SLAC-R-966

MAURIZIO MARTINELLI

SEARCH FOR CP VIOLATION IN SINGLY CABIBBO SUPPRESSED FOUR-BODY D DECAYS

PH. D. THESIS

UNIVERSITÀ DEGLI STUDI DI BARI ALDO MORO

Supervisor: Chiar.mo Prof. A. Palano

Università degli Studi di Bari
Facoltà di Scienze Matematiche Fisiche e Naturali
Dipartimento di Fisica
February 2011

SLAC National Accelerator Laboratory, Menlo Park, CA 94025
Work supported in part by US Department of Energy contract DE-AC02-76SF00515.

Maurizio Martinelli: *Search for* CP *violation in singly Cabibbo suppressed four-body* D *decays,* Ph. D. Thesis, © february 2011.

"From a long view of the history of mankind — seen from, say, ten thousand years from now, there can be little doubt that the most significant event of the 19th century will be judged as Maxwell's discovery of the laws of electrodynamics. The American Civil War will pale into provincial insignificance in comparison with this important scientific event of the same decade."

Richard Phillips Feynman

For Giuliana, My Family and My Friends.

ABSTRACT

We search for CP violation in a sample of 4.7×10^4 singly Cabibbo suppressed $D^0 \to K^+ K^- \pi^+ \pi^-$ decays and $1.8(2.6) \times 10^4$ $D^+_{(s)} \to K^0_S K^+ \pi^+ \pi^-$ decays. CP violation is searched for in the difference between the T-odd asymmetries, obtained using triple product correlations, measured for D and $\bar{D}$ decays. The measured CP violation parameters are $\mathcal{A}_T(D^0) = (1.0 \pm 5.1(\text{stat}) \pm 4.4(\text{syst})) \times 10^{-3}$, $\mathcal{A}_T(D^+) = (-11.96 \pm 10.04(\text{stat}) \pm 4.81(\text{syst})) \times 10^{-3}$ and $\mathcal{A}_T(D^+_s) = (-13.57 \pm 7.67(\text{stat}) \pm 4.82(\text{syst})) \times 10^{-3}$.

PREVIOUS PUBLICATIONS

The analysis on D^0 decays shown in this thesis has been published on [del Amo Sanchez *et al.*, 2010]. A similar analysis on D^+ and D_s^+ decays is going toward publication soon.

ACKNOWLEDGMENTS

First and foremost I wish to thank my supervisor Antimo Palano. He has been incredibly supportive of my scientific pursuits, and I am very grateful for the many opportunities he has afforded during my graduate work. He is an impressive physicist, and his advice and suggestions have been extremely valuable in advancing this analysis and my career.

I wish to thank Daniele Pedrini and Fulvia De Fazio for having reviewed this thesis. Their suggestions and comments greatly improved the quality of this work.

I also wish to thank all the *BABAR* Collaboration's members, especially the Charm Analysis Working Group, and among them Bill Dunwoodie, Jon Coleman, Nicola Neri, Milind Purohit, Bill Lockman and Brian Meadows. Their contribution has been fundamental to define all the details in the analysis. The *BABAR*'s review committee members, Mike Sokoloff, Ray Cowan, David Lange, Owen Long, Chunhui Chen and Elisa Guido, for the attention they put on my analyses and the writing of the article that has been published.

I then wish to thank all the Ph. D. students and post-Doc in *BABAR* Marco Pappagallo, Alfio Lazzaro, Diego Milanes, Vincenzo Lombardo, David Lopes Pegna, José Benitez, Kevin Yarritu, Simone Stracka, Pietro Biassoni, Elisa Guido, Valentina Santoro, Elisabetta Prencipe, Alessandro Gaz and all the others babarians, for their precious suggestions and the funny time spent together.

Last but not least my fellows in Bari, Domenico, Floriana, Stefano, Francesco, Antonella, Irene e Marco, for the precious time spent together and the mutual support with bureaucracy.

Finally, a special thank to Giuliana and my Family, for the extraordinary support, the education and the love I never missed.

Bari, february 2011 M. M.

RINGRAZIAMENTI

Per primo e più di tutti vorrei ringraziare Antimo Palano. Il suo supporto per il mio lavoro di ricerca è stato fondamentale e gli sono estremamente grato per tutte le opportunità che mi ha offerto durante il mio lavoro di dottorato. Le sue qualità come fisico sono eccezionali, e il suo consiglio e i suoi suggerimenti sono stati di enorme impatto nell'avanzamento di questa analisi e del mio percorso professionale.

Vorrei ringraziare Daniele Pedrini e Fulvia De Fazio per avere revisionato questa tesi. I loro suggerimenti e commenti sono stati preziosi per migliorare la qualità di questo testo.

Vorrei ringraziare anche tutti i membri della Collaborazione *BABAR* e in particolare del gruppo di analisi del Charm, tra cui Bill Dunwoodie, Jon Coleman, Nicola Neri, Milind Purohit, Bill Lockman and Brian Meadows. Il loro contributo è stato fondamentale per definire tutti i dettagli dell'analisi. I membri dei comitati di revisione interna a *BABAR* Mike Sokoloff, Ray Cowan, David Lange, Owen Long, Chunui Chen e Elisa Guido, per l'attenzione che hanno dedicato alla mia analisi e alla realizzazione dell'articolo che ne è conseguito.

Vorrei quindi ringraziare tutti i miei colleghi dottorandi e post-Doc in *BABAR* Marco Pappagallo, Alfio Lazzaro, Diego Milanes, Vincenzo Lombardo, David Lopes Pegna, José Benitez, Kevin Yarritu, Simone Stracka, Pietro Biassoni, Elisa Guido, Valentina Santoro, Elisabetta Prencipe, Alessandro Gaz e tutti gli altri babariani, per i preziosi suggerimenti e il piacevole tempo passato insieme.

Infine, ma non per ultimi i miei colleghi dottorandi e amici di Bari, Domenico, Floriana, Stefano, Francesco, Antonella, Irene e Marco, per i divertenti momenti passati insieme e il vicendevole supporto nelle questioni burocratiche.

Per ultimo un ringraziamento speciale va a Giuliana e alla mia famiglia, per lo straordinario supporto, l'educazione e l'amore che non mi hanno fatto mai mancare.

Bari, febbraio 2011 M. M.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

Man has always had a deep interest for symmetries. From the more ancient times the pleasure of building artifacts stimulating our sense of symmetry has represented the main purpose of fine arts.

This interest has reflected in the developing of mathematics and specifically geometry: the former defining the tools to build the latter, both ideally inspired from perfection.

Anyway, it would be unfair to state that the concept of symmetry is innate in human mind: it has been developed from the observation of nature, that provides us many examples of geometrical symmetries that have inspired human beings from the beginning of history. Just think about the wings of a butterfly, the structure of a leaf and even the apparent symmetry among the left and right ends of the human body.

Such a concept is so much well represented in what we see from our first days of life, that is not avoidable in our thinking. That is why we defined it even in Physics, without any suspect of how much fundamental would be for the description of the deep mechanics that rules the world where we live.

The first and more natural kind of symmetry that we defined in Physics is Parity (P), the reflection of space coordinates from $\vec{x}$ to $-\vec{x}$. Similarly we defined another discrete symmetry, the Time Reversal (T), which changes t in $-t$, reverting the motion. As soon as the concept of anti-particle was introduced by the Dirac equation [Dirac, 1928], the Charge conjugation (C), transforming a particle into its anti-particle, has been studied.

We have learnt that nature is largely, but not completely, invariant under these transformations. Although these insights were at first less than eagerly accepted by the Physics community, they form an essential element of what is called the Standard Model (SM) of particle Physics. This model contains 20 odd parameters, a feature that cannot be accidental, but must be shaped by some intriguing New Physics. The scientific community considers it very likely that a comprehensive study of how these discrete symmetries are implemented in nature will reveal the intervention of New Physics [Bigi e Sanda, 2000].

Furthermore, we believe that time reversal and the combined transformation of CP occupy a very unique place in the pantheon of symmetries. Once charge conjugation and parity were found to be maximally violated [Wu *et al.*, 1957], their combination CP was apparently conserved. This point was sustained from the local T invariance derived from Mach's principle and from the CPT invariance that is natural in quantum field theory [Bjorken e Drell, 1965]: no CP violation is then allowed if T violation is not found.

It was then a big shock in Physics' community when the CP violation was found [Christenson *et al.*, 1964], following the prescription of the

CPT theorem, that also means the T is violated, which substantially means that nature makes difference between past and future even in microscopic reactions.

There are other aspects that make CP and T violation worth to be studied in detail:

- from CP violation a characterization of matter and anti-matter that is more general then the one obtained from C violation can be retrieved. C violation allows to separate matter and anti-matter from the reaction that involves left and right neutrinos only. This is anyway a convention that could be reversed until CP is conserved. Once CP is violated, matter and anti-matter can be defined in an absolute way: among the K^0_L decays for example, $K^0_L \to e^+ \nu_e \pi^-$ is favored respect to its CP conjugate [Nakamura *et al.*, 2010], allowing to define the positron as the anti-electron using a hint from nature.
- The anti-unitarity of the T operator of symmetry, together with Kramers' degeneracy allow to observe two eigenvalues for T^2: $T^2 = +1$ corresponds to boson states, $T^2 = -1$ to fermions.
- CP violation is an excellent phenomenological probe. It was discovered in 1964 studying $K^0_L \to \pi^+\pi^-$ and included in the Standard Model by the Kobayashi-Maskawa ansatz [Kobayashi e Maskawa, 1973] in 1973. This ansatz allowed to forecast a third family of leptons and quarks that have been confirmed in the following years.
- We can still not give the notion of maximal CP violation, as well as it can be done for P and C using the right and left-handed neutrino convention. As far as we know, CP violation is allowed in the Standard Model with three quark families, but we cannot assign any meaning to maximal CP violation.
- Finally, any attempt to describe the asymmetry between baryons and anti-baryons in the universe [Sakharov, 1967], is strictly related to CP violation.

From all of the above items, the validity of a wide range Physics program to understand CP and T violation follows up.

The work outlined in this thesis contributes to this program searching for CP violation in two singly Cabibbo suppressed and one Cabibbo favored four-body D decay modes. The specific final states that are subject of the analysis are

- $D^0 \to K^+ K^- \pi^+ \pi^-$ (singly Cabibbo suppressed);
- $D^+ \to K^+ K^0_S \pi^+ \pi^-$ (singly Cabibbo suppressed);
- $D^+_s \to K^+ K^0_S \pi^+ \pi^-$ (Cabibbo favored).

After a brief introduction on the search for CP violation in D decays, the main interesting features of the *BABAR* detector are outlined; then the analysis made on D^0 decays is shown; the work made on D^+ and D^+_s decays is described together in the section following the D^0 analysis.

Finally, the results of these analyses are compared to the results obtained by previous analyses from different experiments, providing the ground for a discussion of the results.

2 CP VIOLATION AND D DECAYS

CONTENTS

2.1 CP VIOLATION

CP is an operation obtained combining the two discrete transformations of Parity P and Charge conjugation C. Even if the conservation of P and C was known to be maximally violated, their combination CP was supposed to be conserved. This assertion was supported by the CPT theorem, that naturally establish the invariance of any quantum field theory when the three discrete symmetries are applied. Following this theorem, observing CP violation, T violation is straightforward and vice-versa. But, if T is violated in fundamental processes, then the assertion that locally past and future are not distinguishable falls down. Without reminding the central role of CP violation within both the Standard Model and any alternative model developed to extend it, in the following some details on CP violation and on the searches for CP violation in D decays will be outlined.

2.1.1 First Observation of CP Violation

CP violation was found in 1964, by means of the observation of $K^0_L \to \pi^+\pi^-$ decay [Christenson *et al.*, 1964]. To understand why this observation has been so crucial, let us suppose that K^0_S and K^0_L are CP $= \pm 1$ eigenstates. K^0_S and K^0_L are named from their highly different mean life ($\tau_{K^0_S} = 0.89 \times 10^{-10}$s and $\tau_{K^0_L} = 5.12 \times 10^{-8}$s) are made of the superposition of K^0 and $\bar{K}^0$ states, that are not CP eigenstates. Unlike K^0 and $\bar{K}^0$, distinguished by their mode of production, K^0_S and K^0_L are distinguished by their decay modes: $\pi^+\pi^-$ and $\pi^0\,\pi^0$ are indeed CP =1

while $\pi^+\ \pi^-\ \pi^0$ and $\pi^0\ \pi^0\ \pi^0$ are CP =-1 states. These statements follow up when considering Bose symmetry, the Q-value of the reaction and the intrinsic Parity of the pions [Perkins, 1982]. The observation of the $K_L^0 \to \pi^+\pi^-$ decay is then a clear signal of CP violation.

2.1.2 Standard Model

For the introduction of the KM ansatz, the two scientists have been awarded the Nobel Prize in Physics for the year 2008

CP violation is included in the Standard Model by means of the Kobayashi-Maskawa (KM) ansatz [Kobayashi e Maskawa, 1973]. In the KM ansatz the CP violation is introduced through a quark mixing matrix V to describe the quark states as a mixture of mass eigenstates. In their paper, the authors try to introduce the CP violation in a model in which the interacting quarks form a quartet: the higher mass quarks c, b and t were still unknown at the time. After struggling to introduce CP violation using a four quark model, they observe that CP violation is naturally introduced in the quark mixing matrix if three families of two quarks are considered. It can be shown that the number of rotational angles and the number of phases to define a matrix that is unitary and orthogonal depend on n, the number of quark families:

$$N_{\text{angles}} = \frac{1}{2}n(n-1), \qquad N_{\text{phases}} = \frac{1}{2}(n-1)(n-2). \tag{2.1}$$

Assuming $n = 2$ there is only one rotational angle, the Cabibbo angle θ_C [Cabibbo, 1963], and no phase term. Introducing another family of quarks, $n = 3$ and the rotational angles become $N_{\text{angles}} = 3$ and an irreducible phase makes its appearance: $N_{\text{phases}} = 1$.

This phase is responsible for the introduction of CP violation in the quark mixing matrix, also known as Cabibbo-Kobayashi-Maskawa (CKM) matrix:

$$\begin{bmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{bmatrix} \Rightarrow$$
$$\begin{bmatrix} 1-\frac{\lambda^2}{2} & \lambda & A\lambda^3(\bar{\rho}-i\bar{\eta}+\frac{i}{2}\eta\lambda^2) \\ -\lambda & 1-\frac{\lambda^2}{2}-i\eta A^2\lambda^4 & A\lambda^2(1+i\eta\lambda^2) \\ A\lambda^3(1-\bar{\rho}-i\bar{\eta}) & -A\lambda^2 & 1 \end{bmatrix} \tag{2.2}$$

here shown in the Wolfenstein representation, approximated at the order $\mathcal{O}(\lambda^5)$, where $\lambda = \sin\theta_C$ [Wolfenstein, 1983]. In this representation, the phase appears in the matrix elements as the complex component and can be identified in V_{ub}, V_{td}, V_{cb} and V_{cs}.

The CP violating effect would show up only if the decay amplitude is the sum of two different parts, whose phases are made of a weak (CKM) and a strong (final state interaction) contribution. The weak phases change sign when going to the CP-conjugate process, while the strong ones do not. A decay amplitude of this type can be written as

$$\mathcal{A} = Ae^{i\delta_1} + Be^{i\delta_2}, \tag{2.3}$$

while the corresponding charged-conjugate is

$$\bar{\mathcal{A}} = A^*e^{i\delta_1} + B^*e^{i\delta_2}. \tag{2.4}$$

For both of them A and B represent the amplitude with the invariant strong phase divided out and δ_i is the strong phase itself. The CP violating asymmetry in the decay rate would be therefore

$$a_{CP} = \frac{|\mathcal{A}|^2 - |\bar{\mathcal{A}}|^2}{|\mathcal{A}|^2 + |\bar{\mathcal{A}}|^2} = \frac{2\Im(AB^*)\sin(\delta_2 - \delta_1)}{|A|^2 + |B|^2 + 2\Re(AB^*)\cos(\delta_2 - \delta_1)}. \quad (2.5)$$

In D decays, there are three kind of processes that can interfere and generate a CP asymmetry:

- $c \to s\bar{s}u$;
- $c \to d\bar{d}u$;
- $c \to u\sum q\bar{q}$.

The last one is represented by a penguin diagram, while the others are charged current ones. All of these processes contribute to singly Cabibbo suppressed D decays; no CP violation is expected in Cabibbo favored or doubly Cabibbo suppressed decays. The effect of CP violation in D decays has been calculated to have a limit of 0.1% [Buccella *et al.*, 1995].

There can be effects of indirect CP violation that need to be accounted for. All the decays involving a neutral kaon are subjected to the asymmetry introduced by the K^0-$\bar{K}^0$ mixing. The effect of this mixing on the asymmetry is $a_{CP}^{K^0 mix} = -3.3 \times 10^{-3}$ [Nakamura *et al.*, 2010].

2.1.3 Alternative Models

The relative smallness of the effect in the Standard Model allows to state that the observation of CP violation at the order of 1% in D decays is a strong evidence for the existence of processes that are not included in the Standard Model.

There are two ways to introduce CP violation in D decays [Grossman *et al.*, 2007] in scenarios beyond the Standard Model:

- Direct CP violation at tree-level: extra quark in Standard Model vector-like representation; supersymmetry without R-parity models; two Higgs doublet models.
- Direct CP violation at one-loop: QCD penguin and dipole operators; Flavor Changing Neutral Currents in supersymmetric flavor models.

While the first group of models can produce an effect that is much less than 1%, the processes having one-loop can even reach the percent level, producing effects that are clearly not expected in the Standard Model.

2.2 SUMMARY OF THE SEARCH FOR CP VIOLATION IN CHARM DECAYS

In the recent years, many searches for CP violation have been performed on charm decays. Four kind of analyses can be recognised:

- search for direct CP violation;
- Dalitz plot analysis;
- time dependent analysis;
- study of T-odd correlations.

The details of each one of these techniques are outlined in the following.

2.2.1 Direct CP violation

The search for direct CP violation is made by looking for an asymmetry in the number of D and $\bar{D}$ events decaying into a specific final state f. The observable for direct CP violation can be then defined as

$$A_{CP} = \frac{N_D - N_{\bar{D}}}{N_D + N_{\bar{D}}}, \tag{2.6}$$

where N_D and $N_{\bar{D}}$ indicate respectively the number of $D \to f$ and $\bar{D} \to \bar{f}$ events.

Many measurements have been performed at the asymmetric B-factories, such as *BABAR* and *BELLE*. In these experiments, there is a well known effect, the forward-backward asymmetry, that could bias these measurement by introducing an asymmetry that is due to the interference between the weak and the electromagnetic $e^+e^- \to c\bar{c}$ production processes. For further information see Appendix A. There are two solutions to remove this bias:

1. evaluate the amount of forward-backward asymmetry itself;
2. delete the effect by normalizing the number of D and $\bar{D}$ events using Cabibbo favored decay modes.

The former solution leads to measure the amount of forward-backward asymmetry A_{FB} and evaluate the direct CP asymmetry as

$$A_{CP} = a_{CP} - A_{FB}, \tag{2.7}$$

being a_{CP} the asymmetry measured from the data. The latter solution can be exploited by measuring the observable

$$A_{CP}^{(norm)} = \frac{R_D - R_{\bar{D}}}{R_D + R_{\bar{D}}}, \qquad R = \frac{N_{CS}}{N_{CF}}. \tag{2.8}$$

2.2.2 Dalitz Plot Analysis

There are many quantities that can be measured in a Dalitz plot analysis to search for CP violation:

- the Dalitz plot histogram can be compared bin-per-bin between D and $\bar{D}$ to search for asymmetries;
- similarly the distribution of the angular moments can be compared;
- once determined the model that best fits the data, the amplitudes of each component can be compared between D and $\bar{D}$.

2.2.3 Time Dependent Analysis

The study of the D decay times and the opportunity to search for CP violation is strictly related to D mixing [Bergmann *et al.*, 2000]. It is indeed known that D mixing affects decay times

$$\begin{aligned}\tau^{+}_{hh} &= \tau_{K\pi}[1 + r_m(y\cos\phi_f - x\sin\phi_f)]^{-1},\\ \tau^{-}_{hh} &= \tau_{K\pi}[1 + r_m^{-1}(y\cos\phi_f - x\sin\phi_f)]^{-1},\end{aligned} \tag{2.9}$$

where $h = K$ or π, $\tau_{K\pi}$ is the lifetime for the Cabibbo favored $D^0 \to K^- \pi^+$ decay. x, y, r_m and ϕ_f are defined considering that the two D^0 mass eigenstates can be represented as

$$\begin{aligned}|D_1> &= p|D^0> + q|\bar{D}^0>,\\ |D_2> &= p|D^0> - q|\bar{D}^0>,\end{aligned} \tag{2.10}$$

where $p^2 + q^2 = 1$. The size of D^0 mixing is traditionally quantified in terms of the parameters

$$x \equiv \frac{\Delta m}{\Gamma} \quad \text{and} \quad y \equiv \frac{\Delta\Gamma}{2\Gamma}, \tag{2.11}$$

where $\Delta m = m_1 - m_2$, $\Delta\Gamma = \Gamma_1 - \Gamma_2$ and $\Gamma = (\Gamma_1 + \Gamma_2)/2$, being m_i and Γ_i the masses and the widths of the two mass eigenstates.

The more relevant parameters for CP violation are

$$r_m \equiv \left|\frac{q}{p}\right| \quad \text{and} \quad \phi_f \equiv \arg\left(\frac{q}{p}\frac{\bar{A}_f}{A_f}\right), \tag{2.12}$$

where $A_f =< f|\mathcal{H}|D^0 >$ ($\bar{A}_f =< f|\mathcal{H}|\bar{D}^0 >$), f being the final state the D is decaying to. Observing $r_m \neq 1$ would indicate CP violation in mixing, while a non-zero value of ϕ_f would indicate CP violation in the interference between mixing and decay.

The observable that is used by time-dependent analysis to search for CP violation is

$$\Delta Y = \frac{\tau_{K\pi}}{\tau_{hh}} A_\tau, \tag{2.13}$$

where $\tau_{hh} = (\tau^{+}_{hh} + \tau^{-}_{hh})/2$ and $A_\tau = (\tau^{+}_{hh} - \tau^{-}_{hh})/(\tau^{+}_{hh} + \tau^{-}_{hh})$. In the Standard Model, ΔY is expected to be zero. Observing a non-zero value would indicate the presence of New Physics.

2.2.4 T-odd correlations

Being the main topic of this thesis, a separate section (the next one) has been left to describe the T-odd correlations.

2.3 T-ODD CORRELATIONS

Analyses that study T-odd correlations are trying to reverse the point of view. While other analysis search directly for CP violation, this kind of analysis are searching for T violation and assume the validity of CPT theorem to state if CP violation is found or not.

The validity of the method of T-odd correlations to search for T violation is inferred by many papers [Bensalem *et al.*, 2002a,b; Bensalem e London, 2001]. The original idea of applying this method to D decays is from I. Bigi [Bigi, 2001]: he suggested to study the four-body D decays building a T-odd observable using the spin or the momentum ($\vec{v}_i$) of the final state particles in the D center-of-mass frame. The observable is

$$A_T = \frac{\Gamma(\vec{v}_1 \cdot (\vec{v}_2 \times \vec{v}_3) > 0) - \Gamma(\vec{v}_1 \cdot (\vec{v}_2 \times \vec{v}_3) < 0)}{\Gamma(\vec{v}_1 \cdot (\vec{v}_2 \times \vec{v}_3) > 0) + \Gamma(\vec{v}_1 \cdot (\vec{v}_2 \times \vec{v}_3) < 0)} \tag{2.14}$$

where Γ represents the number of signal events and is measured on D decays only.

Anyway this is not a true T violating observable, because of the final state interaction effects that can introduce asymmetries. In order to remove these effects, it is needed to measure the charged conjugate of this observable ($\bar{A}_T$) using the $\bar{D}$ decays and evaluate the true T violating observable

$$\mathcal{A}_T \equiv \frac{1}{2}(A_T - \bar{A}_T). \tag{2.15}$$

While the T violating phase changes its sign when taking the CP conjugate, the strong phase (introduced by final state interaction) does not. The difference between A_T and $\bar{A}_T$ then removes the strong phase and the factor 1/2 is placed to normalize the weak phase.

2.3.1 Definition of the T-odd Observables in this work

In this work, the momenta of the final state particles are used to build the variables C_T and $\bar{C}_T$.

$$\begin{aligned} C_T &\equiv \vec{p}_{K^+} \cdot (\vec{p}_{\pi^+} \times \vec{p}_{\pi^-}), \\ \bar{C}_T &\equiv \vec{p}_{K^-} \cdot (\vec{p}_{\pi^-} \times \vec{p}_{\pi^+}). \end{aligned} \tag{2.16}$$

These variables are used to split the D and $\bar{D}$ decays and measure the asymmetry parameters A_T and $\bar{A}_T$:

$$\begin{aligned} A_T &= \frac{\Gamma_D(C_T > 0) - \Gamma_D(C_T < 0)}{\Gamma_D(C_T > 0) + \Gamma_D(C_T < 0)}, \\ \bar{A}_T &= \frac{\Gamma_{\bar{D}}(-\bar{C}_T > 0) - \Gamma_{\bar{D}}(-\bar{C}_T < 0)}{\Gamma_{\bar{D}}(-\bar{C}_T > 0) + \Gamma_{\bar{D}}(-\bar{C}_T < 0)}. \end{aligned} \tag{2.17}$$

Finally the T violating asymmetry parameter is measured as shown in Equation 2.15.

2.3.2 Previous Searches for CP Violation using T-odd Correlations

The first search for CP violation using T-odd correlations in D decays has been made by the FOCUS (E831) experiment at FermiLab [Link *et al.*, 2005].

Their analysis is made on about 800 $D^0 \to K^+ K^- \pi^+ \pi^-$ decays that have been separated among D^0 and $\bar{D}^0$ reconstructing the full $D^{*+} \to D^0 \pi^+$ decay chain. Using the momenta of the final state particles in the D^0 center-of-mass frame, the C_T and $\bar{C}_T$ variables have been retrieved for any D^0 and $\bar{D}^0$ decay. Measuring the number of D^0 and $\bar{D}^0$ when C_T and $\bar{C}_T$ are greater or less than zero, they found

$$\mathcal{A}_T(D^0) = 0.010 \pm 0.057(\text{stat}) \pm 0.037(\text{syst}). \tag{2.18}$$

The sources of systematics identified in their analysis are five:

1. the splitting of the data sample in four independent sub-samples (due to a possible mismatch in the reproduction of the D momentum and the changing experimental conditions during data collection);
2. the fit variant (computed by varying, in a reasonable manner, the fitting conditions on the whole data set);
3. the set of cuts applied to extract the signal (estimated using the standard deviation of the several sets of cuts used);
4. the D^*-tag dilution (actually an erroneous D^*-tag can dilute the measured asymmetry);
5. the limited Monte Carlo statistics.

The total systematic error has been obtained adding in quadrature each source.

The analysis on D^+ and D_s^+ decays has been made as the same as the D^0, the only difference in the events reconstruction. Both the D^+ and D_s^+ signal decays are reconstructed through their $K^+ K_S^0 \pi^+ \pi^-$ final state. This way, about 500 signal events have been reconstructed both for D^+ and D_s^+. The C_T and $\bar{C}_T$ variables have been evaluated as the same as for the D^0 decays and have been used to split the total data sample into four sub-samples. Another difference respect to the D^0 analysis is that both the D^+ and D_s^+ peaks are shown together into the same mass spectrum, so that the fit model has to take care about that. A remarkable contamination from $\Lambda_c^+ \to p\, K_S^0 \pi^+ \pi^-$ and $D^+ \to K_S^0 \pi^+ \pi^+ \pi^-$ has been found in the dataset and has been included into the fit model using a polynomial shape.

The sources of systematics error have been identified as the same as the D^0 sample. The final results are

$$\mathcal{A}_T(D^+) = 0.023 \pm 0.062(\text{stat}) \pm 0.022(\text{syst}); \tag{2.19}$$

$$\mathcal{A}_T(D_s^+) = -0.036 \pm 0.067(\text{stat}) \pm 0.023(\text{syst}). \tag{2.20}$$

The FOCUS experiment did not find any evidence of CP violation in D decays, but their work showed that the T-odd correlations are a clean and an alternative way to search for CP violation in four-body Cabibbo suppressed D decays. Their results then encouraged higher statistics experiment to repeat their measurement with better sensitivity.

2.4 CURRENT EXPERIMENTAL STATUS

In the last few years, the statistics collected by the B-factories and other experiments producing many charmed particles, reached the sensitivity required for the CP violation measurements in the Charm sector. The 0.1% sensitivity has not been reached yet, but 0.5% is already a standard for the latest measurements. Even if the Standard Model effects cannot be probed, the absence of asymmetry at the order of 1% limits the effects of New Physics on this observable.

The measurements made on D^0, D^+ and D_s^+ decays are shown in Tables 1, 2 and 3 respectively. In these tables the best measurements are reported only. To review the full list of measurements, please check the Heavy Flavor Averaging Group website: http://www.slac.stanford.edu/xorg/hfag.

Table 1: Previous searches for CP violation in D^0 decays.

Decay	Experiment (Year)	Measurement
	Direct CP measurements	
$D^0 \to \pi^+\pi^-$	Belle (2008)	$(+4.3 \pm 5.2 \pm 1.2) \times 10^{-3}$
	BABAR(2008)	$(-2.4 \pm 5.2 \pm 2.2) \times 10^{-3}$
$D^0 \to K^+ K^-$	Belle (2008)	$(-4.3 \pm 3.0 \pm 1.1) \times 10^{-3}$
	BABAR(2008)	$(0.0 \pm 3.4 \pm 1.3) \times 10^{-3}$
$D^0 \to \pi^+ \pi^- \pi^0$	*BABAR*(2008)	$(-3.1 \pm 4.1 \pm 1.7) \times 10^{-3}$
	Belle (2008)	$(+4.3 \pm 13.0) \times 10^{-3}$
$D^0 \to K^- \pi^+ \pi^0$	Cleo-c (2007)	$(+0.2 \pm 0.4 \pm 0.8) \times 10^{-2}$
$D^0 \to K^+ \pi^- \pi^0$	Belle (2005)	$(-0.6 \pm 5.3) \times 10^{-2}$
$D^0 \to K_S^0 \pi^+ \pi^-$	Cleo (2004)	$(-0.9 \pm 2.1^{+1.6}_{-5.7}) \times 10^{-2}$
$D^0 \to K^- K^+ \pi^0$	*BABAR*(2008)	$(+10.0 \pm 16.7 \pm 2.5) \times 10^{-3}$
$D^0 \to K^+ \pi^- \pi^+ \pi^-$	Belle (2005)	$(-1.8 \pm 4.4) \times 10^{-2}$
$D^0 \to K^+ K^- \pi^+ \pi^-$	FOCUS (2005)	$(-8.2 \pm 5.6 \pm 4.7) \times 10^{-2}$
	Time dependent CP measurements	
$D^0 \to \pi^+\pi^-$ $(+K^+ K^-)$	*BABAR*(2008)	$(+2.6 \pm 3.6 \pm 0.8) \times 10^{-3}$
	Belle (2007)	$(+1.0 \pm 3.0 \pm 1.5) \times 10^{-3}$
	T-odd correlations CP measurements	
$D^0 \to K^+ K^- \pi^+ \pi^-$	FOCUS (2005)	$(+1.0 \pm 5.7 \pm 3.7) \times 10^{-2}$

Table 2: Previous searches for CP violation in D^+ decays.

