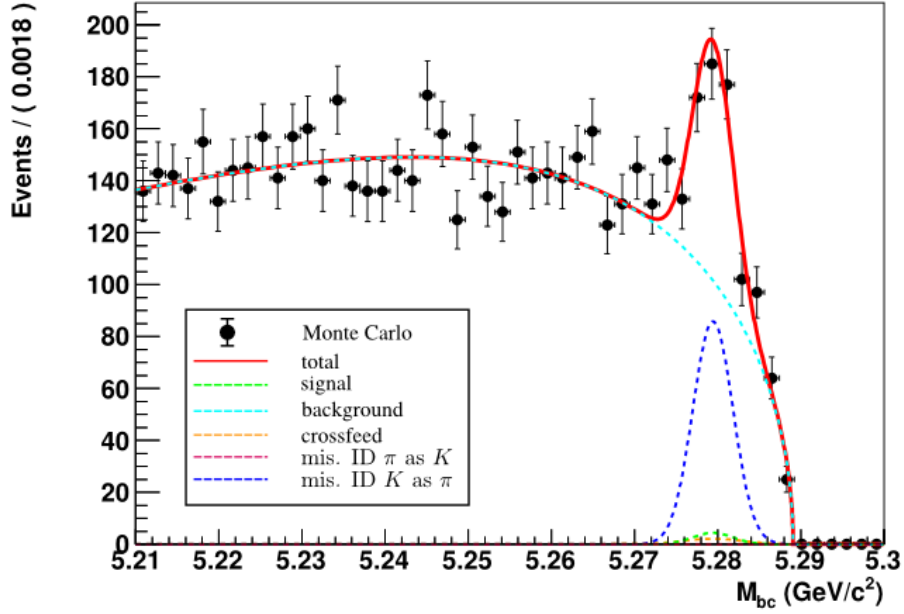
Figure 4.11.: Projection of M_{bc} in ΔE sideband below -40MeV on one stream of generic Monte Carlo
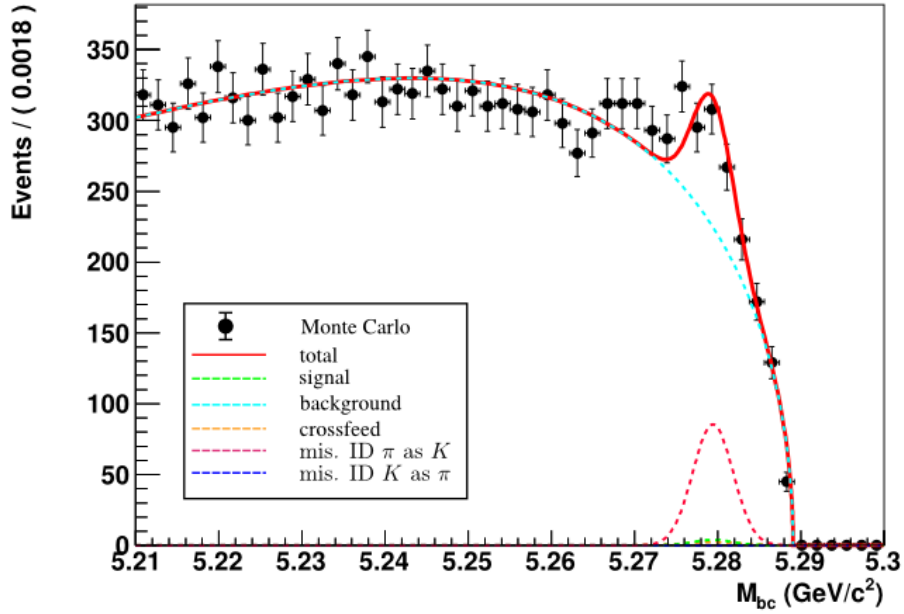
Figure 4.12.: Projection of M_{bc} in ΔE sideband below -40MeV on one stream of generic Monte Carlo
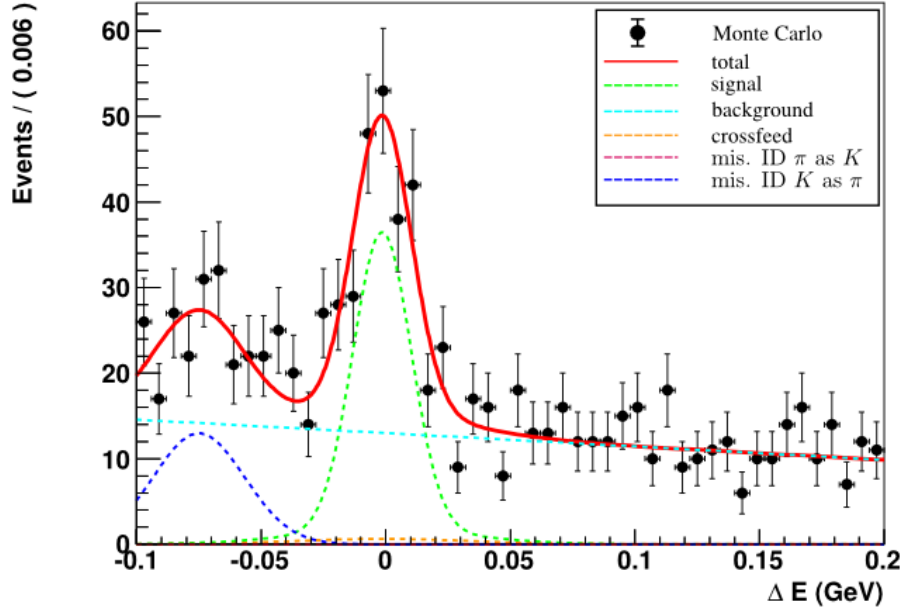
Figure 4.13.: Projection of ΔE in D-mass sideband with $M_{bc} > 5.27\text{GeV}/c^2$ on one stream of generic Monte Carlo
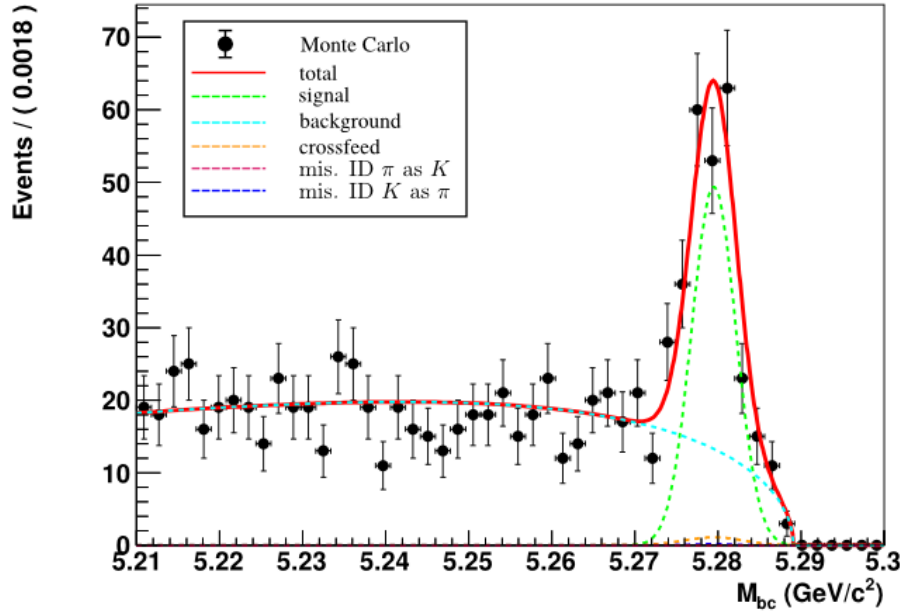
Figure 4.14.: Projection of M_{bc} in D-mass sideband with $|\Delta E| < 30\text{MeV}$ on one stream of generic Monte Carlo

