

Measurement of the Average B_s^0 Lifetime in the Decay $B_s^0 \rightarrow J/\psi \phi$

Thilo Pauly
Keble College, Oxford

Thesis submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy at the University of Oxford

Michaelmas Term, 2003



Abstract

The lifetime difference between the long (CP odd) and short (CP even) lived components of the B_s^0 meson is currently predicted to be of the order of 10 % in the Standard Model. It has been suggested that the decay $B_s^0 \rightarrow J/\psi \phi$ is predominantly CP even and thus the measured average lifetime could be shorter than the lifetime measured in the inclusive decay modes. We present a measurement of the average lifetime of the B_s^0 meson in its decay $B_s^0 \rightarrow J/\psi \phi$, with $J/\psi \rightarrow \mu^+\mu^-$ and $\phi \rightarrow K^+K^-$.

During January 2002 and August 2003 the CDF experiment at the Tevatron has been exposed to about 135 pb^{-1} of $p\bar{p}$ collisions with a centre-of-mass energy of $\sqrt{s} = 1.96 \text{ TeV}$. In the data sample collected with the J/ψ dimuon trigger we fully reconstruct about 125 $B_s^0 \rightarrow J/\psi \phi$ candidates with precision silicon information. This is currently the largest exclusive B_s^0 sample. We perform a fit to the proper decay time information to extract the average B_s^0 lifetime and simultaneously use the mass information to disentangle signal from background. For cross-checks we measure the lifetime in the higher statistics modes $B_u^\pm \rightarrow J/\psi K^\pm$ and $B_d^0 \rightarrow J/\psi K^{*0}$, which both have similar decay topologies and kinematics. We obtain

$$\tau(B_s^0 \rightarrow J/\psi \phi) = (1.31_{-0.13}^{+0.15}(\text{stat.}) \pm 0.02(\text{syst.})) \text{ ps} ,$$

which is currently the best single measurement of the B_s^0 lifetime and is consistent with other measurements. This result is not accurate enough to establish the existence of a possible significant lifetime difference between the CP odd and even states.

Acknowledgements

This thesis would not have been possible without the help and contributions of many friends and colleagues. I want to take the opportunity to thank all of them, and in particular:

- Todd Huffman for giving me the opportunity to work with the CDF group at Oxford and dealing with the administrative side of my D.Phil. I am especially grateful for the opportunity to stay at Fermilab for 1.5 years.
- Jeff Tseng for competent guidance, patient supervision and his constant interest in my work. His sharp physical understanding provided vital input at any time. Together with Farrukh Azfar for the great culinary expeditions all across Chicagoland.
- Joe Boudreau for his great experience in particle physics and many stimulating discussions on B lifetimes.
- The Oxford CDF group (Farrukh Azfar, Sam Harper, Jim Loken, Louis Lyons, Giulia Manca, Matthew Martin, Nicola Pounder, Tracey Pratt, Jonas Rademacker, Armin Reichold, Peter Renton, Aidan Robson) for providing a friendly, stimulating and productive working environment.
- Roman Walczak and Peter Watkins for the final examination and the interesting discussions we had during the viva.
- The MIT group (in particular Christoph Paus, Konstantin Anikeev, Andreas Korn) for letting me participate in their analysis meetings and for all their help with technical issues.
- Michael Muirhead and Gerry Bauer for their interest on my work on time-of-flight reconstruction, and Barry Wicklund, Masa Tanaka and Rolf Oldeman for their support of my work on Level 1 trigger rates.
- Lucio Cerrito, Sinéad Farrington, Anant Gajjar, Helen Hayward, Cigdem Issever, Natasa Miladinovic, Aidan Robson, Martin Turner, Bryan Ward, and Simon Waschke for a great working and social environment at Fermilab.
- Keith Hamilton, Simon Jolly and Michael Ramage for great pub discussions.
- Robert Keeling, Michael Wedow, Stephanie Hunter, Jessica Meade and all members of Keble MCR for good friendship.
- All members of the CDF Collaboration and the Oxford Particle Physics Group for providing a friendly and productive work environment.

