Search for CP violation in $\mathbf{B^0 \to D^{(*)\mp} K^0 \pi^\pm}$ decays

Ph.D. dissertation

submitted

19$^{\text{th}}$ of May 2004

for the degree of

Philosophiae Scientiae Doctor

by

Troels C. Petersen

Jury:

M.	É. Augé (Président)
Mme	B. Bloch-Devaux
M.	J. Iliopoulos
M.	L. Lanceri (Rapporteur)
M.	F. Le Diberder
M.	J. Renner Hansen (Invité)
M.	O. Schneider (Rapporteur)
Mme	M-H. Schune (Invitée)

Laboratoire de l'Accélérateur Linéaire

Universite de Paris XI, Orsay

This Ph.D. thesis is submitted for the title of doctor of science thereby concluding the physics studies of Troels C. Petersen at Université de Paris XI.

The thesis comprises of work on experimental particle physics, along with the treatment of some theoretical aspects. It was commenced in Orsay in June 2001, continued at Berkeley in September 2001 and at SLAC in 2002, and finalized in Paris in April 2004 for *soutenance* (eng: thesis defence) the 19$^{\text{th}}$ of May 2004.

The work was mainly supported by the Danish Ministry of Research through Forskeruddannelsesraadet (FUR) and Laboratoire de l'Accélérateur Linéaire (LAL). Additional funds were granted through UniDanmark Fonden.

This document is set in LaTeX and graphs were made with `xfig`, `PAW`, and `Root`.

Abstract

Following the idea of considering three-body decays with final state $B \to DK\pi$ in the pursuit of the notoriously elusive unitary angle γ [APS03, AP03], the initial steps towards its measurement are presented.

Using a sample of approximately 88 million $B\overline{B}$ pairs collected with the *BABAR* detector at the PEP-II collider, a measurement of the branching fractions $B^0 \to D^{\pm}K^0\pi^{\mp}$ and $B^0 \to D^{*\pm}K^0\pi^{\mp}$ are presented for the entire Dalitz region:

$$
\begin{aligned}
Br(B^0 \to D^{\pm}K^0\pi^{\mp}) &= (4.97 \pm 0.69\,(\text{stat.}) \pm 0.55\,(\text{syst.})) \times 10^{-4}, \\
Br(B^0 \to D^{*\pm}K^0\pi^{\mp}) &= (3.00 \pm 0.66\,(\text{stat.}) \pm 0.29\,(\text{syst.})) \times 10^{-4}.
\end{aligned}
$$

In both decay modes the dominant resonance is the $K^{*\pm}(892)$ between the neutral kaon and the charged pion. Measuring the branching fractions of the corresponding resonant two-body decays yields:

$$
\begin{aligned}
Br(B^0 \to D^{\pm}K^{*\mp}) &= (4.78 \pm 0.58\,(\text{stat.}) \pm 0.53\,(\text{syst.})) \times 10^{-4}, \\
Br(B^0 \to D^{*\pm}K^{*\mp}) &= (3.22 \pm 0.59\,(\text{stat.}) \pm 0.29\,(\text{syst.})) \times 10^{-4}.
\end{aligned}
$$

From these measurements the resonant fractions are determined to be:

$$
\begin{aligned}
f(B^0 \to D^{\pm}K^{*\mp}) &= 0.64 \pm 0.08\,(\text{stat.}) \pm 0.02\,(\text{syst.}), \\
f(B^0 \to D^{*\pm}K^{*\mp}) &= 0.72 \pm 0.13\,(\text{stat.}), \pm 0.02\,(\text{stat.}).
\end{aligned}
$$

Performing a time-dependent CP asymmetries fit of the resonant decay mode $B^0 \to D^{\pm}K^{*\mp}$ yields:

$$
\begin{aligned}
C = 0.93 \pm 0.18 \pm 0.03 \quad &\Longrightarrow \quad |\lambda| = 0.19 \pm 0.25 \pm 0.04 \\
S = 0.18 \pm 0.28 \pm 0.02 \quad & \quad\quad\;\; \Delta S = 0.08 \pm 0.28 \pm 0.02
\end{aligned}
$$

As the amplitude ratio $|\lambda|$ is consistant with zero, no interference sensitive to γ can be established at this point.

Résumé

La mise en oeuvre de la nouvelle méthode [APS03, AP03], reposant sur les canaux à trois corps dans l'état final $B \to DK\pi$, récemment proposée pour permettre la mesure de l'élusif angle γ du triangle d'unitarité est présentée.

Mettant à profit un échantillon de 88 million de paires $B\overline{B}$ collectées avec le détecteur *BABAR* situé auprès du collisionneur PEP-II, les mesures des rapports d'embranchement $B^0 \to D^{\pm}K^0\pi^{\mp}$ et $B^0 \to D^{*\pm}K^0\pi^{\mp}$ sont réalisées sur la totalité des diagrammes de Dalitz.

$$\begin{aligned} Br(B^0 \to D^{\pm}K^0\pi^{\mp}) &= (4.97 \pm 0.69\,(\text{stat.}) \pm 0.55\,(\text{syst.})) \times 10^{-4}, \\ Br(B^0 \to D^{*\pm}K^0\pi^{\mp}) &= (3.00 \pm 0.66\,(\text{stat.}) \pm 0.29\,(\text{syst.})) \times 10^{-4}. \end{aligned}$$

Pour chacun de ces deux modes de désintégration, la contribution dominante se révèle être celle de la résonnance $K^{*\pm}(892)$ entre le kaon neutre et le pion chargé. Les mesures des désintégrations quasi deux corps correspondantes donnent:

$$\begin{aligned} Br(B^0 \to D^{\pm}K^{*\mp}) &= (4.78 \pm 0.58\,(\text{stat.}) \pm 0.53\,(\text{syst.})) \times 10^{-4}, \\ Br(B^0 \to D^{*\pm}K^{*\mp}) &= (3.22 \pm 0.59\,(\text{stat.}) \pm 0.29\,(\text{syst.})) \times 10^{-4}. \end{aligned}$$

Ces valeurs correspondent aux fractions résonantes:

$$\begin{aligned} f(B^0 \to D^{\pm}K^{*\mp}) &= 0.64 \pm 0.08\,(\text{stat.}) \pm 0.02\,(\text{syst.}), \\ f(B^0 \to D^{*\pm}K^{*\mp}) &= 0.72 \pm 0.13\,(\text{stat.}), \pm 0.02\,(\text{stat.}). \end{aligned}$$

L'analyse en temps de l'échantillon $B^0 \to D^{\pm}K^{*\mp}$ conduit à la détermination des paramètres clefs suivants:

$$\begin{aligned} C = 0.93 \pm 0.18 \pm 0.03 \quad &\Longrightarrow \quad |\lambda| = 0.19 \pm 0.25 \pm 0.04 \\ S = 0.18 \pm 0.28 \pm 0.02 \quad & \quad\quad\;\; \Delta S = 0.08 \pm 0.28 \pm 0.02 \end{aligned}$$

Le rapport d'amplitudes $|\lambda|$ étant compatible avec zéro, la statistique disponible ne permet pas encore d'exploiter les effets d'interférence pour mesurer l'angle γ.

Il faut de toute nécessité que des actions dissymétriques président pendant la vie à l'élaboration des vrais principes immédiats naturels dissymétriques. Quelle peut être la nature de ces actions dissymétriques? Je pense, quant à moi, qu'elles sont d'ordre cosmique. L'univers est un ensemble dissymétrique et je suis persuadé que la vie, telle qu'elle se manifeste à nous, est fonction de la dissymétrie de l'univers ou des conséquences qu'elle entraîne. L'univers est dissymétrique.

It is inescapable that asymmetric forces must be operative during the synthesis of the first asymmetric natural products. What might these forces be? I, for my part, think that they are cosmological. The universe is asymmetric, and I am persuaded that life, as it is known to us, is a direct result of the asymmetry of the universe or of its indirect consequences. The universe is asymmetric.

[Louis Pasteur, 1822-1895]

Contents

List of Figures

List of Tables

1 Preface

The following thesis concerns itself with B physics and CP violation at the *BABAR* experiment. It is the result of two and a half years work on the subject, including both theoretical and experimental work. As this thesis rests on many decades of tradition and experience in particle physics, the reader is assumed to have a basic knowledge of this field. Some knowledge of B physics is also advisable. An general introduction to particle physics can be found in textbooks [PS95, HM84], and a thorough treatment of B physics at *BABAR* in [Har98].
A description of published material, writing customs, conventions of units and the various subjects treated can be found below.

1.1 About this thesis

Three parts of the thesis were originally written in other contexts, and though included in rewritten form, they still leave traces. The first two are the theoretical parts on the charged [APS03] and neutral decays [AP03], which were published in PRD and the proceedings of the Durham CKM workshop (submitted to Eur. Phys. J), respectively. The third part is the documentation for the analysis [PS03, PS04] in the final process of internal review, to be submitted to Phys. Rev. Lett.

1.2 Conventions

In the following, abbreviations will be used extensively as is custom in particle physics. At first appearance they will be typed out and the abbreviation written afterward in parenthesis. However, the names of accelerators and experiments will not always be written out, and the unfamiliar reader will refered to Appendix 15.3, where a list of abbreviations can be found.

The convention $c = \hbar = 1$ will be used as is customary, which means that all masses and momenta will be measured in terms of e.g. GeV and not GeV/c^2 and GeV/c, respectively. However, the usual value of c will be retained, when discussing times and distances, which in turn will most often be expressed in ps and μm, as they are the relevant scales for B physics. The symbol B^0 will refer to the B_d^0 meson, and charge conjugate modes are implied throughout the thesis unless otherwise stated[1]. Four vectors are written with normal font, while three vectors are written with an above vector arrow, and matrices are written in bold, unless indices and/or context suggest otherwise.

1.3 Language

In many respects experiments in particle physics, not the least *BABAR*, resemble building the tower of Babel in the sense that many different languages (and cultures) are involved. However, though many beautiful languages are spoken within the *BABAR* collaboration, I have chosen English for this thesis, as it is the scientific standard and by far the most common language in the collaboration.

> *Je ne suis pas comme une dame de la cour de Versailles, qui disait: "C'est bien dommage que l'aventure de la tour de Babel ait produit la confusion des langues; sans cela tout le monde aurait toujours parlé francais".*
>
> I am not like a lady at the court of Versailles, who said: "What a dreadful pity that the bother at the tower of Babel should have got languages all mixed up; but for that, everyone would always have spoken French".
>
> [Voltaire 1694–1778, In letter to Catherine the Great]

Some but far from all quotations are taken from [Mac91] and [ed.01].

[1] This means that e.g. $B^0 \to D^-\pi^+$ implicitly includes the conjugate decay $\overline{B}^0 \to D^+\pi^-$

1.4 Subjects treated

When starting this thesis, the "golden" $\sin(2\beta)$ measurement had already become a standard analysis, and while analyses involving α had (perhaps for historical reasons) been undertaken from the beginning of *BABAR*, measurements of and/or constraints on the notoriously difficult angle γ were still in their infancy.

The prospects at the time were only very few, as the Gronau-Wyler (GW) method [GW91] was almost the only method proposed, and had inherent deficiencies. Also the $B^0 \to D^{(*)\pm}\pi^\mp$ had been proposed [DS88], but whether fully or partially reconstructing the decays, the smallness of the amplitude ratio remains challenging. Finally the $B \to K\pi$ channels and $SU(3)$ symmetries had been considered [HLGR94, GHLR94], but here one faces $SU(3)$ breaking corrections and electroweak penguins.
As a corollary it can be mentioned, that the chapter on extracting γ, despite including B_s decays, is the shortest physics analysis chapter in the *BABAR* Physics Book [Har98].

Building upon an idea conceived by R. Aleksan and F. Le Diberder, the feasibility of an analysis using the threebody decay $B^\pm \to D^0 K^\pm \pi^0$ was studied with Abi Soffer (Fall 2001). This work continued (Spring 2002) and was presented (Blois 2002, Durham 2003) along with a paper accepted by PRD. The channel $B^0 \to D^\mp K_S^0 \pi^\pm$ was also considered, and the feasibility of this time-dependent analysis was studied (Spring 2002) and presented (Durham 2003, SLAC 2003).

After training (Fall 2001), the duty as commissioner for the *BABAR* PID system (the DIRC detector) was fulfilled with T. Hadig, B. Meadows and M. Pivk (Jan. – Jul. 2002), and a measurement of the number of signal photons per track as a function of time was made with J. Schwiening to monitor the endurance of the DIRC.

The $B^0 \to D^\mp K^0 \pi^\pm$ analysis with M.H. Schune and for a period V. Tano[2] was started based on a skim of 56 fb^{-1} (Spring 2002). Through a simple cut and count analysis, which was approved for unblinding (July 2002), a signal was established.
The analysis was then repeated with a larger data sample (82 fb^{-1}), more refined selection techniques and including the $B^0 \to D^{*\pm} K^0 \pi^\mp$ mode (Sep. – Dec. 2002). For extracting the signal an unbinned maximum likelihood fit was constructed, and through the use of a novel statistical technique [PLD04] the signal distribution in the Dalitz plot was established (Jan. – Sep. 2003). Though statistically limited, the initial steps of a time-dependent analysis was taken, yielding the ratio of interfering amplitudes and the CP asymmetry of the $B^0 \to D^\pm K^{*\mp}$ mode (Nov. 2004 – Jan. 2004).

Finally, master student T. Kittelmann was partially supervised during his work on the B^0 mass difference $\Delta\Gamma_B$ (Sep. 2001 – Jan. 2003). For his master thesis [Kit03], he received the highest grade possible at the University of Copenhagen.

1.5 Outline

This thesis consists of three parts. The first is a theoretical introduction to B physics, which goes through mixing and CP violation in general. Then the phenomenology upon which this thesis is based will be treated in some depth, and finally the current knowledge of the CKM matrix and CP violation is summarized.

The second part describes the PEP-II collider and the *BABAR* detector along with the triggers and the performance. Each subdetector will be described, with emphasis on the DIRC, which is both essential to this analysis and for which the author served as commissioner.

The third part contains the analysis. After an outline of the analysis the selection of the data samples is presented. Then follows a thorough discussion of the fitting methods, and finally the branching ratio and time-dependent fitting results are presented along with systematic errors.

[2] Postdoc at LAL from September 2002 to May 2003.

2 Introduction

2.1 Introduction to B physics

B physics is the study of the bottom quark, b, which is done through its bound hadronic states in mainly B mesons. The primary interest is to measure the couplings between the different quarks along with their masses, since this constitutes most of the free parameters of the Standard Model (SM). As many processes involving B mesons arise only from loop effects, also the quantum structure of the theory can be tested. Furthermore, B mesons provide a basis for tests and development of Quantum Chromo Dynamics (QCD), as the heaviness of the b quark allows for perturbative calculations.

Since the discovery and the following experimental accessibility of the b quark, B physics has become a major field in particle physics, and following initial steps by LEP and CLEO, the B factories herald the era of high statistics B physics and precision CP measurements.

The goal and the challenge is to make precise predictions and measurements, thereby scrutinizing the theory in such great detail, that effects unaccounted for by the SM will become apparent, if there. This task is – as most often in science – a challeging but stimulating interplay between theory and experiment.

2.2 History of CP violation, the CKM matrix and B physics

Since the birth of physics, Nature was thought to be symmetric in time and space, and after the discovery of antiparticles, these too were considered as particles' exact opposites (Charge conjugates, C).

In 1956, after reviewing the experimental data then available, Lee and Yang concluded that spatial inversion (Parity, P) was not conserved in weak interactions [LY56]. The following year (1957) Wu *et al.* discovered parity violation in the β decay of Co^{60} to Ni^{60*} [ea57], and this was confirmed by Garvin, Lederman and Weinrich [GLW57], who in a brilliant experiment also discovered the violation of charge conjugation, C. However, the combined symmetry CP was still believed to be a good symmetry of Nature.

In 1964, CP violation was discovered by Christenson *et al.* in the decay of the long-lived neutral kaon [CCFT64], which occasionally (2×10^{-3}) decayed into a state with opposite CP, thereby breaking CP invariance.

Already in 1963, Cabibbo proposed mixing between the d and the s quark [Cab63], and in 1970 Glashow, Iliopoulos and Maiani suggested a fourth quark (the charm quark, c) in this scheme to cancel an unwanted (i.e. unobserved) $\Delta S = 1$ neutral current [GIM70]. An extension of this quark scheme to six quarks was noted as a possibility in 1973, even before the c quark discovery, by Kobayashi and Maskawa [KM73], as a 3×3 complex mixing matrix, now called the Cabibbo-Kobayashi-Maskawa (CKM) matrix, naturally allows for CP violation.

With the discovery in 1977 of the b quark in the Υ resonance [H$^+$77], the extension to three generations of quarks was verified, and a new field of physics – B physics – was born. In 1981 the B meson was discovered at the $\Upsilon(4S)$ resonance [CLEO81], and subsequent experiments measured the lifetime and oscillation frequency of the B meson [ARGUS87a], where the later measurement was used to predict the large top quark mass long before its final discovery [CDF95][3].

The currently running B factories will measure rare branching fractions and possibly time-dependent CP asymmetries with unprecedented statistics, thereby constraining and testing the CKM paradigme.

[3]The $B^0\overline{B}^0$ mixing discovery, which measured $r_d = 0.20\pm0.12$, indicated a top quark mass above ~ 80 GeV, and was announced in the light of a claimed top quark discovery, with a mass around 40 GeV [UA184].

2.3 *CP* violation in cosmological context

In 1966, two years after the discovery of CP violation and just one year after the discovery of the microwave background radiation predicted by Big Bang theory, Sakharov wrote a paper on the possibility of explaining the apparent charge asymmetry of the Universe in terms of particle theory [Sak67]. He argued that three conditions (now known as the Sakharov conditions) had to be satisfied in order to obtain a baryon asymmetry in the Universe:

Baryon non-conservation. If the total baryon number of the Universe were initially zero and baryon number was conserved, then the Universe would remain symmetric.

C **and** CP **violation.** If C or CP were conserved, the reaction rate would be the same for particles and antiparticles, and an initially symmetric Universe would remain symmetric.

Deviations from thermal equilibrium. In thermal equilibrium there is no time dependence, and consequently an initial baryon number of zero would remain that way.

There are of course ways around Sakharov's conditions. It could be that the Universe is only locally asymmetric, but this has now been ruled out locally (20 Mpc) by direct search [Ste76] and globally (visible Universe) by indirect methods [CDG98]. More exotic models also exist, but Sakharov's conditions are generally believed to be true, which makes CP violation extremely interesting also from an existential point of view.
Most interestingly, detailed models of both baryogenesis requires that there are sources of CP violation beyond the Standard Model [DLH$^+$92], and the same is true for leptogenesis[4].

2.4 Experimental situation of B physics

The experimental situation up to the commencement of *BABAR* and Belle in 1999 was that ALEPH and OPAL at LEP and CDF at the Tevatron each had a non-significant $\sin(2\beta)$ measurement, while CLEO-II and to a certain extend CDF had the most precise branching ratio measurements and the best limits on rare B decays.

The next generation of heavy meson experiments besides *BABAR* included, Belle at KEK, CDF/D0 Run-II at the Tevatron, CLEO-III at Cornell and HERA-B at DESY.

However, HERA-B encountered problems, most pronounced with their tracking chambers, which was the main reason why the first level trigger was not able to extract the $B^0 \to J/\psi X$ decays from the minuscule (10^{-6}) fraction of minimum bias $B\overline{B}$ events available, thus yielding only a handful of signal candidates [HERA-B03].

At CLEO-III it was decided to focus on D physics, as the luminosity of CESR would not be able to compete with those of KEK-B and PEP-II, and since neither the detector nor the collider were build asymmetrically to allow for a boosted system.

Run-II at the Tevatron did not start until 2001, and the luminosity has been below expectations. Furthermore, the CDF and D0 detectors have not been working optimally, and so the first results on B physics have yet to be published[5]. However, as the Tevatron is a hadron collider (where the energy of the collisions vary), the B_s meson will also be available for analysis.

This has left Belle as the fiercest competitor. The two detectors are roughly similar in performance, with Belle having slightly better tracking and electromagnetic calorimetry (and lower beam backgrounds), while *BABAR* has better particle identification. Belle started data taking at the same time as *BABAR*[6], and though *BABAR* managed to get ahead with respect to integrated luminosity, Belle has now taken the lead. Interestingly enough, the two datasets

[4] In leptogenesis the particle antiparticle-asymmetry is created through CP violation in the lepton sector, contrary to baryogenesis, where the quark sector is the source of the asymmetry.

[5] A measurement of the mass difference between $D_s^\pm$ and $D^\pm$ was the first published Run-II result [II03].

[6] Almost! *BABAR* saw their first collisions 26th of May 1999, while Belle saw theirs the 4th of June 1999.

have never been further than 25% apart in size, and thus the competition has remained very equal and respectful, even bordering to friendly.

The goal of *BABAR* is to produce at least 5×10^8 $B\overline{B}$ pairs in order to constrain the CKM-matrix as much as possible. A precise $\sin(2\beta)$ measurement is one of the primary goals, but constraints on or measurements of $\sin(2\alpha)$, $\sin(2\beta+\gamma)$ and $\sin(\gamma)$ are also planned. Additionally, V_{ub} and V_{cb}, rare decays, semileptonic and radiative penguins, hadronic B and D physics and Quantum Electro Dynamics (QED) with τ decays are studied.

Part I
Physics

Man masters Nature not by force but by understanding. That is why science has succeeded where magic failed: Because it has looked for no spell to cast on Nature.

[Jacob Bronowski, 1908-1974]

3 The CKM matrix and CP violation

Extending quark mixing to three generations, established by the b quark discovery, gives rise to a CP violating phase, which provides an elegant explanation for the well-established CP violation first observed in the K system.

However, the crucial test of whether CP violation is described by the Cabbibo-Kobayashi-Maskawa (CKM) quark mixing matrix and thus the Standard Model (SM) alone, requires the scrutiny of combining many independent measurements. This gives the B system, rich in phenomenology and the only direct probe of the third quark generation[7], a central role, to be investigated in the following.

First, the CKM matrix will be accounted for on the basis of the Standard Model and the symmetries of the theory investigated. Secondly, the various parametrizations of the CKM matrix and CP violation in general will be discussed. This section is mainly based on [PS95], [Har98] and [Nir02] unless explicitly stated.

3.1 The origin of the CKM matrix

The Standard Model, so far accurately describing all experimental data in particle physics, is based on an $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge theory. The Standard Model Lagrangian is the most general possible, which is renormalisable and consistent with this gauge group.

In order to write gauge invariant mass terms, i.e. assign masses to the particles of the model, one has to spontaneously break the $SU(2)_L \times U(1)_Y$ gauge symmetry. This is done by introducing a scalar field[8], ϕ, which assumes a non-zero vacuum expectation value (VEV), $\langle\phi\rangle = (0, v/\sqrt{2})$. The Lagrangian of the Yukawa couplings, Y, between this Higgs field and the quarks take the form:

$$\mathcal{L}_{\text{Yukawa}} = -Y^u_{ij}\,\overline{q}'^i_L\,\phi\,u'^j_R - Y^d_{ij}\,\overline{q}'^i_L\,\phi^\dagger\,d'^j_R + \text{h.c.}, \qquad \text{where} \quad q^i_L \equiv \begin{pmatrix} u^i_L \\ d^i_L \end{pmatrix}. \tag{3.1}$$

Here, i is the flavor index, L and R the handedness and the prime signifies that the quarks are considered in the weak interaction (flavor) basis, this being *defined* by Eq. 3.1.

The Yukawa couplings, Y, do *not* follow from a gauge principle and they are in general complex-valued matrices. Replacing the Higgs couplings by their VEVs, $M^q_{ij} = (v/\sqrt{2})Y^q_{ij}$, gives rise to mass terms, but the corresponding mass eigenstates will in general *not* be the same as the flavor eigenstates and in fact they are not! By diagonalizing this mass matrix:

$$M^{\text{diag}}_q = V_{qL} M_q V^\dagger_{qR} \qquad (q = u, d), \tag{3.2}$$

we obtain by definition the mass basis, where the entries in the diagonal are real and positive. The mass basis is related to the flavor basis by the transformation:

$$q_{Li} = (V_{qL})_{ij} q'_{Lj}, \quad q_{Ri} = (V_{qR})_{ij} q'_{Rj} \qquad (q = u, d), \tag{3.3}$$

When writing the Lagrangian in the mass basis, we introduce four unitary matrices V_{qL} and V_{qR} into the theory. The right-handed fields all have identical couplings to the gauge fields, thus V_{qR} commute with the covariant derivatives, and they disappear from the theory. The same argument is true for the left-handed fields when considering the $SU(3)_C$ couplings of Quantum Chromo Dynamics (QCD) and the purely kinetic terms, thus only the $SU(2)_L \times U(1)_Y$ couplings remains. In the ElectroMagnetic (EM) couplings they also vanish, as can be seen from a typical term $\bar{u}'^i_L \gamma^\mu u'^i_L \to \bar{u}^i_L (V_{uL})^\dagger_{ij} \gamma^\mu (V_{uL})_{jk} u^k_L = \bar{u}^i_L \gamma^\mu u^k_L$ and equivalently for the Z^0 terms. This is experimentally well founded in that flavor changing neutral currents are not observed at tree level.

[7]The top quark lifetime is too short for hadrons to form, thus no mesons involving top quarks exist.

[8]Other mechanisms for generating mass exists, but they are disfavoured by observation [DRK81, KKW93].

The only place where the V_{qL} matrices do not cancel is in the charged current interactions mediated by the $W^{\pm}$ bosons. In the mass basis the Lagrangian for quarks takes the form:

$$\mathcal{L}_W^q = \frac{g}{\sqrt{2}}\bar{u}_{Li}\gamma^{\mu}(V_{uL}V_{dL}^{\dagger})_{ij}d_{Li}W_{\mu}^{+} + \text{h.c.}. \tag{3.4}$$

The unitary[9] 3×3 matrix, which link the three u_{Li} quarks with the three d_{Li} quarks:

$$V_{\text{CKM}} \equiv V_{uL}V_{dL}^{\dagger}, \tag{3.5}$$

is called the *Cabibbo-Kobayashi-Maskawa (CKM) quark mixing matrix* [Cab63, KM73]. Ordering the couplings by the mass hierarchy yields the entries:

$$V_{\text{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{3.6}$$

As a result of the fact that the physical (mass) eigenstates are not the same as the weak (flavor) eigenstates, one has to introduce a rotation matrix between the two. This is the only source of flavor changing quark interactions, within the Standard Model (SM).

Given N_g generations of particles, the CKM matrix has a priori $2N_g^2$ degrees of freedom (dof.). Out of these N_g^2 dof. are removed by the unitarity condition, and $2N_g - 1$ dof. (number of *relative* phases between quarks) can be absorbed by redefining the quark fields. Thus, one is left with $N_{\text{dof}} \equiv (N_g - 1)^2$ degrees of freedom. In a world with only two generations, $N_{\text{dof}} = 1$, which is just a rotation angle in a real unitary 2×2 matrix. The SM has three generations, thus $N_{\text{dof}} = 4$, which are the three (Euler) angles of a 3×3 rotation matrix and one complex phase[10]. *It is this irreducible complex phase that gives rise to CP violation in the SM.* It is interesting to note, that a third generation of quarks was originally proposed *exactly to provide* a mechanism for CP violation [KM73].

3.1.1 Mixing among leptons

As the CKM matrix arises from breaking the $SU(2)_L \times U(1)_Y$ gauge symmetry, resulting in a mixing matrix between quarks, an equivalent matrix should exist for the leptons, as they also couple to the weak force.

However, if neutrinos have no mass and consequently do not couple to the Higgs particle, the mass matrix will commute with the weak eigenstates, and there will be no mixing matrix for the lepton sector. Since neither neutrino masses nor lepton flavor violation had been observed, it was originally thought that a mixing matrix and CP violation was absent in the lepton sector.

But after the discovery of atmospheric neutrino mixing by Super-Kamiokande [SK98], a mixing matrix for leptons is required and a new source of CP violation is introduced[11]. Unlike in the quark sector, neutrinos can be their own antiparticles (Majorana neutrinos), which results in three complex phases instead of one.
The entries of the lepton mixing matrix, named the Maki-Nakagana-Sakata (MNS) matrix [MNS62], are limited by statistics and in general poorly known. However, this will change in the coming years, and once ν factories start running, (over-)constraining the MNS will surely be pursued.

[9] Local gauge invariance and baryon number conservation requires V_{CKM} to be unitary.
[10] In general, N_g generations will give $\frac{1}{2}N_g(N_g - 1)$ rotation angles and $\frac{1}{2}(N_g - 1)(N_g - 2)$ phases.
[11] Given current data, CP violation in the lepton sector will be insignificant and disregarded in the following.

3.2 *C*, *P* and *T* transformations

Three discrete operators are potential symmetries of field theories. Two are space-time operators (T, P), which are part of the Poincaré group, while the third is an intrinsic transformation (C):

- C – Charge conjugation. Transforms a particle into its antiparticle, $P(\vec{p},\vec{s}) \to \overline{P}(\vec{p},\vec{s})$.
- P – Parity change. Flips the parity (handedness) of space, $P(\vec{p},\vec{s}) \to P(-\vec{p},\vec{s})$.
- T – Time reversal. Interchanges the forward and backward lightcone, $P(\vec{p},\vec{s}) \to P(-\vec{p},-\vec{s})$.

In an unbroken gauge theory, C, P and T are all conserved separately, and therefore the strong and the EM force respect these symmetries. The weak force, however, being of chiral (i.e. handed) form, maximally violates C and P. This has its reason in that each of the two operators interchange particles which couple to the weak force (left-handed particles and right-handed antiparticles) with particles that don't (left-handed antiparticles and right-handed particles). The combined symmetry CP is not violated by flavor mixing in itself, but by the occurrence of an irreducible complex couplings in the Lagrangian, as is the case for the CKM matrix:

$$\frac{g}{\sqrt{2}}\left[\bar{u}_i\gamma^\mu V_{ij}d_jW_\mu^+ + \bar{d}_i\gamma^\mu V_{ij}^\dagger u_jW_\mu^-\right] \stackrel{\text{CP}}{\longrightarrow} \frac{g}{\sqrt{2}}\left[\bar{u}_i\gamma^\mu V_{ij}^* d_jW_\mu^+ + \bar{d}_i\gamma^\mu V_{ij}^T u_jW_\mu^-\right]. \quad (3.7)$$

Thus, for any Lagrangian with complex couplings, CP will generally not be a good symmetry. From Nöther's Theorem[12], CP symmetry induces a quantum number, CP. Invariance under CP implies that this quantum number is conserved in all reactions.

3.2.1 Requirements for *CP* violation and the Jarlskog parameter

CP is not necessarily violated given three generations. In addition one requires that none of the masses of same-type quarks are equal, that none of the three mixing angles are 0 or $\pi/2$ (as the complex phase could then be removed by phase transformations), and that the phase is not 0 or π. The last eight conditions are elegantly summarized in a parameter-independent manner by Jarlskog [Jar85]:

$$Im(V_{ij}V_{kl}V_{il}^*V_{kj}^*) = J\sum_{m,n=1}^{3}\epsilon_{ikm}\epsilon_{jln}, \quad J \neq 0. \quad (3.8)$$

The maximal value J can assume is $1/(6\sqrt{3}) \simeq 0.1$, but the value suggested by data is $J = (3.11^{+0.33}_{-0.46}) \times 10^{-5}$, thus many orders of magnitude smaller, which gives meaning to the statement that CP violation in the quark sector of the SM is small.

3.2.2 The strong *CP* problem

Even with the powerful restrictions of renormalizability and gauge invariance, one can in principle add terms to the Lagrangian, which violate P and T. For the $SU(2)_L$ and $U(1)_Y$ fields this has no observable effect, but for $SU(3)_C$ it implies an electric dipole moment (EDM) for the neutron (a T violating effect), which has been experimentally excluded to an impressive precision [H$^+$99]. Though originating from non-pertubative QCD effects, inherently incalculable, an $\mathcal{O}(1)$ effect is expected, while the EDM limit constrains the coupling to less than $\sim 10^{-10}$, which is unnaturally small.
To solve this problem, known as the strong CP problem, one has either to constrain the Higgs

[12] For each symmetry there exist exactly one corresponding conservation law [Noe31].

couplings or add additional structure to the Higgs sector [PQ77][13], which is not in the spirit of the very general assumptions used in Section 3.1.
In the following, the strong force is assumed to be CP conserving.

3.3 Parametrizations of the CKM matrix

The four parameters of the CKM matrix can be parametrized in many ways of which two will be presented here. The first one is exact and resembles ordinary rotation matrices much. The second is only approximate, but emphasizes the hierarchy of the mixing matrix and is therefore more intuitive.

3.3.1 Standard parametrization

The most obvious way of expressing the CKM matrix is in terms of three rotation angles, $\theta_{12}, \theta_{13}, \theta_{23}$ and one complex phase, δ_{13} as follows [GKR02][14]:

$$V_{\mathrm{CKM}} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix} \quad (3.9)$$

where $c_{ij} \equiv \cos(\theta_{ij})$, $s_{ij} \equiv \sin(\theta_{ij})$ and $i, j = 1, 2, 3$ label the three generations. An advantage of this representation is that if an angle, θ_{ij}, vanishes, then so does the mixing between the generations i and j. The form can be much simplified by using the smallness of $|V_{ub}| \sim 0.003$, so that to an excellent approximation ($\mathcal{O}(10^{-5})$) $c_{13} \simeq 1$ and s_{13} can be neglected compared to terms of order unity.

3.3.2 Wolfenstein parametrization

It happens that the diagonal terms in the CKM matrix are close to unity, while the couplings between different generations decrease as one moves away from the diagonal. No (known) principle dictates this hierarchy (which is not present in the MNS matrix), but along with the mass hierarchy, it is suggestive.

One can choose to emphasize this hierarchy by writing the entries in powers of $\lambda \equiv s_{12} = 0.2229 \pm 0.0022$ [PDG02], the sine of the Cabibbo angle [Cab63], as follows:

$$V_{\mathrm{CKM}} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4), \quad (3.10)$$

where A, ρ and η are real numbers intended to be of order unity. This approximation is called the Wolfenstein parametrization [Wol83], and it is valid up till fourth order of λ, which is accurate enough for most purposes ($\mathcal{O}(10^{-3})$).

When comparing two decay modes, one is said to be Cabibbo suppressed with respect to the other, if the CKM-couplings are one order higher in λ. The branching fraction of the Cabibbo suppressed decay will roughly be $\lambda^2 \simeq \mathcal{O}(5\%)$ of the non-suppressed decay. This notion will be used extensively in later discussions.

[13] This introduces a goldstone boson (the axion), which by now is almost excluded by experiment.
[14] Kobayashi and Maskawa [KM73] originally proposed a parametrization closer to that of Cabibbo [Cab63].

3.4 The unitary triangle(s)

The nine constraints on the CKM matrix from unitarity (see Section 3.1) can be written explicitly as:

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} V_{ud}^* & V_{cd}^* & V_{td}^* \\ V_{us}^* & V_{cs}^* & V_{ts}^* \\ V_{ub}^* & V_{cb}^* & V_{tb}^* \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3.11}$$

Three of these nine linearly independent equations are normalizations and six are linear complex equations written below, divided according to their origin with respect to generation:

$$\begin{array}{ccc} \underline{1\otimes 2 \text{ generation}} & \underline{1\otimes 3 \text{ generation}} & \underline{2\otimes 3 \text{ generation}} \\ (\mathcal{O}(\lambda)+\mathcal{O}(\lambda)+\mathcal{O}(\lambda^5)=0) & (\mathcal{O}(\lambda^3)+\mathcal{O}(\lambda^3)+\mathcal{O}(\lambda^3)=0) & (\mathcal{O}(\lambda^4)+\mathcal{O}(\lambda^2)+\mathcal{O}(\lambda^2)=0) \\ V_{ud}V_{us}^*+V_{cd}V_{cs}^*+V_{td}V_{ts}^*=0 & \boxed{V_{ud}V_{ub}^*+V_{cd}V_{cb}^*+V_{td}V_{tb}^*=0} & V_{us}V_{ub}^*+V_{cs}V_{cb}^*+V_{ts}V_{tb}^*=0 \\ V_{ud}V_{cd}^*+V_{us}V_{cs}^*+V_{ub}V_{cb}^*=0 & V_{ud}V_{td}^*+V_{us}V_{ts}^*+V_{ub}V_{tb}^*=0 & V_{cd}V_{td}^*+V_{cs}V_{ts}^*+V_{cb}V_{tb}^*=0 \end{array} \tag{3.12}$$

These six equations may be represented by triangles in the complex plane, were each of the three complex VV^* terms correspond to a side in the triangle. The triangle representation has the advantage that it is invariant under phase redefinition[15] and therefore physically observable. In fact the sides and angles can be directly related to branching ratios and asymmetries, respectively, and consequently the geometrical interpretation is natural.

The six triangles have the same area, equal to $|J|/2 \sim \lambda^6$, which expresses that there is precisely one CP violating phase. They are pairwise much alike, due to the high degree of symmetry of the CKM matrix, and therefore it is customary to sketch only three triangles, as in Fig. 3.1.

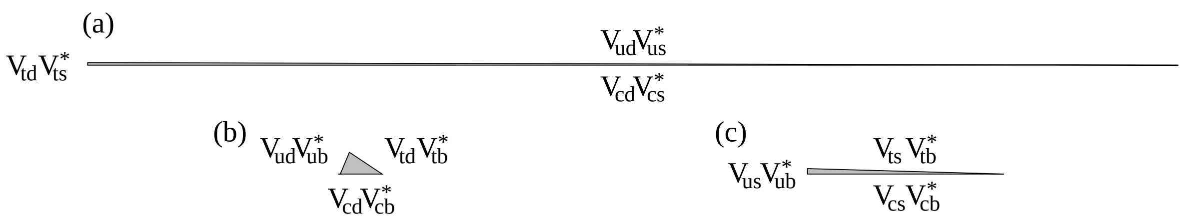


Figure 3.1: The three unitary triangles are geometric interpretations of the three equations $\sum_{q=u,c,t} V_{qa}V_{qb}^* = 0$ with (a) $a=d, b=s$ $(1\otimes 2)$, (b) $a=d, b=b$ $(1\otimes 3)$, (c) $a=s, b=b$ $(2\otimes 3)$. The triangles have equal areas and are here approximately drawn to scale in order to show their shape.

From Figure 3.1 and from the order in λ of the sides, it is clear that only the unitary condition from combining first and third generation couplings of the CKM matrix result in triangles with comparable sides and angles, while the others are "squashed" and therefore very hard to measure.

For historical reasons the triangle corresponding to the equation $V_{ud}V_{ub}^*+V_{cd}V_{cb}^*+V_{td}V_{tb}^*=0$ (framed in Eq. (3.12)), is known as *the* unitary triangle (see Fig. 3.2). None of the side lengths or angles of the second $1\otimes 3$ triangle differs from the unitary triangle by more than 5 %, and it will therefore be a while before the two can be experimentally distinguished.

For the Wolfenstein parametrization to match the graphical representation more precisely (i.e. for the apex of the triangle to coincide with (ρ,η)), one makes the substitution: $\rho \to \overline{\rho} = \rho(1-\lambda^2/2)$ and $\eta \to \overline{\eta} = \eta(1-\lambda^2/2)$.

[15] It is such phase redefinitions that remove $2N_g - 1$ degrees of freedom when counting these (Section 3.1).

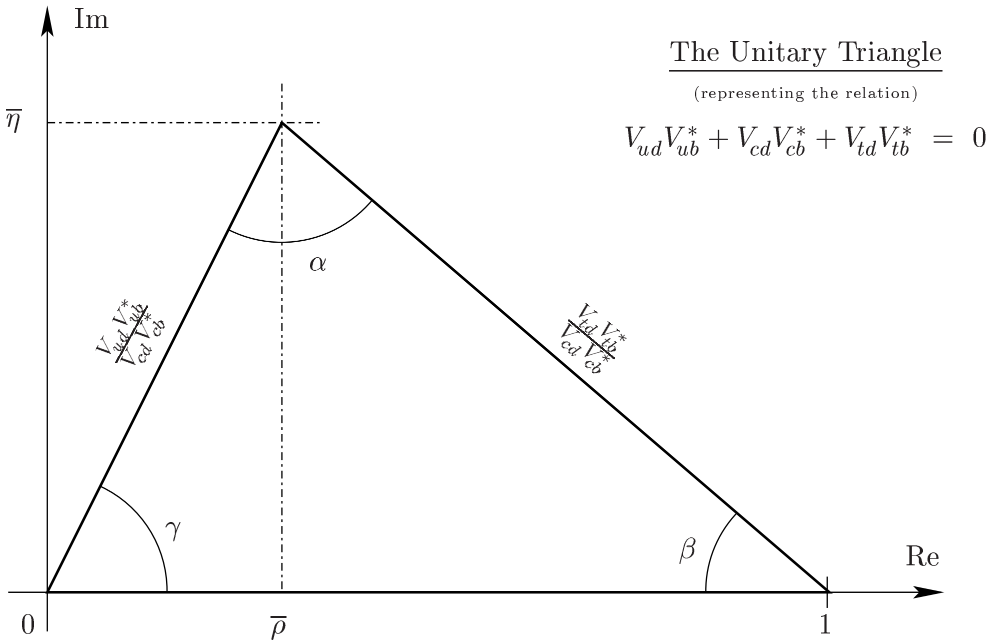


Figure 3.2: The unitary triangle. The triangle corresponding to the equation $V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0$, is for historic reasons known as *the* unitary triangle. Note that the baseline has been normalized by $V_{cd}V_{cb}^*$ to equal one.

The three angles of the unitary triangle expressed in terms of CKM matrix entries are[16]:

$$\alpha \equiv \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right), \qquad \beta \equiv \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right), \qquad \gamma \equiv \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right). \tag{3.13}$$

These physical observables can be measured by CP asymmetries in various B decays, and consistency of the measurements provides in part tests of the SM. The real test, though, is to fit all experimental and theoretical CKM related quantities with *only four parameters.*

3.4.1 The CKM paradigme and physics beyond the Standard Model

> *The Kobayashi-Maskawa phase is, very likely, the dominant source of CP violation in low-energy flavor-changing processes.*
>
> [Y. Nir]

Whereas most other theories have failed to explain CP violation, the Kobayashi-Maskawa mechanism has prevailed. Not only does it explain the very origin of CP violation in a very simple fashion, but as the complex phase only enters when all three generations are involved, it also correctly predicts its smallness in most processes. In light of the results from the B factories, the quotation above is commonly agreed upon, and the CKM mechanism has risen from an explanation to a paradigme.

However, that said, one should keep in mind that the Standard Model fails by many orders of magnitude to explain the baryon asymmetry, $N_B/N_\gamma = (5.5 \pm 0.5) \times 10^{-10}$ [BNT01] observed in the Universe [HS95, GHO$^+$94], and therefore sources of CP violation beyond that of the CKM matrix *must* exist, either in the quark sector or elsewhere (e.g. the lepton sector).

Since the B system contains the dominant part of the SM CP violation and involves many loop processes, which could involve physics beyond the SM and thus alter its predictions, it is an excellent place to search.

[16] Following the original terminology, the angles α, β and γ are denoted ϕ_2, ϕ_1 and ϕ_3 respectively by Belle.

Finally, the Higgs sector may itself contain CP violation, if there are more than one doublet[17], possibly through spontaneous CP violation [Nir02]. Interestingly, CP is an exact gauge symmetry in some string theories, and must be spontaneously broken [DLM92, CKN93]. However, in the following only the SM will be considered, and the above cases will not be discussed any further.

3.5 Section summary and conclusions

Given three generations, the Standard Model not only encompass CP violation, but also correctly predicts its natural smallness. Though well understood (and well tested), some aspects of CP violation remains to be solved. However, the major reason for the interest comes from cosmological arguments.
CP violation is required for the observed matter anti-matter asymmetry in the Universe to arise, but the experimentally determined level of CP violation falls many orders of magnitude short of the required level. This means that sources of CP violation beyond the Standard Model *must* exist.

A search for such sources must begin with a mapping of the SM sources. These are most pronounced in the third quark generation, and since the top quark does not form bound states, the bottom quark systems take a central position. Not only do these decays have loop diagrams, which could involve new physics, but being recently available in very large numbers and for time-dependent measurements, the stage is set for experimentalists to test whether the Standard Model can account for all CP violation measurements.

[17]Then ϕ^u and ϕ^d are not the same (separate Higgs' for the u_i and d_i quarks).

4 Mixing and CP violation in the B system

[Effective field theories] will make hard calculation easy and impossible calculations doable ... If we had to know everything about all the particles, no matter how heavy, we would never get anywhere.

[H. Georgi, Weak Interactions and Modern Particle Theory]

Though abundant in many hadrons, b quarks are produced and studied most easily in the two light B mesons. The neutral B system is complicated by the fact that the flavor eigenstates, with definite quark composition $\bar{b}d$ (B^0 meson) and $b\bar{d}$ ($\overline{B}^0$ meson)[18], are not the same as the physical states of definite masses and lifetimes, which in turn are not the same as the CP eigenstates.

However, it is this complication that gives B physics such a rich phenomenology.

Starting from the Hamiltonian, the various aspects of B physics are explored. First the mixing between the flavor eigenstates as they evolve in time is infered, next the CP eigenstates are time developed, and finally various types of CP violation are considered. Though aimed at the B^0 system, most of the formalism applies to other neutral meson systems, as will be mensioned occasionally and discussed in Section 6.1. The discussion is model independent unless otherwise stated.

4.1 The B-system

The b quark is like other quarks bound in hadrons, but as b baryons are harder to produce, less calculable and phenomelogically inferior, the focus has been on B mesons, of which four types (omitting exited states) exist (see Table 4.1).

B meson	Mass (MeV)	Lifetime (ps)	Mass difference (ps^{-1})	Lifetime difference (ps)
$B_d^0 \equiv \bar{b}d$	5279.4 ± 0.5	1.542 ± 0.016	0.489 ± 0.008	> 0.084 (90% CL)
$B_u^+ \equiv \bar{b}u$	5279.0 ± 0.5	1.674 ± 0.018	–	–
$B_s^0 \equiv \bar{b}s$	5369.6 ± 2.4	1.461 ± 0.057	> 13.1 (95% CL)	> 0.29 (95% CL)
$B_c^+ \equiv \bar{b}c$	6400 ± 400	$0.46^{+0.18}_{-0.16}$	–	–

Table 4.1: The B mesons and measurements of their masses, lifetimes, mass differences and lifetime differences [PDG02], [RNCMV04] and [PDG03]. The two latter only applies to neutral mesons. The lifetime differences are predicted to be -0.005 ± 0.002 [DHKY02] and 0.18 ± 0.09 [A$^+$01].

The main advantage of the two lightest B mesons is that they can be produced in large numbers with relatively small backgrounds, which is why they are the first to be studied in detail. The B_s^0 meson is also very interesting, but has to be produced from $\Upsilon(5S)$ decays (possible at the B factories), Z^0 decays (LEP) or at hadron colliders (Tevatron/LHC), which have so far been restricted the sample sizes.

4.2 Effective field theories

Though particle physics is based on Quantum Field Theory (QFT), effective field theory is the framework in which the physics of B mesons is described, as only the relevant degrees of freedom need to be taken into account.

[18] The notation convention is that the B^0 is the isospin partner of the B^+, which contains a $\bar{b}$ quark.

Integrating out the irrelevant parts leaves one with an effective Hamiltonian, $\hat{H}$, the interaction of which for B^0 mesons has to be time developed on a general state:

$$|\psi(t)\rangle \quad = \quad a(t)|B^0\rangle + b(t)|\overline{B}^0\rangle + \sum_f \int_{PS} \rho_f(E) c_f(E,t)|f(e)\rangle dE, \tag{4.1}$$

where a, b and c are coefficients of the states and one sums over energy-eigenstates, f, and integrate their density, $\rho_f(E)$, over phase space (PS). This leads to a set of coupled first order differential equations, which can be solved for an initial state containing a superposition of B^0 and $\overline{B}^0$ (thus $c_f(0) = 0$), by making the approximation (originally due to Wigner and Weisskopf [WW30]) that the weak force is a perturbation relative to the strong and EM forces, and consequently the time-scale of its reactions governing the evolution of the system much longer. This means that weak Hamiltonian does not couple the final states, f, to each other. The result is a pair of coupled first order differential equations [CK03]:

$$i\frac{d}{dt}\left[\begin{array}{c} a(t) \\ b(t) \end{array}\right] \quad = \quad \mathbf{H}\left[\begin{array}{c} a(t) \\ b(t) \end{array}\right]. \tag{4.2}$$

This Schrödinger-like equation describes, in a simple manner, the time evolution of an arbitrary $B^0/\overline{B}^0$ state. As the relative time-scales between the weak and the strong interactions differ by 8–10 orders of magnitude(!), the approximation made is good for all practical purposes. The actual entries of $\mathbf{H}$ remains to be calculated, which still requires QFT, or effective theories based thereon.

4.3 The Hamiltonian

The Hamiltonian for a system of mixing and decaying particles, is in general not Hermitian. However, it can be uniquely decomposed to a sum of two Hermitian matrices; the dispersive part, $\mathbf{M}$, describing the mixing, and the absorptive part, $\mathbf{\Gamma}$, describing the decay:

$$\mathbf{H} \equiv \mathbf{M} - \frac{i}{2}\mathbf{\Gamma} = \left(\begin{array}{cc} M_{11} & M_{12} \\ M_{21} & M_{22} \end{array}\right) - \frac{i}{2}\left(\begin{array}{cc} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{array}\right) = \left(\begin{array}{cc} \langle B^0|\hat{H}|B^0\rangle & \langle B^0|\hat{H}|\overline{B}^0\rangle \\ \langle \overline{B}^0|\hat{H}|B^0\rangle & \langle \overline{B}^0|\hat{H}|\overline{B}^0\rangle \end{array}\right), \tag{4.3}$$

where the indices $1, 2$ refer to whether the initial and final states were B^0 or $\overline{B}^0$. The dispersive part, $\mathbf{M}$, governs mixing through virtual states while the absorptive part, $\mathbf{\Gamma}$, governs mixing via and decay to real states.

The diagonal terms of the Hamiltonian matrix $\mathbf{H}$ are mostly due to the b-quark mass and strong interactions, whereas the non-diagonal terms are generated by the transitions between B^0 and $\overline{B}^0$ states, thus governed by the weak interaction. For this reason the diagonal terms H_{11} and H_{22} dominate.

4.3.1 Phase conventions

The two B states, B^0 and $\overline{B}^0$, are related through CP transformation, as follows:

$$CP|B^0\rangle = e^{i\xi_1}|\overline{B}^0\rangle, \quad CP|\overline{B}^0\rangle = e^{i\xi_2}|B^0\rangle, \tag{4.4}$$

where the phases $\xi_{1,2}$ are *arbitrary*, i.e. without physical meaning, so a phase transformation:

$$|B^0_\zeta\rangle = e^{-i\zeta}|B^0\rangle, \quad |\overline{B}^0_\zeta\rangle = e^{+i\zeta}|\overline{B}^0\rangle, \tag{4.5}$$

has no physical effects, which is a consequence of the strong forces (b) flavor independence. Without loss of generality one can choose $(CP)^2 = 1$ and likewise $(CPT)^2 = 1$.
One must construct quantities, which do not change under phase rotations (i.e. convention-independent), to measure physically meaningful phases (see Eq. 4.22).

4.3.2 Hamiltonian properties under *CPT* and *CP* transformation

The discrete symmetries C, P and T (see Section 3.2) pose certain constraints on the Hamiltonian. Though none of them are good symmetries by themselves, certain combinations are instructive to consider, as they either constrain the Hamiltonian or reveal which consequences the breaking of that symmetry has.

Invariance under CPT requires that $(CPT)^\dagger H(CPT) = H$, and therefore:

$$\langle B^0_{\text{out}}|H|B^0_{\text{in}}\rangle \overset{CPT}{=} \langle B^0_{\text{out}}|(CPT)^\dagger H(CPT)|B^0_{\text{in}}\rangle \overset{\text{def.}}{=} \langle \overline{B}^0_{\text{in}} e^{-i\nu}|H|e^{i\nu}\overline{B}^0_{\text{out}}\rangle \;\Rightarrow\; H_{11} = H_{22}, \quad (4.6)$$

and from the Hermeticity of $\mathbf{M}$ and $\mathbf{\Gamma}$ follows that $M_{11} = M_{22}$ and $\Gamma_{11} = \Gamma_{22}$ are real ($\in \mathbb{R}$). CPT invariance ensures that particles and antiparticles have equal masses and decay rates. CPT invariance is a general requirement of relativistic field theory[19], and it is believed to be an exact symmetry. As CPT invariance has been tested to extremely high precision, and in many ways [PDG02, P$^+$01] it will be assumed in the following.
CP conservation also yields this result, but in addition it requires that:

$$\langle B^0_{\text{out}}|H|\overline{B}^0_{\text{in}}\rangle \overset{CP}{=} \langle B^0_{\text{out}}|(CP)^\dagger H(CP)|\overline{B}^0_{\text{in}}\rangle \overset{\text{def.}}{=} \langle \overline{B}^0_{\text{out}} e^{-i\xi}|H|e^{-i\xi} B^0_{\text{in}}\rangle \;\Rightarrow\; H_{12} = e^{-2i\xi} H_{21}, \quad (4.7)$$

Along with the Hermiticity requirements, this implies that the off-diagonal elements of the Hamiltonian, H_{12} and H_{21}, are equal in magnitude. CP invariance requires in addition that *partial* decay rates of CP conjugate processes are equal. However, CP is not an exact symmetry of Nature, and violation of Eq. 4.7 is exactly the general requirement for CP violation.

The consequences of the various symmetries are summarized in Table 4.2. The assumption that CPT is a good symmetry reduces the number of parameters of the effective Hamiltonian from seven (one phase removed by phase transformation) to five.

Symmetry	$B^0 \to B^0$	$\overline{B}^0 \to \overline{B}^0$	$B^0 \to \overline{B}^0$	$\overline{B}^0 \to B^0$	Consequences
CPT	$\overline{B}^0 \to \overline{B}^0$	$B^0 \to B^0$			$H_{11} = H_{22}$
CP	$\overline{B}^0 \to \overline{B}^0$	$B^0 \to B^0$	$\overline{B}^0 \to B^0$	$B^0 \to \overline{B}^0$	$H_{11} = H_{22}$ and $\|H_{12}\| = \|H_{21}\|$
T			$\overline{B}^0 \to B^0$	$B^0 \to \overline{B}^0$	$\|H_{12}\| = \|H_{21}\|$

Table 4.2: The consequences of the symmetries CPT, CP and T.

4.4 Time developement

All composite things decay. Strive diligently. [Buddha, ca. 563 - 483 bc., Last words]

The two neutral B meson states can be represented in three different bases, which in the neutral B system are not aligned:

The flavor basis of definite quark content, $|B^0\rangle$ and $|\overline{B}^0\rangle$, in which particle production is to be understood.

Physical basis defining eigenstates of the Hamiltonian, $|B_H\rangle$ and $|B_H\rangle$ with definite mass, lifetime and decays, which evolve in time according to the Schrödinger Equation.

CP basis states of definite CP. If CP was conserved, then the CP eigenstates would also be mass eigenstates.

[19] By the CPT theorem, it requires Lorentz invariance, Locality, usual spin-statistics connection and a hermitian Hamiltonian of the basic theory [Sch53, Lüd57]. At the gravity scale, Lorentz invariance breaks and some string theories are not local, but these are the few exceptions.

As the Hamiltonian governs the physical eigenstates, the time developement of the flavour and CP eigenstates must be done in the physical basis, which requires the eigenvalues and eigenstates of the Hamiltonian.

4.4.1 Hamiltonian eigenvalues and eigenstates

The eigenvalues λ_L and λ_H of the Hamiltonian are:

$$\lambda_H \equiv H_{11} + \sqrt{H_{12}H_{21}} = (M_{11} - \tfrac{i}{2}\Gamma_{11}) + \sqrt{(M_{12} - \tfrac{i}{2}\Gamma_{12})(M_{12}^* - \tfrac{i}{2}\Gamma_{12}^*)} \equiv M_H - \tfrac{i}{2}\Gamma_H, \quad (4.8)$$

$$\lambda_L \equiv H_{11} - \sqrt{H_{12}H_{21}} = (M_{11} - \tfrac{i}{2}\Gamma_{11}) - \sqrt{(M_{12} - \tfrac{i}{2}\Gamma_{12})(M_{12}^* - \tfrac{i}{2}\Gamma_{12}^*)} \equiv M_L - \tfrac{i}{2}\Gamma_L, \quad (4.9)$$

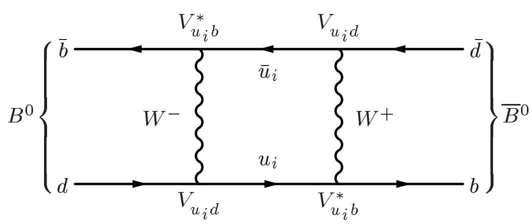
where M and Γ are real. It is customary to introduce the following intuitive definitions:

$$m_B \equiv \tfrac{1}{2}(M_H + M_L), \quad \Delta m_B \equiv m_H - m_L \quad \Leftrightarrow \quad M_{H,L} \equiv m_B \pm \tfrac{1}{2}\Delta m_B, \quad (4.10)$$

$$\Gamma_B \equiv \tfrac{1}{2}(\Gamma_H + \Gamma_L), \quad \Delta\Gamma_B \equiv \Gamma_H - \Gamma_L \quad \Leftrightarrow \quad \Gamma_{H,L} \equiv \Gamma_B \pm \tfrac{1}{2}\Delta\Gamma_B. \quad (4.11)$$

where Δm_B is defined to be positive.

Relating these elements of the Hamiltonian to the basic parameters of the theory through QFT is far from an easy task, which the phenomena of interchanging flavor states, $B\overline{B}$ mixing, governed by Δm_{B_d} (see below), illustrates well. The leading order diagrams for this process are shown in Fig. 4.1:



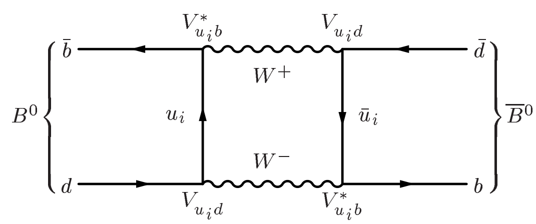


Figure 4.1: Diagrams for $B\overline{B}$ mixing. While the first diagram has a color factor of three, the second diagram has a relative sign difference, so the overall amplitude is twice that of the second diagram. The divergences at large loop momenta are cancelled by the unitarity of the CKM matrix, which is known as the GIM mechanism (see text).

Apart from the relevant CKM factors, the amplitude of the diagrams depends *only* on the masses of the quark involved. If the quark masses were equal, each quark flavor would contribute according to its CKM coupling, and by unitarity they would cancel ($V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0$). In the limit of large loop momenta, the quark masses become negligible, and thus by unitarity the ultra-violet divergences cancel. This is known as the GIM mechanism [GIM70] (named after its inventors Glashow, Illiopoulos and Maiani)[20].

The mass dependence taking all EW contributions into account is described by the Inami-Lim function [IL81]. However, simply from dimensional considerations it can be argued that the mass dependence is linear, thus the top quark by far dominates, and its large mass makes the mixing rate sizable. From measuring the mixing rate, one could therefore imagine extracting $|V_{td}V_{tb}^*|$.

However, while the amplitude of the mixing diagram can be calculated quite accurately, such a calculation treats the b and d quarks as free particles, and therefore ignore the (QCD) forces that bind them into the B^0 meson. The actual B states are not very precisely known,

[20] Interestingly, in the original GIM-scheme it was the unitarity of only two generations that was considered. The idea is by now a general tool to make divergencies cancel, for example in Super Symmetry [SS74, Nil84].

and this lacking knowledge is parametrized in terms of a bag factor, B_{B_d}, and a decay constant, f_{B_d}.
Many schemes have been developed to give more or less precise estimates of these two quantities, e.g. lattice QCD, Heavy Quark Effective Theory (HQET), and sum rules, but none of these can make predictions much better than $\mathcal{O}(20\%)$.

Thus, even though the B_d^0 mixing rate Δm_d has been measured with great precision ($\mathcal{O}(1\%)$), it is still not possible to extract $|V_{td}V_{tb}^*|$ with great precision. Some of the theoretical uncertainties can be cancelled out by considering ratios, such as that of the mixing rates of the B_s^0 and the B_d^0 meson, respectively, but some theoretical uncertainty remains (as does the measurement of Δm_s). With the exception of $\sin(2\beta)$, this is the dominant pattern in extracting CKM parameters.

Although the width difference $\Delta\Gamma_B$ has not been measured (yet), it is expected to be very small, as it arises from decay channels common to B^0 and $\overline{B}^0$, which have branching fractions of $\mathcal{O}(10^{-3})$ or less (either color or Cabibbo suppressed) contributing with alternating sign. Thus in the B system the mass difference is dominating, and it is therefore naturel to label the physical states $|B_L\rangle$ (L for Light) and $|B_H\rangle$ (H for Heavy)[21].

Diagonalizing $\mathbf{H}$ yields the physical eigenstates, which without loss of generality can be written in terms of the flavor eigenstates as:

$$|B_L^0\rangle = p|B^0\rangle + q|\overline{B}^0\rangle, \qquad |B_H^0\rangle = p|B^0\rangle - q|\overline{B}^0\rangle \tag{4.12}$$

where p and q are complex coefficients with arbitrary phase obeying the normalization condition $|p|^2 + |q|^2 = 1$. The eigenvalue equation yields:

$$\frac{q}{p} = -\sqrt{\frac{H_{12}}{H_{21}}} \simeq -\frac{|M_{12}|}{M_{12}}\left(1 - \frac{1}{2}Im\frac{\Gamma_{12}}{M_{12}}\right). \tag{4.13}$$

As both the ratio and the relative phase of Γ_{12}/M_{12} are expected to be small ($\mathcal{O}(m_b^2/m_t^2)$ and $\mathcal{O}(m_c^2/m_b^2)$, respectively), $|q/p|$ will be close to unity, as will be assumed in the following. The same coefficients apply to the two states due to (the assumed) CPT invariance.

The physical states have definite masses and lifetimes and they obey the Schrödinger Equation:

$$|B_{H,L}(t)\rangle = e^{-i\mathbf{H}t}|B_{H,L}\rangle = e^{-(i\mathbf{M}_{H,L}+\Gamma_{H,L}/2)t}|B_{H,L}\rangle. \tag{4.14}$$

From the established Hamiltonian formalism Eq. 4.2 and from inverting Eq. 4.12, the time developement of the flavor eigenstates can be obtained.
Considering initially ($t = 0$) pure flavor states in their Center-of-Mass (CM) frame, the time developed states, denoted $|B^0_{\text{phys}}(t)\rangle$ and $|\overline{B}^0_{\text{phys}}(t)\rangle$, are:

$$|B^0_{\text{phys}}(t)\rangle = g_+(t)|B^0\rangle + \frac{q}{p}\, g_-(t)|\overline{B}^0\rangle, \tag{4.15}$$

$$|\overline{B}^0_{\text{phys}}(t)\rangle = \frac{p}{q}\, g_-(t)|B^0\rangle + g_+(t)|\overline{B}^0\rangle, \tag{4.16}$$

where

$$g_\pm(t) = \tfrac{1}{2}e^{-imt}e^{-\frac{1}{2}\Gamma}(e^{i\frac{1}{2}\Delta mt}e^{\frac{1}{4}\Delta\Gamma} \pm e^{-i\frac{1}{2}\Delta mt}e^{-\frac{1}{4}\Delta\Gamma}). \tag{4.17}$$

Thus, mixing is governed by the phase differences Δm and $\Delta\Gamma$, while the common phases Γ governs decay and e^{-imt} has no physical significance[22].

[21] In the K system (see Section 6.1.1) $\Delta\Gamma_B \ggg \Delta m_B$, hence the terminology K_{short} and K_{long}.

[22] Well, almost! It tests whether gravity attracts or repels particles and antiparticles, which is not the case.

4.4.2 Coherent $\mathbf{B^0\overline{B}^0}$ state

At e^+e^- colliders, B mesons are produced through the $\Upsilon(4S)$ resonance (see Figure 4.2), which is the lightest $b\bar{b}$ state above the threshold for decaying into two B mesons. Due to isospin symmetry, it is in general (and here) assumed that half the B meson pairs are charged and half are neutral.

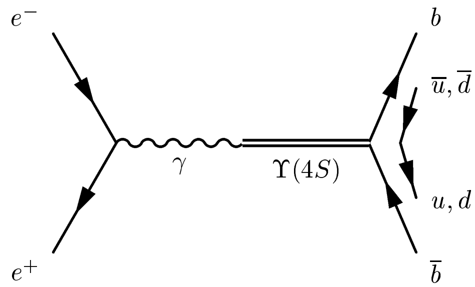


Figure 4.2: Diagram for production of a $B\overline{B}$ pair at the $\Upsilon(4S)$ resonance.

The $\Upsilon(4S)$ has $J=1$, and since the B mesons are pseudo-scalars, conservation of angular momentum requires that the $B\overline{B}$ state function, ψ, is a coherent $L=1$ state, hence with parity $P=(-1)^L=-1$, thus the spatial part is required to be antisymmetric. From Bose-Einstein statistics the overall state must be symmetric, and consequently the flavor-part of ψ must be antisymmetric:

$$\psi_{\text{flavor}} = \frac{1}{\sqrt{2}}\left(|B^0\rangle|\overline{B}^0\rangle - |\overline{B}^0\rangle|B^0\rangle\right). \tag{4.18}$$

At the energies in question the electron mass can be neglected and helicity is conserved, thus the $\Upsilon(4S)$ is produced in a $J_z=\pm1$ state, which (integrating over ψ, which is isotropic) yields the spatial decay distribution:

$$\psi_{\text{space}} = aY_1^1(\theta,\phi) + bY_1^{-1}(\theta,\phi) \quad\Longrightarrow\quad f(\theta) = \int |\psi_{\text{space}}|^2 d\phi = \frac{3}{4}\sin^2(\theta), \tag{4.19}$$

where $|a|^2+|b|^2=1$ are normalization coefficients describing the polarisation of the beams, which is unimportant for this discussion.

The time dependency of the coherent state (Eq. 4.18) is obtained by replacing the flavor states by their time-dependent expressions (Eq. 4.15-4.16):

$$\begin{aligned}
|\psi_{\text{flavor}}(t_1,t_2)\rangle &= \frac{1}{\sqrt{2}}\left(|B^0_{\text{phys}}(t_1)\rangle|\overline{B}^0_{\text{phys}}(t_2)\rangle - |\overline{B}^0_{\text{phys}}(t_1)\rangle|B^0_{\text{phys}}(t_2)\rangle\right)\\
&= \frac{1}{\sqrt{2}}e^{-\frac{1}{2}\Gamma(t_1+t_2)}\left\{\cos\left[(\tfrac{1}{2}\Delta m - i\tfrac{1}{4}\Delta\Gamma)(t_1-t_2)\right]\left(|B^0\rangle|\overline{B}^0\rangle - |\overline{B}^0\rangle|B^0\rangle\right)\right.\\
&\qquad \left. -i\sin\left[(\tfrac{1}{2}\Delta m - i\tfrac{1}{4}\Delta\Gamma)(t_1-t_2)\right]\left(\tfrac{p}{q}|B^0\rangle|B^0\rangle - \tfrac{q}{p}|\overline{B}^0\rangle|\overline{B}^0\rangle\right)\right\},
\end{aligned} \tag{4.20}$$

where the two times, t_1 and t_2, are to be measured in the CM frame of the B mesons, which is close to that of the $\Upsilon(4S)$, as the phase space available is very small ($\gamma^{CM}\simeq 1.002$).

Due to the antisymmetric state, *only the time difference* needs to be measured, as the time sum only enters in the normalization. This is *crucial* for experiments at the $\Upsilon(4S)$ resonance (cf. Section 7.1).

As can be seen from Eq. (4.20), the result of the coherency is that the two B mesons each evolve in time as a single B meson *but in phase*, such that until one B meson decays, there is exactly *one* B^0 meson and *one* $\overline{B}^0$ -meson[23].

[23]This is simply another case of the classic Einstein-Podolsky-Rosen (EPR) situation [EPR35].

However, after the decay of one B meson, the other continues to oscillate. Therefore an event can be either *unmixed*, i.e. the flavor of the two decayed B mesons are opposite, or *mixed*, i.e. the flavor of the two decayed B mesons are the same, as described by Eq. (4.25).

4.4.3 Flavor determination – tagging

To determine the flavor of a decayed B^0 meson (in common terminology, to *tag*[24] it), one needs a decay, which is only allowed by one of the flavors. In principle no such flavor specific decays exist, and the various classifications are almost continuous. Semileptonic decays are to an extremely good approximation flavor specific, but as the associated neutrino escapes undetected, these decays cannot be fully reconstructed, and so they have significant backgrounds.
Hadronic decays can be fully reconstructed and are thus more clean, but they are less flavor specific (by the 'λ^2 theorem' [Pet00]). This has an impact on time-dependent measurements of small CP asymmetries, e.g. in $B^0 \to D^{(*)\pm}\pi^{\mp}$.
However, for the time being (and for illustration) B^0 decays will be divided into two classes:

Flavor decays. If the final state is not allowed (or highly suppressed) for one B flavor, then it is a tag for the other, that is: $\mathcal{A}_{B^0\to f} \gg \mathcal{A}_{\overline{B}^0\to f} \;\Rightarrow\; f$ is a B^0 tag and vice versa.

***CP* decays.** If the final state is allowed for both B flavors, be it a CP eigenstate or not, the decay can be used for CP studies.

While CP decays require to be fully reconstructed, the flavor of a B^0 decay may be infered though missing some tracks and clusters. For experimental reasons (efficiency considerations), most often only one of the two B mesons can be fully reconstructed, while a reasonable estimate can be made about the flavor and decay point of the other decay. Though events where two CP eigenstates are reconstructed are very interesting for solving ambiguities, they will not be considered in the following. Define the amplitudes:

$$\mathcal{A}_f \;\equiv\; \langle f|\hat{H}|B^0\rangle, \qquad \overline{\mathcal{A}}_f \;\equiv\; \langle f|\hat{H}|\overline{B}^0\rangle. \tag{4.21}$$

As discussed in Section 4.3.1, the phase of these amplitudes is arbitrary. However, combining them appropriately with the quantities q and p yields a phase invariant parameter[25], which is unique for each final state f:

$$\lambda_f \;\equiv\; \frac{q}{p}\frac{\overline{\mathcal{A}}_f}{\mathcal{A}_f} \;\stackrel{f=CP}{=}\; \eta_{f_{CP}}\frac{q}{p}\frac{\overline{\mathcal{A}}_{f_{CP}}}{\mathcal{A}_{f_{CP}}}, \tag{4.22}$$

where the last part is defined only if the final f is a CP eigenstate, for which $\eta_{f_{CP}}$ is then the eigenvalue (±1).

If one neutral B meson decays at time t_{rec} into final state f_{rec}, which is fully reconstructed and the other decays at time t_{tag} into a flavor specific state, f_{tag}, where the flavor is determined, then the overall amplitude, $\langle f_{\mathrm{rec}} f_{\mathrm{tag}}|\hat{H}|\psi_{\mathrm{flavor}}(t_{\mathrm{rec}}, t_{\mathrm{tag}})\rangle$, is obtained from Eq. (4.20) by replacing the B states with the corresponding amplitudes. The final (unnormalized) decay time distribution is obtained from the absolute square of this amplitude, and expressed in terms of (u for *unmixed* and m for *mixed*):

$$\mathcal{A}_u = \mathcal{A}_{f_{\mathrm{rec}}}\overline{\mathcal{A}}_{f_{\mathrm{tag}}} - \overline{\mathcal{A}}_{f_{\mathrm{rec}}}\mathcal{A}_{f_{\mathrm{tag}}} \qquad \text{and} \qquad \mathcal{A}_m = \tfrac{p}{q}\mathcal{A}_{f_{\mathrm{rec}}}\mathcal{A}_{f_{\mathrm{tag}}} - \tfrac{q}{p}\overline{\mathcal{A}}_{f_{\mathrm{rec}}}\overline{\mathcal{A}}_{f_{\mathrm{tag}}} \tag{4.23}$$

[24] This expression is used widely in BABAR, both for trivial but also more interesting matters.
[25] The notation is standard but a bit unfortunate, as λ is also used for eigenvalues and as Wolfenstein parameter, but the subscripts and context should disentangle the notations.

which can be regarded as the amplitudes for mixed and unmixed events it yields:

$$e^{-i\Gamma(t_{\rm rec}+t_{\rm tag})}\left[(|\mathcal{A}_u|^2+|\mathcal{A}_m|^2)\cosh(\tfrac{1}{2}\Delta\Gamma t)+2{\rm Re}(\mathcal{A}_u\mathcal{A}_m)\sinh(\tfrac{1}{2}\Delta\Gamma t)\right.$$
$$\left.+(|\mathcal{A}_u|^2-|\mathcal{A}_m|^2)\cos(\Delta m t)+2{\rm Im}(\mathcal{A}_u\mathcal{A}_m)\sin(\Delta m t)\right], \tag{4.24}$$

where $t \equiv t_{\rm rec} - t_{\rm tag}$ have been used[26]. The sin and sinh terms are due to interference between mixed and unmixed states, as they are only present when both $\mathcal{A}_u \neq 0$ and $\mathcal{A}_m \neq 0$.

If $f_{\rm rec}$ is a flavor specific state, then the decay time distribution reduces to:

$$N(f_{\rm rec}, f_{\rm tag})e^{-\Gamma(t_{\rm rec}+t_{\rm tag})}\left[\cosh(\tfrac{1}{2}\Delta\Gamma t) \pm \cos(\Delta m t)\right], \tag{4.25}$$

where the sign is positive for opposite tags ($B^0\overline{B}^0$) and negative for equal tags ($B^0B^0/\overline{B}^0\overline{B}^0$).

If $f_{\rm rec}$is a CP eigenstate, then the decay time distribution is:

$$N(f_{\rm rec}, f_{\rm tag})e^{-\Gamma(t_{\rm rec}+t_{\rm tag})}\left[\frac{1+|\lambda_f|^2}{2}\cosh(\tfrac{1}{2}\Delta\Gamma t) - {\rm Re}(\lambda_f)\sinh(\tfrac{1}{2}\Delta\Gamma t)\right.$$
$$\left.\mp\frac{1-|\lambda_f|^2}{2}\cos(\Delta m t) \mp {\rm Im}(\lambda_f)\sin(\Delta m t)\right]. \tag{4.26}$$

As Δm dominates mixing in the B system, $\Delta\Gamma$ will, though both an interesting and important quantity [Kit03], be omitted in the following.

Anticipating the course of events, the sum of the times is integrated out, as it cannot be measured with any reasonable precision. As both $t_{\rm rec}$ and $t_{\rm tag}$ are greater than zero, $t_{\rm rec} + t_{\rm tag} > |t_{\rm rec} - t_{\rm tag}|$, and one obtains:

$$\begin{aligned} Ne^{-\Gamma(t_{\rm rec}+t_{\rm tag})}f(t_{\rm rec}-t_{\rm tag}) &\rightarrow \int_{|t_{\rm rec}-t_{\rm tag}|}^{\infty} Ne^{-\Gamma(t_{\rm rec}+t_{\rm tag})}f(t_{\rm rec}-t_{\rm tag})d(t_{\rm rec}+t_{\rm tag}) \\ &= Ne^{-\Gamma|(t_{\rm rec}-t_{\rm tag})|}f(t_{\rm rec}-t_{\rm tag}). \end{aligned} \tag{4.27}$$

Thus, by a simple replacement, one obtains decay distributions, which only depend on the (signed) time difference, t, between the two decays.
This simplifies the above expressions, Eqs. (4.25- 4.26), to:

$$N(f_{\rm rec}, f_{\rm tag})e^{-\Gamma\Delta t}\left[1 \pm \cos(\Delta m t)\right] \tag{4.28}$$

$$N(f_{\rm rec}, f_{\rm tag})e^{-\Gamma\Delta t}\left[\frac{1+|\lambda_f|^2}{2} \mp \frac{1-|\lambda_f|^2}{2}\cos(\Delta m t) \mp {\rm Im}(\lambda_f)\sin(\Delta m t)\right]. \tag{4.29}$$

If $f_{\rm rec}$ can be reached by both B^0 and $\overline{B}^0$ but is not a CP eigenstate, the situation is more subtle, and will be discussed in Section 4.6.

4.5 Types of *CP* violation

CP violation can manifest itself in three different ways:

- *CP violation in decay.* The amplitude for a decay (charged or neutral) and its CP conjugate have different magnitudes, that is $|\overline{\mathcal{A}}_{\overline{f}}/\mathcal{A}_f| \neq 1$.
- *CP violation in mixing.* The two mass eigenstates, that are CP conjugates of each other, cannot be chosen to be CP eigenstates, that is $|q/p| \neq 1$.
- *CP violation in interference between decays with and without mixing.* The phase between mixing and decay of neutral CP conjugates is not the same, that is ${\rm Im}(\lambda_{f_{CP}}) \neq 0$.

Each of the three types of CP violation can occur by itself, but they are by no means mutually exclusive.

By phase is meant phase difference, as no phase by itself has any significance as it changes with phase redefinition. Only relative phases can be given any meaning, as they are invariant under phase redefinition (convention independent).

[26] $f_{\rm rec}$is used as reference ($t_{\rm rec}$= 0), which makes the natural and less cluttered notation t instead of Δt.

4.5.1 *CP* violation in decay

CP violation in decay occurs when CP conjugate decay rates differ:

$$\boxed{|\overline{\mathcal{A}}_{\overline{f}}/\mathcal{A}_f| \neq 1 \quad \Longrightarrow \quad CP \text{ violation in } \textbf{decay}} \tag{4.30}$$

and it is a result of interference between two types of phases:

Strong phases, δ, from intermediate on-shell states in the decay, which appear with the same sign in $\mathcal{A}_f$ and the CP conjugated amplitude $\overline{\mathcal{A}}_{\overline{f}}$.

Weak phases, ϕ, from complex parameters (the CKM phase in the SM) in the Lagrangian, which appear with opposite sign (complex conjugated) in $\overline{\mathcal{A}}_{\overline{f}}$.

The amplitude of a process can be written as a sum of amplitudes, one for each diagram contributing to the process:

$$\mathcal{A}_f \;=\; \sum_i |\mathcal{A}_i| e^{i(\delta_i+\phi_i)}, \quad \overline{\mathcal{A}}_{\overline{f}} \;=\; \sum_j |\overline{\mathcal{A}}_j| e^{i(\delta_j-\phi_j)}. \tag{4.31}$$

If only one diagram contributes to an amplitude, the change of sign of the weak phase will not alter the absolute value of this amplitude. CP violation in decay requires that at least two amplitudes with different weak phases acquire different strong phases, which can be seen from considering the difference between the square of the amplitudes:

$$|\mathcal{A}|^2 - |\overline{\mathcal{A}}|^2 \;=\; -2\sum_{i,j} \mathcal{A}_i \mathcal{A}_j \sin(\delta_i - \delta_j)\sin(\phi_i - \phi_j). \tag{4.32}$$

Asymmetries of CP violation in decay defined as:

$$a_f \;=\; \frac{\Gamma(B\to f) - \Gamma(\overline{B}\to \overline{f})}{\Gamma(B\to f) + \Gamma(\overline{B}\to \overline{f})} \;=\; \frac{1-|\overline{\mathcal{A}}_{\overline{f}}/\mathcal{A}_f|^2}{1+|\overline{\mathcal{A}}_{\overline{f}}/\mathcal{A}_f|^2}. \tag{4.33}$$

Such asymmetries are most often measured in charged B decays, as neutral B mesons requires tagging[27], effectively lowering the statistics.

Observing a difference in the number of a final state compared to it's CP conjugate would also constitute evidence for CP violation in decay. However, great care has to be taken that an observed asymmetry is not due to charge asymmetries in the reconstruction.
The possible asymmetries are (contrary to the K system, see Section 6.1.1), expected to be large $\mathcal{O}(10\%)$. However, these channels all have low branching ratios ($\mathcal{O}(10^{-4})$ to $\mathcal{O}(10^{-5})$)[28].

Furthermore, precise calculations of asymmetries are often (but no always) complicated by long distance effects, leading to model dependence and significant theoretical uncertainties, and it is therefore hard to relate any results to the basic parameters of the theory. However, in some cases, when the amplitude of the diagrams involved can all be calculated and/or measured separately, direct asymmetries may provide constraints on the CKM parameters.

CP violation in decay is sometimes refered to as *direct* CP violation, which requires that some of the CP violating phases necessarily appears in the decay ($\Delta F = 1$) amplitude.

[27] This is only true, if the neutral B meson interferes through mixing. There exist (selftagging) modes, for which this is (to a good approximation) not the case (see Section 5.3).

[28] The "uncertainty principle" of direct CP violation: Branching ratio $\times$ asymmetry $\lesssim \mathcal{O}(10^{-6})$ [Jim Smith].

4.5.2 *CP* violation in mixing

CP conservation requires that the physical eigenstates must be CP eigenstates, that is q/p is a pure phase. CP violation in mixing occurs if this is not the case:

$$\boxed{|q/p| \neq 1 \quad \Longrightarrow \quad CP \text{ violation in } \textbf{mixing}} \tag{4.34}$$

that is if the mixing rates between the two eigenstates are not equal.

CP violation was first observed in mixing in the neutral kaon system (see section 6.1.1), and (eventually) measurable asymmetries are expected in the neutral B system, usually in semileptonic decays:

$$a_{sl} = \frac{\Gamma(\overline{B}^0_{\text{phys}}(t) \to X\ell^+\nu) - \Gamma(B^0_{\text{phys}}(t) \to X\ell^-\bar{\nu})}{\Gamma(\overline{B}^0_{\text{phys}}(t) \to X\ell^+\nu) + \Gamma(B^0_{\text{phys}}(t) \to X\ell^-\bar{\nu})} = \frac{1 - |q/p|^4}{1 + |q/p|^4} \tag{4.35}$$

Since $|q/p|$ depends on Γ_{12}, theoretical uncertainties from long distance QCD effects are once again expected to complicate the extraction of CKM parameters from this asymmetry.
To first order only the top quark loop contributes, and q/p is just a pure phase. Contributions from charm quark loops are expected to give $\mathcal{O}(\frac{m_c}{m_t})$ contributions, and so any CP violation in mixing will be of this order.

CP violation in mixing is sometimes refered to as *indirect* CP violation, which requires that all of the CP violating phases can be encompassed in the mixing ($\Delta F = 2$) amplitude.

4.5.3 *CP* violation in the interference between decays with and without mixing

In addition to the conditions $|\overline{\mathcal{A}}_{\overline{f}_{CP}}/\mathcal{A}_{f_{CP}}| = 1$ (no CP violation in decay) and $|q/p| = 1$ (no CP violation in mixing), conservation of CP requires that the relative phase between $\overline{\mathcal{A}}_{\overline{f}_{CP}}/\mathcal{A}_{f_{CP}}$ and q/p vanishes. CP violation in the interference between decays with and without mixing, sometimes abbreviated CP violation in the interference between mixing and decay (or simply CP violation in interference), terms this type of CP violation, and the condition is:

$$\boxed{\text{Im}\,\lambda_{CP} \neq 0 \quad \Longrightarrow \quad CP \text{ violation in } \textbf{interference}} \tag{4.36}$$

This type of CP violation is observed in the time dependent difference between B^0 and $\overline{B}^0$ decays into common final states.

There are two types to distinguish between (assuming no CP violation in mixing and omitting accidental instances):

- If $|\lambda_f| = 1$, one weak phase dominates, thus the difference in Eq. (4.32) is zero and the requirement for CP violation is $\text{Im}\,\lambda_f \neq 0$.
- If $|\lambda_f| \neq 1$, several weak phases interfere, and potentially there could be CP violation in decay.

4.6 Time distribution for non-*CP* final states

If the observed final state f_{rec} (simply denoted f in the following) can be reached by both B^0 and $\overline{B}^0$, but is not a CP eigenstate, four amplitudes and consequently two different amplitude ratios enter the time distributions, as there are now two different final states, which are CP conjugates of each other.
Define the following amplitudes and ratios:

$$\left.\begin{array}{l} A_f = A(B^0 \to f) \\ \overline{A}_f = A(\overline{B}^0 \to f) \end{array}\right\} \; \lambda^{-+} \equiv \frac{q}{p}\frac{\overline{A}_f}{A_f}, \qquad \left.\begin{array}{l} A_{\bar{f}} = A(B^0 \to \bar{f}) \\ \overline{A}_{\bar{f}} = A(\overline{B}^0 \to \bar{f}) \end{array}\right\} \; \lambda^{+-} \equiv \frac{q}{p}\frac{\overline{A}_{\bar{f}}}{A_{\bar{f}}}. \tag{4.37}$$

The reason for the choice of superscript on the λ's will become apparent in Section 5.4. Contrary to the originally defined λ_f (see Eq. (4.22)), the violation of CP is *not* implied by $\lambda^{\pm\mp} \neq 1$, but rather by $\lambda^{-+}\lambda^{+-} \neq 1$.
Given the above definitions, the (unnormalized) time-dependent decay rates can be written as:

$$f(B^0(t) \to f) = Ne^{-\Gamma|t|}\left[1 + \frac{1-|\lambda^{-+}|^2}{1+|\lambda^{-+}|^2}\cos(\Delta mt) - \frac{2\,\mathrm{Im}\,\lambda^{-+}}{1+|\lambda^{-+}|^2}\sin(\Delta mt)\right], \quad (4.38)$$

$$f(\overline{B}^0(t) \to f) = Ne^{-\Gamma|t|}\left[1 - \frac{1-|\lambda^{-+}|^2}{1+|\lambda^{-+}|^2}\cos(\Delta mt) + \frac{2\,\mathrm{Im}\,\lambda^{-+}}{1+|\lambda^{-+}|^2}\sin(\Delta mt)\right], \quad (4.39)$$

$$f(B^0(t) \to \bar{f}) = Ne^{-\Gamma|t|}\left[1 + \frac{1-|\lambda^{+-}|^2}{1+|\lambda^{+-}|^2}\cos(\Delta mt) - \frac{2\,\mathrm{Im}\,\lambda^{+-}}{1+|\lambda^{+-}|^2}\sin(\Delta mt)\right], \quad (4.40)$$

$$f(\overline{B}^0(t) \to \bar{f}) = Ne^{-\Gamma|t|}\left[1 - \frac{1-|\lambda^{+-}|^2}{1+|\lambda^{+-}|^2}\cos(\Delta mt) + \frac{2\,\mathrm{Im}\,\lambda^{+-}}{1+|\lambda^{+-}|^2}\sin(\Delta mt)\right]. \quad (4.41)$$

With no additional weak phases entering the amplitudes, $|\lambda^{-+}\lambda^{+-}| = 1$, and only a phase *difference* between the two implies CP violation. The overall phase of λ^{-+} and λ^{+-} can be written as:

$$\lambda^{-+} = |\lambda^{-+}|\,e^{-i(\delta+\phi)}, \qquad \lambda^{+-} = \frac{1}{|\lambda^{-+}|}\,e^{-i(\delta-\phi)}, \quad (4.42)$$

and thus by measuring the coefficients of the sine terms in Eq. (4.38 – 4.41), the strong phase, δ, can be eliminated and the weak phase, ϕ, extracted (with discrete ambiguities, see Section 5.3).

To shorten the notation and the (otherwise) long expressions in Eqs. (4.38 – 4.41), the following notation is used:

$$C_\pm = \frac{1-|\lambda^{\mp\pm}|^2}{1+|\lambda^{\mp\pm}|^2}, \qquad S_\pm = \frac{2\,\mathrm{Im}\,\lambda^{\mp\pm}}{1+|\lambda^{\mp\pm}|^2}. \quad (4.43)$$

From the assumed (and in the SM given) relation $|\lambda^{-+}\lambda^{+-}| = 1$, one can derive that $C_+ = -C_-$, thus only one parameter, $C \equiv C_+$ (defined to be positive), is required. Furthermore, by defining $\lambda = \min(|\lambda^{+-}|, |\lambda^{-+}|)$, $S_\pm$ can be written as:

$$S_\pm = \frac{2\lambda}{1+\lambda^2}\sin(\delta \pm \phi). \quad (4.44)$$

The parameter λ is the ratio between the interfering amplitudes (traditionally denoted r, see Section 5). The sensitivity to the $S_\pm$ terms and thus to the weak phase, ϕ, increases with λ, and it is therefore *crucial* for the sensitivity that the ratio of amplitudes is sizable. The $B \to DK\pi$ method (see Section 5.2) is based on increasing this ratio.

4.7 *CP* asymmetries and the angles of the unitary triangle

The three angles α, β and γ (aka. ϕ_2, ϕ_1, and ϕ_3, respectively) of the unitary triangle change sign under CP. Consequently, any non-zero (and non-π) value of an angle is a sign of CP violation, and they can in principle be measured directly from the asymmetries of CP violation in interference. While the case for β is straight forward, the angles α and γ are more complicated. As it is instructive to see the examples for β and α, these will be discussed shortly, before engaging in a more detailed discussion on γ (see Section 5).

4.7.1 Determination of β

The angle β in the unitary triangle is the least complicated to measure of the three unitary angles. The reason is that it can be measured in CP eigenstates, which has a single weak phase dominating the amplitude and which has both a relatively large branching fraction and is at the same time easy to reconstruct cleanly.
The states in question are B^0 decay $b \to c\bar{c}s$, at hadronic level e.g. $B^0 \to J/\psi K^0_S$. There are two diagrams contributing to the amplitude (see Fig. 4.3), where the first is a tree diagram and the second is a loop diagram (a so-called penguin diagram[29]).

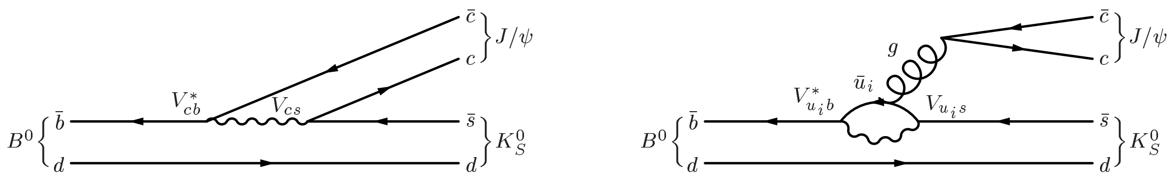


Figure 4.3: Diagrams for the decay $b \to c\bar{c}s$. The tree diagram (left) and penguin diagrams (right) both contribute, but the dominant contributions have the same weak phase, which greatly simplifies the extraction of basic parameters, here $\sin(2\beta)$. It should be noted that the single gluon in the penguin diagram (right) represents (at least) three gluons.

The tree amplitude (T) and (strong) penguin amplitudes ($P^{u,c,t}$) combined with their respective CKM couplings yield:

$$\mathcal{A} \;=\; V_{us}V^*_{ub}P^u + V_{cs}V^*_{cb}(T+P^c) + V_{ts}V^*_{tb}P^t = V_{cs}V^*_{cb}\underbrace{(T+P^c-P^t)}_{\text{“tree”}} + V_{us}V^*_{ub}\underbrace{(P^u-P^t)}_{\text{“penguin”}}, \quad (4.45)$$

where in the second equation, the unitarity of the CKM matrix, specifically $V_{us}V^*_{ub} + V_{cs}V^*_{cb} + V_{ts}V^*_{tb} = 0$, has been used. It is convenient and common (though not in accordance with definition) to refer to the two terms as the tree and penguin term, respectively, as indicated in Eq. (4.45).
The ratio between the first and the second amplitude is of order unity or less, and as $V_{cs}V^*_{cb} = \mathcal{O}(\lambda^2)$ while $V_{us}V^*_{ub} = \mathcal{O}(\lambda^4)$, the first term and its associated weak phase will be dominant, thus to a good approximation:

$$\lambda(b \to c\bar{c}s) \;=\; \eta_{CP}\underbrace{(V^*_{tb}V_{td}/V_{tb}V^*_{td})}_{B^0\text{ mixing}}\underbrace{(V^*_{cs}V_{cb}/V_{cs}V^*_{cb})}_{B^0\text{ decay}}\underbrace{(V^*_{cd}V_{cs}/V_{cd}V^*_{cs})}_{K^0\text{ mixing}} = -\eta_{CP}\, e^{-2i\beta} \quad (4.46)$$

where η_{CP} is the CP eigenvalue of the final state ($\eta_{B^0\to J/\psi K^0_S} = -1$). The approximation that K^0_S is a CP eigenstate (with $CP = +1$) is used. This yields:

$$\mathrm{Im}\,\lambda(B^0 \to J/\psi K^0_S) \;=\; \sin(2\beta) + \mathcal{O}(10^{-3}) \quad (4.47)$$

Higher $c\bar{c}$ resonances and the K^0_L can equally well be used, changing sign according to η_{CP}. As the $B^0 \to J/\psi K^0_S$ mode is abundant, efficient and experimentally clean, this single mode dominates the measurement of $\sin(2\beta)$ and has for this reason been dubbed “the golden channel”.