5. Branching Ratio Measurement

5.1. Corrections between Data and MC

Efficiency Corrections

The efficiencies were measured on signal Monte Carlo for every channel, taking all selection criteria into account. To separate kaons from pions, a cut on $\mathcal{L}_{K/\pi}^{ATC}$, the ATC particle identification described in detail in section 3.1.3, was performed. It is a known problem that the efficiency for a certain cut on this variable is not the same in data and MC. Therefore, it is necessary to determine the correction factors for the efficiencies on data. This is done by using a standard Belle tool, which is described in detail in reference [21]. The tool compares efficiencies for kaons and pions with certain cuts on $\mathcal{L}_{K/\pi}^{ATC}$ for data and Monte Carlo. The efficiencies of kaons and pions were calculated in different momentum bins for the reference channel $D^{*-} \to D^0 \pi^- \to (K^+ \pi^-)$ on data. The efficiencies obtained there are compared with the values obtained on Monte Carlo to calculate the correction factor. Due to the difference in reconstruction efficiency for tracking with the original silicon vertex detector (SVD1) and the new silicon vertex detector (SVD2), the correction factors are calculated for these two different data samples to be weighted afterwards according to the amount of data. The total efficiency correction factor is calculated for every channel separately according to the number of kaons and pions it contains.

$$\epsilon_{cor} = \left(\frac{N_{B\bar{B}}^{(svd1)}}{N_{B\bar{B}}^{all}} \cdot \epsilon_{cor}^{(svd1)} + \frac{N_{B\bar{B}}^{(svd2)}}{N_{B\bar{B}}^{all}} \cdot \epsilon_{cor}^{(svd2)} \right)^{N_K} + \left(\frac{N_{B\bar{B}}^{(svd1)}}{N_{B\bar{B}}^{all}} \cdot \epsilon_{cor}^{(svd1)} + \frac{N_{B\bar{B}}^{(svd2)}}{N_{B\bar{B}}^{all}} \cdot \epsilon_{cor}^{(svd2)} \right)^{N_\pi}$$

	SVD1	SVD2
Pion	0.9868 ± 0.0059	0.9864 ± 0.0110
Kaon	1.0008 ± 0.0092	1.0068 ± 0.0097
$N_{B\bar{B}}$	152	620

Table 5.1.: Efficiency Correction Factors

Correction of Mean Values for ΔE Sideband Peaks

The mean of ΔE was found to be at around -0.5 MeV on generic Monte Carlo. The negative value of the ΔE mean is due to the fact that the detector cannot measure

the whole energy of a certain event. This shift in ΔE can be different between data and Monte Carlo. That is no problem for the mean of the signal in ΔE, as it is a free parameter in the fit. On data, the mean was estimated to be -1.6 MeV. The mean of the two sideband peaks which are fixed are shifted by 1.1 MeV to accommodate for this.

5.2. Results on Data

The whole Belle data set of 772 million $B\bar{B}$-pairs recorded at the Y(4S)-resonance has been reconstructed. All selection cuts and the best candidate selection as described in chapter 3 have been applied. The resulting data sample is fitted with the unbinned extended maximum likelihood fit, described in chapter 4. In doing that the mean values of the ΔE sideband peaks are corrected according to the shift of ΔE between data and Monte Carlo. Figures 5.1 to 5.6 show the results of the fit.

The projection of ΔE for $M_{bc} > 5.27$ can be seen in figure 5.3. The fit function, drawn as a solid red line, describes the data well. Both sideband peaks, the dashed blue and the dashed red line coming from misidentification, are clearly visible. The projection of M_{bc} in the ΔE signal region, plotted in figure 5.4, makes the impact of these misidentification events clearly visible. In figure 5.5, one can see the projection of M_{bc} for $\Delta E < -40$ MeV. The clearly visible peak comes from $D_S^+ \rightarrow K^+ K^- \pi^+$ decays, where a kaon is misidentified as a pion. The peak in the M_{bc} projection for $\Delta E > 40$ MeV, where a pion is misidentified as kaon, is shown in figure 5.6. It is a rather small bump but with 154 ± 32 events still significant.

The fit yields 869 ± 39 signal and crossfeed events. This corresponds to 825 ± 39 signal events with a signal fraction of 95%, as it was determined on Monte Carlo. To be conservative, the 39 events are not scaled down with the signal fraction. The number of direct decays $B^0 \rightarrow D^{*\pm} K_S^0 \pi^\mp$ has to be substracted from the signal events. As on Monte Carlo before, the number of direct decays is estimated by fitting the D-meson mass sideband. Again, a sideband window of $24\ \mathrm{MeV}/c^2 < |m_{D^\pm}| < 60\ \mathrm{MeV}/c^2$ is used. The results of the D-meson mass sideband fit are shown as projections of ΔE and M_{bc} in the signal region in figure 5.7 and 5.8. The sideband fit yields to only 30 ± 11 events. This correspond to a 2.6σ evidence for these sideband events. As the D-meson mass window in the sideband is four times larger than in the signal region, the result of the sideband fit has to be divided by four. This yields to 818 ± 39 signal events in the end. With the corrected efficiencies, the branching ratios of the D-meson decays[1] obtained from PDG and the number of $B\bar{B}$-pairs the branching ratio is calculated.

$$\mathcal{BR}\left(B^0 \rightarrow D^{*\pm} D^\mp\right) = \frac{N_{sig}}{N_{B\bar{B}} \cdot \epsilon_{cor} \cdot \mathcal{BR}(D) \cdot \mathcal{BR}(D^*)} = (5.90 \pm 0.28) \cdot 10^{-4}$$

The given error is of statistical nature.