Contents

1	The Lifetime of the B_s^0 Meson	1
1.1	Introduction	1
1.2	Operator Product Expansion	4
1.3	B Meson Lifetimes	6
1.4	Mixing of Neutral Mesons	11
1.5	Calculation of $\Delta\Gamma_s$	13
1.6	Standard Model Expectation for $\Delta\Gamma_s$	14
1.7	New Physics Contributions to $\Delta\Gamma_s$	20
1.8	Time Evolution of Untagged B Mesons	21
1.9	The Decay Mode $B_s^0 \rightarrow J/\psi \phi$	22
1.10	Fitting the Average Lifetime	25
1.11	Current Experimental Status of B Lifetime Measurements	28
1.12	Current Experimental Status of $\Delta\Gamma_s$ Measurements	28
2	The CDF Experiment at Fermilab	33
2.1	The Tevatron Collider	33
2.2	The CDF II Detector	35
2.3	Silicon Vertex Detectors	37
2.4	Central Outer Tracking Chamber	41
2.5	Time-of-Flight (TOF)	43
2.6	Solenoid Magnet	43
2.7	Calorimeters	44
2.8	Muon Systems	45
2.9	The Trigger System	48
2.10	The Level 1 Trigger	50
3	Reconstruction	55
3.1	Track Parameterisation	55
3.2	COT Track Reconstruction	56
3.3	Outside-In Silicon Tracking	58
3.4	Tracking Corrections	59
3.5	Muon Reconstruction	61
3.6	The J/ψ Dimuon Triggers	64
3.7	The Luminous Region	65

4	Selection of $B_s^0 \rightarrow J/\psi \phi$	67
4.1	Common Selection Cuts	67
4.2	Selection of $B_u^\pm \rightarrow J/\psi K^\pm$	74
4.3	Selection of $B_d^0 \rightarrow J/\psi K^{*0}$	76
4.4	Selection of $B_s^0 \rightarrow J/\psi \phi$	79
4.5	Proper Decay Time Calculation	82
5	Lifetime Analysis	85
5.1	Fit Model	85
5.2	Fitting $B_u^\pm \rightarrow J/\psi K^\pm$	87
5.3	Fitting $B_d^0 \rightarrow J/\psi K^{*0}$	91
5.4	Fitting $B_s^0 \rightarrow J/\psi \phi$	93
6	Systematics and Cross-Checks	97
6.1	Residual Silicon Mis-Alignment	97
6.2	Beam Spot Size	99
6.3	Proper Decay Time Resolution Function	99
6.4	Proper Decay Time Background Parameterisation	101
6.5	Biases from the Fitting Procedure	104
6.6	Selection Cuts	107
6.7	Cross-Check with Realistic Simulation	115
6.8	$K\pi$ Assignment Ambiguity	115
6.9	Cross-Check: Background from Sidebands	116
6.10	Simultaneous Fit to Sideband and Signal Regions	122
6.11	Cross-check: Stability Over Time	122
6.12	Summary	124
7	Summary and Conclusions	127
A	Charged Current Interactions in the Standard Model	131
B	Optimisation of Transverse Momentum Selection Cuts	133
C	Vertex Fit	137
D	Study of a $\Delta\Gamma/\Gamma$ Transversity Analysis	141

Chapter 1

The Lifetime of the B_s^0 Meson

1.1 Introduction

A few neutral mesons show the effect of mixing: the ability to change from their particle to their antiparticle state. This is a remarkable consequence of basic quantum mechanics and the structure of the weak interaction. This oscillation from matter to antimatter can be used to measure fundamental parameters of the Standard Model. In addition, it is closely related to CP violation¹ and might have far reaching effects, such as the matter-antimatter asymmetry observed in the universe.

In the second half of the 1950s, mixing was proposed by Gell-Mann and Pais [3] to occur in the K^0 system and experimentally confirmed a little later [4], [5]. In 1964 violation of CP symmetry was observed in the mixing of K^0 [6]. In 1986 mixing has been observed in the B meson system by the UA1 Collaboration: they measured the time-integrated mixing probability in a sample of B_d^0 and B_s^0 mesons [7]. A little later the same effect was measured only for the B_d^0 system by the ARGUS [8] and CLEO [9] collaborations. In the 1990s time-dependent mixing of B_d^0 mesons has been established by the LEP experiments and SLD [10], and the CDF experiment [11]. To date no experiment has been able to resolve the expected fast oscillation of the B_s^0 meson. In addition the two B_s^0 mesons are expected to have different natural widths. This width difference, which is of main importance to this thesis, has not been experimentally confirmed. Mixing in D^0 mesons as well as the width difference of the B_d^0 system is expected to be a very small effect in the Standard Model, and has not been observed. The K^0 , D^0 , B_d^0 and B_s^0 are the only mesons that can show flavour oscillations. The π^0 is its own antiparticle, the top quark is so heavy that it decays before forming stable hadrons, and excited meson states decay strongly or electromagnetically before any mixing can occur. The properties of the K^0 , D^0 , B_d^0 and B_s^0 such as quark content, mass and lifetime are listed in table 1.1. Table 1.2 gives an overview of the current experimental values of mixing parameters – the mass difference Δm and the width difference $\Delta\Gamma/\Gamma$ [12].

Although the four different neutral meson systems K^0 , D^0 , B_d^0 and B_s^0 can be

¹C denotes the charge conjugation symmetry, and P the parity symmetry. For more details we refer the reader to standard text books on particle physics [1], [2].