[29]The resemblence with a penguin does not come easy, but the story behind the name is entertaining [Vai99]

4.7.2 Determination of α

In principle, $\sin(2\alpha)$ can be determined in the same way as $\sin(2\beta)$, just considering decays of the type $b \to u\bar{u}d$, hadronically being B^0 decays to e.g. $\pi^+\pi^-$, instead.

However, contrary to the $b \to c\bar{c}s$ decay, penguin diagrams with different CKM parameters are *not* Cabibbo suppressed but $\mathcal{O}(1)$ compared to the tree amplitude [AHR02]. Due to non-pertubative long distance effects the relative amplitudes can not be calculated, as neither heavy quark effective theory nor lattice QCD can be effectively applied.
For this reason, measurements of CP asymmetries in this and similar channels can not be directly related to $\sin(2\alpha)$ and thus CKM parameters. The measured asymmetries have been dubbed $\sin(2\alpha_{eff})$, which varies between final states.

The problem can be solved by isospin analysis [LNQS91, SQ93]. The key observation is that the dominant penguin diagrams are purely $\Delta I = \frac{1}{2}$, since they involve a gluon ($\Delta I = 0$), while the tree level diagram have $\Delta I = \frac{3}{2}$ and $\Delta I = \frac{1}{2}$ components. If the $\Delta I = \frac{3}{2}$ part can be isolated, it can be used to determine $\sin 2\alpha$ without theoretical errors.
This procedure requires different final states, as one effectively subtracts the penguins. For this reason the $B^0 \to \rho\pi$ channel is more suitable, since it has three final states.

Another approach is to limit the difference between $\sin(2\alpha_{eff})$ and $\sin(2\alpha)$ by considering the color-suppressed neutral decay – expressed in the Grossman-Quinn bound [GQ98]. For the $B^0 \to \pi\pi$ mode, this does not seem to be a liable path, while in the $\rho\rho$ case, it has proven more effective [AJBY04].

4.8 Section summary and conclusions

Though not fundamental, effective field theory gives the basic framework of the Hamiltonian for the mixing of the B^0 system. Its properties under CP transformation shows exactly the requirements for CP violation to occur, and it can be subdivided into three categories. The largest and most interpretable type of CP violation requires time-dependent analysis. In the coherent $B\overline{B}$ production, characteristic of the $\Upsilon(4S)$ decay, such measurements require knowledge of the flavor of the other B meson and the time between their decay.

The unitary triangle is an illustrative representation of the CP violating quantities measured in the B system, and how they can be related to physics parameters. The three CP violating angles of the triangle correspond to the relative weak phase between possibly interfering amplitudes. Both CP and non-CP eigenstates can be used for measuring CP violation, but only the latter can be used for measuring γ with B_d^0 mesons.

5 Methods for measuring the unitary angle γ

I'll bet you, that γ will be measured before α.

[R. Aleksan, outside the LAL office at SLAC]

While the litterature is rich on methods for measuring the unitary angles β and α, the angle γ has been less courted. It is generally considered to be the most difficult to measure, and so far only limits have been set on γ. Unlike the other angles, no CP eigenstate directly measures this angle, which complicates its extraction[30].
In this section the various methods for extracting γ will be reviewed and their advantages and drawbacks discussed. Then the idea of three-body decays will be presented, the advantages and drawbacks considered, and the feasibility will be discussed in some detail.

5.1 Methods and difficulties

The extraction of γ requires the interference between $b \to c$ and $b \to u$ quark transitions, thus amplitudes, which contain V_{cb} and V_{ub}. Unlike the charmless case for measuring α, there are *no* penguin diagrams (exactly due to the presence of the charm quark in the final state), which greatly simplifies the method and makes it theoretically much cleaner.
The angle measured is $\gamma' = \arg(V_{cs}V_{ub}^*/V_{ud}V_{cb}^*) = \gamma + \xi$, where $\xi = \arg(-V_{cd}V_{cs}^*/V_{ud}V_{us}^*)$ [AKL94]. The angle ξ is the phase of V_{cs} and arises from the unitarity triangle $V_{ud}V_{us}^* + V_{cd}V_{cs}^* + V_{td}V_{ts}^* = 0$, which is very "squashed" ($\lambda : \lambda : \lambda^5$). As ξ is the angle opposite of the small side, it is necessarily small ($\mathcal{O}(\lambda^4)$), and will be neglected in the following, as is costumary.

As the $b \to c$ transition yields a final state containing a charm quark, the $b \to u$ transition must necessarily also do so, for interference to take place. The most abundant $b \to c$ transition leads to the hadronic decay $B^0 \to D^{(*)-}\pi^+/\rho^+/a_1^+$. For a $b \to u$ transition to interfere with such final states, a $\bar{c}d$ quark pair must be produced from the $W^\pm$, making it a Doubly Cabibbo-Suppressed Decay (DCSD), that is the ratio of the amplitudes is of order λ^2, and thus the interference obtained is very small ($\mathcal{O}(\lambda^2)$) [DR86]. The diagrams of these decays are shown in Fig. 5.1. Even if this is not entirely true for the $B^0 \to D^{*\pm}\rho/a_1^\mp$ ($P \to VV/VA$)

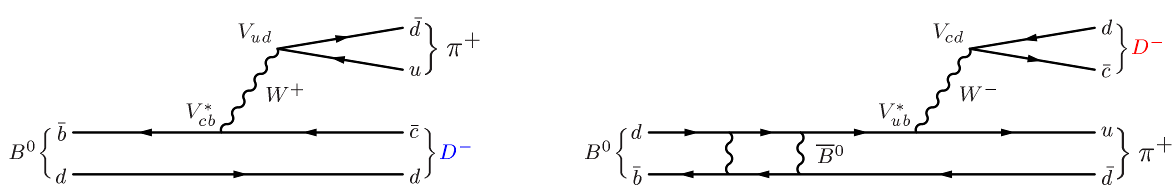


Figure 5.1: Diagrams for the $B^0 \to D^{(*)-}\pi^+/\rho^+/a_1^+$ decays involving the CKM matrix element product $V_{cb}^*V_{ud}$, denoted $b \to c$ transitions (left), and $V_{ub}^*V_{cd}$, denoted $b \to u$ transitions (right). The interference occurs through mixing (sketched) in the right diagram, and the interference term is of order λ^2.

case[31], where only coefficients (but 18(VV)/15(VA) of them) of singly Cabbibo-suppressed interferences are needed [LSS00], the smallness of the amplitude ratio remains troublesome.

This obstacle can be somewhat overcome by considering decays, where both quark transitions are (singly) Cabibbo-suppressed. Whereas the $b \to c$ transition becomes suppressed in producing a $u\bar{s}$ pair from the W, the $b \to u$ transition (already suppressed by this transition)

[30] This is only true for the B_d^0 system. In the B_s^0 system, the CP mode $B_s^0 \to \rho^0 K_S^0$ measures γ [Dun95].

[31] The notation, describing particle spin properties, is Pseudo Scaler (P), Vector (V) and Axial Vector (A).

is not CKM-suppressed any further from producing an s quark, since $V_{cs} \simeq 1$. Thus the two amplitudes are or order λ^3 with an enhanced interference as a result. Hadronically, this corresponds to the decays $B^+ \to \overline{D}^0/D^0 K^+$, the diagrams of which are shown in Fig. 5.2.

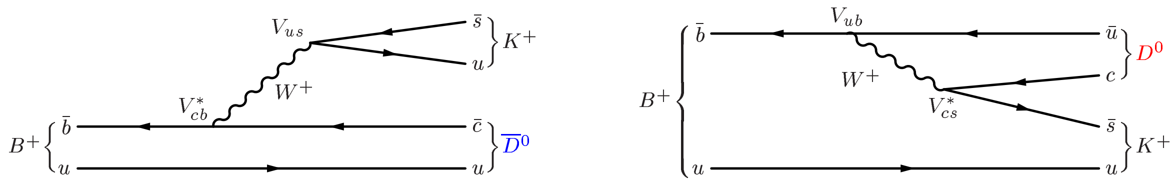


Figure 5.2: Diagrams for the $B^+ \to \overline{D}^0/D^0 K^+$ decay involving the CKM matrix element product $V_{cb}^* V_{us}$, denoted $b \to c$ transitions (left), and $V_{ub}^* V_{cs}$, denoted $b \to u$ transitions (right).

Here the two amplitudes interfere via CP eigenstates of the D^0 meson, defined as:

$$D^0_{CP\,=\,\pm 1} \quad = \quad \frac{1}{\sqrt{2}}\left(D^0 \pm \overline{D}^0\right). \tag{5.1}$$

As mixing and CP violation in the D^0 system is known to be small [PDG02], and in the SM believed to be very small [Nel99, FGLP02], it will be neglected in the following.

Gronau and Wyler (GW) [GW91] originally proposed to use this decay mode. By also measuring the flavor specific amplitudes, the relative phase (containing both a strong phase and γ) can be extracted through a triangle relation (see Figure 5.3). By measuring the CP conjugated mode, for which the weak phase γ changes sign in the amplitude, the strong phase can be eliminated.

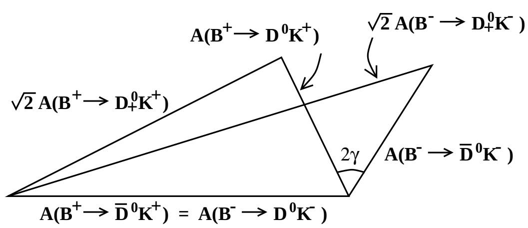


Figure 5.3: Illustration of the original Gronau-Wyler (GW) method. By measuring all the amplitudes of the decay $B^+ \to D^0_{(CP)} K^+$, one can extract γ.

Though of equal order in λ, the $b \to u$ transition amplitude is color suppressed, i.e. the color singlet, $W^\pm$ boson, has to decay into a quark pair with a color-anticolor that matches that of the original B meson, naively suppressing the amplitude by a factor of $1/N_c$. Recent observations of color-suppressed decays [BELLE02, BABAR03b] verify this ratio of amplitudes [NP01], and thus suppression is expected for internal spectator diagrams (see diagram 5.2 (right)).
In addition, the ratio of CKM couplings is $|V_{ub}V_{cs}|/|V_{cb}V_{us}| = \sqrt{\rho^2+\eta^2} = 0.41 \pm 0.07$, which further suppresses the $b \to u$ transition[32]. Given these values, one naively expects the amplitude ratio $r \equiv |A(B^+ \to D^0 K^+)/A(B^+ \to \overline{D}^0 K^+)| = |A(b \to u)/A(b \to c)|$ to be only about $0.41/3 \simeq 0.14$.

[32]This is a case, where the terminology of expressing suppression using orders of λ has its limitations, as the $1 \leftrightarrow 3$ CKM couplings are not simple multiples of λ.

Furthermore, this method requires reconstruction of D^0 mesons into CP eigenstates, which are Cabibbo-suppressed.

However, apart from the above mentioned obstacles, the original GW method suffers from an inherent problem, in that it is not possible to extract the $b \to u$ amplitudes (see Fig. 5.3) from the branching fractions of the decays. In this statement it implied that only hadronic modes are considered. Semileptonic decays can be used, as they are very flavor specific, but their use is experimentally very hard.
The problem is the interference with the DCSD of the D^0 (i.e. $D^0 \to K^+\pi^-$) of the $b \to c$ amplitude. Such a decay yields the same final state as the regular $b \to u$ decay, and therefore interference occurs. Though this is seemingly a small effect, the fact that the $b \to u$ transition is suppressed makes it an $\mathcal{O}(1)$ effect [ADS97]. Thus the original GW method does not work.

In an attempt to avoid this shortcomming of the original GW-method, two suggestions were made. One method is to simply use the two combined rates and two asymmetries [Gro98]:

$$R^{\pm}_{CP} = BR(B^+ \to D^0_{CP\pm}K^+) + BR(B^- \to D^0_{CP\pm}K^-) \tag{5.2}$$
$$A^{\pm}_{CP} = (BR(B^+ \to D^0_{CP\pm}K^+) - BR(B^- \to D^0_{CP\pm}K^-))/R^{\pm}_{CP} \tag{5.3}$$

from which the quantities r, δ and γ can be extracted (four observables and three unknowns). This method does not have the problem with the amplitude interference, but still requires the reconstruction of the D^0 into both $CP = +1$ and $CP = -1$ eigenstates.

Another and more refined proposal is to take the DCSD interferences into account, suggested by Atwood, Dunitz and Soni (ADS) [ADS97]. The method does not require a measurement of these modes in themselves, only their interference. However, at least two modes (one non-CP) are needed to make up for the lack of observables, thus the extraction is less constrained.

Finally, both the (new) GW and the ADS-methods (and any other two-body method) are subject to an eight-fold ambiguity in the solution for γ due to *a priori* unknown strong phases [AKL94, Sof99] (see Section 5.6). As a result, obtaining satisfactory sensitivity requires very high statistics and necessitates the use of as many decay modes, each with different strong phases and therefore different ambiguities (see Fig. 5.8). It may happen, that some of the ambiguities fall on top of each other, independently of the number of particles in the final state. However, this is a favorable scenario, which cannot be expected.

In general, given the $B^- \to D^0/\overline{D}^0K^-$ decay mode, the aim is to reconstruct final states where the overlap between D^0 and $\overline{D}^0$ is the greatest, thus preferably modes, which are equally accessible. While CP modes are the obvious choice, others can also be used, such as e.g. $D^0 \to K^+K^{*-}$ [GGSZ03a], though they have a low branching fraction.

Three-Body decays of the D^0 can also be used [GGSZ03b], though care has to be taken, as their CP content changes across the Dalitz plot. One advantage is that the mode $D^0 \to K^0\pi^+\pi^-$ is very abundant ($5.92 \pm 0.35\%$ [PDG02]), and the overlap between the D^0 and $\overline{D}^0$ significant ($\mathrm{Br}(D^0 \to K^0\rho^0) = 1.47 \pm 0.29\%$). Its use requires a good mapping of the Dalitz distribution, but this can be obtained from other sources (e.g. $D^{*+} \to D^0\pi^+$), for which clean large statistics samples exist. From the variation of the strong phases, some of the ambiguities can also be eliminated (see section 5.6).

Finally, the decay mode $B^0 \to D^0/\overline{D}^0K^{(*)0}$ has been proposed [GL91], as here both the $b \to c$ and $b \to u$ are color suppressed, and hence of similar magnitudes (albeit small), resulting in a larger asymmetry. While the $B^0 \to D^0/\overline{D}^0K^0$ mode requires tagging and may be hard or at least efficiency-costly to fit in time due to longlived final state particles[33], the $K^{*0} \to K^-\pi^+$ mode is selftagging and can measure the amplitude ratio.

The situation is summarized in Figure 5.4, where the amplitudes and branching fractions for the different methods are shown.

[33]Determining the B vertex may be hard, but not impossible, as it has been done for the $K^0_S\pi^0$ mode [Far03].

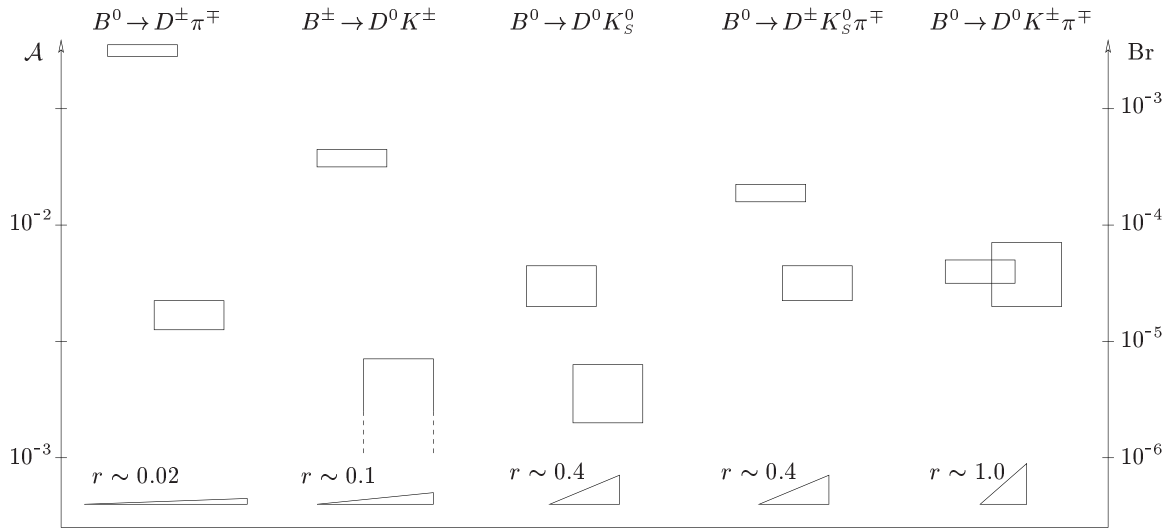


Figure 5.4: The amplitudes (left scale) and branching fractions (right scale) for the various decay modes involved in measuring γ. For each mode (labelled at the top) the size of the $b \to c$ (left) and $b \to u$ (right) amplitude/branching fraction is shown, with the corresponding asymmetry illustrated by a triangle (see Fig. 5.3) at the bottom.

Other methods involving $B \to K\pi$ modes related by $SU(3)$ flavor symmetry have been proposed [GRL94, BF99, Neu99], but these are troubled by electroweak penguin contributions and the reliance on $SU(3)$ symmetry, which is known to be correct only to about $\mathcal{O}(30\%)$ *on average* [CKM03]. Though these modes are interesting, as they can potentially reveal CP violation, they will not be considered in the following, as their reach on γ is most likely limited.

5.2 Extracting γ from B $\to$ DKπ decays

Si nous ne trouvons pas des choses agréables, nous trouverons du moins des choses nouvelles.
If we do not find anything pleasant, at least we shall find something new.
[Voltaire, 1694-1778]

Apart from the problem of not having any CP eigenstate leading to a measurement of γ, the largest difficulty is the difference in amplitude between the final states originating from the $b \to c$ and $b \to u$ transitions, which is caused by the ratio of CKM-factors and color suppression. While the first cause can inherently not be cured, the second can be overcome by considering three-body decays, that is decays of the type $B \to DK\pi$ (see Table 5.1).

Of the nine possible $DK\pi$ final states, there are three for which have interference between the $b \to c$ and the $b \to u$ amplitudes and where the latter is not color suppressed, namely $B^\pm \to D^0 K^\pm \pi^0$, $B^0 \to D^0 K^+ \pi^-$, and $B^0 \to D^\mp K^0 \pi^\pm$.

The $\mathbf{B^\pm \to D^0 K^\pm \pi^0}$ decay is much like the original GW-mode, but the additional π^0 in the final state allows for a $b \to u$ transition, which is not color suppressed, when the $D^0 K^+$ is produced from the W^+.

The $\mathbf{B^0 \to D^\mp K^0 \pi^\pm}$ decay has no predecessor, as there exists no corresponding two-body decay, which is accessible through both $b \to c$ and $b \to u$ transitions[34].

[34] The decay $B^0 \to D^\pm K^\mp$ is very flavor specific, as the opposite charge combination can only be reached through annihilation, which is quite suppressed.

Mode	Features	Advantage/Challenge
$B^+ \to \overline{D}^0/D^0 K^0 \pi^+$	$b \to u$ color suppressed	
$B^+ \to \overline{D}^0/D^0 K^+ \pi^0$	Non-suppressed interference	$D^0_{CP\,=\,\pm 1}$, includes π^0
$B^+ \to D^- K^+ \pi^+$	No interference	
$B^+ \to D^+ K^+ \pi^-$	No interference	
$B^+ \to D^+ K^0 \pi^0$	No interference	
$B^0 \to \overline{D}^0/D^0 K^0 \pi^0$	Color suppressed	
$B^0 \to \overline{D}^0/D^0 K^+ \pi^-$	Non-suppressed $b \to u$ interference	$D^0_{CP\,=\,\pm 1}$, equal amplitudes
$B^0 \to D^\mp K^\pm \pi^0$	No interference	
$B^0 \to D^\mp K^0 \pi^\pm$	Non-suppressed interference	Time dependent analysis

Table 5.1: Possible three-body $B \to DK\pi$ modes. In three of the modes, there is interference and the $b \to u$ amplitude is not color suppressed.

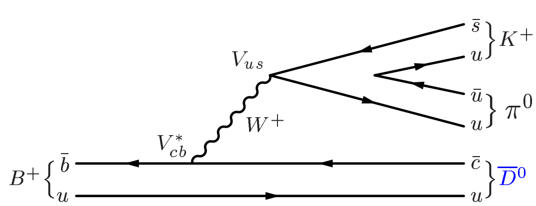
The $\mathbf{B^0 \to D^0 K^+ \pi^-}$ decay is a generalization of the $B^0 \to D^0 K^{*0}$ mode, however with the realization that the $b \to u$ transition is not color suppressed when considering the entire Dalitz plot. The same is not true for the $b \to c$ amplitude, which leaves the possibility that the amplitudes are truely of the same order ($r \sim 1$).

As the $B^0 \to D^0 K^+ \pi^-$ mode is selftagging (the charge of the kaon determines the flavor of the B^0) and interferes through CP eigenstates of the D^0, the analysis required for extracting γ is very similar to the analysis of the $B^\pm \to D^0 K^\pm \pi^0$ mode, and the two modes will be treated concurrently. However, in order not to clutter the formulae, only one mode (here the $B^\pm \to D^0 K^\pm \pi^0$ mode) will be treated, and the other, $B^0 \to D^0 K^+ \pi^-$, will only be mentioned when differences between the two modes occur.
The $B^0 \to D^\mp K^0 \pi^\pm$ mode does not require the reconstruction of D^0 CP modes and does not suffer from the problem of DCSD interference. However, here a time-dependent fit and thus tagging is required, thus this analysis is very different from the two others, and will be treated separately.

Although modes where one or more of the three final state particles is a vector can also be used, for clarity and simplicity only the pseudo-scalar cases are discussed here (including a vector introduces three helicity states, which have to be separated by angular analysis).

5.3 Extracting γ from $\mathbf{B^\pm \to D^0 K^\pm \pi^0}$ and $\mathbf{B^0 \to D^0 K^+ \pi^-}$ decays

The two decay modes, $B^\pm \to D^0 K^\pm \pi^0$ and $B^0 \to D^0 K^+ \pi^-$, are similar in that they have interference through the CP eigenstates of the D^0. However, the diagrams leading to these two final states are different (see Fig. 5.5 and 5.6), resulting in different (and yet unknown) branching fractions and amplitude ratios.



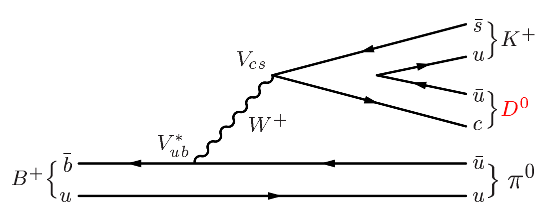


Figure 5.5: Diagrams for the $B^+ \to \overline{D}^0/D^0 K^+ \pi^0$ decay, via the $b \to c$ transitions (left) and the $b \to u$ transitions (right).

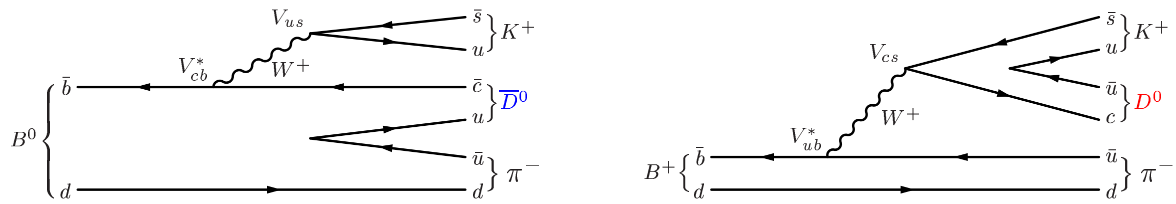


Figure 5.6: Diagrams for the $B^0 \to \overline{D}^{\,0}/D^0K^+\pi^-$ decay, via the $b \to c$ transitions (left) and the $b \to u$ transitions (right).

Due to the absence of color suppression, both interfering amplitudes are large. As a result, the contamination from DCSD of the D^0 meson are *relatively* smaller than for the corresponding two-body modes, though still there. However, should the $b \to u$ amplitude be unexpectedly small, one could still carry out this analysis described by taking DCSD into account [ADS97].

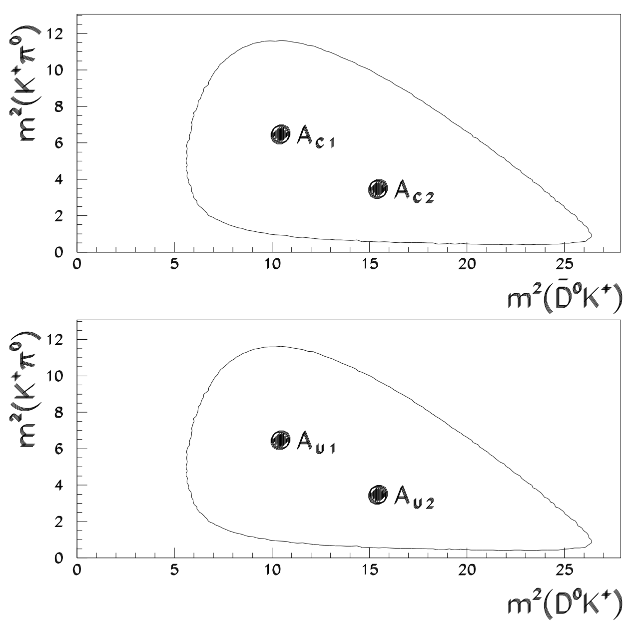


Figure 5.7: Two points on sketch of Dalitz plot. As the dynamics behind these two decays is different, their strong phase is also likely to be so. This can be used to resolve ambiguities (see text).

While the daughters of two-body decays have fixed back-to-back momentum, three-body decays are not as constrained, as the momentum of daughters can vary. However, once the momentum of one daughter has been chosen, the momentum of the others is fixed. Though this is only one degree of freedom, three-body decays are often represented in two dimensional plots (named Dalitz plots [CDD56, Jac02]), spanned by the invariant mass of any two pairs of daughters (see Fig. 5.7). In such Dalitz plots, resonances (i.e. quasi two-body decays) are lines, and the second degree of freedom is the helicity of the resonant state. In the following, the two Dalitz variables are collectively denoted ξ. One advantage of Dalitz plots is, that the size of the phase space is proportional to area in the plot.

Selecting a particular point i in the Dalitz plot (represented by the two invariant masses, e.g. $m^2_{D^0,K^\pm}$ and $m^2_{K^\pm\pi^0}$, see Figure 5.7), Eq. (5.1) implies the relations:

$$\begin{aligned} \mathcal{A}_i(B^+ \to D^0_{CP=\pm1}K^+\pi^0) &= \frac{1}{\sqrt{2}}\left(\mathcal{A}_i(B^+ \to D^0K^+\pi^0) \pm \mathcal{A}_i(B^+ \to \overline{D}^0K^+\pi^0)\right) \\ \mathcal{A}_i(B^- \to D^0_{CP=\pm1}K^-\pi^0) &= \frac{1}{\sqrt{2}}\left(\mathcal{A}_i(B^- \to D^0K^-\pi^0) \pm \mathcal{A}_i(B^- \to \overline{D}^0K^-\pi^0)\right). \end{aligned} \tag{5.4}$$

The amplitudes corresponding to the transitions in Fig. 5.5 can be written as:

$$\begin{aligned} \mathcal{A}_i(B^+ \to \overline{D}^0K^+\pi^0) = A_{\mathcal{C}i}e^{i\delta_{\mathcal{C}i}} \quad &, \quad \mathcal{A}_i(B^+ \to D^0K^+\pi^0) = A_{\mathcal{U}i}e^{i\delta_{\mathcal{U}i}}e^{i\gamma}, \\ \mathcal{A}_i(B^- \to D^0K^-\pi^0) = A_{\mathcal{C}i}e^{i\delta_{\mathcal{C}i}} \quad &, \quad \mathcal{A}_i(B^- \to \overline{D}^0K^-\pi^0) = A_{\mathcal{U}i}e^{i\delta_{\mathcal{U}i}}e^{-i\gamma}, \end{aligned} \tag{5.5}$$

where γ is the relative phase of the CKM matrix elements involved in this decay, and $A_{\mathcal{C}}$ ($A_{\mathcal{U}}$) and $\delta_{\mathcal{C}}$ ($\delta_{\mathcal{U}}$) are the modulus of the amplitude and CP conserving strong interaction phase of the transitions of Fig. 5.5 left (right). The amplitudes, $A_{\mathcal{C}i}$ and $A_{\mathcal{U}i}$, in Eq. (5.5) can be obtained from the measurements of the B meson decay widths

$$\begin{aligned} \Gamma_i(B^+ \to \overline{D}^0K^+\pi^0) &= \Gamma_i(B^- \to D^0K^-\pi^0) = A_{\mathcal{C}i}^2 \\ \Gamma_i(B^+ \to D^0K^+\pi^0) &= \Gamma_i(B^- \to \overline{D}^0K^-\pi^0) = A_{\mathcal{U}i}^2. \end{aligned} \tag{5.6}$$

Eq. (5.4) implies

$$\begin{aligned} 2\Gamma_i(B^+ \to D^0_{CP=\pm1}K^+\pi^0) &= A_{\mathcal{C}i}^2 + A_{\mathcal{U}i}^2 \pm 2A_{\mathcal{C}i}A_{\mathcal{U}i}\cos(\Delta\delta_i + \gamma) \\ 2\Gamma_i(B^- \to D^0_{CP=\pm1}K^-\pi^0) &= A_{\mathcal{C}i}^2 + A_{\mathcal{U}i}^2 \pm 2A_{\mathcal{C}i}A_{\mathcal{U}i}\cos(\Delta\delta_i - \gamma), \end{aligned} \tag{5.7}$$

where $\Delta\delta_i \equiv \delta_{\mathcal{U}i} - \delta_{\mathcal{C}i}$. Thus, by measuring the widths in Eq. (5.6) and (5.7), one extracts $\sin^2\gamma$ from

$$\sin^2\gamma = \frac{1}{2}\left(1 - C\overline{C} \pm \sqrt{(1-C^2)(1-\overline{C}^2)}\right), \tag{5.8}$$

where $C \equiv \cos(\Delta\delta_i + \gamma)$ and $\overline{C} \equiv \cos(\Delta\delta_i - \gamma)$. Thus in the limit of very high statistics, one would extract $\sin^2\gamma$ and hence γ for each point i of the Dalitz plot.

However, in every point of the Dalitz plot, γ is obtained with an eight-fold ambiguity in the $[0, 2\pi]$ interval, which is a consequence of the invariance of the $\cos(\Delta\delta_i \pm \gamma)$ terms in Eq. (5.7) under the three symmetry operations [Sof99]:

$$\begin{array}{lcll} S_{\text{ex}} & : & \gamma \to \Delta\delta_i & \Delta\delta_i \to \gamma \\ S_{\text{sign}} & : & \gamma \to -\gamma & \Delta\delta_i \to -\Delta\delta_i \\ S_\pi & : & \gamma \to \gamma + \pi & \Delta\delta_i \to \Delta\delta_i + \pi. \end{array} \tag{5.9}$$

However, an important benefit is gained from the multiple measurements made in different points of the Dalitz plot. When combining results from different points (or different modes), which are likely to have different strong phases $\Delta\delta_i$, the ambiguities related to the strong phase can be eliminated, as is sketched in Fig. 5.8. This variation can either be due to the presence of resonances or because of a varying phase in the non-resonant (NR) contribution.

In particular, the exchange symmetry S_{ex} is numerically different from one point to the other, which in effect breaks this symmetry and resolves the ambiguity. Similarly, the S_{sign} symmetry is broken if there exists some *a priori* knowledge of the dependence of $\Delta\delta_i$ on the Dalitz plot parameters. This knowledge is provided by the existence of broad[35] resonances, whose Breit-Wigner phase variation is known and may be assumed to dominate the phase

[35] By broad resonance are meant those that can be resolved by the detector resolution (typically $3-5$ MeV).

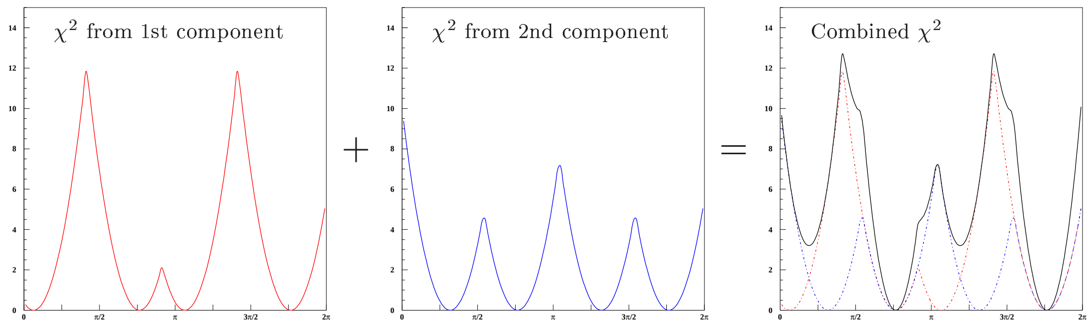


Figure 5.8: Illustration of ambiguity resolving. Given two results for γ (left and middle plots), each with four ambiguities which lie differently, due to different strong phases, the combined result only has two (fully unresolved) ambiguities (due to S_π).

variation over the width of the resonance. To illustrate this, let i and j be two points in the Dalitz plot, corresponding to different values of the invariant mass of the decay products of a particular resonance. For simplicity only one resonance is considered. One then measures $\cos(\Delta\delta_i \pm \gamma)$ at point i and $\cos(\Delta\delta_i + \alpha_{ij} \pm \gamma)$ at point j, where α_{ij} is known from the parameters of the resonance. It is important to note that the sign of α_{ij} is known (contrary to the sign of $\Delta\delta_i$), hence it does not change under S_{sign}. Therefore, should one choose the S_{sign} related solution $\cos(-\Delta\delta_i \mp \gamma)$ at point i, one would get $\cos(-\Delta\delta_i + \alpha_{ij} \mp \gamma)$ at point j. Since this is different from $\cos(\Delta\delta_i + \alpha_{ij} \pm \gamma)$, the S_{sign} ambiguity is resolved. This is illustrated graphically in Eq. 5.10:

$$\begin{array}{ccc} \cos(\Delta\delta_i \pm \gamma) & \overset{S_{\text{sign}}}{\longleftrightarrow} & \cos(-\Delta\delta_i \mp \gamma) \\ BW \downarrow & & \downarrow BW \\ \cos(\Delta\delta_i + \alpha_{ij} \pm \gamma) & \overset{S_{\text{sign}}}{\not\longleftrightarrow} & \cos(-\Delta\delta_i + \alpha_{ij} \mp \gamma) \end{array} \tag{5.10}$$

Thus, broad resonances reduce the initial eight-fold ambiguity to the two-fold ambiguity of the S_π symmetry, which is not broken. Fortunately, S_π leads to the well-separated solutions γ and $\gamma + \pi$, the correct one of which is easily identified when this measurement is combined with other measurements of the unitarity triangle.

5.4 Extracting γ from $\mathbf{B^0 \to D^{\mp} K^0 \pi^{\pm}}$ decays

The $B^0 \to D^{\mp} K^0 \pi^{\pm}$ mode is very different in many respects, and mostly resembles the $B^0 \to D^{\pm}\pi^{\mp}$ mode, though with a far greater asymmetry. The leading (i.e. color-allowed) diagrams for the decay are shown in Fig. 5.9[36].

As the final states interfere through mixing, the weak angle extracted is $2\beta + \gamma$ and requires a time-dependent analysis. This changes the analysis with respect to that of the $B^{\pm} \to D^0 K^{\pm} \pi^0$ and $B^0 \to D^0 K^+ \pi^-$ modes, but the idea is essentially the same and many aspects alike.

Selecting a particular point i in the Dalitz plot, the amplitudes and their CP conjugates corresponding to the $b \to c$ and $b \to u$ transitions shown in Fig. 5.9 can be written as[37]:

$$\begin{array}{rclrcl} \mathcal{A}_i(B^0 \to D^- K^0 \pi^+) & = & A_{\mathcal{C}i} e^{i\delta_{\mathcal{C}i}}, & \mathcal{A}_i(\overline{B}^0 \to D^- K^0 \pi^+) & = & A_{\mathcal{U}i} e^{i\delta_{\mathcal{U}i}} e^{-i\gamma}, \\ \mathcal{A}_i(B^0 \to D^+ K^0 \pi^-) & = & A_{\mathcal{U}i} e^{i\delta_{\mathcal{U}i}} e^{i\gamma}, & \mathcal{A}_i(\overline{B}^0 \to D^+ K^0 \pi^-) & = & A_{\mathcal{C}i} e^{i\delta_{\mathcal{C}i}}, \end{array} \tag{5.11}$$

[36]Throughout, the neutral kaon will be denoted K^0, as its reconstruction is done into the K^0_S state only.

[37]Note that the same notation is used here as in the previous section, though the actual values are not necessarily the same.

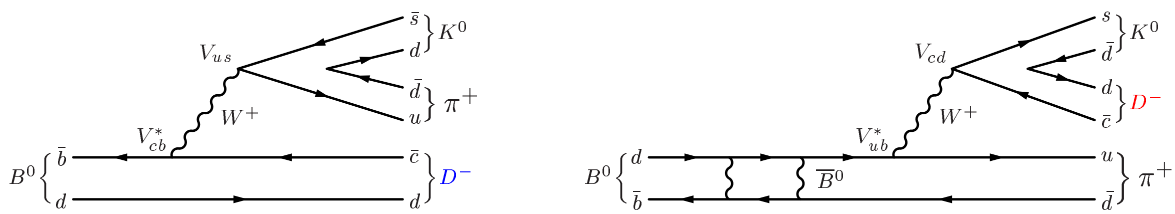


Figure 5.9: Diagrams for the $B^0 \to D^{\mp} K^0 \pi^{\pm}$ decay involving via the $b \to c$ transitions (left) and the $b \to u$ transitions (right).

where γ is the relative phase of the CKM matrix elements involved in this decay, and $A_{\mathcal{C}}$ ($A_{\mathcal{U}}$) and $\delta_{\mathcal{C}}$ ($\delta_{\mathcal{U}}$) are once again the magnitude of the amplitudes and CP conserving strong interaction phases of the $b \to c$ ($b \to u$) transition, respectively. The time dependent Probability Density Functions (PDF) of the four decay types are:

$$
\begin{aligned}
f(B^0 \to D^- K^0 \pi^+)_i &= N e^{-\Gamma|t|} \left[1 + C_i \cos(\Delta m t) + S_{i-} \sin(\Delta m t)\right], && (5.12)\\
f(\overline{B}^0 \to D^- K^0 \pi^+)_i &= N e^{-\Gamma|t|} \left[1 - C_i \cos(\Delta m t) - S_{i-} \sin(\Delta m t)\right], && (5.13)\\
f(B^0 \to D^+ K^0 \pi^-)_i &= N e^{-\Gamma|t|} \left[1 - C_i \cos(\Delta m t) + S_{i+} \sin(\Delta m t)\right], && (5.14)\\
f(\overline{B}^0 \to D^+ K^0 \pi^-)_i &= N e^{-\Gamma|t|} \left[1 + C_i \cos(\Delta m t) - S_{i+} \sin(\Delta m t)\right]. && (5.15)
\end{aligned}
$$

where the amplitude ratio $\lambda_i^{-+} \equiv \mathcal{A}_i(\overline{B}^0 \to D^- K^0 \pi^+)/\mathcal{A}_i(B^0 \to D^- K^0 \pi^+)$ and the coefficients depend on the position in the Dalitz plot, i, as:

$$
C_i \;=\; = \frac{1 - |\lambda_i^{-+}|^2}{1 + |\lambda_i^{-+}|^2}, \qquad S_{i\pm} \;=\; \frac{2|\lambda_i^{\pm\mp}| \sin(\Delta\delta_i \;\pm\; (2\beta + \gamma))}{1 + |\lambda_i^{\pm\mp}|^2} \tag{5.16}
$$

From the total number of events in each of the four final states Eqs. (5.12-5.15) and a global fit of their time dependence, it is possible to extract the four quantities, C_i, S_{i-}, S_{i+}, and the overall normalization. From this one obtains:

$$
\sin^2(2\beta + \gamma) \;=\; \frac{1}{2}\left(1 + S_{i+} S_{i-} \pm \sqrt{(1 - S_{i+}^2)(1 - S_{i-}^2)}\right) \tag{5.17}
$$

Hence, in the limit of very high statistics, one would extract $\sin^2(2\beta+\gamma)$ for each point i of the Dalitz plot. However, for every point of the Dalitz plot, $2\beta + \gamma$ is obtained with an eight-fold ambiguity in the range $[0, 2\pi]$, because of the invariance of the $\sin(\Delta\delta_i \pm (2\beta + \gamma))$ terms in Eq. (5.16) under the three symmetry operations (which are different from the previous case, except for S_π):

$$
\begin{array}{lcl}
S_{\pi/2} & : & 2\beta{+}\gamma \to \pi/2 - \Delta\delta_i \quad , \quad \Delta\delta_i \to \pi/2 - (2\beta{+}\gamma) \\
S_{\pi-} & : & 2\beta{+}\gamma \to -(2\beta{+}\gamma) \quad , \quad \Delta\delta_i \to \pi - \Delta\delta_i \\
S_{\pi} & : & 2\beta{+}\gamma \to 2\beta{+}\gamma + \pi \quad , \quad \Delta\delta_i \to \Delta\delta_i + \pi
\end{array} \tag{5.18}
$$

Once again, when the multiple measurements made in different points of the Dalitz plot are combined, some of the ambiguity will be resolved, in the likely case that the strong phase $\Delta\delta_i$ varies from one region of the Dalitz plot to the other, just as for the $B^{\pm} \to D^0 K^{\pm} \pi^0$ and $B^0 \to D^0 K^+ \pi^-$ decays.

The variation of the strong phase breaks the $S_{\pi/2}$ symmetry and also the $S_{\pi-}$, if the direction of change of $\Delta\delta_i$ is a priori known, which is the case for a (broad) resonance. As always, the S_π symmetry remains unbroken, leaving the well-separated solution $2\beta + \gamma + \pi$.

5.5 The Finite Statistics Case

Since experimental data sets will of course be finite, one cannot extract γ at every point of the Dalitz plot, and thus one is required to describe the variation of amplitudes and strong phases with ξ using a limited set of parameters. The consistency of this approach can be verified by comparing the results obtained from fits of the data in a few different regions of the Dalitz plot, and the systematic error due to the choice of the parameterization may be asserted by using different parameterizations.

A fairly general parameterization assumes the existence of N_R Breit-Wigner resonances, as well as a non-resonant contribution:

$$\begin{aligned}
\mathcal{A}_\xi(\overline{b} \to \overline{c}u\overline{s}) &= \left(A_{\mathcal{C}0}(\xi)\, e^{i\delta_{\mathcal{C}0}} + \sum_j^{N_R} A_{\mathcal{C}j} B_{s_j}(\xi)\, e^{i\delta_{\mathcal{C}j}}\right) e^{i\delta_\mathcal{C}(\xi)} \\
\mathcal{A}_\xi(\overline{b} \to \overline{u}c\overline{s}) &= \left(A_{\mathcal{U}0}(\xi)\, e^{i\delta_{\mathcal{U}0}} + \sum_j^{N_R} A_{\mathcal{U}j} B_{s_j}(\xi)\, e^{i\delta_{\mathcal{U}j}}\right) e^{i\delta_\mathcal{U}(\xi)} e^{i\phi}, \qquad (5.19)
\end{aligned}$$

where ξ represents the Dalitz plot variables, $B_{s_j}(\xi)$ is the Breit-Wigner amplitude for a particle of spin s_j, normalized such that $\int |B_{s_j}(\xi)|^2 d\xi = 1$, $A_\mathcal{U}$ and $\delta_\mathcal{U}$ [$A_\mathcal{C}$ and $\delta_\mathcal{C}$] are the magnitude and CP conserving phase of the non-resonant (subscript 0) or j^{th} resonant (subscript j) $\overline{b} \to \overline{u}c\overline{s}$ [$\overline{b} \to \overline{c}u\overline{s}$] amplitude, and ϕ is the weak phase, be it γ (for the $B^\pm \to D^0 K^\pm \pi^0$ and $B^0 \to D^0 K^+ \pi^-$modes) or $2\beta + \gamma$ (for the $B^0 \to D^\mp K^0 \pi^\pm$ mode), which changes sign under CP conjugation (i.e. interchanging quarks and anti-quarks in Eq. 5.19). The functions $\delta_\mathcal{C}(\xi)$ and $\delta_\mathcal{U}(\xi)$, or rather their difference, $\Delta\delta(\xi)$, as only this can be measured, may be assumed to vary slowly over the Dalitz plot, allowing their description in terms of a small number of parameters.

For the $B^\pm \to D^0 K^\pm \pi^0$ (and $B^0 \to D^0 K^+ \pi^-$) mode(s), Eq. (5.1) again implies

$$\mathcal{A}_\xi(B^\pm \to D^0_{CP=\pm1} K^\pm \pi^0) = \frac{1}{\sqrt{2}}\Big(\mathcal{A}_\xi(B^\pm \to D^0 K^\pm \pi^0) \pm \mathcal{A}_\xi(B^\pm \to \overline{D}^0 K^\pm \pi^0)\Big), \qquad (5.20)$$

from which a PDF can be constructed:

$$P(\xi) = |\mathcal{A}_\xi(f)|^2, \qquad (5.21)$$

where the amplitude $A_\xi(f)$ is given by one of the expressions of Eqs. (5.19) and (5.20), or their CP conjugates, depending on the final state f.

5.6 Resonances and Ambiguities

Resonances may contribute to the three final states in question. Obvious candidates are broad D^{**} and $D_s^{**\pm}$ states. However, only the ones which can decay as $D^{**0/\pm} \to D^0\pi^{0/\pm}$ and $D_s^{**\pm} \to D^{0/\pm}K^{\pm/0}$ are relevant for the final states of interest. This exclude the 1^+ states, which would decay to $D^*\pi/K$ states. Furthermore, since the $D_s^{**\pm}$ is essentially produced through a $W^\pm$, the 2^+ state is forbidden as well, thus one does not expect a large contribution from these states.

Including such resonances in the analysis does not raise particular difficulties and would further enhance the sensitivity of the γ measurement. Similar arguments can be made for higher excited K states.
One also expects narrow resonances, such as the $D^*(2007)^{0/\pm}$ and a narrow $D_s^{**\pm}$ states, to be produced. However, as pointed out in [CLYO+98] and seen in the Dalitz plot of Fig. 5.10, these resonances do not overlap[38], and hence one cannot expect large interferences useful for constraining the phases of the resonances (as in e.g. the $B^0 \to \rho^\pm\pi^\mp$ case).

[38]Overlap is obtained, when including the scalar $K_0^*(1430)$, but it does not seem likely to be very abundant.

In addition, the interference between a very narrow resonance and either a broad resonance or a non-resonant term is suppressed in proportion to the square root of the narrow resonance width. Therefore, narrow resonances contribute significantly to the CP violation measurement only if both the $\overline{b} \to \overline{c}u\overline{s}$ and $\overline{b} \to \overline{u}c\overline{s}$ amplitudes proceed through the same resonance. This scenario is very favorable, but is not necessary for the success of the methods, and will therefore not be focused on in the rest of this study.

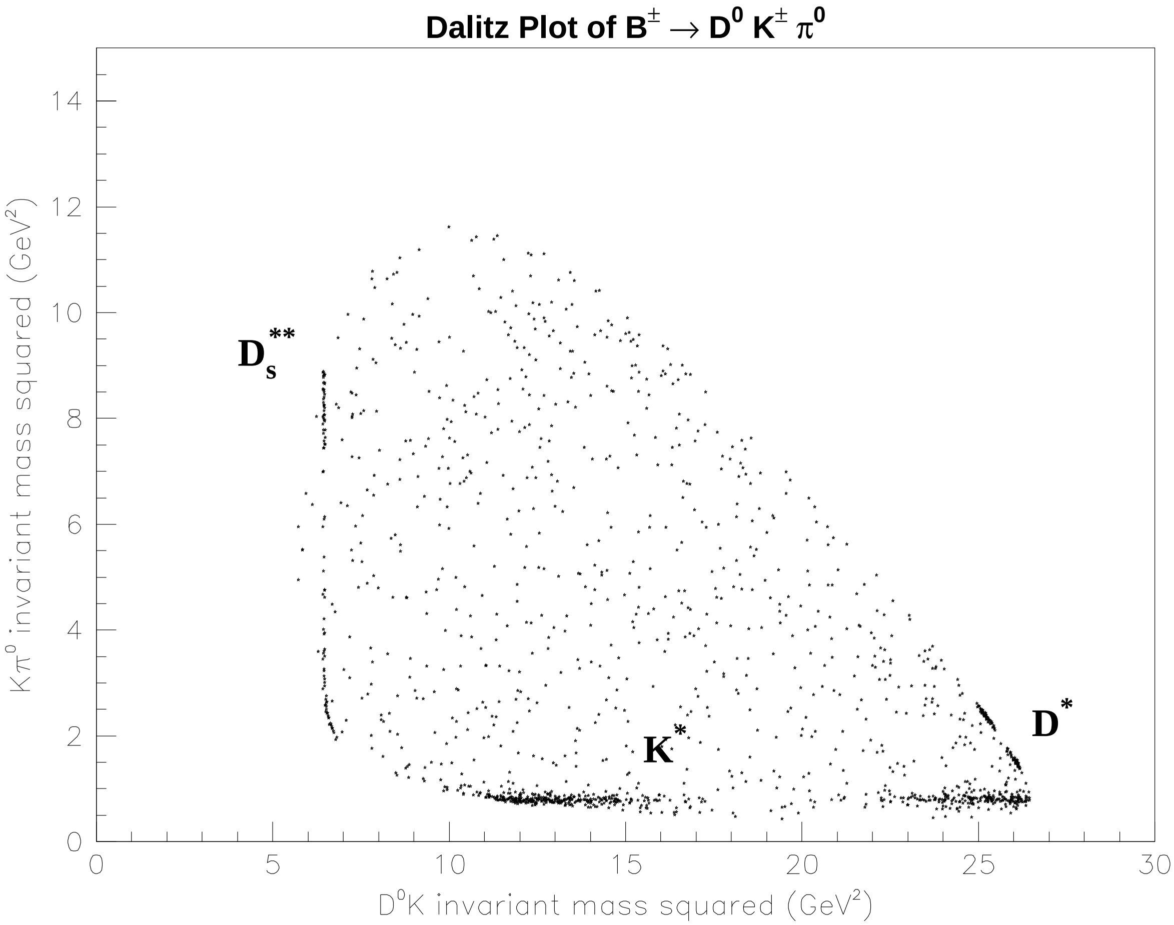


Figure 5.10: Dalitz plots obtained from simulation of B^+ and B^- decays into all final state, $D^0K^{\pm}\pi^0$, $\overline{D}^{\,0}K^{\pm}\pi^0$, and $D^0_{CP=\pm1}K^{\pm}\pi^0$. Along with non-resonant contributions, the resonances K^{*0}, D^{*0}, and D^{**}_s are shown. Unfortunately, the most abundant resonances sit at the very edge of the Dalitz plot, and do not overlap.

While any number of resonances can be included in the analysis, in the following only one will be considered. For concreteness, the resonance is taken to be the $K^{*\pm/0}(892)$. For simplicity, the ξ dependent non-resonant amplitudes and phase difference, $A_{\mathcal{C}/\mathcal{U}_0}$ and $\Delta\delta(\xi)$, were taken to be constant. This is a serious simplification, as one avoids modelling the non-resonant contribution. However, the main point is to see, whether sensitivity can be obtained, given an overlap between the $b \to c$ and the $b \to u$ contributions. For this purpose, a flat non-resonant contribution serves well. While simplifying the analysis, it is at the same time a conservative choice, as this means that the method is not dependent on the size of an interfering resonance, which in reality might enhance the sensitivity significantly.
Including resonances, the discussion of ambiguities gets slightly more complicated. The PDF of the time-independent analysis Eq. (5.21) now depends on four cosine terms that are mea-

sured (in a time-dependent Dalitz fit):

$$\begin{aligned} c^{\pm}_{00} &\equiv \cos(\Delta\delta_0 \pm \gamma) \\ c^{\pm}_{K^*0} &\equiv \cos(\Delta\delta_0 - \Delta\delta_{K^*} - \Delta\delta_{K^*}(\xi) \pm \gamma), \end{aligned} \quad (5.22)$$

where $\Delta\delta_{K^*}(\xi)$ is the ξ dependent K^* Breit-Wigner phase of B_{s_j} (see Eq. (5.19)). The cosines $c^{\pm}_{00}$ ($c^{\pm}_{K^*0}$) arise from interference between the non-resonant (resonant) $\overline{b} \to \overline{c}u\overline{s}$ amplitude and the non-resonant $\overline{b} \to \overline{u}c\overline{s}$ amplitude.

The phases $\Delta\delta_0$, $\Delta\delta_{K^*}$, and γ are all *a priori* unknown. However, it is important to note that $\Delta\delta_{K^*}$ is fully determined from the interference between the resonant and non-resonant contributions to the relatively high statistics $\overline{b} \to \overline{c}u\overline{s}$ mode, thus interference between two components of the *same* decay mode. Therefore, $\Delta\delta_{K^*}$ is obtained with no ambiguities, and with an error much smaller than those of $\delta_{\mathcal{U}0}$ or γ. Consequently, the only relevant symmetry operations are

$$\begin{array}{llll} S_{\rm ex} & : \ \gamma \to \Delta\delta_0 & , & \Delta\delta_0 \to \gamma \\ S_{\rm sign} & : \ \gamma \to -\gamma & , & \Delta\delta_0 \to -\Delta\delta_0 \\ S_{\pi} & : \ \gamma \to \gamma + \pi & , & \Delta\delta_0 \to \Delta\delta_0 + \pi \\ S^{K^*+}_{\rm ex} & : \ \gamma \to \Delta\delta_0 - \Delta\delta_{K^*} & , & \Delta\delta_0 \to \gamma + \Delta\delta_{K^*} \\ S^{K^*-}_{\rm ex} & : \ \gamma \to -\Delta\delta_0 + \Delta\delta_{K^*} & , & \Delta\delta_0 \to -\gamma + \Delta\delta_{K^*} . \end{array} \quad (5.23)$$

As discussed above, only S_π is a symmetry of all four cosines of Eq. (5.22), and is therefore irreducible. Under (combination of) the symmetry operations $S^{K^*+}_{\rm ex}$, $S^{K^*-}_{\rm ex}$, the $c^{\pm}_{K^*0}$ ambiguities can be hard to resolve, as the BW phase at the tails of the K^* resonance only vary slowly and take values around 0 and π. This leads to approximate invariance under these symmetries. The transformation properties of the cosines under combinations of the remaining four operations that can lead to an ambiguity are shown in Table 5.2.

Term	c^+_{00}	c^-_{00}	$c^+_{K^*0}$	$c^-_{K^*0}$	Term	c^+_{00}	c^-_{00}	$c^+_{K^*0}$	$c^-_{K^*0}$
Operation	Non-resonant regime				Operation	Resonant regime			
$S_{\rm ex}$	$\surd$	$\surd$	$\surd$		$S^{K^*+}_{\rm ex}$	$\surd$		$\surd$	$(\surd)$
$S_{\rm sign}$	$\surd$	$\surd$			$S^{K^*-}_{\rm ex}$		$\surd$	$(\surd)$	$\surd$
$S_{\rm ex}S_{\rm sign}$	$\surd$	$\surd$		$\surd$	$S^{K^*+}_{\rm ex}S^{K^*-}_{\rm ex}$			$(\surd)$	$(\surd)$

Table 5.2: Invariance of each of the cosines of Eq. (5.22) under combinations of the symmetry operations of Eq. (5.23), excluding S_π. Full/approximate invariance is indicated by a $\surd/(\surd)$.

Observing that no single operation in the Table 5.2 is a good symmetry of all cosines, one identifies two different regimes: In the non-resonant regime, interference with the non-resonant $\overline{b} \to \overline{c}u\overline{s}$ is dominant, and only $S_{\rm ex}$ and $S_{\rm sign}$ may lead to ambiguities.

In the resonant regime, the K^* amplitude strongly dominates the $\overline{b} \to \overline{c}u\overline{s}$ decay, and $S^{K^*+}_{\rm ex}$ and $S^{K^*-}_{\rm ex}$ become the important ambiguities.
In the transition between these regimes, the operations of Table 5.2 do not lead to clear ambiguities, as has been verified by simulation (See Section 5.7). Thus, while naively one may expect a 2^5-fold ambiguity, in practice the observable ambiguity is no larger than eight-fold, with only the two-fold S_π being fully unresolved, in the likely case of non-negligible resonant contribution.

Furthermore, although one may write down more products of the operations $S_{\rm ex}$, $S_{\rm sign}$, $S^{K^*+}_{\rm ex}$, and $S^{K^*-}_{\rm ex}$, only the products listed in Table 5.2 result in full or partial invariance of both cosines which dominate the same regime. The additional products do not result in any noticeable ambiguities. Similar arguments hold for the time-dependent mode, $B^0 \to D^{\mp}K^0\pi^{\pm}$, only here the trigonometric functions are different, and thus so are the symmetries.

5.7 Simulation Studies and Measurement Sensitivity

The PDFs Eqs. (5.12-5.15) can be used to conduct a full data analysis. Given a sample of N_e signal events, γ and the other unknown parameters of Eq. (5.19) are determined by minimizing the negative log likelihood function

$$\chi^2 \equiv -2\sum_{i=1}^{N_e} \log P(\xi_i). \tag{5.24}$$

where ξ_i are the Dalitz plot variables of event i. However, before engaging into such an enterprise, simulation of the analysis seems in place. By generating the (roughly) estimated number of events, and then fitting these with the models above, not only the sensitivity of the methods can be estimated, but also their dependence on the various parameters can be determined.
In what follows, important properties of the methods are discussed by considering the illustrative case, in which the $\overline{b} \to \overline{u}c\overline{s}$ decay proceeds only via a non-resonant amplitude, and the $\overline{b} \to \overline{c}u\overline{s}$ decay has a non-resonant contribution and a single resonant amplitude.
To study the feasibility of the analysis using Eq. (5.24) and verify the predictions of Section 5.6, simulations of both the time-independent $B^\pm \to D^0K^\pm\pi^0$ and the time-dependent $B^0 \to D^\mp K^0\pi^\pm$analysis were conducted.

As very little is known about three-body decays of the B meson in general, especially those which are Cabibbo-suppressed, the only experimental inputs are the resonant two-body decay branching ratios (see Table 5.3), which have only recently been measured or for which limits have been set [PDG03]. In the simulations conducted, the branching fractions used were consistent with the values listed in Table 5.3.

Final state	Two-Body mode	Br ($\times 10^{-4}$)
$B^\pm \to D^0K^\pm\pi^0$	$B^+ \to \overline{D}^0K^*(892)^+$	6.1 ± 2.3
	$B^+ \to \overline{D}^*(2010)^0K^*(892)^+$	7.2 ± 3.4
$B^0 \to D^0K^+\pi^-$	$B^0 \to \overline{D}^0K^*(892)^0$	0.48 ± 0.12
	$B^0 \to D^0K^*(892)^0$	< 0.18(90%CL)
	$B^0 \to \overline{D}^*(2010)^0K^*(892)^0$	< 0.68(90%CL)
	$B^0 \to D^*(2010)^0K^*(892)^0$	< 0.48(90%CL)
$B^0 \to D^\mp K^0\pi^\pm$	$B^0 \to D^-K^*(892)^+$	3.7 ± 1.8
	$B^0 \to D^*(2010)^-K^*(892)^+$	3.8 ± 1.5

Table 5.3: Branching fractions (or limits) on quasi two-body decays with the same final state as the three-body decays in question [PDG03].

For the simulation, events were generated according to the PDFs of Eq. (5.21) for the $B^\pm \to D^0K^\pm\pi^0$ mode, and Eqs. (5.12–5.15) for the $B^0 \to D^\mp K^0\pi^\pm$ mode, with the base parameter values given in Table 5.4.

Parameters with a tilde denote the "true" values used to generate events, while the corresponding plain symbols represent the "trial" parameters obtained from the simulation. The only non-vanishing amplitudes in the simulation were the non-resonant amplitudes in the $\overline{b} \to \overline{c}u\overline{s}$ and $\overline{b} \to \overline{u}c\overline{s}$ decays, and the K^* resonant $\overline{b} \to \overline{c}u\overline{s}$ amplitude. For simplicity, additional resonances were not included in this demonstration. However, (broad) resonances that are observed in the data should be included in the actual data analysis. This is conservative, as broad resonances tend to constrain the fit. The simulations were conducted with a benchmark integrated luminosity of 400 fb^{-1}, which each of the B factories plan to collect by about 2006.

Parameter	Value	Parameter	Value
$\tilde{\gamma}$	1.2	$\tilde{\beta}$	0.4
$\Delta\tilde{\delta}(\xi)$	0	$\tilde{A}_{\mathcal{U}0}/\tilde{A}_{\mathcal{C}0}$	0.4
$\Delta\tilde{\delta}_{K^*}$	1.8	$\tilde{A}_{\mathcal{C}K^*}/\tilde{A}_{\mathcal{C}0}$	1.0
$\Delta\tilde{\delta}_0$	0.4/1.0	$\tilde{A}_{\mathcal{C}K^*}$	$\sim \sqrt{\text{Br} \times \Gamma_B}$

Table 5.4: Base parameters used to generate events in the simulation. The value of $\tilde{A}_{\mathcal{C}K^*}$ is chosen so as to roughly agree with the measurement of the corresponding branching fraction [PDG03], taking into account the K^{*+} branching fractions. The value of $\tilde{\delta}_{\mathcal{U}0}$ used for the $B^\pm \to D^0 K^\pm \pi^0$ mode was 0.4, while it was 1.0 for the $B^0 \to D^\mp K^0 \pi^\pm$ mode, in order for the ambiguities not to fall on top of each other.

Mode of $B^\pm \to D^0 K^\pm \pi^0$	N_{signal}	Mode of $B^0 \to D^\mp K^0 \pi^\pm$	N_{signal}
$B^+ \to \overline{D}^0 K^+ \pi^0 = B^- \to D^0 K^- \pi^0$	1305	$B^0 \to D^- K^0 \pi^+$	112
$B^+ \to D^0 K^+ \pi^0 = B^- \to \overline{D}^0 K^- \pi^0$	103	$\overline{B}^0 \to D^+ K^0 \pi^-$	111
$B^+ \to D^0_{CP=\pm1} K^+ \pi^0$	186	$B^0 \to D^+ K^0 \pi^-$	33
$B^- \to D^0_{CP=\pm1} K^- \pi^0$	234	$\overline{B}^0 \to D^- K^0 \pi^+$	33

Table 5.5: The numbers of events obtained by averaging 100 simulations using the parameters of Table 5.4 and the reconstruction efficiencies listed in the text for an integrated luminosity of 400 fb^{-1}. Note that the $B^0 \to D^\mp K^0 \pi^\pm$ events are perfectly tagged.

The final state reconstruction efficiencies were calculated based on the approximate capabilities of current $\Upsilon(4S)$ detectors. The efficiency was assumed to be 80% for $\pi^\pm$ and 70% for $K^\pm$, both including particle identification. A reconstruction efficiency of 60% was assumed for π^0 and 50% for K^0_S including branching fraction to $\pi^+\pi^-$. The product of reconstruction efficiencies and branching fractions of the D^0, summed over the final states $K^-\pi^+$, $K^-\pi^+\pi^0$, and $K^-\pi^+\pi^-\pi^+$, was taken to yield an effective efficiency of 6%. Using the CP-eigenstate final states K^+K^-, $\pi^+\pi^-$, $K_S\pi^0$, and $K_S\rho^0$, the efficiency for the sum of the $D_{CP=\pm1}$ final states is 0.8%. Similar considerations for the $D^\pm$, reconstructed only into the final state $K^-\pi^+\pi^+$, yielded 4%.

In addition to these (in)efficiencies, the overall number of events was further reduced by 25% for both modes, in order to approximate the effect of background suppression. Furthermore, the tagging efficiency (ability to determine the flavor of the other side B^0 meson) used for the $B^0 \to D^\mp K^0 \pi^\pm$mode was 27%.

The numbers of signal events obtained in each of the final states with the parameters of Table 5.4 and the above efficiencies are listed in Table 5.5.

The distribution of events in the Dalitz plot can be seen in Fig. 5.10 (for the $B^\pm \to D^0 K^\pm \pi^0$ mode), where additional resonances (between DK and $D\pi$) have been added for illustration. The generated events are then fitted with the corresponding PDFs by minimizing the χ^2 of Eq. (5.24). Generally all parameters are left floating in the fit. It has been verified, that, given enough statistics and/or repeated enough times, the fit obtains the correct input values for all parameters. It has also been checked that the output values of the fit does not depend on any of the other input parameters. Other checks have also been performed to ensure that the simulation works properly.

In Figs. 5.11 through 5.13, is shown the dependence of χ^2 on the values of γ and $\Delta\delta_0$, the weak and strong phase, respectively, for the decay $B^\pm \to D^0 K^\pm \pi^0$.

For each of the figures, a one-dimensional minimum projection $\chi^2(\gamma) = \min\{\chi^2(\gamma, \Delta\delta_0)\}$ is also displayed, showing the smallest value of χ^2 for each value of γ. As the strong phase is unknown, this figure shows the sensitivity to γ. In both types of plots, the smallest value of χ^2 is shown as zero (white), and the edge of the white area roughly corresponds to 2.5σ.

In the first two figures (Fig. 5.11 and 5.12), the resonant and non-resonant $b \to c$ contributions, respectively, have been set to zero, as these examples are very instructive, while in the final simulation (Fig. 5.13) a more realistic simulation is performed with the parameters of Table 5.4.
At each point in these figures, χ^2 is calculated with the generated values of the amplitude ratios $A_{\mathcal{U}0}/A_{\mathcal{C}0} = \tilde{A}_{\mathcal{U}0}/\tilde{A}_{\mathcal{C}0}$ and $A_{\mathcal{C}K^*}/A_{\mathcal{C}0} = \tilde{A}_{\mathcal{C}K^*}/\tilde{A}_{\mathcal{C}0}$. This is done out of time constraints, as scanning the entire plane would otherwise be too time consuming, for this purely illustrative exercise.

When these amplitude ratios are determined by a fit simultaneously with the phases, the correlations between the amplitudes and the phases are generally found to be less than 20%, and it has been tested, that the difference between the two approaches is small. Therefore, the results obtained with the amplitudes fixed to their true values are sufficiently realistic for the purpose of this demonstration. The estimated sensitivity of the method is of course based on fits where all parameters are left floating.

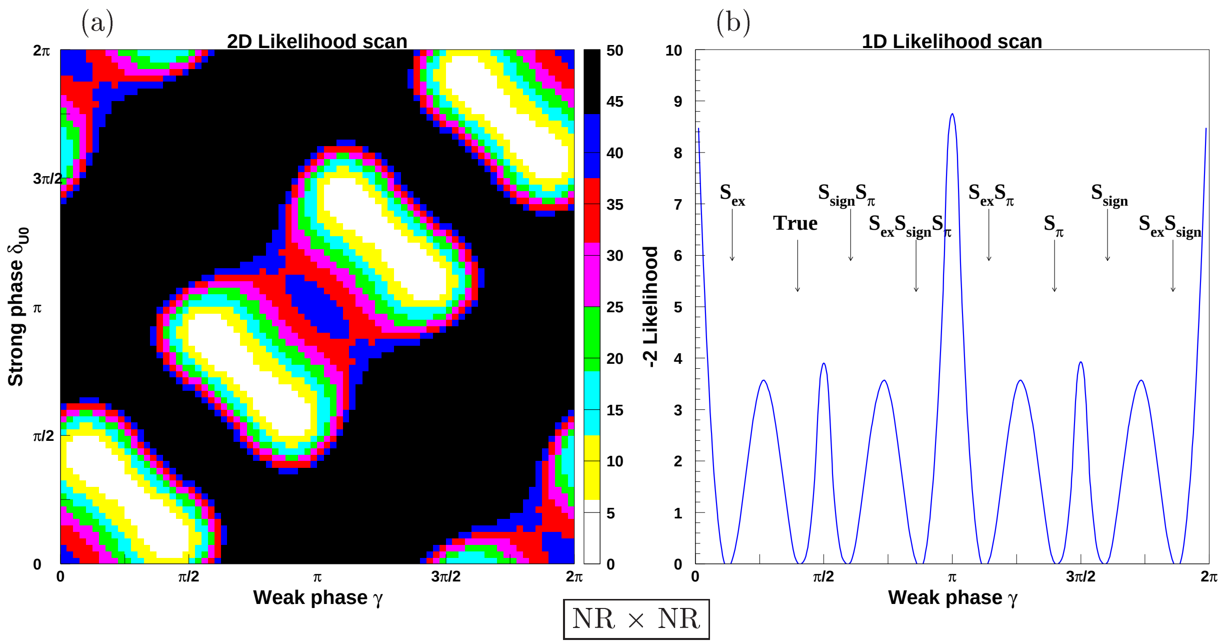


Figure 5.11: (a) χ^2 for the $B^\pm \to D^0 K^\pm \pi^0$ decay, as a function of γ and $\Delta\delta_0$, with the parameters of Table 5.4 and no resonant contribution ($\tilde{A}_{\mathcal{C}K^*} = 0$). (b) Minimum projection of χ^2 onto γ, with the eight degenerate ambiguities clearly visible (and labeled). The edge of the white area roughly corresponds to 2.5σ.

Fig. 5.11 shows results from a simulation obtained with the parameters of Table 5.4, but with $A_{\mathcal{C}K^*} = 0$. With no resonant contribution and thus no changing strong phases, the eight-fold ambiguity of the perfect non-resonant regime is clearly visible. This is the typical case for two-body final states.

Fig. 5.12 is obtained with the parameters of Table 5.4, but with $A_{\mathcal{C}0} = 0$. With no non-resonant $\overline{b} \to \overline{c}u\overline{s}$ contribution, the eight-fold ambiguity of the perfect resonant regime is seen. The ambiguities corresponding to approximate invariance are clearly resolved, with the doubly-approximate $S_{\rm ex}^{K^{*+}} S_{\rm ex}^{K^{*-}}$ ambiguity resolved more strongly.

Fig. 5.13 is obtained with the parameters of Table 5.4 and shows the combination of the two former cases. With equal resonant and non-resonant $\overline{b} \to \overline{c}u\overline{s}$ amplitudes, only the

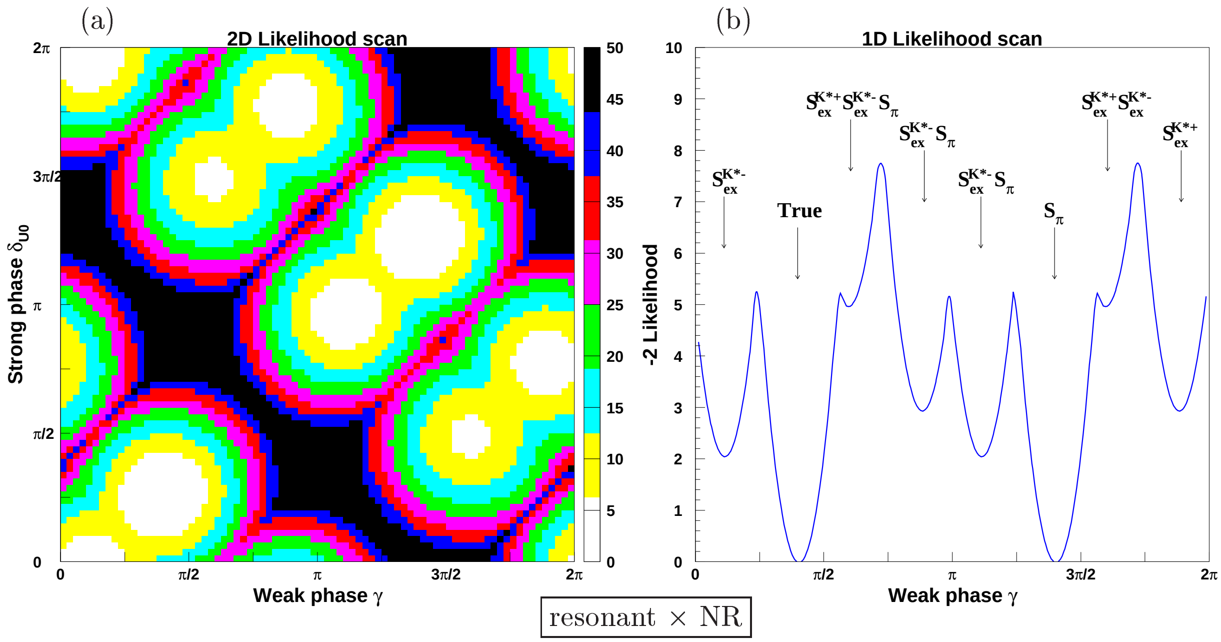


Figure 5.12: (a) χ^2 for the $B^\pm \to D^0 K^\pm \pi^0$ decay, as a function of γ and $\Delta\delta_0$, with no non-resonant $\overline{b} \to \overline{c}u\overline{s}$ contribution ($\tilde{A}_{\mathcal{C}0} = 0$). For illustration, the value $\delta_{K^*} = 1.2$ is used, such that ambiguities do not overlap and thus can be seen. All other parameters are those of Table 5.4. (b) Minimum projection of χ^2 onto γ, with the eight *resonant* and somewhat resolved ambiguities visible (and labeled).

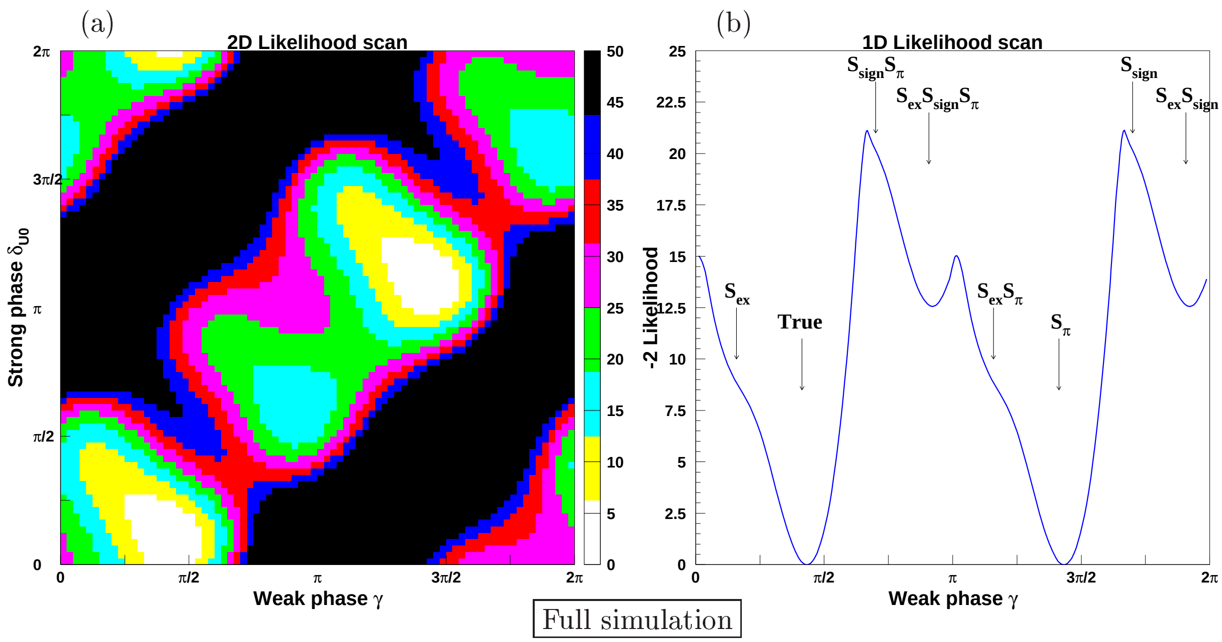


Figure 5.13: (a) χ^2 for the $B^\pm \to D^0 K^\pm \pi^0$ decay, as a function of γ and $\Delta\delta_0$, with the parameters of Table 5.4. (b) Minimum projection of χ^2 onto γ with the eight partially resolved ambiguities indicated.

non-resonant regime ambiguities are observed, due to the relative suppression of the resonant interference terms (discussed in Section 5.6). Nonetheless, the $c^\pm_{K^*0}$ terms are significant enough to resolve all but the S_π ambiguity. $S_{\rm sign}$ is more strongly resolved, since it leaves neither of the $c^\pm_{K^*0}$ terms invariant.

Repeating the exercise for the decay $B^0 \to D^{\mp} K^0 \pi^{\pm}$ with all the parameters of the simulation as in Table 5.4, the result of the time-dependent fit is shown in Fig. 5.14. Once again the minimum projection onto the axis of the weak phase (this time $2\beta + \gamma$) is also shown.

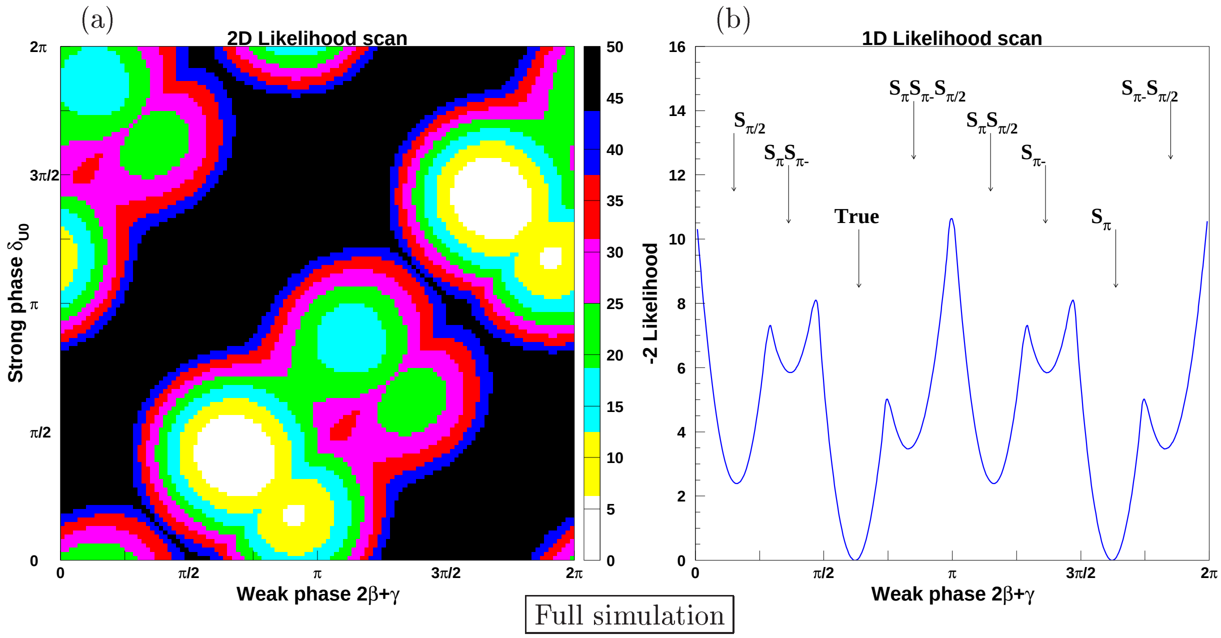


Figure 5.14: (a) χ^2 for the decay $B^0 \to D^{\mp} K^0 \pi^{\pm}$, as a function of $2\beta + \gamma$ and $\Delta\delta_0$, with the parameters of Table 5.4. (b) Minimum projection of χ^2 onto γ, with the eight ambiguities shown and labeled.

In the $B^0 \to D^{\mp} K^0 \pi^{\pm}$ case, the pattern is the same as in the time-independent case, namely that all ambiguities except S_π are resolved, as expected.
In general the cause of the ambiguity resolving is twofold. Both the varying strong phase, and the different contributions (NR $\times$ NR and resonant $\times$ NR), which have different strong phases, help resolve the ambiguities (see Fig. 5.8).

The statistical error, σ_γ, in the measurement of γ, obtained by fitting simulated event samples using the MINUIT package [JR75], as a function of one of the parameters of Table 5.4 is presented in Figs. 5.15 and 5.16 for the $B^{\pm} \to D^0 K^{\pm} \pi^0$ mode. While varying one parameter, all the other parameters used for generating events in the simulation were those listed in Table 5.4. Similarly, the statistical error, $\sigma_{2\beta+\gamma}$, from the time-dependent $B^0 \to D^{\mp} K^0 \pi^{\pm}$ simulation is shown in Figs. 5.17 and 5.18.

Each point in these plots is obtained by repeating the simulation 250 times, to minimize sample-to-sample statistical fluctuations. All the parameters of Table 5.4 were determined by the fit, thus the sensitivity obtained is the most indicative. The arrows in these figures show the value corresponding to the parameters of Table 5.4. The total number of signal events in all final states combined is the same for each of the data points. The error bars describe the statistical error at each point, which is determined by the number of experiments simulated.

From Fig. 5.15 one observes that σ_γ has a mild dependence on $\tilde{\gamma}$ and $\Delta\tilde{\delta}_0$. This is not surprising, as e.g. a value of γ around $\pi/2$ result in a large difference between the $CP = \pm 1$ amplitudes, given the strong phases chosen.

Fig. 5.16 shows that the precision on γ is independent of $\Delta\tilde{\delta}_{K^*}$, but strongly depends on $\tilde{A}_{\mathcal{U}0}/\tilde{A}_{\mathcal{C}0}$, as expected. In fact, this dependence is the reason for considering three-body

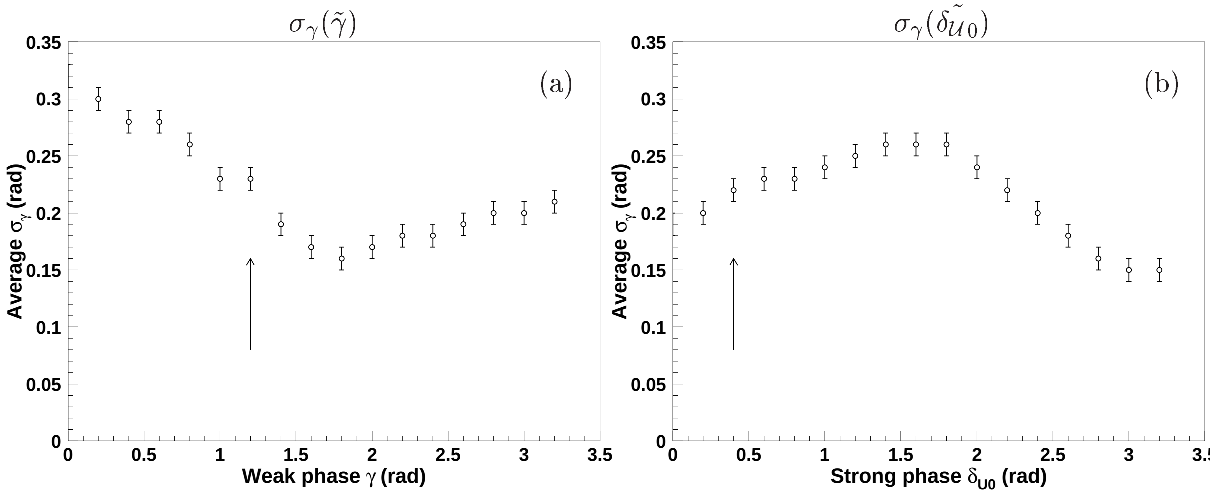


Figure 5.15: The error on γ, σ_γ, as a function of (a) $\tilde{\gamma}$ and (b) $\tilde{\delta}_{\mathcal{U}0}$ in the $B^\pm \to D^0 K^\pm \pi^0$ mode. The arrows indicate the base values, and apart from the variable in question, all other variables are kept constant.

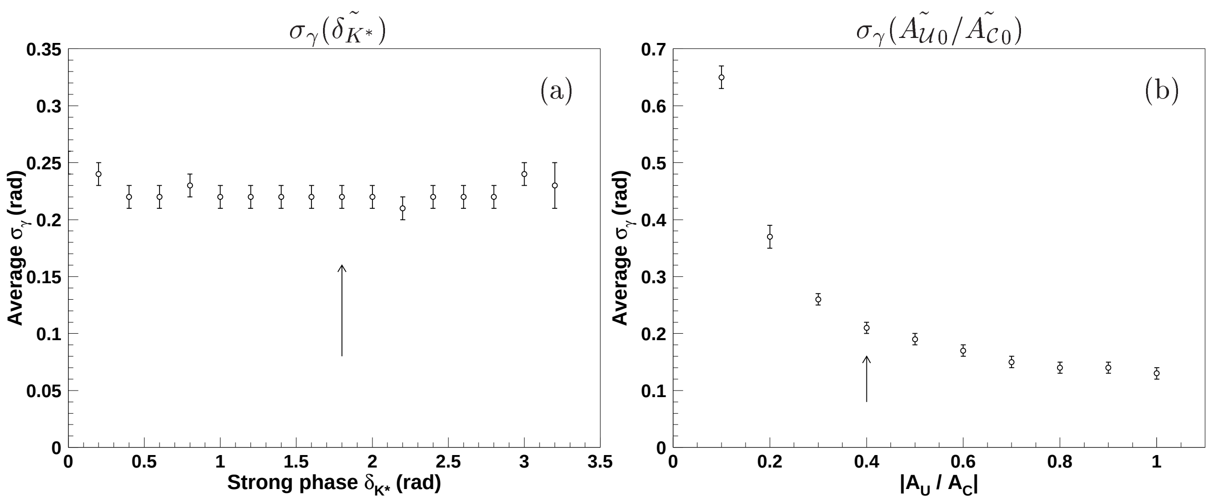


Figure 5.16: The error on γ, σ_γ, as a function of (a) $\tilde{\delta}_{K^*}$ and (b) $\tilde{A}_{\mathcal{U}0}/\tilde{A}_{\mathcal{C}0}$ in the $B^\pm \to D^0 K^\pm \pi^0$ mode. The arrows indicate the base values, and apart from the variable in question, all other variables are kept constant.

decays in the first place, as an increase in the amplitude ratio significantly increases the sensitivity of the mode. At a value of the amplitude ratio of 0.4, the sensitivity is very high, thus the choice of the final state, $B^\pm \to D^0 K^\pm \pi^0$, serves it purpose (if the amplitude ratio is indeed 0.4). However, the $B^0 \to D^0 K^+ \pi^-$ mode, which could have an amplitude ratio around 1.0, would still benefit from this, as the curve in Fig. 5.16 (right) still drops.
The variation of the parameters suggest that a significant sensitivity is obtained over a broad range of parameters. With the parameters of Table 5.4, the precision expected is $\sigma_\gamma \approx 0.23 = 13^\circ$ with an integrated luminosity of 400 fb^{-1}.

As can be seen from Fig. 5.17, the resolution on $2\beta + \gamma$ does not change considerably with the input values $2\tilde{\beta} + \tilde{\gamma}$ and $\Delta\tilde{\delta}_0$. Fig. 5.18 shows once again that the value of the strong phase of the K^* resonance, $\Delta\tilde{\delta}_{K^*}$ does not change the precision of the method, while the ratio of amplitudes $r = \tilde{A}_{\mathcal{U}0}/\tilde{A}_{\mathcal{C}0}$ is of greatest importance.

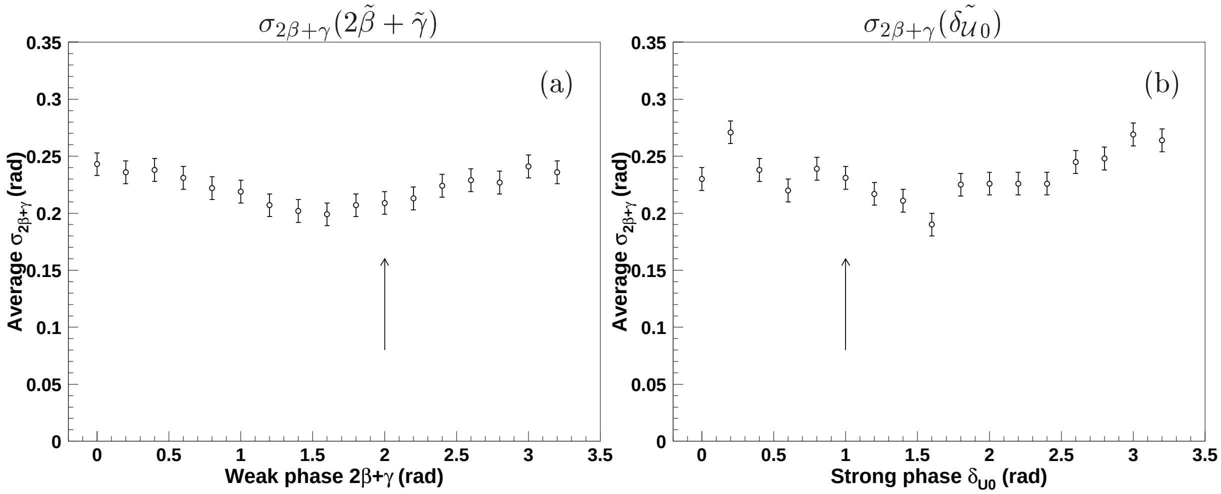


Figure 5.17: The error on $2\beta+\gamma$, $\sigma_{2\beta+\gamma}$, as a function of (a) $2\tilde{\beta}+\tilde{\gamma}$ and (b) $\Delta\tilde{\delta}_0$ in the $B^0 \to D^{\mp}K^0\pi^{\pm}$ mode. The arrows indicate the base values, and apart from the variable in question, all other variables are kept constant.

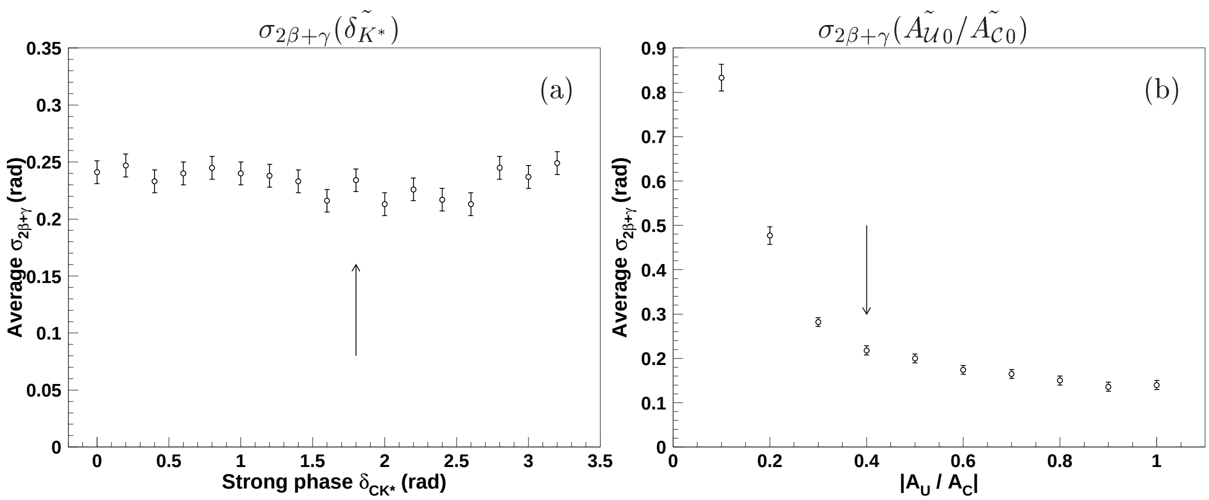


Figure 5.18: The error on $2\beta+\gamma$, $\sigma_{2\beta+\gamma}$, as a function of (a) $\tilde{\delta}_{K^*}$ and (b) $\tilde{A}_{\mathcal{U}0}/\tilde{A}_{\mathcal{C}0}$ in the $B^0 \to D^{\mp}K^0\pi^{\pm}$ mode. The arrows indicate the base values, and apart from the variable in question, all other variables are kept constant.

In general, the time-dependent method seems slightly less sensitive to the input values, than the time-independent, but the ratio of amplitudes remains very dominant. With the parameters of Table 5.4, the precision is $\sigma_{2\beta+\gamma} \approx 0.24 = 14°$ given an integrated luminosity of 400 fb^{-1}.

It is reassuring, that the feasibility of the methods is not very sensitive to the value of the parameters, except for the ratio of amplitudes, which is the primary reason for considering the three-body modes.

It has been checked, that leaving a possible $b \to u\ K^*$ component free in the fit does not change the conclusions, and should the $b \to u$ transition happen to actually have a K^* component (e.g. rescattering), then this improves the sensitivity of the methods, as large interference then occurs in that region. However, as the two strong phases change identically, the phase difference remains constant, and such a component does not help to resolve the ambiguities (essentially the methods then boils down to the two-body case).

5.7.1 General use of three-body modes

After the above analysis using three-body decays had been suggested, it was pointed out that γ can be *limited* by considering general asymmetries in three-body decays [Gro03]. Expanding on an idea from two-body decays, one defines ratios of partial rates, from which limits on γ can be made:

$$\sin^2\gamma \quad \le \quad R_{CP\pm} \;\equiv\; \frac{2[BR(B^- \to D^0_{CP\pm}X_s^-) + BR(B^+ \to D^0_{CP\pm}X_s^+)]}{BR(B^- \to D^0 X_s^-) + BR(B^+ \to \overline{D}^0 X_s^+)}. \tag{5.25}$$

The method has the great advantage, that a Dalitz analysis is not required, and since one is allowed to integrate over *any* part of the phase space, high purity and thereby sensitivity can be obtained. However, the value of the fraction in Eq. 5.25 should be kept blind until the area to consider has been settled upon, to avoid being biased towards statistically large asymmetries.