Decay	Experiment (Year)	Measurement
	Direct CP measurements	
$D^+ \to \mu\nu$	CLEO-c (2008)	$+0.08 \pm 0.08$
$D^+ \to K^0_S\,\pi^+$	*BABAR*(2010)	$(-4.4 \pm 1.3 \pm 1.0) \times 10^{-3}$
	Belle (2010)	$(-7.1 \pm 1.9 \pm 2.0) \times 10^{-3}$
	CLEO-c (2007)	$(-0.6 \pm 1.0 \pm 0.3) \times 10^{-2}$
$D^+ \to K^0_S\,K^+$	Belle (2010)	$(-1.6 \pm 5.8 \pm 2.5) \times 10^{-3}$
$D^+ \to K^-\,\pi^+\,\pi^+$	CLEO-c (2007)	$(-0.5 \pm 0.4 \pm 0.9) \times 10^{-2}$
$D^+ \to K^0_S\,\pi^+\,\pi^0$	CLEO-c (2007)	$(+0.3 \pm 0.9 \pm 0.3) \times 10^{-2}$
$D^+ \to K^+\,K^-\,\pi^+$	CLEO-c (2008)	$(-0.3 \pm 8.4 \pm 2.9) \times 10^{-3}$
	BABAR(2005)	$(+1.4 \pm 1.0 \pm 0.8) \times 10^{-2}$
	FOCUS (2000)	$(+0.6 \pm 1.1 \pm 0.5) \times 10^{-2}$
$D^0 \to K^-\,\pi^+\,\pi^+\,\pi^0$	CLEO-c (2007)	$(+1.0 \pm 0.9 \pm 0.9) \times 10^{-2}$
$D^0 \to K^0_S\,\pi^+\,\pi^+\,\pi^-$	CLEO-c (2007)	$(+0.1 \pm 1.1 \pm 0.6) \times 10^{-2}$
$D^+ \to K^0_S\,K^+\,\pi^+\,\pi^-$	FOCUS (2005)	$(-4.2 \pm 6.4 \pm 2.2) \times 10^{-2}$
	T-odd correlations CP measurements	
$D^+ \to K^0_S\,K^+\,\pi^+\,\pi^-$	FOCUS (2005)	$(+2.3 \pm 6.2 \pm 2.2) \times 10^{-2}$

Table 3: Previous searches for CP violation in D_s^+ decays.

Decay	Experiment (Year)	Measurement
	Direct CP measurements	
$D_s^+ \to \pi^+\,\eta$	CLEO-c (2008)	$(-8.2 \pm 5.2 \pm 0.8) \times 10^{-2}$
$D_s^+ \to \pi^+\,\eta'$	CLEO-c (2008)	$(-5.5 \pm 3.7 \pm 1.2) \times 10^{-2}$
$D_s^+ \to K^0_S\,\pi^+$	Belle (2010)	$(+5.5 \pm 2.5 \pm 0.3) \times 10^{-2}$
$D_s^+ \to K^+\,\pi^0$	CLEO-c (2007)	$(+0.02 \pm 0.29)$
$D_s^+ \to K^+\,\eta$	CLEO-c (2007)	(-0.20 ± 0.18)
$D_s^+ \to K^+\,\eta'$	CLEO-c (2007)	(-0.17 ± 0.37)
$D_s^+ \to K^0_S\,K^+$	Belle (2010)	$(+1.2 \pm 3.6 \pm 2.2) \times 10^{-3}$
	CLEO-c (2008)	$(+4.9 \pm 2.1 \pm 0.9) \times 10^{-2}$
$D_s^+ \to \pi^+\,\pi^+\,\pi^-$	CLEO-c (2008)	$(+2.0 \pm 4.6 \pm 0.7) \times 10^{-2}$
$D_s^+ \to K^+\,\pi^+\,\pi^-$	CLEO-c (2008)	$(+11.2 \pm 7.0 \pm 0.9) \times 10^{-2}$
$D_s^+ \to K^+\,K^-\,\pi^+$	CLEO-c (2008)	$(+0.3 \pm 1.1 \pm 0.8) \times 10^{-2}$
$D_s^+ \to K^0_S\,K^-\,\pi^+\,\pi^+$	CLEO-c (2008)	$(-0.7 \pm 3.6 \pm 1.1) \times 10^{-2}$
$D_s^+ \to K^+\,K^-\,\pi^+\,\pi^0$	CLEO-c (2008)	$(-5.9 \pm 4.2 \pm 1.2) \times 10^{-2}$
	T-odd correlations CP measurements	
$D_s^+ \to K^0_S\,K^+\,\pi^+\,\pi^-$	FOCUS (2005)	$(-3.6 \pm 6.7 \pm 2.3) \times 10^{-2}$

3 THE BABAR EXPERIMENT

CONTENTS

The B-factories have been playing a very important role in the understanding of the CP sector of the Standard Model. In this chapter the *BABAR* Experiment will be outlined in its main features.

3.1 PEP – II COLLIDER

The *BABAR* Detector has been designed to efficiently detect the decay products of the B mesons, which are produced in e^+e^- collisions at the center-of-mass energy of 10.58 GeV, corresponding to the mass of the $\Upsilon(4S)$ resonance, that decays half the times in $B^+ B^-$ and the other half into $B^0 \bar{B}^0$.

The two interacting beams are accumulated in the rings of the PEP-II collider, operated by the SLAC National Accelerator Laboratory. The PEP-II facility also includes the linear accelerator (LINAC) that allows to reach the energies of 9.0 GeV for the electron and 3.1 GeV for the positron beams.

The difference in energy between the two beams produces a boost of the center-of-mass $\beta\gamma = 0.56$. The vertices of the decaying B's have then a separation ($\beta\gamma c\tau \approx 270\,\mu m$) that is longer than the one they would have without any boost ($\beta\gamma = 0.06$, $\beta\gamma c\tau \approx 30\,\mu m$). A longer vertices separation, allows to measure with better definition the decay length of the two B mesons, hence reduce the error on their proper decay time. This effect is then needed to measure the proper time differences between the two decaying particles, a crucial ingredient of some of the golden measurements for which *BABAR* has been designed.

The PEP-II facility provided collisions to the *BABAR* experiment from October 1999 to April 2008. The design luminosity of $3 \times 10^{33}\,cm^{-2}s^{-1}$ has been reached on October 2000, while the *BABAR*'s record luminosity of $1.2 \times 10^{34}\,cm^{-2}s^{-1}$ has been achieved on August 16, 2006.

The collisions were mainly at the energy of the $\Upsilon(4S)$, but throughout the data taking many events were recorded "off-peak", about 40 MeV below the $\Upsilon(4S)$ (about 10% of the "on-peak" data). Since the center-of-mass energy of these events is under the $B\bar{B}$ production threshold (10,560 MeV/c^2), they can be used to study the background from the continuum production of u, d, s and c quarks. The cross-sections for

quarks production from e^+e^- collisions at the energy of the $\Upsilon(4S)$ are shown in Table 4.

Table 4: Hadronic and leptonic cross-sections for e^+e^- collisions at the center-of-mass energy of the $\Upsilon(4S)$.

$e^+e^- \to$	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
$\tau^+\tau^-$	0.94
$\mu^+\mu^-$	1.16
e^+e^-	~ 40

From January 2008, the collider explored the energy ranges of the $\Upsilon(3S)$ (10,335 MeV) and the $\Upsilon(2S)$ (10,023 MeV), collecting the world's largest sample of $\Upsilon(3S)$ decays. Furthermore a scan of the energy range above the $\Upsilon(4S)$ has been performed, allowing to measure the $e^+e^- \to b\bar{b}$ cross section between $\sqrt{s} = 10.54$ GeV and 11.20 GeV [Aubert *et al.*, 2009].

The integrated luminosity for the full data taking period is shown in Figure 1 and a summary of each of the separate runs is shown in Table 5. From the table can be noticed that a small quantity of events has been recorded out of the $\Upsilon(2S)$ and $\Upsilon(3S)$ peaks: the center-of-mass energy of these events is about 5 MeV below the $\Upsilon(3S)$ resonance and a wide range of non-peaking energies below the $\Upsilon(2S)$. These specific events were recorded while searching for the resonance peak.

Table 5: Recorded luminosity for each of the data taking periods.

Run	On Peak (fb^{-1})	Off Peak (fb^{-1})
1 ($\Upsilon(4S)$)	20.6	2.6
2 ($\Upsilon(4S)$)	60.6	6.9
3 ($\Upsilon(4S)$)	32.1	2.5
4 ($\Upsilon(4S)$)	100.7	10.2
5 ($\Upsilon(4S)$)	133.8	14.6
6 ($\Upsilon(4S)$)	76.5	7.3
7 ($\Upsilon(3S)$)	28.5	2.7
7 ($\Upsilon(2S)$)	14.4	1.5
Total	467.2	48.3

3.2 THE DETECTOR

The *BABAR* detector has been designed to satisfy the following needs:

- high and uniform acceptance even at small polar angles respect to the boost's direction;

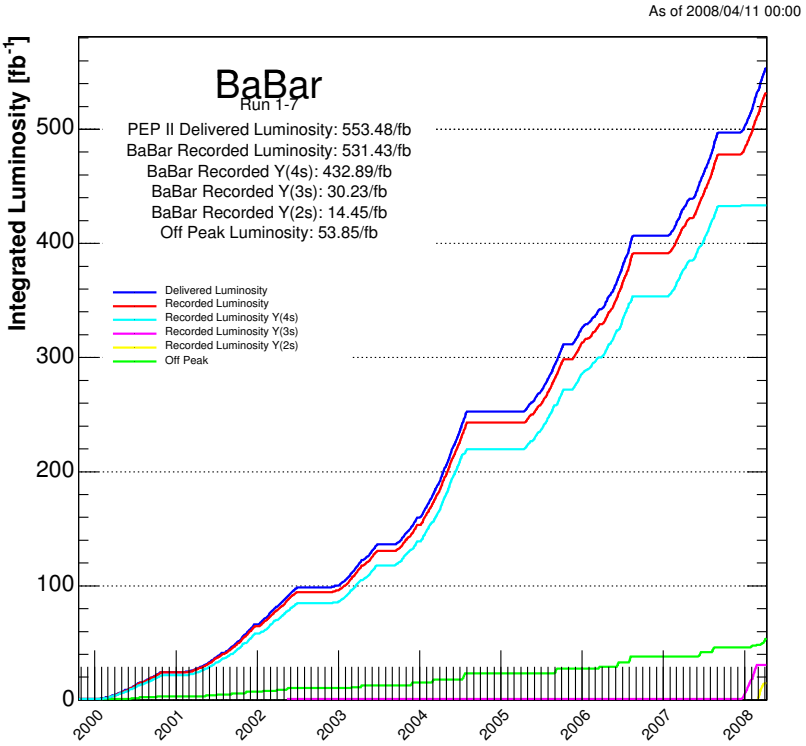


Figure 1: The luminosity delivered by the PEP-II facility (blue) and the luminosity recorded by the *BABAR* detector (red). The contributions to the total recorded luminosity from each resonance studied are shown with different colors.

- excellent reconstruction efficiency for charged particles in the momentum range of 60 MeV/c $< p_T <$ 4 GeV/c and for neutrals with energy in the range 20 MeV $< E <$ 5 GeV;
- very good vertex resolution, both in the beam's transverse and parallel direction;
- efficient discrimination between e, μ, π, K and p over a wide momentum range;
- reconstruction of neutral hadrons.

BABAR have answered to all these needs through a structure of dedicated detectors that can be outlined with the help of Figure 2.

A detailed description of *BABAR* components can be found in [Aubert *et al.*, 2002]. In the following there will be brief description of the major subsystems as they are traversed by a particle produced at the interaction point.

The innermost detector is the Silicon Vertex Tracker (SVT). It is one of the two *BABAR* tracking devices and is designed to perform high resolution measurements of the decay vertices. The resolution for

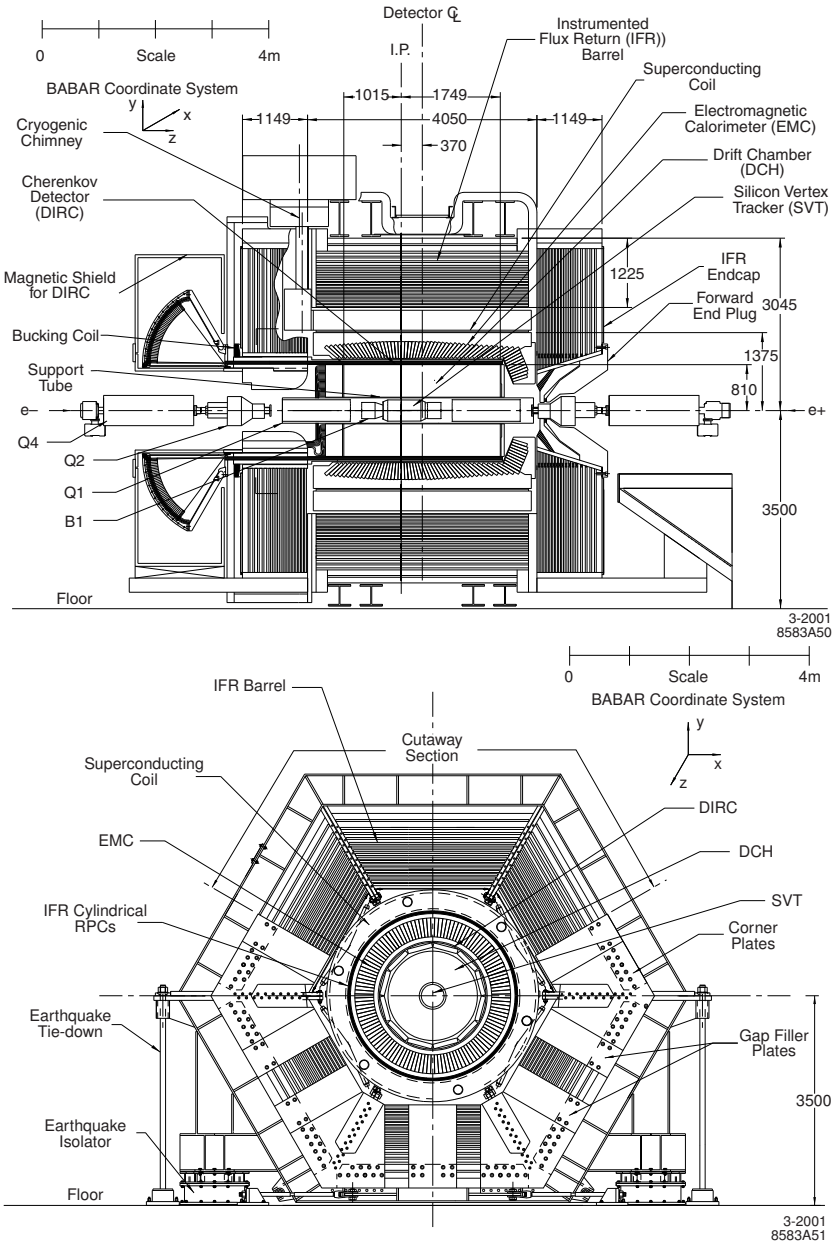


Figure 2: Longitudinal (top) and transverse (bottom) section of *BABAR* detector.

a fully reconstructed B vertex is about 80 μm and 180 μm for a partially reconstructed one. The SVT consists of five cylindrical layers of double-sided silicon micro strip detectors, with a total of approximately 150,000 readout channels. It is the only one device that detects charged particles with a transverse momentum less than 100 MeV/c. The inner radius is 32 mm, the outer radius is 144 mm and the total strip length is 26 cm.

Charged particles with a transverse momentum greater than 100 MeV/c can reach the next layer of *BABAR* detector, the Drift Chamber (DCH). Together with the SVT, the DCH is part of the *BABAR* tracking device. It is a 280 cm long cylinder, with inner radius of 24 cm and outer radius of 81 cm. The gas filling is a mixture of Helium-Isobutane (80:20) that is ionized when crossed by a charged particle. The electron-ion couples produced by the ionizing particle induce an electric signal on the 20 μm diameter anode wires. The anode wires are surrounded by 6 field wires providing the electric field that causes the drift of the electrons. The wires are arranged in 40 layers of drift cells . These design provides the chamber a resolution of about 1 mm, that is needed to reconstruct the K^0_S vertices which decay out of the SVT. Being θ the polar angle respect to the beam axis in the laboratory frame, the acceptance of the drift chamber is $17.2^\circ < \theta < 152.6^\circ$, corresponding to about 90% of the total.

After leaving the DCH, a particle reaches the long synthetic quartz bars of the *BABAR*'s ring imaging Cherenkov (DIRC, Detector of Internally Reflected Cherenkov light). The DIRC is designed to provide excellent particle identification, in particular to separate kaons from pions. The DIRC is most efficient for track momenta up to 4 GeV/c. Charged particles moving faster than speed of light within a radiator of refractive index n ($v > c/n$) emit a cone of Cherenkov light under a well defined angle θ_C so that

$$\cos\theta_C = \frac{1}{n\beta},$$

where $\beta = v/c$ and v is the velocity of the particle. If the particle's momentum is known, the mass of the particle can be inferred by using the velocity from the Cherenkov angle. The DIRC consists of 144 straight fused silica bars ($n = 1.473$) with a rectangular section. Each bar is 4.9 m long, 17 mm thick and 35 mm wide. The light emitted by the crossing particle is multiply reflected and travels inside the radiator bar until it reaches the rear end, where the Cherenkov image is allowed to expand into a stand-off box of stainless steel filled with 6,000 liters of purified water ($n_{H_2O} = 1.346$ provides good optical coupling). The rear end of the DIRC is closely packed with 10,572 photomultipliers tube that detect the Cherenkov light rings, from which pattern recognition algorithms deduce θ_C. Due to the presence of the phototubes, the rear end of the DIRC is out of the magnetic field, that could damage them. Furthermore, the rear end of the DIRC is placed in the backward region of the detector, where particle multiplicity is expected to be less than the forward because of the boost, optimizing the acceptance in the forward region. The geometrical acceptance of the DIRC is 80%.

The next detector component, that a traveling particle reaches, is the ElectroMagnetic Calorimeter (EMC). It is designed to detect photons and electrons within an energy range of $20\,\mathrm{MeV} < E_\gamma < 9\,\mathrm{GeV}$. It is the main device for electron-pion separation, the reconstruction of neutral pions and the detection of neutrons and photons. The calorimeter is made of 6,580 Thallium-doped Cesium iodide crystals. These crystals have a trapezoidal shape, their typical transverse dimension is $5 \times 5\,\mathrm{cm}^2$ at the front, flaring out to $6 \times 6\,\mathrm{cm}^2$ at the back. They have a typical length of 30 cm, corresponding to about $X_0 = 17$ radiation lengths. The calorimeter has a dedicated calibration system, which uses a neutron generator to activate liquid fluorinert. The fluorinert is pumped through tubes in front of the crystals, emitting 6.1 MeV photons from the ^{16}N $\beta^- - \gamma$ cascade. Another monitor system, the light pulser, is also featured: it distributes the light of a single Xenon lamp to each individual crystal and is able to quickly test the readout chain. Each crystal is being read by two photodiodes glued to the rear face. The design and the characteristics of the photodiodes allow to keep them in the magnetic field without being affected.

All of the above components are lying within a strong magnetic field of 1.5T, provided by a super-conducting solenoid. The flux return yoke is instrumented (Instrumented Flux Return - IFR) to identify muons and to detect K_L^0. The yoke consists of 18 layers of iron plates with increasing thickness, representing a total of 65 cm. The 17 gaps between the layers house two kinds of detector systems, Resistive Plate Chambers (RPC) and Limited Streamer Tubes (LST). The initial design of the IFR was entirely based on RPC layers. However some of them had to be replaced between 2004 and 2006 by the LST layers because of the large inefficiency to muon reconstruction shown by the RPCs [Convery *et al.*, 2006]. Only the the IFR end-cap in the forward region of the detector is still filled by RPCs [Anulli *et al.*, 2005].

The RPCs were installed in the gaps of the finely segmented steel of the IFR. They are made of two bakelite (phenolic polymer) sheets, 2 mm-thick and separated by a gap of 2 mm. The gap is kept uniform by polycarbonate spacers, that are glued to the bakelite and are spaced about 10 cm. The bulk resistivity of the bakelite sheets has been especially tuned to $10^{11} - 10^{12}\,\Omega\,\mathrm{cm}$. The external surfaces are coated with graphite to achieve a surface resistivity of about $100\mathrm{k}\Omega/\mathrm{square}$. These two graphite surfaces are connected to high voltage (≈ 8 kV) and ground, and protected by an insulating mylar film. This detectors are operated in limited streamer mode and the signals are read out capacitively, on both sides of the gap, by external electrodes made of aluminum strips on a mylar substrate.

The LSTs replaced the RPCs in the barrel. They are made of cells filled up of a gas mixture ($CO_2 : Ar : C_4H_10 = 89 : 3 : 8$) and containing a high voltage anode wire (5.5 kV). Whenever a charged particle crosses the detector, it ionizes the gas and the high voltage induces a streamer, read by the anode wire and by the silicon strips placed upon the tube. The LSTs work in the region of limited streamer [Iarocci, 1983], characterized by a short relaxing time for the detector and by streamers limited in space around the point of first ionization.

3.3 PARTICLE IDENTIFICATION

In High Energy Physics (HEP) experiments is crucial to correctly identify the particles that are produced in the reaction.

The informations from each one of the detectors that are part of *BABAR* are combined to evaluate a probability for each track to satisfy some identification hypothesis. Once a charged track has been reconstructed from the hits recorded by SVT and DCH, the dE/dx recorded by the two detectors is combined with the angle of the light cone observed by the DIRC, with the energy recorded by the calorimeter and eventually the hits in the IFR.

The combination of all these informations is used by many particle identification algorithms that are based on different statistical techniques such as Likelihood, Global Likelihood, Boosted Decision Trees and Neural Networks. For each of these algorithms there are some criteria to identify the particle with tighter or looser conditions. The analyst is finally intended to choose the algorithm and the criteria that best suit its analysis purpose. In the following, the identification of pions and kaons will be discussed in details, with a focus on the specific selection techniques applied to the dataset.

3.3.1 Pion and Kaon Selection

Both the pion and kaon candidates are retrieved from lists of particles produced at the Event Processing stage. For each track, the selector calculates likelihoods $\mathcal{L}$ for several particle hypotheses: pion, kaon, electron, proton and muon.

The information for each one of the tracking detectors are considered in the calculation: the dE/dx and the number of hits in SVT and DCH; the Cherenkov angle and the number of photons measured by the DIRC; the ratio (E/p) between the energy measured by the EMC and the momentum of the track (measured combining the information of SVT and DCH).

Once the likelihood for each particle hypothesis has been calculated, the selector combines the results into ratios. The selector then applies some previously optimized cuts to the ratios in order to accept or reject the particle hypothesis.

In Figure 3 the reconstruction efficiency in bin of momentum is compared between data and Monte Carlo for the specific pion selector used in the D^0 analysis. In the same figure, the reconstruction efficiency for kaons, i.e. the fake-rate, and the ratio between data and Monte Carlo efficiency are also shown. In Figure 4 the same plots are shown for the kaon selector.

The analysis of the charged $D^+_{(s)}$ decays has been performed about one year later respect to the one on D^0 decays. In *BABAR*, the selectors for particle identification have had an improvement throughout all the life of the collaboration. That is why different selectors have been used for the $D^+_{(s)}$ analysis. The performance plot for these more recent set of selectors are shown in Figures 5 and 6 as the same as the others. The main difference among the selectors used for D^0 and $D^+_{(s)}$ analysis is in the statistics technique used to test the hypotheses. While

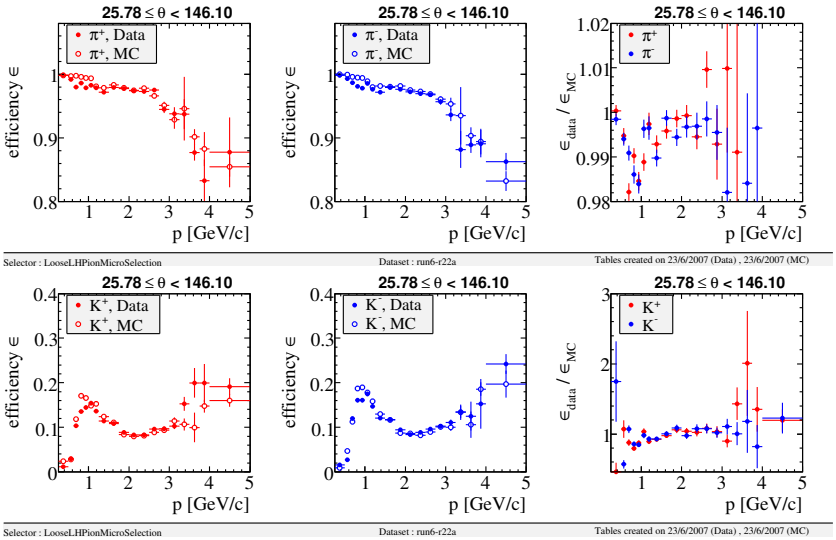


Figure 3: Reconstruction efficiency for pions (top) and kaons fake-rate (bottom) using likelihood pion selector. The reconstruction efficiency for the positively charged particles is shown on the left and for the negatively charge in the middle. The ratio between data and Monte Carlo reconstruction efficiency is shown on the right for both positively and negatively charged particles.

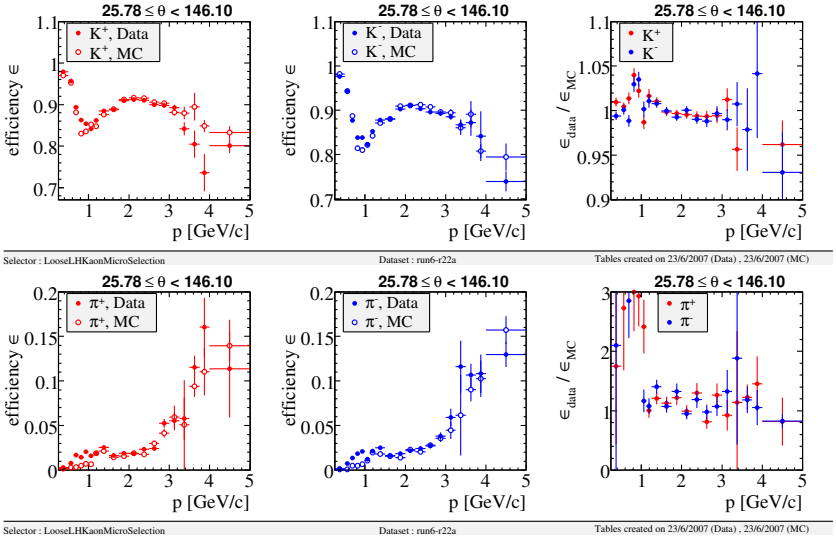


Figure 4: Reconstruction efficiency for kaons (top) and pions fake-rate (bottom) using likelihood kaon selector. The reconstruction efficiency for the positively charged particles is shown on the left and for the negatively charge in the middle. The ratio between data and Monte Carlo reconstruction efficiency is shown on the right for both positively and negatively charged particles.

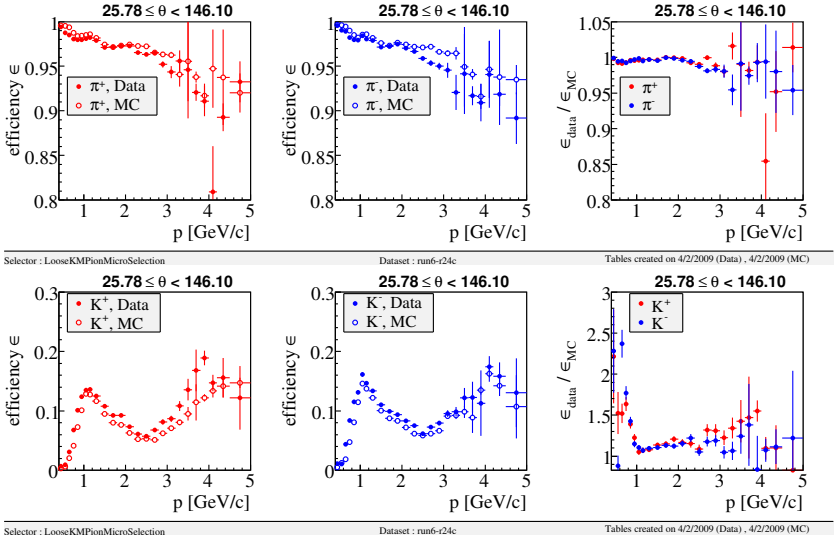


Figure 5: Reconstruction efficiency for pions (top) and kaons fake-rate (bottom) using a pion selector based on likelihood and neural network techniques. The reconstruction efficiency for the positively charged particles is shown on the left and for the negatively charge in the middle. The ratio between data and Monte Carlo reconstruction efficiency is shown on the right for both positively and negatively charged particles.

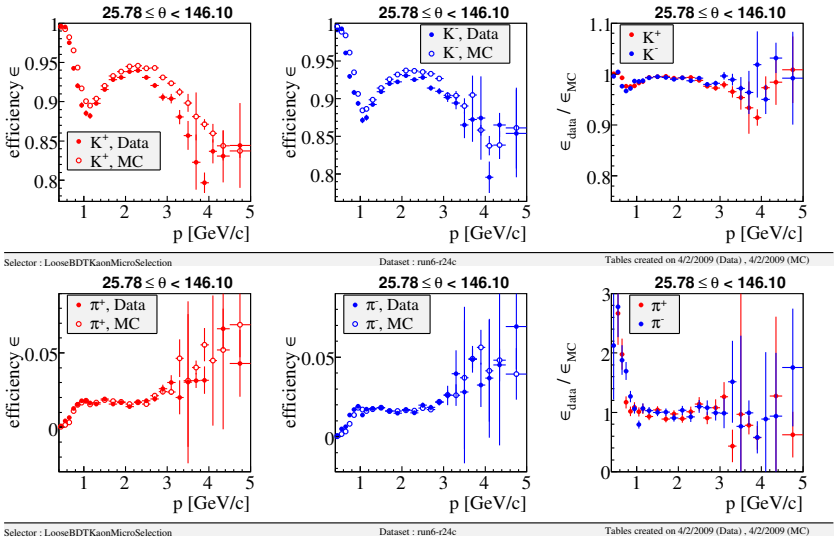


Figure 6: Reconstruction efficiency for kaons (top) and pions fake-rate (bottom) using boosted decision tree kaon selector. The reconstruction efficiency for the positively charged particles is shown on the left and for the negatively charge in the middle. The ratio between data and Monte Carlo reconstruction efficiency is shown on the right for both positively and negatively charged particles.

the selectors used for D^0 analyses simply used likelihood ratios to determine wheter a particle hypothesis could be accepted or not, these new selectors use more advanced techniques, such as Error Correcting Output Code and Boosted Decision Trees, to determine the validity of the hypothesis. For further information upon the different statistical techniques used by *BABAR* selectors, please check Appendix B.

These plots give evidence for the high efficiency and low fake-rate in *BABAR* particle identification. Furthermore, the comparison between data and Monte Carlo shows that the simulation has an excellent agreement to the data. This result is very important for these analyses, since all the reconstruction techniques and the models used to describe the data sample have been tested and developed using the Monte Carlo.

4 $D^0 \to K^+K^-\pi^+\pi^-$ ANALYSIS

CONTENTS

The underlying idea is to measure the asymmetry parameter $\mathcal{A}_T$ directly from a fit to the data sample. In order to perform the fit to the two-dimensional distribution of the events in the $[m(K^+K^-\pi^+\pi^-), \Delta m]$ plane, the fit model has to be carefully defined. In the following, the data set and the selection of the events are briefly outlined; then the procedure to define this model is described in detail; finally the evaluation of the systematic error and the final result are reported.