---

[1]The table with the used branching ratios can be found in appendix A.2.
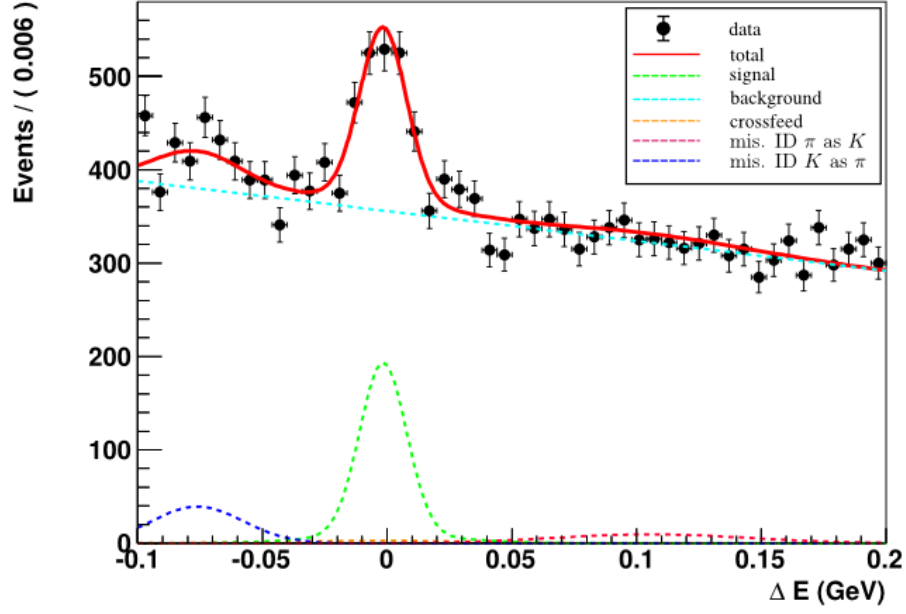
Figure 5.1.: Projection of ΔE in the whole fit range on data
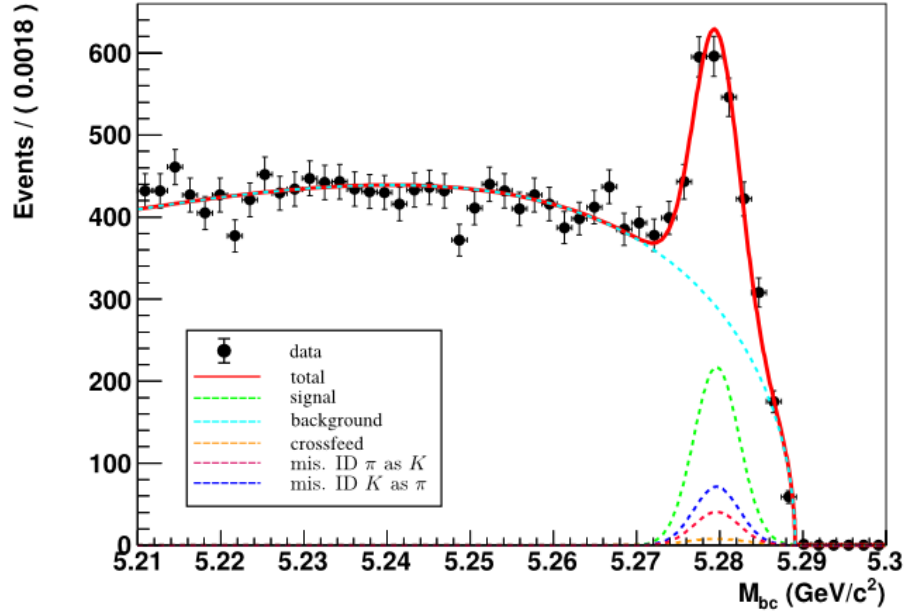
Figure 5.2.: Projection of M_{bc} in the whole fit range on data
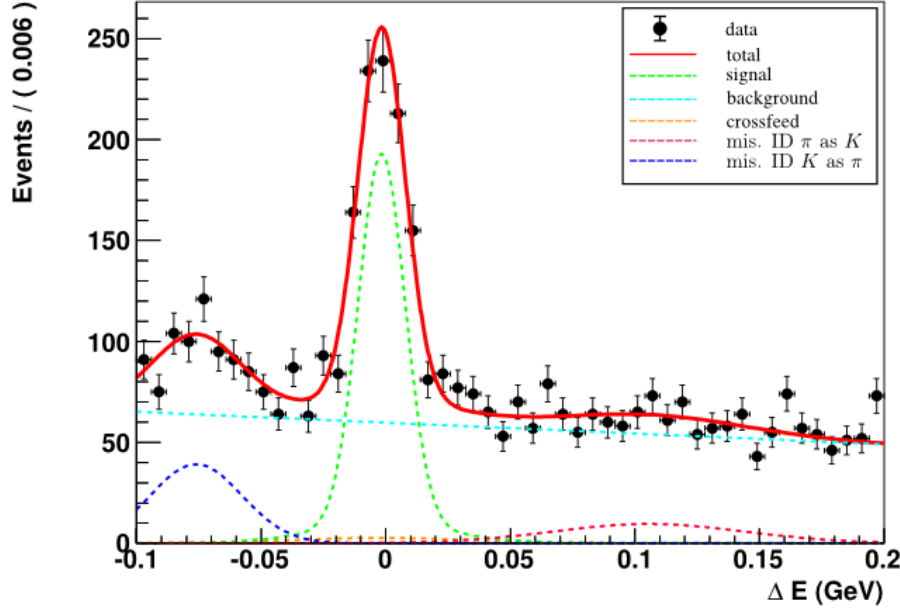
Figure 5.3.: Projection of ΔE with $M_{bc} > 5.27$ GeV$/c^2$ on data
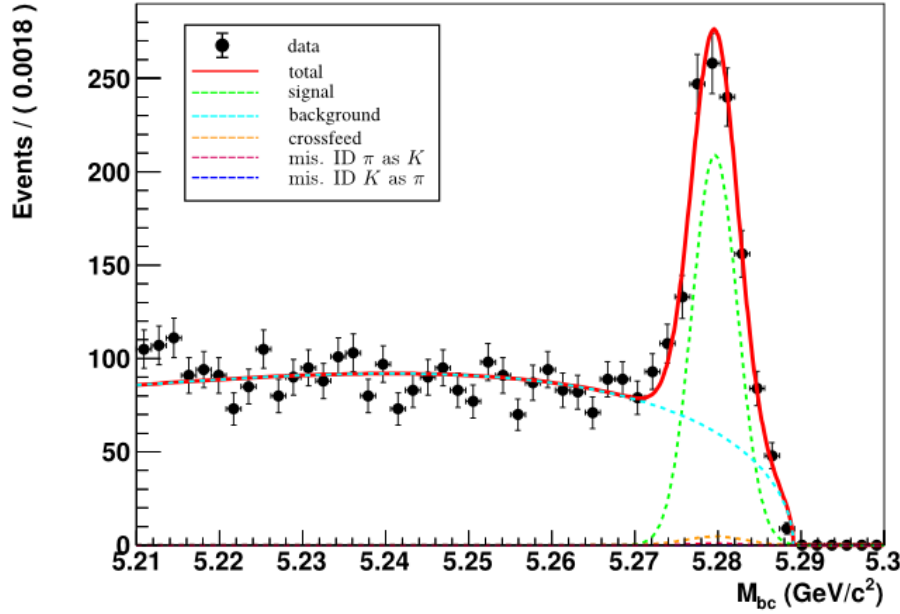
Figure 5.4.: Projection of M_{bc} with $|\Delta E| < 30$ MeV on data