5.8 Section summary and conclusion

The unitary angle γ is difficult to measure. The original GW method does not work, and efforts to avoid its shortcommings are costly in terms of sensitivity. The central problem is the ratio of amplitudes, $r \equiv |A(b \to u)/A(b \to c)|$, which is small due to color-suppression. Furthermore, any measurement of γ in these modes will have an eight-fold ambiguity in $[0, 2\pi]$.

The three three-body decays channels $B^\pm \to D^0K^\pm\pi^0$, $B^0 \to D^0K^+\pi^-$, and $B^0 \to D^\mp K^0\pi^\pm$ are color-allowed and sensitive to the value of γ, both through time-independent ($B^\pm \to D^0K^\pm\pi^0$ and $B^0 \to D^0K^+\pi^-$) and time-dependent ($B^0 \to D^\mp K^0\pi^\pm$) approaches. The absence of color-suppression in the $\overline{b} \to \overline{u}c\overline{s}$ amplitudes is expected to result in relatively large rates and significant CP violation effects, and hence favorable experimental sensitivities.

The largest obstacle is that Dalitz plot analyses are required, which constitutes experimental complication, not the least in terms of background fighting, which has not been included in the simulations conducted. However, the methods are effective for reducing the eight-fold ambiguities that constitute a serious limitation for two-body modes – the reduction being a "by-product" of the Dalitz approach.
A general conclusion is that every effort to increase the interference and the sensitivity to γ leads to lower branching fraction and more complex analysis, and vice versa.

At the risk of stating the obvious, the final answer on the sensitivity of the methods can of course only be determined by carrying out the analyses, but perhaps the statement is more true in this case than others, as so many of the decisive parameters and distributions are unknown. Most likely a large fraction of the $b \to c$ contribution will proceed through the $K^{*\pm}$ resonance, but the real question is where the rarer $b \to u$ contribution falls, and if the two interfere. Indeed, very little is known about three-body B decays, where the bulk of the physics remains to be measured.

6 Current knowledge of the CKM matrix

It's as if kaons represent a 3-note piano, whereas B-mesons give you the whole keyboard. Each tune you can play gives you different information about the source of CP violation.

[K. Peach, Rutherford Appleton Laboratory]

Assuming unitarity, the CKM matrix has only four independent parameters (see Section 3.1), which influence *all* CP asymmetries and electroweak hadronic decay rates. Thus from combining these inherently correlated quantities in a global fit, the original four CKM parameters, whichever parametrization chosen, can be extracted.
The precision with which the CKM matrix elements are known has increased rapidly over the last three years, during the era of B factories, mostly due to $\sin(2\beta)$ measurements, but also thanks to improvements in the measurements of V_{ub}, V_{cb} and the progress in theoretical understanding.
The goal of combining the various CKM related quantities is not only to obtain the most accurate values of the four SM parameters, but more importantly, to test whether CP violation can be described within the SM.

In the following section the other systems in which CP can manifest itself and the current measurements of the relevant parameters will be reviewed and the combined result and thereby knowledge of the CKM matrix summarized. Many groups have pursued the latter task [AL94, PS99, C$^+$01], but the following short description will be based on [HLLLD01b, HLLLD01a].

6.1 *CP* violation in other systems

The short lifetime of the top quark does not allow for it to form bound states, leaving the B system as the only possibility for directly studying the third generation of the CKM matrix. While the B_s^0 meson is hard to produce in large quantities, it is very interesting from a theoretical point of view.

However, other quarks also provide information about the CKM matrix elements, and have neutral systems with possible mixing and CP violation, namely the $K^0\overline{K}^0$ system based on the s quark and the $D^0\overline{D}^0$ based on the c quark.

6.1.1 The K^0 system

The K system is with regards to CP violation a very well studied system, where all three types of CP violation has been observed. The physical states with CP and CPT violating parameters, ϵ and Δ respectively written explicitly, look as follows:

$$|K_S^0\rangle = \frac{1}{\sqrt{2}}\frac{1}{\sqrt{1+|\epsilon+\Delta|^2}}((1+(\epsilon+\Delta))|K^0\rangle + (1-(\epsilon+\Delta))|\overline{K}^0\rangle), \qquad (6.1)$$

$$|K_L^0\rangle = \frac{1}{\sqrt{2}}\frac{1}{\sqrt{1+|\epsilon+\Delta|^2}}((1+(\epsilon-\Delta))|K^0\rangle - (1-(\epsilon-\Delta))|\overline{K}^0\rangle), \qquad (6.2)$$

where the notation S (Short) and L (Long) steams from the dominating lifetime difference ($\tau_{K_S^0} = (0.8935 \pm 0.0008) \times 10^{-10} s$, $\tau_{K_L^0} = (5.17 \pm 0.04) \times 10^{-8} s$ [PDG02]). This feature makes it experimentally easy to separate the two K^0 states, while in the B^0 system one considers B^0_{phys} and $\overline{B}^0_{\text{phys}}$ and tag them according to flavor specific decays.

The K^0 mixing is very similar to B^0 mixing in having u-type quarks in the loops of the mixing diagram (see Fig. 4.1). However, in K^0 mixing top-charm and charm-charm loop contributions cannot be neglected, and this makes it difficult to extract information from observables in the K system, even though they are very well measured ($\Delta m_K = (0.5303 \pm$

$0.0009) \times 10^{10}\ \mathrm{s}^{(-1)}$). The measured CP violating quantities in the K system are the following [PDG02]:

$$\begin{aligned}
|\eta_{00}| &\equiv \left|\mathcal{A}(K^0_L \to \pi^0\pi^0)/\mathcal{A}(K^0_S \to \pi^0\pi^0)\right| = (2.274 \pm 0.017) \times 10^{-3}, && (6.3)\\
|\eta_{+-}| &\equiv \left|\mathcal{A}(K^0_L \to \pi^+\pi^-)/\mathcal{A}(K^0_S \to \pi^+\pi^-)\right| = (2.286 \pm 0.017) \times 10^{-3}, && (6.4)\\
|\epsilon_K| &= |(p-q)/(p+q)| = (2.282 \pm 0.017) \times 10^{-3}, && (6.5)\\
\epsilon'/\epsilon_K &\overset{CPT}{=} \mathrm{Re}(\epsilon'/\epsilon_K) = (19.2 \pm 2.4) \times 10^{-4}, && (6.6)
\end{aligned}$$

where ϵ_K measures CP violation in mixing, while ϵ' (historically written without subscript) measures CP violation in decay, thus superweak theory[39] (which requires $\epsilon' = 0$) is excluded. CPT invariance gives a correlation between η_{00}, η_{+-} and ϵ_K, and therefore CPT invariance can be tested by considering:

$$a_{\mathrm{CPT}} = \frac{P(\overline{K}^0 \to \overline{K}^0(t)) - P(K^0 \to K^0(t))}{P(\overline{K}^0 \to \overline{K}^0(t)) + P(K^0 \to K^0(t))} = \frac{H_{11} - H_{22}}{H_{11} + H_{22}} = 4\mathrm{Re}(\Delta), \qquad (6.7)$$

where P is the (time dependent) probability. Experimental data indicates no CPT violation, since the current measurements of Δ gives [PDG02]:

$$\mathcal{Re}(\Delta) = (2.9 \pm 2.7) \times 10^{-4}, \quad \mathcal{Im}(\Delta) = (-0.8 \pm 3.1) \times 10^{-3}. \qquad (6.8)$$

The rare decays $K^\pm \to \pi^\pm \nu\bar{\nu}$ and $K^0_L \to \pi^0 \nu\bar{\nu}$ are of special interest, since large CP violation is expected with clean interpretation, which expressed in terms of ρ and η looks as follows [Nir98, BF98]:

$$\begin{aligned}
\Gamma(K^\pm \to \pi^\pm \nu\bar{\nu}) &= 8.33 \times 10^{-6} |V_{cb}|^4 X^2(x_t) \left[\eta^2 + (\rho_0 - \rho\ )^2\right] = (3.1 \pm 1.3) \times 10^{-11}, && (6.9)\\
\Gamma(K^0_L \to \pi^0 \nu\bar{\nu}) &= 3.29 \times 10^{-5} |V_{cb}|^4 X^2(x_t)\, \eta^2 = (8.2 \pm 3.2) \times 10^{-11}, && (6.10)
\end{aligned}$$

where $X^2(x_t)$ is a known function of the top mass and $\rho_0 \simeq 1.4$ is calculated from next-to-leading order EW loop contributions. Evidently, ρ and η can be cleanly extracted from measuring these rates, and so the experimental challenge of the low branching ratios is therefore met with proposals of dedicated experiments[40]. The E787 experiment at BNL have seen one charged event [E78797].

6.1.2 The D^0 system

The D^0 system is fundamentally different from the K^0 and the B^0 systems in that the mixing loop (see Fig. 4.1) contains d-type quarks instead of u-type quarks, so the heaviest quark in the loop is b instead of t. Since $m_b \ll m_t$ and since the contribution from the heaviest quark, b, is suppressed by λ^2 ($|V_{ub}V_{cb}|$ compared to $|V_{td}V_{tb}|$ for the B system), mixing in the D system is expected to be very small. In the SM one expects $x_D \equiv \Delta m_D/\Gamma_D \lesssim 0.002$, i.e. the mixing time is ~ 500 times larger than the lifetime, which makes the probability of observing mixing very small.

Most decays of the D^0 meson have flavor specific final states, resulting in a very small expected lifetime difference. Furthermore, direct CP violation in the D^0 system is expected to be insignificant, leaving only CP violation in mixing (though small) and the interference between mixing and decay as possible measurements for probing the theory (and most likely only in the $D^0 \to \pi^+\pi^-, K^\pm\pi^\mp, K^+K^-$ channels).

The currently best experimental limits on D^0 mixing are from BaBar [BABAR03a], and still above the SM prediction. However, the SM prediction has been much debated, as several effects enters in the (time-dependent) amplitude. Therefore, even if D^0 mixing is observed just below the current experimental level, it is not a sure sign of new physics [BSN95].

[39] A theory explaining CP violation as a consequence of an additional superweak force [Wol64].

[40] Kopio (BNL) and Kami (FNAL) for the neutral mode and E949 (BNL) and CKM (FNAL) for the charged.

6.1.3 The B_s^0 system

The B_s^0 system is very much like the B_d^0 system, as only the spectator quark is different, and apart from the CKM couplings, all other quantities are equal, to a fairly good approximation ($\mathcal{O}(5-10\%)$). However, the ratio of CKM couplings to the top quark, which dominates the mixing frequency in the two B systems, is $|V_{ts}/V_{td}|^2 \simeq 25$, which yields a very large (and still unmeasured) mixing frequency for the B_s^0 system.

While it hard to extract any information about the CKM-entries from a Δm_d mixing measurement alone, a ratio with Δm_s can give very precise information, as the theoretical errors are highly correlated and thus tend to cancel out. For this reason, a measurement of Δm_s is highly desirable and therefore persued intensely [PDG03, CDF03].

The different final states allow for both CP eigenstates and Cabibbo-allowed modes, which probes γ. While the CP eigenstate modes, such as e.g. $B_s^0 \to \rho K_S^0$ have very low branching fractions ($\sim 10^{-7}$), decay channels like $B_s^0 \to D_s^- K^+$ have higher branching fraction and good sensitivity to γ. Since the states common to the B_s^0 and $\overline{B}_s^0$ are no longer suppressed, the width difference is expected to be significant (see Table 4.1), though still unmeasured.

6.1.4 Comparison of the neutral systems

The different properties of the b, c, and s-based neutral quark systems, are dictated by the dominant quark decay, as is listed in Table 6.1. The size of potential CP violating effects in the three systems are characterised by the triangles in Fig. 3.1.

System	Dominant quark decay	Decay width	CP violation size
K	$s \to u$	$\propto \lambda^2 \times \mathrm{PS}$	$\propto A^2\lambda^4\eta$
D	$c \to s$	$\propto 1 \times \mathrm{PS}$	$\propto A^2\lambda^6\eta$
B	$b \to c$	$\propto A^2\lambda^4 \times \mathrm{PS}$	$\propto \lambda^2\eta$

Table 6.1: Decay width and (potential) CP violation size for the neutral systems. It is noteworthy that the product of the decay width and the size of CP violation equals the (common) area of the unitary triangles, disregarding phase space (PS).

The product of the decay width and the size of CP violation is fixed by the (common) area of the unitary triangles (disregarding phase space). This is the feature of the B system, which has both a long lifetime (and thus measuring the decay time difference is experimentally possible) and (potentially) large CP violation effects.

6.2 Experimental constraints on CKM related measurements

The parameters, which are sensitive to and can be related to CKM parameters (see Fig. 6.1a) and which have been directly measured are listed in Table 6.2.

It is interesting to note, that relating these measurements to the four CKM parameters is limited by theoretical errors in all cases *except* for the $\sin(2\beta)$ measurement. The theoretical uncertainty can in some cases and to some degree be cancelled out by making ratios and/or making more inclusive measurements (e.g. $\Delta m_d/\Delta m_s$). As many of these theoretical errors are not well understood, their treatment has a significant impact on the final results.

6.3 Combining measurements

The idea is to make a global fit to all measurements constraining the four independent CKM parameters, and then set CL limits on the parameters of interest (call them a), which in the case of the Wolfenstein parametrization are ρ and η, leaving the remaining parameters (call

Parameter	Value ± Error(s)	Reference
$\|V_{ud}\|$	$0.9717 \pm 0.0013 \pm 0.0004$	[HLLLD01a]
$\|V_{us}\|$	$0.2228 \pm 0.0039 \pm 0.0018$	[HLLLD01a]
$\|V_{ub}\|$ (incl.)	$(4.12 \pm 0.13 \pm 0.60) \times 10^{-3}$	[HLLLD01a, PDG03]
$\|V_{ub}\|$ (excl.)	$(3.35 \pm 0.20 \pm 0.50) \times 10^{-3}$	[HLLLD01a]
$\|V_{cd}\|$	0.224 ± 0.014	[HLLLD01a]
$\|V_{cs}\|$	1.04 ± 0.16	[HLLLD01a]
$\|V_{cb}\|$ (incl.)	$(42.0 \pm 0.6 \pm 0.8) \times 10^{-3}$	[HLLLD01a]
$\|V_{cb}\|$ (excl.)	$40.2\,^{+2.1}_{-1.8} \times 10^{-3}$	[PDG03, HKM+02]
$\|\varepsilon_K\|$	$(2.271 \pm 0.017) \times 10^{-3}$	[PDG02]
Δm_d	$(0.502 \pm 0.006)\ \mathrm{ps}^{-1}$	[PDG03]
Δm_s	Amplitude spectrum	[PDG03]
$\sin(2\beta)$	0.739 ± 0.048	[PDG03]
m_c	(1.2 ± 0.2) GeV	[PDG02]
$m_t(\overline{\mathrm{MS}})$	(167.0 ± 5.0) GeV	[PDG02]
m_K	(493.677 ± 0.016) MeV	[PDG02]
Δm_K	$(3.490 \pm 0.006) \times 10^{-12}$ MeV	[PDG02]
m_{B_d}	(5.2794 ± 0.0005) GeV	[PDG02]
m_{B_s}	(5.3696 ± 0.0024) GeV	[PDG02]
m_W	(80.423 ± 0.039) GeV	[PDG02]
G_F	$1.16639 \times 10^{-5}\ \mathrm{GeV}^{-2}$	[PDG02]
f_K	(159.8 ± 1.5) MeV	[PDG02]
B_K	$0.86 \pm 0.06 \pm 0.14$	[Lel01]
$\alpha_s(M_Z^2)$ (in η_{cc})	0.1172 ± 0.0020	[PDG02]
η_{ct}	0.47 ± 0.04	[HN94]
η_{tt}	0.5765 ± 0.0065	[HN94]
$\eta_B(\overline{\mathrm{MS}})$	0.55 ± 0.01	[BBL96]
$f_{B_d\sqrt{B_d}}$	$(228 \pm 30 \pm 10)$ MeV	[Bec03]
ξ	$1.21 \pm 0.04 \pm 0.05$	[Bec03]

Table 6.2: Inputs to the CKM fit. **Upper part:** Experimental determinations of the CKM matrix elements. **Middle upper part:** CP violating and mixing observables. **Middle lower part:** Parameters of the SM predictions obtained from experimental data. **Lower part:** Parameters of the SM predictions obtained from theory.

them μ) free to vary. The choice of parameters is of course arbitrary, but as the $\rho - \eta$ plane displays the unitary triangle directly, it has become standard. As most of the errors are of theoretical origin and therefore not Gaussian (and in most cases ill-defined), great care has to be taken when including these in a likelihood function. The approach is to simply let them vary freely within their range.

A χ^2 is formed, $\chi^2(a,\mu) = -2\ln(\mathcal{L}(a,\mu))$, and the global minimum, $\chi^2(a,\mu)_{min}$, is determined. Then the a space is scanned, finding the offset-corrected minimum, $\Delta\chi^2(a) = \chi^2(\mu)_{min,a} - \chi^2(a,\mu)_{min}$, that is the minimum χ^2 given a fixed value of a. Were the errors Gaussian, the Confidence Level (CL) would be calculated as, $CL(a) = Prob(\Delta\chi^2, N_{\mathrm{dof}})$, where N_{dof} is the number of μ parameters involved. As the errors are not all Gaussian, this simplified approach should be substituted with that obtained from MC simulations.

The result of scanning the $(\overline{\rho}, \overline{\eta})$ plane using the approach described is shown in Fig. 6.1b. Since $\Delta\chi^2$ will always equal zero at the minimum, the CL will always reach one. The blue color (the outermost colored region in b/w versions) approximately equals to the 95 % CL. The central values and CL corresponding to one, two and three sigma of the involved parameters as obtained from the overall fit are shown in Table 6.3.

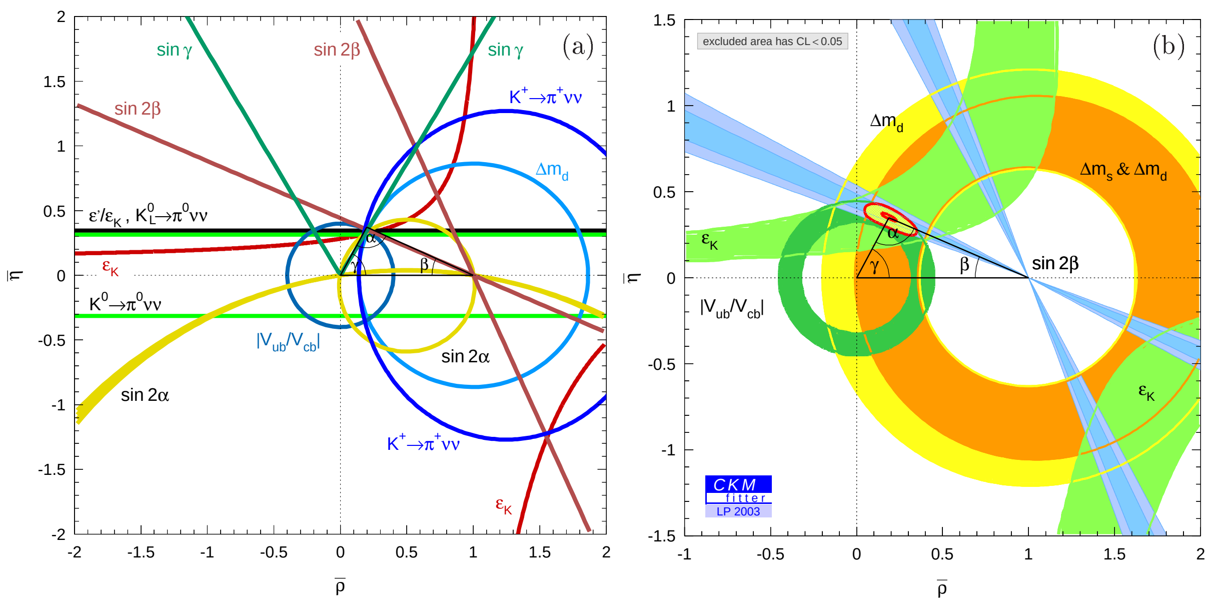


Figure 6.1: Fit to CKM parameters in the $(\overline{\rho}, \overline{\eta})$ plane. (a) Constraints obtained by perfect measurements free of theoretical errors. (b) The result of fitting all parameters that contain information about the CKM matrix elements. The superimposed blue areas correspond to the world average $\sin(2\beta)$ value (here accounted for in the fit) including one and two standard deviations respectively. See text and reference for further explanation [HLLLD01b].

6.4 Section summary and conclusion

In addition to the $B_{u,d}$ system, the light neutral meson systems carry information about the CKM matrix, but these are plagued by theoretical uncertainties. This is not the case for the B_s system, which is very interesting, but experimentally less accessible, and one will have to wait until the start of LHC-B and BTeV for high statistics samples. Combining all measurements and theoretical parameters sensitive to CKM matrix elements, one obtains both the most precise values of these, but also an overall test of the Standard Model. The current status of the overall CKM fit is, that *all measured and theoretical quantities sensitive to the CKM matrix parameters are in agreement*. The fit yields $J = (3.11\,^{+0.33}_{-0.46}) \times 10^{-5}$, in accordance with the notion that CP violation is naturally small in the quark sector.

Observable	central ± CL ≡ 1σ	± CL ≡ 2σ	± CL ≡ 3σ
λ	$0.2265\,^{+0.0025}_{-0.0023}$	$^{+0.0040}_{-0.0041}$	$^{+0.0045}_{-0.0046}$
A	$0.801\,^{+0.029}_{-0.020}$	$^{+0.066}_{-0.041}$	$^{+0.084}_{-0.054}$
$\bar{\rho}$	$0.187\,^{+0.088}_{-0.070}$	$^{+0.182}_{-0.114}$	$^{+0.221}_{-0.156}$
$\bar{\eta}$	$0.356\,^{+0.046}_{-0.042}$	$^{+0.086}_{-0.085}$	$^{+0.118}_{-0.118}$
J $[10^{-5}]$	$3.10\,^{+0.43}_{-0.37}$	$^{+0.82}_{-0.74}$	$^{+1.08}_{-0.96}$
$\lvert V_{ud}\rvert$	$0.97400\,^{+0.00054}_{-0.00058}$	$^{+0.00094}_{-0.00095}$	$^{+0.00106}_{-0.00106}$
$\lvert V_{us}\rvert$	$0.2265\,^{+0.0025}_{-0.0023}$	$^{+0.0040}_{-0.0041}$	$^{+0.0045}_{-0.0046}$
$\lvert V_{ub}\rvert$ $[10^{-3}]$	$3.87\,^{+0.35}_{-0.30}$	$^{+0.73}_{-0.60}$	$^{+0.73}_{-0.76}$
$\lvert V_{ub}\rvert$ $[10^{-3}]$ (meas. not in fit)	$3.87\,^{+0.34}_{-0.31}$	$^{+0.81}_{-0.61}$	$^{+1.27}_{-0.88}$
$\lvert V_{cd}\rvert$	$0.2264\,^{+0.0025}_{-0.0023}$	$^{+0.0040}_{-0.0041}$	$^{+0.0045}_{-0.0046}$
$\lvert V_{cs}\rvert$	$0.97317\,^{+0.00053}_{-0.00059}$	$^{+0.00094}_{-0.00097}$	$^{+0.00106}_{-0.00112}$
$\lvert V_{cb}\rvert$ $[10^{-3}]$	$41.13\,^{+1.36}_{-0.58}$	$^{+2.43}_{-1.16}$	$^{+3.08}_{-1.73}$
$\lvert V_{cb}\rvert$ $[10^{-3}]$ (meas. not in fit)	$41.2\,^{+5.1}_{-5.7}$	$^{+7.9}_{-5.8}$	$^{+9.9}_{-5.8}$
$\lvert V_{td}\rvert$ $[10^{-3}]$	$8.26\,^{+0.72}_{-0.86}$	$^{+1.23}_{-1.79}$	$^{+1.64}_{-2.25}$
$\lvert V_{ts}\rvert$ $[10^{-3}]$	$40.47\,^{+1.39}_{-0.62}$	$^{+2.42}_{-1.21}$	$^{+3.17}_{-1.78}$
$\lvert V_{tb}\rvert$	$0.999146\,^{+0.000024}_{-0.000058}$	$^{+0.000047}_{-0.000104}$	$^{+0.000070}_{-0.000133}$
$\sin 2\alpha$	$-0.14\,^{+0.37}_{-0.41}$	$^{+0.57}_{-0.71}$	$^{+0.74}_{-0.82}$
$\sin 2\beta$	$0.739\,^{+0.048}_{-0.048}$	$^{+0.096}_{-0.095}$	$^{+0.124}_{-0.137}$
$\sin 2\beta$ (meas. not in fit)	$0.817\,^{+0.037}_{-0.222}$	$^{+0.053}_{-0.279}$	$^{+0.067}_{-0.334}$
$\gamma \simeq \delta$ (deg)	$62\,^{+10}_{-12}$	$^{+17}_{-24}$	$^{+23}_{-30}$
$\sin\theta_{12}$	$0.2266\,^{+0.0025}_{-0.0023}$	$^{+0.0040}_{-0.0041}$	$^{+0.0045}_{-0.0046}$
$\sin\theta_{13}$ $[10^{-3}]$	$3.87\,^{+0.35}_{-0.30}$	$^{+0.35}_{-0.60}$	$^{+0.35}_{-0.76}$
$\sin\theta_{23}$ $[10^{-3}]$	$41.11\,^{+1.37}_{-0.58}$	$^{+2.43}_{-1.16}$	$^{+3.08}_{-1.73}$
Δm_d (ps^{-1}) (meas. not in fit)	$0.54\,^{+0.26}_{-0.21}$	$^{+0.62}_{-0.31}$	$^{+0.94}_{-0.34}$
Δm_s (ps^{-1})	$17.8\,^{+6.7}_{-1.6}$	$^{+15.2}_{-2.7}$	$^{+22.1}_{-3.7}$
Δm_s (ps^{-1}) (meas. not in fit)	$16.5\,^{+10.5}_{-3.4}$	$^{+17.7}_{-5.7}$	$^{+23.9}_{-7.2}$
ϵ_K $[10^{-3}]$ (meas. not in fit)	$2.5\,^{+1.6}_{-1.1}$	$^{+2.4}_{-1.4}$	$^{+3.1}_{-1.6}$

Table 6.3: CKM fit results and errors, in terms of CL that correspond to one-, two- and three standard deviations, respectively, using as input the observables listed in Table 6.2 (including the world average on $\sin(2\beta)$). For results marked by "meas. not in fit", the measurement of the corresponding observable has not been included in the fit.

Part II
Accelerator and Detector

The machine does not isolate man from the great problems of nature but plunges him more deeply into them.

[Antoine de Saint-Exupery, 1900-1944]

In order to pursue the program of studying CP violation in the B system, asymmetric high luminosity e^+e^- colliders operating at the $\Upsilon(4S)$ resonance, so-called B factories, were proposed [Odd]. As the branching ratio for interesting B decays is of order $\mathcal{O}(10^{-4})$ or less, one needs to produce at least 10^8 $B\overline{B}$ pairs, requiring very high luminosity and high reconstruction efficiency. Furthermore, the subsequent data analysis requires excellent Particle IDentification (PID), good calorimetry, and precise vertex resolution. The PEP-II collider and the *BABAR* detector were designed for exactly those purposes.

In the following, first the PEP-II collider and second, and more thoroughly, the *BABAR* detector will be described. The accelerator description will include motivations, some of the relevant physics, design and performance of the collider and comparison with other experiments. In the detector description, functions, design and performance will be described and discussed for each subdetector separately, apart from the tracking performance, which naturally is presented after the SVT and DCH sections.

Both the PEP-II collider and the *BABAR* detector are described in detail elsewhere (for the description of the PEP-II collider [PEP93, S+03, WWW03, BABAR02a] and for the *BABAR* detector description mainly [BABAR, BABAR02a, BABAR-DIRC03, BABAR-DIRC04] and private communations with fellow IR2-team members).

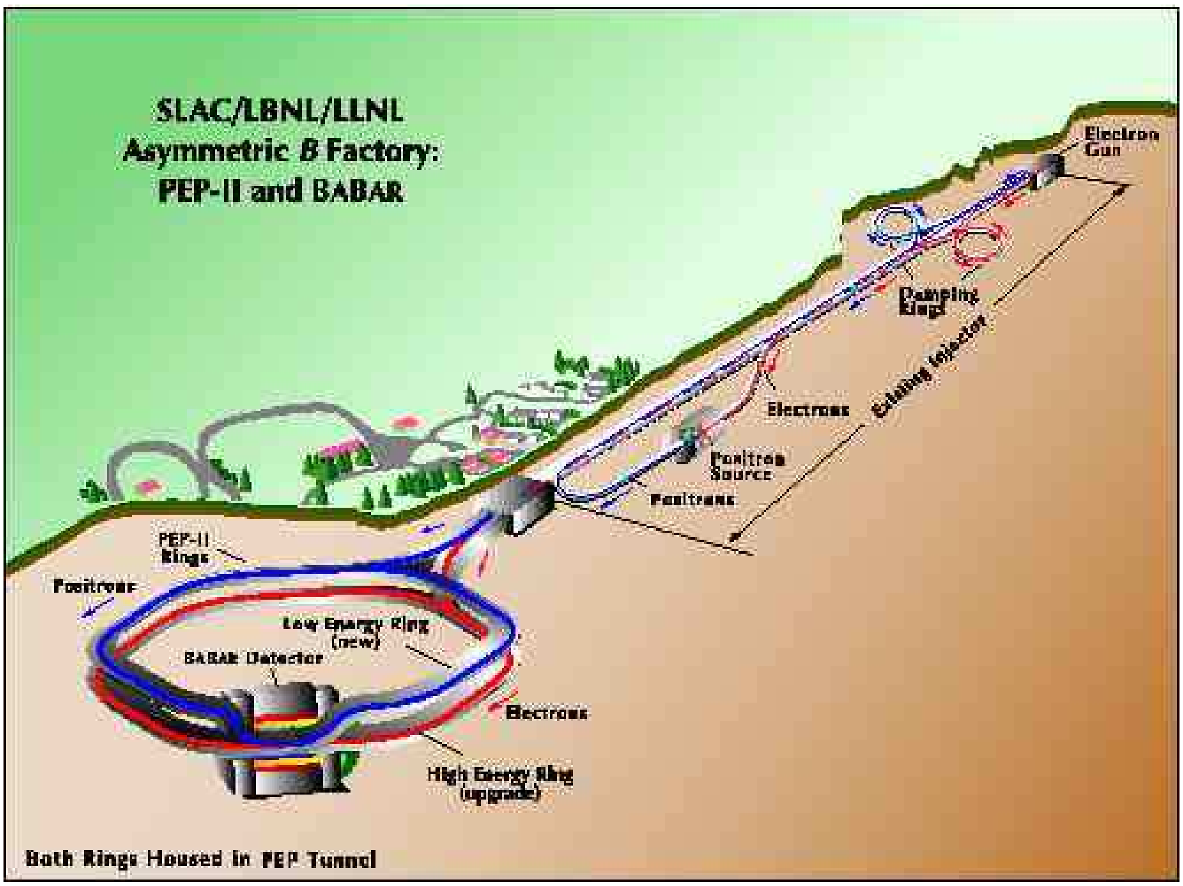


Figure 6.2: Overview of the SLAC accelerator site. The linear accelerator fills the PEP-II rings with electrons and positrons, which are brought to collide at the interaction point where the *BABAR* detector is located.

7 The PEP-II collider

7.1 The reasoning behind an asymmetric collider

One of the most distinct features about the PEP-II collider is the asymmetric beam energies. No other collider before the B factories have had this characteristic[41], but it turns out to be a central part of the design.

To be able to perform time-dependent measurements, the two times $t_{\rm rec}$ and $t_{\rm tag}$ have to be obtained (see Section 4). The most obvious way would be a "stopwatch" approach, that is measure the time of the e^+e^- collision, t_0, which is essentially the time of production for the $B\overline{B}$ pair due the short $\Upsilon(4S)$ lifetime ($\mathcal{O}(10^{-23})$), and then subsequently the two B mesons decay time, $t_{\rm rec}$ and $t_{\rm tag}$.

The first problem encountered is that the collision time can not be measured with sufficient accuracy, as its uncertainty is essentially the time it takes for the two bunches to cross each other. At PEP-II, the bunch length is $\mathcal{O}(1-2\mathrm{cm})$, which yields an uncertainty in the bunch crossing time of ~ 50ps to be compared with $\tau_B \sim 1.5$ps. However, as the time distributions Eqs. (4.25–4.26) show, it is only the time difference that is needed, and thus t_0 does not have to be determined.
The second and unrepairable problem is the determination of the two B meson decay times. The time resolution of present detector technology is $\mathcal{O}(25\mathrm{ps})$[42], which – impressive as it is – is still not in the range of the B meson life and mixing time at $\mathcal{O}(1\mathrm{ps})$. Thus, one cannot from the arrival time of the daughters of the B meson determine the time of decay with adequate precision.

Though not immediately obvious, the solution is to use the vertex position of the two B mesons. At the $\Upsilon(4S)$ resonance[43], each B meson have a boost of $\beta\gamma \simeq 0.06$ in the rest frame of the $\Upsilon(4S)$. Given the lifetime of the neutral B meson, $\tau_{B^0} = (1.542 \pm 0.016)$ps [PDG02], this results in an average flight distance of $\mathcal{O}(30\ \mu m)$, which is not resolvable by todays silicon vertex detector technology[44]. For this reason the PEP-II collider has, unlike previous e^+e^- colliders, been designed with asymmetric beam energies, giving the produced particles a boost of $\beta\gamma \simeq 0.56$, which results in resolvable average decay lengths of $\mathcal{O}(250\ \mu m)$, well within reach of modern silicon vertex detector technology (see section 8.2).
Though the vertex resolution issue applies mostly to the neutral B meson, it is also of interest for other analysis, as it is an important tool in background rejection.

7.1.1 Relation between Δz and Δt

To a good approximation, that is neglecting the modest boost from the $\Upsilon(4S)$ decay, $\beta^{\rm CM}$, and a slight tilt of the beam direction with respect to the z-axis of $\sim 1^\circ$ (see Section 8.7), the relation between the vertex distance, Δz, and the decay time difference, $\Delta t = t_{\rm rec} - t_{\rm tag}$, is linear:

$$\Delta z \quad = \quad \gamma\beta c\Delta t. \tag{7.1}$$

With a conversion factor of $\beta\gamma c = 165\mu$m/ps, this relation provides an almost one-to-one correspondence between Δz and Δt. While the lack of alignment (necessary for beam-orbit stability, see Section 8.7) is a small effect, which is easily corrected for [LeC02], the boost from the $\Upsilon(4S)$ decay is not negligible, as $\beta^{\rm CM}/\beta = 0.14$.

[41] The asymmetric energies of HERA is due to the different type of particles in the beams (e^- and p^+).
[42] The best resolution in time is currently obtained with so-called Pestov spark counters [B$^+$96].
[43] The $\Upsilon(4S)$ is the lightest $b\bar{b}$ state (21.2 ± 3.6 MeV [PDG02]) above the $B\overline{B}$ threshold (see Fig. 7.1).
[44] Emulsion detectors have a precision of a few μm, but emulsion at the vertex would destroy the beams.

Correcting for this additional boost, the correct relation between Δz and Δt becomes:

$$\Delta z = \beta\gamma c\Delta t\left(1+\frac{\beta^{\mathrm{CM}}}{\beta}\cos\theta^{\mathrm{CM}}\frac{t_{\mathrm{rec}}+t_{\mathrm{tag}}}{t_{\mathrm{rec}}-t_{\mathrm{tag}}}\right). \tag{7.2}$$

As it is not possible to measure $t_{\mathrm{rec}}+t_{\mathrm{tag}}$, its value is substituted by its average, which, given a value of Δt, is $\tau_{B^0}+|\Delta t|$. Due to the $\sin\theta$ term in the matrix element (see Eq. (4.19)), the two B mesons from the $\Upsilon(4S)$ decay are mostly emitted perpendicular to the beam direction, and consequently their difference in boost is generally small ($\mathcal{O}(\Delta\beta/\beta)\simeq 0.14$). Therefore, the approximation works quite well, even though it is mathematically inconsistent to apply it (as $t_{\mathrm{rec}}+t_{\mathrm{tag}}$ was already integrated out of the time distribution). It can be shown [Dib90, Kit03], that it is possible to make the time-dependent fit in terms of Δz, but the loss in simplicity overshadows the gain obtained.

However, the correction does not come without a price. Using Eq. (7.2) introduces (slight) correlations between Δt and m_{ES} for continuum background both through the boost and the angle. The choice between the two depends on how dominant the continuum background in the time-dependent fit is.

7.1.2 Cross-sections at the $\Upsilon(4S)$ resonance

Running at the $\Upsilon(4S)$ resonance (see Fig. 7.1) not only produces $B\overline{B}$ pairs, but also much other interesting physics as a useful byproduct. However, seen from the point of view of B physics, these byproducts are considered as backgrounds, collectively refered to as *continuum* background.

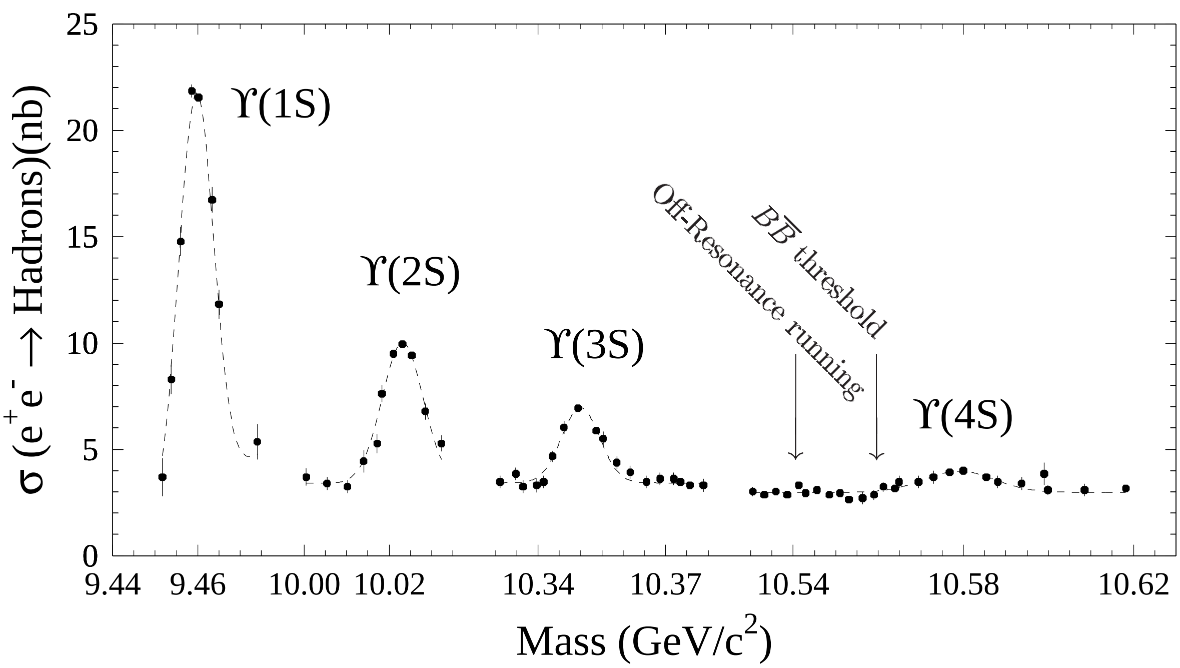


Figure 7.1: Scan of the Υ resonance region, revealing the four lowest S states and the $B\overline{B}$ threshold. The height of the $\Upsilon(4S)$ peak relative to the background level indicates the fraction of continuum background to expect. The content of this background is determined from data taken below the $B\overline{B}$ threshold. (Note: The energy scale on the horizontal axis is not continuous).

The cross-section for production of fermion pairs from e^+e^- collisions with CM energy at the $\Upsilon(4S)$ resonance is given in Table 7.1 [Har98].

Quark pair	Cross-section (nb)
$b\bar{b}$	1.05
$c\bar{c}$	1.30
$s\bar{s}$	0.35
$d\bar{d}$	0.35
$u\bar{u}$	1.39

Lepton pair	Cross-section (nb)
$\tau^+\tau^-$	0.94
$\mu^+\mu^-$	1.16
e^+e^-	~ 40

Table 7.1: Cross-sections for e^+e^- at $\sqrt{s} = m_{\Upsilon(4S)}$. These cross-sections were calculated with the Jetset 7.4 event generator [SB87].

Apart from the cross-section to e^+e^- (mainly Bhabha scattering), having been corrected for detector acceptance as it is highly angular dependent, these numbers reflect to zeroth order the classic cross section for e^+e^- collisions at high energy (where N_c is the number of colors):

$$\sigma(f\bar{f}) \quad = \quad N_c Q^2 \frac{4\pi\alpha^2}{3E_{\rm CM}^2}\sqrt{1-\frac{m_f^2}{E^2}}\left(1+\frac{1}{2}\frac{m_f^2}{E^2}\right) \overset{E_{CM}=m_{\Upsilon(4S)}}{\simeq} N_c Q^2 \; 0.78\text{nb}. \tag{7.3}$$

Taking initial state radiation into account, the *actual* CM energy has a long tail into the lower energy region. The result is that fermion pairs are also created at lower CM energies, where the cross-section is higher due to the inverse scaling with $E_{\rm CM}^2$, thus making the effective cross-section higher. Radiative corrections are expected to be largest for the light quarks, which is the reason for the larger $u\bar{u}$ than $c\bar{c}$ cross-section.
The $b\bar{b}$ cross-section constitutes roughly 25% of the entire quark cross-section.

7.2 Resonance vs. Continuum running

The background "under" the $\Upsilon(4S)$ resonance (see Fig. 7.1), is continuum. In order to make precision measurements, one needs to determine this physics background[45]. This is done by running "off peak", i.e. 40 MeV below the $\Upsilon(4S)$ resonance, where continuum is essentially unchanged, but the $B\overline{B}$ signal is no longer there, as it is below the threshold (see Fig. 7.1).

For decay channels with little background (e.g. $J/\psi K_S^0$) the optimal fraction of data taking with continuum (off-resonance) running will be low. But for channels with more backgrounds (e.g. $B^0 \to \rho^{\pm}\pi^{\mp}$), the uncertainties due to backgrounds become significant, and off-resonance running is needed.

In order to determine the optimal fraction of continuum running, c, for a specific channel with background over signal, $b \equiv B/S$ (assuming all background stems from continuum), one considers the number of events on resonance and in continuum for a fixed amount of integrated luminosity, $\mathcal{L}$, isolates the amount of signal, S, and determines its error:

$$\left.\begin{array}{l} N_{\Upsilon(4S)} = (1-c)\mathcal{L}(1+b)S \\ N_{\rm continuum} = c\mathcal{L}bS \end{array}\right\} \quad \Longrightarrow \quad \sigma(S) = \sqrt{\frac{S}{\mathcal{L}}}\sqrt{\frac{c+b}{c\,(1-c)}}. \tag{7.4}$$

Not surprisingly, the error is proportional to the square root of the signal over the integrated luminosity, which just serves as normalization constants. The minimization of the second square root yields the optimal fraction of continuum running. For no background ($b = 0$) the minimum is obviously at $c = 0$, but with just 5% background, 18% continuum running is optimal, and for $b = 1$ the optimal fraction of continuum running reaches 41%. However, this assumes that there is no other mean of constraining the continuum background parameters. At *BABAR* it has been chosen to let 12% of the luminosity be taken off-resonance.

[45] Background originating from physics processes, contrary to machine background from running the collider.

7.3 The general design of the PEP-II collider

The PEP-II collider is located at the end of the linear accelerator (Linac) situated at the SLAC site (sketched in Fig. 7.2) 40 kilometers south of San Francisco in California, USA. The three kilometer long Linac[46] injects 9.0 GeV electrons into the High Energy Ring (HER) and 3.1 GeV positrons into the Low Energy Ring (LER) of PEP-II, circulating in opposite directions. The two storage rings, which are housed in the former PEP (Positron Electron Project) tunnel of 800 meter diameter, are brought to collide at the interaction point of the *BABAR* detector. The chosen beam energies result in a Center-of-Mass (CM) energy of 10.58 GeV, which corresponds to the $\Upsilon(4S)$ resonance, at a boost of $\beta\gamma \simeq 0.56$, due to the asymmetric beam energies. The choice of boost size is a tradeoff between vertex resolution and background rejection (improved with larger boost) vs. luminosity and acceptance (improved with smaller boost). The chosen boost and thus the time resolution is such that it degrades time measurements by about 5–10% (compared to infinite resolution). A more detailed description of the PEP-II storage rings can be found at [PEP93, S$^+$03].

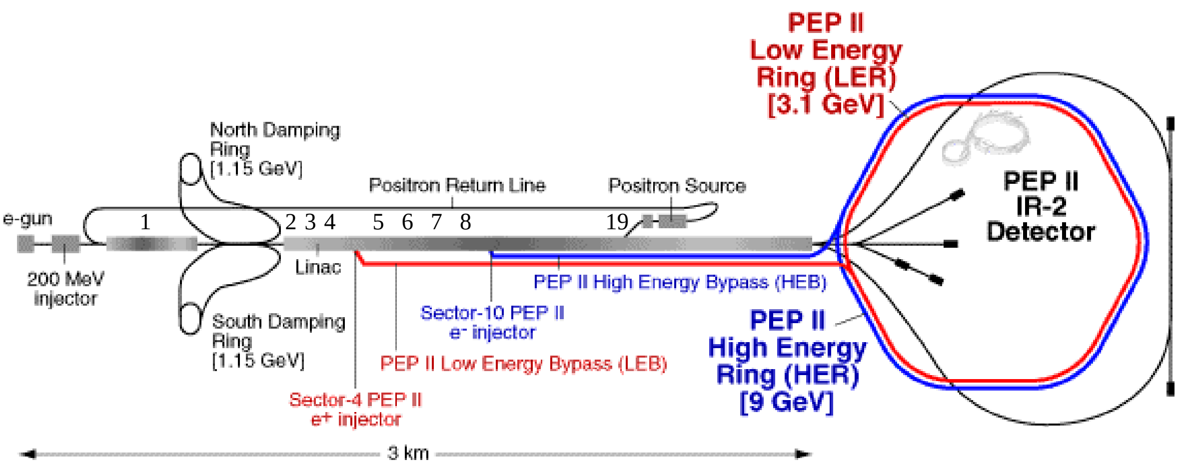


Figure 7.2: Layout of the Linac and the PEP-II collider. Electrons and positrons (created using the electron beam) are accelerated up to their energies of 9.0 and 3.1 GeV, respectively, and then injected into the two PEP-II rings with a frequency of 60 Hz.

PEP-II is designed to have 1658 particle bunches in each ring, resulting in a bunch crossing every 2.1 ns at the interaction point, where *BABAR* is located. Each bunch is designed to consist of 2.1×10^{10} electrons for the HER and 5.9×10^{10} positrons for the LER, which yields a design current of 0.75/2.15 A for the HER/LER respectively. The rings are recharged by the Linac when the luminosity drops below 90 % of its peak value, approximately every second hour. After many tests throughout 2001–2003, continuous injection (describingly termed “trickle” injection) was finally commenced for the LER in December 2003. This eliminates the injection time (though only for the LER), during which no data can be taken, and also decreases the risk of loosing the beam, while at the same time stabilising the running. With discontinuous injection, the temperature of some beam elements changes quite a lot ($\sim 20°$C). Trickle injection for the HER was regarded as more involved and subtle, but after initial successful tests, it has now become standard. With no injection time, the rate of integrating luminosity is increased by about 10-15%.
The luminosity started at an average value of $2.2 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ (Run I), and is currently (Run IV) around $6 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, which is the double of its design value. The total integrated

[46]Originally build in 1966, this linear collider was used for fixed target experiments and later collider experiments, which lead to the discovery of quarks (1968), J/ψ (1974) and the τ (1975).

Parameter	Unit	Design	Best achieved	Typical
Circumference	m	2199.318		
Number of Bunches (HER+LER)		1658	1658	1030–1230
Total Beam Current (HER)	A	0.75	1.20	1.1
Total Beam Current (LER)	A	2.15	2.43	1.6
Horizontal Spot Size Σ_X	μm	220		150
Vertical Spot Size Σ_Y	μm	6.7		5.0
Luminosity	$\mathrm{cm}^{-2}\mathrm{s}^{-1}$	3×10^{33}	7.18×10^{33}	5.5×10^{33}
Integrated Luminosity	pb^{-1}/shift	45	164.6	100
Integrated Luminosity	pb^{-1}/day	135	479.8	280
Integrated Luminosity	fb^{-1}/week	1.0	2.491	1.5
Integrated Luminosity	fb^{-1}/month	4.0	7.334	4.5
Total Int. Lum. (Run I–III)	fb^{-1}		113.27	

Table 7.2: Parameters and performance of PEP-II. The typical numbers are taken from Run-IV. The values are as of New Years Eve 2003–2004, and already significantly outdated.

luminosity as of New Years Eve 2003-2004 is 131.05fb^{-1} of which the *BABAR* detector logged 113.27fb^{-1} on the $\Upsilon(4S)$ resonance and 12.01fb^{-1} 40 MeV below. In Table 7.2 are listed various beam parameters and luminosity performances [BABAR, WWW03].

7.3.1 Determination of luminosity and beam parameters

The integrated luminosity is derived from the known QED processes $e^+e^- \to e^+e^-$ and $e^+e^- \to \mu^+\mu^-$. The measurements are consistent and have negligible statistical error. The systematic error arises from uncertainties in the MC generator and simulation of the detector and its acceptance. With improved understanding of the detector the systematic error drops, and the current level is 1.1%.

The beam energies are determined online from the magnetic bending strength and the acceleration frequency. While the rms on the absolute beam energies is 2.3/5.5 MeV for the LER/HER and the systematic error 5–10 MeV, the relative energy is stable within about 1 MeV. The CM energy, which has a tolerance of about 2 MeV, is kept at the $\Upsilon(4S)$ resonance by monitoring the $B\overline{B}$ production. The best calibration of the CM energy is obtained offline from the momentum (in CM) of fully reconstructed B candidates, which has an error of 1.1 MeV dominated by uncertainty in the B mass and the detector resolution. The spread in the beam energy, caused by minuscule variations in beam particle energies within each bunch, dominates the resolution of the key variable m_{ES} (see Section 9.2).

The beam direction, position and size relative to the *BABAR* detector is determined on a run-by-run basis by considering e^+e^- and $\mu^+\mu^-$ events. The only exception is the very small vertical beam size, which is infered from the measured luminosity. The measurements are checked offline by considering multi-hadron events and also compared to values measured by PEP-II, before stored in a condition database, which is used when (re-)processing the data.

7.3.2 Machine related constraints, backgrounds and radiation

The demand for high luminosity – unprecedented at e^+e^- machines – is achieved by a high number of bunches in the storage rings and a small transverse beam profile at the IP. However, this is not a trivial task to fulfil. With bunches every 4.2ns ($\sim$ 1.2m), the problem of parasitic collisions arises, that is collisions every 0.6m on either side of the IP, where the next bunch crossings occur. To redeem this problem, sweeping magnets (B1) are placed close to the IP, such that the two beams (but especially the LER) are bend in such a way, that they only coincide at the intended IP (see Figure 7.3). To obtain a small transverse beam profile, quadrupole magnets (Q1) are placed very close to the IP, in fact *inside* the detector along with the dipole magnets for bending the beam. This puts stringent requirements on the detector design, as it deprives the *BABAR* detector space and thus acceptance close to the beams.

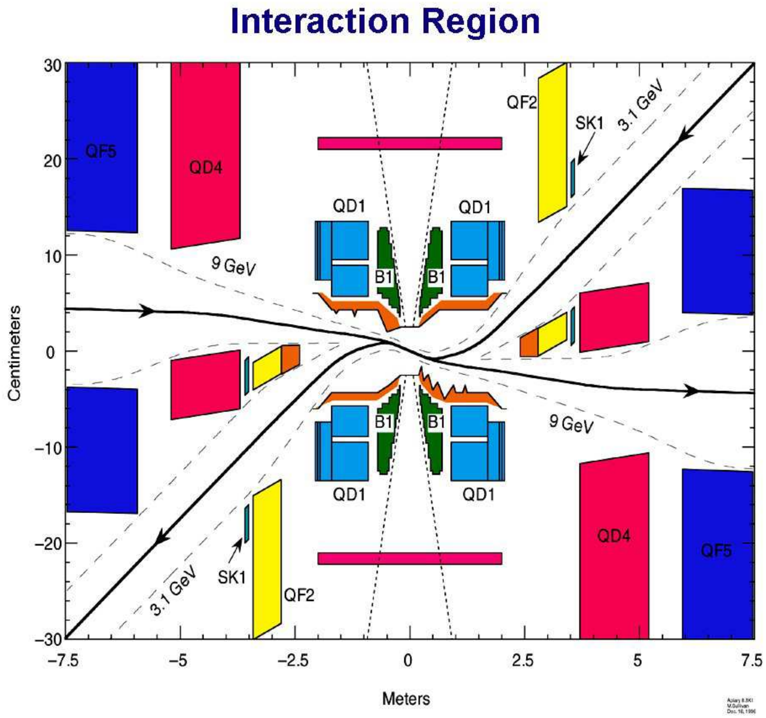


Figure 7.3: Layout of the Interaction Region (IR). Shown are the two beams and the dipole bending (B) and quadrapole focusing (Q) magnets in the IR. Note that the scale of the transverse axis is magnified 20 times.

The largest background is caused by radiative Bhabha scattering generating electromagnetic showers, when off-energy beam particles are swept into the detector. This background is very pronounced during injection, after which it drops to a more reasonable level. As the size of the background depends on the luminosity, it will in the future be even more dominating.
Another large background arises from Coulomb scattering of beam-gas molecules, which along with beam-gas bremsstrahlung is the primary source of radiation damage for the SVT. This background, which scales with the beam currents, is very large after ventilating the beam pipe (due to e.g. maintenance or leakages), but decreases with time, as Synchrotron Radiation (SR) *scrubs* (i.e. cleans) the beam pipe.
The backgrounds may fluctuate very significantly, and for this reason, radiation monitors are installed, such that one is able to abort the beam in case of unacceptably high levels.
The bending of the beams in the Interaction Region (IR) gives rise to SR, which can also damage the detector and induce backgrounds, and therefore radiation masks are installed. The SR from the closest bending magnets is designed to pass through the interaction region without interacting with the beam pipe, while other sources of SR are masked away.

These backgrounds not only cause general degradation and damage, but also increase the occupancy, effectively lowering the efficiency of the detector[47] (most critically for the IFR, cf. Section 8.8) and causing operational difficulties (most critical for the DCH, cf. Section 8.3).

[47] In May 2003 the collider switched to a new configuration, which increased the luminosity but also the background. It was found that going from the low to the higher background configuration caused the number of reconstructed $D^{*\pm}(\rightarrow D^0\pi^{\pm}_{soft})$ to drop from 8.4 pb^{-1} to 7.5 pb^{-1}, thus a sizable effect.

7.4 B-factory vs. other experiments

The advantages of an e^+e^- collider at the $\Upsilon(4S)$ resonance are numerous:

- High ratio of signal to background ($\sigma_{b\bar{b}}/\sigma_{\text{total}} = 0.24$).
- Clean events ($B\overline{B}$ production is exclusive – no associated production).
- Kinematic constraints ($p_{\Upsilon(4S)}$ and $|\vec{p}_B|^{\text{CM}}$ are known).
- Acceptable (i.e. not too low) $b\bar{b}$ cross-section ($\sigma(e^+e^- \to \Upsilon(4S) \to B\overline{B}) = 1.05$nb).
- π^0 and radiative decay capability through γ detection.

The cleanliness of events, especially the fact that no other particles are produced along with a $B\overline{B}$ pair, makes reconstruction and in particular tagging highly efficient compared to other types of experiments. The limiting factor for the B factories is the luminosity. The current luminosity level of PEP-II, allows for reasonable statistics ($\sim 5 \times 10^8$) given a few years of running.

Other types of experiments with B physics potential can be divided into two types: e^+e^- machine running at the Z^0 resonance (LEP experiments) and hadronic machines (HERA-B/Tevatron/LHC experiments). The LEP experiments had the advantage of a somewhat high $b\bar{b}$ cross-section (6 nb) with a much larger boost ($\beta\gamma \simeq 9$) in a fairly clean environment, but the luminosity of LEP did not suffice to provide adequate statistics.
At hadronic machines, the cross section for $b\bar{b}$ production is slightly higher and increases with energy, as does the boost, which is typically also much higher. The main differences are the much higher event rate and backgrounds. The relatively low signal to background ratio requires an extremely good background rejection and complicates the experiment. As the LHC has the highest luminosity, $b\bar{b}$ cross-section and boost, and a dedicated B experiment (LHCb), CERN will have the furthest reach into precision measurement in B physics.
All these other types of experiments have the advantage of B_s capabilities, which can only be attained at B factories by going to the much broader and therefore less advantageous $\Upsilon(5S)$, which lies above the $B_s\overline{B}_s$ threshold. This possibility has been investigated, and in principle it is possible, but the feable boost at the B factories will surely not allow for time-dependent analysis, as the B_s mixing frequency is too high (see Section 6.1.3).

7.5 Section summary and conclusions

In order to perform time-dependent CP measurements in the B system, an asymmetric high luminosity collider is needed. PEP-II at SLAC is exactly such a machine. From the flight distance between the two B decays (measurable due to the boost from the asymmetry), decay times differences below 1 ps can be determined. Since the decays of interest have low branching fractions, PEP-II is build to maximize the luminosity, which is the single most important limitation of B factories.
The impact of the PEP-II design on the *BABAR* detector is mainly through the required magnets close to the interaction region and the background. The latter limits the lifetime of components of the detector and decreases the reconstruction efficiencies, but it is not the leading limiting factor, even at high luminosity.

The main advantages that B factories offer over other B experiments is high ratio of signal to background, clean events, good kinematic constraints, and more efficient neutral particle reconstruction.

8 The *BABAR* detector

The best test-bed is the system itself. [O'Grady's Law[48]]

The study of CP violation in the B system is the primary goal of the *BABAR* experiment, which essentially requires good capabilities of:

- *Reconstruction* of B decays into exclusive final states,
- *Tagging* the other B meson in each event, and
- *Measurement* of the relative time between the two B decays.

The ability to fulfill the above requirements with a high efficiency for decay channels with branching ratios of order $\mathcal{O}(10^{-4})$ or below, requires:

Large acceptance and high reconstruction efficiency, as any lost track or clusters render the decay impossible to fully reconstruct. Furthermore, the tagging efficiency increases significantly with increased tracking efficiency.

Very good energy and momentum resolution both in angle and in magnitude, enabling efficient and clean reconstruction of shorter-lived mesons in the reconstruction chain. The momentum range covered should be 75 MeV to 4 GeV for tracks and 20 MeV to 4 GeV for clusters (photon).

Excellent vertex resolution both transverse and parallel to the beam, for time-dependent measurements and identification of D mesons.

Efficient particle identification of both leptons and hadrons, which is extremely important for both tagging, background rejection and separation of (final) states.

Flexible and redundant trigger, capable of separating out machine noise without loss of signal efficiency.

Detailed monitoring, automated calibration and online processing to realize the "factory" mode needed for obtaining very high volumes of data.

Tolerant detector components both with respect to background and radiation, as the environment will contain high levels of both.

It is the optimization of the above demands with respect to cost and efficiency that has dictated the design of the *BABAR* detector. After a general overview, the various subdetectors will be treated separately, and their function then described in the framework of the entire detector. The *BABAR* detector was tailored for B-physics, but essentially all other types of physics available can also be studied concurrently.

8.1 General detector design

From inside and out the *BABAR* detector consists of a:

- *Silicon Vertex Tracker (SVT)* for impact point and angle of tracks and partial PID,
- *Drift CHamber (DCH)* for absolute momentum of tracks and partial PID,
- *Detector of Internally Reflected Čerenkov light (DIRC)* for charged hadron identification,
- *ElectroMagnetic Calorimeter (EMC)* for photon detection and electron identification,
- *Superconducting coil*, providing a 1.5 T solenoidal magnetic field, and finally
- *Instrumented Flux Return (IFR)*, for muon identification and neutral hadron detection.

[48] The law has the collary: "Modules in the system that are known to work well, should be taken out!" . . . because when all else fails, one will at least be able to reinsert that module and be sure that it works!

The above description of functions is only approximate, as most subdetectors provide additional and often complementary information. Thus, to obtain optimal measurements, information from various subdetectors is combined. The *BABAR* detector is shown in longitudinal and transverse cut drawings in Fig. 8.1.

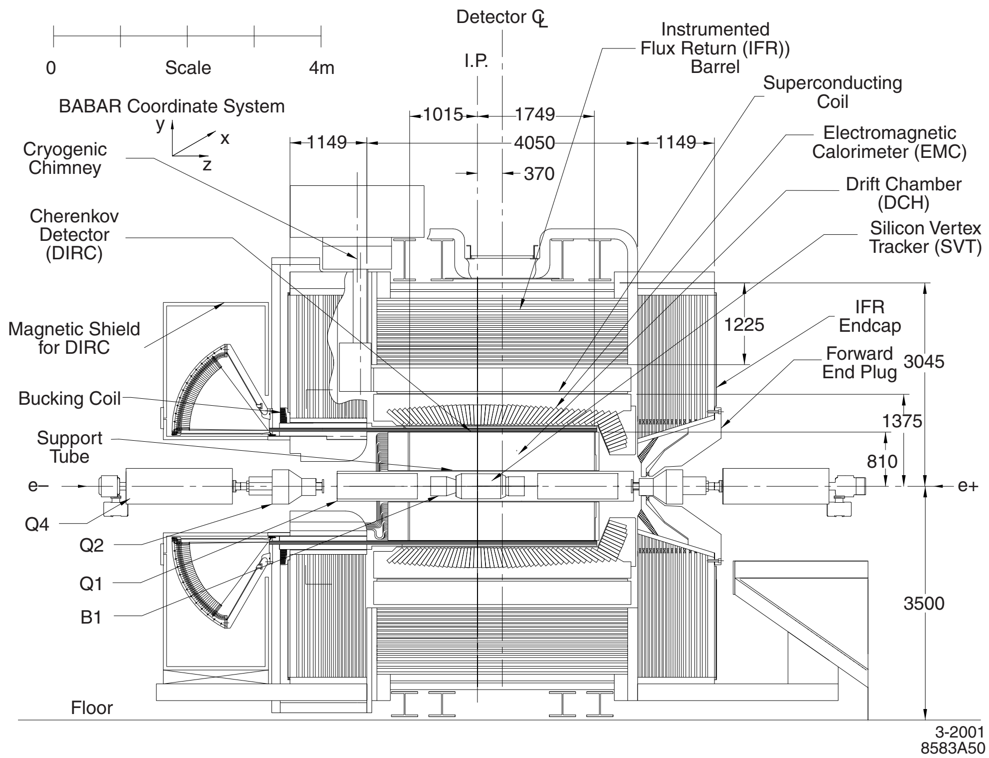
Like most other particle detectors, the *BABAR* detector consists of layers of different independent subdetector systems with complementary functions. Since the particle tracks are smeared by multiple scattering as they pass through material (e.g. the beam pipe and detector), the innermost detector parts should be the ones with the highest spatial precision, and they should consist of the least possible material[49] in order not to degrade the measurements of the outer subdetectors. The various layers of the *BABAR* detector can be found in Table 8.1.

System	θ_1	Radius/cm	ADC/bits	Segmentation	# Layers	Performance
	θ_2	RL/X_0	TDC/ns		Channels	
SVT	20.1°	3.2–14.4	4	50–100 $\mu m\ r\phi$	5	$\sigma_{d_0} = 55\mu m$
	−29.8°	5.1%*	-	100–200 $\mu m\ z$	150k	$\sigma_{z_0} = 65\mu m$
DCH	17.2°	23.6–80.9	8	6–8mm	40	$\sigma_\phi = 1$mrad
	−27.4°	2.9%**	2	drift distance	7104	$\sigma_{\tan\lambda} = 0.001$
						$\sigma_{p_t}/p_t = 0.47\%$
						$\sigma_{dE/dx} = 7.5\%$
DIRC	25.5°	81.0–89.0	-	35×17mm^2	1	$\sigma_{\theta_C} = 2.5$mrad
	−38.6°	17%	0.5	$(r\Delta\phi \times \Delta r)$	10752	per track
				144 bars		
EMC(C)	27.1°	90–136	17-18	47×47mm^2	1	$\sigma_E/E = 3.0\%$
	−39.2°	17	-	5760 crystals	2×5760	$\sigma_\phi = 3.9$mrad
EMC(F)	15.8°			820 crystals	1	$\sigma_\theta = 3.9$mrad
	−27.1°				2×820	
IFR(C)	47°	180–300	1	20–38mm	19+2	90% $\mu^\pm$ eff.
	−57°		0.5		22k+2k	6–8% $\pi^\pm$ mis-id
IFR(F)	20°			28–38mm	18	(loose selection,
	47°				14.5k	1.5-3.0 GeV)
IFR(B)	−57°			28–38mm	18	
	−26°				14.5k	

Table 8.1: Overview of subdetector coverage, thickness, segmentation, and performance. The notation (C), (F), and (B) refers to the central barrel, forward, and backward detector components, respectively. The polar angular coverage θ_1 (forward) and θ_2 (backward) refers to the laboratory frame. The listed radiation lengths of the SVT and the DCH include the beampipe* (1.1% X_0) and the support tube** (0.8% X_0), respectively. The magnet's radial extend between the EMC and the IFR is 1.4–1.7m. The performance is quoted for 1 GeV particles, unless otherwise specified. Table is mainly from [BABAR02a].

All subdetectors use a common electronics architecture, with Front End Electronics (FEE), mounted directly on the detector to minimize cabling. The FEE generally contains amplifier, digitizer, L1 latency buffer and event buffer in that order. After the level 1 trigger (see Section 8.9.1), the data is transfered via optical fibers to the data acquisition system. As essentially all components are very hard to repair/change without significant downtime, all parts of the detector were submitted to a variety of tests before installation.

[49] More precisely, the material should constitute the least amount of radiation lengths (RL), X_0, which is the mean distance over which a high-energy electron looses all but $1/e$ of its energy by bremsstrahlung, and an appropriate scale length for describing high-energy electromagnetic cascades [PDG02].



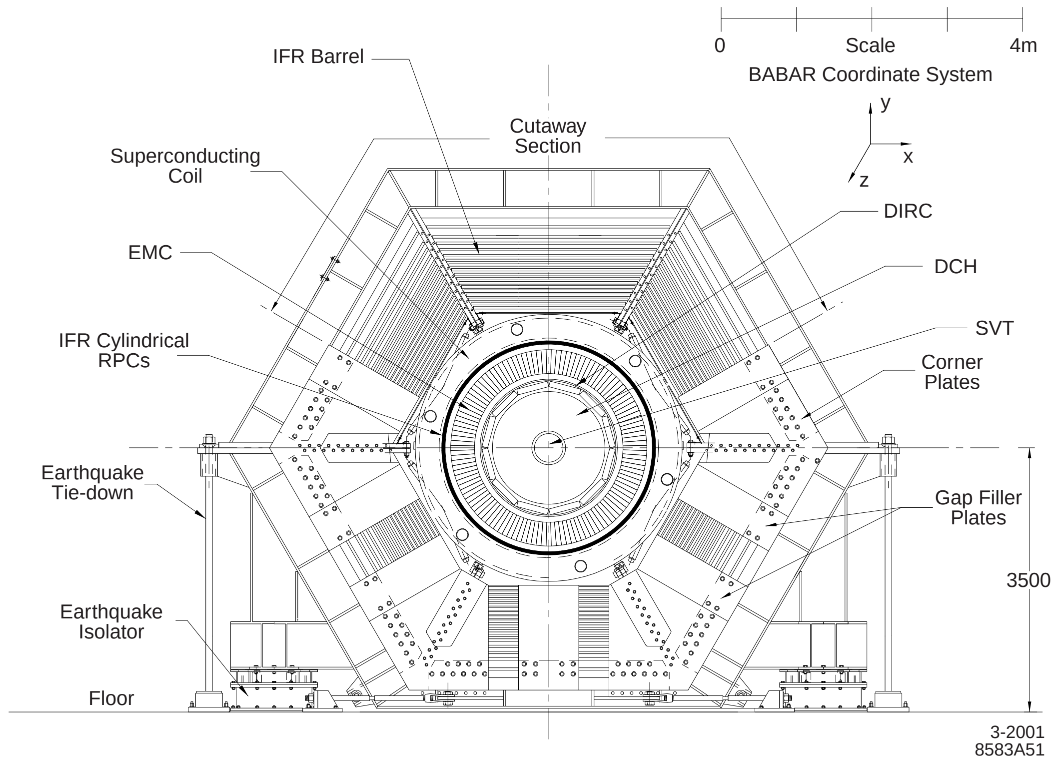


Figure 8.1: Layout of the *BABAR* detector, shown in cut-away drawings from the side (top figure) and from the end (bottom figure). From the inside and out it consists of Silicon Vertex Tracker (SVT), Drift CHamber (DCH), Detector of Internally Reflected Čerenkov light (DIRC), ElectroMagnetic Calorimeter (EMC), Superconducting coil, and Instrumented Flux Return (IFR) (see text).

As the single most expensive subdetector is the calorimeter (EMC), much effort was put into minimizing its volume without unjustified performance degradation of neither the calorimeter nor the tracking system inside of it, leading to a compact design. Due to the asymmetry of the collider, the *BABAR* detector has also been designed asymmetrical, reaching into the forward direction in order to maximize the acceptance. For the same reason, the IP is moved 0.37 m backwards (i.e. towards the HER), and supporting electronics, cryogenics etc. are lead out via the backward end when possible. In Fig. 8.2 is shown the so-called detector protractor, which is the correlation between polar angles in the lab and in the CM reference frame for massless particles.

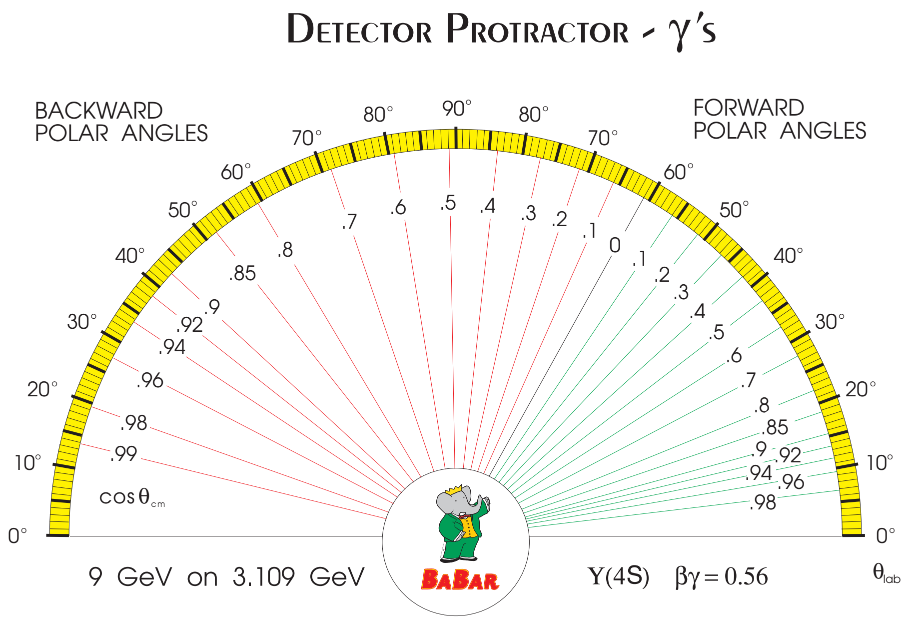


Figure 8.2: Detector protractor. The correlation between polar angles in the laboratory and the CM reference frame at $\beta\gamma = 0.56$ for massless particles. For example a photon emitted perpendicular to the z-axis in the CM frame, will be detected in the lab at about 60° in the *forward* direction.

8.2 The Silicon Vertex Tracker (SVT)

When charged particles pass through semiconducting material they create electron/hole pairs along their path, which are collected by an electric field and thus detected. The energy required to create electron-hole pairs is an order of magnitude smaller than to ionize gas, which gives good spatial resolution (because of increased statistics), and since silicon detectors in addition require little space, they are often the innermost (vertex) subdetector.

8.2.1 SVT functions

The main task of the SVT is to measure the impact point and angle of charged tracks for precise determination of the two B mesons vertices in order to provide the crucial Δz information for time-dependent measurements. In order to efficiently fit time-dependent quantities[50], the Δz resolution should be better than half the average separation of 260 μm, corresponding to about 80 μm resolution for each vertex (in reality this resolution is dominated by the tagging side resolution, see Section 8.4). This precision can be reached, but since better resolution improves background rejection and facilitates the distinction between primary B and secondary

[50]What is meant by "efficiently fit" is that 90% of the statistical power is preserved in time-dependent fits.

D vertices, the aim of this subdetector is the best possible achievable resolution. The point resolution is limited by multiple scattering to $10-15\ \mu$m for the inner layers and $30-40\ \mu$m for the outer ones, thus impact parameter and angular information from the innermost points, provided by the SVT, are the best.

Together with the DCH (see Section 8.3) the SVT constitutes the tracking system, with the SVT dominating the precision for the impact point and angle. The tracking allows exclusive reconstruction of B and D mesons, and is needed by the subsequent subdetectors, most importantly the DIRC, which entirely relies on the track angle (in z), at high momentum best measured by the SVT. Due to the strong magnetic field, charged particles with transverse momentum less than 120 MeV will not reach the DCH, and in this case the SVT provides the only tracking information. This is of particular importance for slow pions from the decay of the $D^{*\pm} \rightarrow D^0\pi^{\pm}_{\rm soft}$, which plays a central role in many contexts (including the one of this thesis). In addition the SVT must be efficient for particles that decay within the active volume, such as K^0_S, and it also contributes to dE/dx measurements for PID.

To perform the above functions, the SVT should have small segmentation and cover much of the solid angle. As the innermost detector, the SVT should cause a minimum of multiple scattering and be radiation hard, in order to stay reliable in the PEP-II environment for many years.

8.2.2 SVT concept and design

The SVT consists of five layers of 300 μm thick double-sided silicon strip detectors; the three inner layers are in standard concentric cylinder design while the two outer layers are in a novel “arch” design in which the barrel ends are folded inwards toward the beam pipe, see Fig. 8.3.

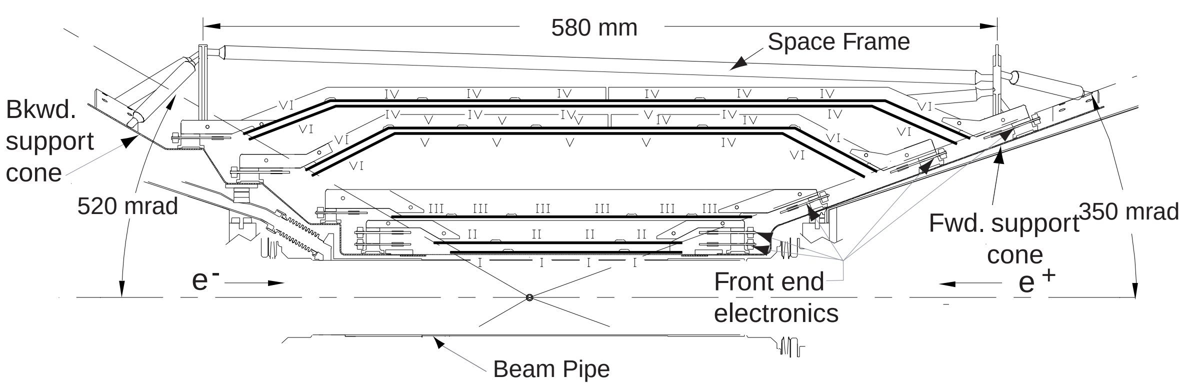


Figure 8.3: The Silicon Vertex Tracker (SVT) shown in the plane of the beam, including principle dimensions and associated structures.

This design maximizes the polar acceptance ($17.2^\circ < \theta_{lab} < 150.2^\circ$) while minimizing the track incidence angles and the amount of material and thereby the multiple scattering. Further coverage is inhibited by beam optics close to the IP. The modules of the inner layers are tilted by 5° and mutually overlap each other, which is also very useful for alignment purposes. The modules on the outer layers are divided into type “a” and type “b” with slightly different radii in order to allow for overlap between the modules, (see figure next to Table 8.2).

The three inner layers perform the impact parameter measurements, and are therefore placed as close as possible to a thin water cooled beryllium beam pipe. The two outer layers are used for low p_t tracking, pattern recognition and linking the tracks to the DCH.

Parameter/Layer	1	2	3	4a/b	5a/b
Radius (mm)	32	40	54	124/127	140/144
Readout channels	18432	24567	30720	32768	36864
Modules/Layer	6	6	6	8	9
Wafers/Module	4	4	6	7	8
Resolution (μm) z	10	10	10	10-12	10-12
Resolution (μm) ϕ	12	12	12	25	25

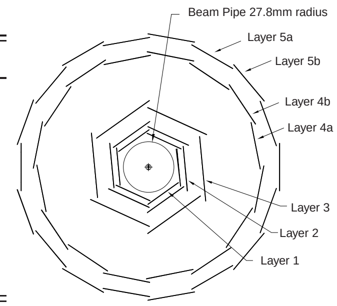


Table 8.2: Dimensions and parameters for the five SVT layers. The resolution quoted is the intrinsic one at normal incidence assuming signal-to-noise of 20 : 1, but multiple scattering makes the actual one higher. The outer layers have a radius at the end of 91mm and 114mm, respectively. The figure shows the SVT in the transverse plane of the beam, where the placing of the five layers can be seen.

On each module silicon strips are placed both parallel (z-strips, on the inner side) and perpendicular (ϕ-strips, on the outer side) to the beam line, giving both a z (polar) and a ϕ (azimuthal) measurement. The SVT covers a total area of 0.94 m^2, and has about 150,000 readout channels. The readout electronics is placed outside the acceptance region to minimize material in the active detector volume, and the modules are supported on ribs placed on the end cones. The material of the SVT constitutes about 4% of a radiation length at normal incidence. The specific dimensions and parameters for the SVT can be found in Table 8.2.

To realize the great precision of the SVT, its internal and relative position must be determined accurately, which is done routinely with particle tracks. Changes in position are mostly due to temperature and humidity changes, which are controlled with water cooling and dry air, respectively, and closely monitored, as they can damage the FEE.
Being the detector closest to the beam, the radiation around the SVT is closely monitored by 12 silicon photodiodes, especially during injection, which the SVT may abort, if the *integrated* dose gets unacceptably high. These diodes are also used by PEP-II for beam tuning, and are to be updated by diamond detectors, as these are more reliable, accurate and tolerant.

8.2.3 SVT operation and performance

During the first year of running, the performance of the SVT has reached all the major design goals. The average hit reconstruction efficiency is 97%, excluding (9 out of 208) defective readout sections but including all other defects. The resolution of the impact parameter is less than 50 μm and typically much lower, and the hit resolution for incidence angles of 90° reach 13 μm, increasing with dip angle as expected. The performance, which can be seen in Fig. 8.4, are in very good agreement with expectations from Monte Carlo studies. As the tracking is done in combination with the DCH, the tracking performance is described in Section 8.4.

In addition, the SVT provides dE/dx measurements for PID with a resolution of 14%, which adds to the PID capabilities, especially for tracks which do not reach the DIRC (or DCH). Given that the average energy loss is solely a function of β, as per the Bethe-Bloch formula [BH34, BN37], the mass and thus identity of charged particles can be established, when the momentum is known.

Thus the inner and outer layers satisfactorily fulfil their functions of performing angle and impact parameter measurements, and pattern recognition and low $p_\perp$ tracking, respectively. The radiation dose is within the planned budget, and no modules have failed due to radiation damage.

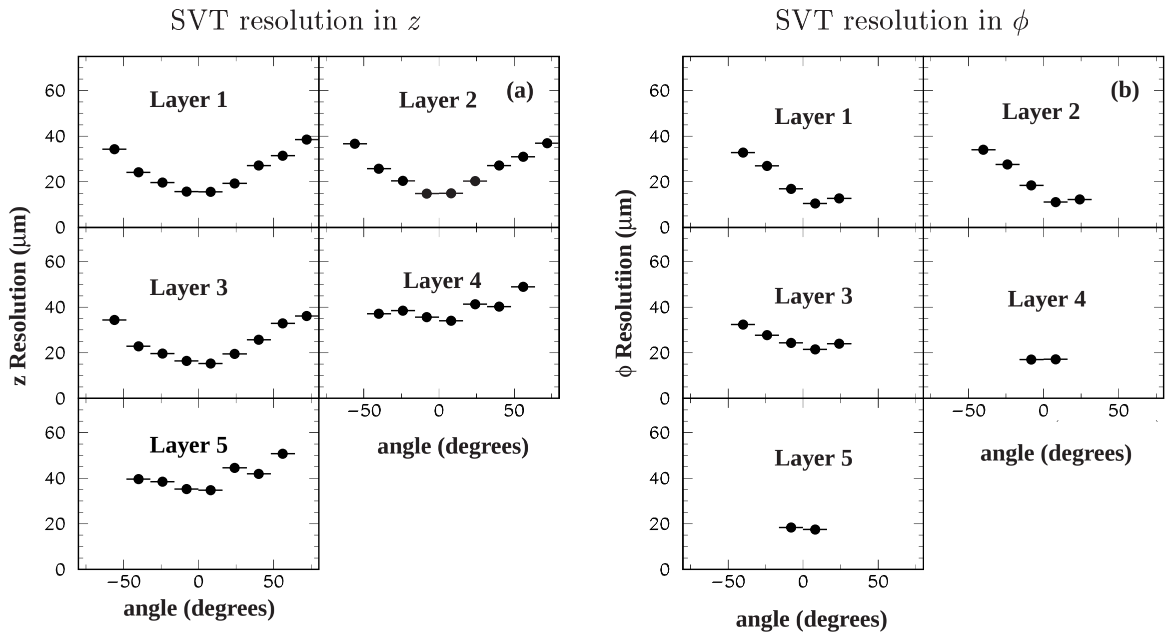


Figure 8.4: The SVT performance. Resolution in (a) z and (b) ϕ for each SVT layer, as a function of dip angle, θ. The uncertainty increases when departing from normal incident, as expected.

8.3 The Drift Chamber (DCH)

Drift chambers use the ionization of gas (or liquid) to reconstruct tracks of charged particles. Anode wires are surrounded by cathode wires in a chamber filled with a suitable gas such that ions and electrons will drift. The distance from the various wires can be deduced from the drift time, and the path of the ionizing particle established.

8.3.1 DCH functions

The DCH plays a central part in the *BABAR* detector, as it serves several important functions required for exclusive and clean B and D reconstruction. Primarily, the DCH is the main tracking device, and must therefore have very good spatial resolution and large angular coverage, especially in the forward direction (due to the boost) in order to provide good momentum resolution and high reconstruction efficiency for particles with transverse momentum down to 120 MeV. The DCH defines the global coordinate system, and like the SVT, it should remain efficient for particles decaying in the active volume (such as K_S^0). To avoid degrading the performance of both the DCH itself, but also outer detectors, the amount of material and thus multiple scattering should be minimized.

In addition to tracking, the DCH should provide PID by measurements of dE/dx from ionization energy loss with 7% precision (assuming passage of 40 DCH layers). This reliably discriminates between pions and kaons up to momenta around 700 MeV, after which PID is mainly provided by the DIRC (see Section 8.5).
Finally, the DCH is expected to provide the charged track Level 1 (L1) trigger, and must therefore be able to trigger with a latency of no more than $(9.5 \pm 0.5)\mu$s, constrained by the SVT data buffering. The DCH should be able to perform the above functions in a 1.5 T magnetic field, and with the large beam-generated backgrounds of $\sim$ 5kHz/cell expected to pass the SVT.

8.3.2 DCH concept and design

In order to perform the above functions, a small-cell, low-mass drift chamber design was chosen (see Fig. 8.5). The cylindrical DCH has a small diameter (inner radius of 23.6cm and outer radius of 80.9cm) but is very long (2.76m).

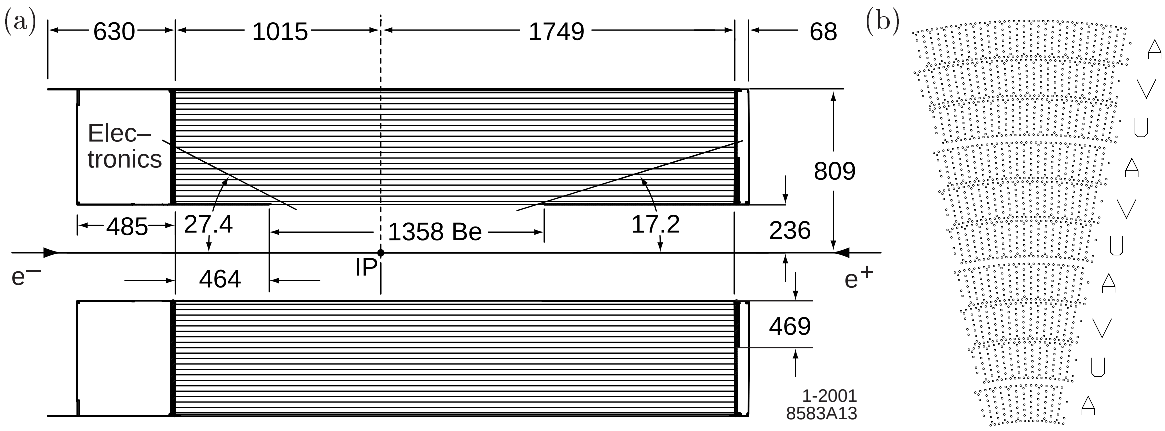


Figure 8.5: The Drift CHamber (DCH). (a) Longitudinal section of the DCH with principal dimensions. The offset from the IP optimizes the acceptance. (b) Transverse section of DCH "slice", showing the alternating Axial (A) and stereo (U,V) superlayers.

The 28768 DCH wires yield 7104 drift cells of typical dimentions 11.9(radial)$\times$19.0(azimuthal)mm^2, which each consists of a sense wire at $\sim$ 1960V surrounded by six grounded field wires, neighboring cells sharing these (see Fig. 8.6a). The cells are arranged into 40 layers, providing up to 40 spatial and ionization measurements. The layers are grouped by four into 10 superlayers, within which each layer have the same number of cells and wire orientation[51]. A forward and backward extension of 1749mm and 1015mm, respectively, ensures that particles emitted at $17.2° < \theta < 152.6°$ will traverse at least half of the layers. The volume is filled with a 80:20 mixture of helium:isobutane, which along with the choice of low-mass aluminum strings limits the multiple scattering inside the DCH to a minimum (0.2% X_0). This mixture also have the advantages of fairly short drift times and good spatial and dE/dx resolution.

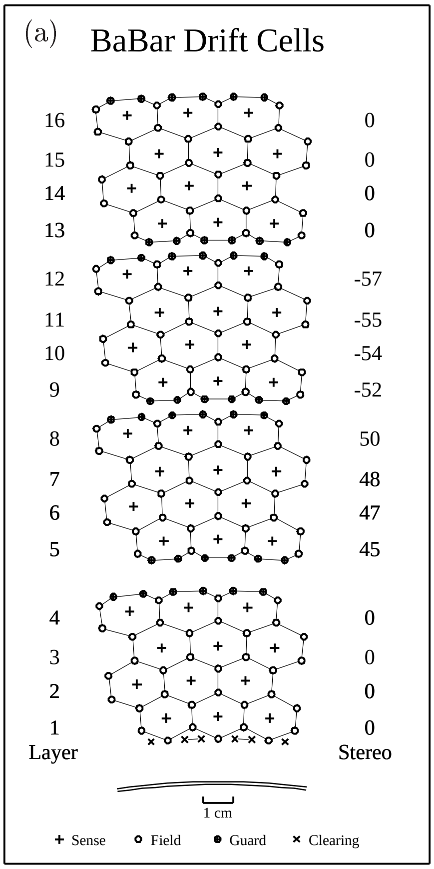
When charged particles pass through the DCH gas, electrons from the ionization drift towards the anode sense wire, causing an avalanche of secondary electrons (5×10^4), resulting in a clear signal. From measuring the time of this signal, the distance of closest approach (doca) is known from the almost circular isochores (see Fig. 8.6b), leaving only a so-called left-right ambiguity, which is resolved by shifting each layer by half a cell. The z-coordinate of the track points is obtained from slightly rotated superlayers (stereo (U,V) layers as opposed to axial (A) layers, see Fig. 8.5b and 8.6a), which alternates. With these stereo-angles going from 45 to 76mrad, the obtained resolution is $\sim 125\mu\mathrm{m}/\sin(50\mathrm{mrad}) = 2\mathrm{mm}$. This is the reason why the SVT dominates the z-component of the momentum, required by the DIRC (see Section 8.5.3).

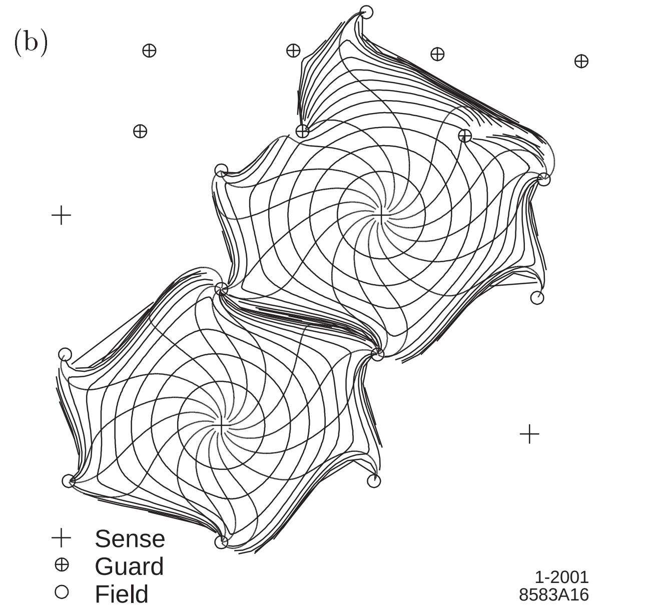
To sweep ions from the fieldless regions between layers and collect charge caused by photon conversions, guarding and field wires at $\sim$340V and $\sim$825V, respectively, are placed at appropriate edges.

To further minimize the amount of material in the forward direction, the outer part ($r >$ 46.9cm) of the forward endplate was made thinner (12 mm) than the rear endplate (24 mm), where also all read-out electronics were mounted. The design and choice of dimensions is a delicate and complicated interplay between wire tensions, material and elasticity along

[51] The superlayer design is due to the L1 trigger requirement of fast reduction of input via segment finding.

with endplate deflection and electromagnetic and gravitational forces. The inner and outer cylindrical walls of the DCH, which carry 40% and 60% of the wire load, respectively, serve to contain the ionization gas and shield the DCH from the RF field from the beam. The total thickness of the DCH at normal incidence is 1.08% X_0.





Type	Material	Diameter (μm)	Voltage (V)	Tension (g)
Sense	W-Re	20	1960	30
Field	Al	120	0	155
Guard	Al	80	340	74
Clearing	Al	120	825	155

Figure 8.6: DCH cell layout, cell isochores, and wire specifications. (a) Inner cell layout, including sense, field, guard and clearing wires. At the bottom of the figure is the 1mm inner DCH beryllium wall. (b) DCH cell isochores around the sense wire. Table: DCH wire specification (all wires gold-plated). Table from [BABAR02a].

Increasing the voltage further improves the resolution, but also induces background and shortens the lifetime of the DCH, and so the voltage setting is a trade-off between the two. The track finding and fitting is performed with the Kalman filter algorithm [Kal60, Bil84], which includes the detailed mapping of magnetic field and distribution of material.

8.3.3 DCH operation and performance

The DCH has from the start of operations performed close to design expectations, and has proven very stable ever since. With the exception of a small number of wires, which were damaged during High Voltage (HV) commissioning and initial running, all cells are fully operational. Due to this unfortunate accident, the voltage was for a period lowered to 1900V and 1930V, but it has now been restored at its design value of 1960V.

From an ensemble of charged tracks, the single cell resolution in the xy-plane, which depends on the distance from the sense wire, was determined to be 125μm on average (see Fig. 8.7a), thus below the 140μm design resolution requirement.