Δm is the difference in mass between the D^{*+} and the D^0 candidate

4.1 THE DATA SET

The analysis is made using the full *BABAR* data set recorded at the $\Upsilon(4S)$ and 40 MeV below the resonance. This corresponds to an integrated luminosity of about 470 fb^{-1}, of which 44 fb^{-1} are recorded off-resonance.

Together with the *BABAR* recorded data set, a large number of Monte Carlo events has been analyzed using the same analysis chain as for the data. These events are generated using the software `Jetset7.4` [Sjostrand, 1994] and their interactions with the detector have been simulated using the `GEANT 4` [Agostinelli *et al.*, 2003] program.

The Monte Carlo events that have been studied in this work are generated from the simulation of $e^+e^- \to c\bar{c}$ interactions. There are two kind of Monte Carlo samples that have been studied:

- simulation of $e^+e^- \to c\bar{c}$ continuum production; the hadronization and decay products are generated according to the cross-section from the Review of Particle Physics [Nakamura *et al.*, 2010] (Generic Monte Carlo);
- events generated requiring that the decay chain that is signal for this analysis is always present (Signal Monte Carlo).

The analysis makes use of about 1.1×10^9 Generic Monte Carlo Events and 5.2×10^6 Signal Monte Carlo Events.

4.2 EVENTS RECONSTRUCTION

Charge-conjugation is implied throughout the document

The $D^0 \to K^+ K^- \pi^+ \pi^-$ candidates are reconstructed together with their D^{*+} parent:

$$D^{*+} \to D^0\pi^+ \\ \hookrightarrow K^+K^-\pi^+\pi^-$$

from all the inclusive $e^+e^- \to D^{*+} X$ decays, where X indicates any system composed of charged and neutral particles. The reconstruction of the full D^{*+} decay chain has two major advantages:

1. the $D^{*+} \to D^0 \pi^+$ decay has about 68% probability and is the unique D^{*+} decay involving a D^0; this allows to obtain a clean D^0 sample through a simple cut on the mass difference $\Delta m = m(D^{*+}) - m(D^0)$;
2. the charge of the slow momentum pion decaying from the D^{*+} uniquely tag the D flavor as a D^0 (π_s^+) or a $\bar{D}^0$ (π_s^-).

The final state consists then of four charged particles with a high efficiency for identification into *BABAR* detector and a slow pion.

The procedure to build the D^{*+} candidate is explained in the following. First a D^0 candidate is built: all well identified K^+, K^-, π^+ and π^- tracks combinations have been considered. If the invariant mass of the four particles is in the range

$$1.81\,\mathrm{GeV}/c^2 < m(K^+K^-\pi^+\pi^-) < 1.92\,\mathrm{GeV}/c^2$$

the four tracks are fitted to a common vertex requiring the χ^2 fit probability to be greater than 0.1%. The momentum of the D^0 candidate in the e^+e^- rest frame is required to be $p^*(D^0) > 2.5\,\mathrm{GeV}/c$.

To search for $D^{*+} \to D^0 \pi_s^+$ candidates, the reconstructed D^0 candidate is combined with all the remaining charged tracks in the event having a momentum in the laboratory frame

$$p_{lab}(\pi_s^+) < 0.65\,\mathrm{GeV}/c.$$

A new fit is therefore performed with the constraint that the new vertex is located in the interaction region and requiring the χ^2 fit probability to be greater than 0.1%.

A couple of comments upon the selections made on the momenta of D^0 and π_s^+:

- Requiring the D^0 momentum in the center-of-mass frame to be greater than 2.5 GeV/c removes any D^0 coming from B decays, together with many combinatorial events, then reducing the background.
- The π_s^+ coming from the D^{*+} decay should be very soft, requiring $p_{lab}(\pi_s^+) < 0.65\,GeV/c$ reduces the background and saves CPU time.

4.3 CONTAMINATIONS IN THE D^0 DATA SAMPLE

Even the more stringent cut cannot completely rule out some contamination to the data sample. These contamination are due to particles that are not correctly identified or from decays with a topology similar to the ones of interest. While combinatoric background is usually uniformly distributed over the phase space, not influencing dramatically the quantity one measures, the contaminations could, in principle, fake a measurement, since they introduce events in the signal data sample that could have different properties respect to the decays that are argument of study. The analyst should then be always aware of all the contaminations that could be into the data sample he is studying and, if they are not avoidable, he should find the best way to deal with them.

At first, the D^{*+} ambiguity has been studied: the selection on Δm does not preserve from taking into account events for which two or more slow pions form a D^{*+} candidate within the signal window. Since the slow pion identification is strictly related to D^* tagging, which allows to separate among D^0 and $\bar{D}^0$ events, it is crucial to verify that there is no ambiguity. The number of slow pions in the signal window has been studied using a sub-sample of the data (about 13%). In this sample the events in the signal window are 15,850 and 21 of them have two or more slow pion candidates. Among them, 12 have slow pions of opposite charge and 9 the same. All these events have been removed, with a relative loss of 1.3×10^{-3} for the full data sample.

An important contamination comes from the decay $D^0 \to K^+ K^- K_S^0$ ($K_S^0 \to \pi^+\pi^-$). Its presence is highlighted by the peak at $0.5\,GeV/c^2$ in the $\pi^+\pi^-$ invariant mass, which is shown in Figure 7 together with a fit and the distribution of the normalized residuals represented by

$$Pull = \frac{(N_{data} - N_{fit})}{\sqrt{N_{data}}}, \tag{4.1}$$

where N_{data} is the number of data events and N_{fit} the number of events expected by the fit for each bin of the mass spectrum.

This contamination is from a Cabibbo favored decay channel ($BR = [4.65 \pm 0.30] \times 10^{-3}$ [Nakamura *et al.*, 2010]) and sits in the signal region. The K_S^0 peak can be represented by a Gaussian distribution with

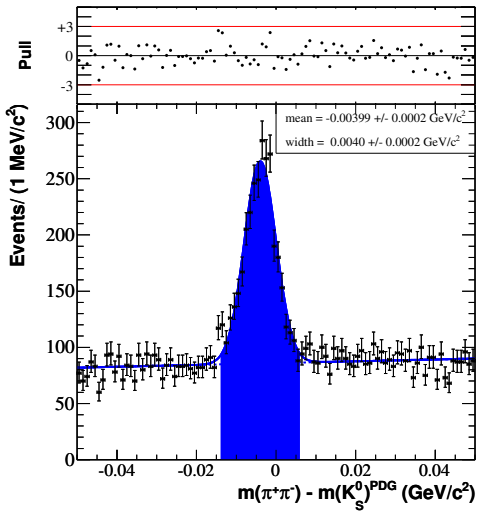


Figure 7: The spectrum of the difference between $\pi^+\pi^-$ invariant mass and K^0_S nominal mass. The fit to a model made of a Gaussian over a line is represented by the blue line. The events removed to get rid of the contamination are shown in the blue filled area. On top the pull distribution of the residuals is shown.

$\sigma = 4.0 \pm 0.2\,\mathrm{MeV/c^2}$, and which accounts for 5.2% of the selected data sample. The fit shows that the $\pi^+\pi^-$ mass distributions peak is $4\,\mathrm{MeV/c^2}$ below the expected value for the K^0_S mass [Nakamura *et al.*, 2010], this is due to the constrain on the two pions to vertex together with the other D^0 decay products. We veto K^0_S candidates within a window of 2.5σ, removing all the events in the blue filled area of Figure 7. This cut, while reducing to negligible level the background from $D^0 \to K^+\,K^-\,K^0_S$, removes 5.8% of the signal events.

The backgrounds from charm decay modes with misidentified pions have been searched by assigning alternatively the pion mass to both kaons. The resulting spectra show no peak both in the four particles and three particles invariant mass. No peak in the four, three and two-body invariant masses which could be associated to charm decays has been found.

A search for electrons which could be misidentified as kaons or pions has been made alternatively assigning the electron mass to pairs of tracks with opposite charge. The resulting e^+e^- invariant mass shows no sign of γ conversions in the SVT, excluding any contamination from radiative modes such as $\rho\gamma$, $\phi\gamma$ or $\bar{K}^*(892)^0\gamma$.

4.4 C_T DISTRIBUTION

The data set after all the selection shows a clear signal from $D^{*+} \to D^0\,\pi^+$, $D^0 \to K^+\,K^-\,\pi^+\,\pi^-$, as it shown in Figure 8. In this plot, the scatter plot of Δm versus $m(K^+\,K^-\,\pi^+\,\pi^-)$ is shown, together with the $m(K^+\,K^-\,\pi^+\,\pi^-)$ spectrum in the Δm signal region and vice versa. The signal regions have been defined from two previous fit that have

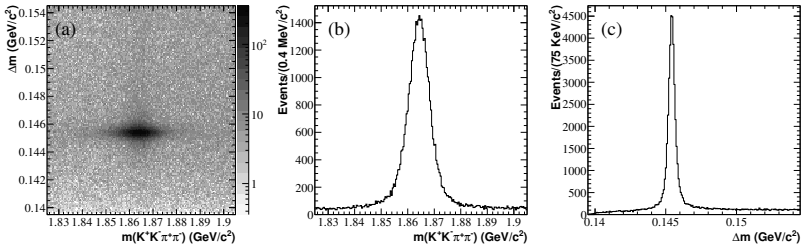


Figure 8: The scatter plot of Δm versus $m(K^+\ K^-\ \pi^+\ \pi^-)$ (a); the $m(K^+\ K^-\ \pi^+\ \pi^-)$ spectrum in Δm signal region (b); the Δm spectrum in $m(K^+\ K^-\ \pi^+\ \pi^-)$ signal region.

been made separately to the $m(K^+\ K^-\ \pi^+\ \pi^-)$ and Δm spectra using a single gaussian over a second order polynomial model. The signal box retrieved from the fits is:

$$144.94\,\mathrm{MeV/c^2} < \Delta m < 145.91\,\mathrm{MeV/c^2},$$
$$1857.11\,\mathrm{MeV/c^2} < m(K^+K^-\pi^+\pi^-) < 1871.47\,\mathrm{MeV/c^2}. \qquad (4.2)$$

Using the signal box defined in Equation 4.2, the C_T and $\bar{C}_T$ distributions have been plot in the D^0 signal box, with the results shown in Figure 9. In the same plots, the distribution of the events when

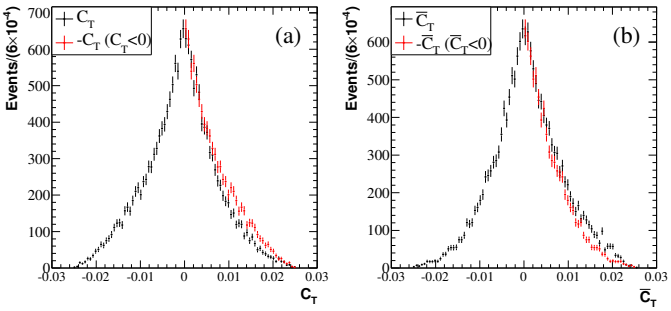


Figure 9: C_T (a) and $\bar{C}_T$ (b) distributions in the D^0 signal box of the data sample. The total distribution is shown in black, while the distribution of $-C_T$ ($-\bar{C}_T$) is overdrawn in red to graphically compare the distribution of the events when $C_T(\bar{C}_T) > 0$ and $C_T(\bar{C}_T) < 0$.

$C_T(\bar{C}_T) > 0$ and $C_T(\bar{C}_T) < 0$ is compared. This has been done drawing the $-C_T$ ($-\bar{C}_T$) distribution over the C_T ($\bar{C}_T$) distribution when C_T ($\bar{C}_T$) is less then zero. From this plot, the asymmetry of the signal events when C_T ($\bar{C}_T$) is greater or less than zero is clear. It is then expected that the A_T and $\bar{A}_T$ asymmetries are different from zero, indicating the effect of final state interaction.

4.5 DEFINITION OF THE FIT MODEL

The fit model has been developed making use of the $e^+e^- \to c\bar{c}$ generic Monte Carlo events. These events have been reconstructed

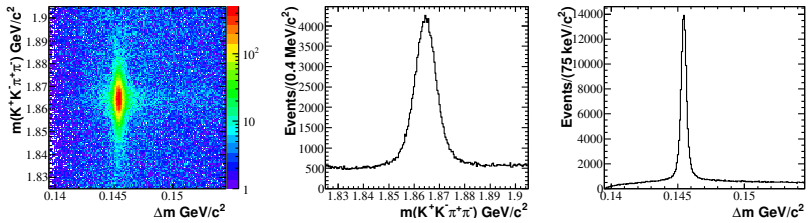


Figure 10: Distribution of Monte Carlo events. On the left the two-dimensional distribution of $m(K^+\ K^-\ \pi^+\ \pi^-)$ and Δm is shown, in the center there is the distribution of $m(K^+\ K^-\ \pi^+\ \pi^-)$ and on the right of Δm.

and analyzed using the same chain as for the real data sample, adding for every event the information on the original decay tree from the generator. This information allows the analyst to determine, after the reconstruction, if the event has been correctly reconstructed as a signal event or not. Furthermore, knowing the generated event, the different sources of background can be divided and studied separately.

The study of the reconstructed events using the generator-level information allowed to define five categories of events into the data sample:

1. True D^0 signal originating from a D^{*+} decay. This component has characteristic peaks in both observables $m(K^+\ K^-\ \pi^+\ \pi^-)$ and Δm.

2. Random π_s^+ events where a true D^0 is associated to an incorrect π_s^+, called D^0 peaking. This contribution has the same shape in $m(K^+\ K^-\ \pi^+\ \pi^-)$ as signal events, but does not peak in Δm.

3. Mis-reconstructed D^0 decays where one or more of the D^0 decay products are either not reconstructed or reconstructed with the wrong particle hypothesis, called Δm peaking. Some of these events show a peak in Δm, but not in $m(K^+\ K^-\ \pi^+\ \pi^-)$.

4. Combinatorial background where the K^+, K^-, π^+, π^- candidates are not fragments of the same D^0 decay, called combinatoric. This contribution does not exhibit any peaking structure in $m(K^+\ K^-\ \pi^+\ \pi^-)$ or Δm.

5. $D_s^+ \to K^+\ K^-\ \pi^+\ \pi^-\ \pi^+$ contamination (BR = $[8.8 \pm 1.6] \times 10^{-3}$), called D_s^+. This background has been studied on Monte Carlo simulations and shows a characteristic linear narrow shape in the two-dimensional $[m(K^+\ K^-\ \pi^+\ \pi^-), \Delta m\ ]$ distribution.

Each one of these categories has been separately studied to determine the model to include into the more general one. The total number of Monte Carlo events that have been reconstructed from the generic $e^+e^- \to c\bar{c}$ Monte Carlo sample described in Section 4.1 is about 2.11×10^5. Their two-dimensional distribution into the plane of $m(K^+\ K^-\ \pi^+\ \pi^-)$ versus Δm and the projections upon these two variables are shown in Figure 10. The number of events for each category identified by the previous description is shown in Table 6.

Table 6: Number of events for each category from Monte Carlo. The contribution to the total of the sample is also shown.

Category	Events	Fraction (%)
Signal	103,561	48.97
D^0 peaking	5,706	2.70
Δm peaking	73,138	34.59
Combinatorial	27,306	12.91
D_s^+	1,753	0.83
Total	211,464	100.00

The plots of the two-dimensional distribution of $m(K^+\ K^-\ \pi^+\ \pi^-)$ and Δm and their projections are shown for each category of events in Figure 11.

In the following, the Probability Density Functions (PDF) defined for each category of events will be described in detail. At first, signal and combinatorial will be studied, then the other categories. This choice has been made because the latter categories show some features that are present in the former ones. All the fits that are shown in this thesis are mad using the code `MINUIT` [James e Roos, 1975] in a `RooFit` [Verkerke e Kirkby, 2003] framework.

4.5.1 Signal Category

The Signal category of events shows a distinct peak both in the $m(K^+\ K^-\ \pi^+\ \pi^-)$ and Δm distribution. However these peaks cannot be simply described by a sum of independent gaussians. While this can be considered a good approximation for the $m(K^+\ K^-\ \pi^+\ \pi^-)$ distribution, the Δm distribution shows a long tail on the higher mass side of the spectrum. Furthermore, a correlation is found between the root mean square (RMS) of Δm and $m(K^+\ K^-\ \pi^+\ \pi^-)$, as it is shown in Figure 12. Here the mean and the RMS of the distribution of Δm is shown in slices of $m(K^+\ K^-\ \pi^+\ \pi^-)$, demonstrating that the more likely the event to be signal, the shorter the variance of Δm distribution. All the features of the Δm distribution have been considered in the fit model.

The probability density functions used in this parametrization are:

- gaussians

$$g(x;\bar{x},\sigma) = \exp\left(-\frac{(x-\bar{x})^2}{2\sigma^2}\right); \tag{4.3}$$

- a two-dimensional gaussian including a correlation factor c

$$s(x,y;\bar{x},\sigma_x,\bar{y},\sigma_y,c) = \exp\left[-\frac{1}{2(1-c^2)}\left(\frac{(x-\bar{x})^2}{\sigma_x^2}\frac{(y-\bar{y})^2}{\sigma_y^2} - 2c\frac{(x-\bar{x})}{\sigma_x}\frac{(y-\bar{y})}{\sigma_y}\right)\right]; \tag{4.4}$$

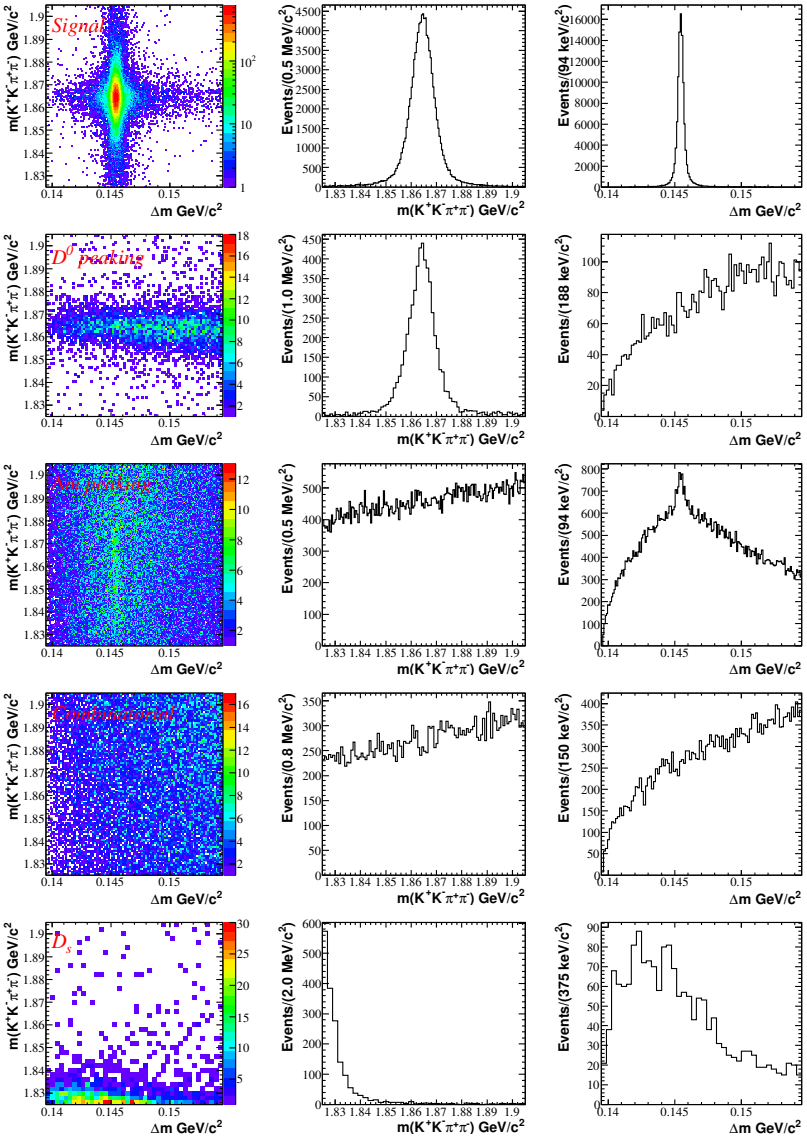


Figure 11: The $m(K^+\ K^-\ \pi^+\ \pi^-)$ versus Δm (left), $m(K^+\ K^-\ \pi^+\ \pi^-)$ (center) and Δm (right) distributions for each category are shown for each recognized category of events, from Signal (Category 1- top) to D_s^+ (Category 5 - bottom). Generic Monte Carlo events.

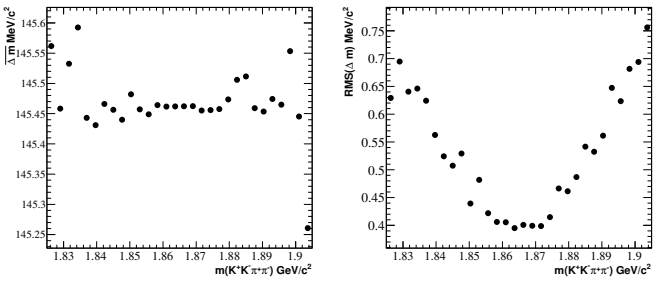


Figure 12: Mean value (left) and RMS (right) of Δm in slices of m(K^+ K^- π^+ π^-).

- a Johnson SU function [Johnson, 1949]

$$J_{SU}(x;\bar{x},\sigma,\gamma,\delta) = \frac{\delta e^{-\frac{1}{2}\left[\gamma+\delta \sinh^{-1}\left(\frac{x-\bar{x}}{\sigma}\right)\right]^2}}{\sigma\sqrt{2\pi}\sqrt{1+\left(\frac{x-\bar{x}}{\sigma}\right)^2}}, \tag{4.5}$$

where $\sinh^{-1}(z) = \ln\left(z+\sqrt{z^2+1}\right)$.

With the help of Figure 13 the combination of these functions into a single PDF can be understood.

The sum of the two distributions made by the product of two independent gaussians is shown in white. This is the main component of the two-dimensional peak:

$$\mathcal{P}_3 = g(m;m_3,\sigma_{m_3}) \times g(\Delta m;\Delta m_3,\sigma_{\Delta m_3}), \tag{4.6}$$
$$\mathcal{P}_4 = g(m;m_4,\sigma_{m_4}) \times g(\Delta m;\Delta m_4,\sigma_{\Delta m_4}). \tag{4.7}$$

The green filled area represent the projection of the distribution which takes care of the long tail in Δm for higher masses. This distribution is made of the product of a gaussian in m(K^+ K^- π^+ π^-) and a Johnson SU in Δm:

$$\mathcal{P}_{JSU} = g(m;m_{JSU},\sigma_{JSU}) \times J_{SU}(\Delta m;\Delta m_{JSU},\sigma_{JSU},\gamma,\delta). \tag{4.8}$$

Finally the red filled area is the projection of the two-dimensional gaussian incorporating a correlation term:

$$\mathcal{P}_{wide} = s(m,\Delta m;m_{CG},\sigma_{CG},\Delta m_{CG},\sigma_{\Delta m_{CG}},c_{CG}). \tag{4.9}$$

The sum of all these contributions forms the distribution used to parametrize the signal events:

$$\mathcal{F}_{sig} = f_{wide}\mathcal{P}_{wide} + (1-f_{wide}) \{f_{JSU}\mathcal{P}_{JSU} + (1-f_{JSU})[f_{core}\mathcal{P}_3 + (1-f_{core})\mathcal{P}_4]\} \tag{4.10}$$

The results of the fit to the distribution of events correctly reconstructed as signal is shown in Figure 13.

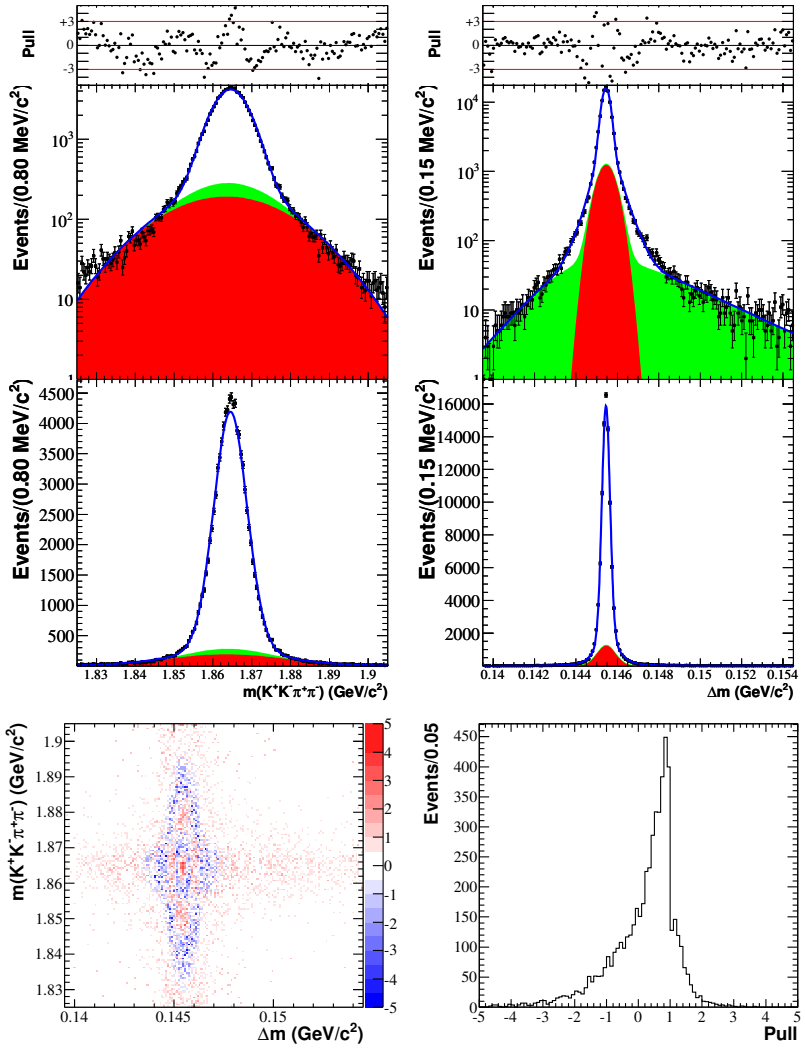


Figure 13: Fit to the distribution of correctly reconstructed signal events from Monte Carlo. On top the projections on m(K^+ K^- π^+ π^-) and Δm of the fit model (blue line) and its components (color filled areas) are shown using a logarithmic scale on the y-axis. In the middle the same plots are shown in linear scale. Over these plots the distribution of the residuals (Pull - see Eq. 4.1) is shown. On the bottom there are the distributions of the residuals in the two-dimensional [m(K^+ K^- π^+ π^-),Δm] plane (left) and integrated (right).

4.5.2 Combinatorial Background

The distribution of the combinatorial background is flat in the $m(K^+ K^- \pi^+ \pi^-)$ variable and shows a distinctive threshold distribution in Δm. Since it is combinatorial, no correlation is expected among $m(K^+ K^- \pi^+ \pi^-)$ and Δm. The general PDF should be then built assuming the two variables to be independent. A flat distribution can be used to describe the $m(K^+ K^- \pi^+ \pi^-)$ spectrum

$$l(x; b) = 1 + b(x - 1.865\,\mathrm{GeV}/c^2). \tag{4.11}$$

The description of the Δm spectrum is made by means of a threshold function with a kinematic endpoint ($x_0 = m_{\pi^+} = 139.57\,\mathrm{MeV}/c^2$):

$$a(x; c) = \sqrt{\left(\frac{x}{x_0}\right)^2 - 1} \cdot \exp\left(-c\left[\left(\frac{x}{x_0}\right)^2 - 1\right]\right). \tag{4.12}$$

The PDF defined to describe the combinatorial background is:

$$\mathcal{F}_{comb}(m, \Delta m; b, c) = l(m; b) \times a(\Delta m; c). \tag{4.13}$$

The results of the fit of this PDF to the combinatorial background events found in the Monte Carlo are shown in Figure 14.

4.5.3 D^0 Peaking Background

The events for which the D^0 has been correctly reconstructed, but the slow pion is randomly assigned, show a peak into the $m(K^+ K^- \pi^+ \pi^-)$ distribution and a threshold distribution identical to the one of combinatorial background for Δm. Given the absence of correlations between the D^0 and the random pion, the two-dimensional PDF can be built as the product of two uncorrelated distributions.

The distribution used to describe the $m(K^+ K^- \pi^+ \pi^-)$ mass spectrum is taken from the integral on Δm of the two-dimensional PDF that describes the signal (see Eq. 4.10). The distribution to describe Δm spectrum is taken from the combinatorial PDF and is the same as Eq. 4.12. The two-dimensional PDF for the description of D^0 peaking category is then:

$$\begin{aligned}\mathcal{F}_{D^0-peak}(m, \Delta m) = \\ (f_{wide} \cdot \mathcal{P}^{D^0}_{wide} + (1 - f_{wide}) \cdot \{f_{JSU} \cdot \mathcal{P}^{D^0}_{JSU} + \\ (1 - f_{JSU}) \cdot [f_{core} \cdot \mathcal{P}^{D^0}_3 + (1 - f_{core}) \cdot \mathcal{P}^{D^0}_4]\}) \times a(\Delta m; c),\end{aligned} \tag{4.14}$$

where

$$\mathcal{P}^{D^0}_{wide} = g(m; \overline{m}_{CG}, \sigma_{m_{CG}}) \tag{4.15}$$

$$\mathcal{P}^{D^0}_{JSU} = g(m; \overline{m}_{JSU}, \sigma_{m_{JSU}}) \tag{4.16}$$

$$\mathcal{P}^{D^0}_3 = g(m; \overline{m}_3, \sigma_{m_3}) \tag{4.17}$$

$$\mathcal{P}^{D^0}_4 = g(m; \overline{m}_4, \sigma_{m_4}). \tag{4.18}$$

The results of the fit to the D^0 peaking category of Monte Carlo events are shown in Figure 15.

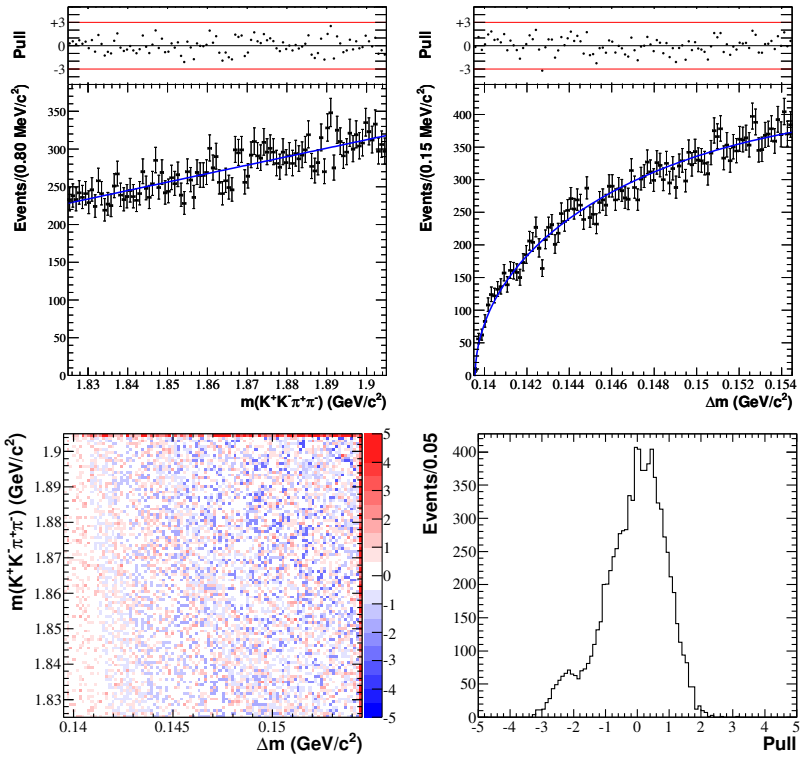


Figure 14: Fit to the distribution of combinatorial background events in the Monte Carlo. On top the projections on $m(K^+\ K^-\ \pi^+\ \pi^-)$ and Δm of the fit model (blue line) are shown. Over each plot the distribution of the residuals (Pull - see Eq. 4.1) is shown. On the bottom there are the distributions of the residuals in the two-dimensional $[m(K^+\ K^-\ \pi^+\ \pi^-),\Delta m\ ]$ plane (left) and integrated (right).