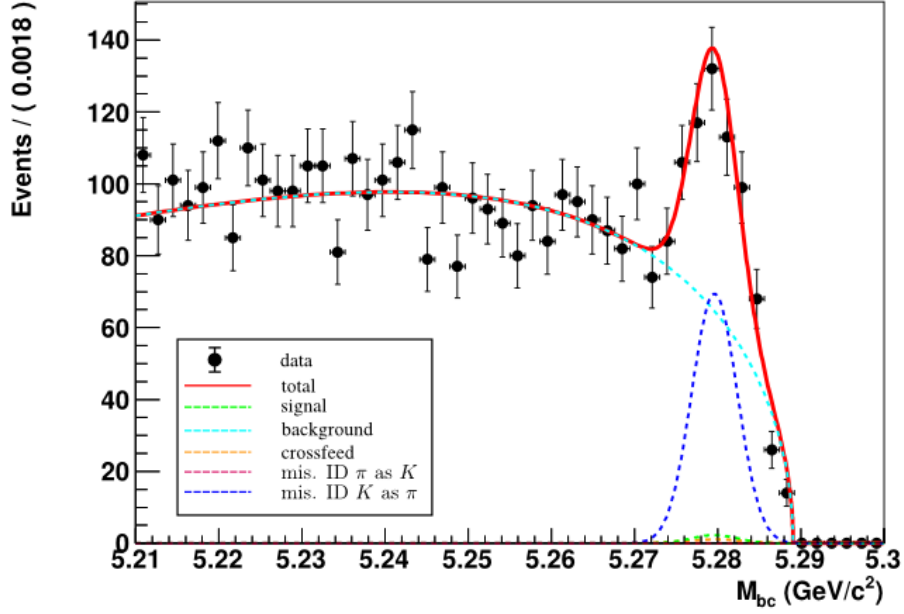
Figure 5.5.: Projection of M_{bc} in ΔE sideband below -40 MeV on data
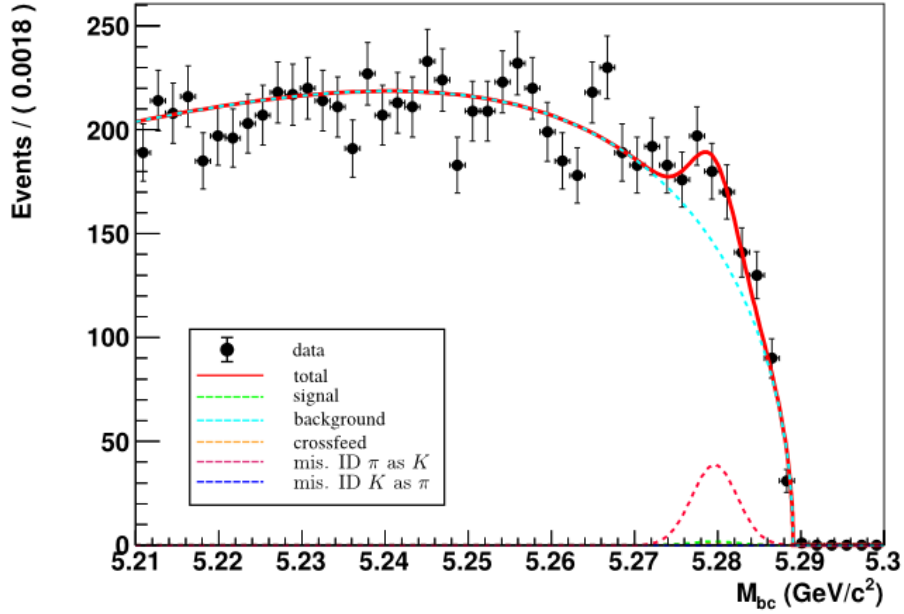
Figure 5.6.: Projection of M_{bc} in ΔE sideband above 40 MeV on data
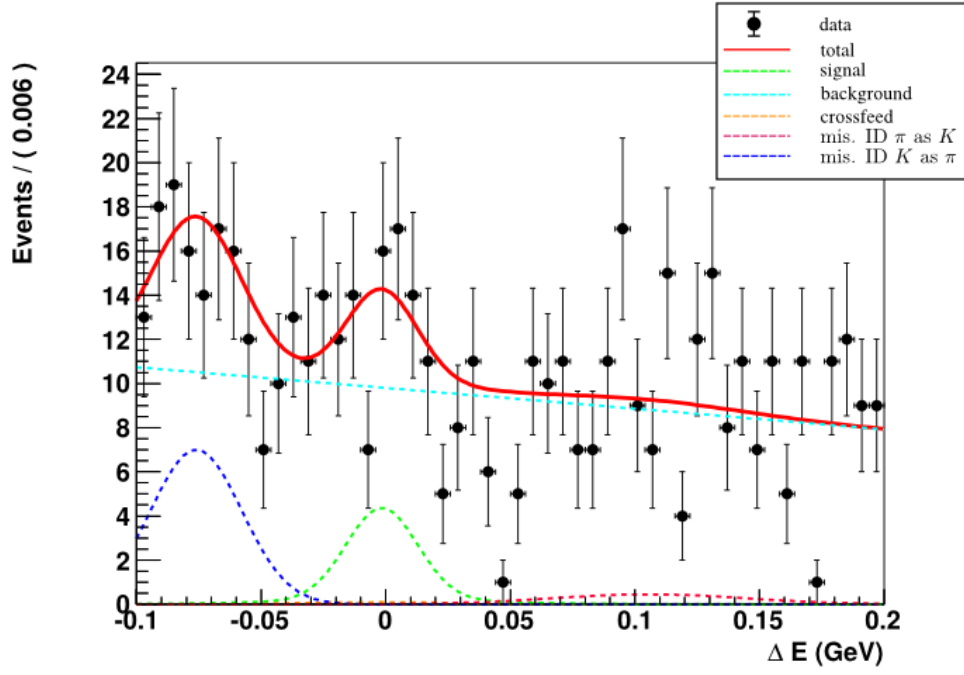
Figure 5.7.: Projection of ΔE in D-mass sideband with $M_{bc} > 5.27$ GeV/c^2 on data
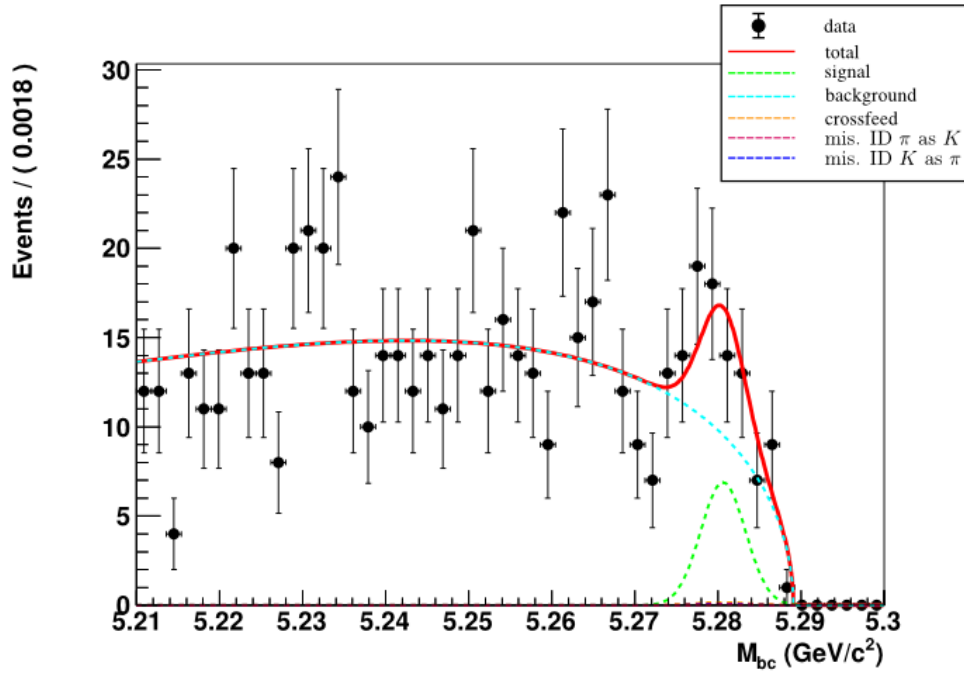
Figure 5.8.: Projection of M_{bc} in D-mass sideband with $|\Delta E| < 30$ MeV on data