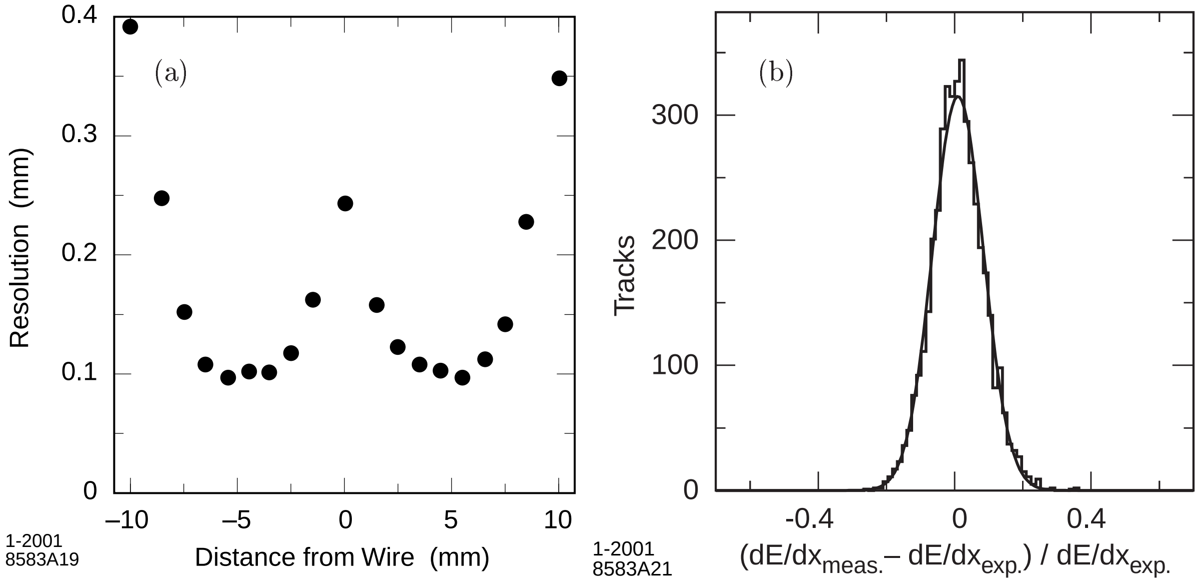


Figure 8.7: DCH performance. (a) The DCH spatial resolution as a function of drift distance (averaged over all cells in layer 18). (b) The dE/dx capabilities of the DCH for Bhabha events.

The dE/dx measurement provided by the DCH is a truncated mean of the 80% lowest individual measurements, which has a resolution 7.5% for Bhabha events (see Fig. 8.7b), limited by the number of samples and the Landau fluctuations. Through further corrections the resolution is expected to reach the design value of 7.0%.

When extrapolating the backgrounds expected with increasing luminosity, the DCH readout is the first limitation one reaches, thus an upgrade of the electronics is needed (and under construction). At increased luminosities, also the HV will have problems. Fortunately, the wire aging does not seem to be significant yet, so a replacement of the actual drift chamber does not seem necessary.

8.4 Tracking performance

The tracking performance, which reflects the combined capabilities of the SVT and the DCH, has been very close to design expectations from the beginning of data taking, and is well reproduced by MC simulations.

The overall tracking efficiency is determined from multi-hadron events, where it is computed as the fraction of tracks reconstructed in the SVT, which are also reconstructed by the DCH. At 1960V, the DCH efficiency is $98 \pm 1\%$ for tracks with transverse momentum greater than 200 MeV and polar angle > 500 mrad, where fake track correction dominates the error. At 1900V the efficiency is reduced to about 94%, which is at the verge of the acceptable. The tracking efficiency can be seen in Fig. 8.8I as a function of momentum and polar angle. Using $\overline{B} \to D^{*+}X$ events with $D^{*+} \to D^0(K^-\pi^+)\pi^+_{\rm soft}$, the tracking efficiency at low transverse momentum was studied. The efficiency remains above 80% down to 70 MeV, after which it drops quickly, in agreement with simulation (see Fig. 8.8II).

The rms width obtained on Δz for a sample where one B meson is fully reconstructed while the other is tagged is 190μm (see Fig. 8.8III), dominated by the tagging side (70μm for the reconstructed side), in accordance with expectations.

The overall tracking resolution is measured by comparing the two half tracks from cosmic rays above 3 GeV, that pass close to the IP. Tracks are parametrized in terms of five variables $(d_0, z_0, \phi_0, \tan\lambda, \omega)$, (transverse and longitudinal point of closest approach (POCA), azimuthal

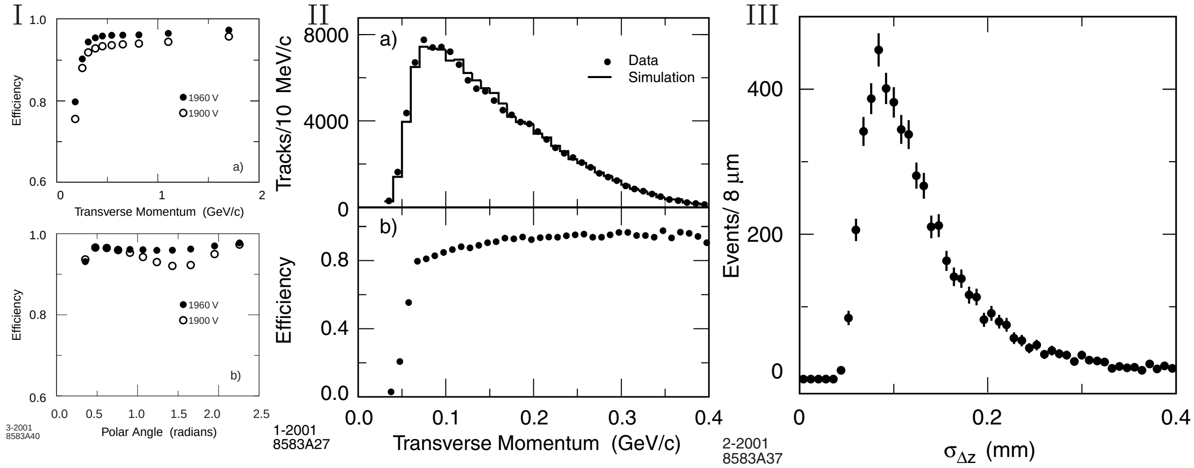


Figure 8.8: Tracking performance of SVT and DCH combined. (I) The tracking efficiency as a function of (a) momentum and (b) polar angle. (II) The low momentum (a) data and MC efficiency correspondance and (b) tracking efficiency. (III) Resolution in Δz for fully reconstructed B decays.

and dip angle and the curvature $\omega = 1/p_\perp$), and their error matrix. While the track POCA (d_0 and z_0) and initial direction are dominated by the SVT, the momentum is primarily measured by the DCH. The average errors are near-Gaussian:

$$
\begin{aligned}
\sigma_{d_0} &= 23\mu\mathrm{m} & \sigma_{z_0} &= 29\mu\mathrm{m} \\
\sigma_{\phi_0} &= 0.43\mathrm{mrad} & \sigma_{\tan\lambda} &= 0.53\times 10^{-3} \\
\sigma_{p_\perp}/p_\perp &= (0.13\pm 0.01)\% \times p_\perp(\,\mathrm{GeV}) + (0.45\pm 0.03)\%
\end{aligned}
\tag{8.1}
$$

Reconstructing $J/\psi \to \mu^+\mu^-$ decays, the invariant mass resolution is found to be $11.4 \pm 0.3\,\mathrm{MeV}$ at 1960V ($13.0\pm 0.3\,\mathrm{MeV}$ at 1900V). The peak falls 0.05% below the expected J/ψ mass, which is believed to be due to residual inaccuracies in alignment and mapping of the magnetic field.

The dE/dx measurement from the DCH is combined with that of the SVT and the information from the DIRC (see Section 8.5) in a maximum likelihood fit, which gives a combined and close to optimal PID.

8.5 The Detector of Internally Reflected Čerenkov light (DIRC)

DIRC is fine^{TM}. [Common saying at the daily *BABAR* operations meeting]

When particles exceed the speed of light in a medium ($\beta > 1/n$), Čerenkov light is emitted in a cone, analogue to the Mach cone created when traveling faster than the speed of sound. From Huygens principle on can calculate the light intensity at time t_0 from a particle passing through the origin at time $t = 0$ with velocity $v = \beta c$ in z-direction by summing the amplitudes:

$$A = \int_{-\infty}^{t_0} \frac{e^{iw(\frac{n}{c}\sqrt{(\beta ct)^2+r^2}-t)}}{\sqrt{(\beta ct)^2 + r^2}}\, dt, \tag{8.2}$$

where r is the transverse direction and n is the refraction index of the medium. This integral vanishes, unless

$$\frac{n}{c}\sqrt{(\beta ct)^2 + r^2} = t \quad \Rightarrow \quad \cos\theta_c = \frac{1}{\beta n}. \tag{8.3}$$

From the index of refraction and measuring the angle of the emitted Čerenkov light, one can determine the velocity β of a particle. Given the track momentum, the mass and thus the identity of the particle can be determined.

8.5.1 DIRC functions

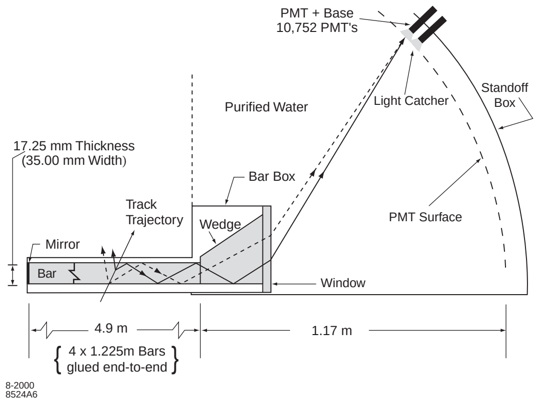
The DIRC subdetector is devoted to particle identification (PID), primarily separating charged kaons and pions at high momentum, but also distinguishing between other stable charged particles. The charged kaon identification is very important, both for tagging purposes (see Section 4.4.3) and for CP studies, where separation between final states (e.g. $B^0 \to \pi^+\pi^-$ and $B^0 \to K^+\pi^-$) is crucial. Thus, the DIRC must be able to separate kaons and pions up to the kinematic limit of 4.2 GeV at large angles in the laboratory frame, since the dE/dx measurement only provides reliable separation between kaons and pions up to 700 MeV. The difference in Čerenkov angle between kaons and pions at the highest momentum is, $\theta_c^\pi - \theta_c^K = 6.5$mrad. Also, the DIRC should provide fast signal response and tolerate high backgrounds, not only from the IR, but also from low energy photons due to off energy electrons (HER), showering in the line components.

8.5.2 DIRC concept and design

The DIRC is a ring-imaging Čerenkov detector, based on the principle that the Čerenkov angle in the ring image is preserved upon reflection from a flat surface (see Fig. 8.9). When charged particles pass through a quartz bar, Čerenkov photons are emitted and transported through total internal reflection to the back end, where their angle relative to the track direction is measured. With the bars serving as both radiators and light guides, much space is preserved, as the radial size used by the DIRC is 80mm in total[52], thereby reducing the size and thus cost of the calorimeter considerably.

There are 144 quartz bars in the DIRC, each 17.25mm thick (Δr), 35.00mm wide ($r\Delta\phi$), and 4.9m long (Δz). Each bar is made up of four 1.225m pieces glued end-to-end (by epotek), as this is the longest possible length attainable for the quality required. The bars are separated by $\sim 150\mu$m air gaps for optical isolation, and grouped in 12 hermetically sealed so-called bar boxes, arranged around the DCH forming a dodecagonal cylinder with radius ~ 85cm around the beam. Mirrors are placed at the front end of each bar, to avoid instrumentation in both ends.

[52] For comparison, the Čerenkov detector used at DELPHI [DELPHI91] used around 70cm.



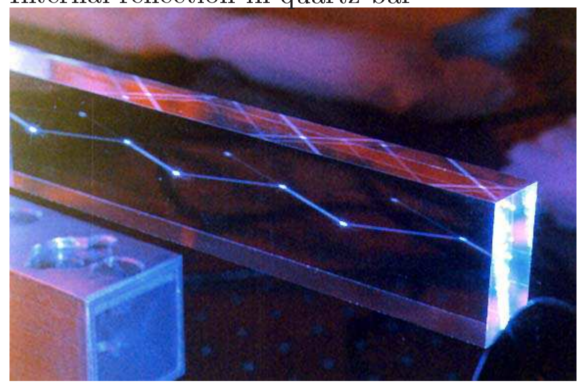


Figure 8.9: Sketch of the working principle behind the DIRC subdetector. Čerenkov photons from charged tracks are through internal reflection transported to one end of the quartz bars, and then projected onto a surface of PMTs. From the direction of the track and the PTM position, the Čerenkov angle and in turn the velocity and particle identity can be established.

This geometry covers angles down to 25.5° (39.6°) in the forward (backward) direction in the lab frame. The material constitutes 17% of a radiation length at normal incidence, the choice of thickness being a trade-off between number of Čerenkov photons and amount of material placed in front of the EMC.

At the back end of each bar a 91mm long fused silica wedge reflects photons at large angles, which reduces the required detection surface and also recovers photons that would otherwise be lost due to internal reflection at the silica/water interface. From the wedges, the photons emerge into the Standoff Box (SOB), which is a water-filled expansion region, instrumented by 10752 Photo-Multiplier Tubes (PMTs) at the approximately toroidal back surface (see Fig. 8.10). The PMTs have a diameter of 29mm, and are arranged in 12 sectors of 896 closely packed arrays, with hexagonal light catchers between, such that the effective area covered is around 90%. Between each sector is placed one *scaler* PMT, which integrates the number of hits over a second, for background monitoring. The PMTs are operated in groups of 16 at HV in the range 0.9–1.3 kV, and they are protected from the fringe field of the 1.5T magnet (see Section 8.7) by a steel shield supplemented by a bucking coil. The correctness of each PMT position was tested using cosmic ray data[53].
Purified water was chosen to fill the 6000 liter SOB, as it is cheap and matches both the index of refraction and chromaticity index of the silica, therefore minimizing internal reflection and dispersion, respectively, at the silica/water interface.

The fused synthetic silica used for the bars offers a variety of advantages, such as long attenuation length, low chromatic dispersion for the wavelengths in question and high index of refraction ($n = 1.473$) increasing the fraction of internally reflected photons. The silica is also resistant to ionizing radiation and the surface can be given an excellent optical finish.
The Čerenkov light cone projected onto the instrumented surface has the original Čerenkov angle modified only by the silica/water interface. The distance from the bars to the PMTs is $\sim$ 1.17m, and given the dimensions of the PMTs and the bars, the intrinsic single-photon geometrical angular resolution is $\sim$ 7mrad, slightly larger than the rms associated with the photon production and transmission dispersion, giving a total of about 10mrad. The overall photon efficiency, which is a function of wave length, is the product of each components efficiency. While most components are almost perfectly efficient, the dominant loss of efficiency is the quantum efficiency of the PMTs ($\sim$ 20%).

[53]Both single PMT and one HV group misconnections were found, as could be expected in 10000 channels.

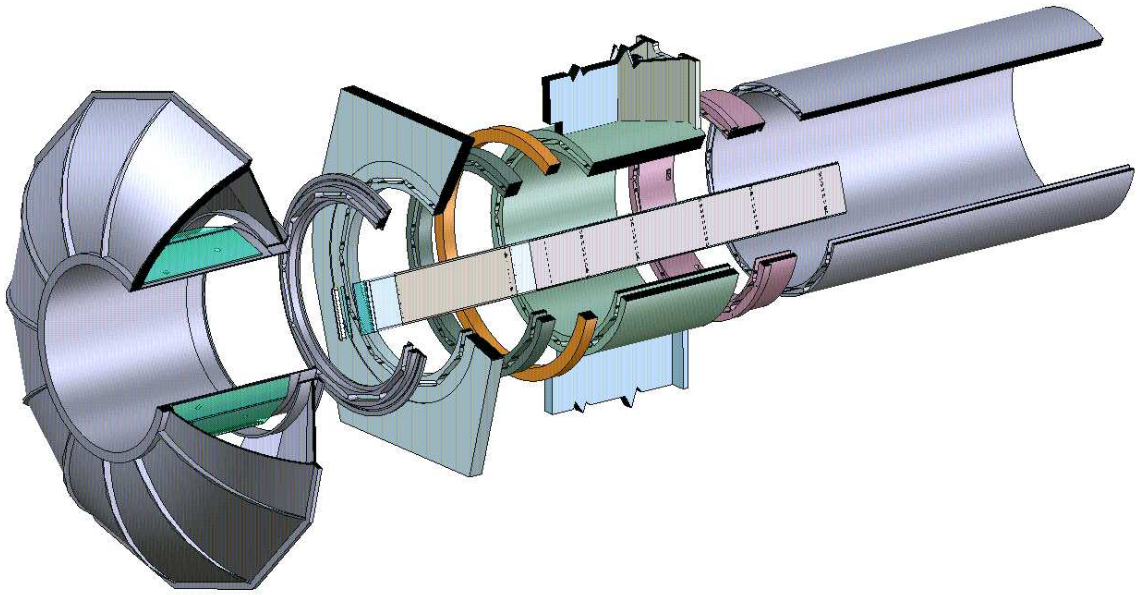


Figure 8.10: Layout of the DIRC. Shown is a quartz bar with its surrounding support structure and the Standoff Box (SOB) at the end.

To ensure high ($> 98\%$) transparency at wavelengths down to 300nm, the water in the SOB is kept ultra-pure and de-ionized by circulating it. As purified water is very reactive, stainless steel and polyvinylidene were used for the parts in contact with the water. The water passes a system consisting of six mechanical filters (1–10μm range), a reverse osmosis unit, a Teflon microtube de-gasser and a 254nm UV laser, which prevents bacteria growth, followed by five additional filters (0.2–1.0μm range and charcoal). The entire volume can be circulated in six hours, though ten hours is the default frequency. Monitors are installed to check for leaks and measure resistivity, pH-value, temperature and flow. To prevent condensation on the quartz bars and to detect leaks, nitrogen gas flows through the barboxes at 100–200cm^3/min, and is filtered through a molecular sieve and three mecanical filters (7μm, 0.5μm and 0.01 μm).

Calibration of the time delay for each channel is done with a light pulser system and from collision data. The two methods are in good agreement, and the values per channel are typically stable to 0.1ns over more than a year of daily calibration. The intrinsic time resolution of the PMTs is 1.5ns, and the electronics pipeline had TDCs, which were designed to operate with backgrounds up to 200 kHz/PMT without deadtime.

The signal Čerenkov photons arrive within $\sim$ 50ns of the 600ns trigger window. For each track and signal photon candidate the Čerenkov angles θ_c and ϕ_c are calculated up to a 16-fold ambiguity; top/bottom, left/right, back/forward and wedge/non-wedge reflection.
The difference between the expected and the actual time of arrival, Δt_γ, has a resolution of 1.7ns (see Fig. 8.14) dominated by the PMTs. Using the time information, the number of ambiguities are typically reduced to 2–3, the correct matching with tracks is substantially improved and the PEP-II induced background is reduced by a factor of 40 (see Fig. 8.11).
An unbinned maximum likelihood fit in the three dimensions, $\theta_c \times \Delta t_\gamma \times N_\gamma$, gives the likelihood value for each of the five stable particle types (e, μ, π, K, p), along with θ_c and the number of signal and background photons for each track.

8.5.3 DIRC operation and performance

Since its successful commissioning, the DIRC has operated close to design goals and been very stable and robust. After five years, 99% of the PMTs channels have nominal performance.

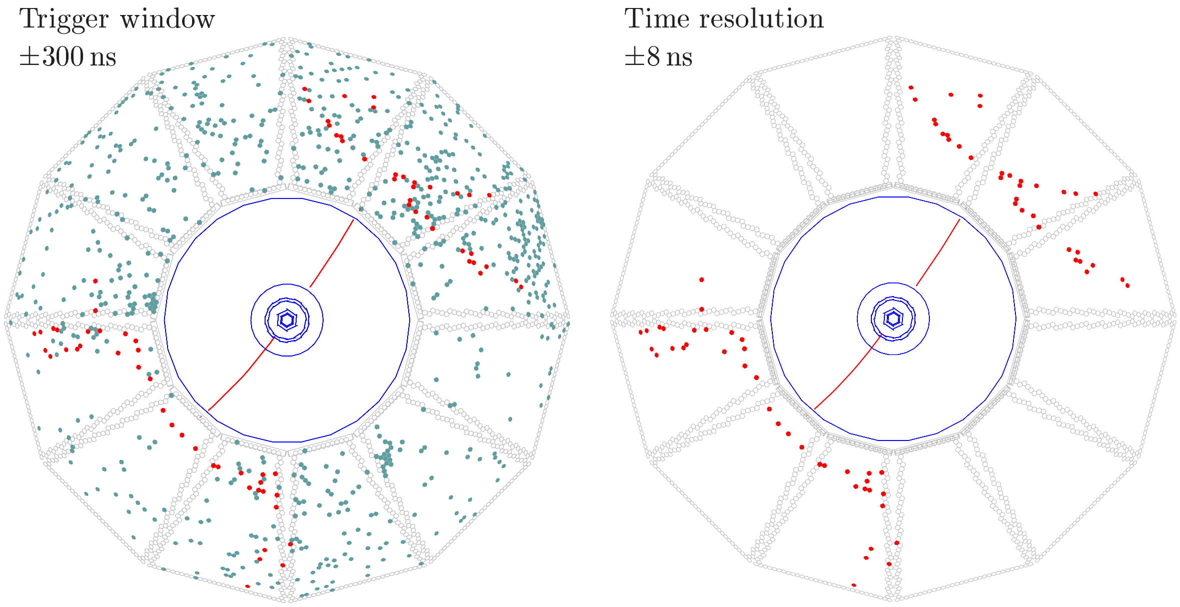


Figure 8.11: DIRC hits from a dimuon event before and after requiring the arrival time of the photons to fall within ±8ns of their expected arrival. The photons shown in the left figure are within the ±300 ns trigger window. The hits in the right figure (shown in red in both figures) are dominantly signal photons.

The primary failure mode is the loss of vacuum in the PMT, causing noise rates at the MHz level and several hundred kHz in the surrounding PMTs[54]. About five PMTs are affected per year. To deal with this problem, the HV group containing the failing PMT is switched off, until the single tube can be disconnected at an access, and the 15 others restored.
The deterioration of the PMT glass is in general not large (few μm/year), except for about 50 PMTs, which were manufactured with incorrect (zincless) glass that has become "frosty" after a while, without impairing their response.

It was realized at an early stage that the DIRC TDCs would not be able to cope with the rise in luminosity and its associated increase of background. Lead shielding around the beam pipe of the HER reduced the background and thus improved the situation (see Fig. 8.12) while, during the Summer shut-down of 2002, new TDCs which could encompass tenfold rates (2MHz/PMT) were installed.

The average number of detected Čerenkov photons for a track of $\beta = 1$ at normal incidence is $\langle N_\gamma \rangle = 28$, increasing by a factor of more than two for a track towards the edge (see Fig. 8.13a).
To monitor the overall efficiency of the DIRC, the number of Čerenkov photons per track is measured with a clean high-statistics sample of di-muon events as a function of time. The average photon loss is about 1.9±0.2%/year (see Fig. 8.13b), but seemingly at changing rates. It is not correlated with barbox nor Čerenkov ring location, which would both be signs of a specific problematic loaction.
This loss, which surpasses the rate of occurance of malfunctioning PMTs ($\sim 0.3\%$/year), is not fully understood, as many explanations (water purity, epotek glue, back mirrors, tracking quality, etc.) have been eliminated. A possibility is beginning deterioration of the reflectors between the PMTs, but it remains unclear. However, the impact on the DIRC performance remains very small.

[54]Given the successive drop in noise level with distance from the center and the subsequent colorful plot in the DIRC display, these occurences have been nick-named "Christmas trees".

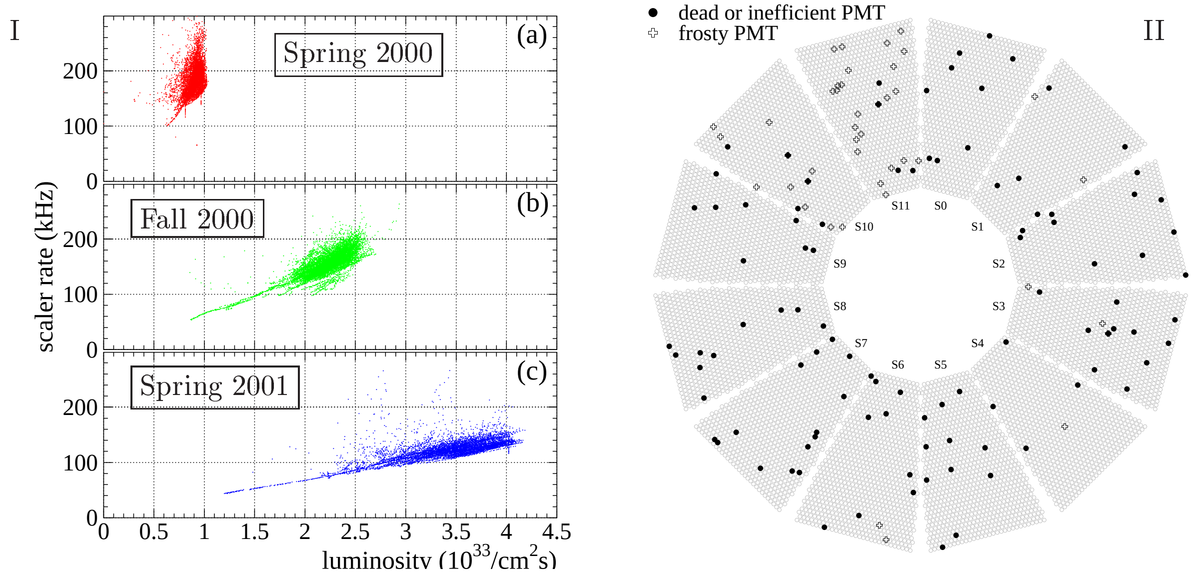


Figure 8.12: Effect of DIRC shielding and status of PMTs. (I) The scaler rates (kHz) for (a) Spring 2000, (b) Fall 2000 and (c) Spring 2001, after installing shielding during the shutdowns (Summer 2000 and Winter 2000–2001). The improvement is substantial, as the shielding has managed to decrease the scaler rates, while the luminosity was increased by a factor of four. (II) Distribution of dead/inefficient and frosty PMTs. No pattern is seen among the dead PMTs, and 99% of the tubes work nominally.

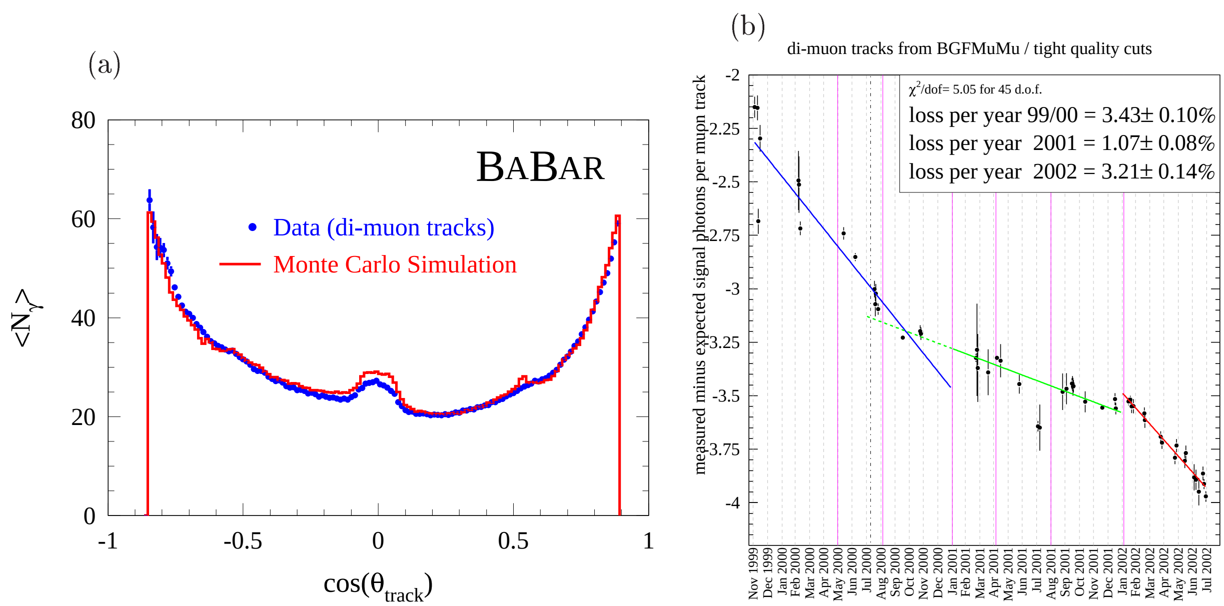


Figure 8.13: Number of photons per track N_γ. (a) N_γ as a function of polar angle. The correspondance between simulation and data is good ($\mathcal{O}(5\%)$). The enhanced number of photons at $\sim 90^\circ$ is due to the fact that around this angle all Čerenkov photons are internally reflected. (b) $\langle N_\gamma \rangle$ monitored 1999–2002, showing a loss of $1.9 \pm 0.2\%$ signal photons per year, but seemingly at changing rates. The purple lines indicate changes in the DIRC reconstruction software. Note that the errors shown are statistical only, and the fits are only indicative. At the current rates, the impact on the DIRC performance is very small.

The resolution on the difference between the measured and expected Čerenkov angle for single photons, $\Delta\theta_{c,\gamma}$, shown in Fig. 8.14, is 9.6mrad (limited by geometrical alignment), in very good agreement with the expected value. In the absence of correlations, the track resolution should then be:

$$\sigma^2_{\text{track}}(\theta_c) \quad = \quad \sigma^2_\gamma(\theta_c)/N_\gamma + \sigma^2_{\text{track}}. \tag{8.4}$$

Using di-muon events, the Čerenkov angle resolution for tracks is measured to be 2.5mrad (see Fig. 8.14), compared to the design goal of 2.2mrad.

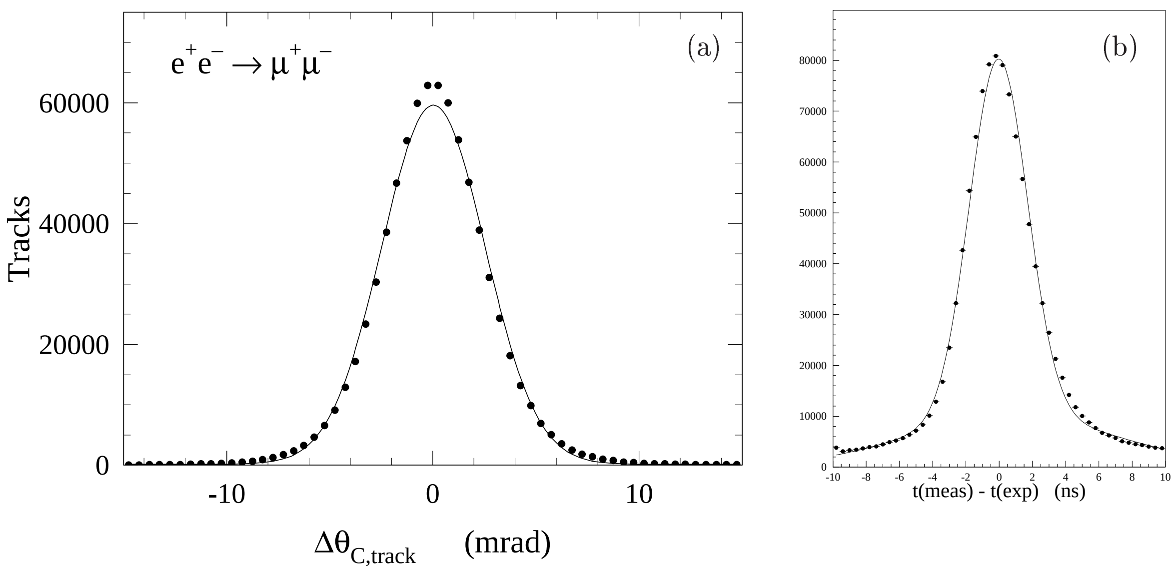


Figure 8.14: The DIRC performance in terms of resolution in (a) $\Delta\theta_c$ and (b) Δt_γ, measured with dimuon events.

From a clean sample of $D^{*+} \to D^0(\to K^+\pi^-)\pi^+$ events, the separation between kaons and pions can be determined over the whole kinematic range. In the absence of tails on the distributions, the separation is simply the difference of the mean Čerenkov angle divided by the track resolution, see Fig. 8.15. However, due to tails, the separation is not quite as significant, and is more accurately defined in terms of efficiency and misidentification probability as a function of momentum, see Fig. 8.15.

Finally, the very precise timing of the DIRC makes it a good source of monitoring device, and is also used for background monitoring by PEP-II. Furthermore, its great sensitivity to tracking makes it a good monitor of the tracking alignment[55].

The DIRC has been operating very reliably, had no major problems and has been most tolerant to high backgrounds. Both kink-finding and PID likelihood are sensitive to the fraction of charged kaons, that decay before reaching the DIRC (15-20%), and so the kaon efficiency is not severely affected. Alignment and code developement are expected to improve the resolution further, but the performance achieved is excellent and within the design requirements.

As the DIRC is very sensible to the tracking, and since tracking is one of the limiting parameters, it was thought of placing an additional tracking layer on the outer part of the DIRC, as this would greatly increase the track and thus the angular precision. However, the DIRC already works very well, and an additional layer of tracking is not needed.

[55] By monitoring the time resolution over time, slight changes in the tracking resolution can be detected. An SVT humidity problem in November 2001, which caused slight misalignment, was first seen by DIRC.

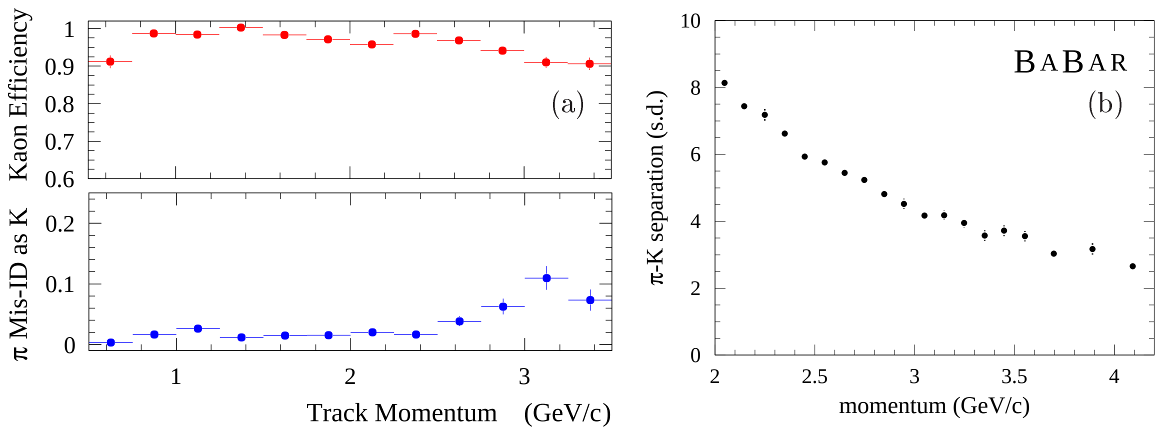


Figure 8.15: The DIRC PID performance measured with a pure sample of $D^{*+} \to D^0(\to K^+\pi^-)\pi^+$ events. (a) Kaon efficiency and pion mis-identification rate and (b) separation between kaons and pions as a function of momentum. The lacking correspondance between mis-identification rates (bottom left) and separation (right) is due to non-Gaussian tails in the Čerenkov angle distribution. Note that the scale on the y-axis on the two left figures are not the same.

8.6 The ElectroMagnetic Calorimeter (EMC)

The detection of photons and charged particles interacting electromagnetically, can be done with "doped" crystals. When electrons and photons pass through crystal, they loose energy by *bremsstrahlung* (electrons) and e^+e^- pair production (photons). Both these processes produce new electrons and photons, thereby creating a *shower* of (low energy) electrons and photons.
These excite the atoms in the crystal, emitting photons in the ultraviolet spectrum, which through a wavelength shifter (the "doped" impurities, which absorb light at one frequency and reemit it at another one) are turned into frequencies compatible with standard photo-diodes acceptance. From the size of the signal, the energy of the interacting particles can be infered.

8.6.1 EMC functions and requirements

The EMC is the main detector for neutral electromagnetically interacting particles, mainly photons. It should be able to measure electromagnetic showers over the energy range 20 MeV up to 9 GeV with excellent efficiency, and energy and angular resolution. Such capabilities enable the reconstruction of π^0 and η decays as well as electromagnetic and radiative processes, ranging from low energy π^0 mesons (from e.g. D^{*0} decays) to high energy $e^+e^- \to e^+e^-(\gamma)$ and $e^+e^- \to \gamma\gamma$ events used for calibration and luminosity determination.

The EMC should in addition be able to identify electrons, used for tagging, reconstruction of vector mesons such as J/ψ, semi-leptonic and rare decays of B and D mesons and τ leptons. The EMC is also required to provide quick event information, as it constitutes one of the two principal triggers. Furthermore, the EMC must be able to operate reliably in a radiation environment over the anticipated ten-year lifetime of the experiment and in the 1.5T magnetic field.
Contrary to the previous detectors, the EMC is not required to minimize material, as the particles detected by the outermost subdetector (the IFR, see Section 8.8) are only slightly affected by the EMC material.

8.6.2 EMC design

To achieve the above requirements, a hermetic, total-absorbtion calorimeter design was chosen. The EMC consists of finely segmented crystals, read out at the end by silicon photodiodes matching the scintillator spectrum[56]. Cesium iodide (CsI) doped with 0.1% thallium (Tl) was chosen for the crystals because of its high light yield and small Moliere radius, which enables high energy and angular resolution, respectively. In addition, its short radiation length reduces the depth required to contain the showers, and its large attenuation length allows the crystals to act as both radiators and transmitters. A total of 6580 crystals pointing towards the vertex are arranged in a barrel section (5760) and a forward conic endcap (820), while a backward endcap is omitted (out of cost issues), see Fig. 8.16. This results in an angular coverage from 15.8° to 141.8° (though effectively slightly less due to shower leakage in the edges), which corresponds to 90% of the CM solid angle.

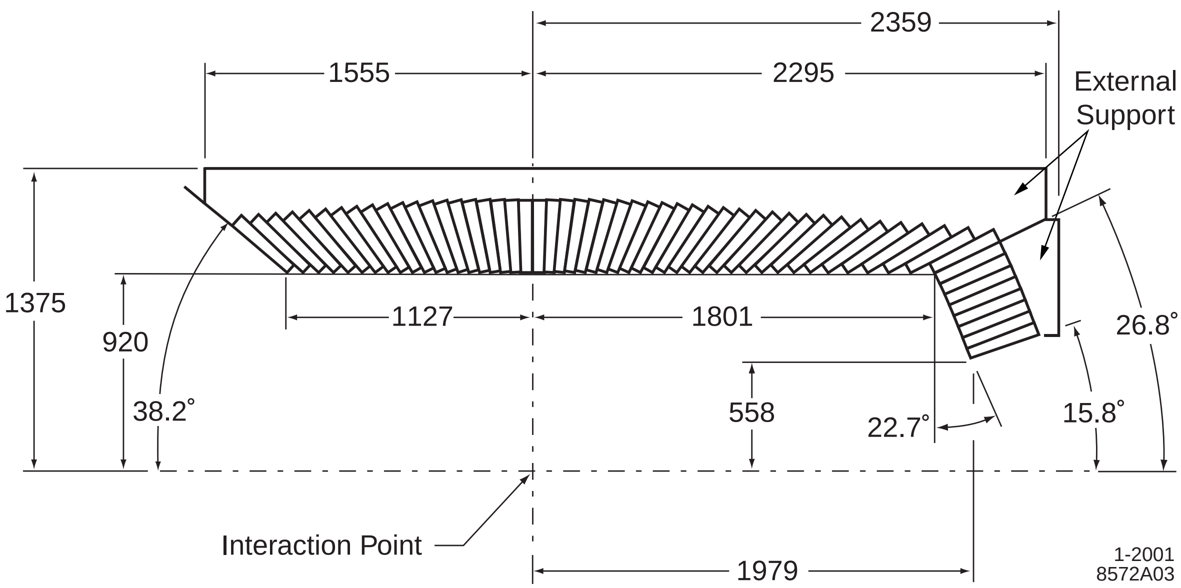


Figure 8.16: Layout of EMC crystals along with principle dimensions and angles. A backward endcap was not deemed necessary, considering its cost.

The dimensions of the crystals are typically 47 × 47 mm at the front face, 60 × 61 mm at the back face and 296–324 mm long, decreasing in transverse dimension, and increasing in length toward the forward direction, for optimizing the usage of a given number of channels. The length corresponds to 16.0–17.5 X_0, enough to contain photons and electron showers, while hadrons (usually) continue to the IFR.

The barrel and outer five endcap rings have less than 0.3–0.6 X_0 in front of them, increasing with absolute polar angle and dominated by the DIRC. More material is in front of the inner three end cap rings, whose main purpose is to contain the showers.

To ensure low transmittance on the sides, magnetic shielding and electric insulation, each crystal is wrapped in reflector and aluminum foil and covered with mylar, before being mounted in the described arrays. The crystals are read out at the rear end by two independent $2 \times 1\mathrm{cm}^2$ photodiodes. The photodiodes reside together with preamplifiers in a shielded housing also used for heat removal. To meet the required energy resolution, the noise from readout electronics should be below an equivalent of 250keV, which is achieved using digital filters.

The EMC is maintained at a low, constant and closely monitored temperature, as changes degrade the precision and can damage the joints.

[56]The design was in part chose because of recent satisfactory experience at CLEO [CLEO83].

To reach the desired energy resolution for the EMC, constant calibration of the energy scale is crucial, in order to monitor both short and long term variations. This is done using the following techniques:

Charge injection into the front end of the amplifiers, in order to monitor the precise response function (0.1% level) of the read-out electronics for each channel.

Light pulsar system injecting light via optical fibers into the rear end of each crystal for tracking short term changes to better than 0.15% (with daily light pulsar runs).

Radioactive source (O^{*16}) producing 6.13 MeV photons[57] producing 6.13 MeV photons for setting the energy scale precisely (0.35%), and monitoring long term variations.

Bhabha events for continuous calibration, to measure the small changes in light yield with integrated radiation (0.35%).

Parameter	Barrel	Endcap
No. Crystals ($17.6X_0$)	840	820
No. Crystals ($17.1X_0$)	840	0
No. Crystals ($16.6X_0$)	840	0
No. Crystals ($16.1X_0$)	3240	0
Total Volume (m^3)	5.2	0.7
No. Readout Channels	11760	1640

Table 8.3: Parameters for the EMC.

Property	Value
Radiation Length	1.85 cm
Density	4.53 g/cm^3
Molière Radius	3.8 cm
Light Yield	$\sim$ 50000 γ/ MeV
Peak Wavelength	565nm
Signal Decay Time	640ns/3.34μs

Table 8.4: Properties of Thallium-doped CsI.

Crystals with energy deposit above 1 MeV (15–20%) are read out, after which clusters and local maxima (bumps) are found by pattern recognition algorithms. In case of several bumps in a cluster, the energy is divided according to an iterative weight calculation. The position of a bump is calculated from a center-of-gravity method with logarithmic weights emphasising lower energy crystals, with an additional forward-backward correction. Bumps, which have no tracks leading to them, are assumed to be due to neutral particles.

To distinguish electromagnetic showers (γ and $e^{\pm}$) from showers of neutral hadrons (K_L^0 and n) which tend to extend over more crystals than the former, the shower shape is characterised by the lateral moment (LAT) [D$^+$85], defined as:

$$LAT \equiv \frac{\sum_{i=3}^{n} E_i r_i^2}{\sum_{i=3}^{n} E_i r_i^2 + (E_1 + E_2) R_0^2}, \tag{8.5}$$

where n is the number of crystals, R_0 is the mean distance between adjacent crystals, r_i is the crystal distance from the shower center and E_i is the energy of crystal i, the crystals being ordered by decreasing energy (i.e. E_1 is the largest crystal energy). Also the regularity of the shower is characterised by Zernike moments [SV97], as EM showers are most regular.

8.6.3 EMC operation and performance

The EMC has in general been operating well. During the second year of running, the cooling of the calorimeter was not entirely stable, which lead to series of minor problems. But having overcome this problem, no major difficulties have been encountered, and only the neutron generator used to produce the calibration source, O^{*16}, has needed replacement.

[57] The full chain of reaction is $F^{19} + n \to N^{16} + He$, $N^{16} \to O^{*16} + e$ and $O^{*16} \to O^{16} + \gamma$, where the initial neutrons are produced by a neutron generator placed next to the detector.

The loss of light from radiation damage is starting to show for the end cap, where a 5–10% degration in energy resolution has been observed. The rate of deterioration will most likely increase with the luminosity, but so far the behavior has been favorable, and the crystals, especially those in the barrel, are likely to last until 2006 without major replacements. The silicon photodiodes are also affected by radiation, and may have to be replaced at some point, but the issue is again not critical.

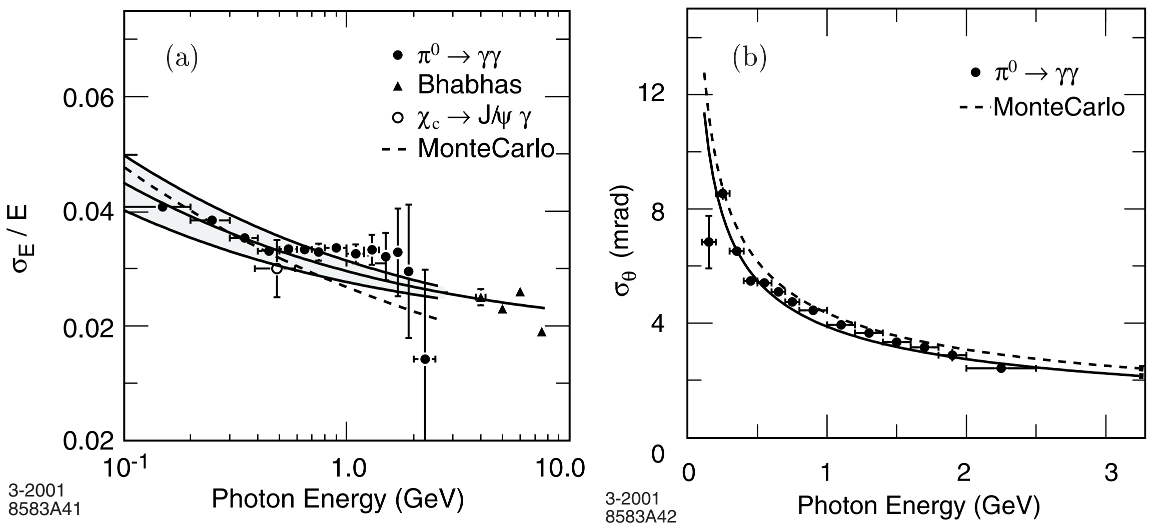


Figure 8.17: Energy and angular resolution of the EMC. (a) The energy resolution is measured using π^0 decays, *radiative* Bhabha events, and $\chi_c \to J/\psi\gamma$ events, while (b) the angular resolution is only determined from π^0 decays. The agreement with MC is good.

The energy resolution is determined by fitting data from the O^{*16} source, $\pi^0 \to \gamma\gamma$ decays, $\chi_{c1} \to J/\psi\gamma$ events and radiative Bhabha events (see Fig. 8.17a), while the angular resolution is measured with approximately equal energy π^0 and η decays to two photons (see Fig. 8.17b). The fit functions each consist of a constant and an energy-dependent term:

$$\sigma_E/E(\mathrm{GeV}) = \frac{2.30 \pm 0.02 \pm 0.30\,\%}{\sqrt[4]{E(\mathrm{GeV})}} \oplus 1.35 \pm 0.08 \pm 0.20\,\%, \tag{8.6}$$

$$\sigma_{\theta,\phi} = \frac{3.87 \pm 0.07\ \mathrm{mrad}}{\sqrt[2]{E(\mathrm{GeV})}} \oplus 0.00 \pm 0.04\ \mathrm{mrad}. \tag{8.7}$$

The obtained energy resolution is in agreement with recent MC studies (but worse than TDR expectations [BABAR], which is thought to be due to too simplistic simulations). The angular resolution is slightly better than the design requirements and simulation expectations.

The e/π separation, is mainly based on the energy-momentum ratio, and studied with $e^+e^- \to e^+e^-\gamma, e^+e^-e^+e^-$ events along with $K_S^0 \to \pi^+\pi^-$ and three-prong τ decays. In the momentum range $0.5 < p < 2.0\,\mathrm{GeV}$, a tight (very tight) selection has an average 94.8% (88.1%) electron efficiency with a pion misidentification probability of 0.3% (0.15%), shown in Fig. 8.18 (left) as a function of momentum and polar angle. The π^0 mass resolution is 6.9 MeV for photons with $E_\gamma > 30\,\mathrm{MeV}$ and $E_{\pi^0} > 300\,\mathrm{MeV}$ (see Fig. 8.18II), which agrees well with MC simulations, and is $\sim 10\%$ lower for isolated π^0 candidates in hadronic events[58].

[58]In terms of calorimeter, the Belle detector is performing slightly better than *BABAR*, this being attributed to the larger distance from the IP (1.2m compared to 1.0m at *BABAR*) and the lower machine backgrounds.

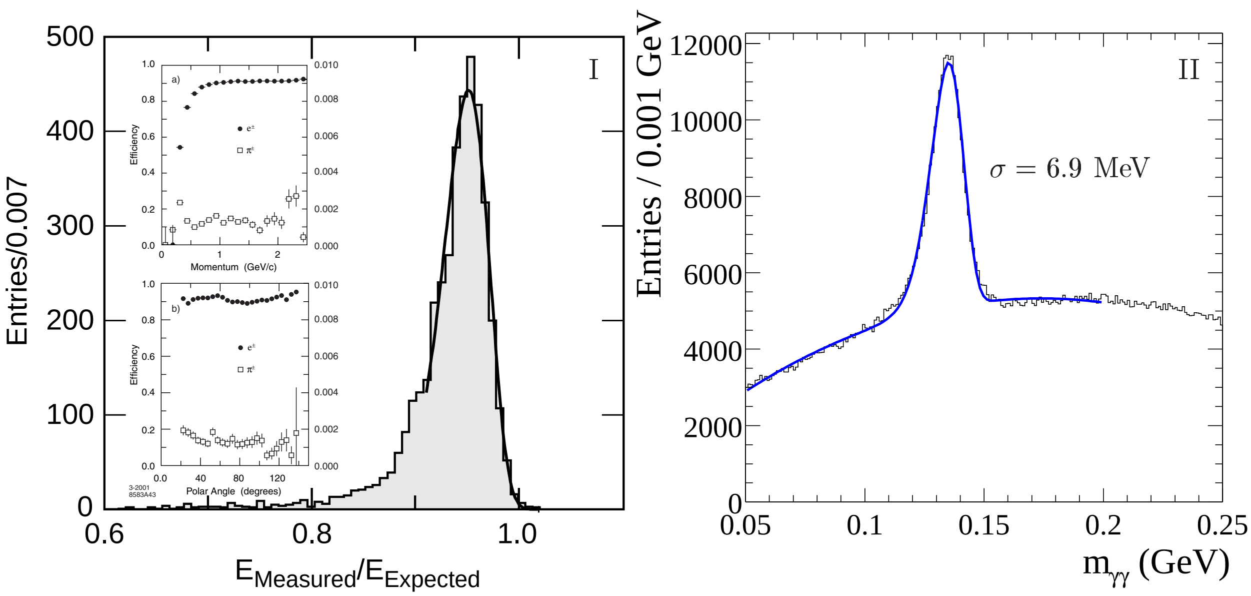


Figure 8.18: Performance of the EMC. (I) The E/p distribution for electrons centered at 0.95. The non-Gaussian tail towards lower values, is due to material in front of the EMC, leakage, and inefficiencies. In the smaller figure is shown the resulting $e^{\pm}$ efficiency (left y-axis) and $\pi^{\pm}$ misidentification (right y-axis) rate as a function of (a) momentum and (b) polar angle. (II) The π^0 mass peak in the invariant mass spectrum of two photons in $B\overline{B}$ events. The photons and π^0 candidates are required to have an energy larger than 30 MeV and 300 MeV, respectively.

8.7 The superconducting solenoidal magnet

The magnetic field needed for measuring the momentum of charged particles is provided by a thin 1.5 T superconducting solenoidal magnet. The inner radius and length of the magnet are determined by the four inner subdetectors, and the thickness by the required muon and K_L^0 efficiencies. The magnitude and uniformity specifications of the field are derived from DCH momentum resolution requirements, which suggest that a 1.5 T field with 2% uniformity in the tracking region is required. Concurrently, the field should interfere with PEP-II beam elements as little as possible. Along with the bucking coil (see Section 8.5), the magnet is controlled by the PEP-II operators, as the field influences the beam. For beam-orbit stability reasons, the bending also requires that the beam-axis is tilted slightly ($\sim 1^\circ$) with respect to the detector magnetic field.

8.8 The Instrumented Flux Return (IFR)

Particles reaching the IFR, such as muons and hadrons, require much material to interact significantly. To attain this, alternating layers of metal and detector plates are used, since the showering in the metal will create charged particles, detectable in the instrumented layers. Though charged hadrons are detected in the IFR, they are also visible to the much better inner detectors, and the IFR information is rarely used for these.

8.8.1 IFR functions and requirements

The primary task of the IFR is high efficiency muon identification over a wide range of angles and momenta, which is very important for reconstruction of vector mesons (such as J/ψ), study of semi-leptonic and rare decays of B and D mesons and not the least, tagging, needed for all time-dependent B^0 analyses. The IFR should be able to efficiently identify muons with momenta ranging from 600 MeV up to several GeV. For muon tagging purposes, the upper limit for a hadron to be misidentified as a muon should be about 5%.

Secondly, the IFR should be able to identify neutral hadrons, mainly the longlived neutral kaon, K_L^0, which is interesting for CP studies, but also neutrons. In addition, the IFR is used to veto on events with missing hadronic energy, and provides a trigger on cosmic ray events used for calibration and vetoing. Finally, and from a construction point of view most importantly, the iron layers of the IFR comprise the flux return for the solenoidal magnet, hence the name Instrumented Flux Return.

8.8.2 IFR design

Contrary to the EMC, a large fraction of the energy absorbed in the IFR is not detected, as only the shower shape and direction of the particle is detected. As muons do not interact strongly, they do not create showers, contrary to hadrons, and this is used for discrimination between the two (mainly μ/π separation).

The IFR, shown in Fig. 8.19, consists of alternating layers of steel and Resistive Plate Chambers (RPC) arranged to form a barrel and two endcaps. The barrel extends radially from 1.78m to 3.01m and is divided into sextants, each 3.75m long and 1.88m to 3.23m wide. The endcaps are hexagonal with a central hole for the beam components, giving the IFR an angular coverage of $20° < \theta < 147°$. The steel plate thickness increases from 2 cm for the nine innermost plates to 10cm for the outermost ones, for a total of 65cm (60cm) of steel and 19 (18) layers for the barrel (endcaps). In addition, two cylindrical RPC layers are placed in the barrel region between the EMC and the cryostat for the solenoidal magnet.

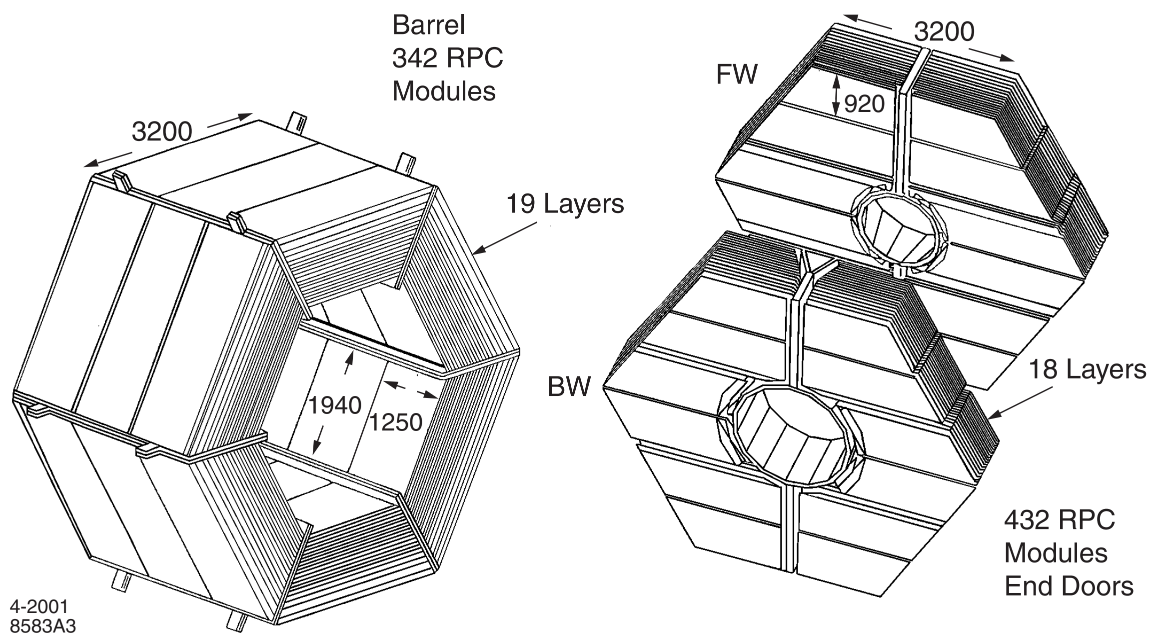


Figure 8.19: Layout of barrel and end caps of the IFR. The many layers of iron serve both as material to stop hadrons and as flux return for the magnet.

The RPCs consist of two thin Bakelite sheets treated with linseed oil, which surround a 2mm gap containing (non-flammable) argon-freon-isobutane gas. High voltage ($\sim$ 7.6 kV) is applied across the gap, thus when charged particles cause ionization in the gas, a discharge (sparks) is triggered. This induces a signal in the thin readout strips on the back of the

Bakelite, which are placed perpendicular to each other, such that they provide ϕz-position in the barrel, and xy-position in the endcaps. Two and three-dimensional clustering algorithms are applied for determining shape and position of clusters, which are also vetoed by tracks, and the resolution obtained is around a few millimetres, depending on the strip segmentation.

8.8.3 IFR operations and performance

Contrary to the other subdetectors, the IFR has not been optimally operating and performing, which has mostly been caused by temperature problems. The IR experimental hall is not temparature controlled, and high temperatures in the IFR ($> 37°$C) resulted in exceedingly high dark currents, which in turn caused HV trips repeatedly. To remedy this problem, water cooling was installed, but some RPCs continued to deteriorate, this thought to be caused by the linseed oil accumulating under the influence of the electric field.

Another problem has been the machine related backgrounds, which do not come from the IP. These cause high rates and fake signals in the outer layers of the IFR (mostly in the end caps), which have caused problems and degraded the capabilities.

The performance of the RPCs are determined using both collision and cosmic ray data, and 75% of the active RPC modules exceed and efficiency of 90%. Calibration with cosmic ray data is performed weekly.

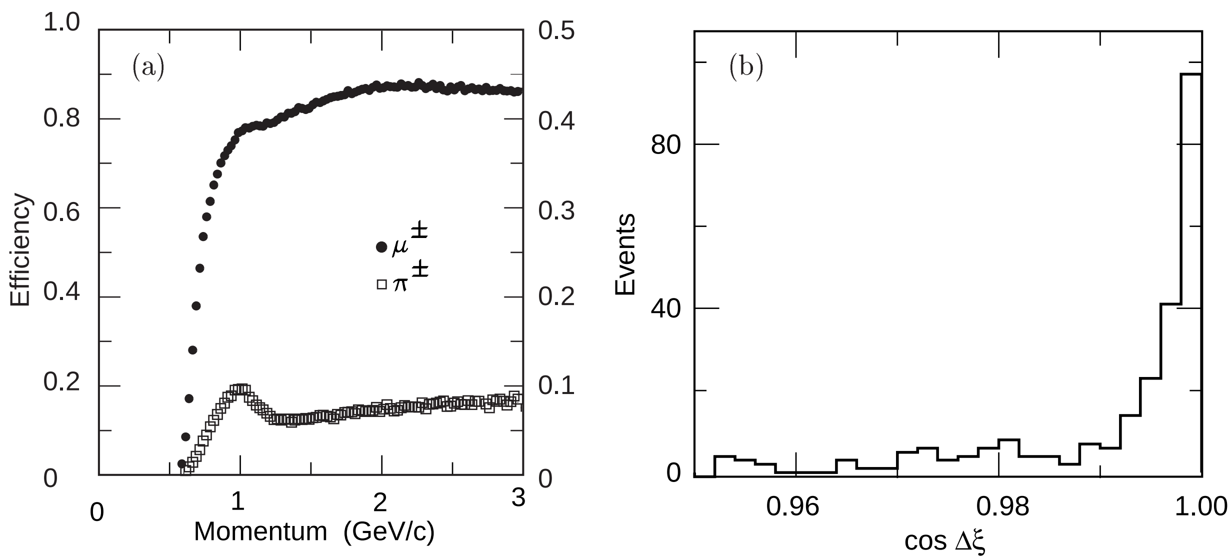


Figure 8.20: Performance of the IFR. (a) Muon efficiency (left y-axis) and pion misidentification rate (right y-axis) as a function of momentum. (b) Angular resolution for K_L^0 mesons.

The muon identification performance is evaluated on kinematically selected muon samples from $\mu\mu ee$ and $\mu\mu\gamma$ events and pion samples from K_S^0 and three-prong τ decays. The average muon efficiency in the momentum range $1.5 < p < 3.0$ GeV is for a loose (tight) selection close to 90% (80%) with a pion misidentification rate of 6–8% (2%). The muon efficiency and pion misidentification rate as a function of momentum and polar angle can be seen in Fig. 8.20.

The K_L^0 detection is based on a combination of the EMC, the two cylindrical RPCs layers and the IFR cluster information. The angular resolution is measured from $e^+e^- \to \phi\gamma \to K_S^0 K_L^0 \gamma$ events, and the difference between the missing momentum and the associated K_L^0 cluster has a resolution of about 60mrad, which improves by a factor of two, if also interacting in the EMC. The K_L^0 efficiency is estimated to be 20–40% increasing linearly with momentum.

8.9 The trigger

While the production rate of interesting physics events is relatively modest ($\sim$ 3Hz for $B\overline{B}$ and $\sim$ 15Hz for others), background events are produced copiously. To bring the data rate down to a reasonable level, such that the data can be stored (and little or no dead-time is associated with the readout) without the loss of interesting events, a selection mechanism has to be employed – this is the trigger.

The trigger must be able to select the events of interest with a very high ($>$ 99% for $B\overline{B}$), stable and precisely known efficiencies, while reducing the background to a level such that the total rate does not exceed 250Hz (120Hz at design luminosity), as required by the online computing system. In addition to the signal events, the trigger should also be able to select events used for calibration, diagnostics and luminosity determination. Furthermore, the trigger should be adjustable and tolerant to endure even extreme backgrounds (many MHz) and several dead/noisy channels, without deadtime above 1%.
To bring the rate down to 250Hz without loss of signal, a detailed analysis of each event is needed, but given the large amount of information and the fast response time needed, this is practically impossible. Therefore the trigger is divided into two levels; the first-level hardware-based trigger (L1) reducing the rate down to 2kHz followed by the third-level software-based trigger (L3)[59].

8.9.1 Level 1 trigger

The L1 trigger relies solely on information from the DCH (DCT), EMC (EMT) and IFR, the latter being used for $e^+e^- \to \mu^+\mu^-$ events and selection/veto on cosmic rays. By demanding that the DCH and the EMC trigger *each* satisfy the above requirements of being independent, redundant, and highly efficient, one obtains an extremely efficient trigger with furthermore the ability to intercalibrate the two triggers.

To obtain the required L1 speed, the detector information is condensed into a more coarse configuration (with 1776 (DCH) + 280 (EMC) + 10 (IFR) channels) with no longitudinal coordinate.
The trigger decision is based on fixed combinations of subtriggers, so-called trigger primitives, which are very simple tracks or clusters, which are above a certain momentum threshold. A global L1 triggering unit attempts to match the angular information from the charged and the neutral triggers, and includes a veto on cosmic ray events from the IFR for the final L1 trigger decision, which is formed within 11–12 μs after the corresponding interaction. While the L1 trigger decision is being made, the data is stored in buffers in the FEE (see Section 8.1).
The typical L1 trigger rate during Run-III was about 2kHz, dominated by beam-induced backgrounds, while actual physics events (see Tab. 8.5) still comprise only a small fraction of the triggers ($>$ 5%). At future higher luminosities, the DCH z-coordinate information will be included in the L1 decision to discard large backgrounds originating away from the IP.

8.9.2 Level 3 trigger

Further reduction of the rate without loss of signal requires reconstruction of the event, which is done on an online computer farm, and again the trigger is divided into two separate filters. The tracking filter requires one high $p_t > 600$ MeV track or two low $p_t > 250$ MeV tracks from around the IP, while the neutral filter demands either two high energy ($E > 350$ MeV) or four clusters and an event mass greater than 1.5 GeV. Table 8.5 shows the various input and

[59] The trigger system was designed to possibly also include a second-level fast software trigger, based on partial detector information, in case the L1 was not capable of reducing the rate enough for the L3 (hence the terminology). However, it has not become necessary, and thus at *BABAR* no second-level trigger exists.

Process	Production Rate	L1 Output	L3 Output
$B\overline{B}$	3.2 Hz	3.2 Hz	3.2 Hz
$q\bar{q}$, $\mu^+\mu^-$, $\tau^+\tau^-$	16.5 Hz	15.6 Hz	15.3 Hz
Bhabha	159 Hz	156 Hz	21 Hz (unidentified)
			45 Hz (for calibration)
Cosmic rays	$\mathcal{O}$(35 kHz)	100 Hz	< 2 Hz
Beam induced backgrounds	$\mathcal{O}$(30 kHz)	750 Hz	30 Hz
Miscellaneous/Random			25 Hz
Total	$\mathcal{O}$(30 kHz)	1000 Hz	130 Hz

Table 8.5: Trigger rates at $\mathcal{L} = 3 \times 10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$. While the physics rates scale linearly with luminosity, the backgrounds depend on more parameters (e.g. the beam-induced background is highly vacuum dependent).

output rates for the various processes. While vetoing on most Bhabha events, some are logged along with random triggers and some are selected uniformly in polar angle, for calibration. The typical output rate at current luminosity, $\mathcal{L} = 8 \times 10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ is ~ 250 Hz, as required. The L3 also monitors luminosity and $B\overline{B}$ fraction, which essentially is for energy scaling (see Section 7.3.1). An example of the L3 trigger (shown online in the control room) is seen in Fig. 8.21

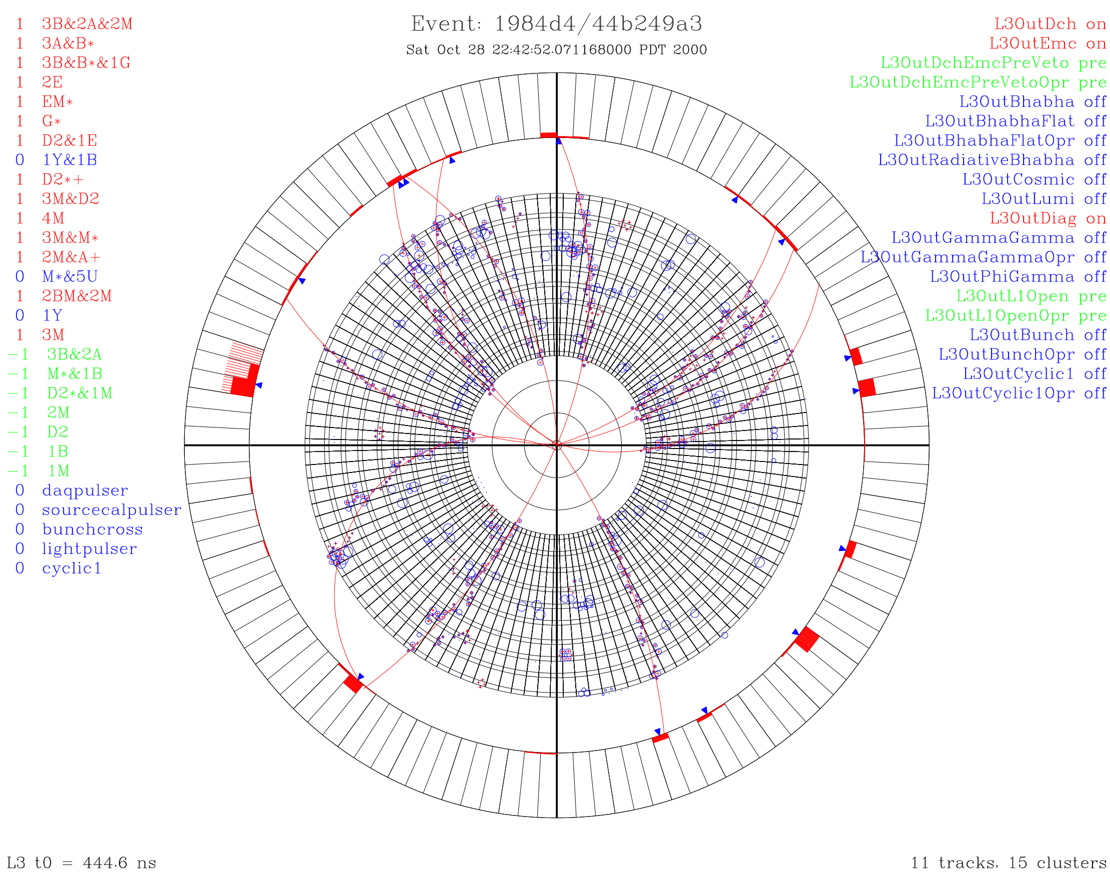


Figure 8.21: Example of event in the online L3 trigger display. To the left the L1 trigger primitives are listed, while the more in depth L3 decision is shown on the right.

8.9.3 Online Prompt Reconstruction

Online Prompt Reconstruction (OPR) of the L3 trigger output event rate is performed on a computing farm (in Padua, Italy!), leaving only the "rolling calibration" (realtime update of calibration constants) at SLAC. The experiment produces almost 1TB of data daily, and the total data volume is closing in on the 1PB (1000 TB) mark, announced at SLAC as the world's largest database. This in turn means, that reprocessing of the data, with new and improved reconstruction algorithms, was most likely done for the last time in 2003, as it has by now become a computationally almost unsurmountable task.

8.10 The *BABAR* Data

The data collected so far (January 2004) is shown in Fig. 8.22. The data taking has been divided into *runs*, of which *BABAR* is currently in the fourth (planned to end in July 2004). The data within each run is taken under stable conditions, to simplify the subsequent processing, quality checks and corrections. The period, recorded data and conditions of each run can be found in Table 8.6.

Period	Run	Int. Luminosity (fb^{-1}) Onpeak	Offpeak	Conditions
Oct. 1999 – Oct. 2000	Run I	20.721	2.602	DCH 1900/1960V
Feb. 2001 – Jun. 2002	Run II	61.158	6.978	DCH 1930V, EMC Prob.
Dec. 2002 – Jun. 2003	Run III	31.393	2.436	DCH 1960V
Sep. 2003 – end of year	Run IV	22.85	2.15	DCH 1960V
Total (New Year 2003–2004)		136.12	14.15	

Table 8.6: Running periods, recorded amount of data, and conditions.

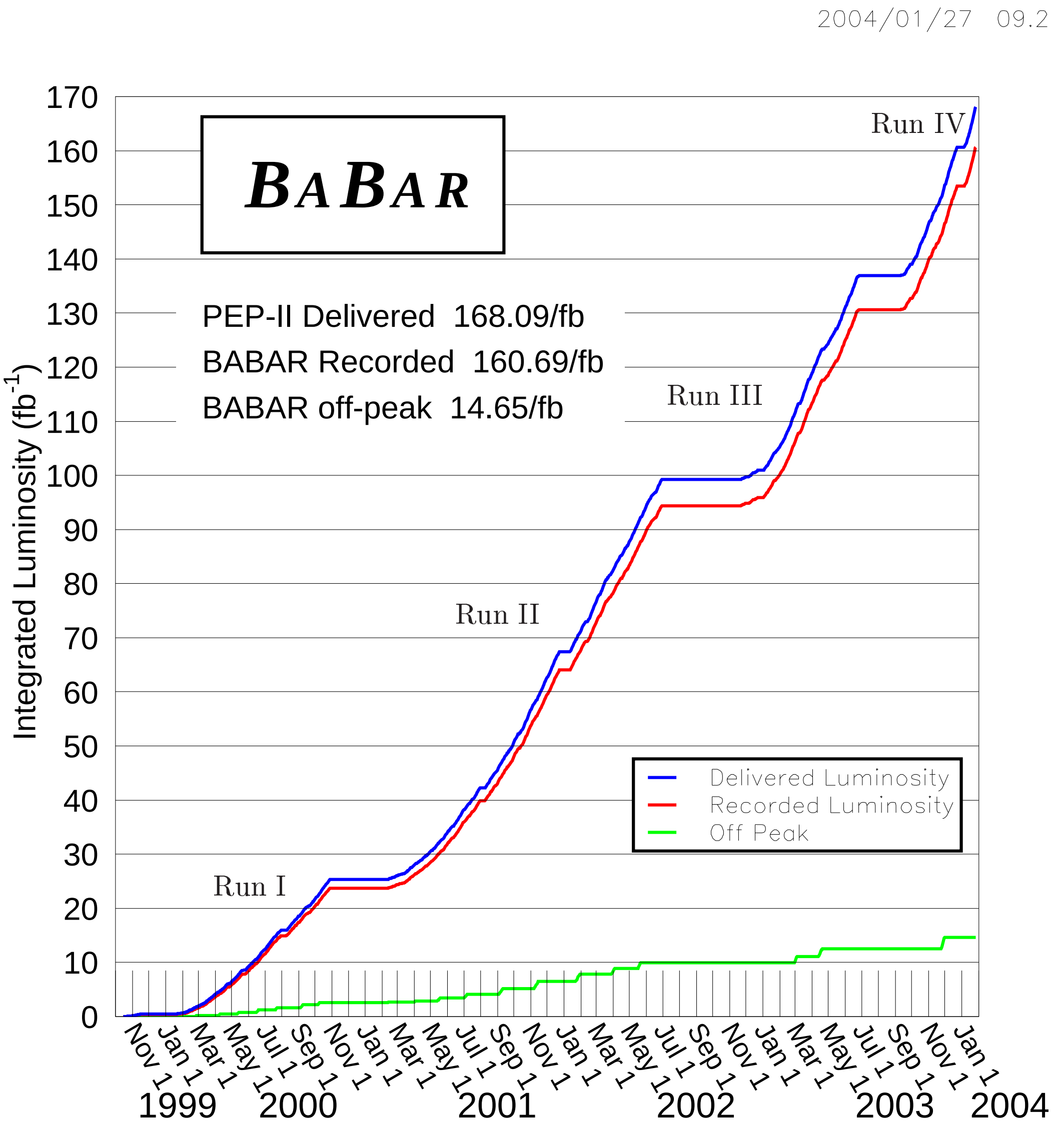


Figure 8.22: Total integrated luminosity for the period 1999–2004. The total integrated PEP-II luminosity expected after finishing Run IV (July 2004) is $\sim 240\,\text{fb}^{-1}$.

The daily integrated luminosity and daily data taking efficiency are shown in Fig. 8.23. The efficiency is generally above 95%.

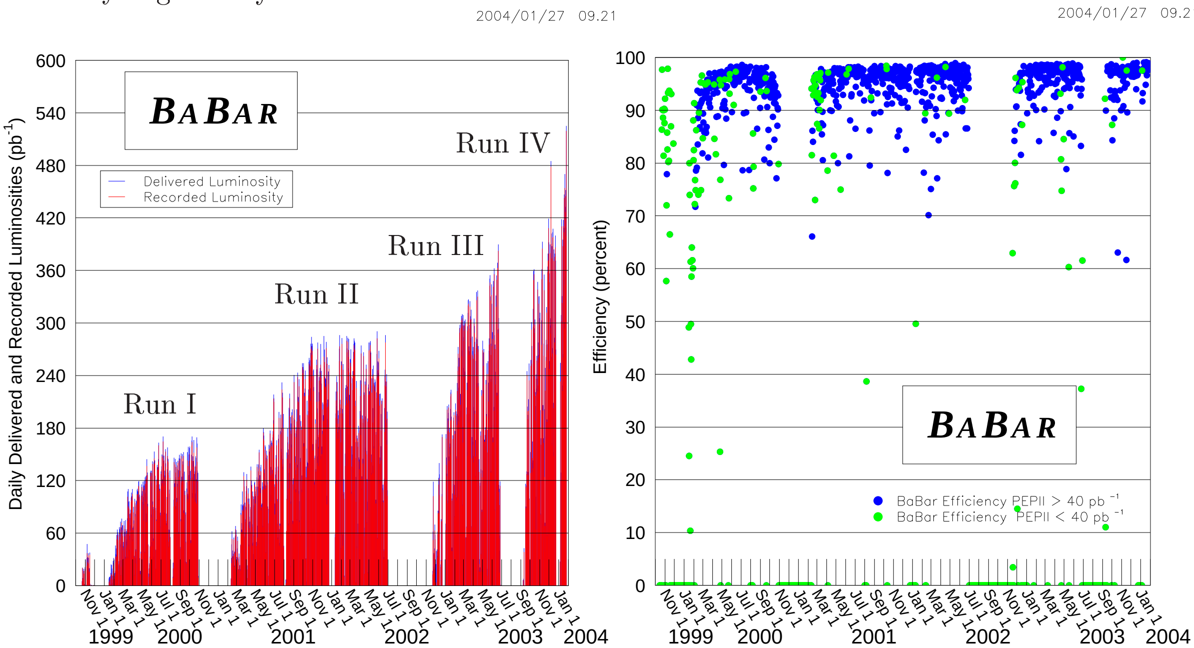


Figure 8.23: Daily integrated luminosity (left) and daily data taking efficiency (right) for the period 1999-2004.

The total number of reconstructed B mesons, using only the most clean and abundant modes (which are $B^0 \to D^{(*)-}(\pi/\rho/a_1)^+$ and $B^+ \to D^{(*)0}\pi^+$) using 81.8 fb^{-1} is $\sim 5.1 \times 10^4$ (see Fig. 8.24). However, by considering additional modes with more background (overall purity $\sim 70\%$), this number can be increased by a factor of two.

The large number of fully reconstructed B decays not only enables the study of a variety of channels, but from the combined sample of these, it allows the inclusive study of the other B meson (termed the recoil side B meson). This offers a very pure and constrained environment, utilized in many analyses (e.g. $B \to X_u \ell \nu$).

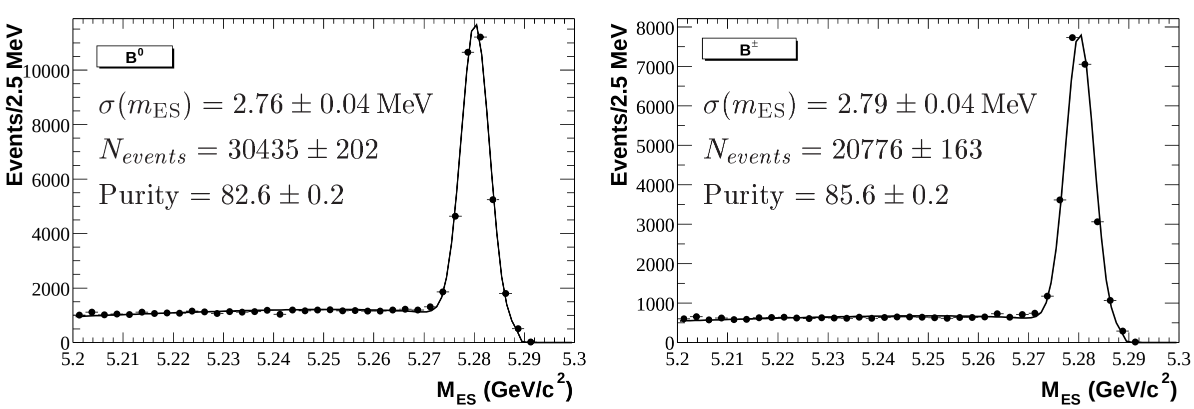


Figure 8.24: The B^0 and $B^\pm$ samples of fully reconstructed events. Shown is the distribution in the variable $m_{\rm ES}$ (see Section 9.2), with the signal peak to the right and basic parameters to the left. Larger (but less pure) samples can be obtained by including additional modes. The samples are used as high statistics control samples, and for inclusive studies in the clean and constrained recoil side.

8.11 Section summary and conclusions

In order to pursue the physics program of measuring CP violation in the B system, the *BABAR* detector should detect tracks and clusters with high efficiency and precision, and both the vertexing and PID system should perform excellently. Other issues are trigger redundancy, close monitoring, precise calibration, and tolerance to high backgrounds.
This is obtained with five consecutive subdetectors. Innermost is the silicon vertex detector (SVT), followed by the drift chamber (DCH), and then the PID system (DIRC). Outermost is the EM calorimeter (EMC), and finally the hadron calorimeter (IFR). All subdetectors were designed and optimized for B physics, though many other subjects are equally well covered.

After four years of running, the *BABAR* detector is living up to its expectations and design goals, as all subsystems are fully operational and all but the IFR, which is not a crucial part, have reached design specifications or beyond. The running has been very stable, and the overall efficiency is greater than 98% (see Fig. 8.22), which is hard (and pointless) to improve. Even with much increased instantaneous luminosity, most of the subdetectors will remain operational.

Having said that, it should be stressed that it is of key importance that each system is monitored and studied closely, such that limiting factors and show-stoppers can be foreseen and dealt with in due time. As the response time (i.e. the time to realize a problem, solve it, design a solution, build it, test it, install it, and commission it) is long, often several years, great foresight is required. The process for the new DIRC TDCs was a good example of this.

A task force created in 2002 was asked to consider the possibilities of improving the *BABAR* detector. Many proposals were considered, including e.g. a $0^{\rm th}$ SVT layer and veto-calorimetry at low polar angles, but the main conclusion of the task force was that it was hard to significantly improve the detector – thus it was built correctly in the first place.

Part III
Analysis

It is a capital mistake to theorize before one has data.

[Sir Arthur Conan Doyle, 1859-1930]

The third part of this thesis contains the data analysis, which combines the theoretical considerations developed in the first part with the experimental possibilities and constraints presented in the second part. The analysis presented here is a first step in the pursuit of a measurement of γ. It was carried out "blindly", that is to say, the signal region was not considered until the selection criteria were fixed, and the CP asymmetries not revealed until the analysis was frozen.

The first objective is to establish a signal and determine its size, purity and distribution in the Dalitz plot. As no assumptions about the Dalitz distribution can be made, the selection criteria have to be common for the entire Dalitz region. Based on the outcome of this analysis, the proportions and prospects of a time-dependent analysis are decided upon, in terms of Dalitz regions included, as too large backgrounds may inhibit the use of certain Dalitz regions.

To execute this strategy, a pre-selection and a subsequent more refined selection of events of interest is made, that is events in the signal region and sidebands, from which the background in the signal region can be determined from interpolation (Sections 9.4 and 9.5). Included in this selection are also *control samples*, that is other samples, which allow for the extraction of features from the data itself and comparison between data and MC with large statistics.

The selected events are then fitted with an unbinned maximum likelihood fit, from which the signal yields and associated branching fractions are extracted. To correct for the varying efficiency in the Dalitz plot, a novel statistical method, ${}_s\mathcal{P}$lot, is employed (Section 10).

Systematic errors are evaluated by several methods, including control samples, general recipies (essentially also from high statistics control samples) and varying unknown parameters within their uncertainties (Section 11). The results obtained are validated using simplified (toy) Monte-Carlo simulations and also the ${}_s\mathcal{P}$lot technology, which provides the true signal distribution with associated errors from the data itself (Section 13).

Though the size of the data sample is not adequate for any decisive conclusions, a time-dependent CP analysis is presented (Section 14) to establish limits and evaluate the sensitivity of such a fit.

9 Data sample and selection

9.1 Data samples

The analysis presented here is based on the Run-I and Run-II data, which together comprise 81.8 fb^{-1} on-resonance and 9.7 fb^{-1} off-resonance data. In addition 132200 Monte Carlo (MC) $D^{\pm}K^0\pi^{\mp}$ and 103200 MC $D^{*\pm}K^0\pi^{\mp}$ signal events generated uniformly in the kinematic region (i.e. Dalitz plot) are used, considering the subdecays $D^{\pm} \to K^{\mp}\pi^{\pm}\pi^{\pm}$, $D^{*\pm} \to \overline{D}^0\pi^{\pm}$ with $\overline{D}^0 \to K^+\pi^-, K^+\pi^-\pi^0, K^+\pi^+\pi^-\pi^-$ and $K^0 \to K^0_S \to \pi^+\pi^-$. Furthermore, 175.9 fb^{-1} generic $B^0\overline{B}^0$ and 143.8 fb^{-1} B^+B^- Monte Carlo simulation[60] are used. The samples are summarised in Table 9.1.

Origin	Sample	Size	Conditions
Data	On-Resonance	81.8 fb^{-1}	May 1999 – July 2002
	Off-Resonance	9.7 fb^{-1}	May 1999 – July 2002
MC	$D^{\pm}K^0\pi^{\mp}$ Signal	132200 events	Feb 2000 – June 2002
	$D^{*\pm}K^0\pi^{\mp}$ Signal	103200 events	Feb 2000 – June 2002
	Generic $B^0\overline{B}^0$	175.9 fb^{-1}	May 1999 – July 2002
	Generic B^+B^-	143.8 fb^{-1}	May 1999 – July 2002

Table 9.1: Data and Monte Carlo samples used in analysis. The data is from Run-I and Run-II, while the Monte Carlo samples are produced with the same beam and detector conditions. Assuming that $Br(B^0 \to D^{*\pm}K^0\pi^{\mp}) = 4 \times 10^{-4}$ the signal MC corresponds to samples of $11.6 \times 10^3\,\text{fb}^{-1}$ for $B^0 \to D^{\pm}K^0\pi^{\mp}$ and $5.0 \times 10^3\,\text{fb}^{-1}$ for $B^0 \to D^{*\pm}K^0\pi^{\mp}$, i.e. two orders of magnitude larger than the data sample used.

9.1.1 General comments on Monte Carlo resemblance to data

Though the detector simulation is quite detailed and involved, and has been tuned to match the data, the correspondance is never complete, and so one should keep in mind that simulated events (generally called Monte Carlo), may have slightly different distributions than the true events, due to the imperfect knowledte of the detector response.
One place where this is particularly pronounced is the central value for ΔE, which is very close to zero for MC, while it is shifted by about $-5\,\text{MeV}$ for data. The cause of this shift is thought to be the same as for the slight shift in the mass peaks, namely tiny discrepancies in tracking (see Section 8.4). The average shift from final state radiation is estimated [PD03] to be around $-1\,\text{MeV}$, which is well reproduced by MC and thus not the reason for the shift.

Furthermore, far from all decay modes of the B meson have been accounted for. In fact, only about 30% of the total branching ratio is known, and so when generating generic B decays, one is forced to use algorithms which choose final states (through a model for hadronization) according to what one observes, e.g. average number of tracks and clusters in $B\overline{B}$ events. In *BABAR*, JetSet [Sjo94] has been chosen to do this task, and ever since data started being recorded, the parameters of the algorithm have been tuned to match data with increasing precision. The overall correspondance is generally good, but certainly has its limitations. Thus, even if the detector responce was accurately simulated, the generic B decays would still not represent reality perfectly, and so one is forced to determine to which level of accuracy the simulation holds and can be used.

The effects are usually relatively modest, but they may be amplified when combined through the use of several-dimensional fits and/or neural networks. For these reasons it has been strived to use data as much as possible, which for background is done with sidebands and for signal with resembling control channels.

[60] The luminosity conversion factor used was $1.1 \times 10^6\, N(B\overline{B})\,\text{fb}$.

9.1.2 Control channels

To extract the signal shape from the data, the $B^0 \to D^{(*)-} a_1^+$ decay channels are selected as control samples (see Section 9.4.1). These channels have *the same final state as the signal*, as the subsequent decay $a_1^+ \to \rho^0(\to \pi^+\pi^-)\pi^+$ matches the final state of the signal with $K_S^0 \to \pi^+\pi^-$, i.e. three charged tracks in addition to the D meson. But contrary to the signal, these channels are not Cabibbo suppressed, and thus these high statistics channels allow for extraction of signal features from the data itself, thus not relying on MC.
Also the channel $B^0 \to D^{*-}\pi^+$ has been selected, as it is very pure, and therefore has fewer requirements in the standard selection, than the signal channels does.

9.1.3 Data topology

The distribution of number of tracks and visible energy is shown for the various event types in Fig. 9.1, while the contents and topology of the data are listed in Table 9.2.

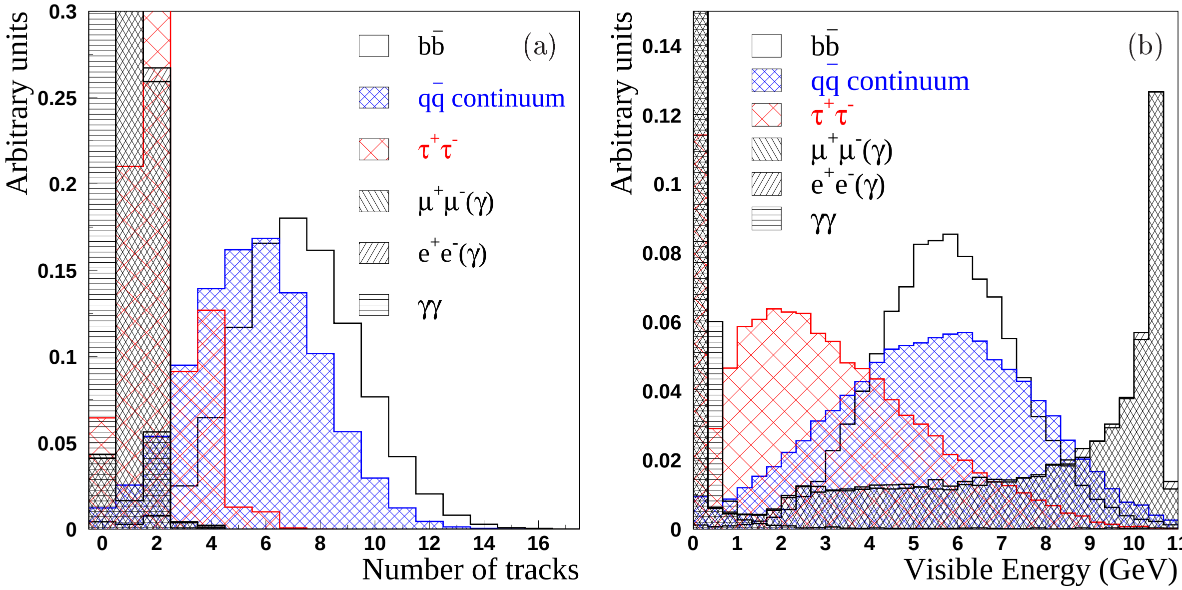


Figure 9.1: (a) Number of tracks and (b) visible energy for processes occuring in e^+e^- collisions at the $\Upsilon(4S)$ resonance. While QED processes are very distinct, $q\bar{q}$ events resemble $B\overline{B}$ decays.

Event type	Topology	Signal contamination
$e^+e^- \to \Upsilon(4S) \to B\overline{B}$	Many tracks mostly hadrons, many $D^{(*)}$	Significant
	Near-isotropic angular dist. in CM	Same angular dist.
$e^+e^- \to q\bar{q}, q = u, d, s, c$	Many tracks mostly hadrons, some $D^{(*)}$	Dominant
	Back-to-back angular dist. in CM	Different angular dist.
$e^+e^- \to \ell^+\ell^-/\gamma\gamma$	Few tracks mostly leptons	Negligible
	Back-to-back angular dist.	(contained in $q\bar{q}$)

Table 9.2: Topology of background event types. Both $B\overline{B}$ and $q\bar{q}$ (continuum) events constitute backgrounds, while $\ell^+\ell^-/\gamma\gamma$ events can be disregarded.

Though continuum events resemble the signal less than $B\overline{B}$ backgrounds, its copious production (cf. Table 7.1) makes it the dominant background. The $\ell^+\ell^-/\gamma\gamma$ background is insignificant, and is under any circumstances included in the general continuum description. $B^0 \to D^{(*)\pm}K^0\pi^\mp$ decays distinguish themselves from other processes in many ways, which will be discussed in the following.

9.2 B meson specific variables

Apart from the distributions of the individual B decay daughters (see Section 9.5), the B meson has features, which makes it distinguishable from the backgrounds. As the beam parameters are very well known, one can within the precision of these infere the magnitude (but not the direction) of the B meson's momentum. Along with the mass of the B meson, this yields the two most discriminating variables.

9.2.1 Energy substituted mass and energy difference

Given a reconstructed B candidate, the *energy substituted mass* is defined as:

$$m_{\mathrm{ES}} \equiv \sqrt{(s/2+\vec{p}_B\cdot\vec{p}_0)^2/E_0^2-p_B^2} \overset{\mathrm{CM}}{=} \sqrt{(E_0^{\mathrm{CM}})^2-(p_B^{\mathrm{CM}})^2} \tag{9.1}$$

where $s=(p_{e^+}+p_{e^-})^2$ is the total invariant mass of the system squared (i.e. essentially $m^2_{\Upsilon(4S)}$, for on-resonance data) and $q=(E,\vec{p})$ is a Lorentz vector, with the subscripts B and 0 refering to the B candidate and the beam, respectively. From the CM expression ($\vec{p}_0=0$) in Eq. (9.1) it is apparent, that m_{ES} is simply the B mass, where the reconstructed energy, which would dominate the error, is substituted with the more precisely known B energy derived from the beam energy.
The reconstructed energy is used in the *energy difference*, ΔE, which is defined as:

$$\Delta E \equiv (2q_Bq_0-s)/2\sqrt{s} \overset{\mathrm{CM}}{=} E_B^{\mathrm{CM}}-E_0^{\mathrm{CM}}, \tag{9.2}$$

where again the CM expression is very intuitive. These two variables optimally use the beam constraint and are minimally (but never-the-less) correlated, due to their common sources of uncertainty.