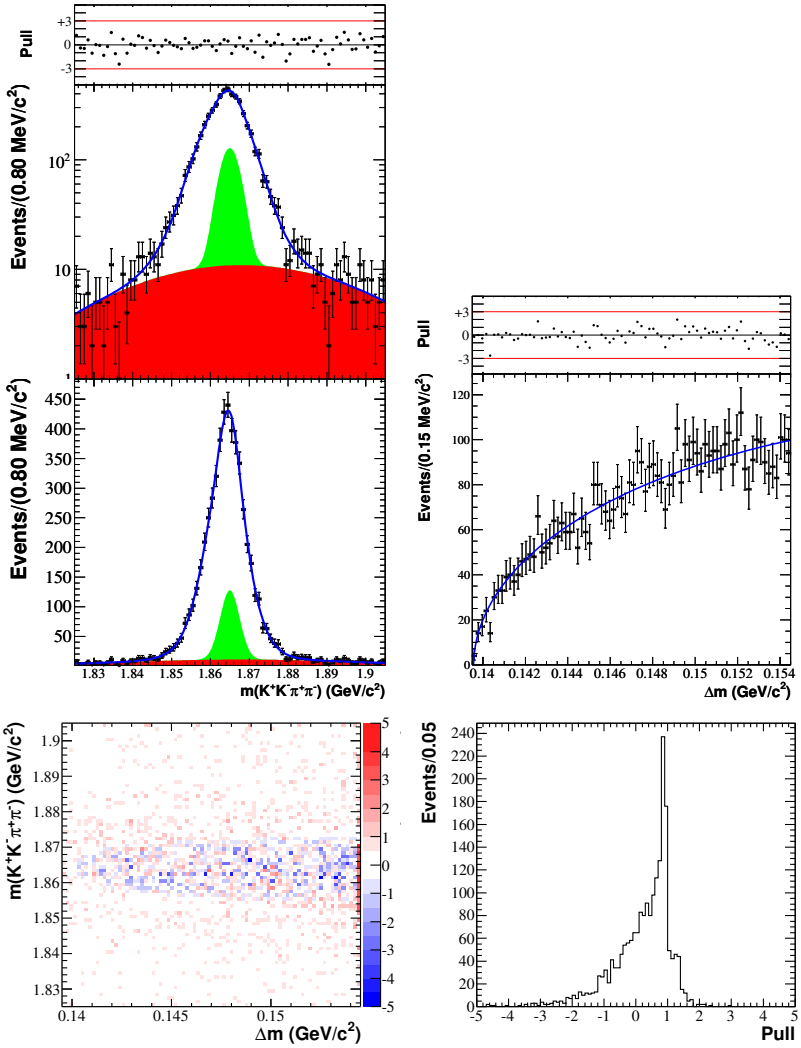


Figure 15: Fit to the distribution of D^0 peaking events from Monte Carlo. On top the projections on m(K^+ K^- π^+ π^-) and Δm of the fit model (blue line) and its components (color filled areas) are shown using a logarithmic scale on the y-axis (m(K^+ K^- π^+ π^-) only). In the middle the same plots are shown in linear scale. Over these plots the distribution of the residuals (Pull - see Eq. 4.1) is shown. On the bottom there are the distributions of the residuals in the two-dimensional [m(K^+ K^- π^+ π^-),Δm] plane (left) and integrated (right).

4.5.4 Δm Peaking Background

The Δm peaking background is due to the reconstruction of events where one or more daughters of the D^0 meson are not correctly reconstructed. On the other side, the slow pion π_s^+ is correctly reconstructed, producing a broad bump into the Δm distribution. The small peak that corresponds to a π_s^+ coming from a D^{*+} decay, is due to the reconstruction of $D^0 \to K^-\,\pi^+\,\pi^+\,\pi^-$, where one of the π^+ has been identified as a kaon: this way the wrong D^0 mass peak is shifted up into the spectrum, out of the [1.82,1.91] GeV/c^2 mass range shown in the plot, while Δm peaks in the right place.

Since the distributions of $m(K^+\,K^-\,\pi^+\,\pi^-)$ and Δm spectra are not correlated in this category of events, the two-dimensional PDF is made of the product of two separate distributions describing $m(K^+\,K^-\,\pi^+\,\pi^-)$ and Δm spectra. The distribution that describes the $m(K^+\,K^-\,\pi^+\,\pi^-)$ spectrum is taken from the combinatorial background and is defined by Eq. 4.11. The description of the Δm spectrum is made by means of a PDF that is composed by two threshold functions and the projection of the signal PDF on Δm to describe the small peak that is on top of the bump:

$$\begin{aligned} dm_{peak}(x;\overline{x};\sigma_x,\eta,c) = f_{\Delta mNar}\cdot\{f_{JSU}\cdot\mathcal{P}^{\Delta m}_{JSU} + (1-f_{JSU})\cdot \\ [f_{core}\cdot\mathcal{P}^{\Delta m}_{3} + (1-f_{core})\cdot\mathcal{P}^{\Delta m}_{4}]\} + (1-f_{\Delta mNar})\cdot \\ \left[f_{\Delta m_2}\cdot g(x;\overline{x};\sigma_x)\cdot a(x;\eta) + (1-f_{\Delta m_2})\cdot a(x;c_{\Delta m})\right], \end{aligned} \tag{4.19}$$

where

$$\mathcal{P}^{\Delta m}_{JSU} = Jsu(\Delta m;\overline{\Delta m}_{JSU},\sigma_{\Delta m_{JSU}},\gamma,\delta) \tag{4.20}$$

$$\mathcal{P}^{\Delta m}_{3} = g(\Delta m;\overline{\Delta m}_3,\sigma_{\Delta m_3}) \tag{4.21}$$

$$\mathcal{P}^{\Delta m}_{4} = g(\Delta m;\overline{\Delta m}_4,\sigma_{\Delta m_4}). \tag{4.22}$$

It can be noticed that the component $\mathcal{P}^{\Delta m}_{wide}$ (see Equation 4.9) has not been included in the peak because it consists on a very small amount of events and the three components $\mathcal{P}^{\Delta m}_{JSU}$, $\mathcal{P}^{\Delta m}_{3}$ and $\mathcal{P}^{\Delta m}_{4}$ are enough to describe this small peak.

The two-dimensional PDF is

$$\begin{aligned} \mathcal{F}_{\Delta m-peak}(m,\Delta m;\overline{\Delta m}_2,\sigma_{\Delta m_2},\eta,c_{\Delta m},b) = \\ (1+b(m-1.865\,\text{GeV}/c^2))\times dm_{peak}(\Delta m;\overline{\Delta m}_2,\sigma_{\Delta m_2},\eta,c_{\Delta m}). \end{aligned} \tag{4.23}$$

The results of the fit to the Δm peaking category from Monte Carlo are shown in Figure 16.

4.5.5 D_s^+ Background Category

The study of the correctly reconstructed Monte Carlo sample showed the presence of $D_s^+ \to K^+\,K^-\,\pi^+\,\pi^-\,\pi^+$ decays among the reconstructed events. Even if this category of events could not easily be identified in the data sample, a component of the fit model has been

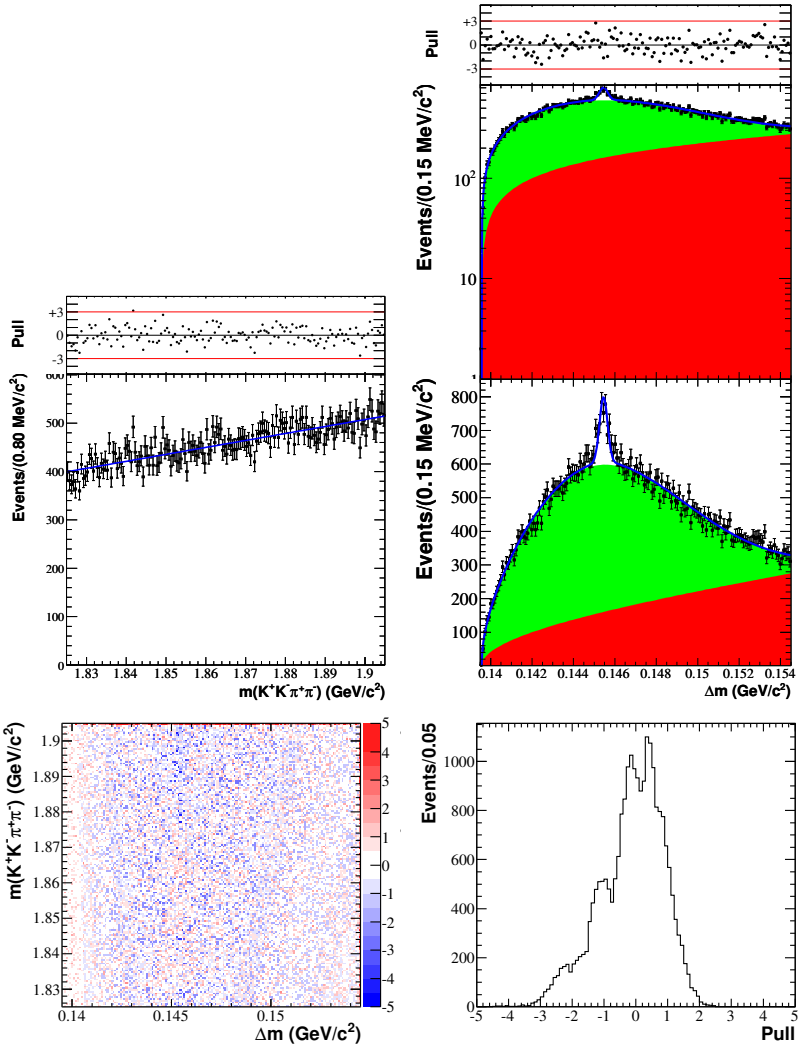


Figure 16: Fit to the distribution of Δm peaking events from Monte Carlo. On top the projections on $m(K^+\ K^-\ \pi^+\ \pi^-)$ and Δm of the fit model (blue line) and its components (color filled areas) are shown using a logarithmic scale on the y-axis (Δm only). In the middle the same plots are shown in linear scale. Over these plots the distribution of the residuals (Pull - see Eq. 4.1) is shown. On the bottom there are the distributions of the residuals in the two-dimensional [$m(K^+\ K^-\ \pi^+\ \pi^-)$,$\Delta m$] plane (left) and integrated (right).

included to describe it. The definition of the PDF for this category of events is based on the fact that

$$m(D_s^+) = m(K^+K^-\pi^+\pi^-\pi_s^+) = m(K^+K^-\pi^+\pi^-) + \Delta m$$

and that the masses of the D_s^+ events should be distributed over a Gaussian shape. On the other side, the Δm distribution should be as the same as combinatorial. One then have

$$\mathcal{F}_{D_s^+}(m, \Delta m; \bar{m}_{D_s^+}, \sigma_{D_s^+}, h) = h + \exp\left(-\frac{(m + \Delta m - \bar{m}_{D_s^+})^2}{2\sigma^2_{D_s^+}}\right) \times a(\Delta m, c), \quad (4.24)$$

where the offset h is included to describe a small amount of continuum events found in this category. This offset has not been included in the full fit model because these events are absorbed by the combinatorial shapes.

The results of the fit to the Δm peaking category from Monte Carlo are shown in Figure 17.

4.6 THE FIT MODEL

The parametrizations of each category of events have been combined to build the model for the events that have been reconstructed. The model is a sum of all the previously defined PDF:

$$\mathcal{F} = N_{sig} \cdot \mathcal{F}_{sig} + N_{comb} \cdot \mathcal{F}_{comb} + N_{D_s^+} \cdot \mathcal{F}_{D_s^+} + N_{D^0-peak} \cdot \mathcal{F}_{D^0-peak} + N_{\Delta m-peak} \cdot \mathcal{F}_{\Delta m-peak}. \quad (4.25)$$

Each PDF is normalized and the fit is done on the binned data set using an extended likelihood to get the number of events for each category. In the model many of the parameters are shared among PDFs referring to different categories of events. The relationship among different PDFs is shown in detail by means of the pictorial view of Figure 18.

In Figure 19, the excellent agreement to the Monte Carlo sample is shown. Different color are used to highlight each category of events and the projections of the fit are shown both in linear and logarithmic scale to show in detail the features of each category. The results of the fit are shown in Table 7.

Given the large number of floating parameters, there might be large correlations between them. When many correlations are present, the error of some of the fit parameters could not be correctly evaluated. To overtake this problem and measure the asymmetry parameter obtaining the correct error, the model is first fitted to the sample leaving all the parameters floating: this fit is used to fix the parameters directly related to the shapes of the distributions. The other parameters, such as the asymmetries and the number of events for each category are left floating into another fit, that returns them with the correct errors.

The asymmetry parameters are directly included in the fit thanks to the specific way the model is designed. Following the relationship

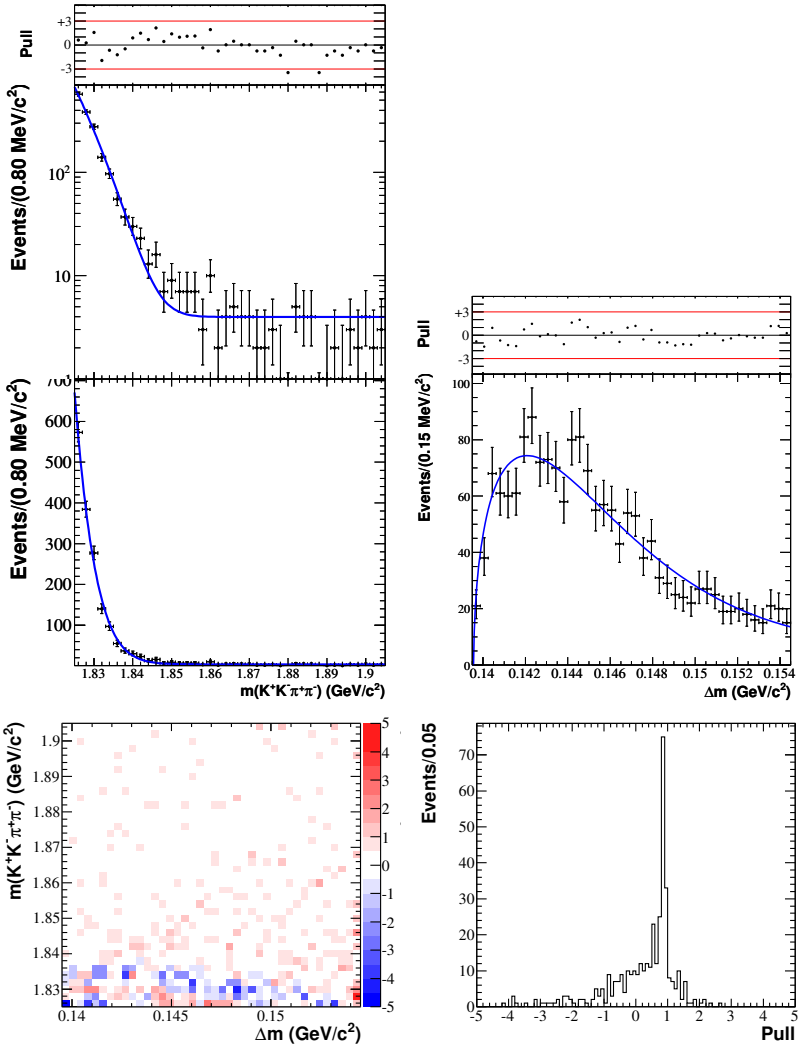


Figure 17: Fit to the distribution of D_s^+ events from Monte Carlo. On top the projections on m(K^+ K^- π^+ π^-) and Δm of the fit model (blue line) are shown using a logarithmic scale on the y-axis (m(K^+ K^- π^+ π^-) only). In the middle the same plots are shown in linear scale. Over these plots the distribution of the residuals (Pull - see Eq. 4.1) is shown. On the bottom there are the distributions of the residuals in the two-dimensional [m(K^+ K^- π^+ π^-),Δm] plane (left) and integrated (right).

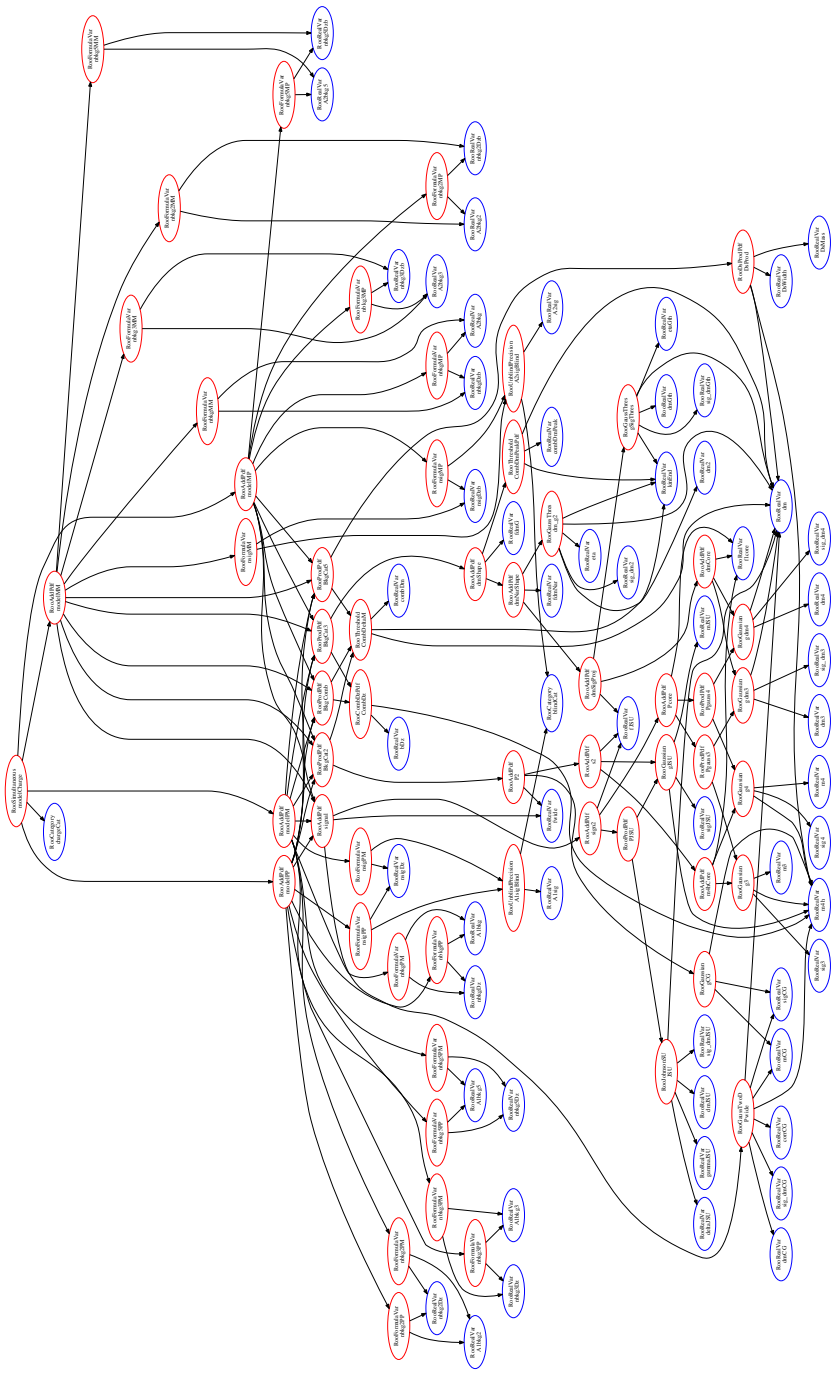


Figure 18: This graphics shows the structure of the fitting model: the blue circles represent the variables, while the red ones the PDFs. The arrows show to which variables the PDFs are related to. It can be seen that some variables are shared between two different PDF.

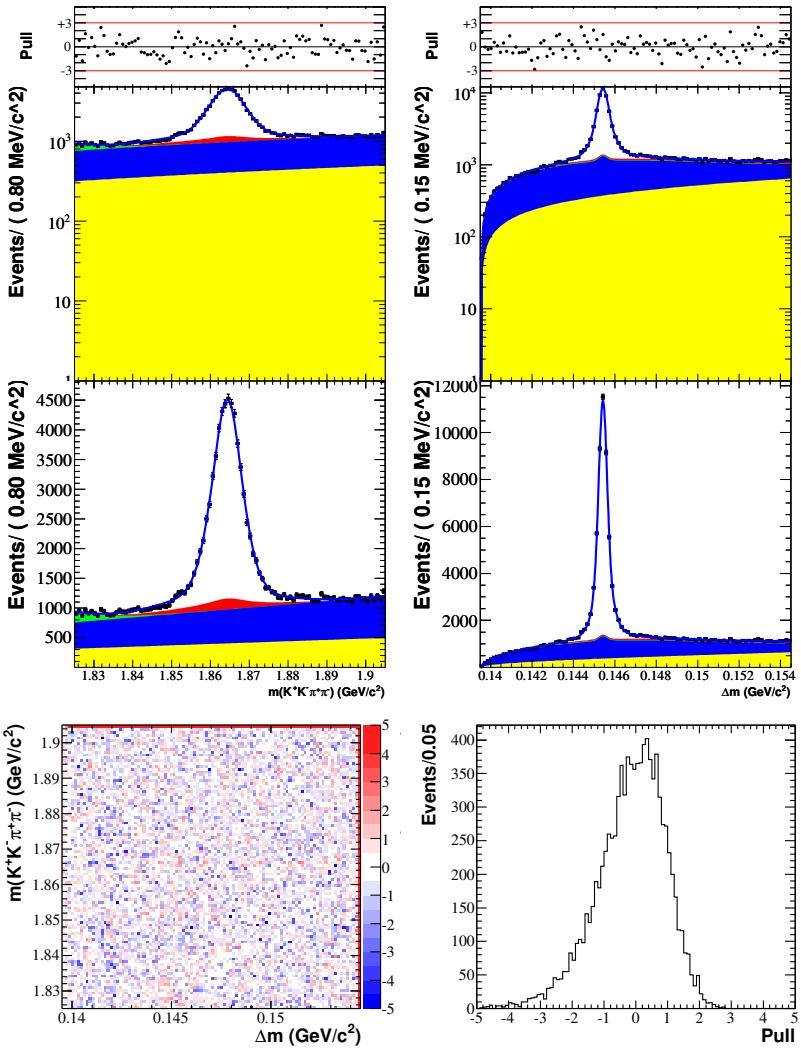


Figure 19: Fit to the distribution of the events from the Monte Carlo sample. On top the projections on m(K^+ K^- π^+ π^-) and Δm of the fit model (blue line) and its components (color filled areas) are shown using a logarithmic scale on the y-axis. The colors refer to the different categories of events: white - signal; yellow - combinatorial; red - D^0 peaking; blue - Δm peaking; green - D_s^+. In the middle the same plots are shown in linear scale. Under each plot the distribution of the residuals (Pull - see Eq. 4.1) is shown. On the bottom there are the distributions of the residuals in the two-dimensional [m(K^+ K^- π^+ π^-),Δm] plane (left) and integrated (right).

Table 7: Results of the fit to the full Monte Carlo sample. The values of the masses and the widths are expressed in terms of MeV/c^2.

Variable	Fit Result	Variable	Fit Result
A_T sig	$(0.40 \pm 0.47)\%$	$\bar{A}_T$ sig	$(-0.06 \pm 0.46)\%$
A_T bkg	$(-5.26 \pm 1.55)\%$	$\bar{A}_T$ bkg	$(3.68 \pm 1.73)\%$
A_T bkg2	$(19.0 \pm 10.0)\%$	$\bar{A}_T$ bkg2	$(-23.5 \pm 8.6)\%$
A_T bkg3	$(0.59 \pm 0.67)\%$	$\bar{A}_T$ bkg3	$(-0.66 \pm 0.65)\%$
A_T bkg5	$(2.85 \pm 7.59)\%$	$\bar{A}_T$ bkg5	$(0.73 \pm 8.49)\%$
N(D^0)	51972 ± 241	N($\bar{D}^0$)	52469 ± 243
N(D^0) cat2	1682 ± 165	N($\bar{D}^0$) cat2	1959 ± 164
N(D^0) cat3	36129 ± 243	N($\bar{D}^0$) cat3	37643 ± 244
N(D^0) cat4	14575 ± 225	N($\bar{D}^0$) cat4	12996 ± 224
N(D^0) cat5	1080 ± 82	N($\bar{D}^0$) cat5	957 ± 81
$\overline{m}_{D_s^+}$	1910.0 ± 3.1	$\sigma_{m_{D_s^+}}$	19.3 ± 0.5
b_1	3.71 ± 0.14	c	$(0.02 \pm 9.23)\%$
$\overline{\Delta m}_2$	145.78 ± 0.05	$\sigma_{\Delta m_2}$	4.83 ± 0.08
η	9.95 ± 0.15	$c_{\Delta m}$	$(0.001 \pm 10.500)\%$
$f_{\Delta mNar}$	$(0.9 \pm 0.3)\%$	$f_{\Delta m_2}$	$(17.0 \pm 0.5)\%$
γ	-1.03 ± 0.03	δ	0.91 ± 0.04
c_{CG}	0.013 ± 0.013		
$\overline{m}_{CG}$	1863.4 ± 0.1	$\sigma_{m_{CG}}$	11.80 ± 0.13
$\overline{m}_3$	1864.6 ± 0.1	σ_{m_3}	4.18 ± 0.02
$\overline{m}_4$	1864.5 ± 0.1	σ_{m_4}	3.86 ± 0.07
$\overline{m}_{JSU}$	1864.3 ± 0.1	$\sigma_{m_{JSU}}$	4.25 ± 0.13
$\overline{\Delta m}_3$	145.45 ± 0.01	$\sigma_{\Delta m_3}$	0.185 ± 0.001
$\overline{\Delta m}_4$	145.52 ± 0.01	$\sigma_{\Delta m_4}$	0.581 ± 0.009
$\overline{\Delta m}_{CG}$	145.45 ± 0.01	$\sigma_{\Delta m_{CG}}$	0.345 ± 0.004
$\overline{\Delta m}_{JSU}$	144.00 ± 0.01	$\sigma_{\Delta m_{JSU}}$	2.08 ± 0.09
f_{core}	$(83.7 \pm 0.3)\%$	f_{JSU}	$(8.8 \pm 0.2)\%$
f_{wide}	$(19.2 \pm 0.3)\%$		

between the number of signal events and the asymmetry parameters, one gets:

$$\begin{aligned}
N(D^0, C_T > 0) &= \frac{N(D^0)}{2}(1 + A_T)\,, \\
N(D^0, C_T < 0) &= \frac{N(D^0)}{2}(1 - A_T)\,, \\
N(\bar{D}^0, \bar{C}_T > 0) &= \frac{N(\bar{D}^0)}{2}\left(1 - \bar{A}_T\right), \\
N(\bar{D}^0, \bar{C}_T < 0) &= \frac{N(\bar{D}^0)}{2}\left(1 + \bar{A}_T\right). \qquad (4.26)
\end{aligned}$$

There is then a mutual correspondence between the following two sets of variables:

$$\begin{gathered}
\left\{N(D^0, C_T > 0), N(D^0, C_T < 0), N(\bar{D}^0, \bar{C}_T > 0), N(\bar{D}^0, \bar{C}_T < 0)\right\} \\
\updownarrow \\
\left\{N(D^0), N(\bar{D}^0), A_T, \bar{A}_T\right\}
\end{gathered}$$

The latter have been used in the fit to evaluate the above yields.

In order to evaluate directly the asymmetry parameters, the data sample has been split into four sub-samples, depending on D flavor and C_T (or $\bar{C}_T$) value. These sub-samples have been fitted simultaneously to the model to get the asymmetry parameters A_T and $\bar{A}_T$ and the number of D^0 and $\bar{D}^0$ decays.

The validation of the fit model has been made using the Monte Carlo sample. At first the yields retrieved from the fit for each category of events are compared to their true number in Table 8: The only category

Table 8: Reconstructed and fitted number of events for each of the categories identified in the Monte Carlo sample. The last column shows the number of σ's between the true and the fitted number of events: $\varepsilon = \frac{n - n_{rec}}{\sigma_{n_{rec}}}$.

Category	MC Events	Fit Results	ε
Signal	103561	104441 ± 342	$+2.57$
D^0 peaking	5705	3641 ± 233	-8.86
Δm peaking	73138	73772 ± 344	$+1.84$
Combinatorial	27306	27571 ± 317	$+0.84$
D_s^+	1753	2037 ± 115	$+2.47$
Signal	211464	211462 ± 635	-0.003

that shows a fair agreement to the fit result is the D^0 peaking one. Since the number of missing events is much less than the number of events in the signal category (about 1%), this effect has been neglected.

In Table 9, the signal yields for each sub-sample are compared to the fit results: every yield has been a little over-estimated by the fit. Anyway, since the trend of the over-estimation is the same for each one of the yields, the final results of the asymmetry parameters should not be affected.

Table 9: Reconstructed and fitted number of signal events for each sub-sample. The last column shows the number of σ's between the true and the fitted number of events: $\varepsilon = \frac{n-n_{rec}}{\sigma_{n_{rec}}}$.

Sub-Sample	MC Events	Fit Results	ε
$\Gamma(D^0, C_T > 0)$	25789	26091 ± 171	+1.77
$\Gamma(D^0, C_T > 0)$	25736	25881 ± 170	+0.85
$\Gamma(\bar{D}^0, \bar{C}_T > 0)$	26041	26250 ± 172	+1.21
$\Gamma(\bar{D}^0, \bar{C}_T > 0)$	25995	26219 ± 171	+1.31

Finally, the asymmetry parameters retrieved from the fit have been compared to their true values reconstructed from Monte Carlo in Table 10. All the parameters retrieved from the fit are consistent with

Table 10: Reconstructed and fitted asymmetry parameters. The last column shows the number of σ's between the true and the fitted parameter: $\varepsilon = \frac{A-A_{rec}}{\sigma_{A_{rec}}}$.

Parameter	Reconstructed ($\times 10^{-3}$)	Fit Results ($\times 10^{-3}$)	ε
A_T	1.03	4.04 ± 4.64	+0.65
$\bar{A}_T$	-0.88	-0.59 ± 4.62	+0.06
$\mathcal{A}_T$	0.95	2.33 ± 3.34	+0.41

the value reconstructed from the Monte Carlo sample. However, the fit bias that has been found

$$\mathcal{A}_T^{fit} - \mathcal{A}_T^{reco} = 1.38 \times 10^{-3}, \tag{4.27}$$

has been taken into account in the evaluation of the systematic error, that will be discussed in the following.