5.3. Systematic Uncertainties

Different sources of systematic uncertainties have to be taken into account. The number of signal events, the number of $B\bar{B}$ pairs, the selection efficiency and the branching ratios of D and D^* meson decays are used to calculate the branching ratio. All these quantities are linearly correlated with the branching ratio. Therefore, they propagate directly to the branching ratio.

The error on the 772 million $B\bar{B}$ pairs was determined to be 1.37% by the Belle Collaboration [22].

Another source of uncertainty is the fit model. To determine the systematic error of the fit model, the fit was repeated on data, varying each fixed parameter[2] one standard deviation up and down. As standard deviation of the parameters, the errors of the fits on Monte Carlo, where the parameters were determined in the first place, are taken. Because the mean values of the lower and upper ΔE peaks are shifted on data, three times the shifting value is taken in addition to the error as variation. For each fit, the deviation of the signal yield to the signal yield of the original fit is calculated. In each fit the deviation is less than 0.5%. The parameter with the highest impact on the number of signal events is the fraction $\lambda_{S/C}$ of signal to crossfeed events. Therefore, this parameter was also examined separately. A fit with free $\lambda_{S/C}$ led to a deviation of 4.3% in the number of signal events. The errors on the other fit parameters are negligible compared to this effect, which one can see if adding up the square of each error and taking the square root of the sum.

Another systematic uncertainty arises from the difference in reconstruction efficiency between data and Monte Carlo. In this analysis, four different sources led to uncertainties in reconstruction efficiency between data and Monte Carlo: Charged track reconstruction efficiency, π^0 reconstruction efficiency, K_S^0 reconstruction efficiency and ATC particle identification[3] efficiency.

The systematic error in charged track reconstruction was estimated by P. Koppenburg et al in 2003 based on a sample of partially reconstructed D^* mesons (see reference [23]). An update of these studies, which has not been published yet, was done by B. Bhuyan. It assigns a 0.4% error to each reconstructed charged track. For each channel, the systematic uncertainty on charge track reconstruction is calculated by linearly adding the error of each charged track, since these uncertainties are 100% correlated. Each error is weighted by branching ratio times efficiency of its specific channel. Then the weighted errors of all channels are summed up. The combined error for charged track reconstruction efficiency is 2.54%.

The systematic error of the π^0 reconstruction was estimated using $\eta \to \pi^+\pi^-\pi^0$ decays by H.C. Huang et al (see reference [24]). An update of these studies, which has not been published yet, was made by Y.Liu and K.Trabelsi. It yields to a systematic error of 2.5% per reconstructed π^0. For each channel, the number of π^0 is again added linearly and weighted with the branching ratio times efficiency of this channel. The sum of all channels yields the total error to be 0.5%.

---

[2]All parameters are listed in table 4.2 in the chapter before.
[3]This is explained in detail in section 3.1.3

The error of the reconstruction efficiency per K_S^0 was found to be 3.21% by E. White [25]. Unfortunately, his result has also not been published yet. The combined error for K_S^0, calculated in the same way as for π^0, is 0.94%.

The difference in efficiencies for the ATC particle information between data and Monte Carlo was calculated in section 5.1 already. But the correction factors have uncertainties themselves. They are shown in table 5.1. Per channel the systematic uncertainty is calculated and weighted with the branching ratio times efficiency of the given channel. The systematic error due to the difference of the efficiency of ATC particle information between data and Monte Carlo sums up to 6.28%.

The branching ratios of D and D^* mesons are taken from PDG. They can be found in appendix A.2. The errors given there are taken as systematic error. The systematic error on the sum of the branching ratios per channel is calculated from the errors of each branching ratio via error propagation. It is 5.27% in total.

The total systematic error on the branching ratio is calculated as square root of the sum of the squared individual errors.

$$\sigma_{sys} = \sqrt{\sum_i \sigma_i^2} = 10.7\%$$

	error
Number of $B\bar{B}$-pairs	1.4%
Track reconstruction efficiency (K/π)	2.5%
K_S^0 reconstruction efficiency	0.9%
π^0 reconstruction efficiency	0.5%
$\mathcal{L}^{ATC}$ reconstruction efficiency	6.3%
$\mathcal{BR}(D)\,\mathcal{BR}(D^*)$	5.3%
Fit model	4.3%
Result	10.7%

Table 5.2.: Systematic Uncertainties

5.4. Conclusion

The goal of this thesis was the measurement of the exclusive branching fraction of the decay $B^0 \to D^{*\pm}D^{\mp}$. Therefore, the decays of B^0 and $\bar{B}^0$ into $D^{*\pm}D^{\mp}$ were studied intensively. A reconstruction of 12 decay channels, six $D^{*\pm}$ and two $D^{\mp}$ modes, was established. This covers 2.46% of all $B^0 \to D^{*\pm}D^{\mp}$ decays. The selection was optimized for all cut variables simultaneously. Thereby the purity in the signal box increased to 60.3%. It was shown that the signal can be clearly extracted and the background is understood.

For the measurement of the branching ratio of $B^0 \to D^{*\pm}D^{\mp}$ decays extensive studies on Monte Carlo were done. This analysis is the first that established a profound ΔE sideband description in $B^0 \to D^{*\pm}D^{\mp}$. It was shown that the fit model performs well in all projections, and that all results are in good agreement with Monte Carlo input values.

Furthermore, it was demonstrated on data the fit procedure works fine and that it can describe the data in all projections. The branching ratio measured on the full Belle data set of 772 million $B\bar{B}$-pairs is:

$$\mathcal{BR}\left(B^0 \to D^{*\pm}D^{\mp}\right) = (5.90 \pm 0.28 \pm 0.63) \cdot 10^{-4}$$

The second term is the statistical and the third term the systematic error. The result agrees with the world average of this branching ratio given by PDG [8]: $(6.1 \pm 1.5)\cdot 10^{-4}$. The world average was calculated from the two previous measurements by Belle and BaBar. The previous Belle analysis [3] of this branching ratio used a data sample of 32 million $B\bar{B}$-pairs and measured a branching ratio of $\left(11.7 \pm 2.6^{+2.2}_{-2.5}\right) \cdot 10^{-4}$.

The previous BABAR measurement [1] used a data sample of 232 million $B\bar{B}$-pairs and measured a branching ratio of $(5.7 \pm 0.7 \pm 0.7) \cdot 10^{-4}$.

The measurement performed in this analysis is the world's single most precise measurement of the branching ratio of $B^0 \to D^{*\pm}D^{\mp}$ and it is more precise than the current world average. The measurement itself is systematically limited. More precise measurements of the branching ratios of D and D^* mesons may decrease this error in the future. But a better understanding of the difference in ATC particle identification efficiency between data and Monte Carlo is most important for a more precise measurement of the branching ratio. With higher statistics, it may be possible to only use the very clean decay modes such as $B^0 \to \left(K^+\pi^-\pi^-\right)\left(K^-\pi^+\right)\pi^+$, which would also result in lower systematics.