The uncertainty of m_{ES} arises from the beam energy spread, σ_0, and the error in the B momentum measurement in the $\Upsilon(4S)$ frame, $\sigma_{p_B^{\mathrm{CM}}}$:

$$\sigma^2_{m_{\mathrm{ES}}} \simeq \sigma^2_{E_0}+\left(\frac{p_B^{\mathrm{CM}}}{m_B}\right)^2\sigma^2_{p_B^{\mathrm{CM}}} \tag{9.3}$$

The beam energy fluctuations are $\sim 2.5\,\mathrm{MeV}$, and the momentum uncertainty is of the same order. As $(p_B/m_B)^2\simeq 0.06$ at the $\Upsilon(4S)$ resonance, the m_{ES} resolution is dominated by the beam energy uncertainty. As a result, all modes have similar m_{ES} uncertainties, in the range $[2.5;2.8]\,\mathrm{MeV}$. For ΔE, the situation is very different, as the resolution varies by more than a factor three between final states.

The uncertainty of ΔE originates from the error on the B energy measurement, $\sigma^2_{E_B}$, and again the beam energy spread, $\sigma^2_{E_0}$:

$$\sigma^2_{\Delta E} = \sigma^2_{E_0}+\sigma^2_{E_B} \tag{9.4}$$

Here, the energy resolution for the measured B energy dominates; for the modes considered in this analysis, σ_{E_B} is $10-15\,\mathrm{MeV}$ and though these are some of the smallest ΔE widths in *BABAR*, they still dominate $\sigma^2_{\Delta E}$.

m_{ES} and ΔE are the two most powerful variables for discriminating against background, and for this reason they are both used in the unbinned maximum likelihood fit. Since the main sources of uncertainty are different for m_{ES} and ΔE, the two variables are largely independent. Correlations stem from the common source of uncertainty – the beam energy spread – which is dominant for m_{ES} while only influences ΔE slightly. As the ΔE resolution decreases, this correlation increases, since the relative influence of the beam energy uncertainty on ΔE increases.

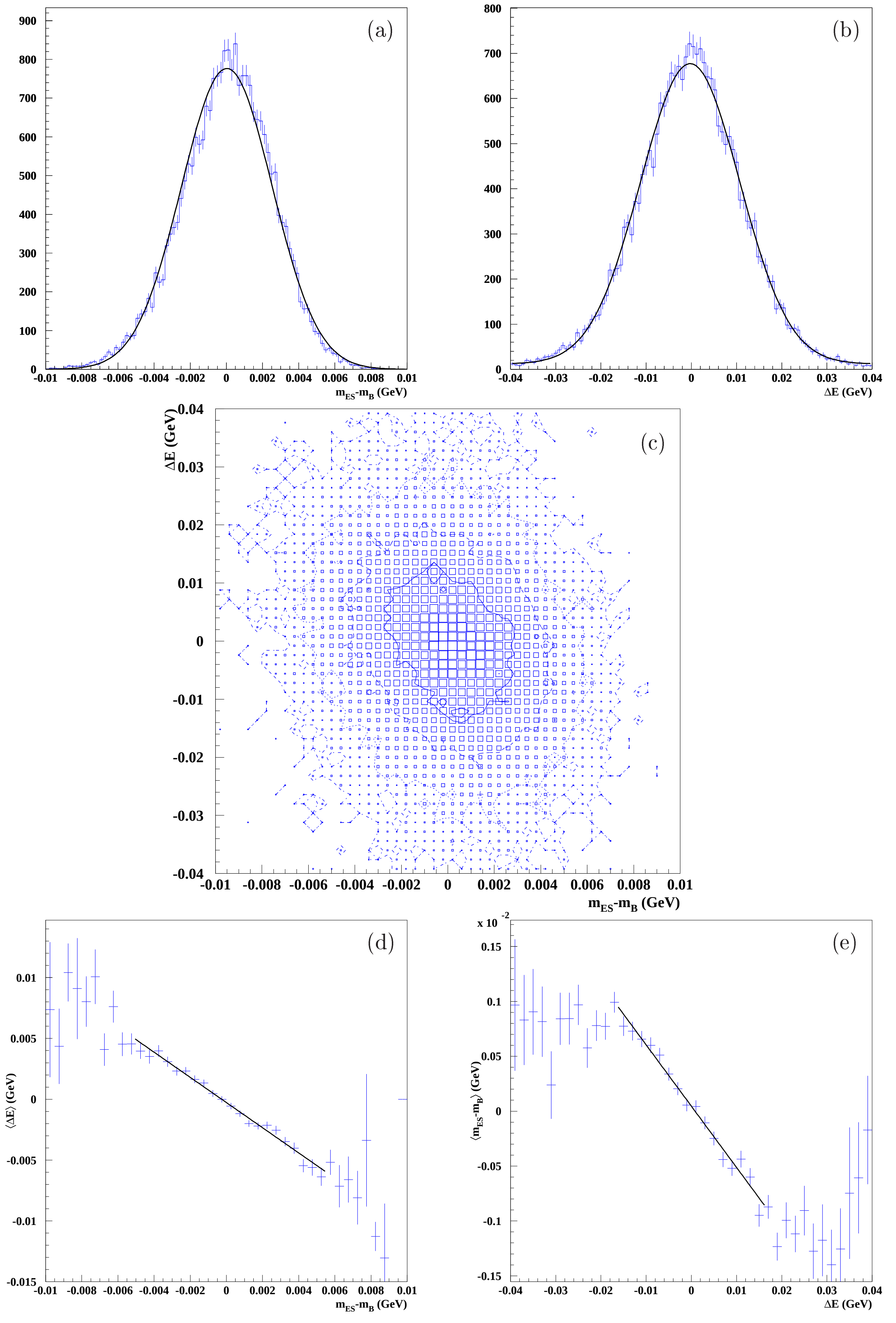


Figure 9.2: Signal MC distributions for $B^0 \to D^{\pm} K^0 \pi^{\mp}$. Distributions of (a) $m_{\rm ES}$ and (b) ΔE. (c) scatter plot of $m_{\rm ES}$ vs. ΔE. Profile plots (average value of one variable vs. another) of (d) $m_{\rm ES}$ vs. ΔE and (e) vice versa (see text). The linear correlation is -15.8%. Note that $m_{\rm ES}$ has been offset by its central value of 5.2795 GeV.

The $B^0 \to D^{(*)\pm}K^0\pi^\mp$ modes have a very well determined ΔE, due to the many mass constraints and relatively large number of tracks, and therefore the correlation is sizable though not large – the linear correlation in signal MC is measured to be -15.8%.

In Fig. 9.2 is shown for $B^0 \to D^\pm K^0\pi^\mp$ signal MC the distribution of m_{ES} and ΔE followed by a scatter plot of the two and finally profile plots which show the average value in bins as a function of the other (note that m_{ES} has been offset by its central value of 5.2795 GeV). Though not large, the correlation is seen in the scatter plot and is revealed very clearly in the two profile plots.
While the m_{ES} distribution is very nearly Gaussian, the ΔE distribution has to be fitted with a double Gaussian with common mean (see Section 10.1.1), due to additional tails from less well reconstructed events. For such events with large absolute values of ΔE, the influence of the beam energy uncertainty is less outspoken, and most likely therefore the correlation decreases for such events, as can be seen in Fig. 9.2e.

In order to have ample sidebands and yet a not too large fitting range and information loss (due to choosing in events with multiple candidates, see Section 9.7.1), only events within $[5.24, 5.29] \times [-0.1, 0.1]$ ($[5.20, 5.288] \times [-0.1, 0.1]$) in m_{ES} (GeV) and ΔE (GeV) for the $B^0 \to D^\pm K^0\pi^\mp$ ($B^0 \to D^{*\pm}K^0\pi^\mp$) mode respectively are accepted. However, to retain larger statistics for the off-resonance data, windows of $[5.20, 5.290] \times [-0.2, 0.2]$ ($[5.20, 5.288] \times [-0.2, 0.2]$) were used in that case.
The upper limits in m_{ES} at 5.29 (5.288) GeV does not influence the fit (the change in signal yield is $\sim \pm 0.04$ event), but ensures that no events lie beyond the endpoint of the background function (see Section 10.1.2), even when changing the endpoint for systematic studies.

9.2.2 Time distribution

In principle the time distribution also discriminates against the background. Correctly reconstruced B decays may have a significant Δz, while continuum events and B decays where the daughters of the two B mesons have been mixed have $\Delta z \sim 0$ modulo the resolution function. Though the resolution significantly dilutes the discrimination power, the information can still be used.
However, having foreseen a time-dependent fit (see Section 14), this information has not been used in the selection.

9.3 Shape and angular variables

The largest background consists of $e^+e^- \to q\bar{q}$ continuum events, mainly $c\bar{c}$ events, as these contain $D^{(*)}$ mesons. Unlike the signal, the $q\bar{q}$ events have considerable phase space available in the decay/hadronization, and the topology of the event will thus tend to have particles in two back-to-back jets (i.e. "jetlike"). This is illustrated in Fig. 9.3.

In addition, angular distributions in the events are discriminating, as both the momentum and thrust directions differs between hadronisation of spin 1/2 particles (continuum) and decay of spin 0 particles created through a spin 1 resonance ($\Upsilon(4S) \to B\overline{B}$).
This means that variables based solely on the shape of the event and angular information of the B candidate and the Rest-Of-Event (ROE) can be used for discrimination against continuum events. By ROE is meant all the tracks and clusters which are *not* used in the reconstruction of the B candidate. For optimal usage, the discriminant variables should be combined into one single variable, easily parametrizable and thereby suitable for a likelihood fit.

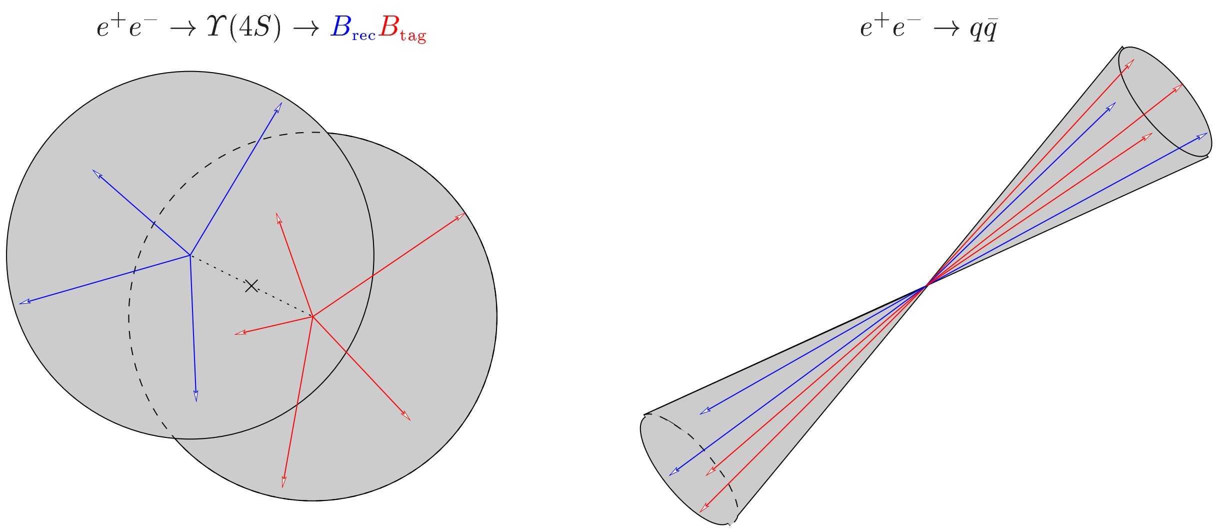


Figure 9.3: Illustration of event shape difference between $B\overline{B}$ and $q\bar{q}$ in CM. While the $B\overline{B}$ is near-uniform, the $q\bar{q}$ is more back-to-back. The difference can be used for $q\bar{q}$ discrimination. The color separates the event into reconstructed B meson (blue) and ROE (red), see text.

9.3.1 Shape variables

The quantification of "jettiness" has been done in many ways, of which the most common are[61]:

Thrust, T, and thrust axis, $\vec{T}$, are general and widely used variables, defined as:

$$T \equiv \mathrm{Max}_{\vec{T}} \frac{\sum_i |\vec{T} \cdot \vec{p_i}|}{\sum_i |\vec{p_i}|}, \tag{9.5}$$

where $\vec{T}$ is the unit direction, which maximizes T (named the *Thrust*), and the sums, i, are taken over all tracks and clusters of interest.

Sphericity, S, defined as:

$$S \equiv \tfrac{3}{2}(\lambda_1 + \lambda_2)/\lambda_3, \quad (0 < \lambda_1 < \lambda_2 < \lambda_3), \tag{9.6}$$

where λ_i are the eigenvalues of the tensor:

$$S^{\alpha\beta} \equiv \frac{\sum_i |\vec{p_{i,\alpha}} \cdot \vec{p_{i,\beta}}|}{\sum_i |\vec{p_i}|^2}, \quad (\alpha, \beta = x, y, z), \tag{9.7}$$

where the sums, i, are taken over all tracks and clusters of interest.

Fox-Wolfram Moments (FWM), H_ℓ, and their ratios, R_ℓ, defined as [ABCLOS79, FW79]:

$$R_\ell^{X_1X_2} \equiv H_\ell^{X_1X_2}/H_0^{X_1X_2}, \quad \text{with} \quad H_\ell^{X_1X_2} \equiv \sum_{i\in X_1, j\in X_2} \frac{|\vec{p_i}||\vec{p_j}|P_\ell(\cos\theta_{ij})}{E_{\text{vis}}}, \tag{9.8}$$

where P_l are the Legendre polynomials, θ_{ij} the angle between the momenta of particles i and j, and E_{vis} is the visible energy of the event. For 2-jet events $H_{\text{even}} \sim 1$ and $H_{\text{odd}} \sim 0$, and the ratio $R_2 \equiv H_2/H_0$ has been used as discriminating variable.

[61] All variables describing shape are evaluated in the CM frame.

Super Fox-Wolfram (SFW) Moments, R_{SFW}**,** introduced by Belle [Belle01], are defined in terms of FWM as:

$$R_{SFW} \equiv r_0 + \sum_{\ell=2,4} \alpha_\ell R_\ell^{B,\mathrm{ROE}} + \sum_{\ell=1,2,3,4} \beta_\ell R_\ell^{\mathrm{ROE,ROE}}. \tag{9.9}$$

$R_\ell^{B,\mathrm{ROE}}$ are much alike the Monomials (see below) with one important difference; the former uses the momenta directly, which may lead to correlations with m_{ES}, while in the latter they enter through the thrust direction only.

Absolute cosine of angle between B **and ROE thrust directions,** $|\cos(\theta_{\vec{T}_B,\vec{T}_{ROE}})|$**,** defined as the angle between the thrust axis of the B candidate and the thrust axis of the ROE. While continuum events tend to have these two axes aligned along the general jet-axis, $|\cos(\theta_{\vec{T}_B,\vec{T}_{ROE}})|$ is almost isotropic for $B\overline{B}$ events[62].

CLEO cones, C^j**,** used by the CLEO collaboration [CLEO96], defined as:

$$C^j \equiv \sum_{i\in\mathrm{ROE}} p_i \delta_i^j(\theta_i), \quad \delta_i^j(\theta_i) = 1 \quad \text{if} \quad |\theta_i| \in [10^\circ\cdot(j-1), 10^\circ\cdot j], \tag{9.10}$$

where p_i is the momentum and θ_i is the angle with respect to the B candidate thrust axis. These CM momentum flow in nine concentric 10° cones around the B candidate thrust axis are then combined liniarly with optimized coefficients α_i as $C = \alpha_0 + \sum \alpha_i C^i$.

Monomials, L_n**,** introduced by the *BABAR* collaboration [BABAR02c], defined as:

$$L_n \equiv \sum_{i\in\mathrm{ROE}} p_i |\cos(\theta_i)|^n, \tag{9.11}$$

where p_i is the momentum and θ_i is the angle with respect to the B candidate thrust axis [BABAR02c], and the sum is taken over the ROE. This is a generalization (and elegantification) of the CLEO cones, as the angular information is made continuous. The most discriminating monomials are L_0 and L_2 (including their correlation).

Not surprisingly these variables are very correlated, as they all describe somewhat the same quantity; the jettiness of the event. However, the differences are important, as the discrimination power and correlation with other variables (mostly m_{ES}) differ. The greatest conceptual difference is whether the shape variable is a sum over the entire event or over the ROE only. Unless only angular information from the B candidate is used (such as e.g. the B momentum and thrust axis), including the tracks from the B candidate can induce unwanted correlations with m_{ES}. For this reason the sphericity and the (super) FW moments are not considered when constructing a fitting variable for continuum suppression, and the thrust axis is only used when dividing the event into B candidate and ROE.

For the analysis, the monomials L_0 and L_2 were chosen, due to their simplicity and high discrimination power. The distribution of these variables can be seen in Fig. 9.4, where they are compared to a reference channel ($B^0 \to D^{*\pm}\pi^\mp$) for signal MC and off-resonance data. The specific control sample, $B^0 \to D^{*-}\pi^+$ is chosen because it is very pure and as it does not have a (default) cut on $|\cos(\theta_{\vec{T}_B,\vec{T}_{ROE}})|$ like the control channels $B^0 \to D^{(*)-}a_1^+$ do (see Section 9.4.1).

The comparison with the control sample distribution is very good, as can be seen in the plots. The L_0 distribution on signal MC (a) exhibits a slight discrepancy due to incomplete truth match in the MC for the $B^0 \to D^{*\pm}\pi^\mp$ mode, but the discrepancy is insignificant in the final combination with other variables (see Section 9.3.3). Using a Kolmogorov test

[62] The lack of perfect isotropicity of the decay is due to the slight boost ($\beta = 0.06$) of the B mesons.

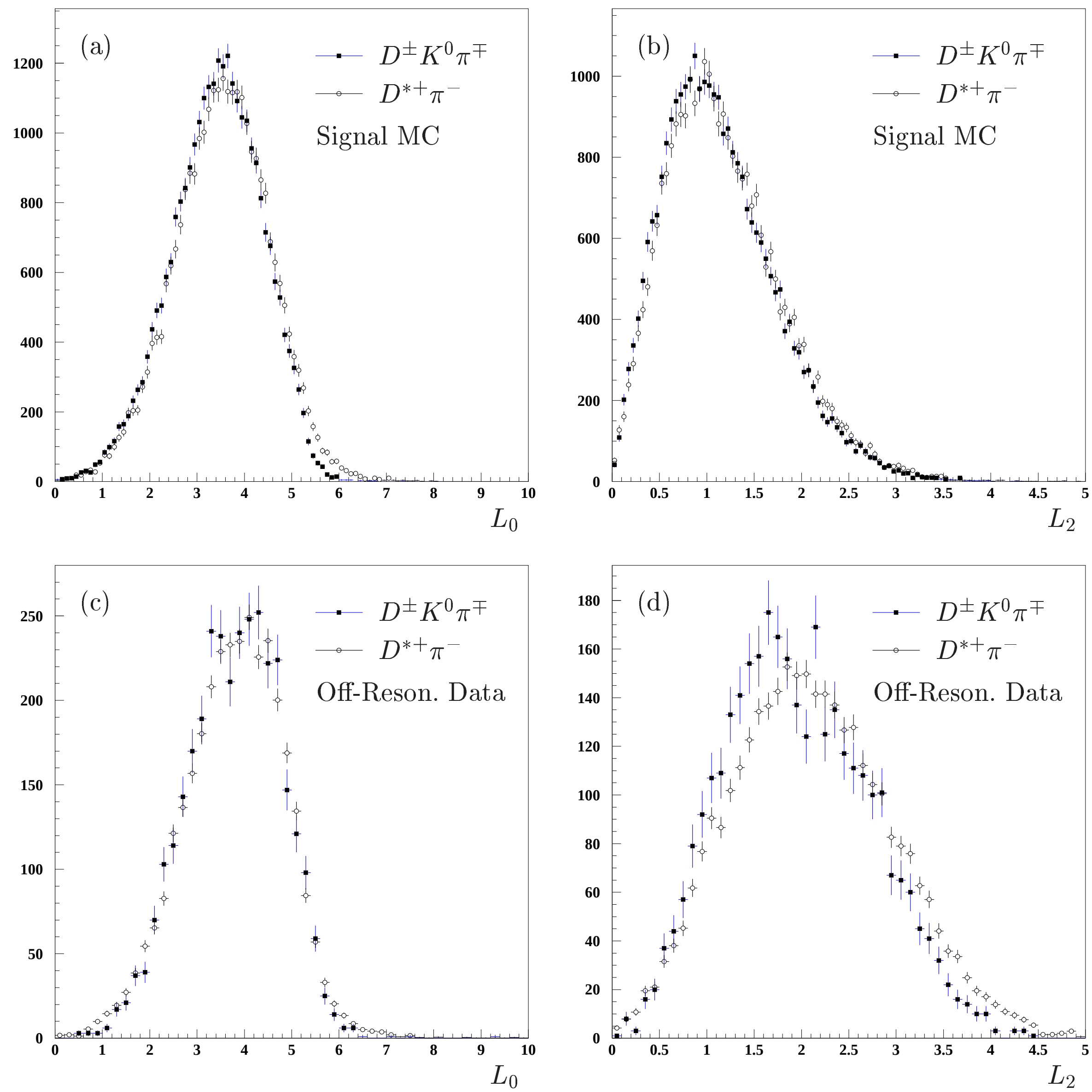


Figure 9.4: The distribution of the shape variables L_0 (left) and L_2 (right) for the signal channel $B^0 \to D^{\pm}K^0\pi^{\mp}$(blue/box/filled) and the reference channel $B^0 \to D^{*\pm}\pi^{\mp}$ (black/circle/open) in signal MC (top) and off-resonance data (bottom). The signal MC distributions are expected to be equal, while the off-resonance data distributions are shown for illustration of separation and not expected to be equal (see text).

(integration of the difference and comparing to statistical uncertainty), the distributions were assured to be in accordance otherwise.
The off-resonance distributions are not expected to be exactly identical, as no real B mesons are reconstructed, and thus two different backgrounds make up the distributions. However, it is noteworthy to see, that despite this difference, the wrongly reconstructed B candidates still show similar distributions for different decay channels.

9.3.2 Angular variables

The $\Upsilon(4S) \to B\overline{B}$ decay ($V \to PP$) follows the distribution $dN/d\cos\theta_{(\vec{p}_B,\vec{z})} = \sin^2\theta_{(\vec{p}_B,\vec{z})}$, where $\theta_{(\vec{p}_B,\vec{z})}$ is the angle between the B candidate momentum and the z-axis in the CM frame, while the $e^+e^- \to q\bar{q}$ has a uniform distribution. In addition, the B candidate thrust axis angle (in CM) with respect to the z-axis, $|\cos\theta_{(\vec{T}_B,\vec{z})}|$, carries discriminating information, as the signal distribution is nearly uniform, while the continuum background follows a $dN/d\cos(\theta_{(\vec{p}_B,\vec{z})}) = 1+\cos^2\theta_{(\vec{p}_B,\vec{z})}$ distribution. Unfortunately this variable is not as discriminating as its analytic expression might suggest, as the region of high discrimination (high $\theta_{(\vec{p}_B,\vec{z})}$) has low acceptance (thus the detector acceptance has already decided on a cut!).
The distribution of these two variables along with that of the reference channel ($B^0 \to D^{*-}\pi^+$) can be seen in Fig. 9.5.

As can be seen from the distributions, the signal shapes are very much in accordance with those of the high statistics control channel. The only exception is $|\cos\theta_{(\vec{T}_B,\vec{z})}|$, where the acceptance for very low thrust angles (which essentially matches the $\pi^\pm$ angle) vanishes for the $B^0 \to D^{*-}\pi^+$ decay, but not entirely for the $B^0 \to D^{(*)\pm}K^0\pi^\mp$decay. It has been checked that this minuscule difference does not influence the distribution of the overall discriminating variables.

9.3.3 Combining variables

While none of the above variables are by themselves very discriminating, a combination of them and their mutual correlations may be. Combining variables corresponds to substituting a crudely selected N-dimensinal box obtained by cutting on each variable separately, with a more refined selection based on the actual distance (however defined) between the signal and the background. There are several methods for combining variables, of which the two most common will be discussed.

A simple and transparent method is a linear combination of the variables in question, x_i, using coefficients, c_i. This combination, called a Fisher discriminant [Fis36], only utilizes the possible *linear* correlations between variables, but often produces signal and background distributions which are easily parametrizable. Furthermore, the impact of each variable is directly visible, and the simplicity of the method renders it commonly used.

A more powerful mean of combining variables is through a Neural Network (NN). The idea is to multiply an input vector of N_0 variables with an $N_0 \times N_1$ matrix of weights, thus yielding a new vector of size N_1 and then apply a non-linear transformation to each of its entries. If $N_1 = 1$ the approach is (apart from the non-linear transformation) exactly that of the Fisher method described above. However, if N_1 is greater than one, the operation yields a new vector with N_1 entries, each being a non-linear function of a linear combinations of the input vector (such an intermediate vector is called a (hidden) layer). The non-linearity of the function is for the NN to utilize the non-linearities in the problem. A typical function used is $F(x) = (1 + e^{\alpha x})^{-1}$.

Repeating this operation with $N_i \times N_{i+1}$ matrices of weights multiplied onto the vector of variables resulting from the last iteration, until $N_{max} = 1$, results in a single number, which is called the neural net output. If the weights of the matrices are tuned correctly, the neural net output will be able to recognize (and use) not only the input variables but also their possible non-linear correlations for background discrimination. This enables a NN to recognize patterns in data, which escape detection using a Fisher discriminant, and which increases its ability to separate signal and background.
The output is not restricted to one variable, as a network can be used to discriminate against several types of backgrounds (and signals), thus having $N_{max} > 1$. But in general one variable is preferred, and only this case will be discussed in the following.

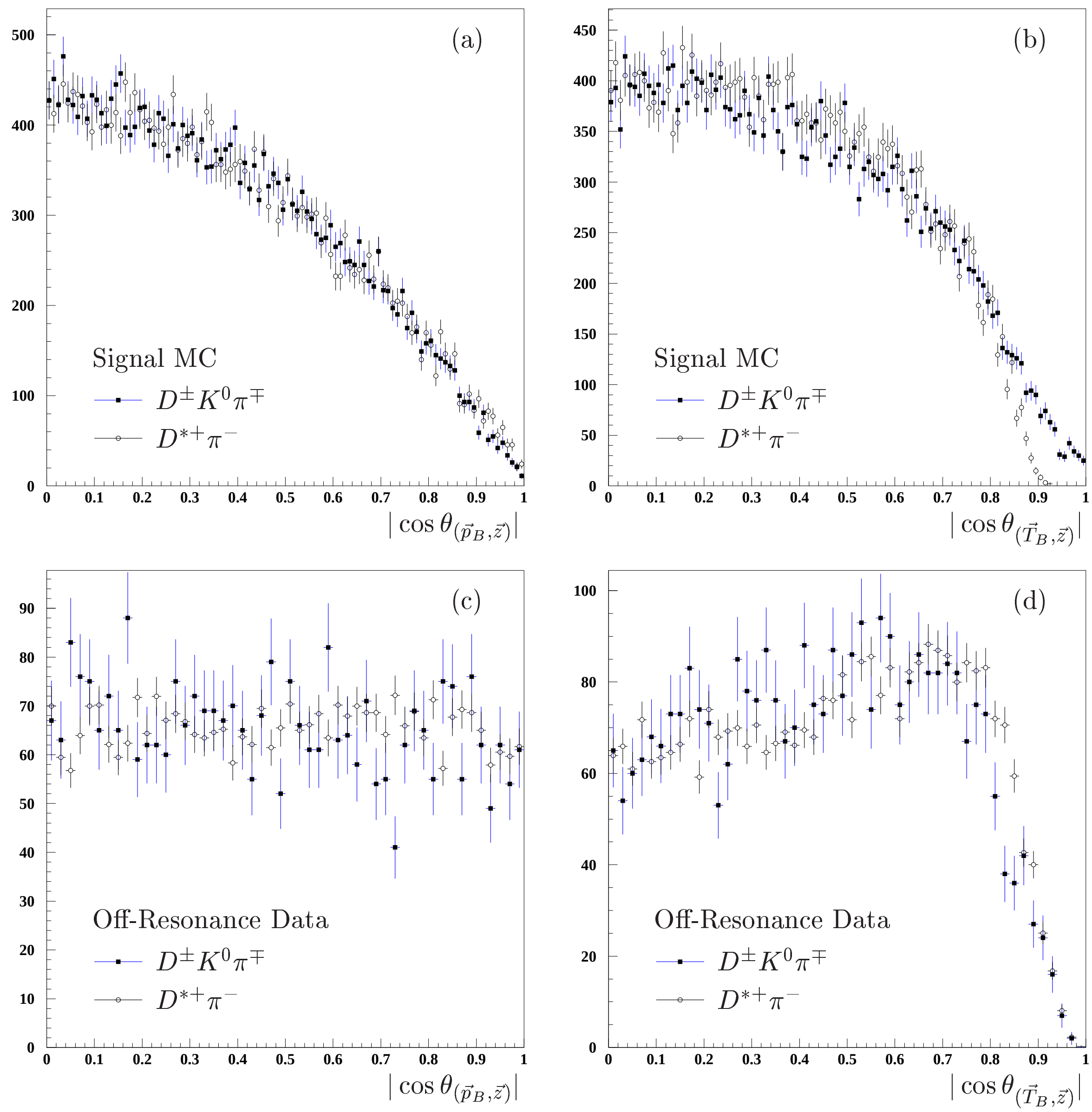


Figure 9.5: The distribution of the angular variables $|\cos\theta_{(\vec{p}_B,\vec{z})}|$ (left) and $|\cos\theta_{(\vec{T}_B,\vec{z})}|$ (right) for the signal channel $B^0 \to D^{\pm}K^0\pi^{\mp}$(blue/box/filled) and the reference channel $B \to D^*\pi$ (black/circle/open) on MC (top) and off-resonance data (bottom). The signal MC distributions are expected to be equal (though acceptance may cause differences, see text), while the off-resonance data distributions are shown for illustration of separation and not expected to be equal.

While the weights of the neural network have to be obtained through iterative optimization ("training"), the coefficients of the Fisher, $\mathcal{F}$, are the result of a direct computation[63]:

$$\mathcal{F} = \sum_i^N c_i x_i, \qquad \text{with} \qquad c_i = \sum_{j=1}^N (V_{ij}^{\text{sig}} + V_{ij}^{\text{bkg}})^{-1}(\langle x_i^{\text{sig}}\rangle - \langle x_i^{\text{bkg}}\rangle), \tag{9.12}$$

where V_{ij}^{sig} and V_{ij}^{bkg} are the covariant matrices and $\langle x_i^{\text{sig}}\rangle$ and $\langle x_i^{\text{bkg}}\rangle$ the mean values of the input variables for signal and background, respectively.

[63]The computation of course depends on how one defines the optimal separation.

In principle, all variables could be included in the combat against the continuum background, however correlated they are, but the work involved in including and validating[64] many variables does not stand in measure with the very little gain to be obtained.
Therefore it was decided, based on previous experience [BABAR03f], to use a combination of the two monomials, L_0 and L_2, along with the two angular variables $|\cos\theta_{(\vec{p}_B,\vec{z})}|$ and $|\cos\theta_{(\vec{T}_B,\vec{z})}|$. The linear correlation among these and relative to other variables can be seen in Fig. 9.6ab.

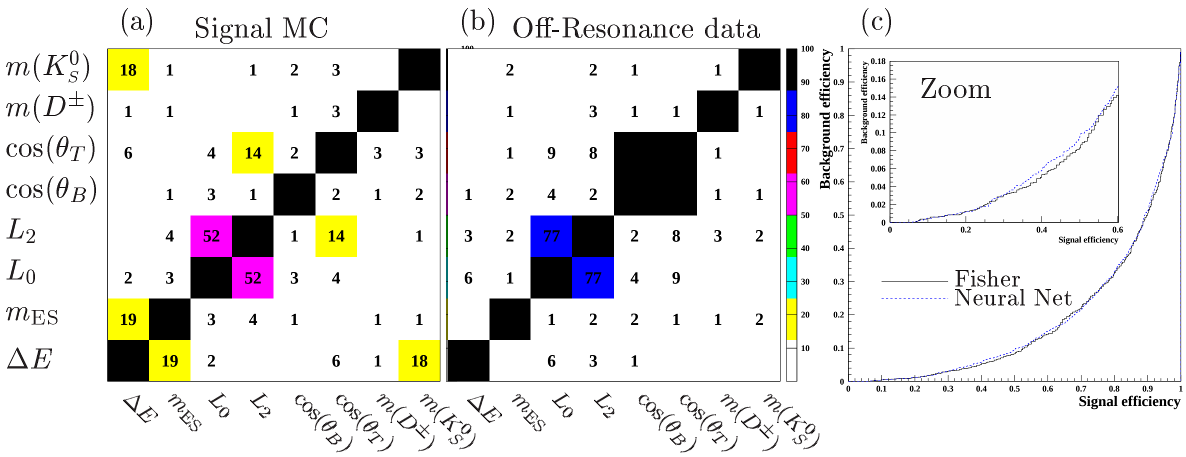


Figure 9.6: Absolute linear correlations of ΔE, $m_{\rm ES}$, the shape variables, and the $D^\pm$ and K^0_S masses for the $B^0 \to D^\pm K^0 \pi^\mp$decay in (a) signal MC and (b) off-resonance data (in %). Those of the $B^0 \to D^{*\pm} K^0 \pi^\mp$channel are similar. (c) The efficiency of the Fisher discriminant and the NN output for signal and background. The Fisher performance is close to that of the NN. Note that the abbreviations $\cos(\theta_B) = |\cos\theta_{(\vec{p}_B,\vec{z})}|$ and $\cos(\theta_T) = |\cos\theta_{(\vec{T}_B,\vec{z})}|$ have been used.

Both the Fisher discriminant, $\mathcal{F}$, and the neural network were tested. The neural network configuration used, consisted of two hidden layers of dimensions 4 and 3, chosen on the grounds that it is fast to train and does not contain too many weights (and is therefore hard to overtrain, i.e. adapt to statistical fluctuations in the training sample), while at the same time preserving the discriminating power. The resulting discrimination power of a cut on the NN output in terms of efficiency for background as a function of efficiency for signal can be seen in Fig. 9.6c.
From the figure it is apparent that the Fisher discriminant is almost as performant as the neural network, and as it is more transparent[65] and yields distributions which can easily be parametrized, it was chosen over the neural network. Other variables ($|\cos(\theta_{\vec{T}_B,\vec{T}_{ROE}})|$ and R_2), other combinations (L_2/L_0), and subdivisions (dividing L_0 and L_2 into charged and neutral components) were also tried, but none of these increased the separation with any significance and were therefore abandoned.
The four variables are combined linearly into one discriminating variable as follows:

$$\mathcal{F} \equiv c_0 + c_1 L_0 + c_2 L_2 + c_3 |\cos\theta_{(\vec{p}_B,\vec{z})}| + c_4 |\cos\theta_{(\vec{T}_B,\vec{z})}|$$

The constants c_i are chosen such that $\mathcal{F}$ maximises the separation between signal MC and off-resonance data. The constant c_0 and the overall scale and sign do not affect the separation, but are simply chosen such that the average of the distribution (of the training samples) is zero, $\langle\mathcal{F}\rangle = 0$, and such that signal in general has a (unit) higher value than background, $0.5 \sim \langle\mathcal{F}_{sig}\rangle > \langle\mathcal{F}_{bkg}\rangle \sim -0.5$. The values of the coefficients can be found in Table 9.3.

[64] I.e. compare distributions with high statistics control samples, for branching fraction measurements.
[65] In *BABAR* (and in HEP in general) there is some cultural (unfounded) reluctance to use NN.

Variable	Mean	L_0	L_2	$\lvert\cos\theta_{(\vec{p}_B,\vec{z})}\rvert$	$\lvert\cos\theta_{(\vec{T}_B,\vec{z})}\rvert$
	c_0	c_1	c_2	c_3	c_4
$D^{\pm}K^0\pi^{\mp}$ coefficients	0.463	0.415	-1.203	-0.179	-0.213
$D^{*\pm}K^0\pi^{\mp}$ coefficients	1.233	0.186	-0.913	-0.803	-0.028

Table 9.3: The Fisher coefficients for $D^{\pm}K^0\pi^{\mp}$and $D^{*\pm}K^0\pi^{\mp}$. The variable L_2 is the most discriminating, though the combination with L_0 increases the separation power significantly, due to their correlation.

To compare the resulting distributions between MC and data using these variables and coefficients, the very clean (98.0%) channel $B^0 \to D^{*\mp}\pi^{\pm}$ has been used. Since the Fisher discriminant only uses variables related to either the rest of the event or variables common to the B candidate (such as the angular distribution), all channels have the same signal Fisher distribution, neglecting possible small differences in acceptance. As can be seen in Figure 9.7 (left), the correspondance between MC and data is good.

The Fisher distribution is shown in Figure 9.7 (right) for the $D^{\pm}K^0\pi^{\mp}$mode for signal MC and off resonance data when the candidate selection has been applied. Also shown are the efficiencies for the two samples, the one as a function of the other. The distributions are very similar for the $D^{*\pm}K^0\pi^{\mp}$ mode.

The two distributions are parametrised by (and fitted with) a Gaussian with different widths above and below the mean (see Section 10.1). The separation is of the order 1σ, mostly due to the combination of the variable L_2 and L_0. $\mathcal{F}$ is used in the unbinned maximum likelihood fit, and therefore no cut is applied on it: $\mathcal{F}$ is only required to lie in the range $[-3, 3]$, to avoid outliers.

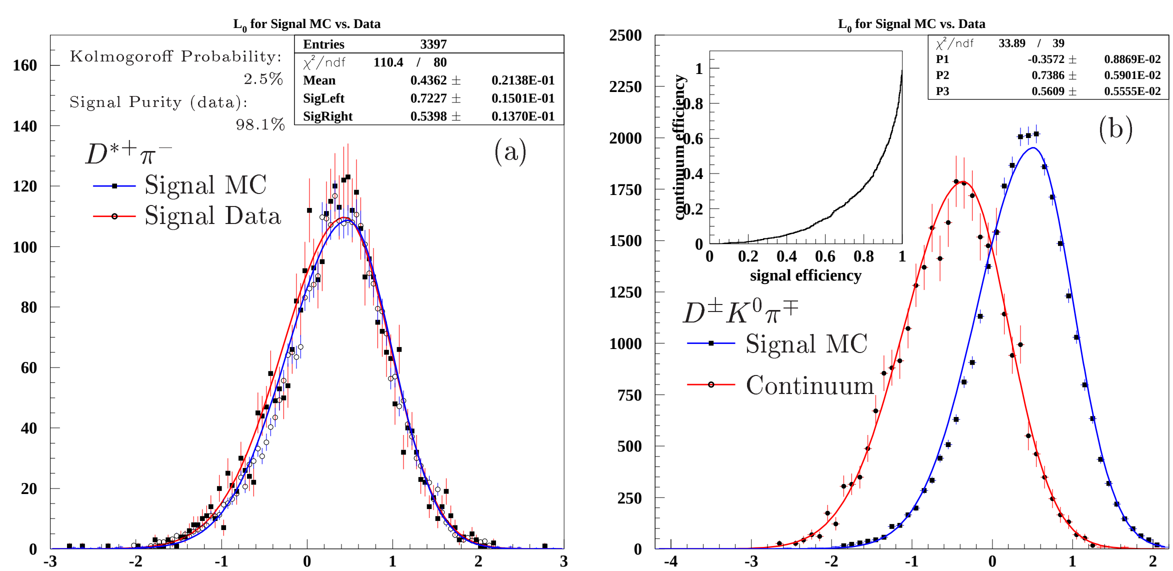


Figure 9.7: Distribution of the Fisher variable (a) compared between $B^0 \to D^{*\mp}\pi^{\pm}$ data and MC for signal, and (b) the separation between off-resonance and MC.

9.4 Pre-selection

In principle, the pre-selection could have been included in the general selection. However, it reflects some of the choices one is faced with when commencing a data analysis, and what sidebands are immediately available.

Ideally, one would like to have all data in a format allowing to directly see the impact of various background rejecting cuts by considering the sidebands. However, given the immense amount of data[66], this is not possible, and one is forced to make a pre-selection (termed a "skim" in *BABAR* jargon) of possible signal candidates and surrounding sidebands. The criteria for the pre-selection should have the highest possible signal efficiency, as later more detailed analysis will decide on when to sacrifice signal events to suppress even further background events for an improved signal significance and purity.
Two pre-selection criteria were designed for the $B^0 \to D^{\pm}K^0\pi^{\mp}$ and $B^0 \to D^{*\pm}K^0\pi^{\mp}$ selection. The submodes and cut values included in the reconstruction are described in Table 9.4 for the intermediate mesons, and in Table 9.5 for the B^0 reconstruction.

The reconstruction is based on two basic quantities, tracks and clusters, which stem directly from the detector. Each of these quantities have subclassifications according to their quality, type and origin. The quality and origin subclasses used in this analysis are the following:

`ChargedTrack` **(CT)** is a list of all charged track found by the tracking algorithm, regardless of origin and PID. By default, the pion mass hypothesis is assigned.

`GoodTrackVeryLoose` **(GTVL)** is a refinement of `ChargedTrack`, requiring that the CM momentum is less than 10 GeV, and that the Distance-Of-Closest-Approach (DOCA) is less than 1.5cm (10cm) in the transverse (longitudinal) direction.

`GoodTrackLoose` **(GTL)** is a (further) refinement of `GoodTrackVeryLoose` adding the (overlapping) requirements of a minimum transverse momentum of 100 MeV and at least 12 hits in the DCH.

`GoodPhotonLoose` **(GPL)** contains all single EMC bumps not matched with any track, which have a minimum energy of 30 MeV and a lateral moment (LAT) below 0.8. By default the photon mass hypothesis is assigned.

The particle identification is likewise discretized, though originating from momentum dependent likelihood fits. The PID criteria used in the following are:

`KMicroNotPion`, uses the SVT ($p < 0.5\,\mathrm{GeV}$), the DCH ($p < 0.6\,\mathrm{GeV}$), and the DIRC ($p > 0.6\,\mathrm{GeV}$) and requires that the K/π probability ratio is greater than 0.1 for $p \leq 0.5\,\mathrm{GeV}$ or 1.0 for $p > 0.5\,\mathrm{GeV}$. It is optimized to reject pions.

`KMicroNotPionGTL`, is the same as `KMicroNotPion`, but applied to GTL.

`KMicroTight`, uses SVT and DCH information up to 0.7 GeV and requires the K/π probability ratio to be greater than 1.0 for $p \leq 2.7\,\mathrm{GeV}$, 80.0 for $p > 2.7\,\mathrm{GeV}$ or 15.0 in the range $0.5 < p < 0.7\,\mathrm{GeV}$. It is optimized to keep π misID below 5% up to 4 GeV.

`eMicroVeryTight`, requiring the electron probability to exceed 95%. With this selection, the misID rate for hadrons is below 0.1%.

Based on these tracks and clusters, more complex candidates are formed, as described in Table 9.4, where also the reconstructed modes are listed. The sizes of the mass windows result from a trade-off between combinatorics and large sidebands. The ones chosen are standard, and roughly $\pm 7\sigma$ around the central value (except for the π^0).

[66] The *BABAR* database was announced as the worlds largest by SLAC and has surpassed the 1000 TB mark.

An enlarged D mass window was considered, as the D mass sideband is used for measuring peaking background from misreconstructed events (charmless or not), but it was dismissed due to unaffordably large combinatorics.

Decay mode	Input lists	Cut value (MeV)
$\pi^0 \to \gamma\gamma$	$2\times$ `GoodPhotonLoose`	$115 < m_{\gamma\gamma} < 150$
$K^0_S \to \pi^+\pi^-$	$2\times$ `ChargedTracks`	$m_{\rm PDG} \pm 15$
$D^+ \to K^-\pi^+\pi^+$	`KMicroNotAPion` $+ 2\times$ `GTVL`	$m_{\rm PDG} \pm 40$
$D^0 \to K^-\pi^+$	$2\times$ `GTVL`	$m_{\rm PDG} \pm 40$
$D^0 \to K^-\pi^+\pi^0$	`KMicroNotAPion` + `GTVL` + π^0	$m_{\rm PDG} \pm 70$
$D^0 \to K^-\pi^+\pi^-\pi^+$	`KMicroNotAPion` $+ 3\times$ `GTVL`	$m_{\rm PDG} \pm 40$
$D^{*+} \to \overline{D}^0\pi^+$	D^0 + `GTVL`	$139 < \delta m < 150$

Table 9.4: Cuts used in the skim for the light and charmed meson reconstruction. The slow pion used for reconstructing the D^{*+} candidate is required to have a momentum in the range $[70\,\text{MeV}, 450\,\text{MeV}]$ in the CM frame.

The charged $D^{(*)}$ candidates are combined with a K^0_S-candidate and a GTVL to form a B^0 candidate. In the preselection, the B^0 candidates are required to have $m_{\rm ES} > 5.15\,\text{GeV}$ and $|\Delta E| < 250\,\text{MeV}$, yielding ample sidebands. To further suppress background, the thrust axes of the B candidate and of the ROE are required *not* to be aligned, that is $|\cos(\theta_{\vec{T}_B,\vec{T}_{ROE}})| < 0.95$, a requirement on the second Fox-Wolfram moment ($R_2 < 0.45$) was also imposed. These quantities are used, as they are available at the pre-selection level (see Section 9.3.1). The reason why the cuts are very loose is that dedicated continuum discrimination is planned for at a later stage. The cuts are summarized in Table 9.5.

Decay	$m_{\rm ES}$	$\lvert\Delta E\rvert$	R_2	$\lvert\cos(\theta_{\vec{T}_B,\vec{T}_{ROE}})\rvert$
$B^0 \to D^{(*)\pm}K^0\pi^\mp$	$> 5.15\,\text{GeV}$	$< 250\,\text{MeV}$	< 0.45	< 0.95

Table 9.5: Cuts used in the skim at the B reconstruction level (see Table 9.15 for efficiency).

The efficiency of the pre-selection (termed the “raw” efficiency) is on signal MC 27.9% for $D^\pm K^0\pi^\mp$, and 27.5%, 10.6%, and 14.3% for $D^{*\pm}K^0\pi^\mp$ into the three decay modes of the D^0, the main losses being due to acceptance[67]. This selection yields 1.216.709 and 320.016 events on-resonance, and 100.906 and 18.722 events off-resonance for the $B^0 \to D^\pm K^0\pi^\mp$ and $B^0 \to D^{*\pm}K^0\pi^\mp$ modes, respectively. These events are considered in the following.
Many other decay modes were included in the skim ($B^\pm \to D^{(*)0}K^\pm(\pi/\rho)^0$ and $B^0 \to D^{(*)0}K^0(\pi/\rho)^0$) for other three-body $DK\pi$ studies and cross checks, but it was decided to focus on the cleanest modes.

9.4.1 Control channels

The decay channels $B^0 \to D^{(*)\mp}a_1^\pm$, which have the exact same final state as the signal channels in question, have also been reconstructed. As they are not Cabibbo suppressed and therefore have large statistics, they serve excellently as control channels for extracting the signal shapes. As the signal channels are thought to contain K^* resonant contributions, even the Dalitz distributions will be somewhat matching, since the a_1 decay proceeds dominantly through a $\rho\pi$ intermediate decay.

[67] A rule of thumb is that 15% of tracks and 25% of clusters are outside the acceptance or lost in background.

The same cuts are applied to these samples as for the signal channels (see Table 9.11), with the K_S^0 cuts as obvious exceptions. As the three tracks originate from an $a_1^\pm$ decay, they are all required to be GTVL, fail the `KMicroTight` requirement, have a common vertex, and an invariant mass in the range $1.0\,\mathrm{GeV} < m_{\pi\pi\pi} < 1.6\,\mathrm{GeV}$. From fitting these samples with the same PDFs as the ones used for the signal (see Section 10), the parameters of the signal shape in m_{ES} and ΔE are obtained and subsequently fixed.

The shape of the Fisher distribution, $\mathcal{F}$, is *not* taken from these samples, as the $B^0 \to D^{(*)\mp}a_1^\pm$ selections (standard in *BABAR*) have a tighter cut on $|\cos(\theta_{\vec{T}_B,\vec{T}_{ROE}})|$ (less than 0.7 for $B^0 \to D^\mp a_1^\pm$ and 0.8 for $B^0 \to D^{*\mp}a_1^\pm$), than the signal channels do ($|\cos(\theta_{\vec{T}_B,\vec{T}_{ROE}})| < 0.95$). As $\mathcal{F}$ and $|\cos(\theta_{\vec{T}_B,\vec{T}_{ROE}})|$ are highly correlated, this cut alters the distributions, and so these modes cannot be used. However, $\mathcal{F}$, being based on angular variables of the B^0 meson and ROE quantities, its distribution is almost independent of the decay mode. Therefore the distribution is taken from the very pure channel $B^0 \to D^{*-}\pi^+$, which does not have any cut on $|\cos(\theta_{\vec{T}_B,\vec{T}_{ROE}})|$ at the pre-selection level. The same cuts on the B and D candidate are applied to this channel as for the signal modes, after which the signal distribution is extracted. The Fisher distribution obtained have been checked thoroughly and compared with MC (see Section 9.3).

9.5 Candidate selection

While the coarse pre-selection is based on general experience without regard to multiple candidates and is meant to include ample sideband in all dimensions, the actual candidate selection is an optimization of the expected signal significance. The selection criteria are obtained essentially by maximising the quantity $S/\sqrt{S+B}$, where S and B are the number of signal and background events, obtained from fitting the most discriminating variable m_{ES} with a Gaussian for the signal, and a generic endpoint function (known as the Argus function [ARGUS87b]) for the background. The Argus function, $\mathcal{A}$, is an empirical endpoint function, which describes the shape of the combinatorial background (i.e. random combinations of tracks and clusters) well (see Eq. 9.13).

$$\mathcal{A}(m_0,\xi,m_{\mathrm{ES}}) \;=\; N_B m_{\mathrm{ES}}\sqrt{1-(m_{\mathrm{ES}}/m_0)^2}\,e^{\xi(1-(m_{\mathrm{ES}}/m_0)^2)}, \quad (m_{\mathrm{ES}} < m_0), \qquad (9.13)$$

where m_0 is the upper kinematic limit (threshold of the beam energy), ξ controls the shape of the function, and N_B is the normalization. A typical example of a fit in m_{ES}, where the background is described by the Argus function can be found in Fig. 8.24.

For the maximisation of the signal significance $S/\sqrt{S+B}$ from a fit in m_{ES}, $|\Delta E| < 20\,\mathrm{MeV}$ is required to render the situation closer to that of the actual fit. The background is defined as the Argus function integrated over the range [5.272,5.288].
As most optimization curves have plateaus, simple values have been chosen in an attempt to keep the cuts standard. Vertex fits are required to either be converged or to have a probability above 0.1%, and PID information is also discretized before use (as described in Sec. 9.4).

The distributions shown in this section are generally blue when showing MC and red when showing data. The fits to the particle mass spectra are only meant for checks and determination of the central value and widths. They are not used otherwise. Unless otherwise specified, all distributions and figures are obtained when applying a set of reference cuts (see Table 9.11), *except on the variable under consideration.*

9.5.1 π^0 selection

π^0 candidates are formed from pairs of `GoodPhotonLoose` with total energy larger than 200 MeV. The mass distribution of the π^0 candidates is shown in Figure 9.8, where it has been fitted with a Gaussian distribution[68] for the signal plus a constant for the background, as suggested by generic MC. The fitting results can be found in Table 9.6.

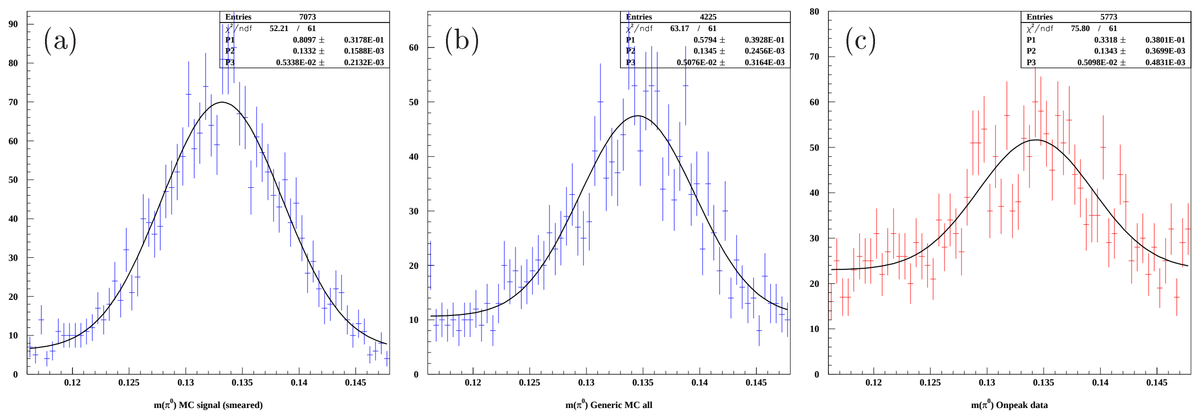


Figure 9.8: The π^0 mass distribution for (a) signal MC, (b) generic MC, and (c) data.

The widths are very much in accordance, and about 10% below the *BABAR* general mass resolution for π^0 of 6.2 MeV as expected (see Section 8.6.3). This narrowness is also a result of constraining the D^0 mass to its nominal value (see Section 9.5.4), as this slightly modifies (within errors) the direction of the two photon candidates and thus the π^0 mass. In a Kolmogorov test, the accordance between generic MC and data was better than 5%, thus both signal and background are well reproduced by the generic MC. The accepted π^0 candidates are mass constrained to the nominal value [PDG02]. No further cut in mass or other variables[69]

Sample	π^0 mass (MeV)	π^0 width (MeV)
Signal MC	134.4 ± 0.1	5.34 ± 0.21
Generic MC	134.5 ± 0.2	5.08 ± 0.32
Data	134.3 ± 0.4	5.10 ± 0.48

Table 9.6: The π^0 mass fit results on signal MC, generic MC and data.

is applied, as later requirements ($\delta m(D^*, D^0)$) greatly purify the $D^0 \to K^-\pi^+\pi^0$ channel (the only channel where π^0 candidates are used).

9.5.2 K_S^0 selection

The K_S^0 candidates are formed from two `ChargedTracks`, with a common vertex ($P_{\rm vtx}(K_S^0) > 0.001$). A significant transverse flight length, L_2, requirement of $L_2/\sigma_{L_2} > 4.0$ is imposed to suppress background from the vastly more abundant $B^0 \to D^{(*)\pm}a_1^{\mp}$ (control) channels.
The K_S^0 mass distribution is fitted with a double Gaussian with common mean plus a constant, as suggested by using the truth information in the generic MC. The results both for generic MC and data can be seen in Fig. 9.9 and Table 9.7. To compare the overall width of data and MC, a weighted average of the two widths, $\sigma_{\rm WA} = f\sigma_{\rm narrow} + (1-f)\sigma_{\rm wide}$ is calculated. The K_S^0 mass spectrum is well reproduced by the MC and required to be within ± 7 MeV of

[68] Since the distribution has a tail below the mean, a Crystal Ball Function [CrBall] is more accurate, but as m_{π^0} is not used in the likelihood fit and have very wide limits, the use of this function is not necessary.
[69] Quantities such as the angular photon distribution and the lateral moments were considered.

the K_S^0 mass. Once accepted, the mass of the K_S^0 candidate is constrained to the PDG value [PDG02].

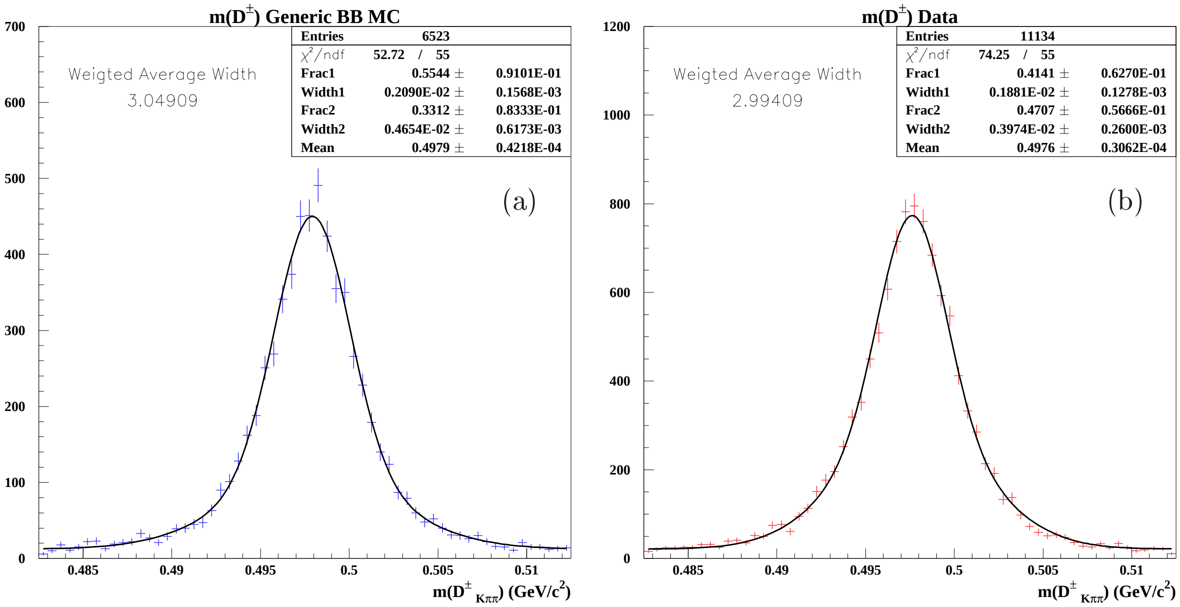


Figure 9.9: The K_S^0 mass distribution for (a) generic MC and (b) data. The distributions are fitted with a double Gaussian with common mean, $fG(m, \sigma_{\text{narrow}}) + (1-f)G(m, \sigma_{\text{wide}})$.

Sample	$m(K_S^0)$ (MeV)	f	σ_{narrow} (MeV)	σ_{wide} (MeV)	σ_{WA} (MeV)
Gen. MC	498.0 ± 0.03	0.63 ± 0.10	2.09 ± 0.16	4.65 ± 0.62	3.05
Data	497.6 ± 0.04	0.47 ± 0.07	1.88 ± 0.13	3.97 ± 0.26	2.99

Table 9.7: The K_S^0 mass fit results on generic MC and data. The weighted average $\sigma_{\text{WA}} = f\sigma_{\text{narrow}} + (1-f)\sigma_{\text{wide}}$ is calculated to enable a comparison of the overall widths.

9.5.3 $D^{\pm}$ selection

The $D^{\pm}$ candidates are reconstructed into the mode $K^{\mp}\pi^{\pm}\pi^{\pm}$ from three GTVL[70], which have a converged vertex fit. The track with the opposite charge of the two others is required to be identified as a kaon (`KMicroNotPionGTL`). The $D^{\pm}$ mass distribution for generic MC and data is shown in Figure 9.10 and the results of fitting the distributions with a single Gaussian plus a first degree polynomical (as suggested by generic MC truth matching) can be found in Table 9.8.

Sample	$D^{\pm}$ mass (MeV)	$D^{\pm}$ width (MeV)
Generic MC	1867.5 ± 0.3	6.1 ± 0.3
Control Sample	1867.8 ± 0.2	6.4 ± 0.2
Data	1867.0 ± 0.5	6.4 ± 0.5

Table 9.8: The $D^{\pm}$ mass fit results on generic MC, control sample and data.

[70]The $D^{\pm} \to K_S^0\pi^{\pm}$ mode was also considered, but though it is equally pure, it only adds approximately 10% to the statistics, and was therefore not included in the data analysis.

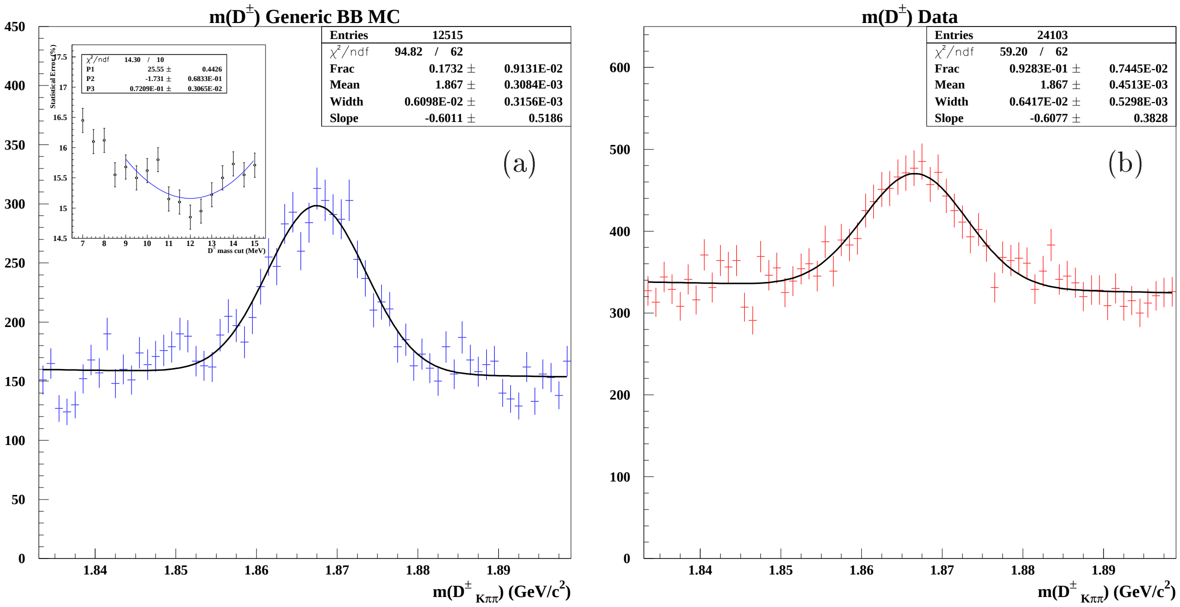


Figure 9.10: The $D^{\pm}$ mass distribution for (a) generic $B^0\overline{B}^0$ MC and (b) data. The plot inserted in figure (a) is the optimization curve, where the error from degenerate peaking background has been included. A cut of $\pm 12\,\mathrm{MeV}$ around the mass peak was chosen.

The $D^{\pm}$ width is in good agreement with that of the generic sample and the control sample, $B^0 \to D^{\mp}a_1$. In the optimization of the selection cut, the effect of peaking background (see Section 10.1.4) was included as this had a non-negligible impact, leading to a slightly tighter cut. The optimization curve is shown (inserted) in Fig. 9.10a. The somewhat tight cut of $\pm 12\,\mathrm{MeV}$ around the mean of the data reflects combinatorial and peaking background which are far from small, as the plot also shows. Finally, the $D^{\pm}$ mass is constrained to that of the PDG value [PDG02].

As much background is due to wrongly reconstructed $D^{\pm}$ mesons, i.e. combinatorial background, other parameters were considered for possible background rejection. With a lifetime of $\tau_{D^{\pm}} = (1051 \pm 13) \times 10^{-15}\,\mathrm{s}$ [PDG02], the average flight length of the $D^{\pm}$ meson is of the same order as the vertex resolution, and thus discrimination can be obtained. In addition to the flight length, the angle between the flight direction and the $D^{\pm}$ momentum carries information.

For each of these variables, a transverse and a three-dimensional quantity is calculated, as the transverse direction is more precisely determined, but the z-direction still carries information. Along with associated errors, this yields eight variables, where the transverse quantities are the most discriminating. Also the vertex fit probability $P_{\mathrm{vtx}}(D^{\pm})$ and the kaon direction in the CM of the $D^{\pm}$ with respect to the $D^{\pm}$ flight direction, $\cos\theta(\vec{p}_K^{\,\mathrm{CM}(D)}, \vec{p}_D)$, have discriminating properties. As the Dalitz distribution of the daughters is somewhat flat (thus does not carry much discriminating power) and not necessarily well modelled (thus possibly biasing the efficiency calculation), it was decided not to use it.

The ten variables discussed above were combined in a neural network, which was trained using data sideband of the $D^{\pm}$ for the background and generic MC for the signal[71]. As no very pure sample containing $D^{\pm}$ exists, alternatives to avoid relying on MC seem distant (e.g. the $B^0 \to D^-\pi^+$ sample is "only" about 80% pure). The resulting separation between signal MC and background data is significant, as can be seen in Fig. 9.11.

[71] The use of two different sources (data/MC) might lead the network to recognize and use these differences.

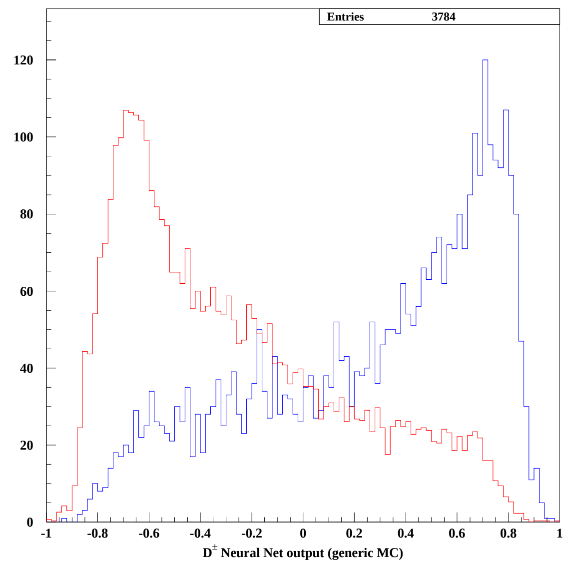


Figure 9.11: The $D^{\pm}$ neural network separation for signal MC (blue/right) and background data (red/left). Significant discrimination between the two is obtained.

Due to possible differences between data and MC, the use of a control channel is essential for estimating the efficiency, if this rejection technique is to be used. As an additional check, the mass distribution could be used, as the NN output should not be correlated with this.
If fake $D^{\pm}$ mesons were the single dominant source of background, the neural network would surely have been included in the analysis, either through a cut or (combined with the mass) as an additional variable in the likelihood fit, as the NN carries significant additional discrimination power. However, this is not the case in the final B^0 selection. Despite the large combinatorial background for the $D^{\pm}$ seen in Fig. 9.10, these are diminished by the (indirect) requirements on m_{ES} and ΔE, and so there is not very much to be gained in applying the network. For these reasons, the neural network for discriminating against fake $D^{\pm}$ candidates has *not* been included, and will *not* be considered in the following.

9.5.4 D^0 selection

The D^0 candidates are reconstructed in the three modes $K^-\pi^+$, $K^-\pi^+\pi^0$, and $K^-\pi^+\pi^-\pi^+$. In the latter two channels, the kaon is required to be identified as such (`KMicroNotPion`). The $K^-\pi^+\pi^0$ candidates are also required to originate from the more densely populated areas of the Dalitz plot ($w_{\mathrm{dalitz}} > 1.4\%$)[72] and the vertex of the two charged tracks must be plausible ($P_{\mathrm{vtx}} > 0.001$). For the two fully charged modes, the vertex fit is simply required to converge. The D^0 mass spectrum has been fitted with a Gaussian and a polynomial of second degree (as suggested by considering the right and wrong truth match of generic MC). The distributions can be seen in Figure 9.12 and the various fitting results in Table 9.9. The MC reproduces well the shapes of the data. In the data, the mass peaks are slightly below the nominal values ($\sim 0.05\%$), which is attributed to remaining tiny residuals in the alignment and parametrization of the magnetic field (the shift is seen for all mass peaks). In the $K\pi\pi^0$ channel, the slight shift is also caused by the π^0, which has a tail in the mass spectrum below the peak. The mass of the D^0 candidates are constrained to the PDG value [PDG02] before further reconstruction.

[72] The Dalitz weight, w_{dalitz}, is essentially the $D^0 \to K^-\pi^+\pi^0$ decay amplitude (maximally 1).

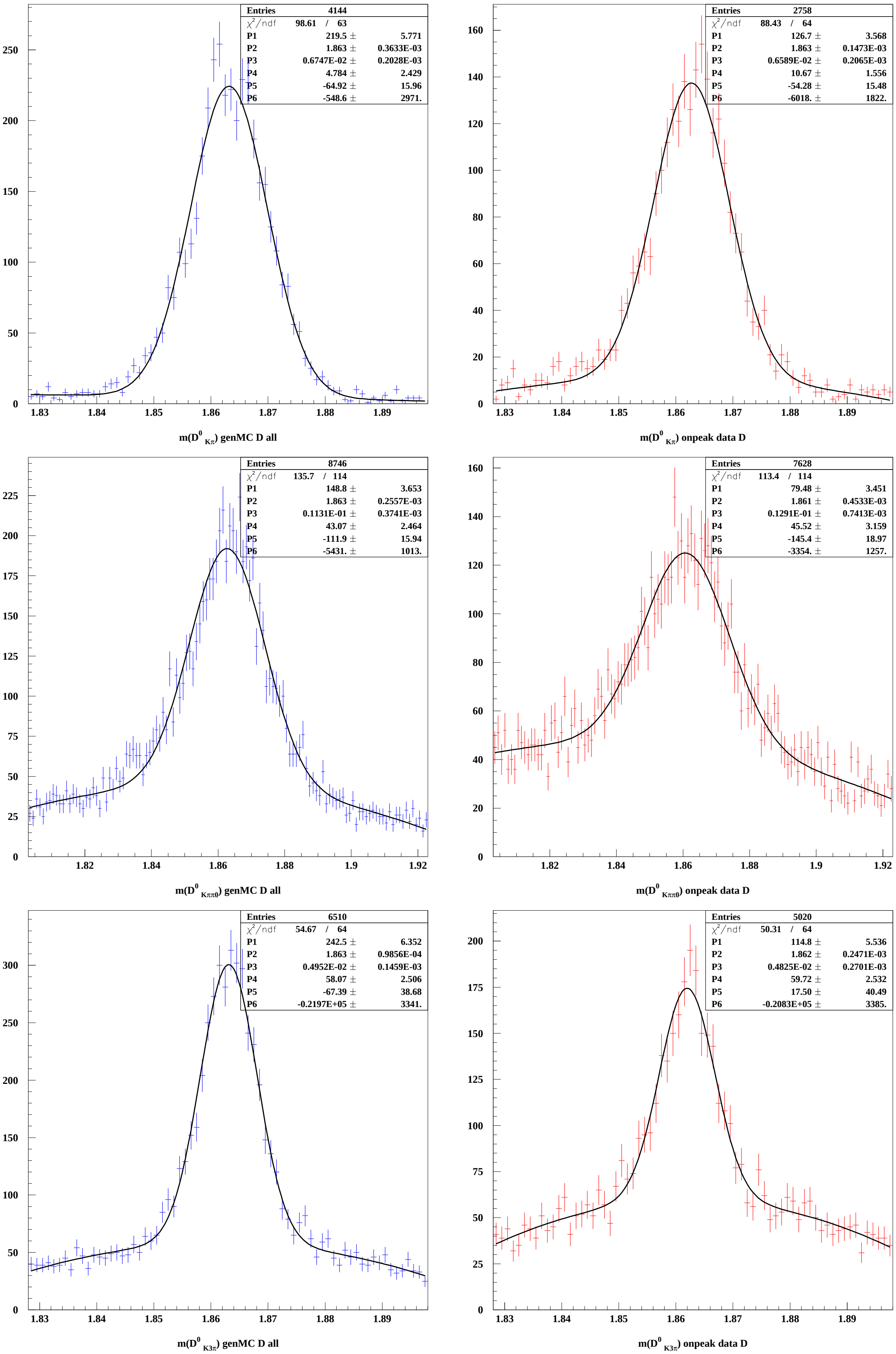


Figure 9.12: The D^0 mass distribution on generic MC (left) and data (right) for $K\pi$ (top), $K\pi\pi^0$ (middle), $K3\pi$ (bottom).

Mode	Sample	D^0 mass (MeV)	D^0 width (MeV)
$K\pi$	Generic MC	1863.3 ± 0.1	6.4 ± 0.2
	Data	1862.9 ± 0.2	6.5 ± 0.3
$K\pi\pi^0$	Generic MC	1862.5 ± 0.4	11.4 ± 0.5
	Data	1861.2 ± 0.6	12.5 ± 1.0
$K3\pi$	Generic MC	1863.1 ± 0.2	5.2 ± 0.2
	Data	1862.0 ± 0.3	6.2 ± 0.5

Table 9.9: Results of fitting the D^0 mass with a single Gaussian plus a second degree polynomial for the background.

9.6 D^* selection

The D^0 candidates are combined with slow $\pi^\pm$ candidates (GTVL with $70 < p_\pi^{\mathrm{CM}} < 450\,\mathrm{MeV}$) to form $D^{*\pm}$ candidates[73]. These candidates are accepted on the basis of the very discriminating mass difference $\delta m_{D^*} \equiv m_{D^*} - m_D$, which does not suffer from the D^0 mass resolution. No requirement on the vertex probability is imposed, as δm_{D^*} is a much better discriminating variable.

In principle the signal distribution of δm_{D^*} should be fitted with a double Gaussian, as the soft pions reaching the DCH will have a much better determined momentum than those reconstructed in the SVT alone. In the MC case, the use of truth matching allows one to fit with a double Gaussian.

This is not possible in data, where the background component mimics the broad Gaussian, and thus one is forced to use only a single Gaussian. In order for a comparison to be possible, the generic MC has also been fitted with a single Gaussian plus a Gaussian for the background component. Knowledge of the true distribution does not play a central role, as δm_{D^*} is not used in the subsequent likelihood fit (see Section 10), and corrections are made using control channels (see Section 11.1.5). The δm_{D^*} distribution for signal MC and data can be seen in Figure 9.13 and the fitting results on data in Table 9.10.

Mode	Sample	Central value of δm_{D^*} (MeV)	Width of δm_{D^*} (MeV)
$K\pi$	Generic MC	145.46 ± 0.01	0.34 ± 0.02
	Data	145.37 ± 0.02	0.32 ± 0.02
$K\pi\pi^0$	Generic MC	145.50 ± 0.02	0.43 ± 0.02
	Data	145.34 ± 0.03	0.46 ± 0.04
$K3\pi$	Generic MC	145.50 ± 0.02	0.38 ± 0.02
	Data	145.38 ± 0.03	0.40 ± 0.03

Table 9.10: Results of fitting δm_{D^*} on generic MC and data.

Candidates with a mass difference, δm_{D^*}, of $\pm 2\,\mathrm{MeV}$ ($\pm 1.5\,\mathrm{MeV}$) around the central value of the mass peak are accepted for the $K\pi$ and $K\pi\pi\pi$ ($K\pi\pi^0$) modes. The reason for the tighter cut in the mode involving a π^0 is, that δm_{D^*} is a better discriminator than the π^0 and D^0 masses, where the cut is very loose despite high backgrounds.

The signal MC distributions are very much the same for all three modes, as it should be, and the cuts applied have a high efficiency for signal. Once again a mass constraint is applied to the accepted D^* candidates.

[73] The $D^\pm\pi^0$ mode was also considered, but it contributes marginally and suffers from high backgrounds.

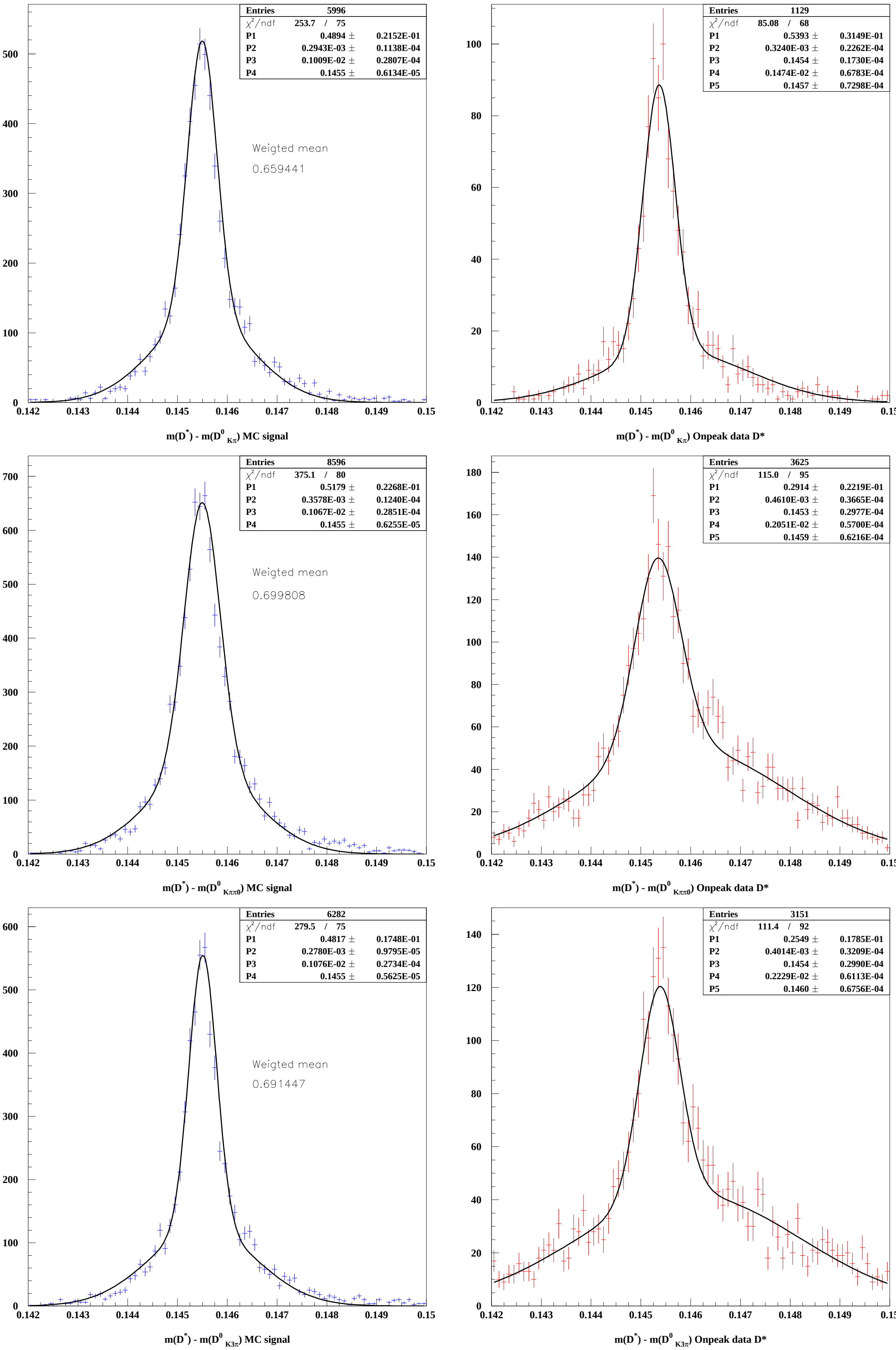


Figure 9.13: The δm_{D^*} mass distribution on signal MC (left) and data (right) for $K\pi$ (top), $K\pi\pi^0$ (middle), $K3\pi$ (bottom).

9.6.1 Summary of selection requirements

There exists no calculations of the $B^0 \to D^{(*)\pm} K^0 \pi^\mp$ branching fractions, only rough estimates [APS03], based on comparison with non-suppressed decays, of the order of 4×10^{-4}. It is the assumption used in the optimization of cuts. The cuts described above are summarized in Table 9.11. They were obtained while blind, and were fixed before unblinding the signal region.

Mode	$B^0 \to D^\pm K^0 \pi^\mp$		$B^0 \to D^{*\pm} K^0 \pi^\mp$	
Submode	$K\pi\pi$	$K\pi$	$K\pi\pi^0$	$K3\pi$
Bachelor track	NOT `VeryTightEl` NOT `TightKaon`	NOT `VeryTightEl` NOT `TightKaon`	NOT `VeryTightEl` NOT `TightKaon`	NOT `VeryTightEl` NOT `TightKaon`
m_{π^0}	–	–	$\pm 20\,\mathrm{MeV}$	–
$L_2/\sigma_{L_2}(K^0_S)$	> 4	> 4	> 4	> 4
$P_{\mathrm{vtx}}(K^0_S)$	> 0.001	> 0.001	> 0.001	> 0.001
$m_{K^0_S}$	$\pm 7\,\mathrm{MeV}$	$\pm 7\,\mathrm{MeV}$	$\pm 7\,\mathrm{MeV}$	$\pm 7\,\mathrm{MeV}$
w_{dalitz}	–	–	$> 1.4\%$	–
Kaon ID	`KMicroNotPion`	–	`KMicroNotPion`	`KMicroNotPion`
$P_{\mathrm{vtx}}(D^{\pm/0})$	Conv. vertex fit	Conv. vertex fit	> 0.001	Conv. vertex fit
$m_D^{\pm/0}$	$\pm 12\,\mathrm{MeV}$	$\pm 15\,\mathrm{MeV}$	$\pm 30\,\mathrm{MeV}$	$\pm 15\,\mathrm{MeV}$
δm_{D^*}	–	$\pm 2\,\mathrm{MeV}$	$\pm 1.5\,\mathrm{MeV}$	$\pm 2\,\mathrm{MeV}$
$P_{\mathrm{vtx}}(B^0)$	> 0.001	Conv. vertex fit	Conv. vertex fit	> 0.001

Table 9.11: Cuts applied to the sample before the B^0 reconstruction (except $P_{\mathrm{vtx}}(B^0)$ cut).

9.7 B^0 selection

The B^0 candidates are formed from $D^{*\pm}$, K^0_S and GTVL candidates, requiring the vertex probability to be greater than 0.1% for all submodes in order to reduce the combinatorics.

9.7.1 Multiple candidates

In some events, several signal candidates are present: 3.5% in signal MC (8.4% in data) for $B^0 \to D^\pm K^0 \pi^\mp$, and 9.2% in signal MC (15.9% in data) for $B^0 \to D^{*\pm} K^0 \pi^\mp$. When selecting all candidates which fulfill the reference cuts and have m_{ES}, ΔE, and $\mathcal{F}$ within $[5.24, 5.29] \times [-0.1, 0.1] \times [-3, 3]$, the number of candidates is distributed as described in Table 9.12.

Mode	Sample	Candidates in event						
		1	2	3	4	5	6	7
$D^\pm K^0 \pi^\mp$	Signal MC	24321	745	61	7	0	0	0
	Data	9330	586	93	17	3	1	0
$D^{*\pm} K^0 \pi^\mp$	Signal MC	6435	623	64	20	6	1	0
	Data	4849	742	112	48	8	9	0

Table 9.12: Distribution of number of candidates in events satisfying the selection criteria for signal MC and on-resonance data. The general overestimation of capabilities in MC causes its lower multiplicities. The $D^{*\pm} K^0 \pi^\mp$ mode has a larger fraction of events with multiple candidates due to the presence of a soft pion.

There are several solutions to the problem, which have been investigated. It was found, that the best *simple* method is to select one candidate on the basis of the $D^{0/\pm}$ mass, as it is not used in the final likelihood fit. One defines a chi-square:

$$\chi^2 \equiv \left(\frac{m_{D^{0/\pm}} - m^{\rm peak}_{D^{0/\pm}}}{\sigma_{m_{D^{0/\pm}}}} \right)^2, \tag{9.14}$$

and retains the candidate with the smallest value. Including ΔE in the χ^2, significantly improves the right candidate fraction, but biases the distribution of this variable. In fitting the ΔE distribution of the off-resonance data, the number of faked signal candidates increased by five events, when including ΔE in the χ^2. Considering this potentially large bias compared to the relative small gain, ΔE was omitted from the χ^2.
The simple χ^2 picks the right candidate in 53.5% of the multi-candidate cases of signal MC for $D^{\pm}K^0\pi^{\mp}$and 58.6% for $D^{*\pm}K^0\pi^{\mp}$, as can be seen in Table 9.13. Adding $m_{K^0_S}$ and/or δm_{D^*}, and the π^0 mass in the $D^0 \to K^-\pi^+\pi^0$ case does not increase the right selection rate. Since most of the cases have two candidates, this is close to a random choice. For candidates sharing the same D candidate, the choice is completely random.

Selection by χ^2	Right cand.	Wrong cand.	Right selection rate (%)
$D^{\pm}K^0\pi^{\mp}$	813	705	53.6 ± 1.3
$D^{*\pm}K^0\pi^{\mp}$	715	521	58.6 ± 1.4

Table 9.13: The right selection rate of the χ^2 choice on signal MC.

9.7.2 Self Cross Feed

Signal events that have been wrongly reconstructed (either due to wrong combination or inclusion of particles from the other B meson) may still fall in the signal region, though the distributions of discriminating variables for such events do not necessarily match that of the correctly reconstructed signal. This is termed self cross feed (SCF). It is most pronounced in final states containing a bachelor π^0, as these are the least constrain ed. If not corrected for, it is a potential source of bias (through the efficiency correction) when measuring branching fractions, as such events are rejected by the truth matching in the signal MC. In Table 9.14 are listed the SCF in each of the four final states considered in this analysis (including submodes) estimated by using the MC truth matching, which has an efficiency better than 0.5%.

Mode	Submode	Fraction of SCF
$B^0 \to D^{\pm}K^0\pi^{\mp}$	$D^{\pm} \to K^{\mp}\pi^{\pm}\pi^{\pm}$	$1.0 \pm 0.1\%$
$B^0 \to D^{*\pm}K^0\pi^{\mp}$	$D^0 \to K^-\pi^+$	$1.7 \pm 0.3\%$
$B^0 \to D^{*\pm}K^0\pi^{\mp}$	$D^0 \to K^-\pi^+\pi^0$	$3.9 \pm 0.5\%$
$B^0 \to D^{*\pm}K^0\pi^{\mp}$	$D^0 \to K^-\pi^+\pi^-\pi^+$	$2.5 \pm 0.4\%$

Table 9.14: Self Cross Feed (SCF) estimates from signal MC using truth matching.

As can be seen from the table, SCF is not very significant for the decay modes considered, as these do not contain bachelor π^0's nor possible interchanges of final state particles[74]. Furthermore, the SCF does not interchange the $D^{(*)\pm}$ and the $\pi^{\mp}$ charges (like the $\rho^{\pm}\pi^{\mp}$ channel), which can have a sizable impact on the time-dependent fit.

[74]Only in the $B^0 \to D^{*\pm}K^0\pi^{\mp}$ mode with the subsequent decay $D^{*+} \to D^0(K^-\pi^+\pi^+\pi^-)\pi^-$ is there potentially a chance of interchanging two pions.

9.7.3 Final samples

The efficiency of the various selection requirements on signal MC are listed in Table 9.15. The requirements have been applied consecutively in the order listed; the quoted efficiencies are with respect to the samples obtained from the pre-selection (see Section 9.4).

Mode	$B^0 \to D^{\pm}K^0\pi^{\mp}$		$B^0 \to D^{*\pm}K^0\pi^{\mp}$					
Submode			$K\pi$		$K\pi\pi^0$		$K3\pi$	
Listing	Singly	Multip.	Singly	Multip.	Singly	Multip.	Singly	Multip.
Raw efficiency		0.275		0.241		0.106		0.143
ΔE	0.992	0.992	0.989	0.989	0.993	0.993	0.992	0.992
m_{ES}	1.000	0.992	1.000	0.989	1.000	0.993	1.000	0.992
$P_{\mathrm{vtx}}(B^0)$	0.939	0.931	0.999	0.989	0.999	0.993	0.970	0.962
NOT VTElectron	0.994	0.925	0.993	0.982	0.995	0.988	0.996	0.958
NOT TKaon	0.986	0.912	0.985	0.967	0.988	0.976	0.991	0.949
δm_{D^*}	–	–	**0.873**	0.844	**0.552**	0.539	**0.825**	0.783
m_D	**0.890**	0.812	**0.929**	0.784	**0.905**	0.488	**0.888**	0.695
$P_{\mathrm{vtx}}(D^{\pm/0})$	0.989	0.803	0.996	0.781	0.999	0.488	1.000	0.695
m_{π^0} and w_{dalitz}	–	–	–	–	0.945	0.461	–	–
$m_{K^0_S}$	0.945	0.759	0.935	0.730	0.935	0.431	0.940	0.653
$P_{\mathrm{vtx}}(K^0_S)$	0.986	0.748	0.971	0.709	0.977	0.421	0.982	0.641
$L_2/\sigma_{L_2}(K^0_S)$	0.965	0.722	0.956	0.678	0.957	0.403	0.958	0.614
Mult. Cand.	0.974	0.703	0.948	0.643	0.913	0.368	0.932	0.572
Overall		0.193		0.155		0.039		0.082

Table 9.15: Singly and multiplicative efficiency of cuts. The raw efficiency is the efficiency after the pre-selection. The $D^{(*)}$ mass requirements (in bold) are the efficiency costly. The last line (Multiple candidates) is the efficiency in dealing with multiple candidates in an event, and thus the chance of choosing the wrong candidate in events with more than one.

The initial samples (see Table 9.1) are at this point much reduced, and the remaining number of events are listed in Table 9.16. The fractions listed in the multiplicative columns are relative to the sample from the preselection.

Origin	Sample	$B^0 \to D^{\pm}K^0\pi^{\mp}$		$B^0 \to D^{*\pm}K^0\pi^{\mp}$	
		Size (events)	Fraction (%)	Size (events)	Fraction (%)
Data	On-Resonance	10030	0.011	5773	0.0064
	Off-Resonance	3274	0.031	732	0.0069
	Sideband	11726	0.013	4307	0.0048
MC	Signal MC	24497	19.33	7150	6.93
	Generic $B^0\overline{B}^0$	6030	0.003	519	0.0003
	Generic B^+B^-	4679	0.003	2493	0.0016

Table 9.16: Reduced data and Monte Carlo samples. The efficiencies quoted are relative to the initial samples.

9.8 Section summary and conclusion

The data samples used for the analysis comprises of 81.8 fb^{-1} on-resonance and 9.7 fb^{-1} off-resonance data. The on-resonance data correspond to roughly 88 million $B\overline{B}$ pairs. In addition to this, simulated samples of both signal and generic decays are used along with similar control samples. Due to possible differences between simulation and reality, the latter samples are used whenever possible.

The two most discriminating variables for B candidates are m_{ES} and ΔE, the latter even more so for the decay modes considered, as the many tracks and mass constraints increase its precision. In addition topological and angular variables can be combined to discriminate against the dominating continuum background.

The event selection applies general requirements to the masses and vertex probabilities of the daughters. It also uses the transverse flight length of the K_S^0 and the decay distribution of $D^0 \to K^-\pi^+\pi^0$. Additional discriminating information exist, but its usage is not straight forward, and the gain it brings turns out to be limited.

The amount of signal mis-reconstruction (SCF) is very low, and the overall efficiencies are 19.3% for $B^0 \to D^{\pm}K^0\pi^{\mp}$, and 15.5%, 3.9%, and 8.2% for $B^0 \to D^{*\pm}K^0\pi^{\mp}$ in the three decay modes of the D^0, respectively.

10 Fitting

After the selection described in the previous section has been applied, each event contains exactly one B^0 candidate. Three variables still haven't been limited to the signal region, and are available for fitting; namely $m_{\rm ES}$, ΔE, and $\mathcal{F}$. They are required to be in the ranges $[5.24, 5.29] \times [-0.1, 0.1] \times [-3, 3]$ and $[5.20, 5.288] \times [-0.1, 0.1] \times [-3, 3]$ for $B^0 \to D^{\pm} K^0 \pi^{\mp}$ and $B^0 \to D^{*\pm} K^0 \pi^{\mp}$ respectively[75]. No requirement on the Dalitz position is made, thus (initially) including the entire Dalitz region in the fit (resonant and non-resonant).

10.1 Signal and background characterisation

Ideally one would like to fit the signal and background with a priori known PDFs leaving all parameters floating in the fit. However, due to lack of statistics, signal parameters have to be obtained from larger data samples (see Section 9.4.1). In principle the control sample could be included in the fit. This has not been done, as the signal sizes are rather small (5–10%) compared to the control samples, and therefore wouldn't add significantly to the precision of the signal parameters. Also, several different control channels are used for the various dimensions and components in the fit, and the increased number of free parameters and events would significantly increase the fitting time. To avoid biases due to differences between data and MC, data samples have been used whenever possible. The shape used for each PDF is discussed below and a summary can be found in Table 10.1.