4.6.1 Fit to the Data Sample

The previously defined and validated model has been used to describe the data sample. The only difference here is that the real value of the asymmetry parameters is kept hidden until the very final stage of the analysis. This general practice, also known as "Blind Analysis", allows to optimize the cuts and evaluate the systematics before obtaining the final results, removing the bias that could be unconsciously (or consciously) introduced modifying the cuts to enhance or suppress the searched effect.

The asymmetry parameters have been masked in this analysis adding an unknown random offset. The result is that the T violating parameter $\mathcal{A}_T$ is shifted of an unknown value. Even if the offset is unknown, it is kept under control assuming the same seed every time it is generated. In this way, the difference between the fit results is the same both in the case that the offset is on or off.

The results of the fit to the data are shown in Figure 20. The yields retrieved from the fit for each category of events are shown in Table 11, while the results for the full set of floating variables are shown in Table 12.

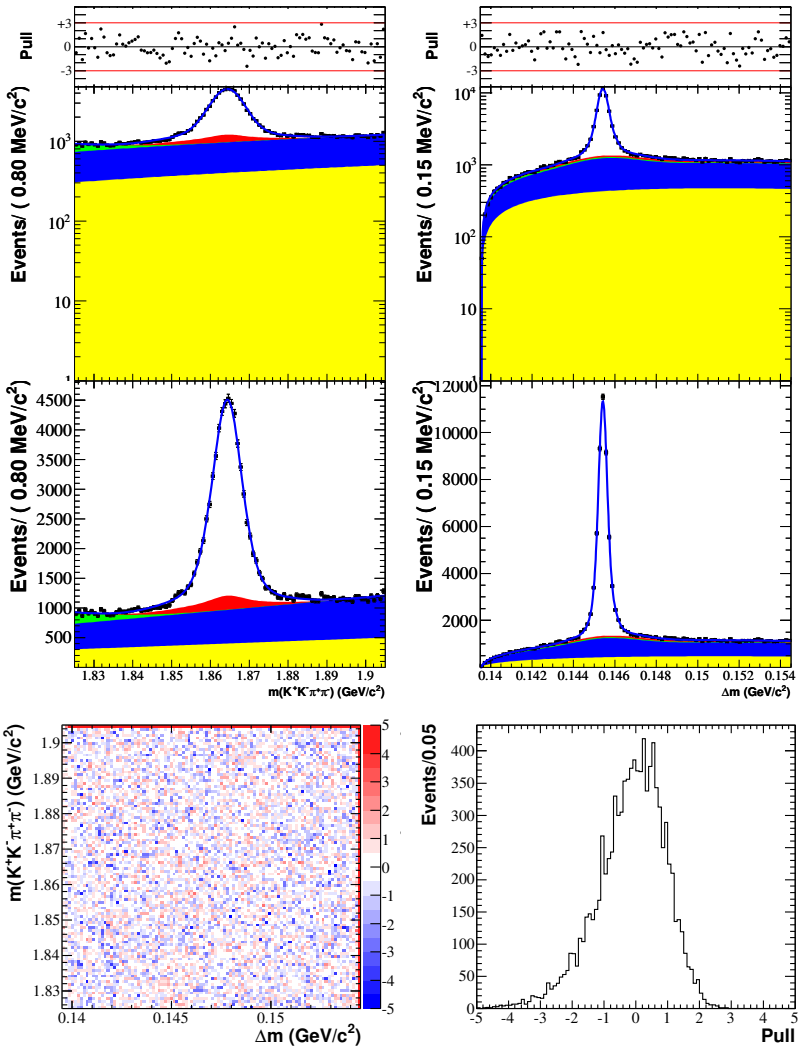


Figure 20: Fit to the distribution of the events from data sample. On top the projections on m(K^+ K^- π^+ π^-) and Δm of the fit model (blue line) and its components (color filled areas) are shown using a logarithmic scale on the y-axis. The colors refer to the different categories of events: white - signal; yellow - combinatorial; red - D^0 peaking; blue - Δm peaking; green - D_s^+. In the middle the same plots are shown in linear scale. Under each plot the distribution of the residuals (Pull - see Eq. 4.1) is shown. On the bottom there are the distributions of the residuals in the two-dimensional [m(K^+ K^- π^+ π^-),Δm] plane (left) and integrated (right).

Table 11: Number of events for each category of events retrieved from the fit to the data sample.

Category	Fit Results
Signal	46691 ± 241
D^0 peaking	5178 ± 331
Δm peaking	57099 ± 797
Combinatorial	40512 ± 818
D_s^+	2023 ± 156
Total	151503 ± 1223

4.7 SYSTEMATICS

Before unblinding the final result, all the systematics have been evaluated. The sources of systematics recognized in this analysis are:

- the fitting model;
- the particle identification;
- the $p^*(D^0)$ cut;
- the identification of the soft pion;
- the effect of forward-backward asymmetry;
- mistag;
- fit bias;
- detector asymmetry.

The procedure to evaluate the contribution to the systematic error associated to each one of these sources can be outlined in a few steps:

1. the specific cut related to the source of systematics is slightly modified;
2. the asymmetry parameter is evaluated again through a new fit to the data;
3. the deviation from the value measured using standard cuts is taken;
4. the largest among all the measured deviations is assumed as the specific contribution to the systematic error.

The way the contributions from each source of systematics are evaluated will be described briefly in the following. At the end the value of all the contributions will be listed and the systematic error is evaluated.

Table 12: Results of the fit to the data sample. The values of the masses and the widths are expressed in terms of MeV/c^2.

Variable	Fit Result	Variable	Fit Result
A_T sig	$(7.33 \pm 0.73)\%$	$\bar{A}_T$ sig	$(8.37 \pm 0.73)\%$
A_T bkg	$(0.39 \pm 2.80)\%$	$\bar{A}_T$ bkg	$(2.07 \pm 2.93)\%$
A_T bkg2	$(20.55 \pm 10.30)\%$	$\bar{A}_T$ bkg2	$(-0.75 \pm 8.39)\%$
A_T bkg3	$(-2.03 \pm 1.99)\%$	$\bar{A}_T$ bkg3	$(-2.59 \pm 1.97)\%$
A_T bkg5	$(-1.99 \pm 10.50)\%$	$\bar{A}_T$ bkg5	$(7.85 \pm 11.59)\%$
N(D^0) sig	23561 ± 171	N($\bar{D}^0$) sig	23129 ± 170
N(D^0) cat4	20857 ± 583	N($\bar{D}^0$) cat4	19655 ± 576
N(D^0) cat2	2382 ± 235	N($\bar{D}^0$) cat2	2797 ± 234
N(D^0) cat3	28497 ± 567	N($\bar{D}^0$) cat3	28603 ± 562
N(D^0) cat5	1060 ± 112	N($\bar{D}^0$) cat5	963 ± 110
$\overline{m}_{D_s^+}$	1910.0 ± 45.3	$\sigma_{m_{D_s^+}}$	25.0 ± 2.0
b_1	5.71 ± 0.52	c	$(0.001 \pm 7.397)\%$
$\overline{\Delta m}_2$	146.08 ± 0.47	$\sigma_{\Delta m_2}$	6.50 ± 0.48
η	8.29 ± 0.86	$c_{\Delta m}$	$(0.001 \pm 19.972)\%$
$f_{\Delta mNar}$	$(2.0 \pm 0.6)\%$	$f_{\Delta m_2}$	$(74.3 \pm 5.1)\%$
γ	-1.33 ± 0.28	δ	2.54 ± 0.18
c_{CG}	0.04 ± 0.03		
$\overline{m}_{CG}$	1862.1 ± 0.3	$\sigma_{m_{CG}}$	13.0 ± 0.7
$\overline{m}_3$	1864.2 ± 0.1	σ_{m_3}	4.36 ± 0.08
$\overline{m}_4$	1864.6 ± 0.1	σ_{m_4}	2.67 ± 0.10
$\overline{m}_{JSU}$	1864.2 ± 0.1	$\sigma_{m_{JSU}}$	4.25 ± 0.20
$\overline{\Delta m}_3$	145.42 ± 0.02	$\sigma_{\Delta m_3}$	0.195 ± 0.003
$\overline{\Delta m}_4$	145.45 ± 0.01	$\sigma_{\Delta m_4}$	0.37 ± 0.03
$\overline{\Delta m}_{CG}$	145.45 ± 0.01	$\sigma_{\Delta m_{CG}}$	0.44 ± 0.03
$\overline{\Delta m}_{JSU}$	144.16 ± 0.22	$\sigma_{\Delta m_{JSU}}$	2.62 ± 0.37
f_{core}	$(74.4 \pm 2.6)\%$	f_{JSU}	$(17.4 \pm 1.5)\%$
f_{wide}	$(19.5 \pm 0.8)\%$		

4.7.1 Fitting Model

The fitting model can introduce a systematic error due to the shapes used to describe each category of events. Another source of systematic error related to the fit is the binning used.

The systematic error due to the signal PDF shape has been estimated changing the $\mathcal{P}_{JSU}$ component, substituting the J_{SU} (see Eq. 4.8) with a Crystal Ball shape[Gaiser, 1982]

$$\mathcal{P}_{CB}(x;\alpha,n,\bar{x},\sigma) = \begin{cases} \exp\left(-\frac{(x-\bar{x})^2}{2\sigma^2}\right) & \text{for } \frac{x-\bar{x}}{\sigma} > -\alpha \\ \left(\frac{n}{|\alpha|}\right)^n \exp\left(-\frac{|\alpha|^2}{2}\right)\left(\frac{n}{|\alpha|} - |\alpha| - \frac{x-\bar{x}}{\sigma}\right)^{-n} & \text{for } \frac{x-\bar{x}}{\sigma} \leqslant -\alpha \end{cases} \tag{4.28}$$

obtaining the distribution

$$\mathcal{P}_{sys}^{JSU} = g(m; m_{JSU}, \sigma_{JSU}) \times \mathcal{P}_{CB}(\Delta m; \alpha^{CB}, n^{CB}, \Delta m^{CB}, \sigma_{\Delta m}^{CB})$$

Another source of systematics has been identified in the description of the category 3 background (Δm peaking): an alternative way to describe it, is to use a Johnson SU with threshold instead of the gaussian with threshold

$$J_{SU}^{th}(\Delta m; \Delta m_{th}, \overline{\Delta m}_{JSU}, \sigma_{\Delta m_{JSU}}, \gamma, \delta, \eta) = J_{SU}(\Delta m; \overline{\Delta m}_{JSU}, \sigma_{\Delta m_{JSU}}, \gamma, \delta) \cdot a(\Delta m; \eta). \tag{4.29}$$

The binning used to measure the reference value has been set to $4 \times 100 \times 100 = 40000$ bins, where 4 is due to the simultaneous fit on the 4 data set built from D^0 flavour and C_T. The choice of the number of bins has been driven from the requirement of a good resolution. Two other sets of binning have been used to estimate the systematic error introduced by this choice: $4 \times 80 \times 80 = 25600$ bins and $4 \times 120 \times 120 = 57600$ bins.

4.7.2 Particle Identification

The pions and the kaons are identified by means of algorithms that combine the informations from the detectors in *BABAR* into selectors using the technique of the likelihood. Two criteria are considered to reconstruct the D^0 events using these selectors: the looser one is the one used in the standard reconstruction of events, the tighter one is used for systematics. The violation parameter is measured again using the tighter criteria to identify kaons or pions alternatively and the larger deviation among the different selections of the data sample has been taken as the systematic's error due to particle identification. The selection using tighter criteria reduces significantly the number of events to be considered in the data set, as shown in Table 13. The deviations of the measurement of $\mathcal{A}_T$ from the standard criteria are then influenced by statistics effects. The results on the contribution to systematics error are not invalidated, but have to be considered a conservative estimate.

Table 13: Number of events in the data set after the application of the selectors using tighter criteria.

Selection	Events
$K_l K_l \pi_l \pi_l$	151,505
$K_l K_l \pi_t \pi_t$	142,449
$K_t K_t \pi_l \pi_l$	125,096
$K_t K_t \pi_t \pi_t$	117,854

4.7.3 $p^*(D^0)$ cut

The selection applied on the D^0 momentum in the center-of-mass frame has been tightened to higher values of momentum to study the variation of the asymmetry parameter. Two selections have been considered:

- $p^*(D^0) > 2.6\,\text{GeV/c}$;
- $p^*(D^0) > 2.7\,\text{GeV/c}$.

The number of the events in the data set is reduced by this cut as shown in Table 14, the deviations of $\mathcal{A}_T$ are then influenced by statistics effect here too. As well as for particle identification, the contribution

Table 14: Number of events in the data set after the application of tighter selections on the momentum of D^0 candidate in the center-of-mass frame.

Selection	Events
$p^*(D^0) > 2.5\,\text{GeV/c}$	151,505
$p^*(D^0) > 2.6\,\text{GeV/c}$	131,569
$p^*(D^0) > 2.7\,\text{GeV/c}$	119,903

to systematics error is considered a conservative estimate.

4.7.4 Soft Pion Charge

The right identification of the charge of the slow momentum pion π_s^+ is crucial to determine the flavor of the D^0 candidate. A systematic error can be introduced if the charge of the reconstructed π_s^+ is different from the charge of the D^* candidate. In order to evaluate the effect, if any, on this analysis, the correctly reconstructed Monte Carlo sample has been used to check how many times the π_s^+ is reconstructed with the wrong charge. Table 15 shows the results of this study for each Run of data taking. The net amount of π_s^+ mis-identification is 4, corresponding to 2×10^{-5} the total.

4.7.5 Effect of Forward-Backward Asymmetry

The effect of forward-backward asymmetry (see Appendix A) on this analysis has been studied evaluating the T asymmetry parameter

Table 15: Study of slow momentum pion mis-identification using correctly reconstructed Monte Carlo.

Run	Wrong D^{*+}	Wrong D^{*-}	Missed	Total	% Total
1	4	2	0	14,995	4×10^{-4}
2	7	7	1	39,309	4×10^{-4}
3	1	4	1	19,280	3×10^{-4}
4	25	26	1	53,517	10×10^{-4}
5	21	24	5	63,707	8×10^{-4}
6	6	5	3	20,656	7×10^{-4}
Sum	64	68	11	211,464	7×10^{-4}

$\mathcal{A}_T$ in regions of $\cos\theta^*$, the cosine of the production angle of D^0 in the center-of-mass frame. The three considered regions are:

- $-1.0 \leqslant \cos\theta^* < -0.3$;
- $-0.3 \leqslant \cos\theta^* < +0.3$;
- $+0.3 \leqslant \cos\theta^* \leqslant +1.0$.

The results of this study are shown in Table 16 and in Figure 21.

Table 16: Number of signal events, deviation from $\mathcal{A}_T$ using standard cuts and error on the T violation parameter measured for each sub-sample of $\cos\theta^*$

$\cos\theta^*$ Region	Signal Events	$\Delta \times 10^{-3}$	$\sigma \times 10^{-3}$
$-1.0 \leqslant \cos\theta^* < -0.3$	20202 ± 158	+6.21	7.79
$-0.3 \leqslant \cos\theta^* < +0.3$	16707 ± 144	-0.74	8.58
$+0.3 \leqslant \cos\theta^* \leqslant +1.0$	10548 ± 120	-11.16	11.38
Total	47457 ± 245	0.94	

The contribution to systematic error cannot be measured simply taking the larger among the three deviations found. This procedure would lead to an overestimation of the error itself due to statistical effects introduced by the splitting of the data set. In order to account for the statistical effects, the deviations are weighted to the error of their measurement:

$$\sigma^{FB}_{(syst)} = \frac{\sum_i \frac{\Delta_i}{\sigma_i^2}}{\sum_i \frac{1}{\sigma_i^2}}.$$

4.7.6 Fit Bias

The difference between the generated and reconstructed T violation parameter has been evaluated using Monte Carlo simulations (see Eq. 4.27).

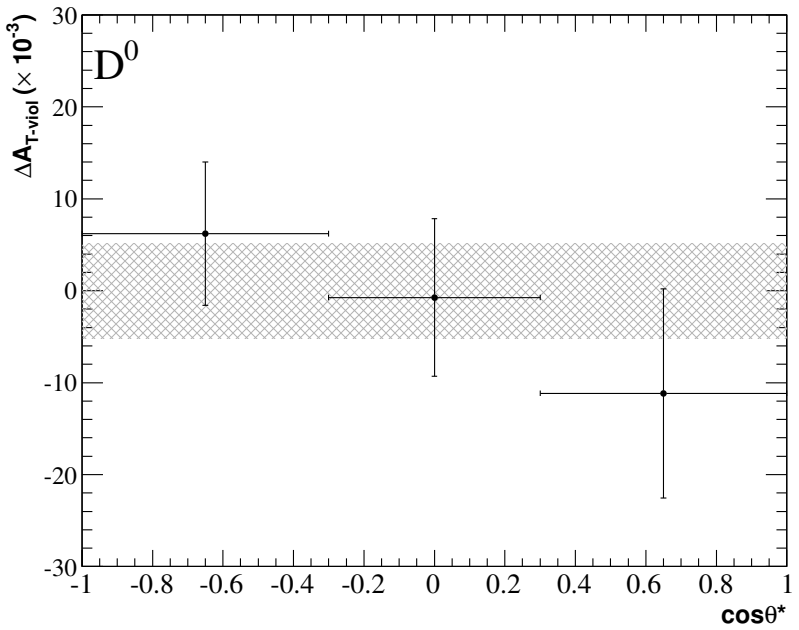


Figure 21: Deviations from $\mathcal{A}_T$ measured in different regions of the detector. The grey shaded area is centered at 0 and corresponds to $\mathcal{A}_T$ measured on the full data set with errors.

4.7.7 Detector Asymmetry

The Monte Carlo sample made of signal events decaying by phase space has been used to test for asymmetries in the reconstruction from the detector. The generated asymmetry has been compared to the one retrieved from correctly reconstructed signal decays. The deviation has been assumed as the systematic error.

4.7.8 Systematics Summary

Finally, the systematic error has been evaluated summing in quadrature all the contributions shown in Table 17. In this table the contributions to systematic error for A_T and $\bar{A}_T$ are quoted as well: they are measured as the same as $\mathcal{A}_T$.

4.8 FINAL RESULT

After unblinding, the following results for the asymmetry parameters have been retrieved:

$$\begin{aligned} A_T &= (-68.5 \pm 7.3(\text{stat}) \pm 5.8(\text{syst})) \times 10^{-3}, \\ \bar{A}_T &= (-70.5 \pm 7.3(\text{stat}) \pm 3.9(\text{syst})) \times 10^{-3}. \end{aligned} \tag{4.30}$$

It can be observed nonzero values of A_T and $\bar{A}_T$ indicating that final state interaction effects are significant in this D^0 decay. This asymme-

Table 17: Systematic uncertainty evaluation on $\mathcal{A}_T$, A_T and $\bar{A}_T$ in units of 10^{-3}.

Effect	$\mathcal{A}_T$	A_T	$\bar{A}_T$
Alternative signal PDF	0.2	0.3	0.2
Alternative mis-reconstructed D^0 PDF	0.5	0.1	0.9
Bin size	0.2	0.4	0.3
Particle identification	3.5	4.2	2.9
$p^*(D^0)$ cut	1.7	1.6	2.4
$\cos\theta^*$ dependence	0.9	0.0	0.2
Fit bias	1.4	3.0	0.3
Mistag	0.0	0.0	0.0
Detector asymmetry	1.1	2.1	0.0
Total	4.4	5.8	3.9

tries are clearly observable from the projections of the fit results into the four sub-samples depending on D^0 flavor and C_T ($\bar{C}_T$). No effect is found, on the other hand, in the analysis of Monte Carlo samples. Final state interaction effects are common in hadronic D decays because of the complex interference patterns between intermediate resonances formed between hadrons in the final states [Oller, 2005].

The final result for the CP violation parameter, $\mathcal{A}_T$, is

$$\mathcal{A}_T = (1.0 \pm 5.1(\text{stat}) \pm 4.4(\text{syst})) \times 10^{-3}. \tag{4.31}$$

In conclusion, this search for CP violation using T-odd correlations in the high statistics sample of Cabibbo suppressed $D^0 \to K^+ K^- \pi^+ \pi^-$ decays gives a T-violationg asymmetry consistent with zero with a sensitivity of the order of 0.5%.

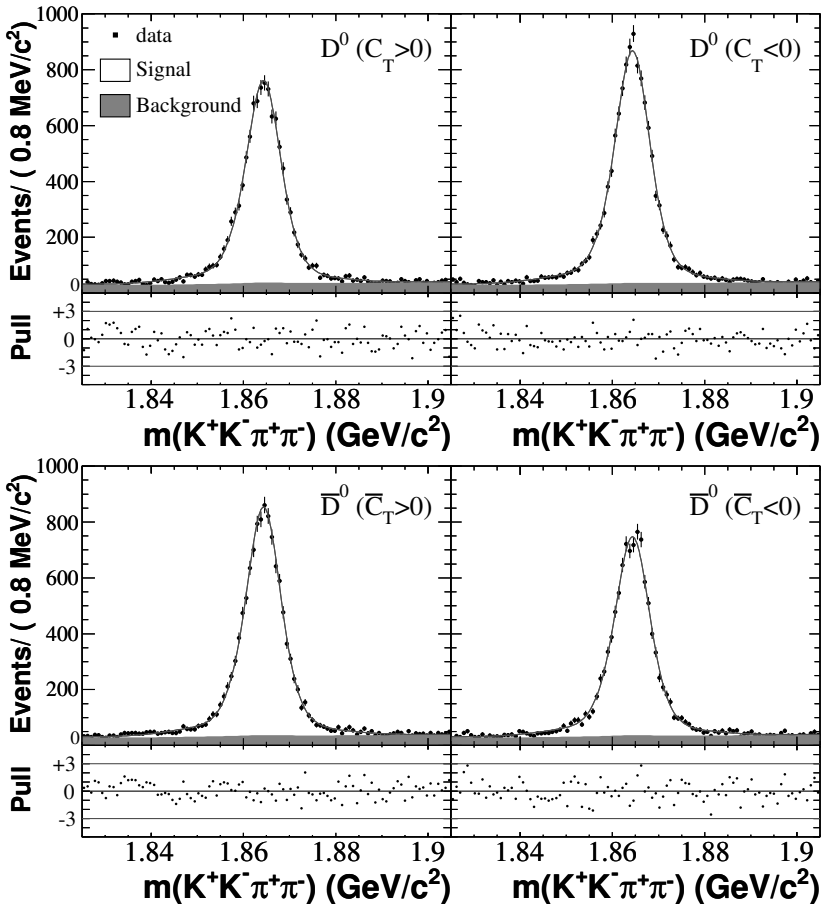


Figure 22: Data and fit projections in Δm signal region. The four sub-samples built using the D^0 flavor and the C_T ($\bar{C}_T$) value are shown.

5 $D^+_{(s)} \to K^0_S K^+ \pi^+ \pi^-$ ANALYSIS

CONTENTS

The analysis of the charged D decay modes is somewhat similar to what has been done on the D^0 data set. The idea is to retrieve the asymmetry parameter directly from a fit to the data sample. The main difference here is that there is no information about the D mother, so a technique based on the topology of the signal decays is needed to enhance the signal to background ratio.

In the following the data set is described; then the procedure used to select the events is outlined; finally the fit model is defined and the results on the data sample are shown.

5.1 THE DATA SET

Together with the full *BABAR* data set recorded at the $\Upsilon(4S)$ and 40 MeV below the resonance, the events recorded at the $\Upsilon(3S)$, $\Upsilon(2S)$ and below these resonances have been included, pushing the total integrated luminosity to about 531 fb^{-1}.

As well as described in Section 4.1, a large number of Monte Carlo events has been analyzed as the same as the data. The Monte Carlo events are again of two kinds: the Generic Monte Carlo, that simulates the continuum production of $c\bar{c}$ from e^+e^- and the Signal Monte Carlo that is made of events generated requiring that the signal decay chain is always present at generator level. The analysis makes use of about

1.1×10^9 Generic Monte Carlo Events and 4×10^6 Signal Monte Carlo Events.

5.2 EVENTS RECONSTRUCTION

The inclusive $D^+_{(s)} \to K^0_S\ K^+\ \pi^+\ \pi^-$ events are reconstructed together with all the $D^+_{(s)}$ decaying to three hadrons and a K^0_S:

$$D^+_{(s)} \to K^0_S h^+ h^+ h^-, \text{ being } h = \pi, K.$$

First a K^0_S candidate is built combining a couple of oppositely charged pions. The invariant mass of the K^0_S candidate is required to be

$$0.35\,\text{GeV}/c^2 < m(\pi^+\pi^-) < 0.65\,\text{GeV}/c^2$$

and the vertex is fitted requiring a probability greater than 0.1% and applying a mass constraint.

The successful K^0_S candidate is then combined to all the combinations of three charged tracks, previously identified as kaons or pions using boosted decision tree and error correcting output code selectors respectively, to form a $D^+_{(s)}$ candidate. This candidate is required to have mass in the range

$$1.80\,\text{GeV}/c^2 < m(K^0_S h^+ h^+ h^-) < 2.05\,\text{GeV}/c^2$$

and momentum in the center-of-mass frame

$$p^*(D^+_{(s)}) > 2.5\,\text{GeV}/c.$$

The probability of the fit of the vertex is again required to be greater than 0.1%.

The $D^+_{(s)}$ candidate passing all these selections is fitted again constraining the vertex into the interaction region. The χ^2 probability retrieved from this fit is recorded for further studies.

5.3 CONTAMINATIONS IN THE $D^+_{(s)}$ DATA SAMPLE

A strong contamination from

$$\begin{aligned} D^{*+} \to\ & D^0\pi^+ \\ & \hookrightarrow K^0_S K^+\pi^- \end{aligned}$$

has been found in the data sample. The D^{*+} peak is visible in the $K^+\ K^0_S\ \pi^+\ \pi^-$ invariant mass spectrum that is shown in Figure 23. This contamination is removed applying a combined cut on the $K^0_S\ K^+\ \pi^-$ invariant mass and on $\Delta m = m(K^+K^0_S\pi^+\pi^-) - m(K^0_S K^+\pi^-)$:

$$\begin{aligned} 144.46\,\text{MeV}/c^2 < \Delta m < 146.47\,\text{MeV}/c^2, \\ 1850.38\,\text{MeV}/c^2 < m(K^0_S K^+\pi^-) < 1878.16\,\text{MeV}/c^2. \end{aligned} \tag{5.1}$$

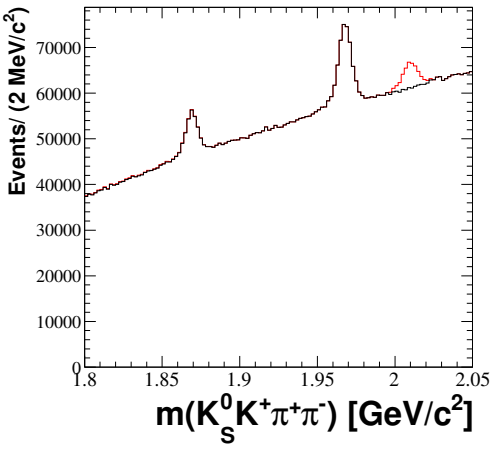


Figure 23: The $K^+\,K^0_S\,\pi^+\,\pi^-$ invariant mass after reconstruction. The D^+ and D^+_s peaks are clearly visible together with the contamination from D^{*+}. The red line shows the spectrum before removing the D^{*+} contamination box, while the black line shows the spectrum after the removal of the contamination.

The sides of this box have been retrieved from two separate fits to the $K^0_S\,K^+\,\pi^-$ invariant mass and to Δm. They correspond to $\pm 2\sigma$ mass region around the mean value of the Gaussian used in the fit. At first the Δm has been fitted to a Gaussian over a second order polynomial background, then the $K^0_S\,K^+\,\pi^-$ invariant mass has been plotted in the D^{*+} signal region retrieved by the fit to Δm spectrum. Finally the $K^0_S\,K^+\,\pi^-$ invariant mass has been fitted to a Gaussian over a second order polynomial background and the D^{*+} signal box is found. The results of these fits are shown in Figure 24. This cut removes all the events that are shown by the red line in Figure 23 and does not affect signal events.

Similarly to what has been found in the $D^0 \to K^+\,K^-\,\pi^+\,\pi^-$ data sample, a contamination from $D^+_{(s)} \to K^+\,K^0_S\,K^0_S$, where both the

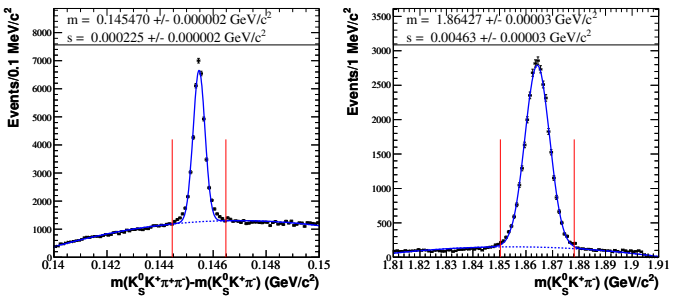


Figure 24: The fit of $\Delta m = m(K^+K^0_S\pi^+\pi^-) - m(K^0_S K^+\pi^-)$ is shown on the left. The results of this fit allowed to define a Δm signal region to plot and fit the $K^0_S\,K^+\,\pi^-$ invariant mass on the right. Both the fits are made to a model of a Gaussian over a second order polynomial background.

two K^0_S decay into $\pi^+\pi^-$ is observed. In Figure 25 the spectrum of $\pi^+\pi^-$ invariant mass is shown in signal region and into the sidebands. The events in the sidebands have been subtracted to the events in the signal region to obtain a clear $K^0_S \to \pi^+\pi^-$ signal peak. This peak has been fitted to a Gaussian over a flat function finding $\sigma = (2.94 \pm 0.02)\,\mathrm{MeV/c^2}$ and a mean $\mu = (494.43 \pm 0.07)\,\mathrm{MeV/c^2}$. The mean value of the K^0_S mass peak is again $3\,\mathrm{MeV/c^2}$ below to the K^0_S mass value reported by the Particle Data Group (PDG) [Nakamura *et al.*, 2010] $m(K^0_S) = (497.614 \pm 0.024)\,\mathrm{MeV/c^2}$: this is due to the constrain of the K^0_S daughters to originate from the D vertex. The contamination is completely removed together with all the events for which

$$488.55\,\mathrm{MeV/c^2} < m(\pi^+\pi^-) < 500.31\,\mathrm{MeV/c^2}.$$

Other sources of contamination have been found looking for wrongly identified particles. A small contaminations from $\Lambda_c^+ \to p\ K^0_S\ \pi^+\ \pi^-$ has been found by substituting the proton mass to the mass of the charged kaon in the final state. The contamination is shown in Figure 26, together with a fit of the Λ_c^+ peak. The number of Λ_c^+ events is estimated to be about 34,000, about 0.6% of the total. Their impact on the $K^+\ K^0_S\ \pi^+\ \pi^-$ spectrum will be studied further in the analysis.