In the future, this selection could be used for a time dependent CP-analysis. The parameters of time dependent CP-violation in $B^0 \to D^{*\pm}D^{\mp}$ decays have been measured by Belle [5] in 2004 on a data set of 152 million $B\bar{B}$-pairs. The BABAR collaboration has published a measurement of the CP-violation parameters in 2008 [4] on their full data set of 467 Million $B\bar{B}$-pairs. Both analyses measure a non-zero mixing induced CP-violation. The parameter corresponding to direct CP-violation is compatible with zero for both analysis. The measurements of Belle and BABAR are in good agreement with each other. An update on the full Belle dataset of 772 million $B\bar{B}$-pairs can be performed on the basis of the selection established in this thesis.

A. Appendix

A.1. PDF

Gaussian function

$$G(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Crystal Ball function

$$C(x; \mu, \sigma, \alpha) = N \cdot \begin{cases} exp\left(-\frac{(x-\mu)^2}{2 \cdot \sigma^2}\right) & \text{for} & \frac{(x-\mu)^2}{2 \cdot \sigma^2} > \alpha \\ \left(\frac{n}{|\alpha|}\right)^n exp\left(-\frac{|\alpha|^2}{2}\right) \cdot \left(\frac{n}{|\alpha|} - \alpha - \frac{x-\mu}{\sigma}\right)^{-n} & \text{for} & \frac{(x-\mu)^2}{2 \cdot \sigma^2} \leq \alpha \end{cases}$$

ARGUS distribution

$$A(x; c, m_0) = x \left(1 - \frac{x^2}{m_0^2}\right)^{\frac{1}{2}} exp\left(c\left(1 - \frac{x^2}{m_0^2}\right)\right) \qquad \forall \quad 0 \leq x \leq m$$

Poisson distribution

$$P(N_x, \eta) = \frac{\eta^{N_x}}{N_x!} e^{-\eta}$$

A.2. D-meson mass fits

(a) $D^+ \rightarrow K^-\pi^+\pi^+$

(b) $D^0 \rightarrow K^-\pi^+\pi^+\pi^-$

(c) $D^0 \rightarrow K^0_S\pi^+\pi^-$
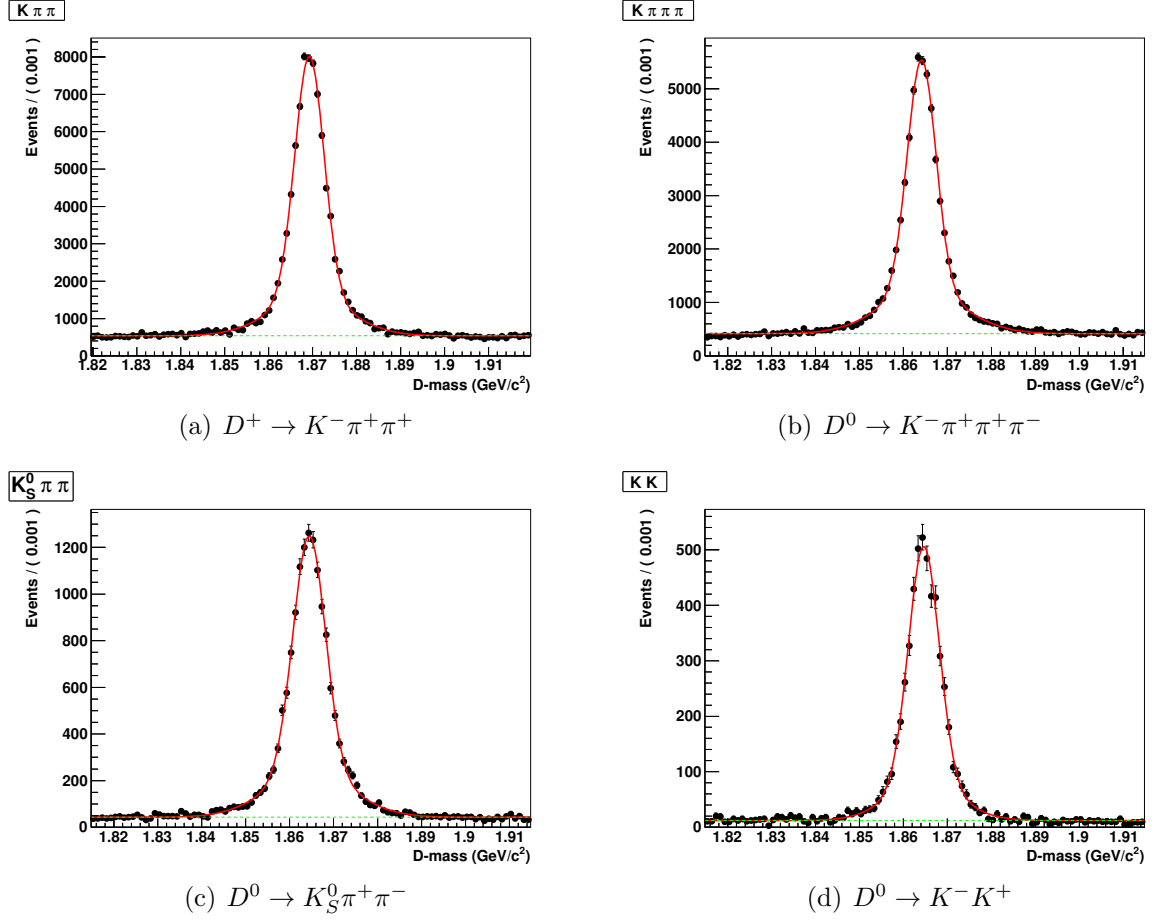
(d) $D^0 \rightarrow K^-K^+$

Figure A.1.: Fit of D-Meson masses and D^*D mass-difference on generic Monte Carlo

A.3. Optimization of channels

(a) Mass of D^+

(b) Mass difference Δm_{D^*D}
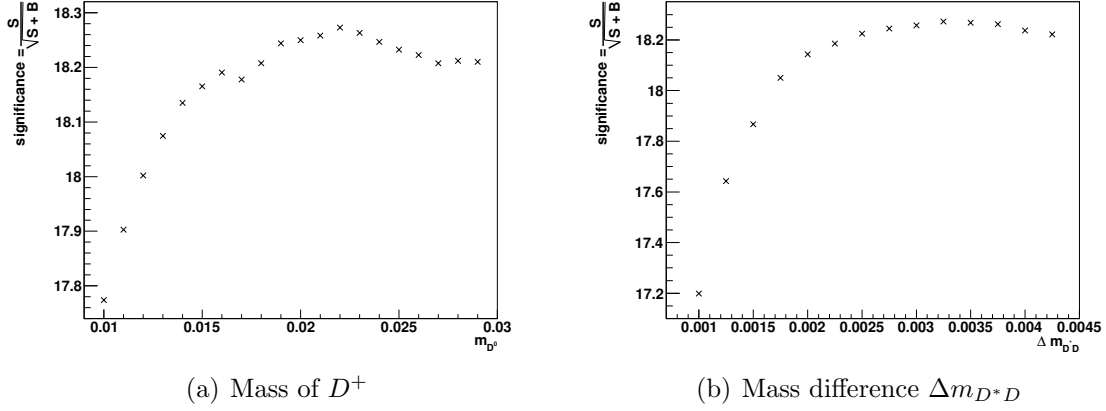
Figure A.2.: Channels: $D^0 \to K^-\pi^+$ and $D^0 \to K^-K^+$. The plots show the significance plotted over the cut of a certain variable, while all other variables are cut at their point of highest significance.