10.1.1 Signal shape

The signal is described by a Gaussian distribution in $m_{\rm ES}$, two Gaussian distributions with common mean in ΔE and a Bifurcated Gaussian[76] in $\mathcal{F}$.
The parameters for the signal shape in $m_{\rm ES}$ and ΔE, obtained from the control channels $B^0 \to D^{(*)\mp} a_1^{\pm}$ (see Section 9.4.1), are the $m_{\rm ES}$ mean and width, the widths of the two Gaussians in ΔE, the fraction of each, and the common mean.
The three parameters for the Bifurcated Gaussian describing the $\mathcal{F}$ distribution are obtained from the $B^0 \to D^{*\mp}\pi^{\pm}$ control sample. To check that the $\mathcal{F}$ distribution is described well by the chosen PDF, and that there are no tails unaccounted for, the number of events beyond the $\pm 3\sigma$ limits was determined on the $B^0 \to D^{*\mp}\pi^{\pm}$ sample. The result was 6 events below and 4 events above the central distribution, which is in good agreement with 4.2 events expected on each side for the $B^0 \to D^{\pm} K^0 \pi^{\mp}$ mode.

10.1.2 Continuum background

Continuum events, especially $c\bar{c}$ events, are the dominant source of background. This background is described by an Argus-function [ARGUS87b] in $m_{\rm ES}$, a polynomial of first degree in ΔE and a Bifurcated Gaussian in $\mathcal{F}$.

The off-resonance data give a handle on the continuum background, and it is fitted in order to constrain the shape. The continuum Fisher shape is fixed to that obtained on the off-resonance data, while the $m_{\rm ES}$ and ΔE parameters are only used as initial values in the likelihood fit, and left floating in the fit. Repeating the exercise of checking the tails (see subsection 10.1.1 above) on the off-resonance data gave no events beyond the 3σ limits, where 1.1 was expected, thus again no long tails were detected.

It should be noted, that considering the $\mathcal{F}$ distribution of the sideband (either in $m_{\rm ES}$ or ΔE) does not give any good handle on the continuum background, as a significant fraction

[75] Before applying this requirement, the off-resonance $m_{\rm ES}$ is corrected (using the beam energy), such that it shares the same distribution as the on-resonance continuum.
[76] A Gaussian with different widths above and below the mean, thus four parameters in total.

of the side band events originates from combinatorial $B\overline{B}$ events, which have very nearly the same $\mathcal{F}$ distribution as signal.

10.1.3 Combinatorial $B\overline{B}$ background

Combinatorial $B\overline{B}$ events also contribute significantly to the background, and this source is evaluated using generic $B^0\overline{B}^0$ and B^+B^- Monte Carlo. The parametrization of combinatorial $B\overline{B}$ background is the same as for the continuum, but with a new set of parameters for all variables except the endpoint of the Argus function, which is common and kept fixed.
To minimize correlation with continuum events when performing the final fit, the Fisher shape has been fixed to that obtained by fitting generic MC, which has been found to describe data well in previous studies (see Section 9.3). All the other parameters are left floating in the likelihood fit, using as initial values the results from fitting the generic $B\overline{B}$ MC sample.

10.1.4 Peaking $B\overline{B}$ background

In addition to the combinatorial backgrounds, two sorts of peaking backgrounds are possible. By peaking is generally meant a tendency to have peaking features in one or more variables, where the random background does not have any such features.

The first type of peaking background comes from decays where a (low momentum) particle from a B decay has been missed (e.g. final state $D^{\pm}K^0_S\pi^{\mp}\pi^0$) or simply exchanged with another belonging to the rest of the event (e.g. final state $D^{\pm}K^0_S X$, where X has been exchanged with a random pion). Such events peak in m_{ES}, but not necessarily with the same shape as the signal. However, such events do not peak in ΔE, as a missing particle (typically a π^0) will shift ΔE by at least the energy of this particle, thus at least the mass of a pion. In the fit, only the interval $[-100, 100]\,\mathrm{MeV}$ is considered, and only a tail due to resolution and exchange with ROE particles will be seen.
In the ΔE sideband of $B^0 \to D^{\pm}K^0\pi^{\mp}$ on-resonance data the peaking background is visible (see Fig. 10.1a), and the mean $(5.278 \pm 0.001\,\mathrm{GeV})$ and width $(3.72 \pm 0.50\,\mathrm{MeV})$ of the peak are in accordance with the values obtained from generic $B\overline{B}$ MC (see Fig. 10.1b), which are used in the likelihood fit (see Table 10.2). Due to the larger combinatorial background, it is hard to identify the exact source of peaking background events. As illustrated in Fig. 10.1b, one can by subtracting the sideband in m_{ES} (blue/dark grey region) from the peaking region (red/light grey region) check, that this background does not peak in ΔE (see Fig. 10.1c).

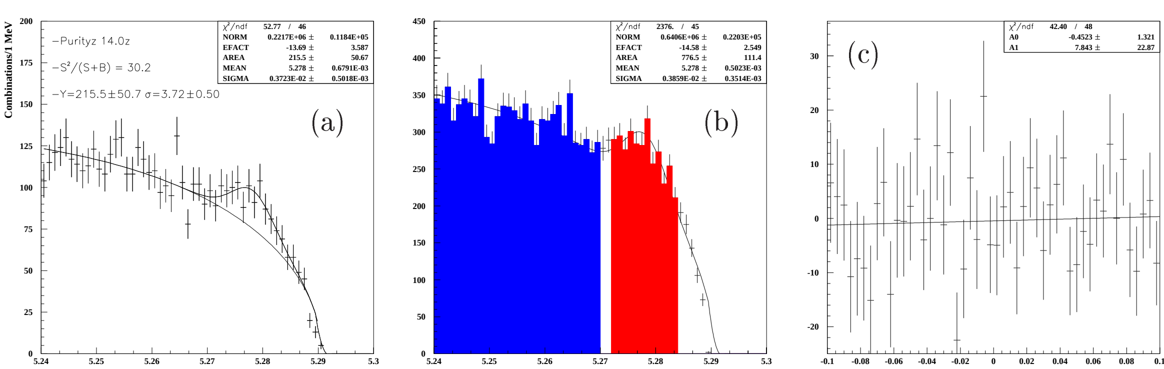


Figure 10.1: (a) Position and width in m_{ES} of peaking background measured in the ΔE sideband of the data. (b) Peaking background and (c) the ΔE distribution of the peaking background in generic $B\overline{B}$ MC. The latter is obtained from the peaking region in m_{ES} (red/light grey region) subtracted the sideband in m_{ES} (blue/dark grey region). None of the background, that peaks in m_{ES}, also peaks in ΔE (see text).

As parametrization for the peaking background in ΔE, polynomials of zeroth, first, second, and third degree along with an exponential and a constant plus an exponential have been tried. Since the exponential is the simplest function that describes the data well, this was chosen (according to Occam's Razor [Occ30]). In addition, it is both liable and robust. Hence, this component is parametrised by a Gaussian in $m_{\rm ES}$, an exponential with coefficient $\alpha_{\rm peak}$ in ΔE, and the *same* Bifurcated Gaussian as for the combinatorial $B\overline{B}$ background for the Fisher variable, $\mathcal{F}$. The size of this background is thus evaluated from the fit.

The second possible type of peaking background originates from B^0 decays with the same particles in the final state (e.g. $K^*(K\pi)K^*(K^0_S\pi)\pi$ or $D^0(K\pi)K^0_S\pi\pi$), which can be reshuffled to mimic a $B^0 \to D^{(*)\pm}K^0\pi^\mp$ candidate. It is thus completely degenerate with the signal shape in terms of variables used in the fit (i.e. it peaks in $m_{\rm ES}$ and ΔE and has the shape of $B\overline{B}$ events). This background is denoted *degenerate peaking background* (also known as *double peaking background*, refering to $m_{\rm ES}$ and ΔE). Its size is evaluated a posteriori by using the m_D and $m_{K^0_S}$ sidebands. This component, being contained in the signal yields, cannot be corrected for in the fit, but must be subtracted afterwards, if present.
An alternative method to incorporate these backgrounds would be to include the D and K^0_S masses and their sidebands in the fit. In this manner, the degenerate peaking background would no longer be degenerate, but could be determined from the fit. This approach was attempted, but due to the enlarged sample and the larger dimensionality of the fit, the time required for the fit to converge increased beyond the acceptable and the approach was abandonned.

10.1.5 PDF and parameter summary

A summary of the above described PDFs are presented in Table 10.1.

Component	Signal	Continuum	$B\overline{B}$	$B\overline{B}$ peak	Total
$m_{\rm ES}$	G_1	Argus$_1$	Argus$_2$	G_2	7
ΔE	GG	$P1_1$	$P1_2$	Exp	7
$\mathcal{F}$	BG_1	BG_2	BG_3	BG_3	9
N parameters	9	6	6	6	23

Table 10.1: PDFs used for signal, continuum, and $B\overline{B}$ background (non-peaking and non-degenerate peaking). Four parameters are common among components, which has to included when summing the bottom line of the table. The abbreviations are G = Gaussian, GG = double Gaussian, $P1$ = polynomial of first degree, Exp = exponential, and BG = Bifurcated Gaussian. The color code is Black: Fixed in fit, Magenta: fixed from MC, and Green: Free.

In addition to the 4 yields of interest, the number of parameters describing the shape of the PDFs are 9 for the signal, and 14 for the background, giving a total of 27 parameters. The all signal and 9 background parameters are fixed in the fit, giving a total of 9 free fit parameters.

10.2 Likelihood fit

The events are fitted with an extended unbinned maximum likelihood fit containing the variables $m_{\rm ES}$, ΔE, and $\mathcal{F}$. A probability product, $P_j(m_{{\rm ES},i}, \Delta E_i, \mathcal{F}_i) = P_j(m_{{\rm ES},i}) \cdot P_j(\Delta E_i) \cdot P_j(\mathcal{F}_i)$, is assigned to each event, i, and an unbinned likelihood is constructed:

$$\ln \mathcal{L} = \sum_i \ln \left(\sum_{j={\rm comp.}} N_j P_j(m_{{\rm ES},i}, \Delta E_i, \mathcal{F}_i) \right) - \sum_{j={\rm comp.}} N_j, \tag{10.1}$$

where the sum j is over the four components of the sample (signal, continuum, combinatorial and peaking background), and N_j is the number of events in each component.

The fitting is done using RooFit [KV01] based on ROOT [Bru95]. To test the fitting routine and obtain initial and possibly fixed values, the off-resonance data, the control samples ($B^0 \to D^{(*)\pm}a_1^{\mp}$ and $B^0 \to D^{*\pm}\pi^{\mp}$) and the generic MC were used. For the signal MC, each dimension in the fit (m_{ES}, ΔE, and $\mathcal{F}$) were fitted separately, while the fit was performed in all relevant dimensions simultaneously for the other samples. The fit converges nicely and the results can be seen in Figure 10.2 for MC samples (signal and generic $B\overline{B}$) and in Figure 10.3 for data samples (off-resonance data and control samples).

In general the PDFs describe the data well. The only exception to this is the $\mathcal{F}$ signal distribution, where the PDF in all three signal cases (signal MC, generic $B\overline{B}$, and $B^0 \to D^- a_1^+$) falls slightly below the actual distribution at its center: The Bifurcated Gaussian used to parametrize this distribution simply does not have enough degrees of freedom to incorporate the features of the distribution.

Several other distributions were tried, and the best fit was obtained with a Bifurcated Gaussian plus a Gaussian. However, this PDF has twice the number of parameters, and as the shape is obtained from the control sample $B^0 \to D^{*-}\pi^+$ (see Fig. 10.3), some would have been poorly determined. In addition, the systematic from changing this PDF is very small (see Section 11.1.7), it was decided to keep the simplest PDF, namely the Bifurcated Gaussian.

To test the fitting routine further, the fit was applied to a test sample consisting of background events (from $m_{D^\pm}$ sideband) and 150 signal MC events: 156 ± 17 events were obtained, thus in accordance within statistical errors.

10.2.1 Result for the $\mathbf{B^0 \to D^{\pm}K^0\pi^{\mp}}$ mode

After having performed the fit on MC, off-resonance, and control samples, it was applied to the $B^0 \to D^{\pm}K^0\pi^{\mp}$ data sample. Once again the fit converged nicely, with values for the background parameters close to the ones obtained from the initial fits. The results of the fit can be seen in Fig. 10.4 (left) and the obtained parameters are listed in Table 10.2.

In addition to the entire Dalitz plot, the $K^*(892)$ resonance region (defined as a mass range centered on the $K^*(892)$ mass, of half width $2\Gamma = 100\,\mathrm{MeV}$) was fitted separately. When applying this cut, the number of background events drop by an order of magnitude, and the subsequent fit cannot determine the same amount of parameters. For this reason the exponential coefficient of the PDF in ΔE for the peaking component was fixed to the value obtained when fitting the entire sample. The fit result is shown in Fig. 10.4 (right) and the parameters are listed in Table 10.3.

The correlation matrix gives the *linear* correlation among the variables. The correlations for the $B^0 \to D^{\pm}K^0\pi^{\mp}$ fit (and the fits in general) are not large, and the signal yield is the least correlated parameter (see Table 10.4).

10.2.2 Result for the $\mathbf{B^0 \to D^{*\pm}K^0\pi^{\mp}}$ mode

Despite containing a π^0 in the final state, the $D^0 \to K^-\pi^+\pi^0$ mode does not have a much larger uncertainty in ΔE, which is due to the mass constrain on the D^0 and the $D^{*\pm}$. From fitting the signal MC with a single Gaussian the width was determined to be $\sigma_{\Delta E}(K^-\pi^+\pi^0) = (12.5 \pm 0.2)\,\mathrm{MeV}$, compared to $\sigma_{\Delta E}(K^-\pi^+) = (11.8 \pm 0.2)\,\mathrm{MeV}$ and $\sigma_{\Delta E}(K^-\pi^+\pi^+\pi^-) = (11.6 \pm 0.2)\,\mathrm{MeV}$.

To include this small effect and possibly other differences (e.g. efficiency corrections), the three D^0 modes were fitted separately in the $B^0 \to D^{*\pm}K^0\pi^{\mp}$ fit, each with the fit parameters obtained from their respective control channels (i.e. the $B^0 \to D^- a_1^+$ channel with the matching decay mode of the D^0). The result of the fit can be seen in Fig. 10.5 (left), where the three D^0 modes have been added together, and the obtained parameters are listed in Table 10.5.

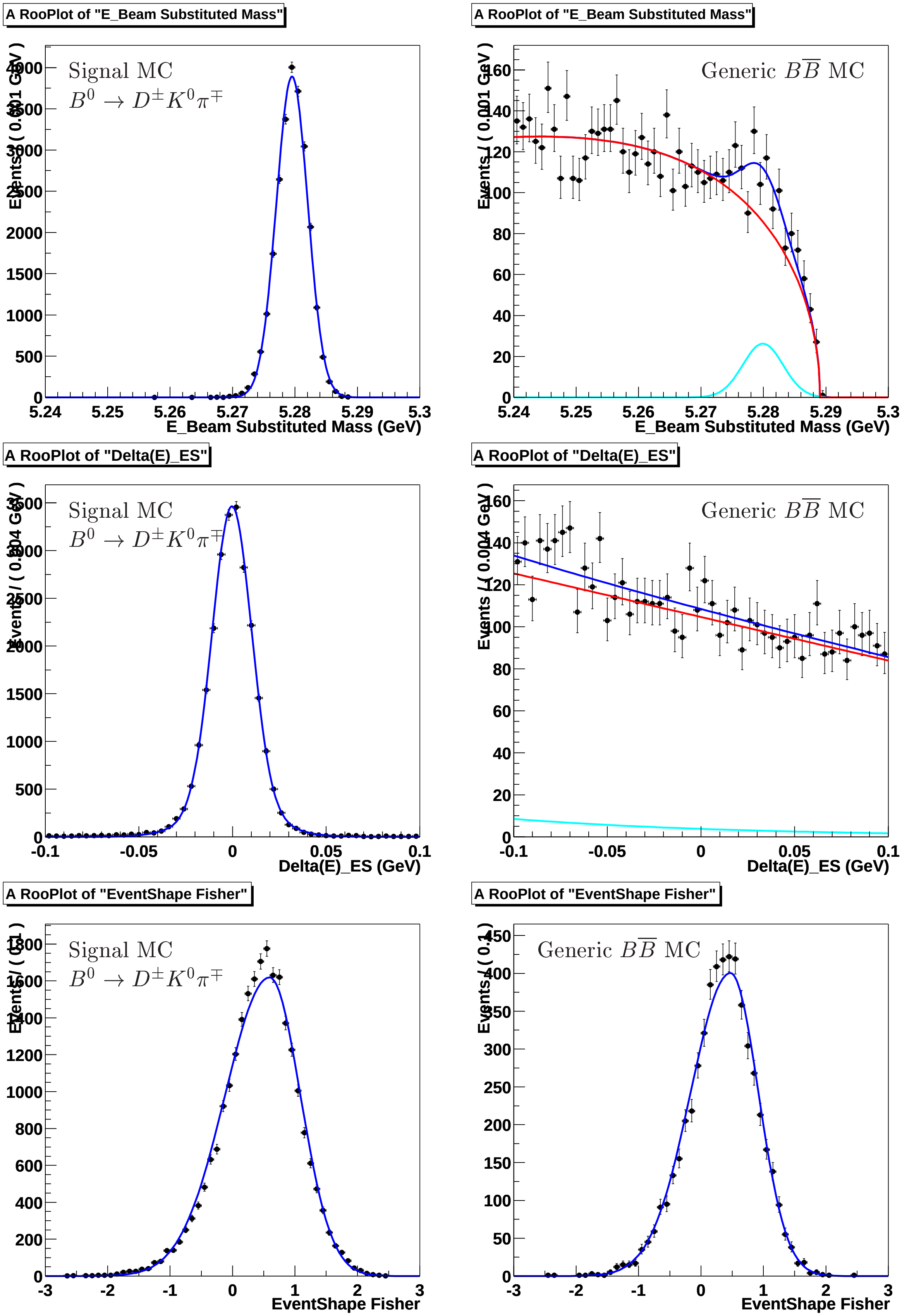


Figure 10.2: Projections of the m_{ES} (top), ΔE (middle) and $\mathcal{F}$ (bottom) distributions with the unbinned maximum likelihood fit PDF plotted on top for signal MC (left) and generic $B\overline{B}$ MC (right) for the decay $B^0 \to D^{\pm}K^0\pi^{\mp}$ (see text). Key for curves: Green: combinatorial $B\overline{B}$, cyan: peaking $B\overline{B}$ and black: signal.

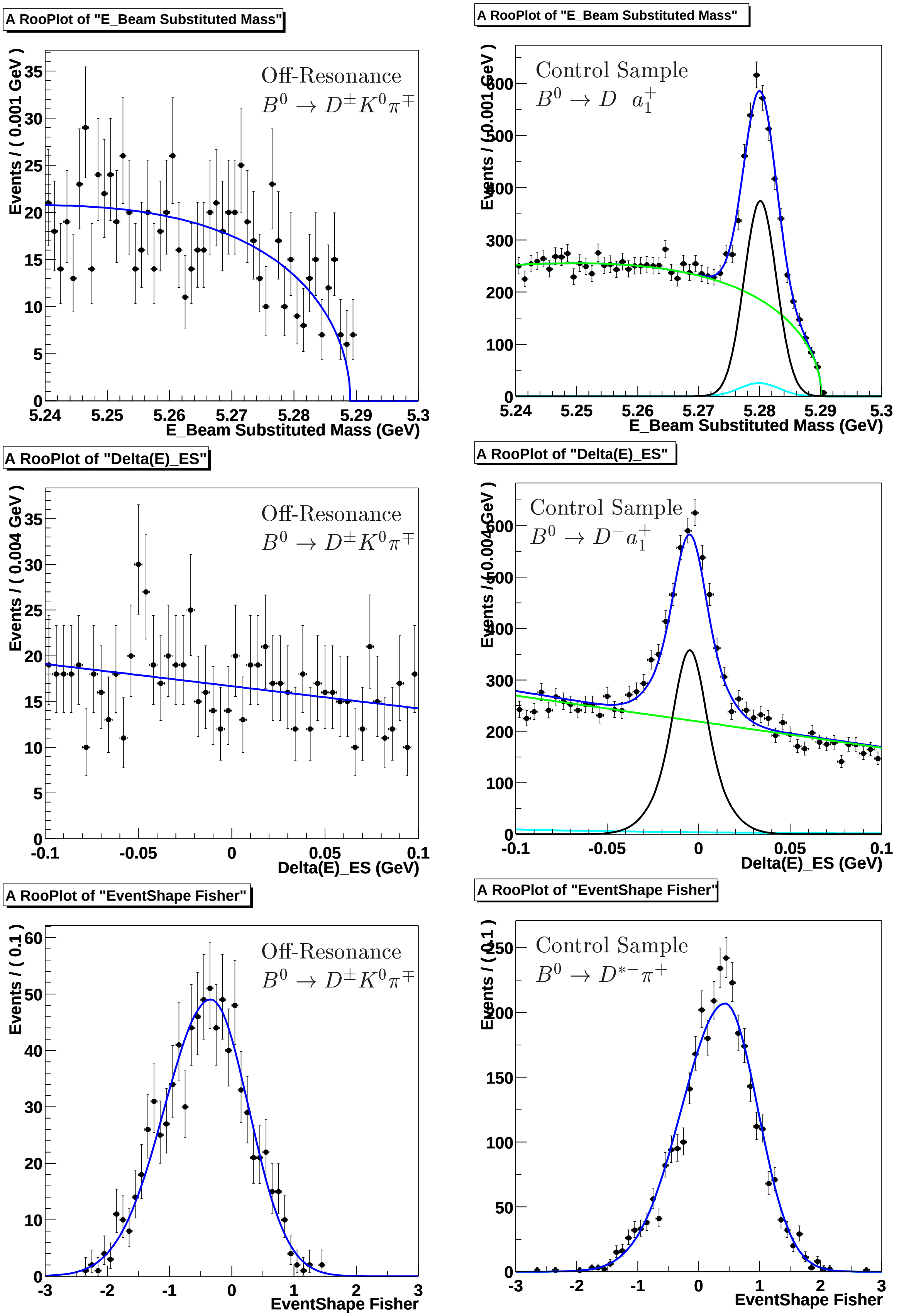


Figure 10.3: Projections of the m_{ES} (top), ΔE (middle) and $\mathcal{F}$ (bottom) distributions with the unbinned maximum likelihood fit PDF plotted on top for off-resonance data (left) and the control samples $B^0 \to D^{\pm}a_1^{\mp}$ for m_{ES} and ΔE and $B^0 \to D^{*\pm}\pi^{\mp}$ for $\mathcal{F}$ (right) for the decay $B^0 \to D^{\pm}K^0\pi^{\mp}$ (see text). Key for curves: Green: continuum and combinatorial $B\overline{B}$, cyan: peaking $B\overline{B}$ and black: signal.

Component	Parameter	Fixed	Value	Unit	Source
Signal	$m_{\rm ES}$ Mean	Yes	5.2801 ± 0.0001	GeV	$D^{\pm}a_1$
	$m_{\rm ES}$ Width	Yes	2.612 ± 0.060	MeV	$D^{\pm}a_1$
	ΔE Mean	Yes	-4.89 ± 0.30	MeV	$D^{\pm}a_1$
	ΔE Width$_1$	Yes	8.18 ± 1.18	MeV	$D^{\pm}a_1$
	ΔE Width$_2$	Yes	15.55 ± 2.11	MeV	$D^{\pm}a_1$
	ΔE $f_{\rm G1}$	Yes	0.443 ± 0.186	–	$D^{\pm}a_1$
	$\mathcal{F}$ Mean	Yes	0.432 ± 0.025	–	$D^{*\pm}\pi$
	$\mathcal{F}$ σ_{Left}	Yes	0.716 ± 0.017	–	$D^{*\pm}\pi$
	$\mathcal{F}$ σ_{Right}	Yes	0.539 ± 0.016	–	$D^{*\pm}\pi$
Continuum	$m_{\rm ES}$ Shape	No	-19.2 ± 3.3	–	
	$m_{\rm ES}$ Endpoint	Yes	5.2903 ± 0.0001	GeV	$D^{\pm}a_1$
	ΔE Slope	No	-1.74 ± 0.29	ev/ GeV	
	$\mathcal{F}$ Mean	Yes	-0.339 ± 0.066	–	Off-Resonance
	$\mathcal{F}$ σ_{Left}	Yes	0.744 ± 0.043	–	Off-Resonance
	$\mathcal{F}$ σ_{Right}	Yes	0.613 ± 0.041	–	Off-Resonance
$B\overline{B}$	$m_{\rm ES}$ Shape	No	-13.3 ± 9.3	–	
	ΔE Slope	No	-0.46 ± 0.72	ev/ GeV	
	$\mathcal{F}$ Mean	Yes	0.468 ± 0.018	–	$B\overline{B}$ Generic MC
	$\mathcal{F}$ σ_{Left}	Yes	0.640 ± 0.012	–	$B\overline{B}$ Generic MC
	$\mathcal{F}$ σ_{Right}	Yes	0.454 ± 0.011	–	$B\overline{B}$ Generic MC
$B\overline{B}$ peak	$m_{\rm ES}$ Mean	Yes	5.2801 ± 0.0008	GeV	$B\overline{B}$ Generic MC
	$m_{\rm ES}$ Width	Yes	3.95 ± 0.56	MeV	$B\overline{B}$ Generic MC
	ΔE Exp. Coef.	No	-8.6 ± 6.3	GeV^{-1}	

Table 10.2: Parameter values from the $B^0 \to D^{\pm}K^0\pi^{\mp}$ fit. The errors on the fixed parameters are the ones obtained from the fit on the off-resonance/control/MC sample. They are varied by $\pm 1\sigma$ to account for systematic errors.

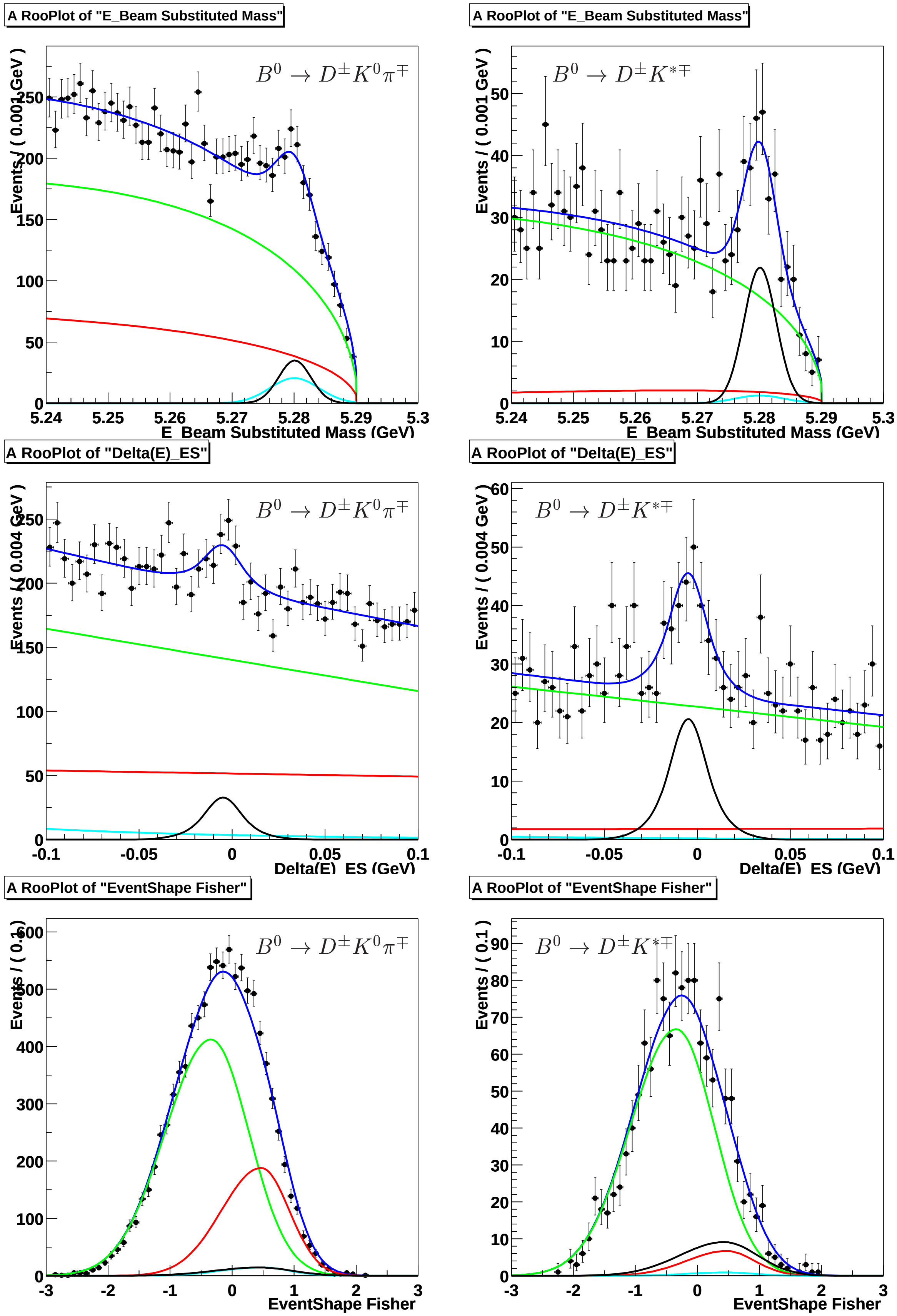


Figure 10.4: Projections of the m_{ES}, ΔE and $\mathcal{F}$ distributions with the unbinned maximum likelihood fit PDF plotted on top for the decays $B^0 \to D^{\pm}K^0\pi^{\mp}$ (left) and $B^0 \to D^{\pm}K(892)^{*\mp}(K^0\pi^{\mp})$ (right). Key for curves: Red: continuum, Green: comb. $B\overline{B}$, cyan: peaking $B\overline{B}$ and black: signal.

Component	Parameter	Fixed	Value	Unit	Source
Signal	$m_{\rm ES}$ Mean	Yes	5.2801 ± 0.0001	GeV	$D^{\pm}a_1$
	$m_{\rm ES}$ Width	Yes	2.612 ± 0.060	MeV	$D^{\pm}a_1$
	ΔE Mean	Yes	-4.89 ± 0.30	MeV	$D^{\pm}a_1$
	ΔE Width$_1$	Yes	8.18 ± 1.18	MeV	$D^{\pm}a_1$
	ΔE Width$_2$	Yes	15.55 ± 2.11	MeV	$D^{\pm}a_1$
	ΔE $f_{\rm G1}$	Yes	0.443 ± 0.186	–	$D^{\pm}a_1$
	$\mathcal{F}$ Mean	Yes	0.432 ± 0.025	–	$D^{*\pm}\pi$
	$\mathcal{F}$ Left Width	Yes	0.716 ± 0.017	–	$D^{*\pm}\pi$
	$\mathcal{F}$ Right Width	Yes	0.539 ± 0.016	–	$D^{*\pm}\pi$
Continuum	$m_{\rm ES}$ Shape	No	-15.8 ± 7.9	–	
	$m_{\rm ES}$ Endpoint	Yes	5.2903 ± 0.0001	GeV	$D^{\pm}a_1$
	ΔE Slope	No	-1.52 ± 0.67	ev/ GeV	
	$\mathcal{F}$ Mean	Yes	-0.339 ± 0.066	–	Off-Resonance
	$\mathcal{F}$ Left Width	Yes	0.744 ± 0.043	–	Off-Resonance
	$\mathcal{F}$ Right Width	Yes	0.613 ± 0.041	–	Off-Resonance
$B\overline{B}$	$m_{\rm ES}$ Shape	No	-55.4 ± 79.0	–	
	ΔE Slope	No	-0.29 ± 5.47	ev/ GeV	
	$\mathcal{F}$ Mean	Yes	0.468 ± 0.018	–	$B\overline{B}$ Generic MC
	$\mathcal{F}$ Left Width	Yes	0.640 ± 0.012	–	$B\overline{B}$ Generic MC
	$\mathcal{F}$ Right Width	Yes	0.454 ± 0.011	–	$B\overline{B}$ Generic MC
$B\overline{B}$ peak	$m_{\rm ES}$ Mean	Yes	5.2801 ± 0.0008	GeV	$B\overline{B}$ generic MC
	$m_{\rm ES}$ Width	Yes	3.95 ± 0.56	MeV	$B\overline{B}$ generic MC
	ΔE Exp. Coef.	Yes	-8.6 ± 6.3	GeV^{-1}	$B\overline{B}$ generic MC

Table 10.3: Parameter values from the $B^0 \to D^{\pm}K^{*\mp}$ fit. The errors on the fixed parameters are the ones obtained from the fit on the off-resonance/control/MC sample. They are varied by $\pm 1\sigma$ to account for systematic errors.

Param.	All	N_{BB}	N_{Peak}	N_{Cont}	N_{sig}	ξ_{BB}	ξ_{Cont}	α_{Peak}	$P1_{BB}$	$P1_{Cont}$
N_{BB}	0.707	1.000	0.474	-0.507	0.097	-0.335	-0.022	-0.296	0.023	-0.013
N_{Peak}	0.824	0.474	1.000	0.004	0.375	-0.587	-0.046	-0.635	0.151	-0.018
N_{Cont}	0.585	-0.507	0.004	1.000	-0.026	0.016	-0.014	0.001	0.037	0.002
N_{sig}	0.455	0.097	-0.375	-0.026	1.000	-0.133	0.042	-0.350	0.121	0.003
ξ_{BB}	0.743	-0.335	-0.587	0.016	-0.133	1.000	-0.405	0.409	-0.145	0.053
ξ_{Cont}	0.553	-0.022	-0.046	-0.014	0.042	-0.405	1.000	-0.018	0.065	-0.039
α_{Peak}	0.744	-0.296	-0.635	0.001	-0.350	0.409	-0.018	1.000	-0.400	-0.006
$P1_{BB}$	0.681	0.023	0.151	0.037	0.121	-0.145	0.065	-0.400	1.000	-0.524
$P1_{Cont}$	0.587	-0.013	-0.018	0.002	0.003	0.053	-0.039	-0.006	-0.524	1.000

Table 10.4: Correlations among fitted variables for the $B^0 \to D^{\pm}K^0\pi^{\mp}$ fit. The leftmost column contains the total correlation, the signal yield having the smallest correlation, mainly with the peaking background, as it resembles signal the most.

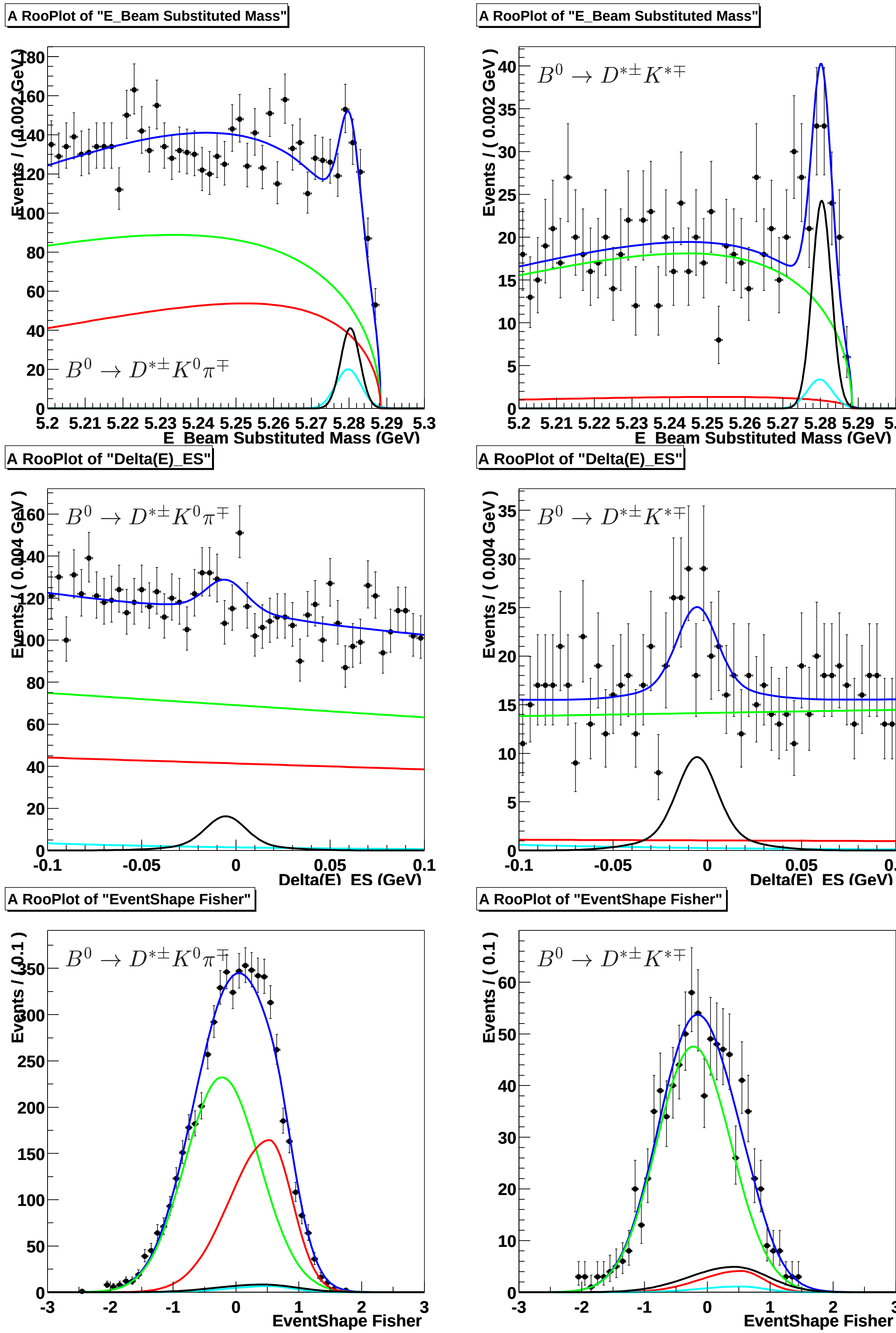


Figure 10.5: Projections of the m_{ES}, ΔE and $\mathcal{F}$ distributions with the unbinned maximum likelihood fit PDF plotted on top for the decays $D^{*\pm}K^0\pi^\mp$(left) and $B^0 \to D^{*\pm}K(892)^{*\mp}(K^0\pi^\mp)$ (right). Key for curves: Red: continuum, Green: comb. $B\overline{B}$, cyan: peaking $B\overline{B}$ and black: signal.

Comp.	Parameter	Fixed	Value $K\pi$	Value $K\pi\pi^0$	Value $K3\pi$	Source
Signal	$m_{\rm ES}$ Mean	Yes	5.2803 ± 0.0001	5.2804 ± 0.0001	5.2803 ± 0.0001	$D^{*\pm}a_1$
	$m_{\rm ES}$ Width	Yes	2.547 ± 0.074	2.623 ± 0.103	2.510 ± 0.090	$D^{*\pm}a_1$
	ΔE Mean	Yes	-4.26 ± 0.47	-6.34 ± 0.66	-6.45 ± 0.57	$D^{*\pm}a_1$
	ΔE Width$_1$	Yes	8.48 ± 0.81	11.42 ± 0.87	11.59 ± 0.67	$D^{*\pm}a_1$
	ΔE Width$_2$	Yes	21.07 ± 2.44	32.69 ± 6.40	40.08 ± 13.12	$D^{*\pm}a_1$
	ΔE $f_{\rm G1}$	Yes	0.517 ± 0.085	0.686 ± 0.078	0.844 ± 0.063	$D^{*\pm}a_1$
	$\mathcal{F}$ Mean	Yes		0.432 ± 0.025 (common for all modes)		$D^{*\pm}\pi$
	$\mathcal{F}$ σ_{Left}	Yes		0.716 ± 0.017 (common for all modes)		$D^{*\pm}\pi$
	$\mathcal{F}$ σ_{Right}	Yes		0.539 ± 0.016 (common for all modes)		$D^{*\pm}\pi$
Cont.	$m_{\rm ES}$ Shape	No	-38.5 ± 10.2	-17.1 ± 5.3	-25.4 ± 4.9	
	$m_{\rm ES}$ Endpoint	Yes	5.2883 ± 0.0002	5.2884 ± 0.0003	5.2884 ± 0.0003	$D^{*\pm}a_1$
	ΔE Slope	No	-0.22 ± 1.42	-1.12 ± 0.73	-0.51 ± 0.70	
	$\mathcal{F}$ Mean	Yes		-0.224 ± 0.081 (common for all modes)		Off-Res.
	$\mathcal{F}$ σ_{Left}	Yes		0.586 ± 0.051 (common for all modes)		Off-Res.
	$\mathcal{F}$ σ_{Right}	Yes		0.601 ± 0.051 (common for all modes)		Off-Res.
$B\overline{B}$	$m_{\rm ES}$ Shape	No	-30.3 ± 14.0	-49.6 ± 8.7	-27.5 ± 9.0	
	ΔE Slope	No	1.01 ± 1.75	-0.54 ± 1.14	-2.14 ± 1.19	
	$\mathcal{F}$ Mean	Yes		0.532 ± 0.019 (common for all modes)		Gen. MC
	$\mathcal{F}$ σ_{Left}	Yes		0.637 ± 0.013 (common for all modes)		Gen. MC
	$\mathcal{F}$ σ_{Right}	Yes		0.367 ± 0.012 (common for all modes)		Gen. MC
Peak	$m_{\rm ES}$ Mean	Yes		5.2798 ± 0.0003 (common for all modes)		Gen. MC
	$m_{\rm ES}$ Width	Yes		3.33 ± 0.45 (common for all modes)		Gen. MC
	ΔE Exp. Coef.	No	-4.5 ± 6.2	-12.9 ± 27.1	-6.1 ± 7.9	

Table 10.5: Parameter values from the $B^0 \to D^{*\pm}K^0\pi^{\mp}$ fit. The errors on the fixed parameters are the ones obtained from the fit on the off-resonance/control/MC sample. They are varied by $\pm1\sigma$ to account for systematic errors.

Once again the fit is repeated in the K^{*+} resonant region only, and for the same reason as when fitting $B^0 \to D^{\pm}K^{*\mp}$, some background parameters are fixed. The resulting distributions and fits can be seen in Fig. 10.5 (right) and the fitted parameters can be found in Table 10.6.

Comp.	Parameter	Fixed	Value $K\pi$	Value $K\pi\pi^0$	Value $K3\pi$	Source
Signal	$m_{\rm ES}$ Mean	Yes	5.2803 ± 0.0001	5.2804 ± 0.0001	5.2803 ± 0.0001	$D^{*\pm}a_1$
	$m_{\rm ES}$ Width	Yes	2.547 ± 0.074	2.623 ± 0.103	2.510 ± 0.090	$D^{*\pm}a_1$
	ΔE Mean	Yes	-4.26 ± 0.47	-6.34 ± 0.66	-6.45 ± 0.57	$D^{*\pm}a_1$
	ΔE Width$_1$	Yes	8.48 ± 0.81	11.42 ± 0.87	11.59 ± 0.67	$D^{*\pm}a_1$
	ΔE Width$_2$	Yes	21.07 ± 2.44	32.69 ± 6.40	40.08 ± 13.12	$D^{*\pm}a_1$
	ΔE $f_{\rm G1}$	Yes	0.517 ± 0.085	0.686 ± 0.078	0.844 ± 0.063	$D^{*\pm}a_1$
	$\mathcal{F}$ Mean	Yes	0.432 ± 0.025 (common for all modes)			$D^{*\pm}\pi$
	$\mathcal{F}$ σ_{Left}	Yes	0.716 ± 0.017 (common for all modes)			$D^{*\pm}\pi$
	$\mathcal{F}$ σ_{Right}	Yes	0.539 ± 0.016 (common for all modes)			$D^{*\pm}\pi$
Cont.	$m_{\rm ES}$ Shape	No	-38.7 ± 12.5	-22.6 ± 7.6	-32.4 ± 6.9	
	$m_{\rm ES}$ Endpoint	Yes	5.2883 ± 0.0002	5.2884 ± 0.0003	5.2884 ± 0.0003	$D^{*\pm}a_1$
	ΔE Slope	No	0.27 ± 1.82	-0.78 ± 1.05	1.18 ± 1.01	
	$\mathcal{F}$ Mean	Yes	-0.224 ± 0.081 (common for all modes)			Off-Res.
	$\mathcal{F}$ σ_{Left}	Yes	0.586 ± 0.051 (common for all modes)			Off-Res.
	$\mathcal{F}$ σ_{Right}	Yes	0.601 ± 0.051 (common for all modes)			Off-Res.
$B\overline{B}$	$m_{\rm ES}$ Shape	Yes	-30.3 ± 14.0	-49.6 ± 8.7	-27.5 ± 9.0	$D^{*\pm}K_S^0\pi$
	ΔE Slope	Yes	1.01 ± 1.75	-0.54 ± 1.14	-2.14 ± 1.19	$D^{*\pm}K_S^0\pi$
	$\mathcal{F}$ Mean	Yes	0.532 ± 0.019 (common for all modes)			Gen. MC
	$\mathcal{F}$ σ_{Left}	Yes	0.637 ± 0.013 (common for all modes)			Gen. MC
	$\mathcal{F}$ σ_{Right}	Yes	0.367 ± 0.012 (common for all modes)			Gen. MC
Peak	$m_{\rm ES}$ Mean	Yes	5.2798 ± 0.0003 (common for all modes)			Gen. MC
	$m_{\rm ES}$ Width	Yes	3.33 ± 0.45 (common for all modes)			Gen. MC
	ΔE Exp. Coef.	Yes	-4.5 ± 6.2	-12.9 ± 27.1	-6.1 ± 7.9	$D^{*\pm}K_S^0\pi$

Table 10.6: Parameter values from the $B^0 \to D^{*\pm}K^{*\mp}$ fit. The errors on the fixed parameters are the ones obtained from the fit on the off-resonance/control/MC sample. They are varied by $\pm 1\sigma$ to account for systematic errors.

From the figures with PDFs superimposed on top of the projected data distributions one can see that the parametrizations and fits describe the data well.

Param.	All	N_{BB}	N_{Peak}	N_{Cont}	N_{sig}	ξ_{BB}	ξ_{Cont}	$P1_{BB}$	$P1_{Cont}$
N_{BB}	0.707	1.000	-0.459	-0.535	-0.002	-0.220	0.067	-0.116	-0.029
N_{Peak}	0.713	-0.459	1.000	0.015	-0.128	0.496	-0.013	0.093	0.066
N_{Cont}	0.601	-0.535	0.015	1.000	-0.034	-0.031	-0.059	0.061	-0.006
N_{sig}	0.223	-0.002	-0.128	-0.034	1.000	0.016	0.037	-0.028	0.007
ξ_{BB}	0.735	-0.220	0.496	-0.031	0.016	1.000	-0.520	-0.173	0.181
ξ_{Cont}	0.606	0.067	-0.013	-0.059	0.037	-0.520	1.000	0.171	-0.119
$P1_{BB}$	0.624	-0.116	0.093	0.061	-0.028	-0.173	0.171	1.000	-0.584
$P1_{Cont}$	0.599	-0.029	0.066	-0.006	0.007	0.181	-0.119	-0.584	1.000

Table 10.7: Correlations among fitted variables for the $B^0 \to D^{\pm}K^{*\mp}$ fit. The leftmost column contains the total correlation, the signal yield having the smallest correlation, mainly with the peaking background, as it resembles signal the most.

10.3 Yields

The yields from the fits shown in Figure 10.4 for $B^0 \to D^{\pm}K^0\pi^{\mp}$ and in Figure 10.5 for $B^0 \to D^{*\pm}K^0\pi^{\mp}$are listed in Table 10.8.

Component	$D^{\pm}K^0\pi^{\mp}$	$D^{\pm}K^{*\mp}$	$D^{*\pm}K^0\pi^{\mp}$	$D^{*\pm}K^{*\mp}$
Signal	230 ± 24	143 ± 14	134 ± 17	78 ± 10
Continuum	7009 ± 117	1135 ± 44	3464 ± 97	708 ± 38
$B\overline{B}$	2578 ± 112	92 ± 34	2072 ± 91	52 ± 28
$B\overline{B}$ peak	202 ± 66	12 ± 17	69 ± 26	12 ± 9
Purity (%) box	40	84	45	83
Purity (%) error	39	73	44	71

Table 10.8: Raw yields and purities obtained from the likelihood fit (degenerate peaking background has not been subtracted). The errors are statistical only, and the purities are defined either within a certain signal region (box) or directly from the uncertainties (error).

As can be seen from the yields, a significant fraction of the events lies in the K^* band. Not surprisingly, both the background and the peaking background are significantly reduced when considering only the $K(892)^{*\pm}$ band. The purity listed in Table 10.8 is defined in two ways; first as the signal fraction in the region $[5.27, 5.29] \times [-0.02, 0.02] \times [-0.5, 3]$ in m_{ES}, ΔE, and $\mathcal{F}$ (box) and second as $N_{sig}/\sigma^2_{N_{sig}}$ (error). The latter definition shows to what degree the statistical error corresponds to that of an equally large perfectly clean sample.

10.4 Degenerate peaking background

The degenerate peaking background is evaluated using the sidebands of m_D and $m_{K^0_s}$. The sidebands consist of events which satisfy all selection criteria, but whose D or K^0_S masses lie outside the signal region. For the D meson sideband the intervals $20 < |m_D - m_D^{\mathrm{peak}}| < 35\,\mathrm{MeV}$ around the central value were used for the fully charged modes ($35 < |m_D - m_D^{\mathrm{peak}}| < 65\,\mathrm{MeV}$ for the $D^0 \to K^-\pi^+\pi^0$ mode) and for the K^0_S the intervals $10 < |m_{K^0_s} - m_{K^0_s}^{\mathrm{peak}}| < 15\,\mathrm{MeV}$ were used. The result can be found in Table 10.9.

Mode	Signal	m_D sideb.	$m_{K^0_s}$ sideb.
$D^{\pm}K^0\pi^{\mp}$	229.5 ± 24.2	21.7 ± 13.7	-6.4 ± 4.5
$D^{\pm}K^{*\mp}$	142.5 ± 14.1	10.8 ± 5.5	3.5 ± 3.2
$D^{*\pm}K^0\pi^{\mp}(K\pi)$	42.9 ± 10.0	9.4 ± 7.5	3.7 ± 4.6
$D^{*\pm}K^{*\mp}\ (K\pi)$	31.8 ± 8.0	3.1 ± 3.0	-2.0 ± 2.9
$D^{*\pm}K^0\pi^{\mp}(K\pi\pi^0)$	59.4 ± 11.8	2.6 ± 4.6	5.7 ± 4.7
$D^{*\pm}K^{*\mp}\ (K\pi\pi^0)$	24.6 ± 5.5	1.0 ± 1.0	-0.8 ± 1.7
$D^{*\pm}K^0\pi^{\mp}(K3\pi)$	31.9 ± 6.9	6.7 ± 3.4	2.1 ± 2.2
$D^{*\pm}K^{*\mp}\ (K3\pi)$	23.1 ± 5.9	3.6 ± 2.8	1.6 ± 2.4

Table 10.9: Signal and degenerate peaking background yields evaluated using the m_D and $m_{K^0_S}$ sideband.

For the K^0_S mass spectrum, the sideband window (5 MeV on each side) has a smaller size than the signal window (7 MeV on each side), and as the background is assumed (and tested on generic MC and by fits) to be uniformly distributed, a scaling factor of $7/5 = 1.4$ has been applied. For the $D^{\pm}$ sideband the window is 15 MeV, which is larger than the signal window of 12 MeV, so here a scaling factor of $12/15 = 0.8$ is used. The D^0 sidebands have

the same size as the signal window, thus no scaling factor is required. In the selection, the D and K_S^0 sidebands are made mutually exclusive to avoid double counting (due to multiple candidates).
The yields found in the sidebands can have two origins. Either they are due to misreconstructed signal events or they are from B^0 decays with the same final state. It is the fraction of the yields which exceed the misreconstruction estimates, that has to be subtracted.
It is noteworthy that only very little degenerate peaking background (if any) originates from fake K_S^0 mesons, thus the potentially troublesome $B^0 \to D^{(*)\pm} a_1^{\mp}$ channels (used as control channels) do not contribute to the $B^0 \to D^{(*)\pm} K^0 \pi^{\mp}$ signal. The degenerate peaking background from off-resonance data is consistent with zero, as one would expect.

10.5 Section summary and conclusion

The selected data sample is fitted with an unbinned maximum likelihood fit in the three variables m_{ES}, ΔE, and $\mathcal{F}$, considering four distinct PDF components: Signal, along with continuum, combinatorial $B\overline{B}$, and peaking $B\overline{B}$ background. The fit is tested on both simulated events and control samples. The signal shape is determined from the latter and frozen in the fit, while the background shape is mostly left floating.

The fit converges close to the initial values obtained from MC samples, and the yields are 230 ± 24 for $B^0 \to D^{\pm} K^0 \pi^{\mp}$ and 134 ± 17 for $B^0 \to D^{*\pm} K^0 \pi^{\mp}$, with a purity around 40%. Repeating the fit in the $K^{*\pm}$ resonant region yields 143 ± 14 and 78 ± 10 events, respectively.

11 Systematic errors

Everything is vague to a degree you do not realize till you have tried to make it precise.

[Bertrand Russell, 1872-1970]

Systematic errors is a common expression for sources of uncertainty which are associated with the (subjective) choices made in an analysis. They arise from the factors used for turning an event yield into a branching fraction, such as efficiency corrections, PDF parametrizations, theoretical uncertainties and total number of B decays considered.
Their treatment can be controversial, as no fixed set of rules exists on how to determine their size. Furthermore, as they are often *not* of statistical origin and therefore not Gaussian in distribution, great care has to be taken in their interpretation.
Below are listed the various systematic errors considered in this analysis. For the most part they are of quite standard origin and their size relatively easily calculable.

11.1 Efficiency corrections

One of the most important and central corrections to make is the efficiency correction, and it is not surprisingly also the one that requires the most time and effort. Basically, the efficiency for signal events is calculated using signal MC. Knowing how many decays went into the simulation, and how many were reconstructed, an efficiency can be calculated.
However, though the detector simulation is both detailed and tuned to data, the correspondance is never complete (see Section 9.1.1). As a result, one is forced to quantify the level of accuracy to which the simulation holds and consequently the size of eventual corrections needed to restore identical properties between data and MC, which evidently have associated (systematic) errors.
To determine the size of the corrections and their errors, one uses large and sufficiently clean data samples obtained from other sources. Along with corresponding MC samples, the differences between the two can be quantified for each type of particle/phenomena.

11.1.1 Tracking efficiency correction

The tracking efficiency correction is determined from $e^+e^- \rightarrow \tau^+\tau^-$ events, where (apart from neutrinos) one tau decays to a lepton ($Br(\tau^- \rightarrow \ell^-\bar{\nu}_{\ell^-}\nu_{\tau^-}) = (35.21 \pm 0.09)\%$) and the other tau decays to three charged pions ($Br(\tau^- \rightarrow \pi^-\pi^+\pi^-\nu_{\tau^-}) = (9.22 \pm 0.10)\%$) *of which only two are detected.* Thus the signature and hence selection criteria of such events is an isolated lepton with two recoiling tracks. From knowing that a fourth track has to exist (simply from charge conservation), the following ratio (for that track) can be establised:

$$\Delta = \frac{(\epsilon A)_{\mathrm{Data}} - (\epsilon A)_{\mathrm{MC}}}{(\epsilon A)_{\mathrm{MC}}}, \tag{11.1}$$

where ϵ is the efficiency and A is the acceptance of MC and data. To a very good approximation the acceptance, cancels in this ratio, and consequently the difference between the simulated and the actual tracking efficiency, denoted Δ, can be determined (the correction to be made is $1 - \Delta$). Note, that not the absolute, but only the relative tracking efficiency between data and Monte Carlo is measured.
In general, the tracking efficiency correction is a function of momentum, direction, and track quality, and therefore a mapping according to these quantities may be needed. With a binned mapping, the division has to be coarse enough for each bin to contain enough statistics, such

that the (systematic) error on the correction remains small compared to the correction itself[77].
A set of corrections is computed for each setting of the DCH high voltage, since it changes with these. But as the Monte Carlo samples are produced with the same conditions as the data, the corrections do not change significantly for the different voltage settings.

The tracking efficiency corrections and systematic errors are computed *on the average* for each track quality involved, as no significant dependencies are seen in other dimensions [Tra03]. GoodTracksVeryLoose (GTVL) do not require any correction (the correction is consistent with unity), and a systematic error of 1.3% is assigned for each track. GTVL with low momentum (i.e. the soft pion from $D^{*\pm}$ decays) do not require any correction either, but entail a systematic error of 1.6%. The track identified as a kaon for the reconstruction of $D^{\pm}$ candidates is required to be a GoodTrackLoose (GTL). For these, there is a correction of 0.8% and a systematic error (which is entirely correlated with those of GTVL) of 2.0%. Overall, the total systematic error due to tracking efficiency corrections is estimated to be 5.9% for $B^0 \to D^{\pm}K^0\pi^{\mp}$, and 5.5%, 5.5%, and 8.1% for the three D^0 modes of $B^0 \to D^{*\pm}K^0\pi^{\mp}$.

11.1.2 K^0_S efficiency correction

The K^0_S efficiency correction has to be computed separately from the tracking efficiency correction, due to the depency on the decay distance and the additional requirements demanded in the analyses. The corrections are studied using high-statistics data and MC samples, and calculated for the most common set of requirements. The efficiency correction is applied on an event-by-event basis [KsC03].
The overall correction factor obtained is 0.971 ± 0.018 for $B^0 \to D^{\pm}K^0\pi^{\mp}$ and 0.971 ± 0.019 for $B^0 \to D^{*\pm}K^0\pi^{\mp}$ averaged over different DCH voltage settings and D^0 decay modes. In addition, a correction of 0.981 ± 0.002 has to be applied, due to the K^0_S mass cut.

11.1.3 π^0 efficiency correction

The π^0 efficiency correction is obtained from tau decays, where one tau decays to a lepton ($\tau^+ \to \ell^+\nu_\ell\bar{\nu}_\tau$) and the other to one charged hadron and N π^0's ($\tau^- \to h^- N\pi^0\nu_\tau$). The ratio between events with $N = 1$ and $N = 2$ depends only on the branching ratio (known to better than 1%), the MC kinematics (assumed well-modelled) and the π^0 efficiency. Finally, a large clean sample of such events can be obtained from the data. This sample is also used to measure differences between the π^0 mass resolution and the energy scale difference between data and MC.
The systematic errors due to efficiency differences is estimated to be 3% and the uncertainty due to mass resolution 3% also. The overall systematic error comes out to be 5.0%, which also includes a small bias error from the tau method itself.

11.1.4 PID efficiency correction

The PID efficiency correction is obtained from a pure sample of $D^{*+} \to D^0(\to K^-\pi^+)\pi^+_{\rm soft}$ decays, where the identity of the kaon and pion from the D^0 decay is determined from the charge of the soft pion. By comparing the PID performance for this selection with that of an equivalent MC sample, the correction and its uncertainty can be obtained as a function of momentum and selection criteria.
For the three decay modes of the D^0, the kaon list used is `KMicroNotPion`, for which the corrections comes out to be 0.977, 0.965, and 0.961, respectively. In the reconstruction of $D^{\pm}$ candidates, `KMicroNotPionGTL` are used, for which no mapping of the efficiency correction

[77] A more refined approach is to parametrize the efficiency correction and then fit it on the data samples.

exists. Therefore the corrections relevant to KMicroNotPion were used, which yielded a correction of 0.957. The associated systematic error is multiplied by the efficiency ratio between the two selections[78] to account for possible differences. Included in the above corrections is also the demand that the bachelor track is not a KMicroTight or a eMicroVeryTight.
The systematic error due to the PID corrections was taken to be 2.0% for the $B^0 \to D^{\pm}K^0\pi^{\mp}$ mode, as suggested by the PID analysis [PID03], and 2.2% for the $B^0 \to D^{*\pm}K^0\pi^{\mp}$ mode.

11.1.5 $D^{(*)}$ efficiency corrections and branching ratios uncertainties

The cuts applied on the $D^{0/\pm}$ masses and D^* mass difference are also potentially different for data and MC. In order to estimate the corrections needed and their associated systematic error, the efficiencies were determined from a fit in $m_{\rm ES}$ of the control channels $B^0 \to D^{(*)\pm}a_1^{\mp}$ on both data and MC samples. The *ratio* of the data and MC efficiencies were computed for each mode, and the results can be found in Table 11.1 along with the errors associated with the branching fractions of the $D^{(*)}$ mesons [PDG02].

Mode	$B^0 \to D^{\pm}K^0\pi^{\mp}$		$B^0 \to D^{*\pm}K^0\pi^{\mp}$	
Submode	$K\pi\pi$	$K\pi$	$K\pi\pi^0$	$K3\pi$
$D^{\pm/0}$ mass cut	0.981 ± 0.004	0.983 ± 0.009	0.980 ± 0.011	0.976 ± 0.012
Br of $D^{\pm/0}$	± 0.066	± 0.024	± 0.069	± 0.042
$D^{*\pm}$ mass cut	–	1.004 ± 0.004	0.999 ± 0.012	0.998 ± 0.005
Br of $D^{*\pm}$	–	± 0.007 (common for all modes)		

Table 11.1: Efficiency corrections and associated systematic errors for $D^{(*)}$ mesons. The corrections are calculated by applying the same cuts as used in the analysis to data and MC samples of the control channel $B^0 \to D^{(*)\mp}\pi^{\pm}$.

The impact of the vertex probability cut was studied in the same manner, but as all corrections were consistent with one with very small errors (typically 0.998 ± 0.002), no corrections were applied. The branching fraction uncertainties were taken from [PDG02].

11.1.6 Efficiency modelling

In three-body decays the momenta of the daughters is not fixed and therefore the efficiency varies from event to event, contrary to the usual two-body case. In order to correct for the efficiency in the branching ratio calculation, one is obliged to map the efficiency across the Dalitz plot. This is done using signal MC generated with a uniform distribution in the Dalitz plot.
Both a continuous and a binned approach were attempted. The continuous mapping consisted of a generic function, which expanded around every statistically significant second derivative. This yielded a function, which was sufficiently flexible to map the changing efficiency across the Dalitz plot, while at the same time not being too sensitive to statistical fluctuations. The binned mapping was done by dividing the Dalitz plane (defined as $m(D,K) \in [0,30]\,\mathrm{GeV}^2$ and $m(K,\pi) \in [0,15]\,\mathrm{GeV}^2$) into an $N \times N$ grid. The choice of number of bins, N_{bins}, was a trade-off betweeen granularity of the mapping and the statistics in each bin. Below 10×10 bins the binning was too coarse to describe the features of the efficiency variations, and above 100×100 bins, statistical fluctuations started to show.
Due to resolution effects, events can lie beyond the Dalitz limit, and the question of where to limit the Dalitz region arises. In the following the energies and invariant masses were strictly

[78] Using signal MC, this ratio was evaluated to be 1.082.

required to lie within their kinematic limits, while the cosine of the angle between any pair of daughters was required to lie in the range $[-1-\delta, 1+\delta]$, where $\delta = 0.05$ accounts for the resolution. These requirements were found to be adequate for the purpose, though more involved criteria were considered (see Section 12.3.2).

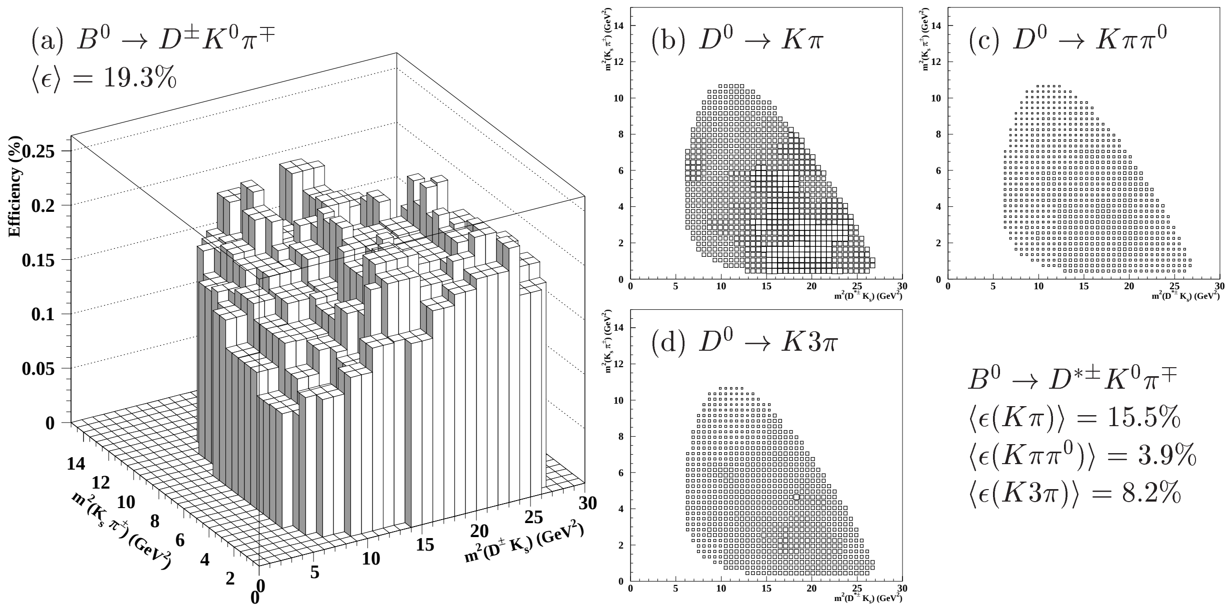


Figure 11.1: Efficiency as a function of Dalitz plot position for (a) $B^0 \to D^{\pm}K^0\pi^{\mp}$ and (b)-(d) $B^0 \to D^{*\pm}K^0\pi^{\mp}$ into the D^0 final states $K\pi$, $K\pi\pi^0$, and $K3\pi$, respectively. The general distribution is alike for all channels, but the efficiency varies much between them. Note the grouped binning, which avoids edge effects.

Both the continuous and the binned approach suffer from edge effects, where the efficiency drops rapidly, but has tails as discussed above. This results in regions with very low and quite uncertain efficiencies, which can have a sizable impact on the branching fraction measurement, when correcting for them.
Whereas this problem is hard to redeem for the continuous method, it is easier for the binned approach, where it can be fixed by the use of "smart" binning. One demands that every bin has a minimum number of events in them ($N < 100$ for $D^{\pm}K^0\pi^{\mp}$ and $N < 25$ for each D^0 mode of $D^{*\pm}K^0\pi^{\mp}$). If this is not the case, then the bin is averaged with its surrounding neighbours (i.e. merged) till the statistics is increased beyond the minimum requirement (see Fig. 11.1a). For each bin the efficiency is calculated as the number of reconstructed signal MC events divided by the number generated in that bin (correcting for the size of the bin, if it lies on the edge).
For kinematic reasons, the invariant masses should lie in the plane $m^2_{DK} + m^2_{D\pi} + m^2_{K\pi} = m^2_{B^0} + m^2_D + m^2_{K^0} + m^2_{\pi^{\pm}} \equiv M^2$, which simply corresponds to demanding that $\Delta E = 0$. The invariant masses are therefore corrected by the factor $(m^2_{DK} + m^2_{D\pi} + m^2_{K\pi})/M^2$, which essentially realizes a linear projection onto the Dalitz plane. None of the corrections are larger than 4%.

To quantify the systematic error in modelling the efficiency, the parametrization of the latter is varied. The resolution parameter δ was varied between 0.025 and 0.075, the minimum required number of events, N_{min}, was varied between 50 and 150 (12 and 50), and the binning N_{bins} was varied between 10 and 100 (10 and 50) for the $B^0 \to D^{\pm}K^0\pi^{\mp}$ ($B^0 \to D^{*\pm}K^0\pi^{\mp}$) mode. The impact of the last variation can be seen in Fig. 11.2.

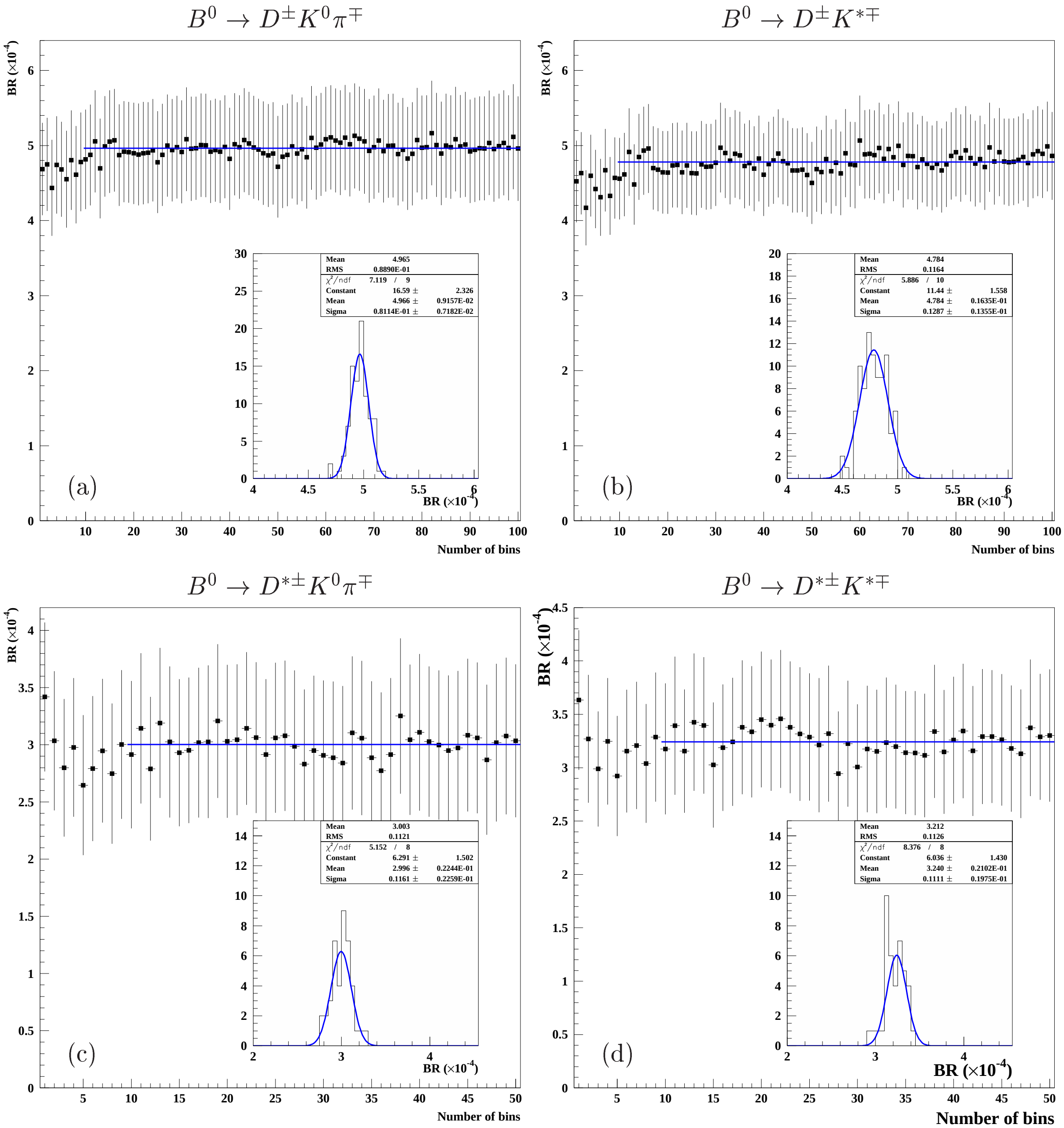


Figure 11.2: Result for branching ratios using different binnings for the four modes (a) $B^0 \to D^\pm K^0 \pi^\mp$, (b) $B^0 \to D^\pm K^{*\mp}$, (c) $B^0 \to D^{*\pm} K^0 \pi^\mp$, and (d) $B^0 \to D^{*\pm} K^{*\mp}$. Binnings are required to be 10×10 or greater, and they are fitted with a constant. The inserted figures show the distribution of results, and the width is taken as a systematic error.

The systematic error from normalization was 3.0% (4.1%), the one from the minimum required number of events in a bin was 0.8% (1.3%) and the RMS of the results using different binnings were 1.7% (2.3%) for the $B^0 \to D^\pm K^0 \pi^\mp$ ($B^0 \to D^\pm K^{*\mp}$) mode. The last error is mostly of statistical origin in lacking signal MC events in the most contributing areas.
For the $B^0 \to D^{*\pm} K^0 \pi^\mp$ ($B^0 \to D^{*\pm} K^{*\mp}$) mode, the numbers are 0.2%, 0.2% and 0.6% (0.5%, 1.4% and 1.4%) from normalization, 2.0%, 2.0% and 2.3% (2.8%, 1.5% and 3.2%) from the minimum required number of events in a bin and 5.7%, 6.9% and 5.6% (5.0%, 4.0% and 5.4%) from changing the binning, for the three D^0 modes, respectively.

11.1.7 PDF shape

To calculate the systematic due to the PDF shape, one varies all fixed parameters in turn by $\pm 1\sigma$, as listed in Tables 10.2 and 10.5, and note the impact it has on the number of signal events, N_{sig}. The systematic error is taken to be these variations in quadrature taking correlations (obtained from the control samples) among the variables into account:

$$\Delta N = \sqrt{\sum_{i\leq j} \frac{\partial N}{\partial x_i}\frac{\partial N}{\partial x_j}\sigma(x_i)\sigma(x_j)\,\rho_{ij}}. \tag{11.2}$$

As the coefficient of the exponential PDF in ΔE for the peaking background is only fixed in the fit of the $K^{*\pm}$ band, it is of course only varied here.

Sample	Entire Dalitz Plot		K^* Band (± 100 MeV)	
Parameter	$\Delta N_{\text{All}}(+1\sigma_x)$	$\Delta N_{\text{All}}(-1\sigma_x)$	$\Delta N_{K^*}(+1\sigma_x)$	$\Delta N_{K^*}(-1\sigma_x)$
m_{ES} Mean	-0.29	0.42	0.03	-0.03
m_{ES} Width	1.34	-1.36	0.49	-0.51
m_{ES} Max	0.13	-0.01	-0.02	0.02
ΔE Mean	0.35	-0.18	-0.20	0.21
ΔE Width1	**3.44**	**-4.38**	**1.04**	**-1.02**
ΔE Width2	**7.25**	**-8.57**	**4.20**	**-4.62**
ΔE Frac	**-10.96**	**8.86**	**-4.91**	**3.57**
Correlated variation	8.28		3.59	
$\mathcal{F}$ Mean Sig	**-3.25**	**3.03**	**-0.83**	**0.73**
$\mathcal{F}$ SigmaLeft Sig	1.54	-1.57	0.47	-0.48
$\mathcal{F}$ SigmaRight Sig	-1.74	1.81	-0.33	0.32
Correlated variation	3.93		0.86	
$\mathcal{F}$ Mean Cont	0.47	-0.30	-0.21	0.25
$\mathcal{F}$ SigmaLeft Cont	0.84	-0.70	0.17	-0.16
$\mathcal{F}$ SigmaRight Cont	0.01	-0.23	0.13	-0.34
Correlated variation	1.16		0.16	
$\mathcal{F}$ Mean BB	0.07	-0.35	0.02	-0.10
$\mathcal{F}$ SigmaLeft BB	-0.09	0.09	0.01	-0.01
$\mathcal{F}$ SigmaRight BB	0.20	0.08	-0.01	-0.06
Correlated variation	0.22		0.09	
m_{ES} Mean Peak	**4.03**	**-2.19**	0.57	-0.45
m_{ES} Width Peak	**2.77**	**-2.35**	0.20	-0.20
ΔE Peak Exp. Coef.	–	–	**1.35**	**-2.08**
Correlated variation	4.31		1.81	
Total systematic (events)	10.2		4.1	
Total systematic (%)	4.5		2.9	

Table 11.2: Systematics from the PDF shapes for the $B^0 \to D^{\pm}K^0\pi^{\mp}$ and $B^0 \to D^{\pm}K^{*\mp}$ modes. Each fixed parameter is varied by $\pm\sigma$, and the change in number of signal events, ΔN_{sig}, is determined. Including correlations (i.e. when the parameters are obtained from the same control sample), the total impact on the signal yield is determined. The exponential coefficient α_{peak} is not varied for the $DK\pi$ sample, as it is left floating in the fit. Correlations are included in the calculation of the overall impact on the yield (see Eq. (11.2)). The variables with the largest impact are written in bold.

Sample	$D^0 \to K^-\pi^+$		$D^0 \to K^-\pi^+\pi^0$		$D^0 \to K^-\pi^+\pi^+\pi^-$	
Dalitz Region	ΔN_{All}	ΔN_{K^*}	ΔN_{All}	ΔN_{K^*}	ΔN_{All}	ΔN_{K^*}
m_{ES} Mean	-0.27	-0.17	-0.29	-0.07	-0.03	-0.11
m_{ES} Width	0.33	0.44	0.25	0.16	0.52	0.44
m_{ES} Max	-0.02	0.07	-0.67	-0.05	0.12	0.01
ΔE Mean	0.03	-0.05	0.21	0.09	-0.05	-0.14
ΔE Width1	**0.53**	**0.21**	**1.14**	**0.24**	**1.21**	**0.20**
ΔE Width2	**1.15**	**0.38**	**-1.52**	**0.61**	**1.37**	**0.66**
ΔE Frac	**-1.15**	**-0.45**	**-1.74**	**-0.65**	**-2.49**	**-1.26**
Correlated variation	1.48	0.71	2.82	0.88	2.86	1.37
$\mathcal{F}$ Mean Sig	0.13	-0.03	-0.21	-0.06	-0.41	0.02
$\mathcal{F}$ SigmaLeft Sig	-0.12	-0.02	0.02	-0.01	0.05	-0.03
$\mathcal{F}$ SigmaRight Sig	-0.09	-0.04	-0.18	-0.03	-0.46	-0.05
Correlated variation	0.11	0.08	0.34	0.09	0.77	0.05
$\mathcal{F}$ Mean Cont	0.06	-0.22	2.14	-0.03	-0.04	-0.09
$\mathcal{F}$ SigmaLeft Cont	-0.08	0.11	-0.79	0.04	-0.31	0.02
$\mathcal{F}$ SigmaRight Cont	-0.05	-0.07	0.75	-0.05	-0.24	-0.10
Correlated variation	0.07	0.22	2.26	0.06	0.50	0.17
$\mathcal{F}$ Mean BB	-0.14	0.00	0.19	0.02	0.06	-0.12
$\mathcal{F}$ SigmaLeft BB	-0.03	0.00	-0.16	-0.01	-0.03	0.03
$\mathcal{F}$ SigmaRight BB	-0.09	0.01	-0.03	-0.04	0.02	-0.09
Correlated variation	0.24	0.01	0.15	0.03	0.06	0.18
m_{ES} Mean Peak	0.32	0.00	0.00	-0.05	-0.30	0.15
m_{ES} Width Peak	0.10	0.00	0.00	-0.12	-0.17	-0.70
Correlated variation	0.35	0.00	0.00	0.14	0.35	0.71
Total systematic (events)	2.39	0.75	3.63	0.90	3.02	1.56
Total systematic (%)	7.5	3.0	8.5	3.9	5.1	4.8

Table 11.3: Systematics from the PDF shapes for the $B^0 \to D^{*\pm}K^0\pi^\mp$ and $B^0 \to D^{*\pm}K^{*\mp}$ modes. Each fixed parameter is varied by $\pm 1\sigma$, and the change in number of signal events, ΔN_{sig} is determined. Including correlations (i.e. when the parameters are obtained from the same control sample), the total impact on the signal yield is determined. The exponential coefficient α_{peak} is not varied for the $DK\pi$ sample, as it is left floating in the fit. Correlations are included in the calculation of the overall impact on the yield (see Eq. (11.2)). The variables with the largest impact are written in bold.

11.1.8 Resonant correction

In the absence of interference, resonances follow a Breit-Wigner (BW) shape with a caracteristic central value and width for each resonance. The width of the $K^{*\pm}$ is 50 MeV, which means that it has tails beyond the required mass of $m(K^{*\pm}) \pm 100$ MeV. To estimate these tails, and thus the selection efficiency of the mass requirement, dedicated $B^0 \to D^{(*)\mp}K^{*\pm}$ signal MC was used. The selection efficiency was found to be $93 \pm 1\%$ for both the $B^0 \to D^{\pm}K^{*\mp}$ and the $B^0 \to D^{*\pm}K^{*\mp}$ decays, which is in agreement with numerical estimates.

In addition, non-resonant contributions (or tails of other resonances) can potentially be present under the BW shape. Such contributions are estimated from fitting the invariant mass

$m(K^0, \pi^\pm)$ in the range [0.8;1.3] GeV on efficiency, background, and correlation corrected signal events in data[79]. The results of such a fit are shown in figure 11.3. The fractions fitted, $f^{[0.8;1.3]}$, are the fractions of the K^* resonant component in the range [0.8;1.3] GeV, and we obtain: $f^{[0.8;1.3]}_{D^\pm K^0\pi^\mp} = 0.950 \pm 0.081$ and $f^{[0.8;1.3]}_{D^{*\pm} K^0\pi^\mp} = 1.217 \pm 0.148$. From these one can extract the same fractions for the interval [0.8;1.0] GeV, which are the correction factors with systematic uncertainties $\sigma(f)$ needed. This is done as follows:

$$f = \frac{\epsilon_{\text{res}} f^{[0.8;1.3]}}{\epsilon_{\text{res}} f^{[0.8;1.3]} + \epsilon_{\text{non-res}}(1 - f^{[0.8;1.3]})}, \qquad \sigma(f) = \frac{\epsilon_{\text{res}}\epsilon_{\text{non-res}}}{(\epsilon_{\text{res}} f^{[0.8;1.3]} + \epsilon_{\text{non-res}}(1 - f^{[0.8;1.3]}))^2} \tag{11.3}$$

where $\epsilon_{\text{res}} = \int_{0.8}^{1.0} BW / \int_{0.8}^{1.3} BW$ and $\epsilon_{\text{non-res}} = (1.0 - 0.8)/(1.3 - 0.8) = 0.4$ are the fractions of resonant and non-resonant in the range [0.8;1.3] GeV, which falls in the range [0.8;1.0] GeV. This yields the corrections and associated systematic errors, $f_{D^\pm K^0\pi^\mp} = 0.978 \pm 0.036$ and $f_{D^{*\pm} K^0\pi^\mp} = 1.083 \pm 0.050$. Since the value of $f_{D^{*\pm} K^0\pi^\mp}$ is outside the physical region (a negative non-resonant contribution does not make any sense), but consistent within errors, the resonant fraction is chosen to be unity with the same error.
The additional corrections and systematics to $Br(B^0 \to D^{(*)\pm} K^{*\mp})$ are thus 0.950 ± 0.037 and 0.930 ± 0.051 for the modes $B^0 \to D^\pm K^{*\mp}$ and $B^0 \to D^{*\pm} K^{*\mp}$, respectively.

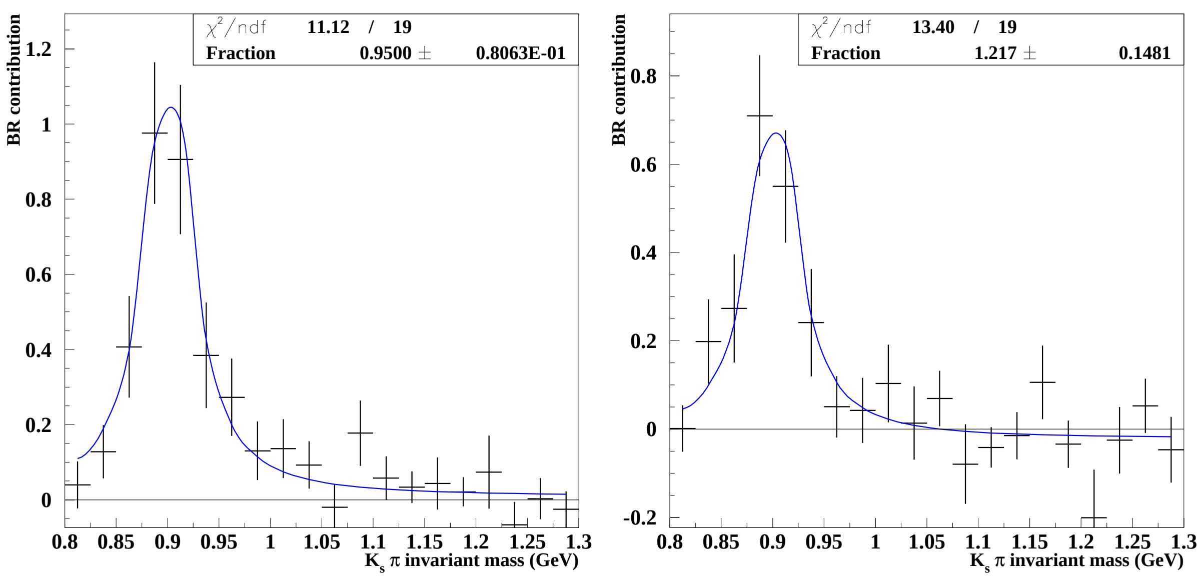


Figure 11.3: Fit of $K^{*\pm}$ invariant mass $m(K^0, \pi^\pm)$ in the range [0.8;1.3] GeV on efficiency, background, and correlation corrected signal. The PDF contains a BW (resonant) and a flat (non-resonant) part both normalized to unity in the range [0.8;1.3] GeV, and the fraction $f^{[0.8;1.3]}$ can be used to extract the fraction of (non-)resonant in the K^* range.

The result of the above fits can be cross checked by fitting the helicity distribution of the $K^{*\pm}$ resonance (see Fig. 12.4). The distribution is fitted with the characteristic $dN/d\cos(\theta) = \cos^2(\theta)$ distribution for the resonant part and a flat component for the non-resonant part, and the result of such a fit is $f = 0.993 \pm 0.058$, in good agreement with the invariant mass fit. The error is slightly larger, as only the range [0.8,1.0] GeV is fitted.

[79]In fact the $m(K^0, \pi^\pm)$ distribution of data weighted by W_{sig} defined in Eq. 12.2.

11.1.9 Sideband subtraction

The degenerate peaking background is measured using the $D^{\pm/0}$ and K^0_S mass sidebands (see Section 10.4). The sideband data is fitted in the exact same manner as the signal, and subtracted from the signal. The associated errors are considered statistical.
However, there is a small amount of signal in the sidebands, which has to be corrected for. The fraction is determined from the control channels, which yields $3.2 \pm 0.6\%$ for the $B^0 \to D^\pm a_1^\mp$ channel and $4.2 \pm 1.1\%$ for the $B^0 \to D^{*\pm} a_1^\mp$ channel. This is in accordance with signal MC for the same channels ($3.6 \pm 0.5\%$ and $3.7 \pm 0.4\%$, respectively), and signal MC for the $B^0 \to D^{(*)\pm} K^0 \pi^\mp$ modes ($3.9 \pm 0.2\%$ and $4.6 \pm 0.3\%$, respectively).
This means that the fraction of events in the $D^{\pm/0}$ sidebands, which exceeds $\sim 3-4\%$ is not accounted for by the signal, and thus other contributions are there. However, the statistics available deprive the situation of a decisive conclusion.

For obvious reasons, the K^0_S can not be investigated in the control channels, but in signal MC signal fractions of $1.5 \pm 0.2\%$ for $B^0 \to D^\pm K^0 \pi^\mp$ and $1.6 \pm 0.3\%$ for $B^0 \to D^{*\pm} K^0 \pi^\mp$ were found. The K^0_S mass sideband contributions are consistent with zero, as can be seen in Table 10.9, but the statistics are not adequate for determining effects of the order suggested by signal MC.

To account for any true signal that falls in the sidebands, corrections of the order suggested by the control channels for the $D^{\pm/0}$ and by the signal MC for the K^0_S are made, and the errors in these corrections are considered systematic.

11.1.10 Systematic error from luminosity determination

The total luminosity used in the analysis is evaluated by considering QED processes (see Section 7.3.1), and the determination carries with it a systematic error of 1.1%.

11.1.11 Summary of corrections and systematic errors

The efficiency corrections are summarised in Table 11.4 and the systematic errors are summarised in Table 11.5.

Mode	$B^0 \to D^\pm K^0 \pi^\mp$	$B^0 \to D^{*\pm} K^0 \pi^\mp$		
Submode	$K\pi\pi$	$K\pi$	$K\pi\pi^0$	$K3\pi$
K^0_S reconstruction	0.971 ± 0.018	0.971 ± 0.019 (same for all three modes)		
K^0_S mass cut	0.981 ± 0.002	0.981 ± 0.002 (same for all three modes)		
$K^{*\pm}$ mass cut	0.930 ± 0.010	0.930 ± 0.010 (same for all three modes)		
$D^{\pm/0}$ mass cut	0.981 ± 0.007	0.991 ± 0.009	0.987 ± 0.010	0.994 ± 0.011
$D^{*\pm}$ mass cut	–	1.002 ± 0.004	1.020 ± 0.007	0.996 ± 0.008
Tracking	0.992 ± 0.059	1.000 ± 0.055	1.000 ± 0.055	1.000 ± 0.081
PID	0.957 ± 0.020	0.968 ± 0.020 (same for all three modes)		
$D^{\pm/0}$ sideband subtraction	0.968 ± 0.006	0.958 ± 0.011 (same for all three modes)		
K^0_S sideband subtraction	0.985 ± 0.002	0.984 ± 0.003 (same for all three modes)		
Total	0.769 ± 0.076	0.803 ± 0.081	0.814 ± 0.082	0.800 ± 0.101

Table 11.4: Efficiency corrections from cuts, reconstruction, PID, and sideband subtraction. The correction from the requirement on $m_{K^{*\pm}}$ only applies to the resonant modes.

Mode	$B^0 \to D^{\pm}K^0\pi^{\mp}$	$B^0 \to D^{*\pm}K^0\pi^{\mp}$			
Submode	$K\pi\pi$	$K\pi$	$K\pi\pi^0$	$K3\pi$	
Tracking	5.9	5.5	5.5	8.1	corr.
PID	2.2	2.0	2.0	2.0	corr.
π^0 reconstr.	–	–	5.0	–	
BR of K_S^0	0.4	0.4	0.4	0.4	corr.
K_S^0 reconstr.	1.8	1.9	1.9	1.9	corr.
K_S^0 mass cut	0.2	0.2	0.2	0.2	corr.
$K^{*\pm}$ mass cut	– (1.0)	– (1.0)	– (1.0)	– (1.0)	corr.
BR of $D^{\pm/0}$	6.6	2.4	6.9	4.2	
$D^{\pm/0}$ mass cut	0.7	0.9	1.1	1.2	
BR of D^*	–	0.7	0.7	0.7	corr.
D^* mass cut	–	0.4	0.7	0.8	
PDF	4.5 (2.9)	7.5 (3.0)	8.5 (3.9)	5.1 (4.8)	
Normalization	3.0 (4.1)	0.2 (0.5)	0.2 (1.4)	0.6 (1.4)	
Nmin	0.8 (1.3)	2.0 (2.8)	2.0 (1.5)	2.3 (3.2)	
Modeling	1.8 (2.4)	5.7 (5.0)	6.9 (4.0)	5.6 (5.4)	
Sideband	0.6	1.1	1.1	1.1	corr.
Luminosity	1.1	1.1	1.1	1.1	corr.
Total Systematic Error	11.0 (11.1)	11.8 (9.5)	15.5 (12.4)	12.6 (12.7)	

Table 11.5: Summary of systematic errors for measuring $Br(B^0 \to D^{\pm}K^0\pi^{\mp})$ and $Br(B^0 \to D^{*\pm}K^0\pi^{\mp})$. In parenthesis are shown the systematic errors, for the $K^{*\pm}$ resonant mode, when different. The last column states if the error is considered correlated among submodes. The errors are in percent.

12 Results

12.1 Branching fraction calculation

The standard method of calculating the branching fraction is:

$$Br = \frac{N_{signal} - N_{peaking}}{N_{B\overline{B}}\ \epsilon\ \Pi_i Br_i} \tag{12.1}$$

where N_{signal} is the number of signal events, $N_{peaking}$ is the estimated number of degenerate peaking background events, N_{B^0} is the total number of B^0 mesons, ϵ is the signal efficiency and Br_i are the branching fractions of the subsequent decays. It is assumed that all $\Upsilon(4S)$ decays to $B\overline{B}$ pairs of which half are neutral (i.e. $N_{B^0} = N_{B\overline{B}}$).