5.4 OPTIMIZATION OF THE DATA SAMPLE

A logarithmic likelihood ratio (see Appendix B) has been defined in order to optimize separately the D^+ and D_s^+ peak in our data sample. The likelihood ratio is built once identified the variables that show the best separation between the signal and background distribution. The choice has been done following the prescription of a previous *BABAR* analysis [Aubert *et al.*, 2005], where three variables are considered:

- the momentum of the $D^+_{(s)}$ in the center-of-mass frame, $p^*(D^+_{(s)})$;
- the difference in the probability between the fits of the vertex without and with the constrain in the interaction region (IR),

$$\Delta p = \mathrm{prob_{fit}}(D^+_{(s)}) - \mathrm{prob^{IR}_{fit}}(D^+_{(s)});$$

- the distance of flight of the $D^+_{(s)}$ candidate in the (x, y) plane,

$$d_{xy} = \frac{\vec{p}_{xy}(D^+_{(s)}) \cdot \vec{v}_{xy}(D^+_{(s)})}{|\vec{p}_{xy}(D^+_{(s)})|}, \left(\vec{v} = \vec{D}_{vtx} - \vec{IP}\right).$$

$\vec{D}_{vtx}$ *and* $\vec{IP}$ *are the vectors pointing respectively to the fitted* D *vertex and to the Interaction Point in the BABAR coordinates system*

The distributions of these three variables are drawn in the sidebands subtracted signal region an in the sidebands. For any event in the data sample the likelihood ratio is evaluated as follows:

$$\log \mathcal{L} = \log \mathcal{P}^{sig}_{p^*} + \log \mathcal{P}^{sig}_{\Delta p} + \log \mathcal{P}^{sig}_{d_{xy}} - \log \mathcal{P}^{bkg}_{p^*} - \log \mathcal{P}^{bkg}_{\Delta p} - \log \mathcal{P}^{bkg}_{d_{xy}}, \tag{5.2}$$

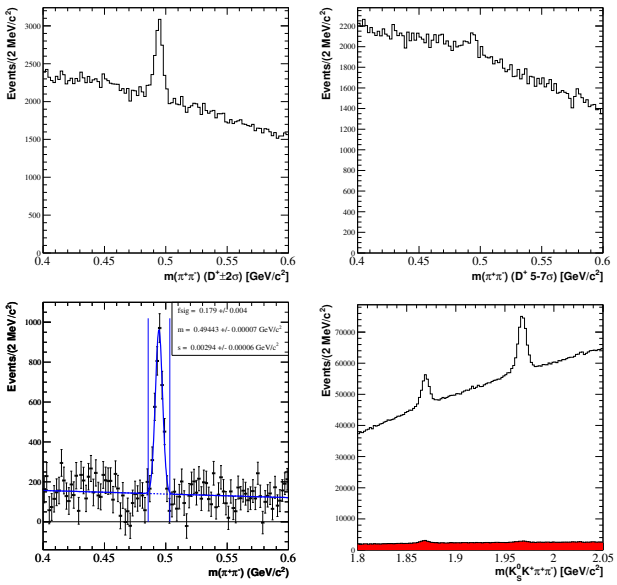


Figure 25: Study of $D^+ \to K^0_S\ K^0_S\ K^+$ contamination. On the top left the $m(\pi^+\pi^-)$ spectrum is shown in D^+ signal region; in the top right the same spectrum is shown in D^+ sidebands region; on the bottom left the subtraction of the two spectra has been fitted to a Gaussian over a flat function; on the bottom right the events removed applying the cut on $D^+ \to K^0_S\ K^0_S\ K^+$ are shown in the red filled area compared to the spectrum before the application of this cut.

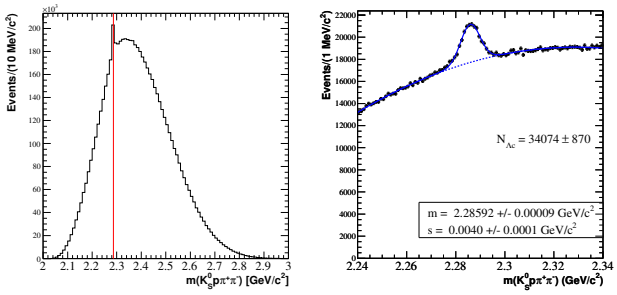


Figure 26: p K^0_S π^+ π^- mass spectrum is shown on the left. The nominal Λ_c^+ mass from PDG is indicated by a red line ($m(\Lambda_c^+)$=2286.46 MeV/c^2). The fit of the Λ_c^+ peak using a Gaussian over a second order polynomial background is shown on the right.

where the value of $\mathcal{P}^{sig(bkg)}_{var}$ is obtained from the distributions defined above.

In order to obtain the better definition for the signal and background distributions, they have been retrieved from the D^+ and D^+_s peaks that have the higher yield among all the $D^+_{(s)} \to K^0_S\, h^+h^+h^-$ decays. They are

- $D^+ \to K^0_S\, \pi^+\, \pi^+\, \pi^-$;
- $D^+_s \to K^-\, K^0_S\, \pi^+\, \pi^+$.

These decays have been reconstructed as the same as the $D^+_{(s)} \to K^+\, K^0_S\, \pi^+\, \pi^-$ and the distribution of the kinematic variables should be the same between the D^+ and D^+_s peaks reconstructed in each one of the different $D^+_{(s)} \to K^0_S\, h^+h^+h^-$ decays. Furthermore, the high statistics of these decay modes gives an excellent resolution for the signal and background distributions, which allows to obtain a continuos distribution for the likelihood ratio.

The signal and sidebands regions for D^+ and D^+_s spectra are shown in Figure 27 and are defined in Table 18. They have been retrieved

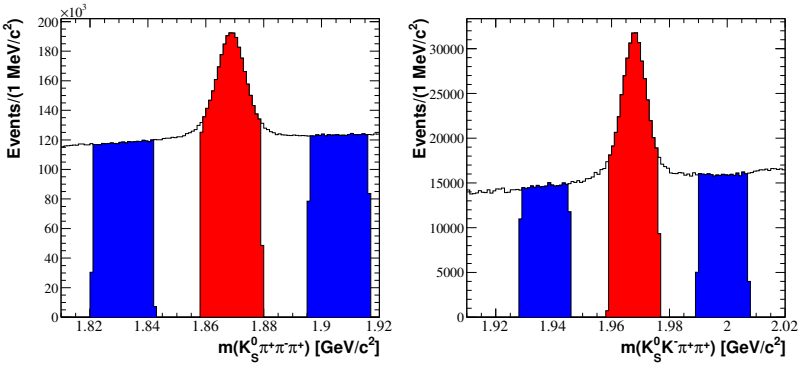


Figure 27: Signal (red) and sidebands (blue) regions for $D^+ \to K^0_S\, \pi^+\, \pi^+\, \pi^-$ (left) and $D^+_s \to K^-\, K^0_S\, \pi^+\, \pi^+$ mass spectra.

Table 18: Definition of signal and sidebands mass regions. The regions are shown in intervals of mass measured in MeV/c^2.

Decay Mode	Signal (MeV/c^2)	Sidebands (MeV/c^2)
$D^+ \to K^0_S\, \pi^+\, \pi^+\, \pi^-$	[1858.07,1879.35]	[1820.83,1842.11] [1895.31,1916.59]
$D^+ \to K^-\, K^0_S\, \pi^+\, \pi^+$	[1959.0,1976.6]	[1928.2,1945.8] [1989.8,2007.4]

from a fit to the mass spectrum using a model made of a Gaussian over a second order polynomial: the signal region is defined as the mass region within 2 standard deviations from the mean value of the Gaussian, while the sidebands regions are made of the events with invariant mass between 5 and 9 standard deviations from the Gaussian's mean. The sidebands region has double width with respect to

the signal region: this choice has been made in order to guarantee the continuity of the likelihood ratio distribution. If an event is found out of the phase space defined by the signal and sidebands distributions outlined in this section, given the definition of the likelihood ratio, it would not be possible to evaluate the value of the likelihood ratio itself for that event because it would produce a divergence in the logarithm. Two solutions have been found:

1. enlarge the sidebands regions to include as many events as possible to define the distributions;
2. set a standard value of the likelihood ratio whenever the variables are out of the range of the distributions.

Both the two solutions have been applied in this analysis. Since the effect of the second solution is to create regions of the likelihood ratio that are separated from the main peak, the application of the first significantly reduces the number of events out of range and, consequently, aggregates all the events into a continuous distribution.

The distributions retrieved from the signal and sidebands regions previously defined are shown in Figure 28. The distribution of the

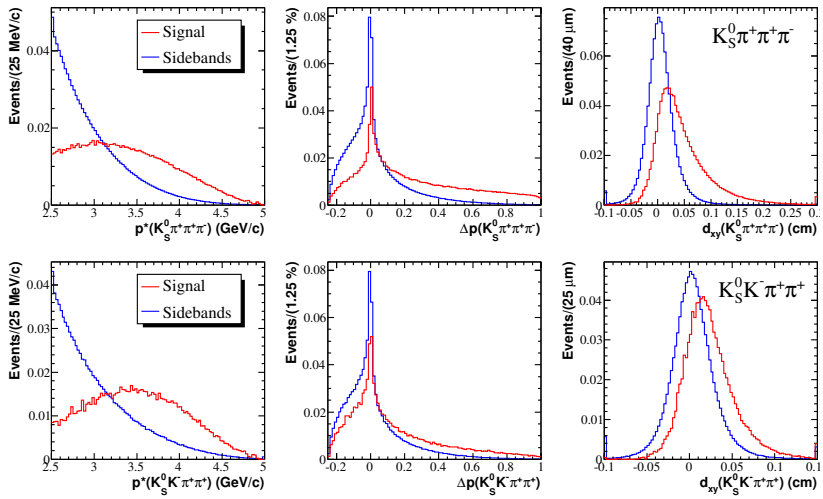


Figure 28: Signal (red) and sidebands (blue) distributions for the D^+ (top) and D_s^+ (bottom) peaks. From left to right $p^*(D^+_{(s)})$, Δp and d_{xy}.

likelihood ratios defined by Equation 5.2 and obtained from the events in the $D^+_{(s)} \to K^0_S\ K^+\ \pi^+\ \pi^-$ data sample using the distributions in Figure 28 are shown in Figure 29.

The best cut to apply on Likelihood ratio has been determined studying the significance of $D^+_{(s)}$ peak using many different cuts. The significance of the peak is defined as

$$\mathcal{S} = \frac{S}{\sqrt{S+B}}, \tag{5.3}$$

where S is the number of signal events, obtained subtracting the events in the sidebands from the ones in the signal region; B is the number of

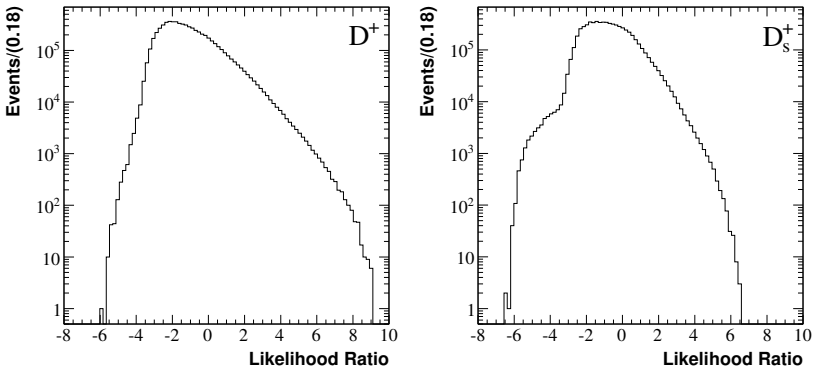


Figure 29: Plot of the distribution of D^+ (left) and D_s^+ (right) likelihood ratio evaluated on $D^+_{(s)} \to K^0_S\ K^+\ \pi^+\ \pi^-$ data sample.

background events, evaluated as half the events in the sidebands (considering that the sidebands region is two times wide respect to signal region). The cut for which the best significance is found is regarded as the best to optimize the data sample. Following the results shown in Figure 30, $\mathcal{L}_{D^+} > 1.5$ is used to optimize D^+ peak, while $\mathcal{L}_{D_s^+} > 0.0$ is used for D_s^+, obtaining the spectra shown in Figure 31.

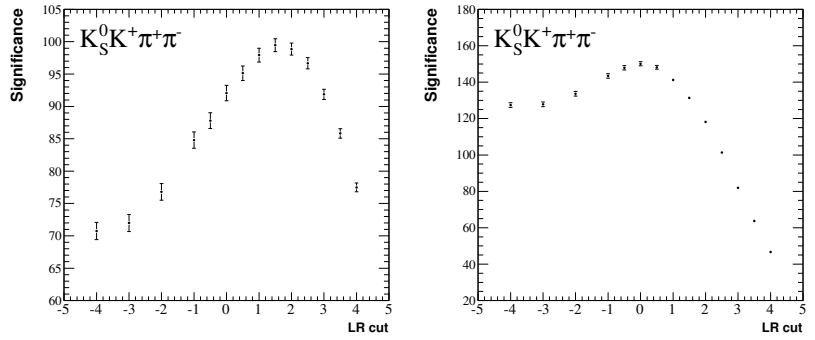


Figure 30: Variation of $D^+_{(s)}$ peak significance depending on the selection criteria applied on D^+ (left) of D_s^+ (right) likelihood ratio.

5.4.1 Potential Bias from Likelihood Ratio Selection

The potential bias introduced by the selection made on likelihood ratio has been studied considering three data sets:

- data sidebands;
- generic Monte Carlo sample without $D^+_{(s)}$ signal events;
- $D^+_{(s)}$ signal Monte Carlo samples generated by phase space.

Many tests have been performed on these data sets:

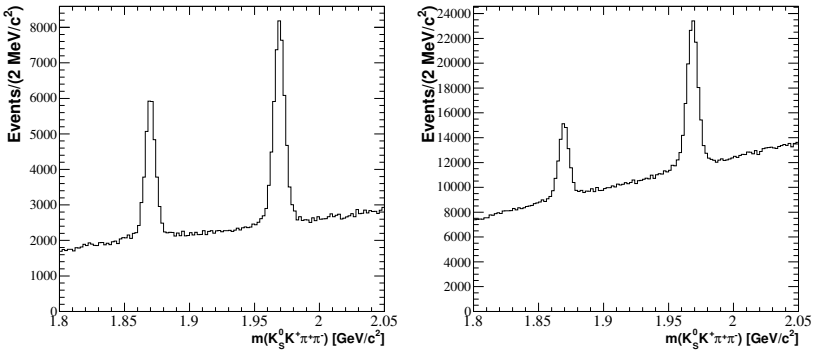


Figure 31: $K^+ K^0_S \pi^+ \pi^-$ mass spectrum optimized using D^+ (left) and D^+_s (right) likelihood ratios.

1. The likelihood ratio output has been compared for the three samples described above for D^+ (D^-) when C_T ($\bar{C}_T$) is greater or less than zero (Figure 32). All the distributions show a difference in χ^2 about equal to 1, demonstrating that they are consistent.

2. The $K^+ K^0_S \pi^+ \pi^-$ mass spectrum is reconstructed from generic Monte Carlo for D^+ when $C_T > 0$ and $C_T < 0$. The same has been done for D^- using $\bar{C}_T$. The results are shown in Figure 33.

3. All the selections developed on the data sample have been applied to signal Monte Carlo, then $\mathcal{A}_T$ has been evaluated by counting the $D^+_{(s)}$ ($D^-_{(s)}$) events with $C_T(\bar{C}_T) > 0$ and $C_T(\bar{C}_T) < 0$. The results are shown in Figure 34 for many cuts on likelihood ratio. The deviations observed between generated and reconstructed $\mathcal{A}_T$ (before the selection on likelihood ratio) have been regarded as a statistical effect.

For some of the previously mentioned tests a χ^2 has been evaluated to determine whether a bias is found or not:

$$\chi^2 = \sum_i \frac{N_{1,i} - N_{2,i}}{\sigma^2_{1,i} + \sigma^2_{2,i}} \qquad i = \text{bin}, \tag{5.4}$$

where $N_{j,i}$ and $\sigma_{j,i}$ are the number of events and its relative error in bin i of histogram j.

5.5 C_T RESOLUTION

The distribution of C_T and $\bar{C}_T$ for both D^+ and D^+_s sideband subtracted signal region is shown in Figure 35. This test, similar to the one made on D^0 signal sample and shown in Figure 9, shows the effects of final state interaction. Also for D^+ and D^+_s, it can be observed an asymmetry in the C_T distribution from which a non-zero value of A_T and $\bar{A}_T$ can be expected.

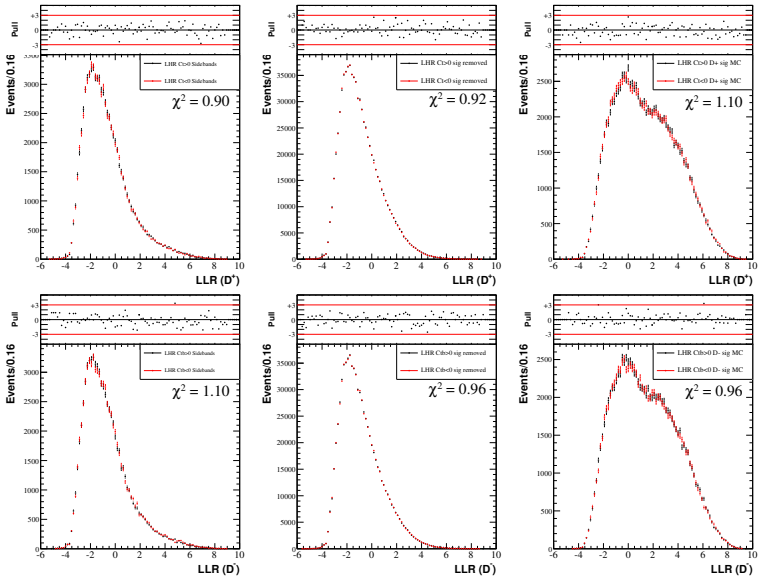


Figure 32: On top the plots of likelihood ratio for D^+ when C_T is greater (black) or less (red) than zero. On the bottom the same plot made on D^- depending on $\bar{C}_T$.

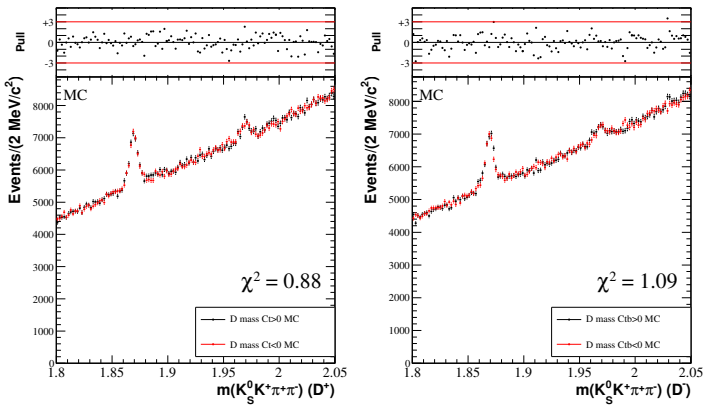


Figure 33: The K^0_S K^+ π^+ π^- mass spectrum is compared when $C_T > 0$ (black) or $C_T < 0$ (red) using the generic Monte Carlo sample. On the left is shown the comparison between D^+ candidates, on the right between D^- candidates. Both the D candidates distributions are consistent between $C_T(\bar{C}_T) > 0$ and $C_T(\bar{C}_T) < 0$.

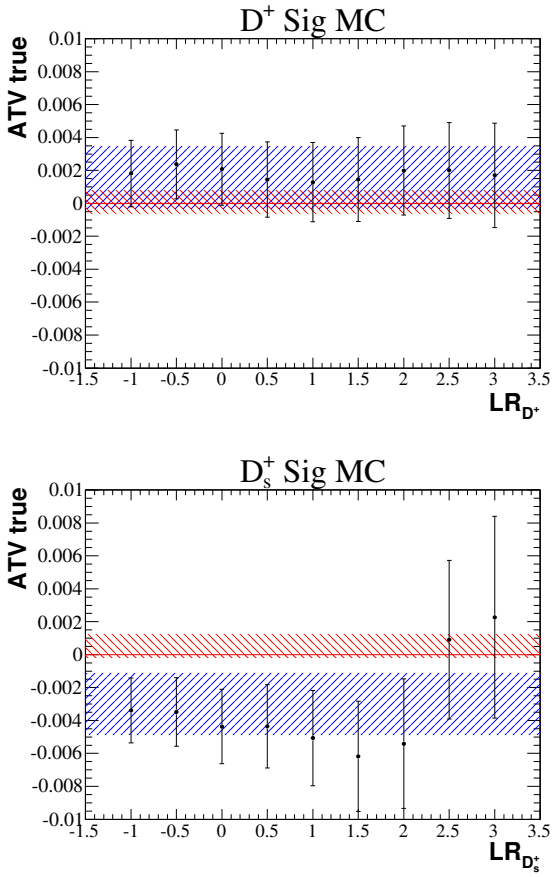


Figure 34: The $\mathcal{A}_T$ observable measured on signal Monte Carlo after the application of all the selections using different cuts on likelihood ratio. The red shaded area shows the $\mathcal{A}_T$ value and error retrieved from generated Monte Carlo. The blue shaded area shows the $\mathcal{A}_T$ value and error after events reconstruction and before the cut on likelihood ratio.

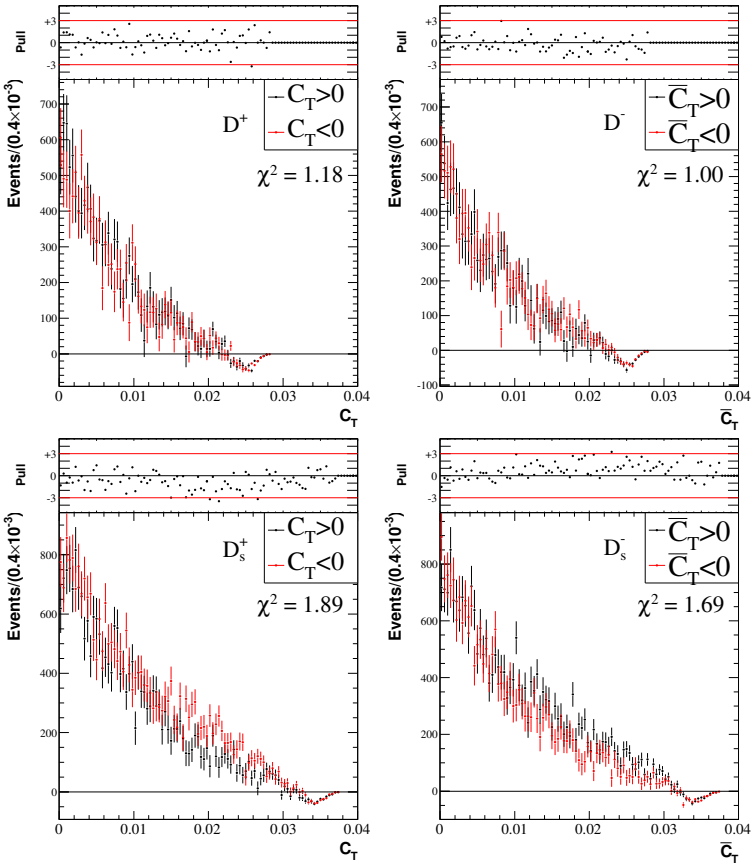


Figure 35: The distribution of the events in the signal region with the sidebands subtracted with C_T ($\bar{C}_T$) greater or less than zero is compared on top left for D^+ (D^-- right) and on the bottom left for D_s^+ (D_s^-- right) data sample.

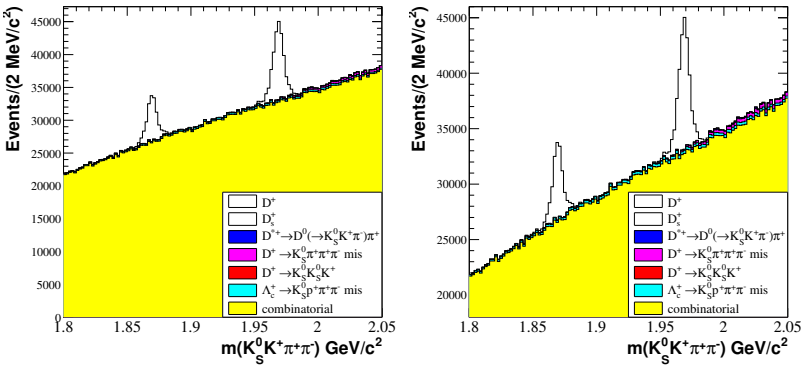


Figure 36: Composition of the $K^+\ K^0_S\ \pi^+\ \pi^-$ mass spectrum after reconstruction. On the right, the same plot that is on the left is shown limiting the range of the y-axis to [18,000:46,000]. The highlight allows to show the events categories in detail.

5.6 DEFINITION OF THE FIT MODEL

Since the $\mathcal{A}_T$ measurement is made separately on D^+ and D^+_s signal events, the $K^+\ K^0_S\ \pi^+\ \pi^-$ mass spectrum has been split in two regions:

$$\begin{aligned} 1.81 < m(K^+K^0_S\pi^+\pi^-) < 1.92\,\text{GeV}/c^2 \\ 1.91 < m(K^+K^0_S\pi^+\pi^-) < 2.02\,\text{GeV}/c^2, \end{aligned} \tag{5.5}$$

the former mass region to measure $\mathcal{A}_T$ on D^+, the latter on D^+_s.

The model to be used to describe the mass spectrum has been developed on the generic $e^+e^- \to c\bar{c}$ Monte Carlo.

5.6.1 Study of the Events Categories

Using the informations from the generator, the correctly reconstructed Monte Carlo allowed to identify the following categories of events in this analysis:

- signal, made of events with all the final state particles correctly identified, coming from a $D^+_{(s)}$ decay.
- $D^+ \to K^0_S\ \pi^+\ \pi^+\ \pi^-$ reflection, where the π^+ has been identified as a K^+. This category of events manifests itself into a bump for masses higher than $1.94\,\text{GeV}/c^2$.
- $\Lambda^+_c \to p\ K^0_S\ \pi^+\ \pi^-$ reflection where the proton has been identified as a K^+. These events are widely distributed on the spectrum as shown in Figure 36.
- combinatorial background;

Applying the selection on likelihood ratio a strong suppression of $\Lambda^+_c \to p\ K^0_S\ \pi^+\ \pi^-$ reflection and combinatorial background has been observed, as shown in Figure 37. In the same figure can be observed that the $D^+ \to K^0_S\ \pi^+\ \pi^+\ \pi^-$ reflection seems to be enhanced respect

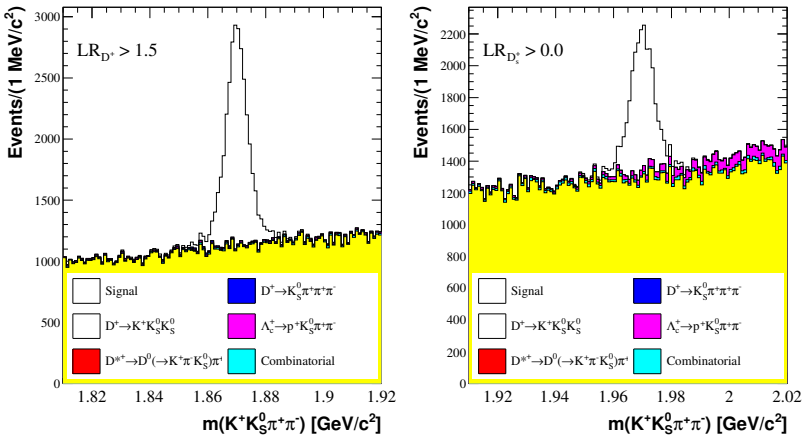


Figure 37: Composition of the $K^+ K^0_S \pi^+ \pi^-$ mass spectrum after the selection on likelihood ratio. The plot on the left shows the D^+ mass range; the one on the right shows the D^+_s range.

to the background: this effect is due to its kinematic topology, equivalent to the D^+ signal events.

5.6.2 Definition of the Shapes

At first, a shape for the background description has been searched for. A line, a second order and a third order polynomial have been fitted to the generic Monte Carlo sample once removed the correctly reconstructed signal events. The results for both the two mass regions are shown in Figure 38.

As a preliminary approach a line is assumed for the description of the background distribution for both the two ranges of mass. Using a model of two Gaussians with common mean to describe the signal distribution, the number of signal events retrieved from the fit has been compared to the number of correctly reconstructed signal events in order to validate the fit model. The results of the fit to the Monte Carlo distribution are shown in Figure 39.

The model is validated comparing the retrieved number of signal events to their true number, as shown in Table 19. While the results

Table 19: Comparison between the number of events retrieved by the fit and by Monte Carlo truth. The last column shows the number of σ's between true and fitted number of events: $\varepsilon = \frac{n - n_{true}}{\sigma_n}$

Category	Events	True	ε
D^+ signal	17068 ± 281	16848	+0.78
D^+ background	104998 ± 408	105219	-0.54
D^+_s signal	37615 ± 396	39679	-5.22
D^+_s background	426189 ± 738	424125	+2.80

for the D^+ mass range are quite consistent within the errors, the fit to

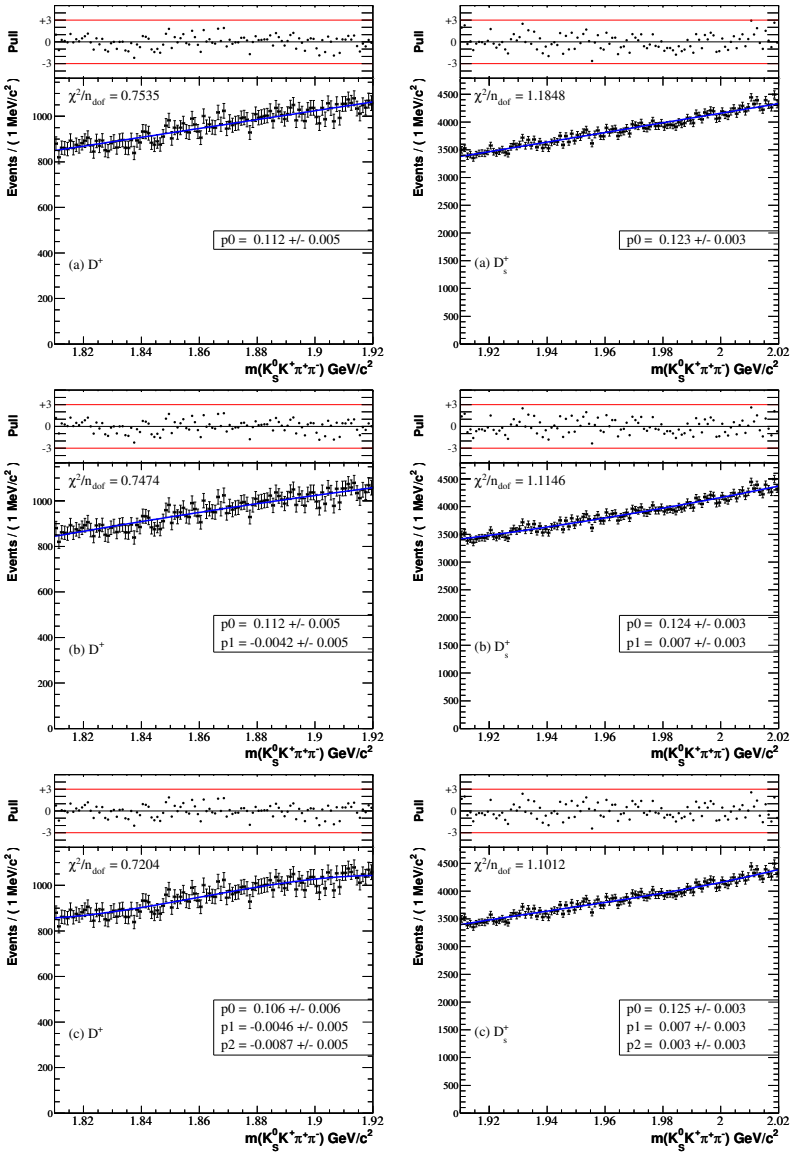


Figure 38: Fit of the background categories using three different polynomial shapes: a line (a), a second order (b), a third order (c). On left the D^+ mass range is shown, on the right the one of D_s^+. The χ^2 reduced to the number of degrees of freedom is shown on top left of each plot. The normalized residuals shown on top of each plot are evaluated as outlined in Equation 4.1.

the D_s^+ mass range returns some values that are not consistent to the Monte Carlo.