(a) Mass of D^+

(b) Mass difference Δm_{D^*D}
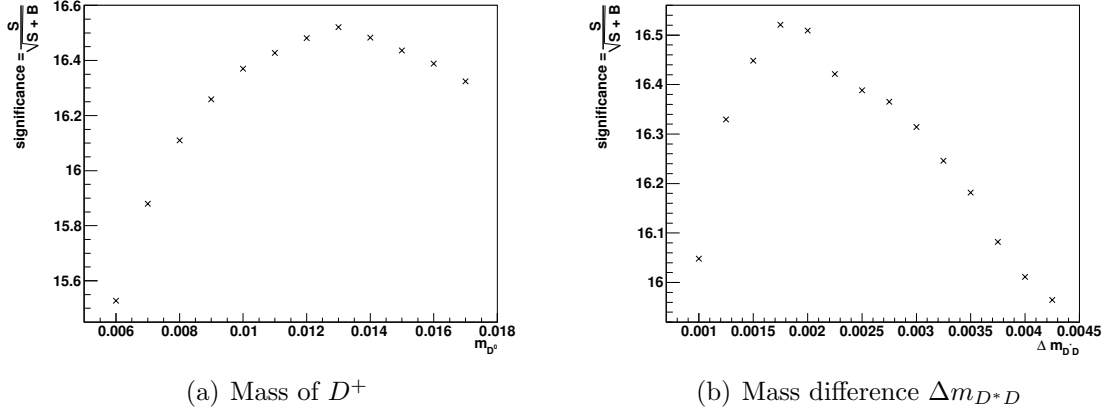
Figure A.3.: Channel: $D^0 \to K^-\pi^+\pi^-\pi^+$. The plots show the significance plotted over the cut of a certain variable, while all other variables are cut at their point of highest significance.

(a) Mass of D^+

(b) Mass difference Δm_{D^*D}

Figure A.4.: Channel: $D^0 \to K_S^0 \pi^+ \pi^-$.

(a) Mass of D^+

(b) Mass difference Δm_{D^*D}

(c) Mass of π^0
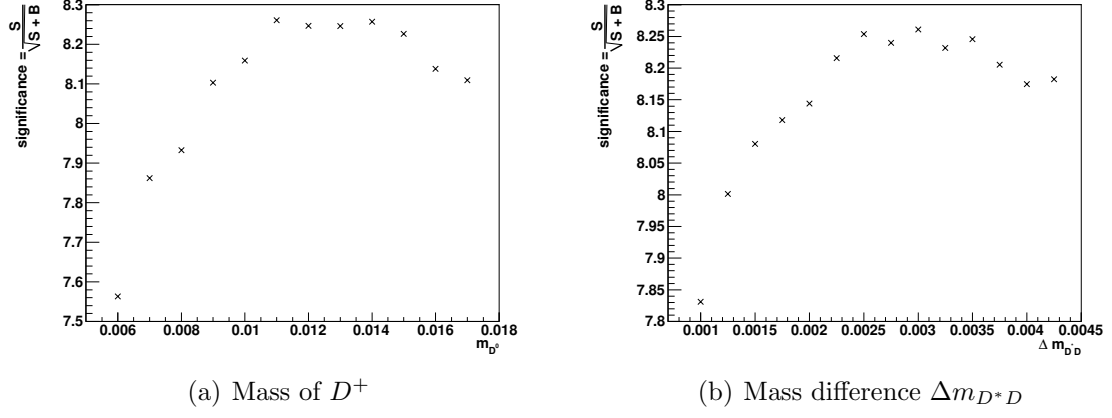
(d) Minimal γ energy of π^0
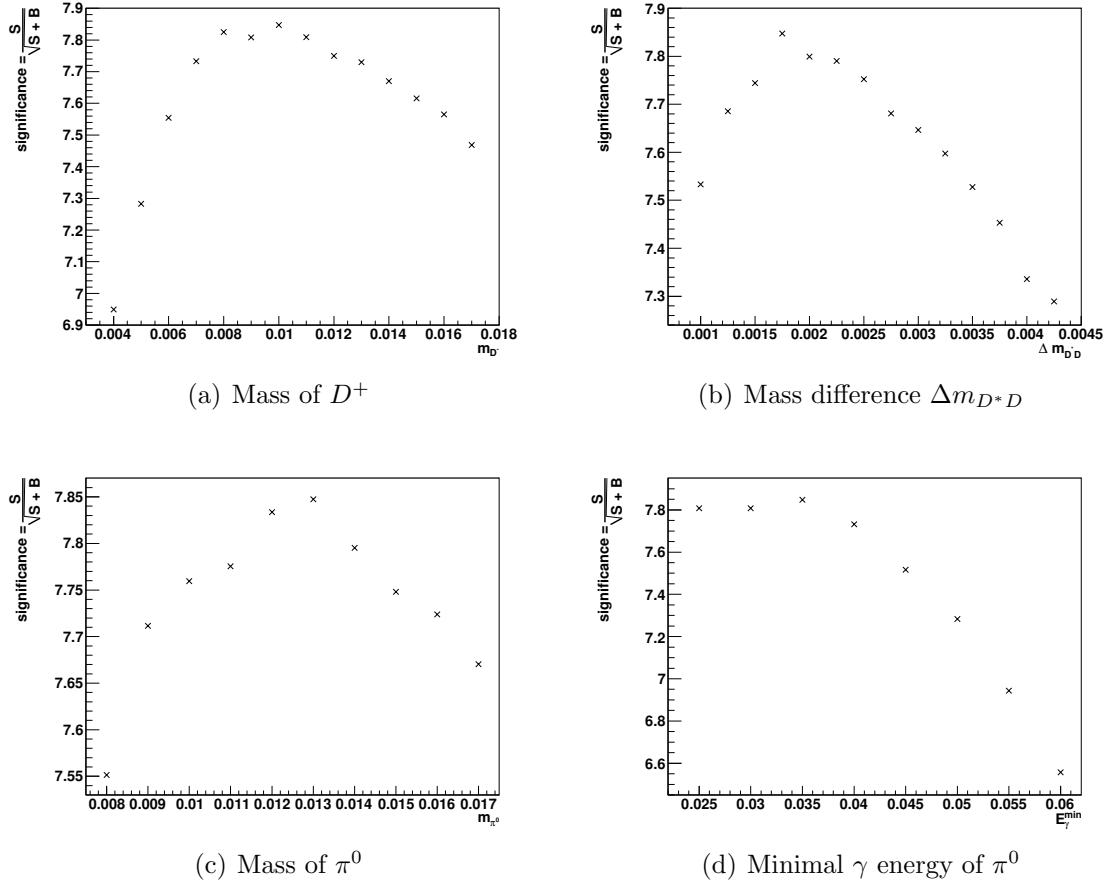
Figure A.5.: Channel $D^0 \to K^- \pi^+ \pi^0$. The plots show the significance plotted over the cut of a certain variable, while all other variables are cut at their point of highest significance.