However, as the efficiency of a three-body decay varies from event to event (see Section 11.1.6), the branching fraction becomes the sum of each event's contribution corrected for the efficiency at the point of the event in the Dalitz plot.
The crucial question is what to use for the event contribution in the numerator of Eq. 12.1. A natural suggestion seems to be the signal probability of each event, simply defined as $N_{\rm sig}P_{\rm sig}/\sum_j N_j P_j$, and this choice would be a possibility. Nevertheless, this usual probability suffers from several deficiencies. First of all it fails to include the inevitable correlations among the yields of the various components in the likelihood fit. This means that the subsequent signal distributions (e.g. Dalitz plot) derived from these signal probabilities will not be entirely correct. Secondly and equally important, it is not clear how to calculate the statistical error on such a sum.
An alternative, which solves both of these problems in the most simple of manners, is ${}_s\mathcal{P}$lot weights [PLD04], where one introduced a weight defined on an event-by-event basis as:

$$W_{\rm sig} \equiv \frac{\sum_j \mathbf{V}_{{\rm sig},j}\ P_j}{\sum_j N_j\ P_j}, \tag{12.2}$$

where N_j and P_j are the number of events and the probability (PDF) of the j^{th} component, and $\mathbf{V}_{{\rm sig},j}$ is the signal row of the covariance matrix of the component yields obtained from the likelihood fit. Weighting each event by $W_{\rm sig}$, which in the absence of correlations is the signal usual probability defined as $N_{\rm sig}P_{\rm sig}/\sum_j N_j P_j$, yields the data signal distribution of any quantity (e.g. Dalitz plot distribution). Note that these weights can take values on both sides of the usual probability range $[0, 1]$.
Using these weights, the branching fraction becomes the sum of each event's contribution corrected by its efficiency:

$$Br = \sum_i \frac{W_{\rm sig}(m_{{\rm ES},i}, \Delta E_i, \mathcal{F}_i)}{N_{B\overline{B}^0}\ \epsilon_i\ \Pi_k Br_k}, \tag{12.3}$$

where the sum i is over all events and ϵ_i the efficiency at the point of the Dalitz plot where the event lies. In this manner, the varying efficiency is correctly accounted for. Another way of interpreting this is to calculate an efficiency corrected event yield as $N_{\rm corr} = \sum_i W_{{\rm sig},i}/\epsilon_i$, which can be entered into the numerator of Eq. (12.1). The degenerate peaking background is subtracted in the same manner, using the fits on the sidebands.

Finally the branching fractions are calculated applying the corrections (see Table 11.4) and subtracting the peaking background (see Table 10.9), and the results can be found in Table 12.1. Once again the procedure is applied both for the entire Dalitz plot and then for the $K^{*\pm}$ resonant region. To ensure that the procedure is the exact same, the same efficiency correction was applied to the resonant decay, but afterwards cross-checked with dedicated $B^0 \to D^{(*)\mp}K^{*\pm}$ signal MC (see Section 11.1.8).
The weights used for combining the three D^0 submodes are calculated from the statistical error with the *uncorrelated* systematic error added in quadrature, that is weight $= (\sigma^2_{\rm stat} + \sigma^2_{\rm syst,uncorr.})^{-1}$. The $D^0 \to K\pi$ dominates the average, which is a result of its purity.

Channel	Submode (weight)	Branching fraction	Stat. Error	Syst. Error
$B^0 \to D^{\pm}K^0\pi^{\mp}$		**4.97**	**0.69**	**0.55**
$B^0 \to D^{\pm}K^{*\mp}$		**4.78**	**0.58**	**0.53**
$B^0 \to D^{*\pm}K^0\pi^{\mp}$	$D^0 \to K\pi$ (0.62)	2.35	0.85	0.28
$B^0 \to D^{*\pm}K^0\pi^{\mp}$	$D^0 \to K\pi\pi^0$ (0.16)	2.83	1.66	0.44
$B^0 \to D^{*\pm}K^0\pi^{\mp}$	$D^0 \to K3\pi$ (0.22)	4.96	1.40	0.63
$B^0 \to D^{*\pm}K^0\pi^{\mp}$	$\chi^2 = 2.6$, $P(\chi^2) = 0.28$	**3.00**	**0.66**	**0.29**
$B^0 \to D^{*\pm}K^{*\mp}$	$D^0 \to K\pi$ (0.48)	3.33	0.85	0.32
$B^0 \to D^{*\pm}K^{*\mp}$	$D^0 \to K\pi\pi^0$ (0.23)	2.97	1.21	0.37
$B^0 \to D^{*\pm}K^{*\mp}$	$D^0 \to K3\pi$ (0.29)	3.23	1.09	0.41
$B^0 \to D^{*\pm}K^{*\mp}$	$\chi^2 = 0.1$, $P(\chi^2) = 0.97$	**3.22**	**0.59**	**0.29**

Table 12.1: Calculated branching fractions (10^{-4}) with statistical and systematic uncertainties. $P(\chi^2)$ reflects the compatibility when combining the modes, using the weights indicated in parenthesis. The $B^0 \to D^{(*)\pm}K^{*\mp}$ channels have been corrected for $Br(K^{*\pm} \to K^0\pi^{\pm})$.

12.2 Resonant fraction calculation

To calculate the fraction of $K^{*\pm}$ resonant contribution, denoted $f(B^0 \to D^{(*)\pm}K^0\pi^{\mp})$, and the uncertainty on this ratio, the branching fraction for the $K^{*\pm}$ resonant decay, Br_{res}, has to be compared to the branching fraction for the entire Dalitz region, Br_{all}. From the branching fraction for the three-body decay excluding the resonant region, i.e. the non-resonant decay[80], Br_{non}, the resonant fraction can be determined as:

$$f \equiv \frac{Br_{\mathrm{res}}}{Br_{\mathrm{res}} + Br_{\mathrm{non}}}, \qquad \sigma_f^2 = \sigma_{Br_{\mathrm{res}}}^2 \left(\frac{1-f}{Br_{\mathrm{res}} + Br_{\mathrm{non}}}\right)^2 + \sigma_{Br_{\mathrm{non}}}^2 \left(\frac{f}{Br_{\mathrm{res}} + Br_{\mathrm{non}}}\right)^2 \tag{12.4}$$

The non-resonant fraction and its uncertainty is computed as $Br_{\mathrm{non}} = Br_{\mathrm{all}} - Br_{\mathrm{res}}$ and $\sigma_{Br_{\mathrm{non}}}^2 = \sigma_{Br_{\mathrm{all}}}^2 - \sigma_{Br_{\mathrm{res}}}^2$, respectively, and the validity of this computation has been checked with fits of the non-resonant region. The result for both the $B^0 \to D^{\pm}K^0\pi^{\mp}$ and the $B^0 \to D^{*\pm}K^0\pi^{\mp}$ mode can be found in Table 12.2.

Channel	Br_{res}	Br_{non}	f
$B^0 \to D^{\pm}K^0\pi^{\mp}$	3.19 ± 0.39	1.78 ± 0.57	$0.64 \pm 0.08 \pm 0.02$
$B^0 \to D^{*\pm}K^0\pi^{\mp}$	2.15 ± 0.39	0.85 ± 0.53	$0.72 \pm 0.13 \pm 0.02$

Table 12.2: Resonant and non-resonant branching fraction and resonant fractions. The calculation of the latter is based on the resonant contribution to the total branching fraction, thus the resonant branching fraction is *not* corrected for $Br(K^{*\pm} \to K^0\pi^{\pm}) = 2/3$.

The systematic uncertainty in these results are due to systematic differences in the efficiency correction and the correction for the $K^{*\pm}$ mass requirement and the (possible) non-resonant fraction under the $K^{*\pm}$ BW peak. The first is estimated to be smaller than 2.0% of the fractions, since it is not the systematic error from the efficiency correction but only systematic differences that are of importance. The second contribution was determined to be 1.0% for the mass cut and 3.6% (5.0%) for the non-resonant contribution correction for the $B^0 \to D^{\pm}K^0\pi^{\mp}$ ($B^0 \to D^{*\pm}K^0\pi^{\mp}$) mode.
Overall this means that the systematic error on the fractions is 0.04 and 0.05, respectively.

[80]The name "non-resonant" is chosen out of ease, as the statistics is not adequate for determining the origin of signal outside the $K^{*\pm}$ resonance.

12.3 Dalitz distributions

It takes about 1000 events to fit a Dalitz plot. [Brian Meadows]

While almost all two-body B decays are measured, little is known about three-body decays. Efforts have started, but true signal Dalitz plot distributions are still in their infancy. Among the reasons is the problem of background subtraction, which is delicate both from a physics and statistical point of view, due to correlations, efficiency corrections and propagation of errors. A simple, efficient, and transparent method of solving all of the above problems is the use of ${}_s\mathcal{P}$lots.

12.3.1 Properties of ${}_s\mathcal{P}$lots

The ${}_s\mathcal{P}$lot weights $W_{\rm sig}(m_{\rm ES}, \Delta E, \mathcal{F})$ were introduced in Section 12.1, where they were used for the branching fraction calculation. Their difference from ordinary probabilistic weights is that they include correlations between the components, which has the consequence that they can take values outside the usual probability range $[0, 1]$.
Their advantage is that by weighting each event by this weight, the "true" data signal distribution is obtained for *any* quantity, which can be used both for illustration and checks. A special case of this is for checking variables, which are used in the likelihood fit. For these, the fit is repeated *excluding* this variable, and weights are recalculated. The data signal distribution using these weights can then be compared to the signal PDF of this variable (from the original fit). This gives an unbiased and visual comparison between the "true" data signal distribution and its PDF description[81] (see Section 13.2). Such plots are extremely useful for detecting contributions, which are not accounted for by the PDF, but which are "hidden" in the background distribution. In addition, the value and statistical error of any subregion (e.g. bin) can be calculated simply from the sum and the square-root of the sum of the squares, respectively. This makes ${}_s\mathcal{P}$lots a simple but powerful tool, without which three-body analyses would be much more complicated.

12.3.2 Resolution in Dalitz plot

The resolution in the Dalitz plot is examined using signal MC. For each event the difference between the generated and reconstructed value in m^2_{DK} and $m^2_{K\pi}$ are recorded. These residuals are shown as a function of m^2_{DK} and $m^2_{K\pi}$ in Fig. 12.1.

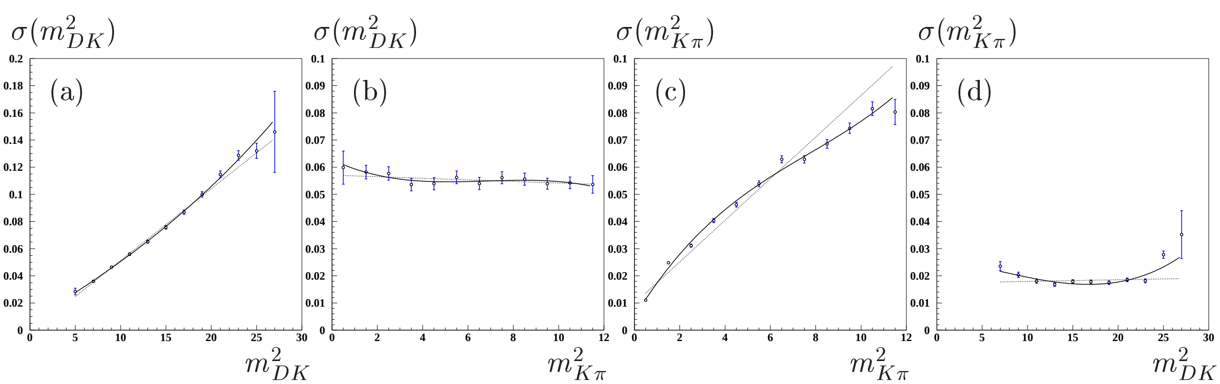


Figure 12.1: Resolution on invariant mass squared in Dalitz plot as a function of m^2_{DK} and $m^2_{K\pi}$. The resolution depends approximately linearly on the invariant mass squared in question ((a)+(c)) as expected (see text), while it is independent of the other invariant mass ((b)+(d)).

[81] If the variable in question was included in calculating the weights, it would bias these to resemble its PDF.

As can be seen from the figures, the resolution in m^2_{DK} and $m^2_{K\pi}$ are to a fair approximation a linear function of the square of the invariant masses themselves and independent of the other invariant mass. The reason for this is the approximate rule (due to equal spatial tracking resolution), $\sigma(1/p_\perp) \sim \sigma_0$, where σ_0 is a constant resolution specific for each tracking system (see Eq. (8.1)). Given this, one has that:

$$p_i = \frac{p_{i\perp}}{\sin\theta_i}, \qquad \sigma_{p_i} = \frac{1}{\sin\theta_i}\sigma(p_{\perp i}) = \frac{1}{\sin\theta_i}\frac{1}{(1/p_{\perp i})^2}\sigma(1/p_{\perp i}) = \sin\theta_i p_i^2 \sigma_0. \tag{12.5}$$

With the invariant mass expression $m^2_{12} \simeq m^2_1 + m^2_2 + 2p_1p_2(1-\cos\theta_{12})$ and neglecting the uncertainty on the masses and angles, this yields:

$$\sigma(m^2_{12}) \;\simeq\; 2(1-\cos\theta_{12})(p_2\sigma_{p_1} \oplus p_1\sigma_{p_2}) \;\simeq\; m^2_{12}\,\langle\sqrt{p^2_{\perp 1}+p^2_{\perp 2}}\rangle\,\sigma_0. \tag{12.6}$$

In general, the position resolution in the Dalitz plot is quite good due to the lack of unconstrained neutral particles. In addition, essentially no SCF falls in a different position of the Dalitz plot than the true decay. Since the resolution is smaller than the resonant structures looked for, it can essentially be neglegted. However, for narrow resonances the resolution is essential, and should one wish to fit for interferences in the Dalitz plot involving sharp resonances, the resolution should be included.

12.3.3 Signal Dalitz distribution

Weighting each event by the weight $W_{sig}(m_{\rm ES}, \Delta E, \mathcal{F})$ described in Eq. (12.2) and dividing by the efficiency at each given point, the Dalitz distribution of the *signal* can be obtained.
In Figure 12.2 is shown the efficiency, background, and correlation corrected $B^0 \to D^\pm K^0 \pi^\mp$ signal distribution in the Dalitz plot along with projections onto the three invariant masses, to display the errors (which are hard to show in the Dalitz plot), and reveal potential resonant substructures, care being taken of reflection effects.
The Dalitz distribution clearly shows that the $K^{*\pm}$ resonance is dominant, as expected, and the characteristic spin-1 helicity shape ($dN/d\cos\theta \propto \cos^2\theta$, cf. Section 12.3.4) is clearly seen. Other contributions are somewhat scattered, the largest part falling in the bottom left corner (large $D\pi$ invariant mass), but it is not significant. For the sake of clarity, bins in the Dalitz plot where the sum of the event weights is negative are indicated as if there were no entries. The lack of statistical errors in the Dalitz plot makes it hard to evaluate the significance of seeming features.
No significant structure is seen in the m_{DK} distribution, as the two "peaks" at 3.6 and 5.1 GeV in Fig. 12.2b are simply due to the helicity structure of the dominant $K^{*\pm}$ resonance projected onto the m_{DK} axis (so-called reflections). The contribution at high $D\pi$ invariant mass is visible in the projection. If significant, it would be the signature of the decay $\overline{B}^0 \to D^{**0}K^0_S$, where the D^{**0} is a wide heavy resonance decaying to $D^+\pi^-$. The $m_{K\pi}$ distribution clearly shows the $K^{*\pm}$ peak, while other structures are not significant. A fit with a relativistic Breit-Wigner to the peak yielded $m_{K^{*\pm}} = 900.2 \pm 6.3$ MeV and $\Gamma_{K^*} = 51.5 \pm 10.5$ MeV, in accordance with the parameters of the $K^{*\pm}$ resonance values.

For the $B^0 \to D^{*\pm}K^0\pi^\mp$ mode, the efficiency, background, and correlation corrected signal Dalitz plot and its projections are shown in Figure 12.3.
Once again the Dalitz distribution clearly shows that the $K^{*\pm}$ resonance is dominant. An interesting contribution at very high $K^0_S\pi$ invariant mass can be seen, however not significant. Contrary to the $B^0 \to D^\pm K^0\pi^\mp$ Dalitz plot, no sign of contributions at high $D\pi$ invariant mass is seen, while the lower end of the spectrum is more populated.
No significant structures are seen in the m_{DK} distribution, nor in the $m_{D\pi}$ distribution, where the hint of a (more narrow) D^{**} resonance seemed suggested in the Dalitz plot.

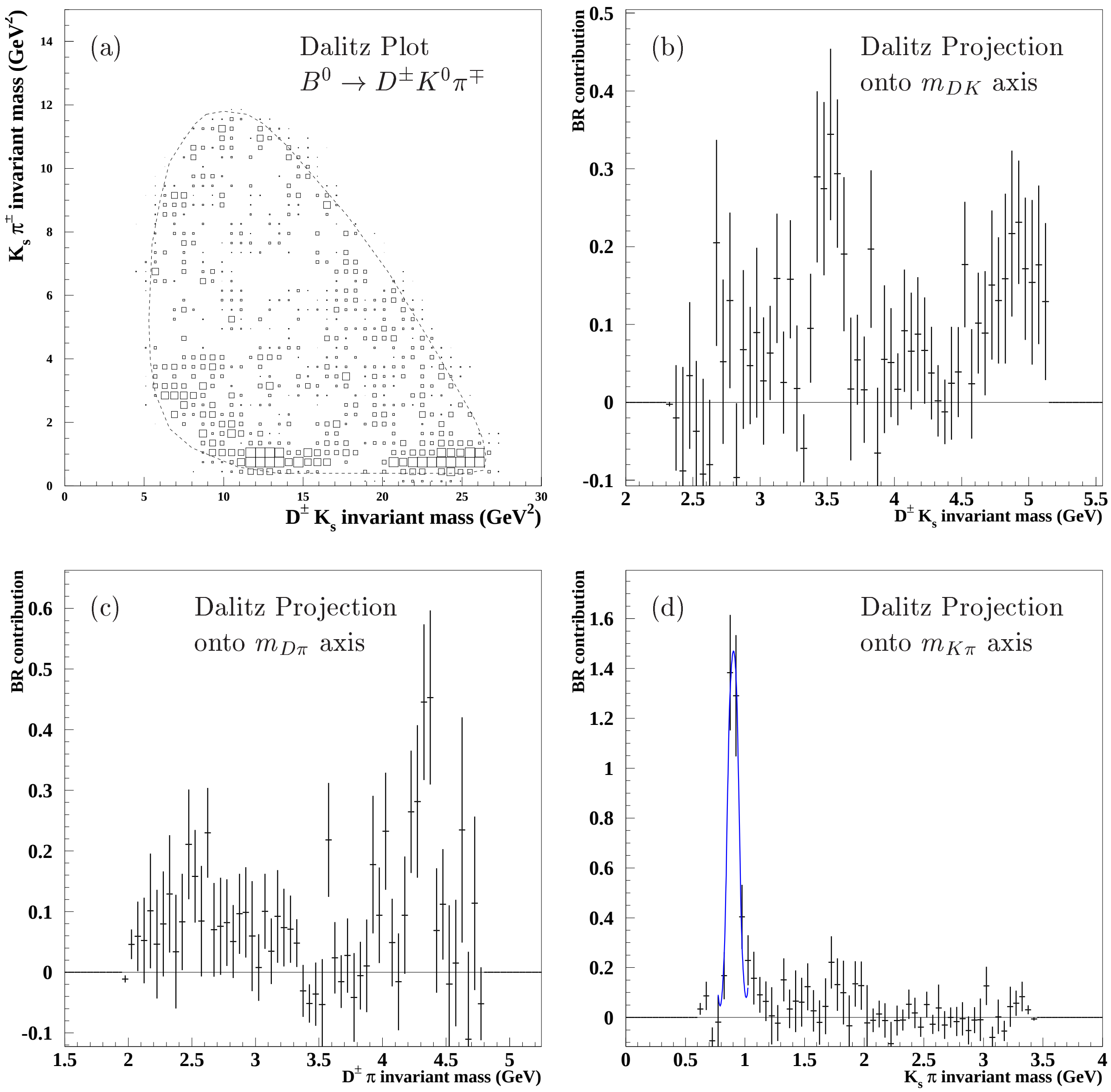


Figure 12.2: $B^0 \to D^\pm K^0 \pi^\mp$ data signal Dalitz distribution and projections onto the three invariant masses after efficiency, background, and correlation corrections. (a) The Dalitz distribution shows the dominant $K^{*\pm}$ resonant contribution. The dashed line shows the approximate Dalitz region limit. The projections of the Dalitz plot are shown in (b) onto m_{DK}, (c) onto $m_{D\pi}$, and onto $m_{K\pi}$ axis. The fit in the last plot is a relativistic Breit-Wigner fit to the $K^{*\pm}$ resonance.

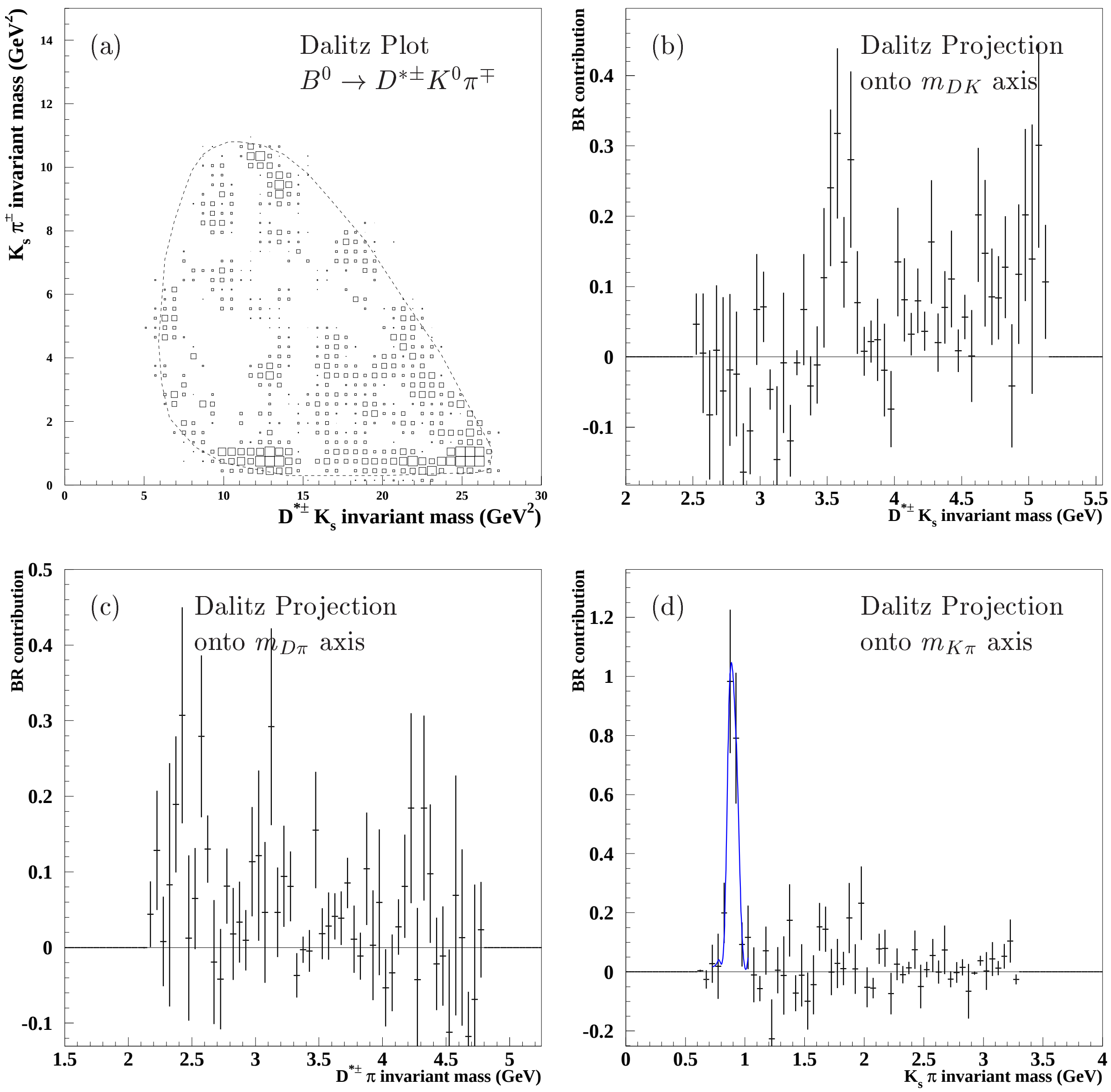


Figure 12.3: $B^0 \to D^{*\pm} K^0 \pi^\mp$ data signal Dalitz distribution and projections onto the three invariant masses after efficiency, background, and correlation corrections. (a) The Dalitz distribution exhibits the dominant $K^{*\pm}$ resonant contribution. The dashed line shows the approximate Dalitz region limit. The projections of the Dalitz plot are shown in (b) onto m_{DK}, (c) onto m_{DK}, and onto $m_{K\pi}$ axis. The fit in the last plot is a relativistic Breit-Wigner fit to the $K^{*\pm}$ resonance.

12.3.4 Partial wave analysis

In the $B^0 \to D^{\pm}K^0\pi^{\mp}$ channel, the helicity of the spin-1 $K^{*\pm}$ resonance follows the distribution $dN/d\cos\theta \propto \cos^2\theta$, where θ is the angle between $K^{*\pm}$ and the K^0 in the $K^{*\pm}$ CM frame, since both the B^0 and the $D^{\pm}$ mesons have spin 0. The $B^0 \to D^{*\pm}K^{*\mp}$ channel is more complicated, as there are three helicity amplitudes to be separated by angular analysis (mentioned in Section 5.2), and it will therefore not be regarded here.
If the $K^{*\pm}$ resonance interferes with other contributions, which have other spin components, then the helicity distribution of the $K^{*\pm}$ is altered from its original distribution. One can therefore look for such components by considering the helicity distribution of the $K^{*\pm}$. The signal helicity distribution of the $K^{*\pm}$ and the higher mass region $1.0 < m(K_S^0\pi) < 2.0\,\text{GeV}$ is shown in Fig. 12.4 along with fits to $dN/d\cos\theta \propto \cos^2\theta$ and a general polynomial of second order.

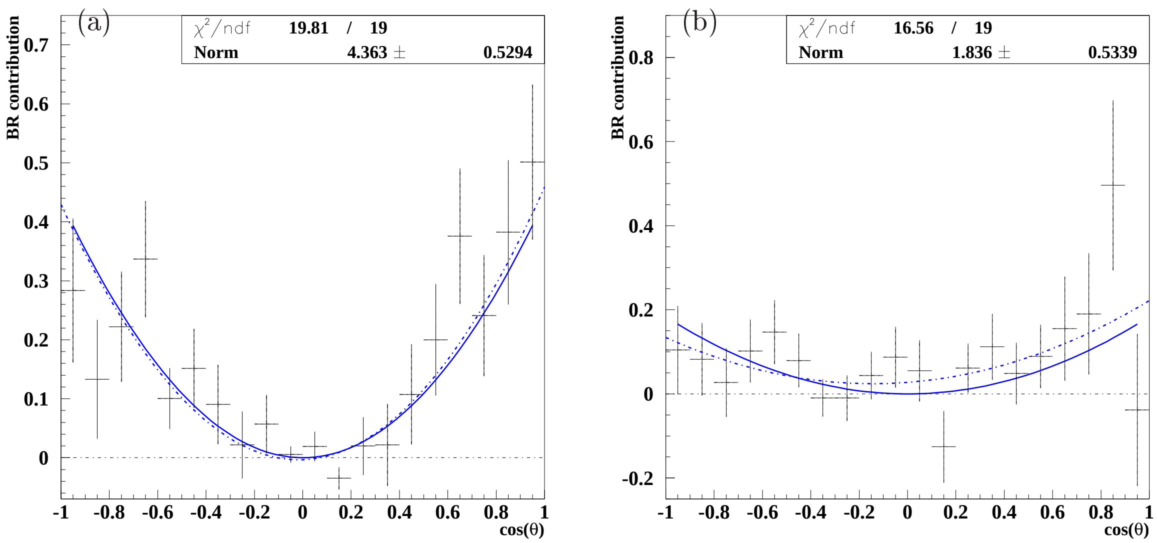


Figure 12.4: Signal helicity distribution for the $K^{*\pm}$ and the higher mass region $1.0 < m(K_S^0\pi) < 2.0\,\text{GeV}$. The solid curves are fits to the spin 1 distribution $dN/d\cos\theta \propto \cos^2\theta$, while the dashed curves are fits to general polynomial of second degree.

As can be seen from the fits, the $K^{*\pm}$ helicity distribution follows the spin-1 prediction with no sign of interference. The higher mass region distribution has too little statistics to reveal anything about the underlying spin structure. A further investigation of the $K^{*\pm}$ helicity distribution and its possible interference with other components require the division into slices of mass, for which the statistics are not sufficient.
To check that the helicity distribution isn't affected by an artifact of low momenta K_S^0 mesons ($\cos\theta < 0$), the width of $m(K_S^0)$ is measured for each end of the spectrum. No difference is found.

13 Validation

Now we can conclude that the number of sign errors is even, that the sum of the biases is small, and that the bugs do accidentally not influence the results too much.

[Upon having completed the validation]

In addition to the various test and checks before applying the analysis to the data, a posteriori validations of the results are performed. The validation is done in two way.
The first is the classic method of using toy MC studies, which through repeating the fit many times can detect irregularities in the fitting parameters and can provide a basis of comparison for the value of the data likelihood.
The second method used is not classic (yet), and uses a novel statistical method called ${}_s\mathcal{P}$lot, which was introduced in Section 12.1. Using the covariance matrix from the likelihood fit, the signal (and background) distributions are extracted from the data itself (in an optimal manner), such that it can be compared to the PDFs from the fit.

13.1 Toy MC studies

One way to test a likelihood fit is by generating many similar distributions (named toy MC) and then repeat the fit on these samples to see if the values and errors obtained distribute themselves correctly, and if the likelihood of the fit to data is probable.
Given the PDFs and yields obtained from the likelihood fit, 250 toy MC experiments were generated and fitted in the same manner as the actual fit: The number of events of each component were chosen according to a Poissonian distribution with the mean being the yield extracted from the fit, and their distribution in the three fitting variables following that of the PDFs obtained from the fit. For each toy experiment, the result in terms of yields, fitting parameter values and errors, and the likelihood were recorded.

Since the likelihood value does not in itself carry any information (one can always add a constant to the likelihood), it is very useful to repeat the fit on toy MC and obtain a distribution of likelihood values with which the value from the fit to data can be compared. The distribution of likelihood values obtained from the fitted toy MC is shown in Figure 13.1.

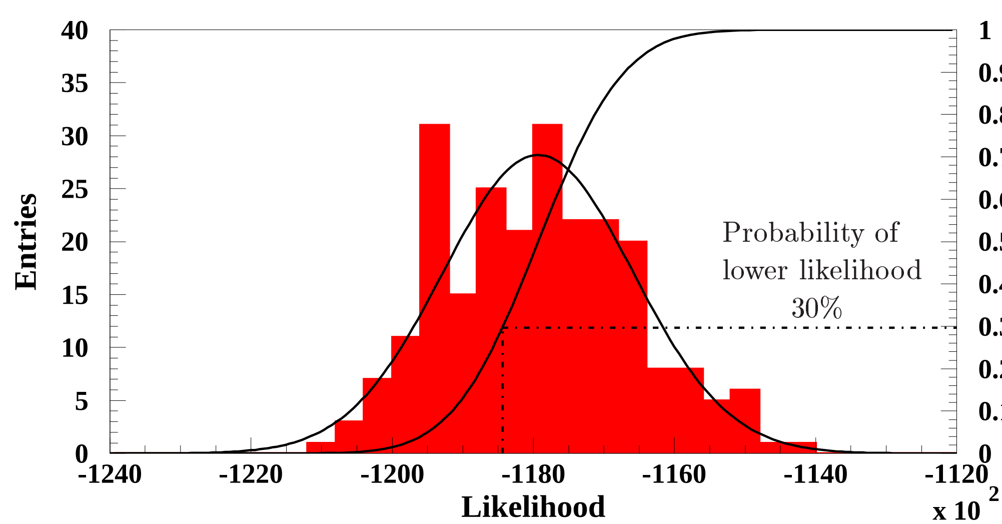


Figure 13.1: Distribution of likelihood from toy MC simulations. The distribution has been fitted with a Gaussian. The value obtained in the fit to data was -118439 (vertical line), which 30% of the toy MC distribution falls below (horizontal line).

The value of the likelihood from the fit on data falls in the central part of the distribution, which indicates that the PDFs describe the data well. The chance of obtaining a smaller likelihood is 30%.

The results for the parameters and yields are shown using pull distribution, that is the difference between the fitted and the true value divided by the error on the fitted value, $(x_{fit} - x_{true})/\sigma(x_{fit})$. Such distributions should be unit Gaussians, that is Gaussian distributions with a mean of zero and a width of one. If the mean of the pull distribution, μ, falls different than zero, it means that the parameter is biased in the fit, i.e. it consistently takes a value different from the true value. If the width of the pull distribution, σ, is not one, it means that the error obtained from the fit is overestimated ($\sigma < 1$) or underestimated ($\sigma > 1$). The pull distributions for the five floating background parameters are shown in Fig. 13.2.

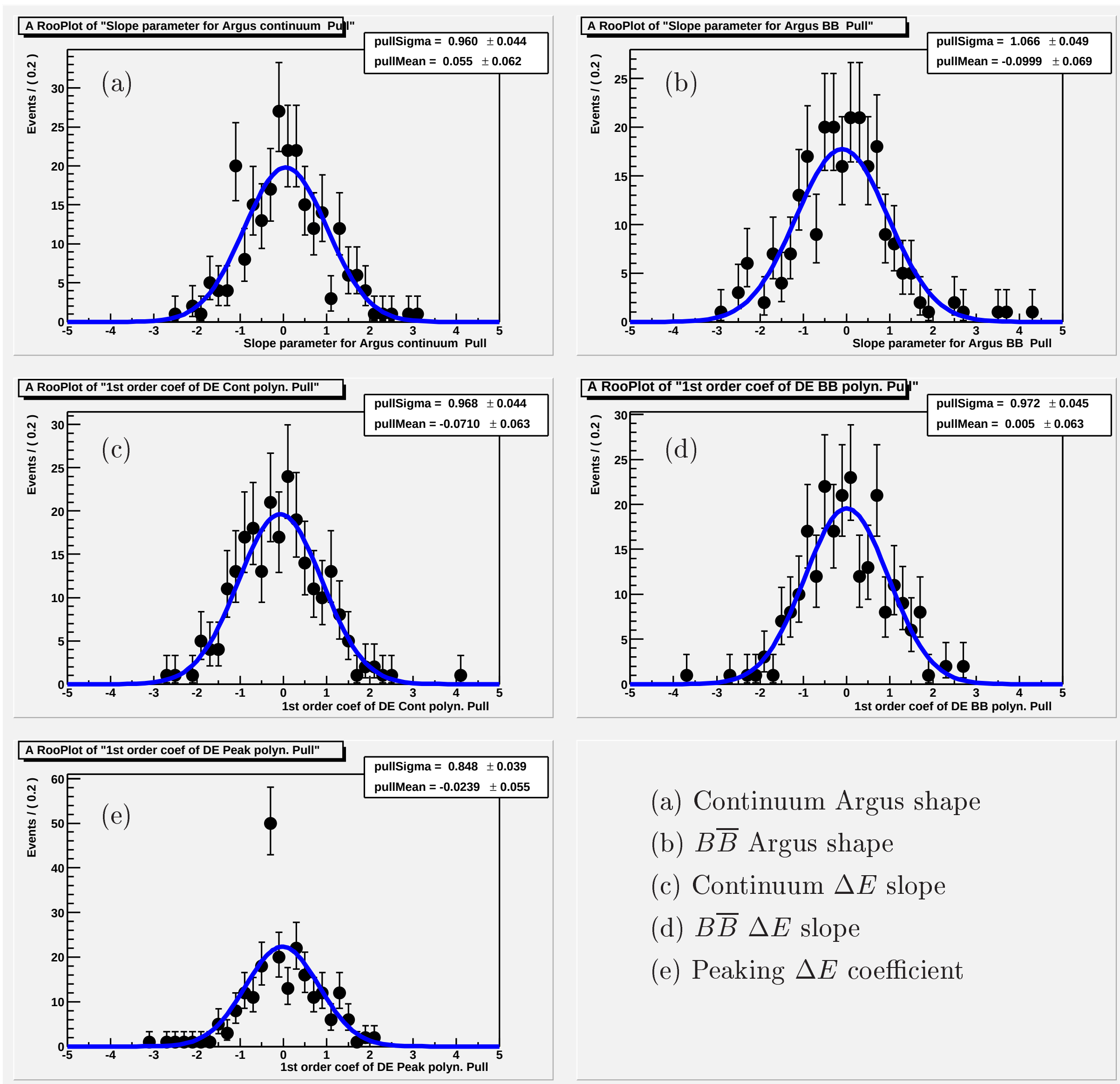


Figure 13.2: Pull distribution for the floating background parameters. The plots are (a) continuum Argus shape, (b) $B\overline{B}$ Argus shape, (c) continuum ΔE slope, (d) $B\overline{B}$ ΔE slope, and (e) peaking ΔE coefficient. All means are consistent with zero (no biases) and all widths are consistent with unity (correct error estimation). The one deviating bin in the peaking ΔE coefficient pull distribution (e) is due to a computational artifact.

From the toy MC pull distributions of the parameters describing the background, which are left free in the fit, it is apparent that these are all well determined with correct errors in the fit. The one deviating bin in the peaking ΔE coefficient pull distribution (see Fig. 13.2e) is due to a computational artifact (initial value sometimes not floated!).

The pull distributions of the four yields are shown in Fig. 13.3.

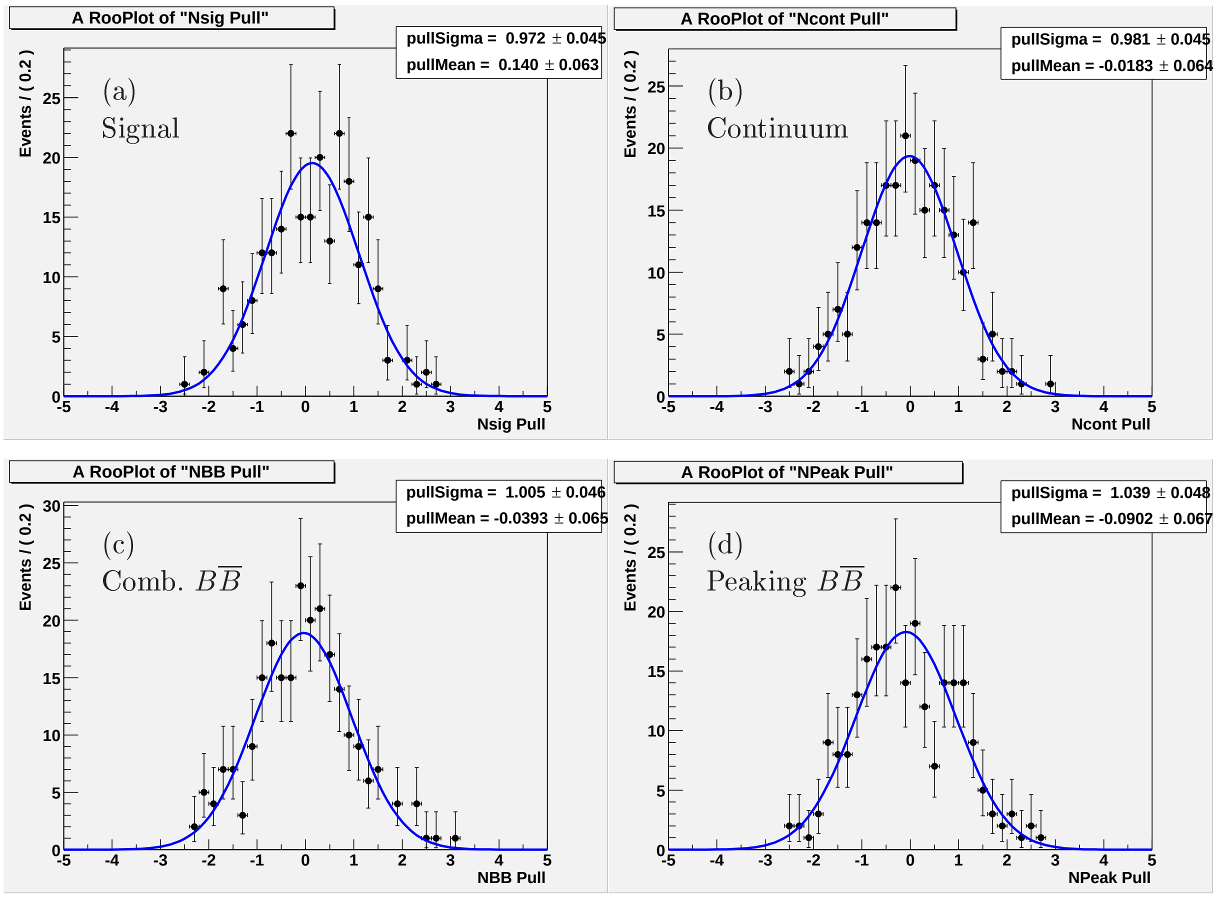


Figure 13.3: Pull distributions for the four yields. The plots are for (a) signal, (b) continuum, (c) combinatorial $B\overline{B}$, and (d) peaking $B\overline{B}$. All means are consistent with zero (no biases) and all widths are consistent with unity (correct error estimation).

The pull distributions for the event yields of each component again show the behaviour of unbiased fit parameters with a correct error evaluation. Even the peaking background yield, which is small and has a large error does not show any sign of misbehaviour. The pull distribution for the signal yield is most important, as the branching fraction measurement depend directly on the signal yield.

In order to test possible correlations in the signal, the exercise was repeated, but this time using signal MC for the signal events, while the backgrounds were still generated from PDFs. In this way one can test if any correlations not accounted for by the fit (e.g. the correlation between m_{ES} and ΔE) biases the result. The result of this study can be found in Fig. 13.4, and no biases are found in the pull distributions of the four yields.
However, as each event contributes roughly the same to the likelihood, the distribution of these is dominated by the background, and it will be less sensitive to signal discrepancies.
Once again the pull distributions for the yields are in accordance with unit Gaussian distributions, and no biases or excessive tails are seen.

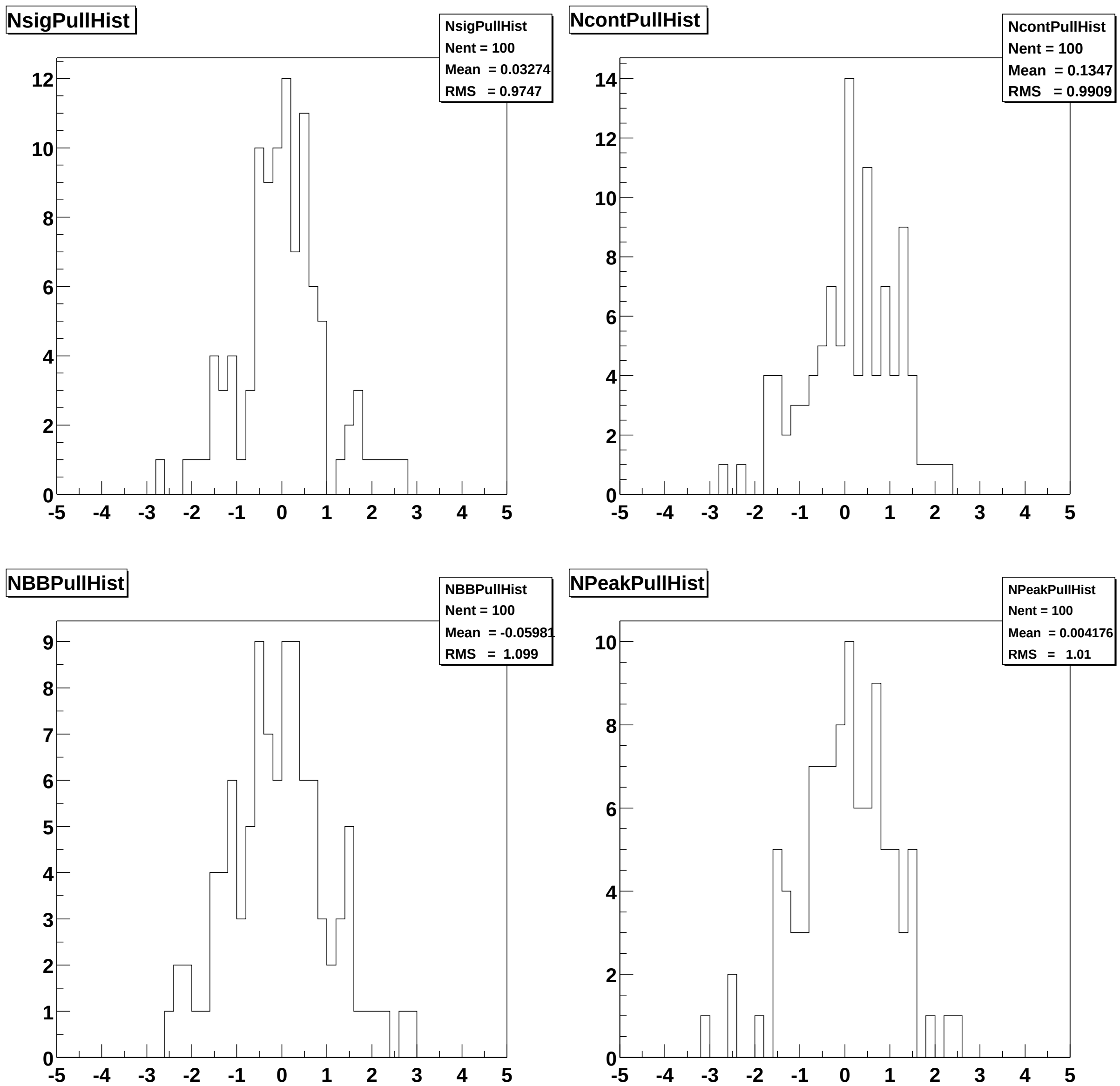


Figure 13.4: Pull distributions for the four yields when using SP4 signal MC for the signal. The plots are for signal (upper left), continuum (upper right), $B\overline{B}$ (lower left) and peaking (lower right).

13.2 ${}_s\mathcal{P}lot$ validation

As mentioned in Section 12.1, the ${}_s\mathcal{P}$lot weights $W_{\rm sig}(m_{\rm ES}, \Delta E, \mathcal{F})$ have the property, that any quantity weighted by them will show the background subtracted, correlation corrected data signal[82] distribution (including corresponding errors). Such a feature can be used for checking that the PDFs indeed describe the data correctly or that other distributions come out as expected.
To check the match between a distribution of a variable and the corresponding PDF, one calculates the ${}_s\mathcal{P}$lot weight *omitting the variable in question* (if the variable is included, it will bias the data distribution to match the PDF), and plots the distribution with the PDF overlayed. In this manner all other (background) components have been subtracted correctly, and the comparison becomes direct and visual. Such plots, called ${}_s\mathcal{P}$lots, are shown in Fig. 13.5 and 13.6 for $B^0 \to D^\pm K^0 \pi^\mp$ and $B^0 \to D^{*\pm} K^0 \pi^\mp$, respectively.

As can be seen from the figures, the distributions have the expected shapes. Especially $m_{\rm ES}$ and ΔE exhibit very clearly the shapes foreseen and put into the PDFs, which is encouraging, as they are the most discriminating variables.
When fitting the $m_{\rm ES}$ distributions, the mean, width and yield are in agreement with those obtained from the fit in the analysis. The same is true when fitting the peaking component of the background.

Finally, the ${}_s\mathcal{P}$lot weights are calculated for the masses of the $D^\pm$ and K^0_S, which are not in the fit, and thus all variables in the fit can be included in the calculation. The results are shown in Figure 13.7. The distributions clearly show Gaussian behavior with resolutions in agreement with the expected values, as the overlayed fits verify. This in turn serves as a test of the ${}_s\mathcal{P}$lot weights. The fit does not have any information about the signal distribution of the $D^\pm$ and K^0_S masses, but from the separation obtained from the fit, these can be extracted from the data.

A common problem in displaying data form a likelihood fit in several dimensions, is that a simple projection onto each variable (see Figs. 10.4 and 10.5) does not show the discriminating properties of all the other variables. If a cut is applied to these, then the power of the likelihood fit is somewhat sacrificed visually, and some signal events will not be included in the plot.
With ${}_s\mathcal{P}$lots, these difficulties are avoided, as one simply shows the distribution of the ${}_s\mathcal{P}$lot weights with the corresponding PDF overlayed. However, though the purity can in principle be extracted from comparing the signal size with the errors, it does not allow for a visual demonstration. Therefore, plots with both the ${}_s\mathcal{P}$lot and the projection with a likelihood cut are shown in Fig. 13.8 for $m_{\rm ES}$.
Such plots *visually* give the signal distribution and the quality of its description by the PDF along with the level of background and thus the purity. In the limit of very high purity, the two approaches coincide, as can be seen for the pure resonant samples.

13.3 Section summary and conclusion

Systematic errors arise from a variety of corrections and variations of parameters known with a limited accuracies. The dominant sources of error are the $D^{\pm/0}$ branching fractions, tracking efficiency, and the PDF shapes, while other corrections have minor effects. Given the yields with the statistical errors along with the systematic uncertainties, the branching fractions of the three-body decays $B^0 \to D^\pm K^0 \pi^\mp$ and $B^0 \to D^{*\pm} K^0 \pi^\mp$ and their $K^{*\pm}$ resonant parts were calculated. In order to correct for the varying efficiency in the Dalitz plot, so-called ${}_s\mathcal{P}$lot weights were employed, but the scope of these weights go far beyond such a calculation. The weights are very useful for extracting the Dalitz distribution of the signal, which were also shown. For the validation both the usual method using toy MC and the new method using ${}_s\mathcal{P}$lots were used. Both methods showed no sign of biases or incorrect error estimation.

[82] Though most often applied to signal, it can equally well be done for any other component in the fit.

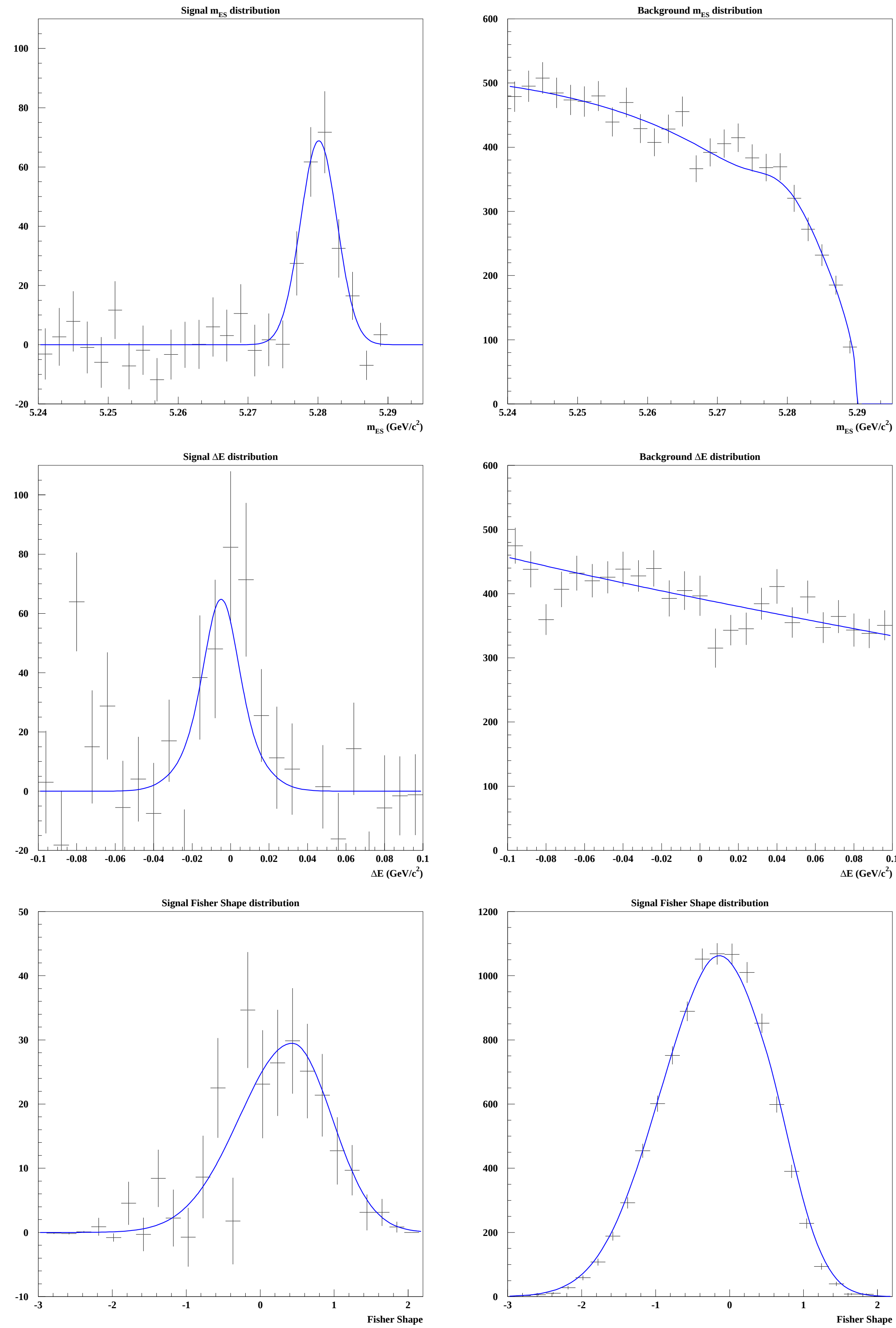


Figure 13.5: The signal (left) and background (right) ${}_s\mathcal{P}$lot $B^0 \to D^{\pm}K^0\pi^{\mp}$ distributions for the variables $m_{\rm ES}$ (top), ΔE (middle) and $\mathcal{F}$ (bottom). Superimposed is the PDF obtained from the fit.

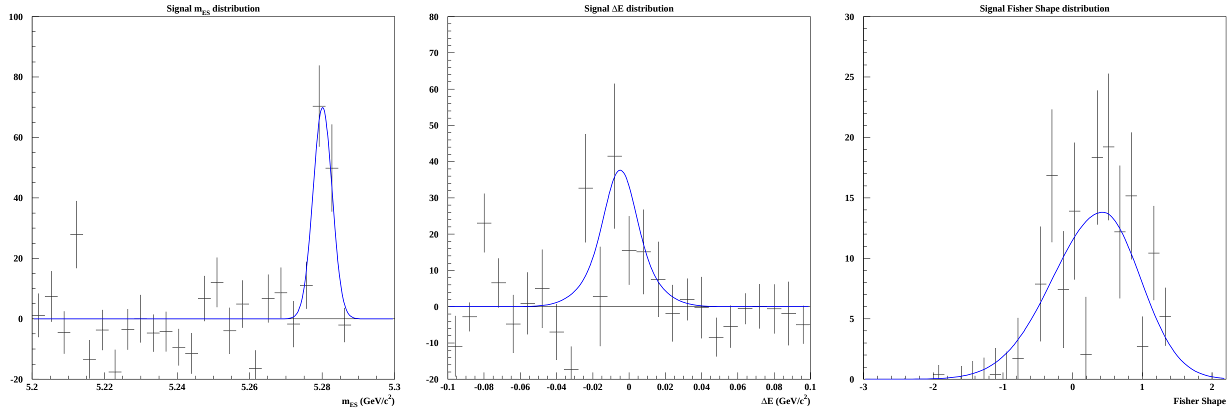


Figure 13.6: The $B^0 \to D^{*\pm} K^0 \pi^\mp$ ${}_s\mathcal{P}$lot distribution for the variables $m_{\rm ES}$ (left), ΔE (middle), and $\mathcal{F}$ (right). Superimposed is the PDF obtained from the fit.

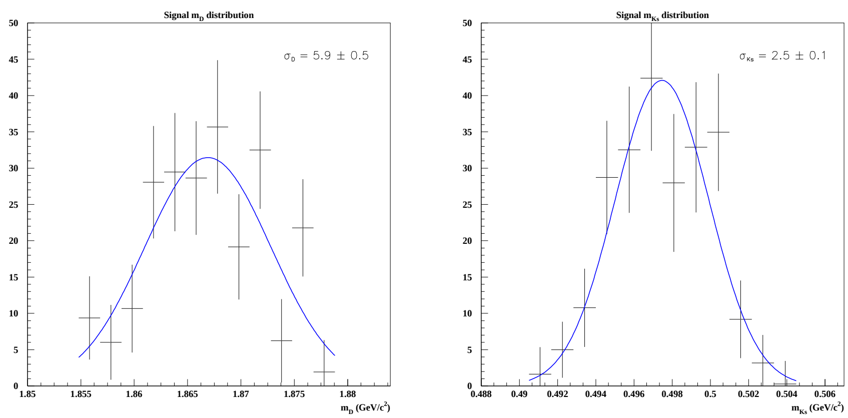


Figure 13.7: The signal ${}_s\mathcal{P}lot$ distributions for the $D^\pm$ (left) and K^0_S (right) masses, fitted with a Gaussian distribution.

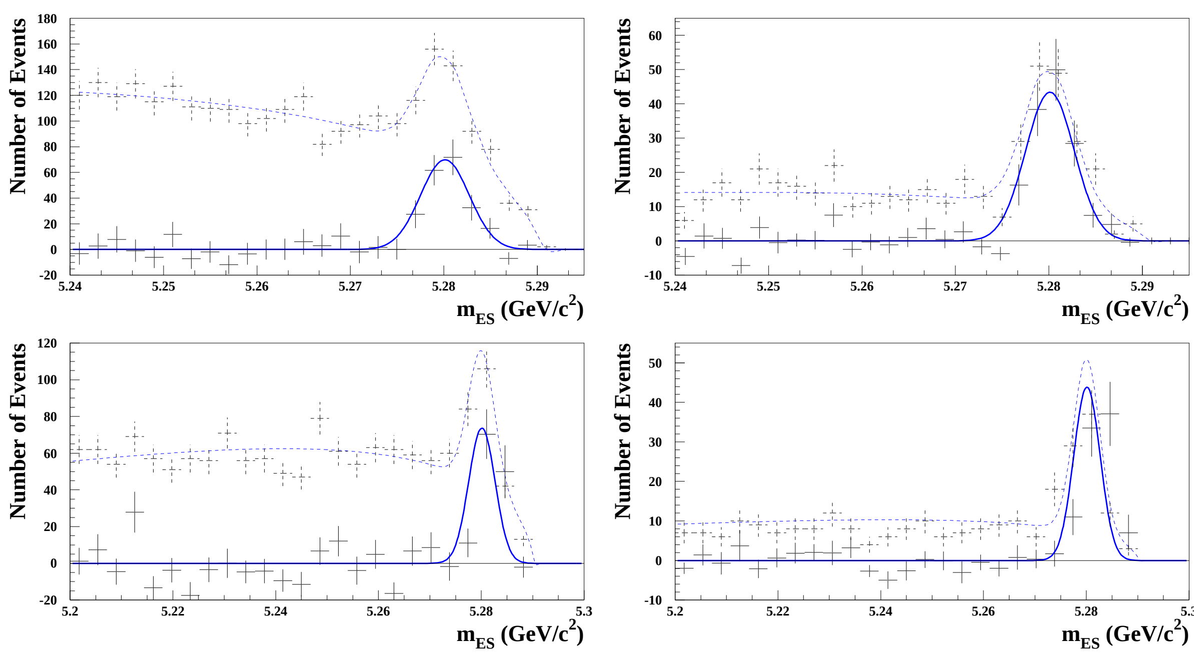


Figure 13.8: The $B^0 \to D^{(*)\pm} K^0 \pi^\mp$ ${}_s\mathcal{P}$lot distribution in $m_{\rm ES}$. The solid entries and line shows the signal contribution and PDF in $m_{\rm ES}$ and serves as a visual check of this PDF component, while the dashed entries and line is a fit to the $m_{\rm ES}$ distribution after applying cut on ΔE (see text).

14 Time-dependent analysis

The biggest difference between time and space is that you can't reuse time.

[Merrick Furst]

14.1 Time-dependent analyses ingredients

The time distributions of B^0 and $\overline{B}^0$ decays to a common final state may reveal CP violation through the interplay between mixing and decay (see Section 4.5.3). However, since the final state of interest can be reached by both b quark flavors (otherwise there would be no interference), one is required to determined that of the decaying B^0 or $\overline{B}^0$ by other means. To obtain this crucial information, the flavor of the other B^0 meson in the $\Upsilon(4S)$ decay has to be determined along with the time difference between the two decays, since the correlation between the two flavors is time-dependent (see Section 4.4.2). As only a small fraction of the B decays are fully reconstructed, one has to determine these two quantities from incomplete information[83], which of course induces errors.

The situation is schematically illustrated in Fig. 14.1. From an asymmetric e^+e^- collision at the $\Upsilon(4S)$ resonance, one creates a boosted $B^0\overline{B}^0$ pair. One of the two B mesons decays into the final state of interest (here $D^-K^0_S\pi^+$), and from the remaining tracks and clusters, both the flavor and the decay vertex of the other B^0 meson has to be determined. The drawing is simplified, as generic B decays on average have 5.5 charged and an equivalent number of neutral particles in the final state [Har98].

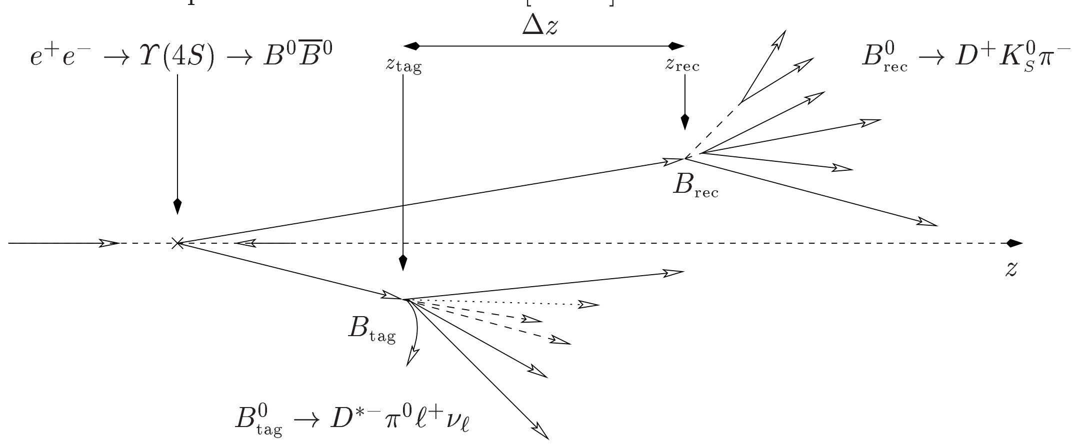


Figure 14.1: Example of $B^0\overline{B}^0$ decay with $D^-K^0_S\pi^+$ final state. To fit the CP asymmetry in the lifetime distributions, both the flavor and the decay vertex of the other (tagging side) B^0 meson has to be determined. Note that the drawing is schematic and not to scale.

14.2 Determination of B^0 flavor – tagging

Given a $B^0\overline{B}^0$ event, where one B meson is fully reconstructed (denoted B_{rec}), the ability to determine the flavor and decay vertex of the other neutral B meson (commonly referred to as the tagging side B meson and denoted B_{tag}) plays a central role in time-dependent analyses. If the tagging side B meson could be fully reconstructed, the task would be almost trivial. But the efficiency of full reconstruction is quite low ($\sim 10^{-3}$), which means that one is forced to take an inclusive approach. Thus, the decay vertex and flavor of the tagging side B meson has to be determined from considering the remaining tracks, whichever part of the tagging side B decay they may constitute.

[83] Mostly due to limited acceptance, but also because of particle misidentification and machine backgrounds.

Though the variety of B decays is very rich, and not all decay modes carry flavor information, some characteristic properties offer general flavor features, which can be recognized by algorithms. Due to the incompleteness of the information available, the outcome is not always correct, and in general two quantities characterize such an algorithm:

- The efficiency of yielding a B^0 ($\overline{B}^0$) flavor tag, ϵ ($\overline{\epsilon}$).
- The probability of a B^0 ($\overline{B}^0$) tag indicating the *wrong* flavor, ω ($\overline{\omega}$).

Due to differences in the interaction cross section of particles and anti-particles with the detector, the tagging efficiency and the wrong tag (mistag) fraction need not be exactly the same for B^0 and $\overline{B}^0$. This is parametrized as $\langle\omega\rangle = \frac{1}{2}(\omega + \overline{\omega})$ and $\Delta\langle\omega\rangle = \omega - \overline{\omega}$. A common notation is $D \equiv (1 - 2\langle\omega\rangle)$ and thus $\Delta D = -2(\omega - \overline{\omega})$.
The key figure of merit for the tagging performance is $Q = \epsilon(1 - 2\langle\omega\rangle)^2$, which is called the effective tagging power[84]. This expression can be derived by taking the derivative of the likelihood of time-dependent fit, and represents the quantity with which the statistical uncertainty in time-dependent asymmetry measurements scales: $\sigma \propto 1/\sqrt{Q}$. Q essentially measures what fraction of the events before tagging can actually be used as being perfectly tagged[85], hence the common name "effective tagging efficiency".
In order to measure CP asymmetries, a high but also well determined and well understood tagging power is required, as its uncertainty propagates directly to the asymmetry error.

14.2.1 Sources of B flavor information

The flavor information of B mesons is contained in the correlations between charge, particle identity and kinematic properties of the decay products. The main sources of information are the charge of primary leptons (from b quark decays) and/or kaons, but other indicators are the charge of slow pions (from $D^{*\pm}$ decays) and secondary leptons (from c quark decays).

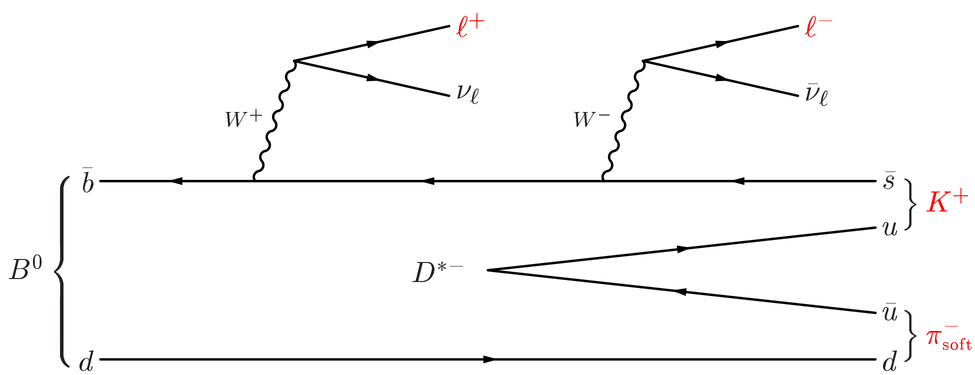


Figure 14.2: Diagram of decay B^0 decay showing general sources of flavor information. The decay $B^0 \to D^{*-}\ell^+\nu$ with the subsequent decays $D^{*-} \to \overline{D}^0\pi^-_{\text{soft}}$ and $\overline{D}^0 \to K^+\ell^-\bar{\nu}$ is an example of a useful tagging decay, as B flavor information can be obtained from the charge of the direct lepton, the cascade lepton, the kaon, and the soft pion (see text). Though an example, the features here shown (in red) are common for many B decays.

Each of these sources can be used as an algorithm by itself or the information can be combined into a more powerful tagging algorithm. However, as there is a certain correlation between the various sources of information, the combination is desirable.

[84] As it is a central quantity, it has many names, e.g. quality factor, absolute separation, and tagging efficiency.
[85] That is N events before tagging will correspond to QN perfectly tagged events.

14.2.2 Leptons

The semi-leptonic B^0 decays, $B^0 \to X\ell^+\nu_\ell$, $(\ell = e, \mu)$, constitute $21.0 \pm 1.6\%$ [PDG02] of the total B^0 width, and represents the main source of lepton tagging power. The $\bar{b}$ quark decays via a W^+ boson to a positive lepton ℓ^+ signifying a B^0 decay see Fig. 14.2.
Other sources of leptons, which may increase the mistag fraction, $\langle\omega\rangle$, are secondary leptons from D meson decays, and (more rarely) leptons from vector meson decays (e.g. $J/\psi \to \mu^+\mu^-$), and kaon/pion decays. However, these have much softer momentum spectra (essentially stopping at $p_\ell^{\mathrm{CM}} = 1.4\,\mathrm{GeV}$), which algorithms can use not only to purify the primary lepton sample, but also to gain additional tagging power from secondary leptons. The momentum of the lepton in the CM frame is combined with the energy measured in its hemisphere and the angle between the lepton and the missing momentum to form a lepton tag.
Hadrons mis-identified as leptons (so-called *fake leptons*) can potentially increase $\langle\omega\rangle$ for leptons, but with tight selections the problem becomes close to negligible, especially for electrons.

14.2.3 Kaons

The majority of b quark decays follow the (cascade) decay chain $b \to c \to s$ (see Fig. 14.2), which leads to a correlation between the kaon charge and the B flavor (i.e. a K^+ signifies a B^0). Neutral kaons do not carry any specific B flavor information. As opposed to these *right sign* kaons, charged kaons of both right and wrong sign are produced in other processes as well. The multiplicities are $n(B^0 \to K^+X) = 0.570 \pm 0.025 \pm 0.024$ and $n(B^0 \to K^-X) = 0.187 \pm 0.017 \pm 0.010$ [Tri01], which means that in most cases the charge of a kaon will give a correct tag. Contrary to the lepton case, the momentum distributions of right and wrong kaons are almost identical and carry no useful information.
In the presence of several charged kaons, the information of these is combined. The kaon mis-identification rate also has an impact on the tagging performance, and therefore PID based on likelihood ratios (i.e. continuous) is used.

14.2.4 Soft pions

Slow pions from $\overline{B}^0 \to D^{*+}(\to D^0\pi^+_{\mathrm{slow}})X$ decays are another source of tagging information, this time with a positive charge signifying a $\overline{B}^0$ meson (see Fig. 14.2). As the slow pion and the D^0 are produced nearly at rest in the D^{*+} CM, the slow pion direction in the B frame should be along the line of direction of the D^0 daughters and the rest of the B decay products. This direction is to a good approximation that of the thrust axis of the B meson, and thus the angle between the slow pion momentum and the thrust axis can be used for discrimination. This information is combined with the CM momentum and the PID of the soft track, to decrease the substantial background and infer the most tagging information possible.

14.2.5 Other sources

In addition to the above defined tags, the charge of the highest momentum track can be used to recognize pions from two-body decays, e.g. $B^0 \to D^{(**)-}\pi^+$ and recover high momentum leptons missed by the more exclusive lepton tag. The tracks used for this very inclusive tag are required to have a transverse impact parameter of less than 1mm, as the tracks should be prompt.

14.2.6 Combining information

The information from these sub-algorithms (which in the implementation are neural nets) are combined via a second global neural net algorithm (the so-called Moriond tagger) to form a tag [Ber02]. If the kaon tag does not agree with the lepton tag (i.e. they have opposite

charges), the event is retained for further analysis. The kaon and slow pion tag have a strong angular correlation. If they agree (have opposite charges), the angle between the two are used to refine the selection.
The output of each (possibly combined) sub-algorithm is evaluated in terms of their wrong tag fraction, $\langle\omega\rangle$, and those with similar values are grouped into one of four hierarchical mutually exclusive categories with the names `Lepton` (`L`), `Kaon I` (`K`$_1$), `Kaon II` (`K`$_2$), and `Inclusive` (`I`).

14.2.7 Why tagging categories?

The wrong tag fractions, $\langle\omega\rangle_j$, are expected to vary noticeably between the four tagging categories. This is the reason for introducing the categories in the first place, as an averaging over all tags lowers the effective tagging efficiency, $Q = \sum_j Q_j$, which can be seen as follows. Consider two types of analysis: One which uses only one tagging category, and one which uses two. Let N be the total number of CP events; N_1 in the first tagging category and N_2 in the second. The sum of the efficiencies of the two tagging categories ε_1 and ε_2 will then be the overall efficiency ε. Let $\langle\omega\rangle_i$ be the wrong tag probability of each individual event, i, and define:

$$\text{One tagging category:} \qquad D \equiv \frac{1}{N}\sum_i^N (1 - 2\langle\omega\rangle_i), \tag{14.1}$$

$$\text{Two tagging categories:} \qquad D_1 \equiv \frac{1}{N_1}\sum_{i\in N_1} (1 - 2\langle\omega\rangle_i), \quad D_2 \equiv \frac{1}{N_2}\sum_{i\in N_2} (1 - 2\langle\omega\rangle_i). \tag{14.2}$$

Obviously, $ND = N_1D_1 + N_2D_2$. The effective tagging efficiency Q for each analysis is then:

$$\begin{aligned} Q_{one} &\equiv \varepsilon D^2 = \varepsilon \left(\frac{N_1D_1 + N_2D_2}{N}\right)^2 = \frac{1}{\varepsilon}(\varepsilon_1^2 D_1^2 + \varepsilon_2^2 D_2^2 + 2\varepsilon_1\varepsilon_2 D_1 D_2), && (14.3)\\ Q_{two} &\equiv \varepsilon_1 D_1^2 + \varepsilon_2 D_2^2. && (14.4)\end{aligned}$$

The claim is that Q_{two} is greater than Q_{one}, which can be proven as follows:

$$\begin{aligned} Q_{two} - Q_{one} &= \epsilon_1 D_1^2 + \epsilon_2 D_2^2 - \frac{1}{\epsilon}(\epsilon_1^2 D_1^2 + \epsilon_2^2 D_2^2 + 2\epsilon_1\epsilon_2 D_1 D_2)\\ &= \frac{1}{\epsilon}[\epsilon_1(\epsilon - \epsilon_1)D_1^2 + \epsilon_2(\epsilon - \epsilon_2)D_2^2 - 2\epsilon_1\epsilon_2 D_1 D_2]\\ &= \frac{\epsilon_1\epsilon_2}{\epsilon}(D_1^2 + D_2^2 - 2D_1D_2) = \frac{\epsilon_1\epsilon_2}{\epsilon}(D_1 - D_2)^2 \geq 0. \quad \square \end{aligned} \tag{14.5}$$

By induction this argument can be extended to include an arbitrary number of tagging categories, and thus, if subdivision into classes with significantly different wrong tag fractions is possible, it should be done. However, there are limits to this procedure. Given finite statistics, the statistical uncertainty will at a certain point make the various dilutions overlap, and one can no longer gain information by subdividing.
In addition and very importantly, the grouping into categories should also reflect the underlying physics process, as subtle effects either from the process itself or the detector response to it needs to be well understood. An example of the former is the effect of possible (small) CP asymmetries *on the tagging side*, which are present in the hadron based tagging categories only[86]. An example of the latter is the strong (linear) correlation between the error in Δt and the wrong tag fraction in the kaon tagging categories, which is a geometrical effect of the vertexing resolution.
Note that the gain ($Q_{two} - Q_{one}$) is proportional to the square of the difference between the wrong tag fractions, thus small differences between dilution factors are not important.

[86] In the $B^0 \to D^{\mp}\pi^{\pm}$ mode (see Section 5.1) the tagging side asymmetries needs to be accounted for.

14.2.8 Performance of B flavor tagging algorithm

The performance of the Moriond tagging algorithm was evaluated on data and MC. For reasons explained earlier (see Section 9.1.1), but in particular due to the $B^0 \rightarrow D^{*-}X$ rate, which is 30% higher in the *BABAR* MC than the PDG value, the tagging performance on MC is better than that of data, and care has to be taken, when comparing the two. The performance on data was evaluated using a sample of fully reconstructed B^0 decays into states of definite flavor, applying the tagging algorithm to the ROE. From an unbinned time-dependent maximum likelihood fit to the mixing rate (Eq. (4.25)) including resolution (see Section 14.3.3), the mistag rate $\langle\omega\rangle$ and its difference $\Delta\langle\omega\rangle$ were extracted. The determination of the mistag fraction or rather $D = 1 - 2\langle\omega\rangle$ corresponds exactly to measuring the amplitude in the time-dependent mixing rate, $(N_{\text{unmixed}} - N_{\text{mixed}})/(N_{\text{unmixed}} + N_{\text{mixed}})$, modulo resolution.

Sample	Category	ϵ (%)	$\langle\omega\rangle$ (%)	$\Delta\langle\omega\rangle$ (%)	Q (%)
MC	Lepton	10.0 ± 0.1	2.8 ± 0.1	-0.7 ± 0.3	9.0 ± 0.1
	Kaon I	17.6 ± 0.1	9.2 ± 0.2	-0.9 ± 0.4	11.7 ± 0.1
	Kaon II	19.9 ± 0.1	21.2 ± 0.3	-3.0 ± 0.5	6.6 ± 0.1
	Inclusive	20.1 ± 0.1	30.9 ± 0.3	-2.3 ± 0.6	2.9 ± 0.1
	Total	67.7 ± 0.2			30.2 ± 0.2
Data	Lepton	9.3 ± 0.2	3.3 ± 0.8	-1.0 ± 1.3	8.1 ± 0.3
	Kaon I	16.4 ± 0.3	11.2 ± 0.9	-0.8 ± 1.4	9.9 ± 0.5
	Kaon II	19.8 ± 0.3	22.1 ± 1.0	-2.9 ± 1.5	6.2 ± 0.4
	Inclusive	20.2 ± 0.3	31.3 ± 1.0	-4.6 ± 1.6	2.8 ± 0.3
	Total	65.7 ± 0.6			27.0 ± 0.8

Table 14.1: Performance of the tagging algorithm on MC and data [Ber02]. The latter is obtained from fitting the time-dependent mixing rate of a sample of fully reconstructed B into a flavor specific state. Apart from the general slight overestimation, the MC resembles the data well. It is assumed that $\epsilon = \overline{epsilon}$.

The numbers listed in Table 14.1 are worthy of a few remarks. While about 2/3 of reconstructed B^0 candidates are assigned a flavor tag, the overall effective tagging power, Q, is 27%. The Lepton tagging category has the smallest wrong tag fraction, as expected, but even so, the Kaon I category has the greatest tagging power, due to its higher efficiency. The tagging categories Kaon II and Inclusive show a significant difference between the mistag fractions for B^0 and $\overline{B}^0$, due to No significant differences are seen between Run I and Run II, except perhaps a slight decrease in the lepton efficiency for Run II due to the degrading IFR and therefore muon identification capabilities.

14.2.9 Possible further improvements

The tagging algorithm clearly uses the main features of flavor specific information. Further improvements thought of but not included (yet), are the following:

- K^0_S candidates, as they signify undetected strangeness.
- Leptons from cascade decays, by fully using the momentum spectrum.
- Electron veto for slow pions, to remove machine backgrounds.
- Optimization (and updating) of PID for tagging purposes.
- Semi-inclusive reconstruction to recognize D mesons and other features.

However, even when including these or possibly other features which carry information, the impact on the overall tagging power is limited, as the tagging performance is asymptotically approaching the limit of what information can be extracted.

14.3 Determination of decay time difference

The decay time between the two B decays is infered from the distance between the vertices (see Section 7.1.1), thus the vertex resolution therefore propagates into the time resolution. While the vertex of the fully reconstructed B meson B_{rec} is usually well determined, the vertex of the tagging side B meson B_{tag} is more problematic and therefore dominates the time resolution. As the resolution in time is comparable to the lifetime of the B^0, time-dependent fits require its proper description in data. This is obtained from fitting the well known lifetime distribution on a sample of fully reconstructed B^0 decays.

14.3.1 Vertex of reconstructed B meson

The decay vertex of B_{rec} is determined by fitting all (charged) decay products in the final state. Intermediate states with non-negligible lifetimes, such as K_S^0 and D mesons, replace their daughters in the fit (after improving their resolution with a kinematic fit). The typical z vertex resolution is $\sim 65\,\mu$m, slightly better for CP final states ($\sim 45\,\mu$m), and slightly worse for the final states in question ($\sim 75\,\mu$m), as these contain both a D and a K_S^0 meson (see Fig. 14.3a). However, this is not of great importance, as the time resolution is dominated by the tagging side vertex[87].

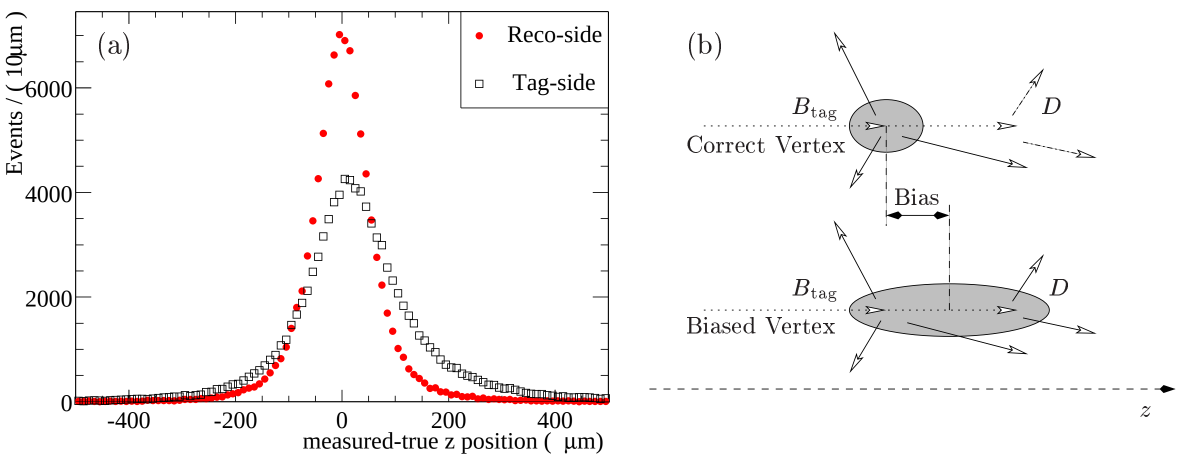


Figure 14.3: (a) Residuals for z_{rec} and z_{tag} in MC, showing the more precise determination of z_{rec} than z_{tag} and the bias on z_{tag}. (b) The bias origins from including particles from secondary vertices, mainly from D mesons. If excluded (shown with dash-dot linestyle), an unbiased value of z_{tag} can be obtained (see text).

14.3.2 Vertex of tagged B meson

The decay vertex of B_{tag} is determined from the charged tracks not used in the reconstruction of B_{rec}. Once again longlived intermediate states, such as K_S^0 and Λ^0, replace their daughters to avoid biases from secondary vertices, and tracks consistent with photon conversion ($\gamma \to e^+e^-$) are excluded from the fit.
From the reconstructed B meson, the momentum of B_{tag} can be derived, and together with the beam spot position, which has dimensions [MC01] $(\sigma_x \times \sigma_y \times \sigma_z) = 150\mu\text{m} \times 5\mu\text{m} \times 1\text{cm}$, a kinematic and geometrically constrained fit can be performed. Due to remaining secondary vertices (mostly from D mesons) and the forward boost, the tagging vertex is biased (in the negative direction due to the definition $\Delta z \equiv z_{\text{rec}} - z_{\text{tag}}$. To minimize these biases, any track

[87] For this reason, the resolution function will approximately be the same for all fully reconstructed B decays. Exceptions are $K_S^0\pi^0$, which using the beam spot attain comparable precision, and $\pi^0\pi^0$, which has no vertex.

contributing to the vertex χ^2 by more than 6 are excluded from the fit, and the fit is repeated, which is illustrated in Fig. 14.3b. Fig. 14.3a shows the residuals ($\delta z \equiv z^{\rm meas} - z^{\rm true}$) for $z_{\rm rec}$ and $z_{\rm tag}$. In the latter distribution, the remaining bias from secondary vertices can be seen. The value of Δz is determined directly from the $B_{\rm tag}$ vertex fit in order to incorporate correlations. The resulting resolution on Δz is 190 μm, dominated by the uncertainty in $B_{\rm tag}$.

14.3.3 Time resolution function

From the longitudinal distance between the reconstructed and the tagging vertex, Δz, the time difference, Δt, can be infered (see Section 7.1.1). Due to finite resolution, the measured and the true values of Δt differ. This detector response, called the resolution function, is parametrized by three Gaussian distributions (core, tail, and outlier components) as a function of the residual, $\delta_t = \Delta t_{\rm meas} - \Delta t_{\rm true}$ as:

$$\mathcal{R}(\delta_t, \hat{a}) = \sum_j^{\rm core,tail} f_j \, G(\delta_t - b_j \, \sigma_{\Delta t}, S_j \, \sigma_{\Delta t}) + f_{\rm outl} \, G(\delta_t, 8{\rm ps}), \tag{14.6}$$

where $G(x, \sigma_x)$ is a Gaussian and $\mathcal{R}$ is normalized by $f_{\rm core} + f_{\rm tail} + f_{\rm outl} = 1$. This parametrization is flexible enough to encompass the features of the resolution using the current statistics. The third Gaussian describing the so-called outliers is included to account for the 0.3% misreconstructed vertices. Therefore its width is fixed to 8ps with no scale factor or bias.
The vertex fit provides the Δt uncertainty, $\sigma_{\Delta t}$ on an event-by-event basis, which for simulated events is plotted as a function of the RMS of the residual in Fig. 14.4a. Scale factors, defined as the coefficient relating the per-event error to the uncertainty, $\sigma_j = S_j \sigma_{\Delta t}$, are used to account for underestimates ($S_j > 1$) or overestimates ($S_j < 1$) of $\sigma_{\Delta t}$. Due to large correlations (with $f_{\rm tail}$), $S_{\rm tail}$, is fixed (to 3) in the fit.

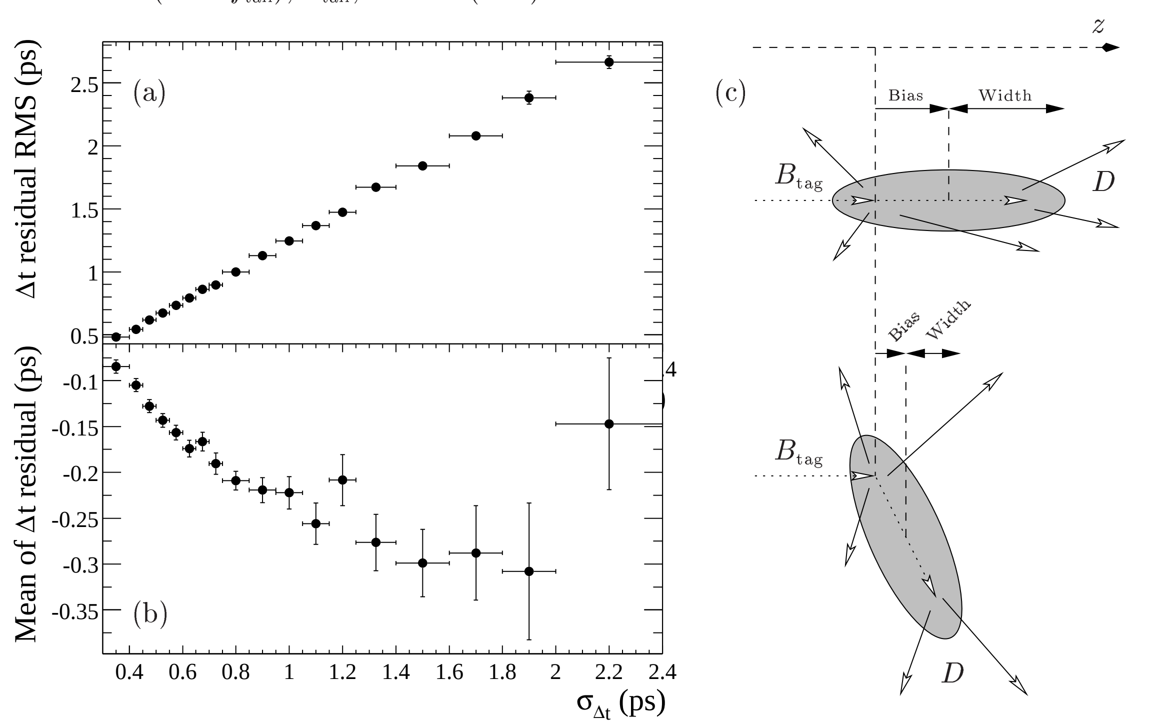


Figure 14.4: Correlation between $\sigma_{\Delta t}$ and (a) RMS and (b) mean of the residual for simulated events. (c) Illustration of the reason why events with a better defined z position (bottom diagram) also tend to be associated with a smaller bias, in accordance with (b). The black arrows above the decays describe the bias and width in z, respectively (see text).

The non-zero offsets due to the secondary vertices, are linearly correlated with $\sigma_{\Delta t}$, as can be seen for simulated events in Fig. 14.4b. The origin of this correlation is their common dependence on longlived intermediate particles (mostly D mesons) in the vertex fit. When emitted transversely (thus with a large sum of the transverse momentum squared, $\sum_i p_{t_i}^2$), a precise and unbiased z-position of the vertex is determined. However, the more longitudinally they are emitted the more they bias *and* increase the uncertainty of $z_{\rm tag}$, illustrated in Fig. 14.4c. The offsets, denoted δ_j^0, are therefore parametrized as $\delta_j^0 = b_j\sigma_{\Delta t}$, and as the biases for the core Gaussian varies with tagging category, k, these have separate biases, $b_{{\rm core},k}$. The reason for this variation is the source of tagging information. While leptons are from the primary vertex, kaons are mostly from secondary vertices, which creates a difference, as can be seen from Table 14.2. Note that the kaon and inclusive categories have very similar biases. The mistag fraction also has a correlation with the sum of the transverse momentum squared, $\sum_i p_{t_i}^2$, as lower momentum spectra (e.g. from $B^0 \to DDK$ events) have higher mistag rates [CC02]. The correlation is largest for tagging categories involving kaons. Studies have shown, that the impact of this correlation on CP asymmetries is of the order 0.5% [Rah02], and for this reason it has so far been included in the systematics.

With the exception of the tail scale factor, $S_{\rm tail}$, all core and tail parameters are determined by fitting the decay time distribution (a double-sided exponential function) of a large sample of fully reconstructed B mesons with an unbinned maximum likelihood fit. The determined resolution parameters are listed in Table 14.2. As a consistency check, the fitted lifetime is compared to the world average [PDG02].

Parameter	Value	Status	Parameter	Value	Status
	Signal ($\hat{a}$)			Background ($\hat{b}$)	
$f_{\rm core}(\%)$	0.893 ± 0.014	From fit	$f^{\rm bkg}(\%)$	Given by $f_{\rm outl}^{\rm bkg}$	
$S_{\rm core}$	1.096 ± 0.046	From fit	$S^{\rm bkg}$	1.33 ± 0.14	From fit
$b_{{\rm core},L}({\rm ps})$	0.040 ± 0.043	From fit	$b^{\rm bkg}({\rm ps})$	-0.2 ± 0.06	From fit
$b_{{\rm core},K_1}({\rm ps})$	-0.238 ± 0.069	From fit			
$b_{{\rm core},K_2}({\rm ps})$	-0.230 ± 0.083	From fit			
$b_{{\rm core},I}({\rm ps})$	-0.229 ± 0.063	From fit			
$f_{\rm tail}(\%)$	Given by $f_{\rm core}$ and $f_{\rm outl}$				
$S_{\rm tail}({\rm ps})$	3.0	Fixed			
$b_{\rm tail}({\rm ps})$	-1.039 ± 0.016	From fit			
$f_{\rm outl}(\%)$	0.003 ± 0.001	From fit	$f_{\rm outl}^{\rm bkg}(\%)$	0.007 ± 0.003	From fit
$\sigma_{\rm outl}({\rm ps})$	8.0	Fixed	$\sigma_{\rm outl}^{\rm bkg}({\rm ps})$	8.0	Fixed
$b_{\rm outl}({\rm ps})$	0	Fixed	$b_{\rm outl}^{\rm bkg}({\rm ps})$	0	Fixed

Table 14.2: The parameters of the resolution function for signal and background.

The resolution function used for all background components is described simply by two Gaussian distribution, where the dominant Gaussian describes the core distribution and the other describes outliers and have a fixed width and mean of 8 and 0 ps, respectively. For the core distribution, a common scale factor and bias offset is used for all tagging categories, as the data show no appreciable difference between these.

$$\mathcal{R}^{\rm bkg}(\delta_t, \hat{b}) \quad = \quad f_{\rm core}^{\rm bkg}\, G(\delta_t - b_{\rm core}\,\sigma_{\Delta t}, S_{\rm core}\,\sigma_{\Delta t}) \;+\; f_{\rm outl}^{\rm bkg}\, G(\delta_t, 8{\rm ps}), \tag{14.7}$$

The background parameters for the resolution function, $\hat{b}$, are obtained from the sidebands of a large data sample of fully reconstructed events, and it is assumed that the background resolution function (not the background time distribution) is the same for this sample and the one in question.

14.4 Time-dependent fit

Having determined the tagging performance and the resolution function, these experimental features have to be included in the time distributions of interest. The time-dependent fit here presented follows that of the $\sin(2\beta)$ analysis [BABAR02b], since this is well founded and has been subject to numerous tests and checks.

In principle all the samples discussed above can be fitted. However, time-dependent analysis of high background samples require quite detailed studies of these background's properties in time. Furthermore, when not performing a Dalitz analysis, only smaller areas of the Dalitz region should be fitted in order to be able to extract any information from the fit results. For these reasons only the $B^0 \to D^{\pm}K^{*\mp}$ sample has been fitted.