In order to retrieve the correct number of events from the fit, some variations to the model are applied. A second order polynomial for the background significantly improves the quality of the fit, but is not enough: the cut on the likelihood ratio has been tightened to $\mathcal{L}_{D_s^+} > 1.5$ in order to obtain a number of fitted signal events consistent to the true value from Monte Carlo. The results of this fit are shown in Figure 40 and the relative difference in the number of D_s^+ signal events is shown in Table 20.

Table 20: Comparison between the number of events retrieved by the fit and by Monte Carlo truth. The last column shows the number of σ's between true and fitted number of events: $\varepsilon = \frac{n - n_{true}}{\sigma_n}$

Category	Events	True	ε
D_s^+ signal	17599 ± 324	18063	-1.43
D_s^+ background	99911 ± 423	99447	+1.10

In Table 21 the results for each parameter in the fit are shown both for the D^+ and D_s^+ fit models. It can be noticed that the width of the

Table 21: Results of the fits to the Monte Carlo sample requiring $\mathcal{L}_{D^+} > 1.50$ and $\mathcal{L}_{D_s^+} > 1.50$.

(a) $D^+ \to K^0_S\ K^+\ \pi^+\ \pi^-$

Parameter	Final Value
N(D^+)	17068 ± 281
N(bkg)	104998 ± 408
mean D^+	$(1870.0 \pm 0.1)\,\mathrm{MeV/c^2}$
σ_1	$(3.30 \pm 0.11)\,\mathrm{MeV/c^2}$
σ_2/σ_1	(2.69 ± 0.33)
f_{G_1}	$(75.34 \pm 4.06)\%$
p_0	$(1.09 \pm 0.05) \times 10^{-1}$

(b) $D_s^+ \to K^0_S\ K^+\ \pi^+\ \pi^-$

Parameter	Final Value
N(D_s^+)	17598 ± 296
N(bkg)	426189 ± 738
mean D_s^+	$(1969.8 \pm 0.1)\,\mathrm{MeV/c^2}$
σ_1	$(3.47 \pm 0.27)\,\mathrm{MeV/c^2}$
σ_2/σ_1	(1.78 ± 0.24)
f_{G_1}	$(65.21 \pm 16.50)\%$
p_0	$(1.43 \pm 0.06) \times 10^{-1}$
p_1	$(1.83 \pm 0.03) \times 10^{-2}$

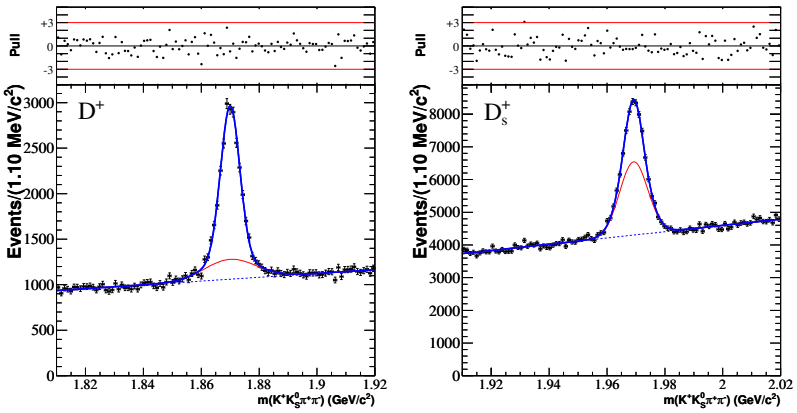


Figure 39: Model made of two gaussians and a linear background fitted to the Monte Carlo sample in the D^+ (left) and D_s^+ (right) mass region. The projection of the model on the data is shown by the blue line; the second Gaussian in the signal shape is shown in red, the background shape is shown by a dashed blue line. On top of the plots, the normalized residuals (Pull) distribution is shown.

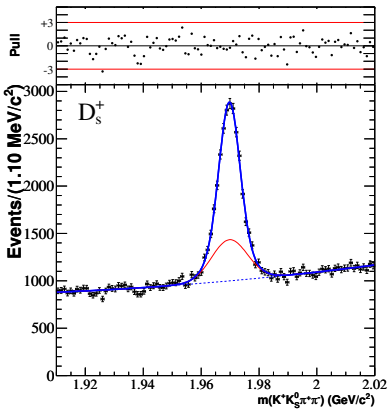


Figure 40: Model made of two gaussians and a second order polynomial background fitted to the Monte Carlo sample in the D_s^+ mass region. The projection of the model on the data is shown by the blue line; the second Gaussian in the signal shape is shown in red, the background shape is shown by a dashed blue line. On top of the plots, the normalized residuals (Pull) distribution is shown.

second Gaussian is expressed in multiples of the width of the main Gaussian in the model.

5.7 DATA ANALYSIS AND SYSTEMATICS

5.7.1 Fit procedure

The model developed in the previous section has been used to perform an unbinned maximum likelihood simultaneous fit to the four data samples obtained dividing the data set depending on D charge and C_T value. Each histogram is fitted simultaneously to the same model and the yield of each peak is evaluated separately in the same fit. Given the relationship between the following sets of observables

$$\begin{bmatrix} N(D^+_{(s)}, C_T > 0) & N(D^+_{(s)}, C_T < 0) \\ N(D^-_{(s)}, \bar{C}_T > 0) & N(D^-_{(s)}, \bar{C}_T < 0) \end{bmatrix}$$
$$\updownarrow$$
$$\left\{ N(D^+_{(s)}), N(D^-_{(s)}), A_T, \bar{A}_T \right\},$$

the signal yields for each sub sample can be expressed as:

$$\begin{aligned} N(D^+_{(s)}, C_T > 0) &= \frac{N(D^+)}{2}(1 + A_T) \\ N(D^+_{(s)}, C_T < 0) &= \frac{N(D^+)}{2}(1 - A_T) \\ N(D^-_{(s)}, \bar{C}_T > 0) &= \frac{N(D^-)}{2}(1 - \bar{A}_T) \\ N(D^-_{(s)}, \bar{C}_T < 0) &= \frac{N(D^-)}{2}(1 + \bar{A}_T). \end{aligned}$$

The latter set of variables has been used in the fit: this way $\mathcal{A}_T$ can be calculated directly from fit results as

$$\begin{aligned} \mathcal{A}_T &= \frac{1}{2}(A_T - \bar{A}_T) \\ \sigma_{\mathcal{A}_T} &= \frac{1}{2}\sqrt{\sigma^2_{A_T} + \sigma^2_{\bar{A}_T}}. \end{aligned}$$

The fit to the D^+ and D^+_s data samples are shown in Fig. 41. Similarly to what has been done for the D^0 analysis (see Section 4.6.1), these fits are "blind", because the central values of asymmetry parameters are masked by means of the addition of an unknown random offset. The quality of the fit is excellent.

5.7.2 Fit Validation on the Generic Monte Carlo sample

The same procedure has been applied to fit the generic Monte Carlo sample. This is done to check some of the results that have been obtained from the data sample. In particular, it can be observed from Table 22 that the asymmetries A_T and $\bar{A}_T$ in the background events of the

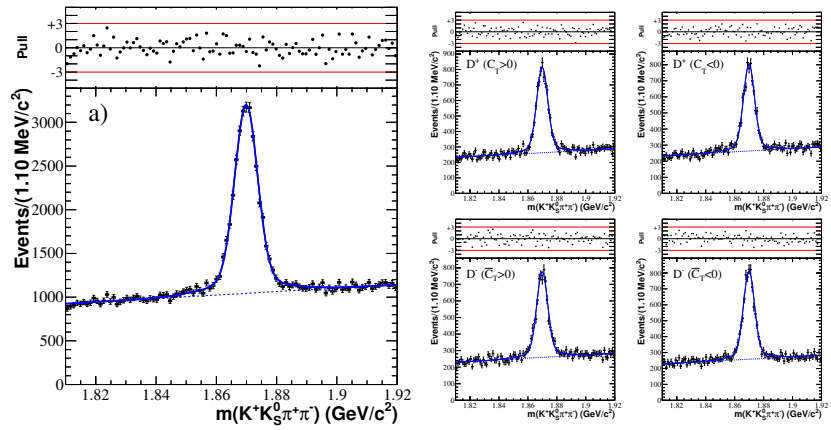
Table 22: Results of the fits to the data sample requiring $\mathcal{L}_{D^+_{(s)}} > 1.50$ (A_T and $\bar{A}_T$ are blind for both D^+ and D^+_s peaks).

(a) $D^+ \to K^0_S\, K^+\, \pi^+\, \pi^-$

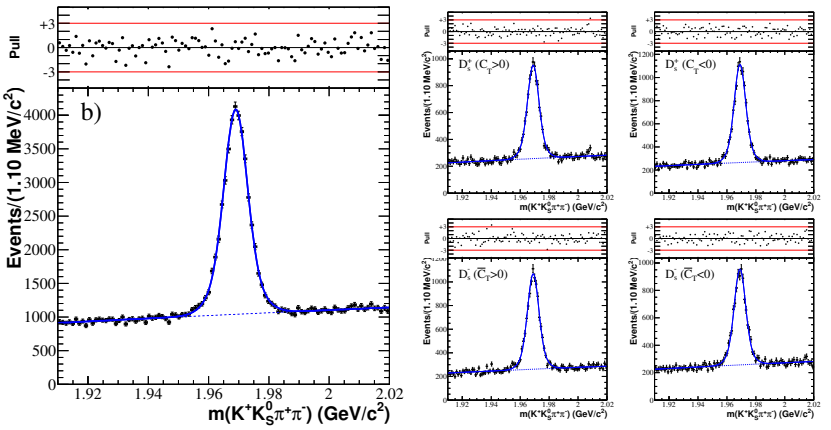
Floating Parameter	Final Value
A_T D^+	$(x.xx \pm 1.41) \times 10^{-2}$
A_T bkg	$(-1.71 \pm 4.86) \times 10^{-3}$
$\bar{A}_T$ D^+	$(x.xx \pm 1.42) \times 10^{-2}$
$\bar{A}_T$ bkg	$(-2.48 \pm 4.90) \times 10^{-3}$
$N(D^+)$	10701 ± 257
$N(D^+)$ bkg	52118 ± 328
$N(D^-)$	10526 ± 254
$N(D^-)$ bkg	51245 ± 324
mean D^+	$(1869.9 \pm 0.1)\,\mathrm{MeV/c^2}$
σ_1	$(3.75 \pm 0.07)\,\mathrm{MeV/c^2}$
σ_2/σ_1	3.40 ± 0.47
f_{G_1}	$(81.50 \pm 2.32)\%$
p_0	$(1.05 \pm 0.05) \times 10^{-1}$

(b) $D^+_s \to K^0_S\, K^+\, \pi^+\, \pi^-$

Floating Parameter	Final Value
A_T D^+_s	$(x.xx \pm 1.34) \times 10^{-2}$
A_T bkg	$(-14.54 \pm 4.85) \times 10^{-3}$
$\bar{A}_T$ D^+_s	$(x.xx \pm 1.09) \times 10^{-2}$
$\bar{A}_T$ bkg	$(-9.33 \pm 4.87) \times 10^{-3}$
$N(D^+_s)$	15083 ± 233
$N(D^+_s)$ bkg	51732 ± 301
$N(D^-_s)$	14717 ± 232
$N(D^-_s)$ bkg	51315 ± 300
mean D^+_s	$(1969.0 \pm 0.1)\,\mathrm{MeV/c^2}$
σ_1	$(3.65 \pm 0.20)\,\mathrm{MeV/c^2}$
σ_2/σ_1	$1.76 \pm 0.15\,\mathrm{MeV/c^2}$
f_{G_1}	$(65.20 \pm 12.00)\%$
p_0	$(1.12 \pm 0.05) \times 10^{-1}$
p_1	$(-2.31 \pm 7.65) \times 10^{-3}$



(a) $D^+ \to K^0_S\ K^+\ \pi^+\ \pi^-$



(b) $D^+_s \to K^0_S\ K^+\ \pi^+\ \pi^-$

Figure 41: Fit to the $D^+_{(s)} \to K^0_S\ K^+\ \pi^+\ \pi^-$ data sample. The fit has been performed using a model made of a double gaussian over a linear (second order polynomial) background. On the left the fit to the overall sample is shown; on the right the projections on the four sub-samples are shown. For each plot, the distribution of the normalized residuals (Pull) is shown on the top.

D_s^+ peak are significantly different from zero, while the asymmetries on D^+ are not. This effect is not present in the fit to the generic Monte Carlo sample, shown in Tab. 23. However, the *BABAR* events generator is not well tuned to final state interaction effects. It has indeed been observed in the D^0 analysis that the values of A_T and $\bar{A}_T$ found in the data sample are completely different from the ones retrieved from the fit to the Monte Carlo. From Table 23, it can also be observed that the asymmetries are mainly consistent to zero, suggesting the things to be similar in this analysis.

5.7.3 Systematics

Many checks have been applied to evaluate the systematic errors. The sources of systematic error identified are:

- events selection;
- forward-backward asymmetry;
- likelihood ratio selection;
- fit model;
- particle identification criteria.

For each source of systematic error identified a new fit on blind data will be done little modifying the variables that are linked to that systematic. The higher deviation from $\mathcal{A}_T$ will be regarded as the error due to that source of systematics.

To test the systematic effect of each source listed above, they have been studied as follows.

5.7.4 Events selection and Forward-Backward asymmetry

The signal Monte Carlo data set has been used to study the effects of events reconstruction on these measurements. Considering that these events have been generated by phase space, while resonances are present in the real data, these effects should be somewhat included: to do so, the detector reconstruction efficiency, the resonant structure and the forward-backward asymmetry need to be included in the Monte Carlo generated by phase space.

The reconstruction efficiency has been measured comparing the number of correctly reconstructed signal events and the generated number of signal events for each bin of the cosine of the polar angle $\cos\theta^*$, as shown in Fig. 42.

All the events in the generated Monte Carlo sample have been weighted using this efficiency plot in order to obtain a simulation of the events reconstructed by the *BABAR* detector.

The resonant structure has been included in this simulated data set by means of relativistic Breit-Wigner distributions whose parameters have been retrieved from a fit to the data sample. The $K_S^0\,\pi^+$, $K_S^0\,\pi^-$, $K^+\,\pi^-$ and $\pi^+\pi^-$ invariant mass spectra have been drawn into the signal region of the data sample. These spectra, shown together with

Table 23: Results of the fits to the generic Monte Carlo sample requiring $\mathcal{L}_{D^+_{(s)}} > 1.50$.

(a) $D^+ \to K^0_S\ K^+\ \pi^+\ \pi^-$

Floating Parameter	Final Value
A_T D^+	$(0.61 \pm 1.61) \times 10^{-2}$
A_T bkg	$(3.40 \pm 4.76) \times 10^{-3}$
$\bar{A}_T$ D^+	$(-2.98 \pm 1.63) \times 10^{-2}$
$\bar{A}_T$ bkg	$(1.04 \pm 4.81) \times 10^{-3}$
$N(D^+)$	8631 ± 205
$N(D^+)$ bkg	53024 ± 294
$N(D^-)$	8438 ± 198
$N(D^-)$ bkg	51973 ± 288
mean D^+	$(1870.0 \pm 0.1)\,MeV/c^2$
σ_1	$(3.30 \pm 0.11)\,MeV/c^2$
σ_2/σ_1	2.69 ± 0.35
f_{G_1}	$(74.21 \pm 4.39)\%$
p_0	$(1.09 \pm 0.05) \times 10^{-1}$

(b) $D^+_s \to K^0_S\ K^+\ \pi^+\ \pi^-$

Floating Parameter	Final Value
A_T D^+_s	$(-0.95 \pm 1.54) \times 10^{-2}$
A_T bkg	$(5.11 \pm 4.86) \times 10^{-3}$
$\bar{A}_T$ D^+_s	$(-0.46 \pm 1.53) \times 10^{-2}$
$\bar{A}_T$ bkg	$(-0.17 \pm 4.91) \times 10^{-3}$
$N(D^+_s)$	8770 ± 228
$N(D^+_s)$ bkg	50531 ± 306
$N(D^-_s)$	8829 ± 229
$N(D^-_s)$ bkg	49380 ± 305
mean D^+_s	$(1969.8 \pm 0.1)\,MeV/c^2$
σ_1	$(3.47 \pm 0.35)\,MeV/c^2$
σ_2/σ_1	$1.78 \pm 0.31\,MeV/c^2$
f_{G_1}	$(65.18 \pm 21.90)\%$
p_0	$(1.43 \pm 0.06) \times 10^{-1}$
p_1	$(1.83 \pm 0.80) \times 10^{-2}$

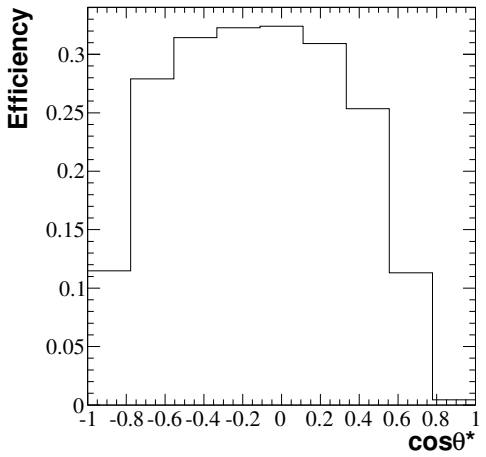


Figure 42: Reconstruction efficiency in bins of the cosine of the $c\bar{c}$ production angle.

the fits in Fig. 43, indicate the presence of the resonances $K^{*-} \to K^0_S \pi^-$, $K^{*0} \to K^+ \pi^-$ and $\rho \to \pi^+\pi^-$. The results of the fits are shown in Tab. 24 and are reasonably according to the PDG values [Nakamura *et al.*, 2010]. The structure observed in the $K^0_S \pi^+$ invariant mass has not been considered significant.

Table 24: Results of the fits shown in Fig. 43. All the resonances are fitted to a relativistic Breit-Wigner shape. The resulting masses and widths are shown in MeV/c^2.

Resonance	Mass	Width
$K^{*-} \to K^0_S \pi^-$	886.74	55.50
$K^{*0} \to K^+ \pi^-$	894.80	42.64
$\rho \to \pi^+\pi^-$	730.65	144.41

The Monte Carlo events have been weighted making use of a function that depends on $m(K^0_S\pi^-)$, $m(K^+\pi^-)$ and $m(\pi^+\pi^-)$. This function is made of normalized relativistic Breit-Wigner functions having the mass and widths shown in Table 24.

Finally the effect of the forward-backward asymmetry has been included into this data set. Forward-backward asymmetry is an effect due to the interference between the electromagnetic and the weak currents in the $e^+e^- \to c\bar{c}$ process. A detailed study of this effect can be found in Appendix A.

Given the boost of the e^+e^- Center of Mass System (CMS) in PEP-II beam and the construction asymmetry of *BABAR* detector, the net effect of this interference is that the ratio $c/\bar{c}$ is not constant over the cosine of the polar angle θ^*, the angle formed by the c quark momentum and the e^- momentum.

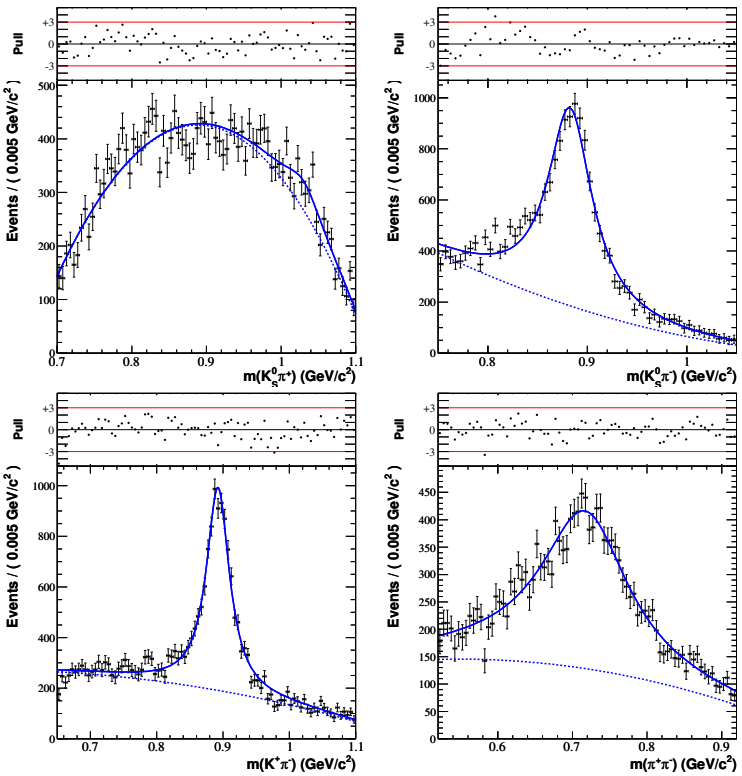


Figure 43: Fits to the two-body invariant mass distributions for D^+ signal events. From left to right, top to down: $K^0_S\ \pi^+$, $K^{*-} \to K^0_S\ \pi^-$, $K^{*0} \to K^+\ \pi^-$ and $\rho \to \pi^+\pi^-$.

This effect has been included weighting the simulated events by the function (shown in Fig. 44):

$$w_{FB}(\cos\theta^*) = 1 + \frac{8}{3} a_{FB} \frac{\cos\theta^*}{1+\cos^2\theta^*}, \tag{5.6}$$

where $a_{FB} = -2.49\%$ has been retrieved, together with the function, from an internal note of the *BABAR* Collaboration [Chen *et al.*, 2010].

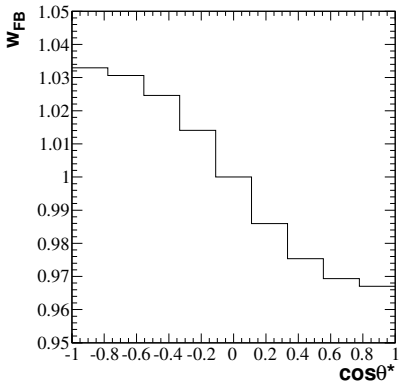


Figure 44: Distribution of the weights used to simulate the forward-backward asymmetry effect.

The asymmetry parameters are then compared in the plots shown in Fig. 45. In these plots, the original value of the asymmetry with

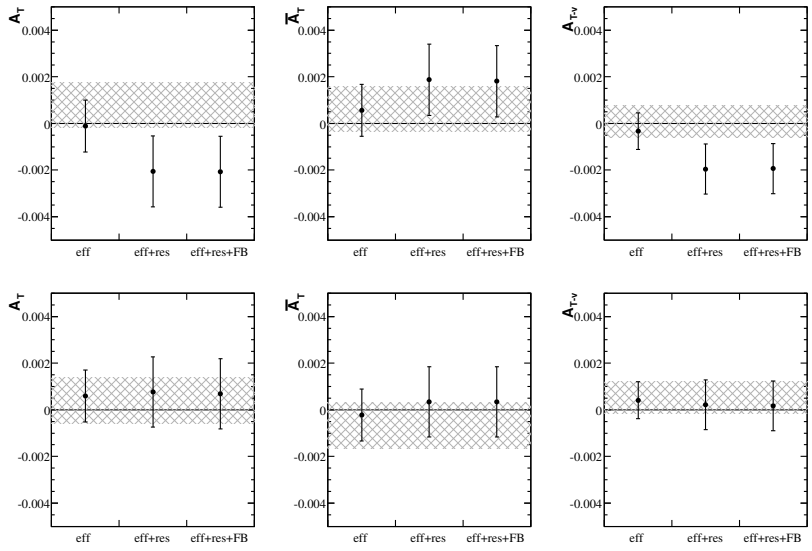


Figure 45: The asymmetry parameters A_T (left), $\bar{A}_T$ (center) and $\mathcal{A}_T$ (right) are compared with their value before the simulation (grey shaded area) and after the introduction of the efficiency, the resonant structure and the forward-backward asymmetry. On top, the results on D^+ peaks, in the bottom, the D_s^+.

its error is represented by the grey shaded area, while the value ob-

tained after each step of the weighting described into this section is represented with its error by the graph points.

Some considerations can be retrieved from these plots:

- the reconstruction efficiency and the resonant structures introduce a bias of about 3×10^{-3} into the A_T, that reflects into $\mathcal{A}_T$, just in the D^+ sample;
- the introduction of a weighting factor depending from the forward-backward asymmetry does not affect much the asymmetry parameters.

Following these considerations, a systematics error should be included related to the reconstruction bias in the D^+ asymmetries, that is the deviation from the generated value:

$$\sigma^{reco}_{syst}(A_T, D^+) = -2.84 \times 10^{-3},$$
$$\sigma^{reco}_{syst}(\bar{A}_T, D^+) = +1.26 \times 10^{-3},$$
$$\sigma^{reco}_{syst}(\mathcal{A}_T, D^+) = -2.05 \times 10^{-3}.$$

Since all the deviations in the D^+_s measurements are within the statistical fluctuations, a minimum bias of 1×10^{-3} has been assumed for each asymmetry parameter:

$$\sigma^{reco}_{syst}(A_T, D^+_s) = +1.00 \times 10^{-3},$$
$$\sigma^{reco}_{syst}(\bar{A}_T, D^+_s) = +1.27 \times 10^{-3},$$
$$\sigma^{reco}_{syst}(\mathcal{A}_T, D^+_s) = +1.00 \times 10^{-3}.$$

5.7.5 Likelihood ratio selection

Tighter and looser cuts on likelihood ratio have been applied and the difference of $\mathcal{A}_T$ respect to the optimal condition ($\mathcal{L}_{D^+_{(s)}} > 1.5$) has been measured. The result of this study is shown in Tab. 25.

Table 25: Statistical error (σ) on $\mathcal{A}_T$, A_T and $\bar{A}_T$ and their deviation (Δ) from the best $LR_{D^+_{(s)}}$ cut. All the values are multiples of $\times 10^{-3}$.

	LR cut	$\sigma_{\mathcal{A}_T}$	$\Delta_{\mathcal{A}_T}$	σ_{A_T}	Δ_{A_T}	$\sigma_{\bar{A}_T}$	$\Delta_{\bar{A}_T}$
D^+	1.30	10.09	-0.80	14.21	+1.34	14.32	+2.95
	1.50	10.04	—	14.13	—	14.26	—
	1.70	10.04	+1.08	14.16	-3.41	14.24	-5.58
D^+_s	1.30	7.44	+2.46	10.43	+0.96	10.62	-3.95
	1.50	7.67	—	10.73	—	10.93	—
	1.70	7.95	+0.21	11.13	-7.77	11.36	-8.16

Assuming the larger deviation as the systematic error to be assigned, one gets

$$\sigma^{\mathcal{L}}_{syst}(\mathcal{A}_T[D^+]) = \pm 1.10 \times 10^{-3},$$
$$\sigma^{\mathcal{L}}_{syst}(\mathcal{A}_T[D^+_s]) = \pm 2.41 \times 10^{-3}.$$

The same procedure applied to A_T and $\bar{A}_T$ gives

$$\sigma^{\mathcal{L}}_{syst}(A_T[D^+]) = \pm 3.41 \times 10^{-3},$$
$$\sigma^{\mathcal{L}}_{syst}(\bar{A}_T[D^+]) = \pm 5.58 \times 10^{-3},$$
$$\sigma^{\mathcal{L}}_{syst}(A_T[D_s^+]) = \pm 7.77 \times 10^{-3},$$
$$\sigma^{\mathcal{L}}_{syst}(\bar{A}_T[D_s^+]) = \pm 8.16 \times 10^{-3}.$$

5.7.6 Fit model

The fit model has been changed in many ways and the difference of $\mathcal{A}_T$ respect to the model chosen in Sec. 5.6: a double gaussian has been used for both peaks of for the D^+ peak only; a second order polynomial or a third order polynomial have been used for background description. The results of these tests are shown in Tabs. 26, 27.

Table 26: Deviation of $\mathcal{A}_T$, A_T and $\bar{A}_T$ from the best model to fit the distribution. All the values are multiples of $\times 10^{-3}$.

	Fit Model (sig + bkg)	$\Delta_{\mathcal{A}_T}$	Δ_{A_T}	$\Delta_{\bar{A}_T}$
D^+	double gaussian + line	—	—	—
	double gaussian + 2^{nd} ord. poly	-1.30	-1.14	+1.46
	three gaussians + 2^{nd} ord. poly	-0.53	-0.44	+0.62
	three gaussians + line	+0.19	+0.22	-0.16
	sum two gaussians + line	+0.05	+0.00	-0.01
D_s^+	double gaussian + 2^{nd} ord. poly	—	—	—
	sum two gaussians + 2^{nd} ord. poly	+0.04	-0.11	-0.11
	single gaussian + 2^{nd} ord. poly	-0.04	-0.72	-0.63
	single gaussian + 3^{nd} ord. poly	-0.03	-0.78	-0.70
	double gaussian + 3^{nd} ord. poly	+0.02	+0.11	+0.10
	three gaussians + 2^{nd} ord. poly	-0.01	+0.06	+0.11

Table 27: Statistical error of $\mathcal{A}_T$, A_T and $\bar{A}_T$ depending on the model used to fit the distribution. All the values are multiples of $\times 10^{-3}$.

	Fit Model (sig + bkg)	$\sigma_{\mathcal{A}_T}$	σ_{A_T}	$\sigma_{\bar{A}_T}$
D^+	double gaussian + line	10.04	14.13	14.26
	double gaussian + 2^{nd} ord. poly	10.27	14.46	14.59
	three gaussians + 2^{nd} ord. poly	10.15	14.29	14.42
	three gaussians + line	10.04	14.13	14.25
	sum two gaussians + line	10.06	14.13	14.26
D_s^+	double gaussian + 2^{nd} ord. poly	7.66	10.73	10.94
	sum two gaussians + 2^{nd} ord. poly	7.67	10.76	10.93
	single gaussian + 2^{nd} ord. poly	7.74	10.84	11.05
	single gaussian + 3^{nd} ord. poly	7.75	10.85	11.06
	double gaussian + 3^{nd} ord. poly	7.66	10.73	10.93
	three gaussians + 2^{nd} ord. poly	7.66	10.72	10.93

Assuming the larger deviation as the systematic error to be assigned, one gets

$$\sigma^{model}_{syst}(\mathcal{A}_T[D^+]) = \pm 1.30 \times 10^{-3},$$
$$\sigma^{model}_{syst}(\mathcal{A}_T[D^+_s]) = \pm 0.04 \times 10^{-3}.$$

The same procedure applied on A_T and $\bar{A}_T$ gives

$$\sigma^{model}_{syst}(A_T[D^+]) = \pm 1.14 \times 10^{-3},$$
$$\sigma^{model}_{syst}(\bar{A}_T[D^+]) = \pm 1.46 \times 10^{-3},$$
$$\sigma^{model}_{syst}(A_T[D^+_s]) = \pm 0.78 \times 10^{-3},$$
$$\sigma^{mode}_{syst}(\bar{A}_T[D^+_s]) = \pm 0.70 \times 10^{-3}.$$

5.7.7 Particle identification criteria

Particle identification criteria have been modified to looser or tighter conditions and the difference of $\mathcal{A}_T$ has been measured respect to the chosen selection. The results are shown in Tabs. 28, 29.