A.4. Branching Fractions

Channel	Branching Ratio	Error
$D^+ \to K^- \pi^+ \pi^+$	9.50%	0.03%
$D^+ \to K_S^0 \pi^+$	1.48%	0.01%
$D^0 \to K^- \pi^+$	3.8%	0.02%
$D^0 \to K^- \pi^+ \pi^+ \pi^-$	7.47%	0.03%
$D^0 \to K^- \pi^+ \pi^0$	13.39%	0.04%
$D^0 \to K_S^0 \pi^+ \pi^-$	3.13%	0.02%
$D^0 \to K^- K^+$	0.39%	0.01%
$D^{*+} \to D^0 \pi^+$	67.7%	0.5%
$D^{*+} \to D^+ \pi^0$	30.7%	0.5%

Table A.1.: Branching Fractions on MC

Channel	Branching Ratio	Error
$D^+ \to K^- \pi^+ \pi^+$	9.4%	0.4%
$D^+ \to K_S^0 \pi^+$	1.49%	0.04%
$D^0 \to K^- \pi^+$	3.89%	0.05%
$D^0 \to K^- \pi^+ \pi^+ \pi^-$	8.09%	0.2%
$D^0 \to K^- \pi^+ \pi^0$	13.9%	0.5%
$D^0 \to K_S^0 \pi^+ \pi^-$	2.94%	0.16%
$D^0 \to K^- K^+$	0.39%	0.01%
$D^{*+} \to D^0 \pi^+$	67.7%	0.5%
$D^{*+} \to D^+ \pi^0$	30.7%	0.5%

Table A.2.: Branching Fractions of PDG [8]

A.5. Signal Monte Carlo decay table

```
 #Aliases

Alias myB0 B0
Alias myAntiB0 anti-B0
ChargeConj myB0 myAntiB0

Alias myD*- D*-
Alias myD*+ D*+
ChargeConj myD*- myD*+

Alias my1D- D-
Alias my1D+ D+
ChargeConj my1D- my1D+

Alias my2D+ D+
Alias my2D- D-
ChargeConj my2D+ my2D-

Alias myD0 D0
Alias myantiD0 anti-D0
ChargeConj myD0 myantiD0

Decay Upsilon(4S)
1.0 B0 anti-B0 myB0 myAntiB0 VSS_BMIX dm;
Enddecay

Decay myB0
1.0 myD*- my1D+ PHOTOS SVS_CP beta_cdcd dm -1 1.0 0.0 1.0 0.0;
Enddecay

CDecay myAntiB0

Decay myD*-
0.215381 myantiD0 pi- PHOTOS VSS;
0.031426 my2D- pi0 PHOTOS VSS;
Enddecay

CDecay myD*+

Decay myantiD0
0.0389  K+ pi- PHOTOS PHSP;
0.0809 K+ pi- pi- pi+ PHOTOS PHSP;
```

```
0.1390 K+ pi- pi0 PHOTOS D_DALITZ;
0.0294 K_S0 pi+ pi- PHOTOS PHSP;
0.00394 K+ K- PHOTOS PHSP;
Enddecay

CDecay myD0

Decay my1D+

0.094 K- pi+ pi+ PHOTOS D_DALITZ;
0.0149 K_S0 pi+ PHOTOS PHSP;
Enddecay

CDecay my1D-

Decay my2D+

0.094 K- pi+ pi+ PHOTOS D_DALITZ;
Enddecay

CDecay my2D-

End
```

List of Figures

List of Tables

Bibliography

[1] Aubert, B. Measurement of branching fractions and CP-violating charge asymmetries for B-meson decays to $D^{(*)}D^{(*)}$, and implications for the Cabibbo-Kobayashi-Maskawa angle γ. *Phys. Rev. D*, 73(11):112004, Jun 2006.

[2] K. Abe et al. Observation of the decay $B^0 \to D^{\pm}D^{*\mp}$. *Phys. Rev. Lett.*, 89:122001, 2002.

[3] B. Aubert et al. Observation of CP violation in the B^0 meson system. *Phys. Rev. Lett.*, 87(9):091801, 2001 `arXiv:hep-ex/0107013v1`.

[4] Bernard Aubert et al. Measurements of time-dependent CP asymmetries in $B^0 \to D^{(*)}D^{(*)}$ decays. *Phys. Rev.*, D79:032002, 2009.

[5] T. Aushev et al. Search for CP violation in the decay $B^0 \to D^{*\pm}D^{\mp}$. *Phys. Rev. Lett.*, 93:201802, 2004.

[6] S. Kurokawa and E. Kikutani. Overview of the KEKB Accelerators. *Nucl. Inst. Meth. A*, 499(1):1–7, 2003.

[7] A. Abashian et al. The Belle Detector. *Nucl. Inst. Meth. A*, 479(1):117–232, 2002.

[8] K. Nakamura et al. (Particle Data Group), Journal of Phyxics G 37 075021 (2010). `http://pdg.lbl.gov`, 2010.

[9] Rith Povh. *Teilchen und Kerne*, volume 8. Springer, 2008.

[10] D. Griffiths. *Introduction to Elementary Particles*, volume 2. WILEY-VCH, 2008.

[11] C. Berger. *Teilchenphysik*, volume 2. Springer, 2001.

[12] A. Zupanc. Measurement of D^0 mixing in $D^0 \to \phi K_S^0$ decays, 2009.

[13] M. Prim. Neural Network based B^0 Flavor Tagging at the Belle Experiment, 2010.

[14] K. Furukawa et al. Towards reliable acceleration of high-energy and high-intensity electron beams. 2000.

[15] KekB public relations website. `http://belle.kek.jp/belle/transparency/`, 2011.

[16] Luminosity - KEKb website. `http://www-acc.kek.jp/kekb/Commissioning/Record/Total/total_luminosity.html`, 2011.

[17] F. Fang. Study of $K_S^0 \rightarrow \pi^+\pi^-$ Selection. *Belle internal note 323*, 2000.

[18] D. Kirkby et al. A User's Guide to the RooFitTools Package for Unbinned Maximum Likelihood Fitting. 2001.

[19] V.Blobel/E.Lohrmann. *Statistische und numerische Methoden der Datenanalyse*, volume 2. B.G. Teubner, 1998.

[20] T. Nelson H. Nelson. A Guide to Unbinned Maximum Likelihood Fits. 1998.

[21] S. Nishida. Study of Kaon an Pion Identification Using Inclusive D^* Sample. *Belle internal note 779*, 2005.

[22] Belle internal website. `http://belle.kek.jp/secured/nbb/nbb.html`, 2011.

[23] P. Koppenburg. A measurement of the Track Finding Efficiency Using Partially Reconstructed D^* Decays. *Belle internal note 621*, 2003.

[24] H.C. Huang. Uptdate of π^0 Systematics Using Inclusive η (78/fb). *Belle internal note 645*, 2003.

[25] E. White. Study of K_S^0 Systematics. *Belle Analysis Meeting*, 2010.

Danksagung

Erklärung zur eigenständigen Anfertigung meiner Diplomarbeit

Hiermit erkläre ich, dass ich die vorliegende Diplomarbeit

Measurement of the Branching Fraction
of the Decay $B^0 \to D^{*\pm} D^{\mp}$
at the Belle Experiment

selbständig verfasst und alle verwendeten Quellen und Hilfsmittel angegeben habe.

Daniel Stemmer   Karlsruhe, den 31.05.2011