14.4.1 Signal time distribution

Including the tagging performance, the time distributions Eqs. (4.38–4.41) become:

$$f'(t) = (1-\omega)f(t) + \varpi\overline{f}(t) \quad (14.8)$$
$$\overline{f}'(t) = \omega f(t) + (1-\varpi)\overline{f}(t), \quad (14.9)$$

where $f(t)$ ($\overline{f}(t)$) is the time distribution for events with a *true* B^0 ($\overline{B}^0$) on the tagging side, and the $f'(t)$ ($\overline{f}'(t)$) are for events with a *tagged* B^0 ($\overline{B}^0$).
Applying these modifications to Eqs. (4.38 - 4.41), rearranging the terms, and using the notation with C and $S_{\pm}$ introduced in Section 4.6, yields:

$$\mathcal{F}_{B_{\rm rec}=D^-K^0_S\pi^+,B_{\rm tag}=B^0,k}(t) = Ne^{-\Gamma|t|}\left[1+\tfrac{1}{2}\Delta D_k - D_k(C\cos(\Delta mt) - S_-\sin(\Delta mt))\right] \quad (14.10)$$
$$\mathcal{F}_{B_{\rm rec}=D^-K^0_S\pi^+,B_{\rm tag}=\overline{B}^0,k}(t) = Ne^{-\Gamma|t|}\left[1-\tfrac{1}{2}\Delta D_k + D_k(C\cos(\Delta mt) - S_-\sin(\Delta mt))\right] \quad (14.11)$$
$$\mathcal{F}_{B_{\rm rec}=D^+K^0_S\pi^-,B_{\rm tag}=B^0,k}(t) = Ne^{-\Gamma|t|}\left[1+\tfrac{1}{2}\Delta D_k + D_k(C\cos(\Delta mt) + S_+\sin(\Delta mt))\right] \quad (14.12)$$
$$\mathcal{F}_{B_{\rm rec}=D^+K^0_S\pi^-,B_{\rm tag}=\overline{B}^0,k}(t) = Ne^{-\Gamma|t|}\left[1-\tfrac{1}{2}\Delta D_k - D_k(C\cos(\Delta mt) + S_+\sin(\Delta mt))\right] \quad (14.13)$$

The inclusion of the finite vertex resolution is obtained by convoluting the time distributions with the resolution function, $\mathcal{R}$.

The signal PDF is in short written as $\mathcal{F}_{Q_D,{\rm tag},k}$, where Q_D refers to the charge of the reconstructed $D^{(*)}$ meson, $_{\rm tag}$ refers to the tag being either B^0 or $\overline{B}^0$, and k is the tagging category. One defines $\xi = 1$ for events with $B_{\rm rec} = D^+K^0_S\pi^-$ and $B_{\rm tag} = B^0$ or $B_{\rm rec} = D^-K^0_S\pi^-$ and $B_{\rm tag} = \overline{B}^0$, and $\xi = -1$ for the other two combinations. Using this notation, the above formulae Eqs. (14.10- 14.13) can be contracted into:

$$\mathcal{F}_{+,{\rm tag},k}(t) = Ne^{-\Gamma|t|}\left[1 \pm \tfrac{1}{2}\Delta D_k + \xi D_k(C\cos(\Delta mt) - S_+\sin(\Delta mt))\right], \quad (14.14)$$
$$\mathcal{F}_{-,{\rm tag},k}(t) = Ne^{-\Gamma|t|}\left[1 \pm \tfrac{1}{2}\Delta D_k + \xi D_k(C\cos(\Delta mt) + S_-\sin(\Delta mt))\right], \quad (14.15)$$

where the $\pm$ refers to the tag ($+$ for B^0). The variables S_+ and S_- are substituted in the fit by:

$$S \equiv \tfrac{1}{2}(S_+ + S_-), \qquad \Delta S \equiv \tfrac{1}{2}(S_+ - S_-) \quad (14.16)$$

With the definition of Eq. (4.44), this means that the parameter ΔS contains all CP violating effects, while S is a measure of strong phases.
The distributions $\mathcal{F}_{Q_D,{\rm tag},k}$ are normalized according to:

$$\int_{-\infty}^{\infty}\sum_{\rm tag}(\mathcal{F}_{+,{\rm tag},k}(\Delta t) + \mathcal{F}_{-,{\rm tag},k}(\Delta t))d\Delta t = 1. \quad (14.17)$$

14.4.2 Background time distribution

The background time distribution needs a description for each component, since their origin and therefore properties in time are different. However, as the background consists of many different sources, it has to be described by a general time distributions. In general there are two issues to address, namely whether the component has a lifetime and whether it is mixing. Since the background is not overwhelming, the sensitivity to some of the terms may not be great, but they should be included anyway, to make sure that there are enough degrees of freedom to incorporate the features of the background time distribution. From the distribution of the Fisher shape variable, the background components are divided according to whether it is a continuum $c\bar{c}$ event or a $B\overline{B}$ event.

The continuum background (abbreviated *cont*) is in principle prompt, in the sense that the hadronization happens immediately, and thus all tracks should originate from the same vertex, resulting in a delta function distribution in time, $\mathcal{B}^{\mathrm{cont},1}_{Q_D,\mathrm{tag},k} = \frac{1}{2}\delta(\Delta t_{\mathrm{true}}) \otimes \mathcal{R}^{\mathrm{bkg}}(\delta_t, \hat{b})$.
However, the longlived D mesons may mimick a lifetime component, and therefore the continuum background description should contain both a prompt and a non-zero lifetime component, where the lifetime of the latter is an effective one, simply parametrizing the typical bias, $\mathcal{B}^{\mathrm{cont},2}_{Q_D,\mathrm{tag},k} = \frac{1}{4}\Gamma^{\mathrm{eff}}_{\mathrm{cont}} \exp(-\Gamma^{\mathrm{eff}}_{\mathrm{cont}}|\Delta t_{\mathrm{true}}|) \otimes \mathcal{R}^{\mathrm{bkg}}(\delta_t, \hat{b})$.
The prompt fraction of continuum background is obtained from the off-resonance data, and the effective lifetime is set to 1ps. It is fixed in the fit as it is very correlated with the prompt fraction.

The $B\overline{B}$ combinatorial background (abbreviated *comb*) comes both from neutral as well at charged pairs of B mesons. Both of these have a non-zero lifetime, but only the neutral component mixes. Nevertheless, the lifetime is not necessarily present in the time distribution of the $B\overline{B}$ background, as daughters from both B mesons may be used in the (background candidate) reconstruction, which means that the lifetime should be compatible with zero.
For B^+B^- events, this interchange is unavoidable, as the charge would otherwise not match, leaving the prompt non-mixing distribution $\mathcal{B}^{\mathrm{comb},1}_{Q_D,\mathrm{tag},k} = \frac{1}{2}\delta(\Delta t_{\mathrm{true}}) \otimes \mathcal{R}^{\mathrm{bkg}}(\delta_t, \hat{b})$. For $B^0\overline{B}^0$ events the degree of interchange may vary, leaving either a prompt non-mixing component (described by $\mathcal{B}^{\mathrm{comb},1}$) or a non-prompt mixing component, described by $\mathcal{B}^{\mathrm{comb},2}_{Q_D,\mathrm{tag},k} = \frac{1}{4}\Gamma_B \exp(-\Gamma_B|\Delta t_{\mathrm{true}}|)(1 \pm \Delta D_k + \xi D_k \cos(\Delta m \Delta t_{\mathrm{true}})) \otimes \mathcal{R}^{\mathrm{bkg}}(\delta_t, \hat{b})$, with the lifetime and mixing frequency of the B^0 meson, where the $\pm$ refers to the B tag.

The non-degenerate peaking background derives from misreconstructed $B\overline{B}$ events, which are similar to the signal events, thus with a lifetime. The component is described by $\mathcal{B}^{\mathrm{peak},1}_{Q_D,\mathrm{tag},k} = \frac{1}{4}\Gamma_B \exp(-\Gamma_B|\Delta t_{\mathrm{true}}|) \otimes \mathcal{R}^{\mathrm{bkg}}(\delta_t, \hat{b})$.
To incorporate possible CP violation in the peaking component, an additional component is included. It is described by a distribution similar to signal, including an *effective* CP asymmetry in terms of C^{eff} and S^{eff}, that is $\mathcal{B}^{\mathrm{peak},2}_{Q_D,\mathrm{tag},k} = \frac{1}{4}\Gamma_B \exp(-\Gamma_B|\Delta t_{\mathrm{true}}|)(1 \pm \Delta D_k + \xi D_k[C^{\mathrm{eff}} \cos(\Delta m \Delta t_{\mathrm{true}}) + S^{\mathrm{eff}} \sin(\Delta m \Delta t_{\mathrm{true}})]) \otimes \mathcal{R}(\delta_t, \hat{a})$, where the $\pm$ again refers to the B tag. This component is most likely *not* sensitive to any CP asymmetries, but incorporates possible systematic effects. Note that this component is convoluted with the *signal* resolution function, due to its close resemblance.
All of these three backgrounds are normalized as:

$$\int_{-\infty}^{\infty} \sum_{\mathrm{tag},b} (\mathcal{B}^{\mathrm{bkg},b}_{+,\mathrm{tag},k}(\Delta t_{\mathrm{true}}) + \mathcal{B}^{\mathrm{bkg},b}_{-,\mathrm{tag},k}(\Delta t_{\mathrm{true}})) d\Delta t \quad = \quad 1, \tag{14.18}$$

where $b = 1, 2$ is the index of the background. These background descriptions are summarized in Table 14.3.

Component	Features	PDF
Continuum	P, NM	$\mathcal{B}^{\mathrm{cont},1}_{Q_D,\mathrm{tag},k} = \frac{1}{2}\delta(\Delta t_{\mathrm{true}})$
	L, NM	$\mathcal{B}^{\mathrm{cont},2}_{Q_D,\mathrm{tag},k} = \frac{1}{4}\Gamma^{\mathrm{eff}}_{\mathrm{cont}} \exp(-\Gamma^{\mathrm{eff}}_{\mathrm{cont}}\|\Delta t_{\mathrm{true}}\|)$
$B\overline{B}$	P, NM	$\mathcal{B}^{\mathrm{comb},1}_{Q_D,\mathrm{tag},k} = \frac{1}{2}\delta(\Delta t_{\mathrm{true}})$
	L, M	$\mathcal{B}^{\mathrm{comb},2}_{Q_D,\mathrm{tag},k} = \frac{1}{4}\Gamma_B \exp(-\Gamma_B\|\Delta t_{\mathrm{true}}\|)(1 \pm \Delta D_k + \xi D_k \cos(\Delta m \Delta t_{\mathrm{true}}))$
Peaking	P, NM	$\mathcal{B}^{\mathrm{peak},1}_{Q_D,\mathrm{tag},k} = \frac{1}{4}\Gamma_B \exp(-\Gamma_B\|\Delta t_{\mathrm{true}}\|)$
	L, M, CPV	$\mathcal{B}^{\mathrm{peak},2}_{Q_D,\mathrm{tag},k} = \frac{1}{4}\Gamma_B \exp(-\Gamma_B\|\Delta t_{\mathrm{true}}\|)(1 \pm \Delta D_k + \xi D_k$ $[C^{\mathrm{eff}} \cos(\Delta m \Delta t_{\mathrm{true}}) + S^{\mathrm{eff}} \sin(\Delta m \Delta t_{\mathrm{true}})])$

Table 14.3: The PDFs describing the various Δt components of the background. The PDFs are normalized and convoluted with the background resolution function (except $\mathcal{B}^{\mathrm{peak},2}_{Q_D,\mathrm{tag},k}$) before entering the likelihood fit. The feature abbreviations are Prompt (P), Lifetime (L), Mixing (M), Non-Mixing (NM), and CP violating (CPV).

14.4.3 Likelihood function

Given the signal and background parametrizations above, a likelihood function can now be build as a sum over events i:

$$\ln \mathcal{L} = \sum_i \ln \mathcal{L}_{Q_{D,i},\mathrm{tag}_i,k_i}(\Delta t_i) \tag{14.19}$$

where k denotes a tagging category, and $\mathcal{L}_{Q_D,\mathrm{tag},k}$ is the sum:

$$\mathcal{L}_{Q_D,\mathrm{tag},k} = f_k^{\mathrm{sig}} \mathcal{F}_{Q_D,\mathrm{tag},k} + f_k^{\mathrm{cont}} \mathcal{B}^{\mathrm{cont}}_{Q_D,\mathrm{tag},k} + f_k^{\mathrm{comb}} \mathcal{B}^{\mathrm{comb}}_{Q_D,\mathrm{tag},k} + f_k^{\mathrm{peak}} \mathcal{B}^{\mathrm{peak}}_{Q_D,\mathrm{tag},k} \, . \tag{14.20}$$

The component fractions satisfy the relation $f_k^{\mathrm{sig}} + f_k^{\mathrm{cont}} + f_k^{\mathrm{comb}} + f_k^{\mathrm{peak}} = 1$.

14.4.4 Likelihood fit

For the time-dependent fit the events are required to have:

- A tag
- $|\Delta t| < 20\,\mathrm{ps}$
- $\sigma_{\Delta t} < 2.5\,\mathrm{ps}$

The tag is necessary in order to extract CP information. This can be seen from Eqs. (14.10-14.13), where the C and $S_\pm$ terms vanish, when $D_k = 0$ (corresponding to $\langle\omega\rangle = 0.5$, that is equal probability of being a B^0 or a $\overline{B}^0$). The requirement on $|\Delta t|$ is very loose considering the lifetime of the B^0. Along with the $\sigma_{\Delta t}$ requirement it rejects misreconstructed events. This ensures that no events lie beyond what is accounted for by the outlier component of the resolution function. Events with $\sigma_{\Delta t} > 2.5\,\mathrm{ps}$ provide essentially no sensitivity anyway.
The time-dependent fit is applied to the $B^0 \to D^\pm K^{*\mp}$ resonant region, as this has a high purity. A fit of the region of higher invariant mass in $K^0_S\pi$, $1.0 < m(K^0_S\pi) < 2.0\,\mathrm{GeV}$, was also attempted, but the fit had essentially no sensitivity to C, S, and ΔS.

The dominant background in this region is not surprisingly from continuum, as this background more easily resembles (quasi) two-body decay modes. The $B\overline{B}$ backgrounds are not very large (see Table 10.8), and their influence on the fit results is therefore limited.

The result of the fit can be seen for each tagging category in Fig. 14.5, and the parameters and their correlations with C, S, and ΔS can be found in Table 14.4.

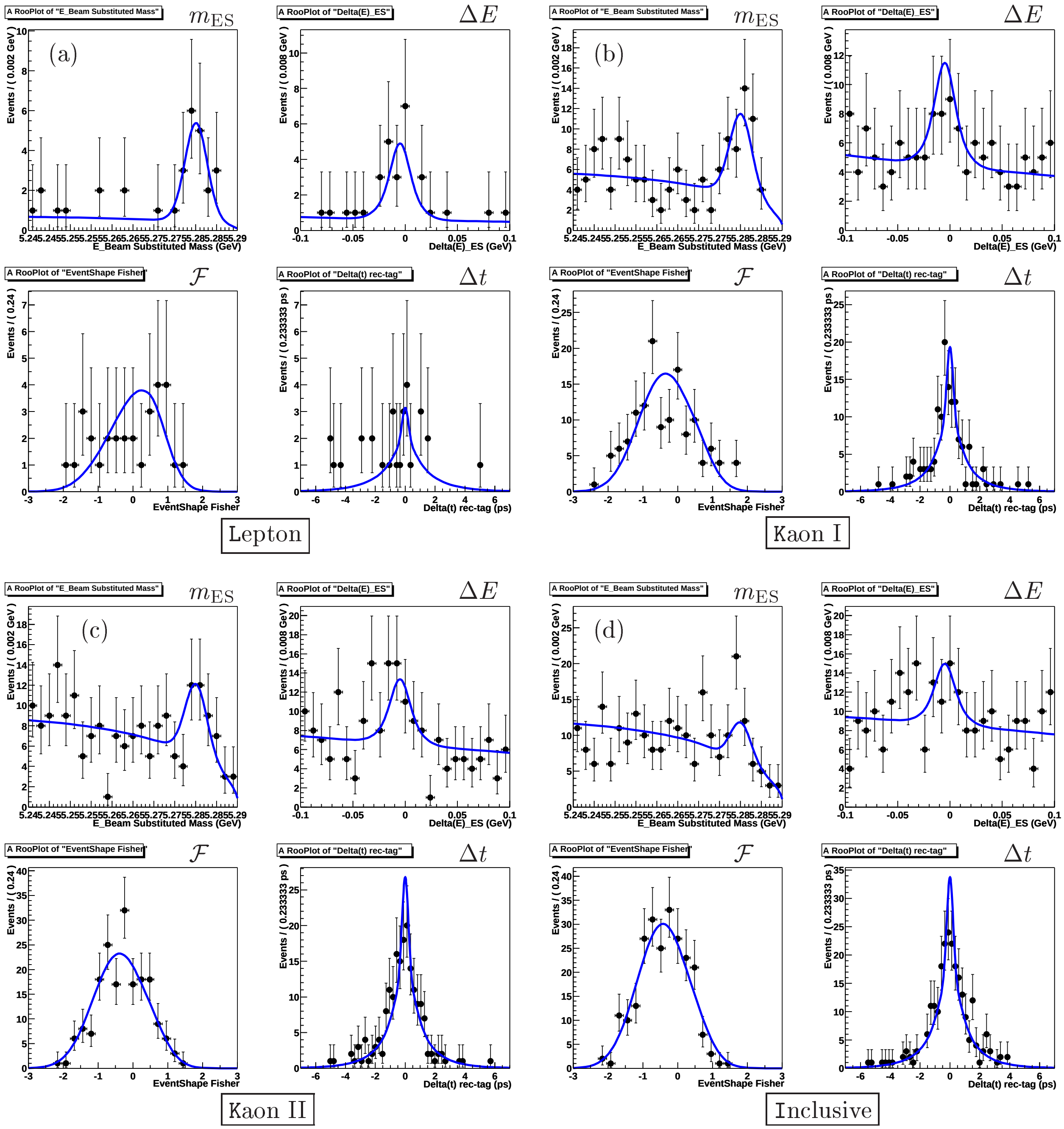


Figure 14.5: Data distribution and fit PDF of the four tagging categories (a) Lepton, (b) Kaon I, (c) Kaon II, and (d) Inclusive in the variables m_{ES}, ΔE, $\mathcal{F}$, and Δt (see text).

14.4.5 Systematic errors

The study of systematic errors is still preliminary and has not been fully performed yet, thus the following is only an outline of initial studies. However, the systematic errors of the time-dependent fit are not expected to be of the same magnitude as the statistical uncertainties. The reason for this is that most systematic errors (e.g. tracking, branching fractions, efficiency corrections) tend to cancel out, since their impact on reconstruction efficiency can be factorized out of the likelihood function. For the same reason, conservative estimates are used.

Parameter	Value	ρ_C	ρ_S	$\rho_{\Delta S}$
$f_{\text{sig},1}$	0.50 ± 0.10	-0.01	0.02	0.02
$f_{\text{sig},2}$	0.18 ± 0.03	-0.05	-0.06	0.00
$f_{\text{sig},3}$	0.13 ± 0.03	-0.05	0.01	-0.02
$f_{\text{sig},4}$	0.09 ± 0.02	-0.01	0.01	-0.01
$f_{\text{cont},1}$	0.34 ± 0.10	-0.01	-0.01	0.00
$f_{\text{cont},2}$	0.76 ± 0.05	0.01	0.02	-0.01
$f_{\text{cont},3}$	0.81 ± 0.05	0.00	0.01	0.01
$f_{\text{cont},4}$	0.86 ± 0.05	0.00	0.00	-0.01
$f_{\text{comb},1}$	0.10 ± 0.08	0.04	-0.01	-0.02
$f_{\text{comb},2}$	0.03 ± 0.04	-0.01	0.00	0.01
$f_{\text{comb},3}$	0.07 ± 0.05	-0.01	0.00	0.00
$f_{\text{comb},4}$	0.08 ± 0.05	0.01	-0.01	0.01
$f_{\text{cont}}^{\text{prompt}}$	0.61 ± 0.05	Fixed (from off-resonance data)		
τ_{cont}	1.0		Fixed	
$f_{\text{comb}}^{\text{prompt}}$	0.01 ± 0.08	0.00	0.00	0.00
$f_{\text{peak}}^{\text{prompt}}$	0.90 ± 0.18	0.00	0.00	0.00
C^{eff}	4.76 ± 6.73	-0.00	0.00	0.00
S^{eff}	2.93 ± 2.26	-0.00	0.00	0.00
C	$\mathbf{0.93 \pm 0.18}$	**1.00**	**0.03**	**-0.07**
S	$\mathbf{0.18 \pm 0.28}$	**0.04**	**1.00**	**-0.31**
ΔS	$\mathbf{0.08 \pm 0.28}$	**-0.07**	**-0.31**	**1.00**

Table 14.4: Component fractions, background parameters, and signal coefficients. In addition to the value of the parameters, the correlations with the three signal coefficients C, S, and ΔS are shown.

The systematic errors from the uncertainty in the mixing frequency and lifetime of the B^0 are obtained by varying these within the uncertainties on their values [PDG02]. The impact on the three parameters C, S, and ΔS is noted, and this change is taken as a systematic error. Likewise, the systematic uncertainty from the tagging performance is obtained by varying the tagging parameters of Table 14.1 within their errors. The changes noted are added in quadrature for each parameter, as their correlations were not available.
The systematic error from the resolution function is twofold. One regards the shape used to describe its PDF, and the other the actual values (and uncertainties) in the parameters entering the PDF. The uncertainty from varying the parameters is divided into that of the signal resolution function and that of the background, and the errors are added in quadrature. The outlier width was varied ±2ps, the scale factor of the tail by ±0.5, and the effective lifetime of the continuum by ±0.2ps. To quantify the error associated with the model for the signal resolution function, an alternative model is used. The model, similar to the one used in the fit, incorporates the biases from D decays not explicitly but by introducing a convolution of a Gaussian with a one-sided exponential, hence the name GExp:

$$\mathcal{R}^{\text{GExp}}(\delta_t, \hat{a}) \;=\; f_{\text{core}}\, G(\delta_t, S_j\, \sigma_{\Delta t}) \otimes (f_i^G \delta(\delta_t) + (1 - f_i^G)\tfrac{1}{\tau} e^{-\delta_t/\tau}) \;+\; f_{\text{outl}}\, G(\delta_t,\, 8\text{ps}), \quad (14.21)$$

where the exponential is one-sided, i.e. the function takes the value zero for $\delta_t < 0$. The effective lifetime τ is common for all tagging categories, but the fraction of pure Gaussian f_i^G is a function of tagging category, to account for varying D meson biases (cf. Table 14.2).
The possibility of DCSD of B_{tag} leads to interference terms, which may influence the time distribution [LBCK03]. Essentially, DCSD introduce a sine and a cosine term. The sine term simply reduces the amplitude by $(1 - r^2)$, where r is the ratio of the DCS amplitude and the

Cabibbo allowed amplitude ($\sim \lambda^2$), and this effect is already included in the mistag fractions. The impact on the cosine term depends on the strong and weak phases of the backgrounds. These are unknown and vary from mode to mode, and the size of the effect depends on the coherency of their adding, which makes it impossible to directly calculate. The systematic uncertainty from this effect is conservatively taken to be 0.02.

Additional errors are the effect of the beamspot position used when determining the vertex positions, the tracking alignment also affecting the vertex precision, and efficiency differences in reconstructing $D^- K^0_S \pi^+$ compared to $D^+ K^0_S \pi^-$, which can bias all of the three terms. The first effect has been studied elsewhere [BABAR02b] by varying the beam spot position within its known position (determined from e^+e^- and $\mu^+\mu^-$ events and the luminosity). The uncertainty from the tracking alignment is dominated by the “local” SVT alignment, where “local” refers to the relative position of the SVT wafers [BABAR01]. Both of these effects are expected to contribute to the systematic error by about 1.0%.
The efficiency difference between positive and negative GTVL has a systematic error of 0.30%. As the final states $D^{(*)+} K^0_S \pi^-$ and $D^{(*)-} K^0_S \pi^+$ have the same number of charged tracks, there should be no efficiency difference between the two modes, unless the charge asymmetry is significantly different for kaons compared to pions. The size of such an effect would most likely not exceed the general charge asymmetry uncertainty, and therefore no large corrections and systematics are expected from this source.

Parameter	C	ΔS	S
$\Delta m \pm 0.009\,\mathrm{ps}^{-1}$	0.004	0.005	**0.010**
$\tau_{B^0} \pm 0.016\,\mathrm{ps}$	0.001	0.000	0.000
Tagging	**0.017**	**0.010**	0.005
Resolution model (signal)	0.002	**0.009**	0.004
Resolution function (signal)	0.002	0.002	0.003
Resolution function (background)	0.005	**0.014**	**0.009**
DCSD	**0.020**	0.000	0.000
Beam spot	0.010	0.010	0.010
SVT Alignment	0.010	0.010	0.010
Total	0.031	0.024	0.021

Table 14.5: Preliminatry list of systematic errors for time-dependent fit. The largest systematic errors are written in bold.

The systematic uncertainties are in general small as expected. The largest source of error is from tagging, which is not surprising, as the tagging coefficients have a very central role in the signal PDF and as they have to be obtained from fitting data samples, which are statistically limited. Also the resolution function model and parameters play a role. The large impact of the background resolution is mostly due to the uncertainty in the background (core) scale factor.
While the B^0 mixing frequency also has a sizable impact, the lifetime does not alter the results in any significant way. Finally, the impact on the systematic error on C from DCSD is large, which is most likely due to the conservative approach.

14.4.6 Section summary and conclusions

Though the data sample is still too limited to yield decisive results, a first pass at a time-dependent analysis was employed. The tagged $B_{\rm tag}$ mesons are divided into categories according to the source and quality of the flavor information, as this increases the sensitivity. The detector response is parametrized in terms of resolution functions, which included the

various biases and other effects due to longlived particles in the decay products of the B^0 mesons. The parameters of both the tagging algorithm and the resolution functions are obtained from a large sample of fully reconstructed B^0 mesons. The effective tagging efficiency is $Q = 27.0 \pm 0.9\%$ and most tagged events have well determined values of Δt.
Given a theoretical signal time distribution and an empirical background time distribution description, the parameters $C = 0.93 \pm 0.18$, $S = 0.18 \pm 0.28$, and $\Delta S = -0.08 \pm 0.28$ are extracted from an unbinned maximum likelihood fit. The systematic errors are evaluated by varying the parameters involved, and they are not very large compared to the statistical errors, as expected. From Eq. (4.43) the value of C can be translated into an amplitude ratio, λ, between the $b \to u$ and the $b \to c$ amplitudes. The value obtained is $\lambda = 0.19 \pm 0.25 \pm 0.04$.

15 Conclusions

Using approximately 88 million $B\overline{B}$ pairs, the decay modes $B^0 \to D^{\pm}K^0\pi^{\mp}$ and $B^0 \to D^{*\pm}K^0\pi^{\mp}$ have been studied. Their branching fractions have been measured and their Dalitz distributions extracted. In addition, the branching fractions of the $K^{*\pm}$ resonant channels $B^0 \to D^{\pm}K^{*\mp}$ and $B^0 \to D^{*\pm}K^{*\mp}$ have been measured. The results are:

$$
\begin{aligned}
Br(B^0 \to D^{\pm}K^0\pi^{\mp}) &= (5.0 \pm 0.7_{\text{stat}} \pm 0.6_{\text{syst}}) \times 10^{-4} \\
Br(B^0 \to D^{\pm}K^{*\mp}) &= (4.8 \pm 0.6_{\text{stat}} \pm 0.5_{\text{syst}}) \times 10^{-4} \\
Br(B^0 \to D^{*\pm}K^0\pi^{\mp}) &= (3.0 \pm 0.7_{\text{stat}} \pm 0.3_{\text{syst}}) \times 10^{-4} \\
Br(B^0 \to D^{*\pm}K^{*\mp}) &= (3.2 \pm 0.6_{\text{stat}} \pm 0.3_{\text{syst}}) \times 10^{-4}
\end{aligned}
$$

A time-dependent analysis of the $K^{*\pm}$ resonant region of the $B^0 \to D^{\pm}K^0\pi^{\mp}$ channel was performed. The values of the interference and CP sensitive variables are:

$$
\begin{aligned}
C &= 0.93 \pm 0.18_{\text{stat}} \pm 0.03_{\text{syst}}, \\
S &= 0.18 \pm 0.28_{\text{stat}} \pm 0.02_{\text{syst}}, \\
\Delta S &= -0.08 \pm 0.28_{\text{stat}} \pm 0.02_{\text{syst}}
\end{aligned}
$$

A comparison of these results with previous measurements and an interpretation of the results are presented in Section 15.1. The possible improvements of the analysis and the prospects of increased integrated luminosity is discussed in Section 15.2. Finally, the outlook on measuring γ and the future of B physics are concluded upon in Section 15.3.

15.1 Comparison and interpretation of results

The three-body branching fractions have been measured for the first time and their size of the order expected. The $K^{*\pm}$ resonant branching fractions have been measure before [PDG02].

$$
\begin{aligned}
Br(B^0 \to D^{\pm}K^{*\mp}) &= (3.7 \pm 1.5_{\text{stat}} \pm 1.0_{\text{syst}}) \times 10^{-4}, \\
Br(B^0 \to D^{*\pm}K^{*\mp}) &= (3.8 \pm 1.3_{\text{stat}} \pm 0.8_{\text{syst}}) \times 10^{-4}.
\end{aligned}
$$

The comparison is good, and the errors have been reduced by more than a factor of two, even though the selection is not optimized for these channels.

The results can also be compared to the isospin-related resonant two-body channels $B^+ \to D^0K^{*+}$ [BABAR04a] and $B^+ \to D^{*0}K^{*+}$ [BABAR03c]. However, the comparison cannot be done directly, as the charged channels have an additional diagram, which is color suppressed. A naive addition of these would suggest a total enhancement of the branching fraction of $|1 + 1/3|^2 = 16/9$.

Taking this factor at face value, the $B^0 \to D^{\pm}K^0\pi^{\mp}$ measurement is in good agreement with the measured value of the isospin-related decay, while the $B^0 \to D^{*\pm}K^0\pi^{\mp}$ result is slightly low:

$$
\begin{aligned}
Br(B^+ \to D^0K^{*+}) &= (6.3 \pm 0.7_{\text{stat}} \pm 0.5_{\text{syst}}) \times 10^{-4}, \\
Br(B^+ \to D^{*0}K^{*+}) &= (8.3 \pm 1.1_{\text{stat}} \pm 1.0_{\text{syst}}) \times 10^{-4}.
\end{aligned}
$$

A possible explanation would be destructive interference between the $b \to c$ and the $b \to u$ transitions, as both the strong phase and the weak phase are potentially large. However, given the errors, the discrepancy is not significant.

The Dalitz plot distributions shows that the $K^{*\pm}$ resonance is dominant, but in both the $B^0 \to D^{\pm}K^0\pi^{\mp}$ and the $B^0 \to D^{*\pm}K^0\pi^{\mp}$ mode approximately one third of the contribution is from other sources. However, the determination of additional structures requires more statistics. From the helicity distribution of the $K^{*\pm}$ resonance, possible other interfering contributions do not seem to have a spin 0 nor spin 2 structure, as their interference would then alter the distinct $\cos^2(\theta)$ helicity distribution of the $K^{*\pm}$.

The time-dependent analysis of the resonant $B^0 \to D^{\pm}K^{*\mp}$ region does not have the required sensitivity to determine if there is any interference between the $b \to c$ and $b \to u$ transition. This was to be expected, but the time-dependent fit sets limits on the parameters, and in addition it shows the feasibility and precision of such a fit. The value of the coefficient $C = 0.93 \pm 0.18 \pm 0.03$ can be translated into a ratio of amplitudes, and one obtains $\lambda = 0.19 \pm 0.25 \pm 0.04$, which leaves the question of interference open. The CP violating parameter ΔS is consistent with zero, as one would expect when no significant interference is observed. For the region of higher invariant mass in $K^0_S\pi$, $1.0 < m(K^0_S\pi) < 2.0\,\text{GeV}$, no constraints on C, S, and ΔS are obtained.

15.2 Improvements and prospects of the analysis

Given additional time and/or data, the analysis presented could be improved in several ways of which a few are discussed below.

The dominating systematic errors are those from tracking and the uncertainty in the branching fractions of the subsequent D meson decays. While the latter is hard to improve (requires a dedicated analysis and is dominated by systematic errors), the former can with some effort be diminished. Also the PDF shapes involve significant uncertainty, but this is of statistical origin, and could be included as such by simultaneously fitting the control samples. However, the systematic errors are not dominant especially not for the time-dependent analysis, which renders the discussion somewhat academic.

The real limitations are the slightly low branching ratios and the seemingly irreducable backgrounds from combinatorics. The former can of course be redeemed with increased statistics, while the latter decreases the signal sensitivity to some degree. Even if employing neural networks for increasing the purity of the daughter selection, the impact on the overall purity is limited, as the larger part of the background consists of correctly reconstructed $D^{(*)}$, K^0_S, and π mesons, which mimick a B^0 meson decay. Nevertheless, signal can be extracted both inside and outside the $K^{*\pm}$ resonant region.

Given no a priori knowledge of or handles on the Dalitz distribution, conducting a full time-dependent Dalitz analysis does not seem feasible at this time[88].

Regarding ${}_s\mathcal{P}$lots, this method has proven to be a great force of the analysis, and the approach has subsequently been adapted by other three-body analyses of B mesons, mainly charmless (e.g. [Ola04]). The ${}_s\mathcal{P}$lot weights also give the capability of separating problems without sacrificing statistical power, e.g. allowing Dalitz plot fits without a simultaneous fit in m_{ES}, ΔE, and $\mathcal{F}$, which is very demanding in terms of fitting complexity and computing power.

15.3 Outlook on γ and B physics

Since the commencement of this thesis, γ has proven worthy of its reputation as notoriously elusive.

The improved GW method of considering asymmetries and branching ratios (see Eq. (5.3) of Section 5.1) has only given very weak constraints on γ [Sto04]. The ADS method has recently been attempted in the $D^0 \to K\pi$ mode [BABAR04b], which essentially consists of looking for the decay $B^+ \to [K^-\pi^+]_D K^+$. No events were seen, which yields an upper limit on the amplitude ratio of $r < 0.224$ ($r < 0.196$ assuming $48^\circ < \gamma < 75^\circ$) at 90% CL. This means that the prospects for measuring γ in the $B^{\pm} \to D^0K^{\pm}$ mode are diminished. Combining

[88]Other analysis with the same level of background have at least 4-5 times the number of signal events.

the result with the weak limits obtained from the improved GW method yields no preferred value for γ (the likelihood curve is flat!).
The only promising method is that of the $B^{\pm} \to [K_S^0\pi^+\pi^-]K^{\pm}$ mode, where the Dalitz plot of the D^0 meson lays the basis for the interference between the $b \to c$ and the $b \to u$ transitions. The main advantage is that the Dalitz plot can be obtained from the $D^{*+} \to D^0(\to K_S^0\pi^+\pi^-)\pi^+$ decay, which has very large statistics. Using this channel, limits of $61° < \gamma < 142°$ (90% CL) has been set [Belle03], and contrary to the two-body methods, there is only the $\pm\pi$ ambiguity in the solution.
In addition, the time-dependent $B^0 \to D^{(*)\pm}\pi^{\mp}$ analysis have also set limits on $\sin(2\beta + \gamma)$ [BABAR03d, BABAR03e], but the prospects for actual measurements γ are not clear, not the least because of ambiguities.
The situation is generally that the sensitivity to γ of any method is low, most notably because of the small amplitude ratio. Whether the amplitude ratio and the interference is large in the $B^0 \to D^{(*)\pm}K^0\pi^{\mp}$ modes can still not be concluded upon. However, if seen, the modes have prospects of high sensitivity without the eight-fold ambiguity.

In the future, most likely the large statistics and B_s^0 capabilities of first CDF and D0, and afterwards LHCb and BTeV will in time give fairly precise measurements of γ, since decays measuring this angle are of the same order in λ in a time-dependent analysis similar to that of $B^0 \to D^{\pm}\pi^{\mp}$. Nevertheless, until then γ will most likely not be measured with any significant precision, and its value will have to be infered indirectly from fitting the CKM matrix.
It is interesting to note, that since the $B^0 \to D^{(*)\pm}K^0\pi^{\mp}$ mode is fully charged and has many constraints for rejecting background, of which continuum is the dominant, it might be suitable for hadronic machines.

References

[A^+01] K. Anikeev et al., *B Physics at the Tevatron: Run II and Beyond*, (2001), hep-ph/0201071.

[ABCLOS79] K. W. J. Barnham et al. (Aachen-Bonn-CERN-London-Oxford-Saclay Collaboration), *Spericity and Thrust Distributions in High-Energy Neutrino Interactions*, Phys. Lett. **B85**, 300 (1979).

[ADS97] D. Atwood, I. Dunietz and A. Soni, *Enhanced CP violation with $B \to KD^0(\overline{D}^0)$ modes and extraction of the CKM angle γ*, Phys. Rev. Lett. **78**, 3257–3260 (1997), hep-ph/9612433.

[AHR02] M. P. A. Höcker, H. Lacker and L. Roos, Interpreting CP Asymmetries in $B^0 \to \pi^+\pi^-$ Decays, BAD#519, LAL 02-103, 2002.

[AJBY04] C. T. A. J. Bevan and C. Yeche, Measurement of CP Violating asymmetries in $B^0(\overline{B}^0) \to \rho^+\rho^-$ decays and constraints on the CKM angle α at *BABAR*, BAD#795, 2004.

[AKL94] R. Aleksan, B. Kayser and D. London, *Determining the quark mixing matrix from CP violating asymmetries*, Phys. Rev. Lett. **73**, 18–20 (1994), hep-ph/9403341.

[AL94] A. Ali and D. London, *An Update of the CKM Matrix*, (1994), hep-ph/9409399.

[AP03] R. Aleksan and T. C. Petersen, *Measuring 2β + gamma with Color-allowed $B^0 \to D^\pm K^0\pi^\mp$ Decays*, ECONF **C0304052**, WG414 (2003), hep-ph/0307371.

[APS03] R. Aleksan, T. C. Petersen and A. Soffer, *Measuring the Weak Phase γ in Color Allowed $B \to DK\pi$ Decays*, Phys. Rev. **D67**, 096002 (2003), hep-ph/0209194.

[ARGUS87a] H. Albrecht et al. (ARGUS Collaboration), *Observation of $B^0\overline{B}^0$ Mixing*, Phys. Lett. **B192**, 245 (1987).

[ARGUS87b] H. Albrecht et al. (ARGUS Collaboration), *Reconstruction of B mesons*, Phys. Lett. **B185**, 218 (1987).

[B^+96] E. Badura et al., *Pestov spark counter prototype development for the CERN LHC ALICE experiment*, Nucl. Instrum. Meth. **A372**, 352–358 (1996).

[BABAR] D. Boutigny et al. (BABAR Collaboration), *BaBar technical design report*, SLAC-R-0457.

[BABAR-DIRC03] J. Schwiening et al. (BABAR-DIRC Collaboration), *The DIRC detector at the SLAC B factory PEP-II: Operational experience and performance for physics application*, Nucl. Instrum. Meth. **A502**, 67–75 (2003).

[BABAR-DIRC04] I. Adam et al. (BABAR-DIRC Collaboration), *The DIRC Particle Identification System for the BaBar Experiment*, (2004), To be submitted to Nucl. Instrum. Meth.

[BABAR01] C. Bozzi et al. (BABAR Collaboration), *First year experience with the BaBar silicon vertex tracker*, Nucl. Instrum. Meth. **A473**, 7–16 (2001).

[BABAR02a] B. Aubert et al. (BABAR Collaboration), *The BaBar detector*, Nucl. Instrum. Meth. **A479**, 1–116 (2002), hep-ex/0105044.

[BABAR02b] B. Aubert et al. (BABAR Collaboration), *Measurement of the CP violating asymmetry amplitude* $\sin 2\beta$, Phys. Rev. Lett. **89**, 201802 (2002), hep-ex/0207042.

[BABAR02c] B. Aubert et al. (BABAR Collaboration), *Measurements of branching fractions and CP violating asymmetries in $B^0 \to \pi^+\pi^-, K^+\pi^-, K^+K^-$ decays.*, Phys. Rev. Lett. **89**, 281802 (2002), hep-ex/0207055.

[BABAR03a] B. Aubert et al. (BABAR Collaboration), *Limits on $D^0\overline{D}^0$ mixing and CP violation from the ratio of lifetimes for decay to $K^-\pi^+$, K^-K^+ and $\pi^-\pi^+$*, Phys. Rev. Lett. **91**, 121801 (2003), hep-ex/0306003.

[BABAR03b] B. Aubert et al. (BABAR Collaboration), *Measurement of branching fractions of color-suppressed decays of the $\overline{B}^0$ meson to $D^{(*)0}\pi^0, D^{(*)0}\eta, D^{(*)0}\omega$, and $D^0\eta'$*, (2003), hep-ex/0310028.

[BABAR03c] B. Aubert et al. (BABAR Collaboration), *Measurement of the branching fraction and polarization for the decay $B^- \to D^{0*}K^{*-}$*, (2003), hep-ex/0308057.

[BABAR03d] B. Aubert et al. (BABAR Collaboration), *Measurement of time-dependent CP asymmetries and constraints on* $\sin(2\beta+\gamma)$ *with partial reconstruction of $B^0 \to D^{*\pm}pi^\pm$ decays*, (2003), hep-ex/0310037.

[BABAR03e] B. Aubert et al. (BABAR Collaboration), *Measurement of time-dependent CP asymmetries in $B^0 \to D^{(*)\pm}pi^\mp$ decays and constraints on* $\sin(2\beta+\gamma)$, (2003), hep-ex/0309017.

[BABAR03f] B. Aubert et al. (BABAR Collaboration), *Measurements of branching fractions and CP violating asymmetries in $B^0 \to \rho^\pm h^\mp$ decays*, Phys. Rev. Lett. **91**, 201802 (2003), hep-ex/0306030.

[BABAR04a] B. Aubert et al. (BABAR Collaboration), *Measurement of the branching fraction for $B^- \to D^0K^{*-}$*, Phys. Rev. **D69**, 051101 (2004), hep-ex/0312051.

[BABAR04b] B. Aubert et al. (BABAR Collaboration), *Search for $B^\pm \to (K^\mp\pi^\pm)_D K^\pm$ and upper limit on the $b \to u$ amplitude in $B^\pm \to DK^\pm$*, (2004), hep-ex/0402024.

[BBL96] G. Buchalla, A. J. Buras and M. E. Lautenbacher, *Weak Decays Beyond Leading Logarithms*, Rev. Mod. Phys. **68**, 1125–1144 (1996), hep-ph/9512380.

[Bec03] D. Becirevic, *Status of the computation of $f(B_{s,d})$, ξ and g*, (2003), hep-ph/0310072.

[Belle01] K. Abe et al. (Belle Collaboration), *A measurement of the branching fraction for the inclusive $B \to X_s$ gamma decays with Belle*, Phys. Lett. **B511**, 151–158 (2001), hep-ex/0103042.

[BELLE02] K. Abe et al. (BELLE Collaboration), *Observation of color-suppressed $\overline{B}^0 \to D^0\pi^0, D^{*0}\pi^0, D^0\eta$, and $D^0\omega$ decays*, Phys. Rev. Lett. **88**, 052002 (2002), hep-ex/0109021.

[Belle03] K. Abe et al. (Belle Collaboration), *Measurement of the angle ϕ_3 with Dalitz analysis of three-body D^0 decay from $B \to D^0K$ process*, (2003), hep-ex/0308043.

[Ber02] J. Beringer, BTagger – A Multivariate Tagging Algorithm with Categories Based on the Physics of the B_{tag} Decay, BAD#317, 2002.

[BF98] A. J. Buras and R. Fleischer, *Quark mixing, CP Violation and Rare decays after the Top Quark Discovery*, Adv. Ser. Direct. High Energy Phys. **15**, 65–238 (1998), hep-ph/9704376.

[BF99] A. J. Buras and R. Fleischer, *A general analysis of gamma determinations from $B \to \pi K$ decays*, Eur. Phys. J. **C11**, 93–109 (1999), hep-ph/9810260.

[BH34] H. Bethe and W. Heitler, *On the Stopping of fast particles and on the creation of positive electrons*, Proc. Roy. Soc. Lond. **A146**, 83–112 (1934).

[Bil84] P. Billoir, *Track Fitting with Multiple Scattering: A new method*, Nucl. Instr. Meth. **A225**, 352 (1984).

[BN37] F. Bloch and A. Nordsieck, *Note on the Radiation Field of the electron*, Phys. Rev. **52**, 54–59 (1937).

[BNT01] S. Burles, K. M. Nollett and M. S. Turner, *Big-Bang Nucleosynthesis Predictions for Precision Cosmology*, Astrophys. J. **552**, L1–L6 (2001), astro-ph/0010171.

[Bru95] R. Brun, `http://root.cern.ch`, 1995.

[BSN95] G. Blaylock, A. Seiden and Y. Nir, *The Role of CP violation in $D^0\overline{D}^0$ mixing*, Phys. Lett. **B355**, 555–560 (1995), hep-ph/9504306.

[C$^+$01] M. Ciuchini et al., *2000 CKM-triangle analysis: A critical review with updated experimental inputs and theoretical parameters*, JHEP **07**, 013 (2001), hep-ph/0012308.

[Cab63] N. Cabibbo, *Unitary Symmetry and Lepton decays*, Phys. Rev. Lett., **10**, 531 (1963).

[CC02] P. B. C.H.C̃heng, Correlation between mistag on kaon tag and calculated error on Δt, BAD#349, 2002.

[CCFT64] J. H. Christenson, J. Cronin, V. Fitch and R. Turlay, *Evidence for 2π decay of the K_2^0 meson*, Phys. Rev. Lett. **13**, 138 (1964).

[CDD56] L. Castillejo, R. H. Dalitz and F. J. Dyson, *Low's scattering equation for the charged and neutral scalar theories*, Phys. Rev. **101**, 453–458 (1956).

[CDF95] F. Abe et al. (CDF Collaboration), *Observation of Top Quark Production in $\overline{p}p$ Collisions*, Phys. Rev. Lett. **74**, 2626–2631 (1995), hep-ex/9503002.

[CDF03] D. Lucchesi (CDF Collaboration), *B_s physics and Prospects at the Tevatron*, ECONF **C0304052**, WG207 (2003), hep-ex/0307025.

[CDG98] A. Cohen, A. DeRujula and S. Glashow, *A Matter-Antimatter Univers?*, Astrophys. J. **495**, 539 (1998).

[CK03] R. N. Cahn and D. Kirkby, A General Model for Neutral B Decay-time Distributions, BAD#188, 2003.

[CKM03] October 2003, This was pointed out in the general discussed at the CKM Workshop at SLAC.

[CKN93] K.-w. Choi, D. B. Kaplan and A. E. Nelson, *Is CP a Gauge Symmetry?*, Nucl. Phys. **B391**, 515–530 (1993), hep-ph/9205202.

[CLEO81] K. Chadwick et al. (CLEO Collaboration), *Decay of b flavored Hadrons to Single muon and dimuon Final States*, Phys. Rev. Lett. **46**, 88–91 (1981).

[CLEO83] D. Andrews et al. (CLEO Collaboration), *The CLEO Detector*, Nucl. Instr. Meth. **211**, 47 (1983).

[CLEO96] D. M. Asner et al. (CLEO Collaboration), *Search for exclusive charmless hadronic B decays*, Phys. Rev. **D53**, 1039–1050 (1996), hep-ex/9508004.

[CLYO$^+$98] J. Charles, A. Le Yaouanc, L. Oliver, O. Pene and J. C. Raynal, *$B_d^0(t) \to DPP$ time-dependent Dalitz plots, CP violating angles 2β, $2\beta+\gamma$, and discrete ambiguities*, Phys. Lett. **B425**, 375 (1998), hep-ph/9801363.

[CrBall] T. Skwarnicki (Crystal Ball Collaboration), *A study of the Radiative Cascade Transitions between the Υ' and the Υ Resonances*, DESY F31-86-02.

[D$^+$85] A. Drescher et al., *The argus electron-photon calorimeter III. Electron-hadron separation*, Nucl. Instrum. Meth. **A237**, 464–474 (1985).

[DELPHI91] P. A. Aarnio et al. (DELPHI Collaboration), *The DELPHI detector at LEP*, Nucl. Instrum. Meth. **A303**, 233–276 (1991).

[DHKY02] A. S. Dighe, T. Hurth, C. S. Kim and T. Yoshikawa, *Measurement of the Width Difference of B_d Mesons*, Nucl. Phys. Proc. Suppl. **111**, 267–269 (2002), hep-ph/0202070.

[Dib90] F. L. Diberder, CP Violation as seen from the Δz Distribution, *BABAR* Note 42, 1990.

[DLH$^+$92] M. Dine, R. G. Leigh, P. Y. Huet, A. D. Linde and D. A. Linde, *Towards the Theory of the Electroweak Phase Transition*, Phys. Rev. **D46**, 550–571 (1992), hep-ph/9203203.

[DLM92] M. Dine, R. G. Leigh and D. A. MacIntire, *Of CP and other Gauge Symmetries in String Theory*, Phys. Rev. Lett. **69**, 2030–2032 (1992), hep-th/9205011.

[DR86] I. Dunietz and J. L. Rosner, *Time-dependent CP Violation Effects in $B^0\overline{B}^0$ Systems*, Phys. Rev. **D34**, 1404 (1986).

[DRK81] S. Dimopoulos, S. Raby and G. L. Kane, *Experimental Predictions from Technicolor Theories*, Nucl. Phys. **B182**, 77 (1981).

[DS88] I. Dunietz and R. G. Sachs, *Asymmetry Between Inclusive Charmed and Anticharmed Modes in $B^0\overline{B}^0$ Decay as a Measure of CP Violation*, Phys. Rev. **D37**, 3186 (1988).

[Dun95] I. Dunietz, *$B_{(s)}\overline{B}^0_{(s)}$ mixing, CP violation and extraction of CKM phases from untagged $B_{(s)}$ data samples*, Phys. Rev. **D52**, 3048–3064 (1995), hep-ph/9501287.

[E78797] S. C. Adler et al. (E787 Collaboration), *Evidence for the decay $K^+ \to \pi^+\nu\bar{\nu}$*, Phys. Rev. Lett. **79**, 2204–2207 (1997), hep-ex/9708031.

[ea57] C. S. W. et al., *Experimental Test of Parity Violation in Beta Decay*, Phys. Rev. **105**, 1413 (1957).

[ed.01] K. E. ed., *The Oxford Dictionary of Quotations*, Oxford University Press, 2001.

[EPR35] A. Einstein, B. Podolsky and N. Rosen, *Can Quantum Mechanical description of physical reality be considered complete?*, Phys. Rev. **47**, 777–780 (1935).

[Far03] A. Farbin, Measrement of Time-dependent CP asymmetries in $B^0 \to K_s\pi^0$ decays, BAD#714, 2003.

[FGLP02] A. F. Falk, Y. Grossman, Z. Ligeti and A. A. Petrov, *SU(3) breaking and $D^0\overline{D}^0$ mixing*, Phys. Rev. **D65**, 054034 (2002), hep-ph/0110317.

[Fis36] R. A. Fisher, *The use of multiple measurements in taxonomic problems*, Annals Eugen. **7**, 179–188 (1936).

[FW79] G. C. Fox and S. Wolfram, *Event Shapes in e^+e^- Annihilation*, Nucl. Phys. **B149**, 413 (1979).

[GGSZ03a] A. Giri, Y. Grossman, A. Soffer and J. Zupan, *Determination of the angle gamma using multibody D decays in $B^\pm \to DK^\pm$*, ECONF **C0304052**, WG424 (2003), hep-ph/0306286.

[GGSZ03b] A. Giri, Y. Grossman, A. Soffer and J. Zupan, *Determining gamma using $B^\pm \to DK^\pm$ with multibody D decays*, Phys. Rev. **D68**, 054018 (2003), hep-ph/0303187.

[GHLR94] M. Gronau, O. F. Hernandez, D. London and J. L. Rosner, *Decays of B mesons to two Light Pseudoscalars*, Phys. Rev. **D50**, 4529–4543 (1994), hep-ph/9404283.

[GHO+94] M. B. Gavela, P. Hernandez, J. Orloff, O. Pene and C. Quimbay, *Standard Model CP Violation and Baryon Asymmetry. Part 2: Finite Temperature*, Nucl. Phys. **B430**, 382–426 (1994), hep-ph/9406289.

[GIM70] S. L. Glashow, J. Iliopoulos and L. Maiani, *Weak Interactions with lepton-hadron Symmetry*, Phys. Rev. **D2**, 1285–1292 (1970).

[GKR02] F. J. Gilman, K. Kleinknecht and B. Renk, *The Cabibbo–Kobayashi–Maskawa Quark-mixing Matrix*, Phys. Rev. **D66**, 113–117 (2002).

[GL91] M. Gronau and D. London, *How to Determine all the Angles of the Unitarity Triangle from* $B^0 \to DK_s$ *and* $B_s^0 \to D^0\phi$, Phys. Lett. **B253**, 483–488 (1991).

[GLW57] R. L. Garwin, L. M. Lederman and M. Weinrich, *Observation of the Failure of Conservation of Parity and Charge Conjugation in Meson Decays: the Magnetic Moment of the Free Muon*, Phys. Rev. **105**, 1415 (1957).

[GQ98] Y. Grossman and H. R. Quinn, *Bounding the effect of penguin diagrams in* $a(CP)(B^0 \to \pi^+\pi^-)$, Phys. Rev. **D58**, 017504 (1998), hep-ph/9712306.

[GRL94] M. Gronau, J. L. Rosner and D. London, *Weak coupling phase from decays of charged* B *mesons to* πK *and* $\pi\pi$, Phys. Rev. Lett. **73**, 21–24 (1994), hep-ph/9404282.

[Gro98] M. Gronau, *Weak phase* γ *from color-allowed* $B \to DK$ *rates*, Phys. Rev. **D58**, 037301 (1998), hep-ph/9802315.

[Gro03] M. Gronau, *Improving bounds on* γ *in* $B^\pm \to DK^\pm$ *and* $B^{\pm,0} \to DX_s^{\pm,0}$, Phys. Lett. **B557**, 198–206 (2003), hep-ph/0211282.

[GW91] M. Gronau and D. Wyler, *On Determining a Weak Phase from* CP *Asymmetries in Charged* B *Decays*, Phys. Lett. **B265**, 172–176 (1991).

[H+77] S. W. Herb et al., *Observation of a Dimuon Resonance at 9.5* GeV *in 400* GeV *proton-nucleus Collisions*, Phys. Rev. Lett. **39**, 252–255 (1977).

[H+99] P. G. Harris et al., *New Experimental Limit on the Electric Dipole Moment of the Neutron*, Phys. Rev. Lett. **82**, 904–907 (1999).

[Har98] *The BaBar physics book: Physics at an Asymmetric B factory*, 1998.

[HERA-B03] I. Abt et al. (HERA-B Collaboration), *Measurement of the* $b\bar{b}$ *Production Cross Section in 920-GeV Fixed-target Proton Nucleus Collisions*, Eur. Phys. J. **C26**, 345–355 (2003), hep-ex/0205106.

[HKM+02] S. Hashimoto, A. S. Kronfeld, P. B. Mackenzie, S. M. Ryan and J. N. Simone, *Lattice calculation of the zero recoil form factor of* $\overline{B}^0 \to D^*\ell\bar{\nu}$*: Toward a model independent determination of —V(cb)—*, Phys. Rev. **D66**, 014503 (2002), hep-ph/0110253.

[HLGR94] O. F. Hernandez, D. London, M. Gronau and J. L. Rosner, *Measuring Strong and Weak Phases in Time-independent B Decays*, Phys. Lett. **B333**, 500–506 (1994), hep-ph/9404281.

[HLLLD01a] A. Hocker, H. Lacker, S. Laplace and F. Le Diberder, *Advanced Profile of the CKM Matrix*, (2001), In preparation.

[HLLLD01b] A. Hocker, H. Lacker, S. Laplace and F. Le Diberder, *A new approach to a global fit of the CKM matrix*, Eur. Phys. J. **C21**, 225–259 (2001), hep-ph/0104062.

[HM84] F. Halzen and A. D. Martin, *An Introduction Course in Modern Particle Physics*, John Wiley & Sons, Inc. 1984.

[HN94] S. Herrlich and U. Nierste, *Enhancement of the $K_L^0 K_S^0$ mass difference by short distance QCD corrections beyond leading logarithms*, Nucl. Phys. **B419**, 292–322 (1994), hep-ph/9310311.

[HS95] P. Huet and E. Sather, *Electroweak Baryogenesis and Standard Model CP Violation*, Phys. Rev. **D51**, 379–394 (1995), hep-ph/9404302.

[II03] D. Acosta et al. (CDF II Collaboration), *Measurement of the Mass Difference $m(D_s^+) - m(D^+)$ at CDF II*, Phys. Rev. **D68**, 072004 (2003), hep-ex/0310043.

[IL81] T. Inami and C. S. Lim, *Effects of Superheavy Quarks and Leptons in Low-energy Weak Processes $K_L^0 \to \mu\bar{\mu}$, $K^+ \to \pi^+\nu\bar{\nu}$ $K^0 \leftrightarrow \overline{K}^0$*, Prog. Theor. Phys. **65**, 297 (1981).

[Jac02] J. D. Jackson, *Kinematics*, Phys. Rev. **D66**, 248–251 (2002).

[Jar85] C. Jarlskog, *Commutator of the Quark Mass Matrices in the Standard ElectroWeak Model and a Measure of Maximal CP Violation*, Phys. Rev. Lett. **55**, 1039 (1985).

[JR75] F. James and M. Roos, *MINUIT a System for Function Minimization and Analysis of the Parameter errors and Correlations*, Comput. Phys. Commun. **10**, 343–367 (1975).

[Kal60] R. E. Kalman, *A New Approach to Linear Filtering and Prediction Problems*, Transactions of the ASME–Journal of Basic Engineering **82**, 35–45 (1960).

[Kit03] T. H. Kittelmann, Putting a Limit on the Width Difference of the Neutral B Mesons with the *BABAR* Experiment, Master Thesis, 2003.

[KKW93] G. L. Kane, C. F. Kolda and J. D. Wells, *Calculable Upper Limit on the Mass of the Lightest Higgs Boson in any Perturbatively valid Supersymmetric Theory*, Phys. Rev. Lett. **70**, 2686–2689 (1993), hep-ph/9210242.

[KM73] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49**, 652 (1973).

[KsC03] `http://www.slac.stanford.edu/BFROOT/www/Physics/TrackEfficTaskForce/Recipe/KsTracks-R10.html`, 2003.

[KV01] D. Kirkby and W. Verkerke, `http://roofit.sourceforge.net`, 2001.

[LBCK03] O. Long, M. Baak, R. N. Cahn and D. Kirkby, *Impact of tag-side interference on time dependent CP asymmetry measurements using coherent B0 anti-B0 pairs*, Phys. Rev. **D68**, 034010 (2003), hep-ex/0303030.

[LeC02] C. M. LeClerc, *Measurement of the $B^0\overline{B}^0$ Oscillation Frequency Using Fully-Reconstructed Semileptonic Decays*, PhD thesis, University of California, Berkeley, 2002.

[Lel01] L. Lellouch, *Light hadron weak matrix elements*, Nucl. Phys. Proc. Suppl. **94**, 142–158 (2001), hep-lat/0011088.

[LNQS91] H. J. Lipkin, Y. Nir, H. R. Quinn and A. Snyder, *Penguin trapping with isospin analysis and CP asymmetries in B decays*, Phys. Rev. **D44**, 1454–1460 (1991).

[LSS00] D. London, N. Sinha and R. Sinha, *Extracting weak phase information from $B \to V(1)V(2)$ decays*, Phys. Rev. Lett. **85**, 1807–1810 (2000), hep-ph/0005248.

[Lüd57] G. L. Lüders, *Proof of the TCP theorem*, Ann. Phys. 2 (1957).

[LY56] T. D. Lee and C. N. Yang, Phys. Rev. **104**, 254 (1956).

[Mac91] A. L. Mackay, *Dictionary of Scientific Quotations*, Institute of Physics Publishing, 1991.

[MC01] F. M.-V. M.C̃arpinelli, The *BABAR* vertexing, BAD#102, 2001.

[MNS62] Z. Maki, M. Nakagawa and S. Sakata, *Remarks on the Unified Model of Elementary Particles*, Prog. Theor. Phys. **28**, 870 (1962).

[Nel99] H. N. Nelson, *Compilation of $D^0 \to \overline{D}^0$ mixing predictions*, (1999), hep-ex/9908021.

[Neu99] M. Neubert, *Model-independent analysis of $B \to \pi K$ decays and bounds on the weak phase gamma*, JHEP **02**, 014 (1999), hep-ph/9812396.

[Nil84] H. P. Nilles, *Supersymmetry, Supergravity and Particle Physics*, Phys. Rept. **110**, 1 (1984).

[Nir98] Y. Nir, *Flavor physics and CP violation*, (1998), hep-ph/9810520.

[Nir02] Y. Nir, Heavy Flavour Physics, in *Proceedings of the Fifty-Fifth Scottish Universities Summer School in Physics*, edited by C. T. H. Davies and S. M. Playfer, Institute of Physics Publishing, 2002.

[Noe31] E. A. Noether, *Normalbasis bei Körpern ohne höhere Verzweigung*, J. reine angew. Math. **167**, 147–152 (1931).

[NP01] M. Neubert and A. A. Petrov, *Comments on color suppressed hadronic B decays*, Phys. Lett. **B519**, 50–56 (2001), hep-ph/0108103.

[Occ30] W. Occam, 1330, Occam's Razor states that *Entities should not be multiplied beyond necessity.*

[Odd] P. Oddone, *An asymmetric B factory based on PEP*, Prepared for Second International Symposium on The 4th Family of Quarks and Leptons, Santa Monica, California, 23- 25 Feb 1989.

[Ola04] M. Olaiya, Measurement of the $B^0 \to K^0\pi^+\pi^-$ Branching Fraction, BAD#840, 2004.

[P$^+$01] D. F. Phillips et al., *Limit on Lorentz and CPT Violation of the Proton using a Hydrogen Maser*, Phys. Rev. **D63**, 111101 (2001), physics/0008230.

[PD03] M. Pivk and F. L. Diberder, Radiative corrections in charmless decays, (despite the title, the approach is general for all B decays), 2003.

[PDG02] K. Hagiwara et al. (Particle Data Group Collaboration), *Review of particle physics*, Phys. Rev. **D66**, 010001 (2002).

[PDG03] K. Hagiwara et al. (Particle Data Group Collaboration), *Off-year partial opdate of PDG*, (2003), (URL: http://pdg.lbl.gov/).

[PEP93] *PEP-II: An Asymmetric B Factory. Conceptual Design Report*, (1993), SLAC-418.

[Pet00] T. C. Petersen, The λ^4 theorem, For a hadronic decay of one flavor, the amplitude for the decay to the same final state for the other flavor is at most suppressed by λ^2., 2000.

[PID03] `http://www.slac.stanford.edu/BFROOT/www/Physics/Tools/Pid/PidKilling/PidKilling.html`, 2003.

[PLD04] M. Pivk and F. R. Le Diberder, *sPlot: a statistical tool to unfold data distributions*, (2004), physics/0402083.

[PQ77] R. D. Peccei and H. R. Quinn, *Constraints Imposed by CP Violation in the Presence of Instantons*, Phys. Rev. **D16**, 1791–1797 (1977).

[PS95] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*, Addison-Wesley Publishing Company, 1995.

[PS99] S. Plaszczynski and M.-H. Schune, *Overall determination of the CKM matrix*, (1999), hep-ph/9911280.

[PS03] T. C. Petersen and M. H. Schune, Measurement of the $B^0 \to D^{(*)\pm} K^0 \pi^\mp$ Branching Fractions and Dalitz Distribution, BAD#630, 2003.

[PS04] T. C. Petersen and M. H. Schune, Measurement of the $B^0 \to D^{(*)\pm} K^0 \pi^\mp$ Branching Fractions and Dalitz Distribution, BAD#820, 2004.

[Rah02] S. Rahatlou, *Observation of matter-antimatter asymmetry in the B^0 meson system*, PhD thesis, University of California, San Diego, 2002.

[RNCMV04] D. K. R. N. Cahn, F. Forti and F. Martinez-Vidal, Limits on the Lifetime Difference of Neutral B Mesons and on CP, T, and CPT Violation in $B^0\overline{B}^0$ Mixing, BAD#625, 2004.

[S$^+$03] J. Seeman et al., *Progress of the PEP-II B factory*, ICFA Beam Dyn. Newslett. **31**, 66–71 (2003).

[Sak67] A. D. Sakharov, JETP Lett. **5**, 24 (1967).

[SB87] T. Sjostrand and M. Bengtsson, *The Lund Monte Carlo for Jet Fragmentation and e^+e^- Physics: JETSET version 6.3: An update*, Comput. Phys. Commun. **43**, 367 (1987).

[Sch53] J. S. Schwinger, *The Theory of Quantized Fields II*, Phys. Rev. **91**, 713–728 (1953).

[Sjo94] T. Sjostrand, *High-energy physics event generation with PYTHIA 5.7 and JETSET 7.4*, Comput. Phys. Commun. **82**, 74–90 (1994).

[SK98] Y. Fukuda et al. (Super-Kamiokande Collaboration), *Evidence for Oscillation of Atmospheric Neutrinos*, Phys. Rev. Lett. **81**, 1562–1567 (1998), hep-ex/9807003.

[Sof99] A. Soffer, *Discrete ambiguities in the measurement of the weak phase gamma*, Phys. Rev. **D60**, 054032 (1999), hep-ph/9902313.

[SQ93] A. E. Snyder and H. R. Quinn, *Measuring CP asymmetry in $B \to \rho\pi$ decays without ambiguities*, Phys. Rev. **D48**, 2139–2144 (1993).

[SS74] A. Salam and J. Strathdee, *Supersymmetry and non-Abelian gauges*, Phys. Lett. **B51**, 353–355 (1974).

[Ste76] G. Steigman, Ann. Rev. Astr. Ap. **14**, 339 (1976).

[Sto04] Private discussions and considerations with A. Stocchi., 2004.

[SV97] R. Sinkus and T. Voss, *Particle identification with neural networks using a rotational invariant moment representation*, Nucl. Instrum. Meth. **A391**, 360–368 (1997).

[Tra03] `http://www.slac.stanford.edu/BFROOT/www/Physics/TrackEfficTaskForce/TrackingTaskForce-2003.html#Recipes`, 2003.

[Tri01] S. Trincaz, PhD thesis, Universite Paris XII, Paris, 2001.

[UA184] G. Arnison et al. (UA1 Collaboration), *Associated Production of an Isolated Large Transverse Momentum Lepton (electron or muon), and Two Jets at the CERN $\overline{p}p$ Collider*, Phys. Lett. **B147**, 493 (1984).

[Vai99] A. I. Vainshtein, *How penguins started to fly*, Int. J. Mod. Phys. **A14**, 4705–4719 (1999), hep-ph/9906263.

[Wol64] L. Wolfenstein, *Violation of CP Invariance and the Possibility of very Weak Interactions*, Phys. Rev. Lett. **13**, 562–564 (1964).

[Wol83] L. Wolfenstein, *Parametrization of the Kobayashi–Maskawa matrix*, Phys. Rev. Lett. **51**, 1945 (1983).

[WW30] V. Weisskopf and E. P. Wigner, *Calculation of the Natural Brightness of Spectral Lines on the basis of Dirac's Theory*, Z. Phys. **63**, 54–73 (1930).

[WWW03] BaBar Operations, 2003, Available at `http://www.slac.stanford.edu/BFROOT/www/Detector/Operations/Operations.html`.

Abbreviations

Abbreviations are very frequently used in the experimental high energy physics community, as the field is somewhat technical, and so this thesis contains many as well. They are written in full the first time they occur with the abbreviation in parenthesis, after which only the abbreviation is used. However, abbreviations of accelerators and detectors/experiments are rarely written in full. Below is a list of the abbreviations used.

BAD	BaBar Analysis Document.
CDF	Collider Detector at Fermilab. Detector at Tevatron.
CERN	The European Laboratory for Particle Physics.
CKM	Cabibbo, Kobayashi and Maskawa. Inventors of the CKM quark mixing matrix.
CL	Confidence Limit.
CM	Center-of-Mass.
CP	Charge conjugation and Parity. Abstract operators transforming particles.
CPT	Charge conjugation, Parity and Time reversal.
D0	Detector at Tevatron, Fermilab.
DCH	Drift CHamber. Subdetector of BaBar.
DIRC	Detector of Internally Reflected Ĉerenkov light. Subdetector of BaBar.
EMC	ElectroMagnetic Calorimeter. Subdetector of BaBar.
FEE	Front End Electronics. Common electronics architecture placed on the subdetectors.
GIM	Glashow, Iliopoulos and Maiani. Inventors of the GIM-mechanism.
HER	High Energy Ring. The electron storage ring at PEP-II.
HERA-B	B physics at the HERA-accelerator. Experiment at DESY in Hamburg.
HV	High Voltage.
IFR	Instrumented Flux Return. Subdetector of BaBar.
IP	Interaction Point.
IR	Interaction Region.
LAL	Laboratoire de l'Accélérateur Linéaire, Orsay, France.
LEP	Large Electron-Positron collider, CERN, Switzerland.
LER	Low Energy Ring. The positron storage ring at PEP-II.
LHC	Large Hadron Collider. Accelerator in progress at CERN.
MC	Monte Carlo. Technique for numerical integration and estimation.
MNS	Maki, Nakagana, Sakata. Inventors of the MNS lepton mixing matrix.
NBI	Niels Bohr Institute, Copenhagen.
PDF	Probability Density Function.
PEP	Positron Electron Project. Storage rings at SLAC.
PID	Particle IDentification.
PMT	Photo Multiplier Tube. Instrument for detecting photons.
POCA	Point Of Closest Approach. Point of shortest distance to point or track (e.g. beam spot).
PS	Phase Space.
QCD	Quantum Chromo Dynamics. Theory of the strong force.
QED	Quantum Electro Dynamics. Theory of the electromagnetic force.
QFT	Quantum Field Theory.
RPC	Resistive Plate Counter. Type of detector used in the IFR.
SCF	Self Cross Feed.
SLAC	Stanford Linear Accelerator Center, Stanford, USA.
SM	Standard Model.
SR	Synchrotron Radiation.
SVT	Silicon Vertex Tracker. Subdetector of BaBar.
TDC	Time to Digital Converter.
TDR	Technical Design Report.

The *BaBar* Collaboration

B. Aubert, R. Barate, D. Boutigny, F. Couderc, J.-M. Gaillard, A. Hicheur, Y. Karyotakis, J. P. Lees, V. Tisserand, A. Zghiche

Laboratoire de Physique des Particules, F-74941 Annecy-le-Vieux, France

A. Palano, A. Pompili

Università di Bari, Dipartimento di Fisica and INFN, I-70126 Bari, Italy

J. C. Chen, N. D. Qi, G. Rong, P. Wang, Y. S. Zhu

Institute of High Energy Physics, Beijing 100039, China

G. Eigen, I. Ofte, B. Stugu

University of Bergen, Inst. of Physics, N-5007 Bergen, Norway

G. S. Abrams, A. W. Borgland, A. B. Breon, D. N. Brown, J. Button-Shafer, R. N. Cahn, E. Charles, C. T. Day, M. S. Gill, A. V. Gritsan, Y. Groysman, R. G. Jacobsen, R. W. Kadel, J. Kadyk, L. T. Kerth, Yu. G. Kolomensky, G. Kukartsev, C. LeClerc, G. Lynch, A. M. Merchant, L. M. Mir, P. J. Oddone, T. J. Orimoto, M. Pripstein, N. A. Roe, M. T. Ronan, V. G. Shelkov, W. A. Wenzel

Lawrence Berkeley National Laboratory and University of California, Berkeley, CA 94720, USA

K. Ford, T. J. Harrison, C. M. Hawkes, S. E. Morgan, A. T. Watson

University of Birmingham, Birmingham, B15 2TT, United Kingdom

M. Fritsch, K. Goetzen, T. Held, H. Koch, B. Lewandowski, M. Pelizaeus, M. Steinke

Ruhr Universität Bochum, Institut für Experimentalphysik 1, D-44780 Bochum, Germany

J. T. Boyd, N. Chevalier, W. N. Cottingham, M. P. Kelly, T. E. Latham, F. F. Wilson

University of Bristol, Bristol BS8 1TL, United Kingdom

T. Cuhadar-Donszelmann, C. Hearty, N. S. Knecht, T. S. Mattison, J. A. McKenna, D. Thiessen

University of British Columbia, Vancouver, BC, Canada V6T 1Z1

A. Khan, P. Kyberd, L. Teodorescu

Brunel University, Uxbridge, Middlesex UB8 3PH, United Kingdom

V. E. Blinov, A. D. Bukin, V. P. Druzhinin, V. B. Golubev, V. N. Ivanchenko, E. A. Kravchenko, A. P. Onuchin, S. I. Serednyakov, Yu. I. Skovpen, E. P. Solodov, A. N. Yushkov

Budker Institute of Nuclear Physics, Novosibirsk 630090, Russia

D. Best, M. Bruinsma, M. Chao, I. Eschrich, D. Kirkby, A. J. Lankford, M. Mandelkern, R. K. Mommsen, W. Roethel, D. P. Stoker

University of California at Irvine, Irvine, CA 92697, USA

C. Buchanan, B. L. Hartfiel

University of California at Los Angeles, Los Angeles, CA 90024, USA

J. W. Gary, B. C. Shen, K. Wang

University of California at Riverside, Riverside, CA 92521, USA

D. del Re, H. K. Hadavand, E. J. Hill, D. B. MacFarlane, H. P. Paar, Sh. Rahatlou, V. Sharma

University of California at San Diego, La Jolla, CA 92093, USA

J. W. Berryhill, C. Campagnari, B. Dahmes, S. L. Levy, O. Long, A. Lu, M. A. Mazur, J. D. Richman, W. Verkerke

University of California at Santa Barbara, Santa Barbara, CA 93106, USA

T. W. Beck, A. M. Eisner, C. A. Heusch, W. S. Lockman, T. Schalk, R. E. Schmitz, B. A. Schumm, A. Seiden, P. Spradlin, D. C. Williams, M. G. Wilson

University of California at Santa Cruz, Institute for Particle Physics, Santa Cruz, CA 95064, USA

J. Albert, E. Chen, G. P. Dubois-Felsmann, A. Dvoretskii, D. G. Hitlin, I. Narsky, T. Piatenko, F. C. Porter, A. Ryd, A. Samuel, S. Yang

California Institute of Technology, Pasadena, CA 91125, USA

S. Jayatilleke, G. Mancinelli, B. T. Meadows, M. D. Sokoloff

University of Cincinnati, Cincinnati, OH 45221, USA

T. Abe, F. Blanc, P. Bloom, S. Chen, W. T. Ford, U. Nauenberg, A. Olivas, P. Rankin, J. G. Smith, J. Zhang, L. Zhang

University of Colorado, Boulder, CO 80309, USA

A. Chen, J. L. Harton, A. Soffer, W. H. Toki, R. J. Wilson, Q. L. Zeng

Colorado State University, Fort Collins, CO 80523, USA

D. Altenburg, T. Brandt, J. Brose, T. Colberg, M. Dickopp, E. Feltresi, A. Hauke, H. M. Lacker, E. Maly, R. Müller-Pfefferkorn, R. Nogowski, S. Otto, A. Petzold, J. Schubert, K. R. Schubert, R. Schwierz, B. Spaan, J. E. Sundermann

Technische Universität Dresden, Institut für Kern- und Teilchenphysik, D-01062 Dresden, Germany

D. Bernard, G. R. Bonneaud, F. Brochard, P. Grenier, S. Schrenk, Ch. Thiebaux, G. Vasileiadis, M. Verderi

Ecole Polytechnique, LLR, F-91128 Palaiseau, France

D. J. Bard, P. J. Clark, D. Lavin, F. Muheim, S. Playfer, Y. Xie

University of Edinburgh, Edinburgh EH9 3JZ, United Kingdom

M. Andreotti, V. Azzolini, D. Bettoni, C. Bozzi, R. Calabrese, G. Cibinetto, E. Luppi, M. Negrini, L. Piemontese, A. Sarti

Università di Ferrara, Dipartimento di Fisica and INFN, I-44100 Ferrara, Italy

E. Treadwell

Florida A}M University, Tallahassee, FL 32307, USA

R. Baldini-Ferroli, A. Calcaterra, R. de Sangro, G. Finocchiaro, P. Patteri, M. Piccolo, A. Zallo

Laboratori Nazionali di Frascati dell'INFN, I-00044 Frascati, Italy

A. Buzzo, R. Capra, R. Contri, G. Crosetti, M. Lo Vetere, M. Macri, M. R. Monge, S. Passaggio, C. Patrignani, E. Robutti, A. Santroni, S. Tosi

Università di Genova, Dipartimento di Fisica and INFN, I-16146 Genova, Italy

S. Bailey, G. Brandenburg, M. Morii, E. Won

Harvard University, Cambridge, MA 02138, USA

R. S. Dubitzky, U. Langenegger

Universität Heidelberg, Physikalisches Institut, Philosophenweg 12, D-69120 Heidelberg, Germany

W. Bhimji, D. A. Bowerman, P. D. Dauncey, U. Egede, J. R. Gaillard, G. W. Morton, J. A. Nash, G. P. Taylor

Imperial College London, London, SW7 2AZ, United Kingdom

G. J. Grenier, U. Mallik

University of Iowa, Iowa City, IA 52242, USA

J. Cochran, H. B. Crawley, J. Lamsa, W. T. Meyer, S. Prell, E. I. Rosenberg, J. Yi
Iowa State University, Ames, IA 50011-3160, USA

M. Davier, G. Grosdidier, A. Höcker, S. Laplace, F. Le Diberder, V. Lepeltier, A. M. Lutz, T. C. Petersen, S. Plaszczynski, M. H. Schune, L. Tantot, G. Wormser
Laboratoire de l'Accélérateur Linéaire, F-91898 Orsay, France

C. H. Cheng, D. J. Lange, M. C. Simani, D. M. Wright
Lawrence Livermore National Laboratory, Livermore, CA 94550, USA

A. J. Bevan, J. P. Coleman, J. R. Fry, E. Gabathuler, R. Gamet, R. J. Parry, D. J. Payne, R. J. Sloane, C. Touramanis
University of Liverpool, Liverpool L69 72E, United Kingdom

J. J. Back, C. M. Cormack, P. F. Harrison,[89] G. B. Mohanty
Queen Mary, University of London, E1 4NS, United Kingdom

C. L. Brown, G. Cowan, R. L. Flack, H. U. Flaecher, M. G. Green, C. E. Marker, T. R. McMahon, S. Ricciardi, F. Salvatore, G. Vaitsas, M. A. Winter
University of London, Royal Holloway and Bedford New College, Egham, Surrey TW20 0EX, United Kingdom

D. Brown, C. L. Davis
University of Louisville, Louisville, KY 40292, USA

J. Allison, N. R. Barlow, R. J. Barlow, P. A. Hart, M. C. Hodgkinson, G. D. Lafferty, A. J. Lyon, J. C. Williams
University of Manchester, Manchester M13 9PL, United Kingdom

A. Farbin, W. D. Hulsbergen, A. Jawahery, D. Kovalskyi, C. K. Lae, V. Lillard, D. A. Roberts
University of Maryland, College Park, MD 20742, USA

G. Blaylock, C. Dallapiccola, K. T. Flood, S. S. Hertzbach, R. Kofler, V. B. Koptchev, T. B. Moore, S. Saremi, H. Staengle, S. Willocq
University of Massachusetts, Amherst, MA 01003, USA

R. Cowan, G. Sciolla, F. Taylor, R. K. Yamamoto
Massachusetts Institute of Technology, Laboratory for Nuclear Science, Cambridge, MA 02139, USA

D. J. J. Mangeol, P. M. Patel, S. H. Robertson
McGill University, Montréal, QC, Canada H3A 2T8

A. Lazzaro, F. Palombo
Università di Milano, Dipartimento di Fisica and INFN, I-20133 Milano, Italy

J. M. Bauer, L. Cremaldi, V. Eschenburg, R. Godang, R. Kroeger, J. Reidy, D. A. Sanders, D. J. Summers, H. W. Zhao
University of Mississippi, University, MS 38677, USA

S. Brunet, D. Côté, P. Taras
Université de Montréal, Laboratoire René J. A. Lévesque, Montréal, QC, Canada H3C 3J7

H. Nicholson
Mount Holyoke College, South Hadley, MA 01075, USA

[89]Now at Department of Physics, University of Warwick, Coventry, United Kingdom

N. Cavallo, F. Fabozzi, C. Gatto, L. Lista, D. Monorchio, P. Paolucci, D. Piccolo, C. Sciacca

Università di Napoli Federico II, Dipartimento di Scienze Fisiche and INFN, I-80126, Napoli, Italy

M. Baak, H. Bulten, G. Raven, L. Wilden

NIKHEF, National Institute for Nuclear Physics and High Energy Physics, NL-1009 DB Amsterdam, The Netherlands

C. P. Jessop, J. M. LoSecco

University of Notre Dame, Notre Dame, IN 46556, USA

T. A. Gabriel

Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

T. Allmendinger, B. Brau, K. K. Gan, K. Honscheid, D. Hufnagel, H. Kagan, R. Kass, T. Pulliam, A. M. Rahimi, R. Ter-Antonyan, Q. K. Wong

Ohio State University, Columbus, OH 43210, USA

J. Brau, R. Frey, O. Igonkina, C. T. Potter, N. B. Sinev, D. Strom, E. Torrence

University of Oregon, Eugene, OR 97403, USA

F. Colecchia, A. Dorigo, F. Galeazzi, M. Margoni, M. Morandin, M. Posocco, M. Rotondo, F. Simonetto, R. Stroili, G. Tiozzo, C. Voci

Università di Padova, Dipartimento di Fisica and INFN, I-35131 Padova, Italy

M. Benayoun, H. Briand, J. Chauveau, P. David, Ch. de la Vaissière, L. Del Buono, O. Hamon, M. J. J. John, Ph. Leruste, J. Ocariz, M. Pivk, L. Roos, S. T'Jampens, G. Therin

Universités Paris VI et VII, Lab de Physique Nucléaire H. E., F-75252 Paris, France

P. F. Manfredi, V. Re

Università di Pavia, Dipartimento di Elettronica and INFN, I-27100 Pavia, Italy

P. K. Behera, L. Gladney, Q. H. Guo, J. Panetta

University of Pennsylvania, Philadelphia, PA 19104, USA

F. Anulli, M. Biasini, I. M. Peruzzi,[2] M. Pioppi

Università di Perugia, Dipartimento di Fisica and INFN, I-06100 Perugia, Italy

C. Angelini, G. Batignani, S. Bettarini, M. Bondioli, F. Bucci, G. Calderini, M. Carpinelli, V. Del Gamba, F. Forti, M. A. Giorgi, A. Lusiani, G. Marchiori, F. Martinez-Vidal, M. Morganti, N. Neri, E. Paoloni, M. Rama, G. Rizzo, F. Sandrelli, J. Walsh

Università di Pisa, Dipartimento di Fisica, Scuola Normale Superiore and INFN, I-56127 Pisa, Italy

M. Haire, D. Judd, K. Paick, D. E. Wagoner

Prairie View A&M University, Prairie View, TX 77446, USA

N. Danielson, P. Elmer, Y. P. Lau, C. Lu, V. Miftakov, J. Olsen, A. J. S. Smith, A. V. Telnov

Princeton University, Princeton, NJ 08544, USA

F. Bellini G. Cavoto, R. Faccini, F. Ferrarotto, F. Ferroni, M. Gaspero, L. Li Gioi, M. A. Mazzoni, S. Morganti, M. Pierini, G. Piredda, F. Safai Tehrani, C. Voena

Università di Roma La Sapienza, Dipartimento di Fisica and INFN, I-00185 Roma, Italy

S. Christ, G. Wagner, R. Waldi

Universität Rostock, D-18051 Rostock, Germany

T. Adye, N. De Groot, B. Franek, N. I. Geddes, G. P. Gopal, E. O. Olaiya

Rutherford Appleton Laboratory, Chilton, Didcot, Oxon, OX11 0QX, United Kingdom

R. Aleksan, S. Emery, A. Gaidot, S. F. Ganzhur, P.-F. Giraud, G. Hamel de Monchenault, W. Kozanecki, M. Langer, M. Legendre, G. W. London, B. Mayer, G. Schott, G. Vasseur, Ch. Yèche, M. Zito

DSM/Dapnia, CEA/Saclay, F-91191 Gif-sur-Yvette, France

M. V. Purohit, A. W. Weidemann, F. X. Yumiceva

University of South Carolina, Columbia, SC 29208, USA

D. Aston, R. Bartoldus, N. Berger, A. M. Boyarski, O. L. Buchmueller, M. R. Convery, M. Cristinziani, G. De Nardo, D. Dong, J. Dorfan, D. Dujmic, W. Dunwoodie, E. E. Elsen, S. Fan, R. C. Field, T. Glanzman, S. J. Gowdy, T. Hadig, V. Halyo, C. Hast, T. Hryn'ova, W. R. Innes, M. H. Kelsey, P. Kim, M. L. Kocian, D. W. G. S. Leith, J. Libby, S. Luitz, V. Luth, H. L. Lynch, H. Marsiske, R. Messner, D. R. Muller, C. P. O'Grady, V. E. Ozcan, A. Perazzo, M. Perl, S. Petrak, B. N. Ratcliff, A. Roodman, A. A. Salnikov, R. H. Schindler, J. Schwiening, G. Simi, A. Snyder, A. Soha, J. Stelzer, D. Su, M. K. Sullivan, J. Va'vra, S. R. Wagner, M. Weaver, A. J. R. Weinstein, W. J. Wisniewski, M. Wittgen, D. H. Wright, A. K. Yarritu, C. C. Young

Stanford Linear Accelerator Center, Stanford, CA 94309, USA

P. R. Burchat, A. J. Edwards, T. I. Meyer, B. A. Petersen, C. Roat

Stanford University, Stanford, CA 94305-4060, USA

S. Ahmed, M. S. Alam, J. A. Ernst, M. A. Saeed, M. Saleem, F. R. Wappler

State Univ. of New York, Albany, NY 12222, USA

W. Bugg, M. Krishnamurthy, S. M. Spanier

University of Tennessee, Knoxville, TN 37996, USA

R. Eckmann, H. Kim, J. L. Ritchie, A. Satpathy, R. F. Schwitters

University of Texas at Austin, Austin, TX 78712, USA

J. M. Izen, I. Kitayama, X. C. Lou, S. Ye

University of Texas at Dallas, Richardson, TX 75083, USA

F. Bianchi, M. Bona, F. Gallo, D. Gamba

Università di Torino, Dipartimento di Fisica Sperimentale and INFN, I-10125 Torino, Italy

C. Borean, L. Bosisio, C. Cartaro, F. Cossutti, G. Della Ricca, S. Dittongo, S. Grancagnolo, L. Lanceri, P. Poropat,[90] L. Vitale, G. Vuagnin

Università di Trieste, Dipartimento di Fisica and INFN, I-34127 Trieste, Italy

R. S. Panvini

Vanderbilt University, Nashville, TN 37235, USA

Sw. Banerjee, C. M. Brown, D. Fortin, P. D. Jackson, R. Kowalewski, J. M. Roney

University of Victoria, Victoria, BC, Canada V8W 3P6

H. R. Band, S. Dasu, M. Datta, A. M. Eichenbaum, M. Graham, J. J. Hollar, J. R. Johnson, P. E. Kutter, H. Li, R. Liu, F. Di Lodovico, A. Mihalyi, A. K. Mohapatra, Y. Pan, R. Prepost, A. E. Rubin, S. J. Sekula, P. Tan, J. H. von Wimmersperg-Toeller, J. Wu, S. L. Wu, Z. Yu

University of Wisconsin, Madison, WI 53706, USA

H. Neal

Yale University, New Haven, CT 06511, USA

[90]Deceased

Acknowledgement

Scientific Acknowledgement

First and foremost I'm endlessly endepted to my thesis advisor François Le Diberder. From the moment you accepted to have me as a visting master student in the Spring of 2000 at Jussieu until now, you have been a source of knowledge and inspiration. It is a rare but most exhilarating when sorceror and apprentice share temperament and humor. Thank you for all you have taught me both of physics, statistics, and beyond, and thank you for all the enjoyable discussions, wagers, dinners, etc. we've shared.
However, the analysis here presented would not have been, were it not for Marie-Helene Schune, whom I've had the pleasure of working with for most of my thesis' duration. I'm grateful not only for your efforts, but also for you direct yet subtle attitude, and the quality of your approaches. Our mixture of temperament seem to have struck a both pleasant and productive cocktail.
I'm thankful to Valeria Tano who contributed to the analysis in its initial phase and gave it German efficiency and Italian style. People are fortunate in Munich.
Thanks to Malcolm John, who caried out the computational work of skimming the data used. It has been a pleasure to work, discuss, and share long hours with you.

In addition I'm very much indebted to the *BABAR* group at LAL. Thank you Anne-Marie for your kindness and help, especially when facing French administration. Thank you Andreas for being a rich source of answers and for introducing me to neural networks. Thank you Achille for exciting discussions on γ.
Thank you Sandrine for so many enjoyable years of working side by side[91]. Thank you Laurant for joyful sharing of SLAC office and for your introduction to French slang and costumes. Thank you to Gilbert, Vincent, Michel, and Patrick for an enjoyable and inspiring atmosphere.
In also I owe thanks to Guillaume, Julie, Jacques, Jose, and Lydia from LPNHE, and not the least Muriel Pivk for the implementation of the beautiful ${}_s\mathcal{P}$lots.

Beyond the realms of France, I'm grateful for the pleasant and inspiring collaboration I've had with Roy Aleksan and Abi Soffer while working on the $B \to DK\pi$ idea. In addition the quality of the analysis have been improved by the insight and judgement of reviewers Justin Albert, Fred Blanc, and Claudio Campagnari.
I'm very much indebted to Jean Iliopoulos, Brigitte Bloch-Devaux, and Étienne Augé for having taken the burden of serving as thesis jury upon their shoulders, and even more so to Livio Lanceri and Olivier Schneider for committing to the task of being main opponents (fr: rapporteurs) in the most kind and friendly fashion. Your perspective and feedback have been extremely valuable, and you have rendered the thesis submission and defence a most pleasant experience.

It was an unforseen but truely rewarding pleasure to serve as DIRC commissioner, and I would like to thank especially Thomas Hadig, Joshen (Joe) Schweining, Blair Ratcliff, Brian Meadows, and Jerry Vav'ra for this wonderful experience.
I'm proud to have been part of the IR-2 team, and grateful to have worked with people as knowledgeble as Chris O'Grady, Livio Piemonte, Giovani Caldrini, Reiner Bartoldous, and countless others.
The SLAC environment has been a source of inspiration, not the least because of Vivek Sharma, Shahram Rahatlou, Nick Danielson, Danielle Del Re, Gabrielle Simi, and Francesca Bucci. However, days would not have been the same without the humor and hospitality of my friends at *BABAR* Alex Olivas, Christiano Borean, Domenique Lavin, and Ed Hill. Also thanks to Moishe Pripstein and Bob Cahn for having me (and Thomas Kittelmann) at LBNL.

[91] At the time of writing, Sandrine owes me a dinner at Tour d'Argent!

Throughout my distant studies, the HEP group at the Niels Bohr Institute in Copenhagen has always provided me with a base to return to, for which I'm extremely grateful. Thank you John R. Hansen, for accepting the position as my contact person, and thanks to the rest of the group for the most pleasant of work environments. A special thank to Thomas Kittelmann, who trusted me in choosing *BABAR* for his master thesis. Thank you for the quality of your work and the joyful (late) hours of working at LBNL.

I'm very grateful to have worked on the *BABAR* experiment with the people there, who not only encompases talent indescribable in this short acknowledgement, but also a rich mixture of languages, cultures, and traditions. Though endepted countless helpful and guiding persons, solely I stand to account for the faults and errors of this thesis, as it was written by my hand.

Finally a – perhaps clumpsy but certainly well meant – thanks to Nature for the many wonder she holds, and an appreciation of living in a world, which have the resources and will to support researchers venturing into the unknown.

Personal Acknowledgement

Ph.D. studies in experimental particle physics is a demanding task, which takes a toll on social life and tests the patience of others. Were it not for the support and encouragement from family and friends, this thesis would never have been.

Starting a new life in San Francisco would have been much more dull, where it not for especially Thomas, Torben, and Malene, but also Joe and Charles. I'm also indepted to George Rygg for much sailing on the bay, and to all of those, who visited my beloved lairs on Bernal Hill.

The feast in Paris would have been much less lively, were it not for all of those I've shared Parisian life with in the small boheme apartment in Rue de la Ferronnerie, most notably Kaare and Adam, but also Anders, Slag, Pernille, Monica, Nicolai, Tina, Johan, and Andreas along with a many others. It has been an almost poetic pleasure to indulge in Paris' pleasures with you. I'm also very thankful to my French friends, especially Loan, Sebastien, and Nico, who have opened Paris and France to me, and who gives me all the excuses I need to return. In addition I've had the pleasure of spending many hours at Fondation Danoise, who have always shown me great hospitality.

Though living abroad, the longing for my native city of dreams has never been allayed. And when I've had the opportunity to return, friends, not the least those from Regensen, have always made it blissful.

Encompassing the most enjoyable places on earth, the behilled city by the bebridged bay, the city of lights, and the city of dreams, have laid the scene for a tale of three cities.

Finally, I would like to thank my wonderful family. Without, I would not have had the courage to follow the path I have. Not only have I always had their support, but they've also inspired and encouraged me to persue my passion for science.

Troels C. Petersen
19th of May 2004