Assuming the larger deviation as the systematic error to be assigned, one gets

$$\sigma^{PID}_{syst}(\mathcal{A}_T[D^+]) = \pm 3.70 \times 10^{-3},$$
$$\sigma^{PID}_{syst}(\mathcal{A}_T[D^+_s]) = \pm 2.14 \times 10^{-3}.$$

The same procedure applied on A_T and $\bar{A}_T$ gives

$$\sigma^{PID}_{syst}(A_T[D^+]) = \pm 3.33 \times 10^{-3},$$
$$\sigma^{PID}_{syst}(\bar{A}_T[D^+]) = \pm 4.08 \times 10^{-3},$$
$$\sigma^{PID}_{syst}(A_T[D^+_s]) = \pm 2.47 \times 10^{-3},$$
$$\sigma^{PID}_{syst}(\bar{A}_T[D^+_s]) = \pm 6.73 \times 10^{-3}.$$

5.7.8 Summary of systematic errors

The systematic errors assigned to each identified source are listed in Tab. 30. The final systematic error is represented by the sum in quadrature of all the contributions.

5.8 FINAL RESULT

The final results for the CP asymmetry measured through T-odd correlation in $D^+_{(s)}$ decays are:

$$\mathcal{A}_T(D^+) = (-11.96 \pm 10.04_{stat} \pm 4.81_{syst}) \times 10^{-3},$$
$$\mathcal{A}_T(D^+_s) = (-13.57 \pm 7.67_{stat} \pm 4.82_{syst}) \times 10^{-3}.$$

Table 28: Deviation of $\mathcal{A}_T$, A_T and $\bar{A}_T$ from the used particle identification criteria. All the values are multiples of $\times 10^{-3}$.

	PID	$\Delta_{\mathcal{A}_T}$	Δ_{A_T}	$\Delta_{\bar{A}_T}$
D^+	kaonBDTTight + piKMLoose	–	–	–
	kaonBDTLoose + piKMLoose	-1.08	-1.86	+0.30
	kaonBDTVeryTight + piKMLoose	-3.70	-3.33	+4.08
	kaonBDTTight + piKMTight	+1.56	+3.76	+0.65
	kaonBDTTight + piKMVeryTight	+0.58	-0.43	-1.58
D_s^+	kaonBDTTight + piKMLoose	–	–	–
	kaonBDTLoose + piKMLoose	+1.09	+0.60	-1.57
	kaonBDTVeryTight + piKMLoose	+0.27	+0.33	-0.19
	kaonBDTTight + piKMTight	+1.07	-0.56	-2.68
	kaonBDTTight + piKMVeryTight	+2.14	-2.47	-6.73

Table 29: Statistical error of $\mathcal{A}_T$, A_T and $\bar{A}_T$ using different particle identification criteria. All the values are multiples of $\times 10^{-3}$.

	PID	$\sigma_{\mathcal{A}_T}$	σ_{A_T}	$\sigma_{\bar{A}_T}$
D^+	kaonBDTTight + piKMLoose	10.04	14.13	14.26
	kaonBDTLoose + piKMLoose	10.06	14.17	14.28
	kaonBDTVeryTight + piKMLoose	10.04	14.15	14.26
	kaonBDTTight + piKMTight	10.24	14.37	14.54
	kaonBDTTight + piKMVeryTight	10.65	14.95	15.15
D_s^+	kaonBDTTight + piKMLoose	7.66	10.73	10.93
	kaonBDTLoose + piKMLoose	7.67	10.74	10.94
	kaonBDTVeryTight + piKMLoose	7.66	10.73	10.94
	kaonBDTTight + piKMTight	7.80	10.93	11.12
	kaonBDTTight + piKMVeryTight	8.14	11.43	11.59

Table 30: Summary of the contributions to systematic error. All the values are multiples of $\times 10^{-3}$.

	D^+			D_s^+		
Source	$\sigma(\mathcal{A}_T)$	$\sigma(A_T)$	$\sigma(\bar{A}_T)$	$\sigma(\mathcal{A}_T)$	$\sigma(A_T)$	$\sigma(\bar{A}_T)$
Fit Bias	1.52	0.92	3.97	3.52	2.94	4.10
Reconstruction	2.05	2.84	1.26	1.00	1.00	1.27
Likelihood ratio	1.10	3.41	5.58	2.41	7.77	8.16
Fit model	1.30	1.14	1.46	0.11	0.78	0.70
PID	3.70	3.33	4.08	2.22	2.47	6.73
Total	4.81	5.74	8.20	4.82	8.76	11.41

These values are obtained uncovering the central values of A_T and $\bar{A}_T$:

$$A_T(D^+) = (+11.22 \pm 14.13_{stat} \pm 5.74_{syst}) \times 10^{-3},$$
$$\bar{A}_T(D^+) = (+35.13 \pm 14.26_{stat} \pm 8.20_{syst}) \times 10^{-3},$$
$$A_T(D^+_s) = (-99.24 \pm 10.73_{stat} \pm 8.48_{syst}) \times 10^{-3},$$
$$\bar{A}_T(D^+_s) = (-72.09 \pm 10.93_{stat} \pm 12.43_{syst}) \times 10^{-3}.$$

In conclusion, this search for CP violation using T-odd correlations in the high statistics sample of D^+ and $D^+_{(s)}$ decays to $K^+ K^0_S \pi^+ \pi^-$ gives a T-violating asymmetry that is consistent to zero with a sensitivity of the order of 1% for D^+ and 0.8% for D^+_s.

6 CONCLUSIONS

The work outlined in this thesis has allowed to measure the observables for CP violation in four-body D decays using T-odd correlations. Three decay modes have been considered:

- $D^0 \to K^+ K^- \pi^+ \pi^-$;
- $D^+ \to K^+ K^0_S \pi^+ \pi^-$;
- $D^+_s \to K^+ K^0_S \pi^+ \pi^-$.

Among them, the first two are Cabibbo suppressed, while the D^+_s one is Cabibbo favored.

The analysis of the D^0 decays showed that A_T and $\bar{A}_T$ are significantly different from zero:

$$A_T(D^0) = (-68.5 \pm 7.3_{(stat)} \pm 5.8_{(syst)}) \times 10^{-3},$$
$$\bar{A}_T(D^0) = (-70.5 \pm 7.3_{(stat)} \pm 3.9_{(syst)}) \times 10^{-3}.$$

indicating the effects of final state interaction. Although this two asymmetries are different from zero, their difference is not

$$\mathcal{A}_T(D^0) = (1.0 \pm 5.1_{(stat)} \pm 4.4_{(syst)}) \times 10^{-3}.$$

CP violation is then not found in this decay mode, according to the Standard Model. Furthermore, the high sensitivity of this result constrains the possible effects of new physics in this observable. In Figure 46 the comparison of this result to the other results for CP violation in D^0 decays is shown.

The search for CP violation in $D^+ \to K^+ K^0_S \pi^+ \pi^-$ decays showed that the A_T and $\bar{A}_T$ asymmetries are consistent to zero

$$A_T(D^+) = (+11.2 \pm 14.1_{(stat)} \pm 5.7_{(syst)}) \times 10^{-3},$$
$$\bar{A}_T(D^+) = (+35.1 \pm 14.3_{(stat)} \pm 8.2_{(syst)}) \times 10^{-3}.$$

This is an interesting result, since the same asymmetries are found different from zero in the D^0 and D^+_s analysis. The resulting T-violating asymmetry is

$$\mathcal{A}_T(D^+) = (-12.0 \pm 10.0_{(stat)} \pm 4.8_{(syst)}) \times 10^{-3},$$

consistent to zero with a sensitivity of 1%.

The similar search in $D^+_s \to K^+ K^0_S \pi^+ \pi^-$ decays gave

$$A_T(D^+_s) = (-99.2 \pm 10.7_{(stat)} \pm 8.5_{(syst)}) \times 10^{-3},$$
$$\bar{A}_T(D^+_s) = (-72.1 \pm 10.9_{(stat)} \pm 12.4_{(syst)}) \times 10^{-3},$$

resulting into a T-violating asymmetry of

$$\mathcal{A}_T(D_s^+) = (-13.6 \pm 7.7_{(stat)} \pm 4.8_{(syst)}) \times 10^{-3},$$

consistent to zero with a sensitivity of 0.8%.

The results are expected to be published in the first half of 2011 and look competitive to what has been previously obtained in the D^+ and D_s^+ decays, as well as shown in Figures 47 and 48. In particular for D_s^+, the measurement is among the ones having the best sensitivity.

All the measurements are consistent to zero as well as expected by the Standard Model. Nevertheless the results on the A_T and $\bar{A}_T$ asymmetries are quite interesting: while these asymmetries are different from zero in the D^0 and D_s^+ decays studied, they are consistent to zero in the D^+ decays, indicating a smaller contribution from final state interactions.

In conclusion, this search for CP violation showed that the T-odd correlations are a powerful tool to measure the CP violating observable $\mathcal{A}_T$. The relative simplicity of an analysis based on T-odd correlations and the high quality results that can be obtained, allow to consider this tool as fundamental to search for CP violation in four-body decays.

Even if the CP violation has not been found, excluding any New Physics effect to the sensitivity of about 0.5%, it is still worth to search for CP violation in D decays. The high statistics that can be obtained at the LHC or by the proposed high luminosity B-factories, make this topic to be considered in high consideration by experiments such as LHCb, SuperB or SuperBelle. The results outlined in this thesis strongly suggest to include a similar analysis into the Physics program of these experiments.

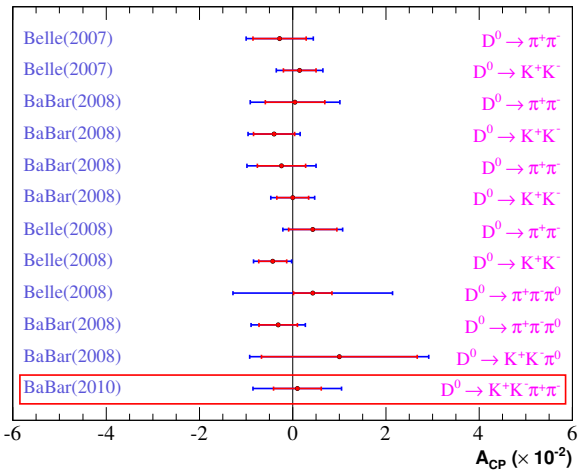


Figure 46: Comparison of the best recent results for CP violation in D^0 decays. The result of the analysis outlined in this thesis is highlighted into a red box.

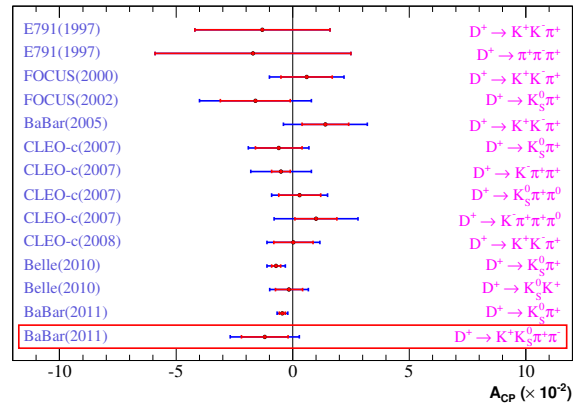


Figure 47: Comparison of the best recent results for CP violation in D^+ decays. The result of the analysis outlined in this thesis is highlighted into a red box.

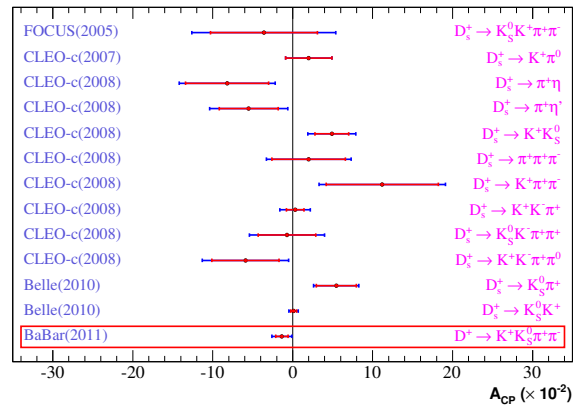


Figure 48: Comparison of the best recent results for CP violation in D_s^+ decays. The result of the analysis outlined in this thesis is highlighted into a red box.

A THE $c\bar{c}$ ASYMMETRIC PRODUCTION IN BABAR AND *BELLE*

The production of $c\bar{c}$ pairs in e^+e^- collisions at $10.58\,\mathrm{GeV}/c^2$ center of mass energy is mainly mediated by the electromagnetic process

$$e^+e^- \to \gamma \to c\bar{c}, \tag{A.1}$$

but e^+e^- could also produce $c\bar{c}$ pairs by the neutral weak current process [Halzen e Martin, 1984]

$$e^+e^- \to Z^0 \to c\bar{c}. \tag{A.2}$$

Let's see how this process can interfere in the analyses made at asymmetric detectors, such as *BABAR* and *Belle*. The electromagnetic process amplitude is

$$\mathcal{M}^{EM} \approx \frac{e^2}{k^2}, \tag{A.3}$$

where $k^2 = s = E^2_{CM}$. The weak process interfere with the electromagnetic one at the level

$$\frac{|\mathcal{M}^{EM}\mathcal{M}^{NC}|}{|\mathcal{M}^{EM}|^2} \approx \frac{G}{e^2/k^2} \approx \frac{10^{-5}k^2}{m_N^2}. \tag{A.4}$$

Assuming the mass of the neutron $m_N \approx 1\,\mathrm{GeV}/c^2$ and $k^2 \approx E^2_{CM} \approx 100(\,\mathrm{GeV}/c^2)^2$, one gets a 1% interference effect.

The differential cross section of the interference process, calculated using Feynman rules, is

$$\frac{dq}{d\Omega}(e^+_L e^-_R \to c^-_L c^+_L) = \frac{3\alpha^2}{4s}(1+\cos\theta)^2|2/3 + r c^c_L c^e_L|^2, \tag{A.5}$$

with

$$r = \frac{\sqrt{2}GM_Z^2}{s - M_Z^2 + iM_Z\Gamma_Z}\left(\frac{s}{e^2}\right). \tag{A.6}$$

The result is that the final production ratio of D and $\bar{D}$ differs from 1 depending on θ, the quarks production angle.

BABAR and *Belle* are asymmetric detectors and the center of mass frame is boosted to enhance separation of B's vertices, so that the final yield of D and $\bar{D}$ events has an intrinsic asymmetry [Aubert *et al.*, 2008].

B STATISTICAL TECHNIQUES USED BY BABAR SELECTORS

CONTENTS

The *BABAR* collaboration developed many statistical techniques to test the identification of the particles from the detector. A brief introduction to the techniques that are used for the selectors related to the analysis shown in this thesis are outlined in the following.

B.1 LIKELIHOOD RATIO

The technique of the likelihood ratio is based on the probability distribution that an event is signal or background depending upon a variable. Let suppose that an event is represented by n variables

$$\vec{x} = \{x_1, x_2, ..., x_n\}. \tag{B.1}$$

Using signal and background control samples, for each one of these variables a probability density function can be retrieved:

$$\mathcal{S}_i(x_i), \mathcal{B}_i(x_i), \quad \text{for } i = 1, .., n. \tag{B.2}$$

The likelihood ratio is defined as the product of the signal distribution of each variable, divided by the product of the relative background distributions:

$$\mathcal{L}(\vec{x}) = \prod_{i=1}^{n} \frac{\mathcal{S}_i(x_i)}{\mathcal{B}_i(x_i)}. \tag{B.3}$$

This technique allows to maximize the number of signal events respect to the background by a selection on a single variable. Usually the selection is chosen in order to maximize this ratio.

The advantage of this technique is that there are not developed as many selections as the number of variables studied, but there is only one selection, limiting the number of contributions to the systematics error.

This technique is widely used in this work, not only for the particle identification, but even for reducing the background in the $D^+_{(s)}$ data sample and to find the model that best fits the resulting distributions, both for D^0 and $D^+_{(s)}$ analysis.

B.2 ERROR CORRECTING OUTPUT CODE

Error Correcting Output Code (ECOC) [Dietterich e Bakiri, 1995] is an algorithm for making a multi-class classifier from binary classifiers. It was invented around 1995 by T. Dietterich and G. Bakiri.

Let's consider a simple example to understand how it works. Suppose the need to classify the data in four classes

$$\{A, B, C, D\} \tag{B.4}$$

by means of two binary classifiers

$$\alpha \text{ that separates } \{A, B\} \text{ from } \{C, D\}, \tag{B.5}$$
$$\beta \text{ that separates } \{A, C\} \text{ from } \{B, D\}. \tag{B.6}$$

At first, the indicator matrix can be obtained. The rows and columns

Table 31: Indicator matrix of a multi-class classifier from the combination of two binary classifiers. t_α and t_β represent how the given class is treated in training a sample for classifier α and β respectively. 1 means signal and 0 means background.

Class	t_α	t_β
A	1	1
B	1	0
C	0	1
D	0	0

of the indicator matrix shown in Table 31 can be interpreted as follows:

- each row represents the ideal output from each classifier for a given class. That is the template to compare the output from the classifier.
- Each column represents the ideal output for each class for a given classifier. That is the way the analyst should train each classifier.

Let's now suppose that an event d needs to be classified into one of the four classes defined in Equation B.4: the output of each classifier can be 0 or 1. In the case that

$$O_\alpha = 1 \quad O_\beta = 0, \tag{B.7}$$

the total output can be written as $O = 10$.

To determine which class is the more likely to identify the event, the Hamming distance can be defined: this is the number of bits that are different between the output and the corresponding row in the indicator matrix. The lower the Hamming distance, the more likely the class. In Table 32 the Hamming distance for the output of Equation B.7 together with the indicator matrix. The lowest Hamming distance is related to the class B, which would have returned the both the right answers from the two indicators if the event belonged to it. But what if an indicator had failed? In this hypothesis, the classes A and C become

Table 32: Hamming distance obtained comparing the output in Equation B.7 with the indicator matrix of Table 31.

Class	t_α	t_β	Hamming distance
A	1	1	1
B	1	0	0
C	0	1	2
D	0	0	1

the most likely. If both the two indicators are failing, then the class D is the more likely to be the right one.

The multi-class classified built from only 2 binary classifiers would then mis-classify the event if one or both the classifiers make a mistake. In order to fix this problem, one can add more classifiers and try to make sure that when one of them is failing, their sum can still recover the correct answer.

Let's first count how many different classifiers can be built for a fixed number n of classes. Each classifier can be represented by a binary string of length n, returning the ideal output for each one of the classes. Considering that there are no permutations among the elements of the string and that each element is a binary number, the total number of classifiers is 2^n. All these classifiers are shown in Table 33, that can also be followed to understand the next few paragraphs.

Table 33: All possible classifiers for a four-classes problem. Among them, there are only 7 valid classifiers.

A	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
B	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0
C	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0
D	1	1	1	1	0	0	0	0	1	1	1	1	0	0	0	0
d	all 1	Good Ones							Complement							

However, some classifiers are doubly accounted this way: the classifiers that are complement to each other actually are the same classifier. See for example 1100 and 0011, they are both trained to separate $\{A, B\}$ from $\{C, D, \}$. The number of valid classifiers is then 2^{n-1}.

Among these, there is also the "all 1" classifier (1111): this classifier does not add any separating power to the technique, since all the classes are the same for it. It can then be removed.

The final number of valid classifiers is $2^{n-1} - 1$ and they are shown in Table 34 for a four-classes problem. The indicator matrix made of all

Table 34: Exhaustive Matrix for a four-classes problem.

Class	t_0	t_1	t_2	t_3	t_4	t_5	t_6
A	1	1	1	1	1	1	1
B	0	1	0	1	0	1	0
C	1	0	0	1	1	0	0
D	1	1	1	0	0	0	0

the valid combinations of classifiers is called the "exhaustive matrix".

Let's draw an example to understand the power of this technique. If the event belongs to the class A, the ideal output is

$$O = 1111111. \tag{B.8}$$

Let's suppose that one, two or three classifiers are failing. The resulting Hamming distance for each one of the three cases is shown in Tables 35-37.

It can be observed that the class indication is correct up to two classifiers failing. The reason for this property is the row hamming distance. For this four-class exhaustive matrix, the hamming distance between any two rows is 4. That means classifiers can make two mistakes, in the worst case, at the bits that are not the same between two rows and there is still the chance to recover the correct answer.

As shown in Table 37, there is the chance to run into a tie situation. In the actual *BABAR* implementation, the problem is avoided by using real numbers for the classifiers output and generalize the hamming distance to the sum of the squared differences between each bit.

The drawback of this technique, is that the exhaustive matrix contains all the possible classifiers, and their number grows exponentially with the number of classes. This technique then cannot be easily developed for a large number of classes. However, a discussion about the feasibility of this technique into a many-classes case is outlined in Section 3 of [Dietterich e Bakiri, 1995].

In the *BABAR* implementation of the technique, a real output in $[-1, 1]$ is assigned to each classifier in place of the binary one. The four classes of events are related to the particle identification and are K, π, p, e. The exhaustive matrix is shown in Table 38.

B.3 BOOSTED DECISION TREE

Boosted Decision Tree (BDT) [Breiman *et al.*, 1984] is a technique of multi-variate analysis. This kind of technique is useful when the Likelihood ratio is failing due to the large correlations between the input variables or to the complexity of a probability density function into a multi-variate space.

In Boosted Decision Trees, the selection is done on a majority vote on the result of several decision trees, which are all derived from the same training sample by supplying different event weights during the training.

The decision trees are made of successive decision nodes that are used to categorize the events out of the sample as either signal or background. Each node is associated to a single discriminating variable to decide if the event is signal-like or background-like. This forms a tree like structure with "baskets" at the end (leave nodes), and an event is classified as either signal or background according to whether the basket where it ends up has been classified signal or background during the training.

During the training, the "cut criteria" for each node are defined. Starting from the root node, the full training event sample is taken and

Table 35: Hamming distance for the exhaustive matrix of a four-classes problem when there is one classifier failing (O=0111111).

Class	t_0	t_1	t_2	t_3	t_4	t_5	t_6	Hamming distance
A	1	1	1	1	1	1	1	1
B	0	1	0	1	0	1	0	3
C	1	0	0	1	1	0	0	5
D	1	1	1	0	0	0	0	5

Table 36: Hamming distance for the exhaustive matrix of a four-classes problem when there are two classifiers failing (O=0011111).

Class	t_0	t_1	t_2	t_3	t_4	t_5	t_6	Hamming distance
A	1	1	1	1	1	1	1	2
B	0	1	0	1	0	1	0	4
C	1	0	0	1	1	0	0	4
D	1	1	1	0	0	0	0	6

Table 37: Hamming distance for the exhaustive matrix of a four-classes problem when there are three classifiers failing (O=0001111).

Class	t_0	t_1	t_2	t_3	t_4	t_5	t_6	Hamming distance
A	1	1	1	1	1	1	1	3
B	0	1	0	1	0	1	0	3
C	1	0	0	1	1	0	0	3
D	1	1	1	0	0	0	0	7

Table 38: Exhaustive Matrix for the *BABAR* particle identification.

Class	t_0	t_1	t_2	t_3	t_4	t_5	t_6
K	1	1	1	1	1	1	1
π	-1	1	-1	1	-1	1	-1
p	1	-1	-1	1	1	-1	-1
e	1	1	1	-1	-1	-1	-1

the variables and the corresponding cut values that give the best separation between signal and background are selected. Using this cut criterion, the sample is then divided into two sub-samples, a signal-like and a background-like sample. Two new nodes are then created for each of the two sub-samples and they are constructed using the same mechanism as described for the root node. The division is stopped once a certain node has reached either a minimum number of events, or a minimum or maximum signal purity. These leave nodes are then called "signal" or "background" if they contain more signal respective background events from the training sample.

The decision tree is boosted by giving a weight to signal events in the training sample that end up into a background node (and vice versa) larger than the events falling into the correct leave node. This results in a re-weighted training event sample, with which a new decision tree can be developed. The boosting can be applied several times (typically 100-500 times) and one ends up with a set of decision trees (a forest). The boosting stabilizes the response of the decision trees respect to fluctuations in the training sample and is able to considerably enhance the performance with respect to a single tree. As a drawback, there is the problem of the over training of the decision tree, that can fake the results.

BIBLIOGRAPHY

Agostinelli, S. *et al.* (2003), «G4–a simulation toolkit», *Nucl. Instrum. Methods Phys. Res., Sect. A* , vol. 506 (3), p. 250 – 303.

Anulli, F. *et al.* (2005), «*BABAR* forward endcap upgrade», *Nucl. Instrum. Methods Phys. Res., Sect. A* , vol. 539 (1-2), p. 155–171.

Aubert, B. *et al.* (2002), «The *BABAR* detector», *Nucl. Instrum. Methods Phys. Res., Sect. A* , vol. 479, p. 1–116.

Aubert, B. *et al.* (2004), «Measurement of the $B^0 \to \phi K^{*0}$ Decay Amplitudes», *Phys. Rev. Lett.* , vol. 93, p. 231 804.

Aubert, B. *et al.* (2005), «A Search for CP violation and a measurement of the relative branching fraction in $D^+ \to K^- K^+ \pi^+$ decays», *Phys. Rev. D* , vol. 71, p. 091 101.

Aubert, B. *et al.* (2008), «Search for CP Violation in the Decays $D^0 \to K^- K^+$ and $D^0 \to \pi^- \pi^+$ », *Phys. Rev. Lett.* , vol. 100, p. 061 803.

Aubert, B. *et al.* (2009), «Measurement of the $e^+e^- \to b\bar{b}$ cross section between $\sqrt{s}$ = 10.54 GeV and 11.20 GeV», *Phys. Rev. Lett.* , vol. 102, p. 012 001.

Bensalem, W., Datta, A. e London, D. (2002a), «New physics effects on triple product correlations in Lambda(b) decays», *Phys. Rev. D* , vol. 66, p. 094 004.

Bensalem, W., Datta, A. e London, D. (2002b), «T violating triple product correlations in charmless Λ_b decays», *Phys. Lett. B* , vol. 538, p. 309–320.

Bensalem, W. e London, D. (2001), «T odd triple product correlations in hadronic b decays», *Phys. Rev. D* , vol. 64, p. 116 003.

Bergmann, S., Grossman, Y., Ligeti, Z., Nir, Y. e Petrov, A. A. (2000), «Lessons from CLEO and FOCUS measurements of mixing parameters», *Phys. Lett. B* , vol. 486, p. 418 – 425.

Bigi, I. I. (2001), «Charm physics: Like Botticelli in the Sistine Chapel», in «Invited talk given at KAON 2001, the 3th International Conference on CP Violation, Pisa, Italy», .

Bigi, I. I. Y. e Sanda, A. I. (2000), *CP Violation*, vol. 9, Cambridge Monographs in Particle Physics, Nuclear Physics and Cosmology.

Bjorken, J. D. e Drell, S. D. (1965), *Relativistic quantum fields*, International Series in Pure and Applied Physics, McGraw-Hill.

Breiman, L. *et al.* (1984), *Classification and Regression Trees*, Waldsworth International.

Buccella, F., Lusignoli, M., Miele, G., Pugliese, A. e Santorelli, P. (1995), «Nonleptonic weak decays of charmed mesons», *Phys. Rev. D* , vol. 51, p. 3478–3486.

Cabibbo, N. (1963), «Unitary Symmetry and Leptonic Decays», *Phys. Rev. Lett.* , vol. 10, p. 531–533.

Chen, C., Cenci, R. e Jawahery, H. (2010), *Search for* CP *violation in the decay* $D^+ \to K^0_S \pi^+$, *BABAR* internal note.

Chen, K.-F. *et al.* (2005), «Measurement of Polarization and Triple-Product Correlations in $B \to \phi K^*$ Decays», *Phys. Rev. Lett.* , vol. 94, p. 221 804.

Christenson, J. H., Cronin, J. W., Fitch, V. L. e Turlay, R. (1964), «Evidence for the 2π Decay of the K^0_2 Meson», *Phys. Rev. Lett.* , vol. 13, p. 138–140.

Convery, M. *et al.* (2006), «A novel technique for the production of large area Z-coordinate readout planes for the *BABAR* muon system», *Nucl. Instrum. Methods Phys. Res., Sect. A* , vol. 556 (1), p. 134 – 139.

del Amo Sanchez, P. *et al.* (2010), «Search for CP violation using T-odd correlations in $D^0 \to K^+ K^- \pi^+ \pi^-$ decays», *Phys. Rev. D* , vol. 81, p. 111 103.

Dietterich, T. e Bakiri, G. (1995), «Solving Multiclass Learning Problems via Error-Correcting Output Codes», *Journal of Artificial Intelligence Research*, vol. 2, p. 263–286.

Dirac, P. A. M. (1928), «The Quantum Theory of Electron», *Proceedings of Royal Society London*, vol. A (117), p. 610–624.

Gaiser, J. (1982), *Charmonium Spectroscopy From Radiative Decays of the* J/ψ *and* ψ', Tesi di dottorato, SLAC, sLAC-0255.

Grossman, Y., Kagan, A. L. e Nir, Y. (2007), «New physics and CP violation in singly Cabibbo suppressed D decays», *Phys. Rev. D* , vol. 75, p. 036 008.

Halzen, F. e Martin, D. (1984), *Quarks And Leptons: An Introductory Course In Modern Particle Physics*, New York, Usa: Wiley.

Iarocci, E. (1983), «Plastic streamer tubes and their applications in high energy physics», *Nucl. Instr. Methods Phys. Res.*, vol. 217, p. 30–42.

Itoh, R. *et al.* (2005), «Studies of CP Violation in $B \to J/\psi\ K^*$ Decays», *Phys. Rev. Lett.* , vol. 95, p. 091 601.

James, F. e Roos, M. (1975), «Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations», *Comput. Phys. Commun.* , vol. 10, p. 343–367.

Johnson, N. L. (1949), «Systems of frequency curves generated by methods of translation», *Biometrika*, vol. 36, p. 149.

Kobayashi, M. e Maskawa, T. (1973), «CP Violation in the Renormalizable Theory of Weak Interaction», *Prog. Theor. Phys.* , vol. 49, p. 652–657.

Link, J. *et al.* (2005), «Search for T violation in charm meson decays», *Phys. Lett. B* , vol. 622, p. 239–248.

Nakamura, K. *et al.* (2010), «Review of Particle Physics», *J. Phys.* **G**, vol. 37, p. 075 021.

Oller, J. A. (2005), «Final state interactions in hadronic D decays», *Phys. Rev. D* , vol. 71, p. 054 030.

Perkins, D. H. (1982), *Introduction To High-energy Physics*, Reading, Usa: Addison-wesley.

Sakharov, A. D. (1967), «Violation of CP Invariance, C Asymmetry, and Baryon Asymmetry of the Universe», *Pisma Zh. Eksp. Teor. Fiz.*, vol. 5, p. 32–35.

Sjostrand, T. (1994), «High-energy physics event generation with PYTHIA 5.7 and JETSET 7.4», *Comput. Phys. Commun.* , vol. 82, p. 74–90.

Verkerke, W. e Kirkby, D. P. (2003), «The RooFit toolkit for data modeling», ArXiv:physics/0306116.

Wolfenstein, L. (1983), «Parametrization of the Kobayashi-Maskawa Matrix», *Phys. Rev. Lett.* , vol. 51, p. 1945–1947.

Wu, C. S., Ambler, E., Hayward, R. W., Hoppes, D. D. e Hudson, R. P. (1957), «Experimental Test of Parity Conservation in Beta Decay», *Phys. Rev.* , vol. 105, p. 1413–1415.

COLOPHON

This thesis has been typeset using the *ArsClassica* LaTeX style, made by Lorenzo Pantieri slightly modifying the style *ClassicThesis* from André Miede. Both of them can be found in the Comprehensive TeX Archive Network (CTAN).

Final Version: February 2011.