System	Particle	Antiparticle	Mass (GeV/ c^2)	Lifetime (ps)
K^0	$\bar{s}d$	$s\bar{d}$	0.497672 ± 0.000031	$\tau_{K_S} = 89.35 \pm 0.08$
				$\tau_{K_L} = 51,700 \pm 400$
D^0	$c\bar{u}$	$\bar{c}u$	1.8645 ± 0.0005	0.4117 ± 0.0027
B_d^0	$\bar{b}d$	$b\bar{d}$	5.2794 ± 0.0005	1.542 ± 0.016
B_s^0	$\bar{b}s$	$b\bar{s}$	5.3696 ± 0.0024	1.461 ± 0.057

Table 1.1: Quark content for particle and antiparticle, average mass and lifetime of the K^0 , D^0 , B_d^0 and B_s^0 systems. Since the two physical states of the K^0 system have very different lifetimes, the individual lifetimes of the long-lived (K_L) and short-lived state (K_S) are listed (see section 1.4 for details).

System	Δm ($\hbar s^{-1}$)	$\Delta\Gamma/\Gamma$
K^0	$(0.5303 \pm 0.0009) \times 10^{10}$	$\tau_{K_S} = (0.8935 \pm 0.0008) \times 10^{-10}s$
		$\tau_{K_L} = (5.17 \pm 0.04) \times 10^{-8}s$
D^0	$< 7 \times 10^{10}$	-0.003 ± 0.022
B_d^0	$(0.489 \pm 0.008) \times 10^{12}$	< 0.08
B_s^0	$> 13.1 \times 10^{12}$	< 0.31

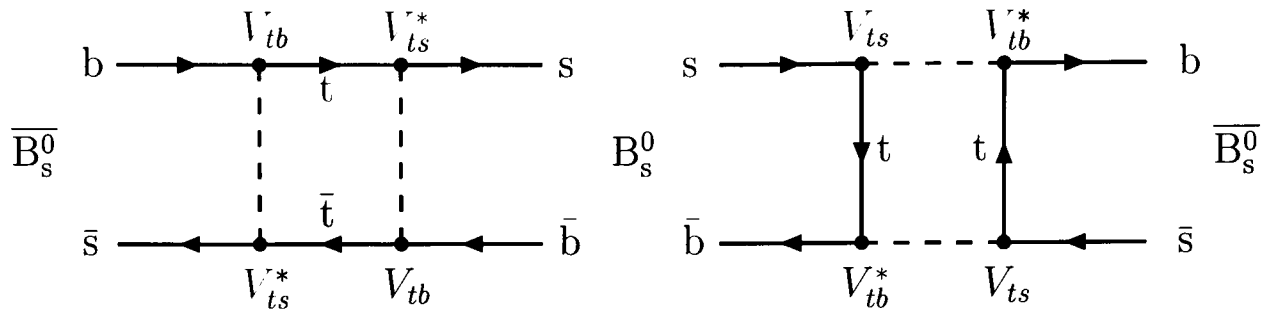
Table 1.2: Current experimental status of mixing parameters for the different neutral meson systems. Limits are at 95% C.L. For the K^0 system the individual lifetimes are reported.

described with the same formalism, they differ widely from one another in decay rates, mixing and CP violation. In the Standard Model these differences come mainly from the hierarchy between CKM matrix elements² and quark masses, but also phase space and long distance rescattering effects play a role.³

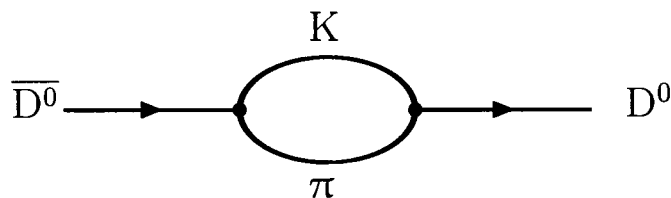
The description of mixing through box diagrams as in figure 1.1 leads to estimates of the mixing frequency Δm and the width difference $\Delta\Gamma$ by simply counting powers of quark masses and CKM parameters. The quark mass of the dominating internal quark lines contributes with the second power, and with the Wolfenstein parameterisation of the CKM matrix (see appendix A) enters the Wolfenstein parameter $\lambda = \sin\theta_c \approx 0.22$ [13]. The different mixing frequencies Δm for the K^0 , D^0 , B_d^0 and B_s^0 systems are proportional to $\lambda^2 m_c^2$, $\lambda^2 m_c^2$, $\lambda^6 m_t^2$ and $\lambda^4 m_t^2$, respectively. Note in particular the different dynamical origin of Δm in the K^0 and $B_{d,s}^0$ systems. In the latter case the high top quark mass compensates for the CKM suppression. For the D^0 and K^0 systems the important momenta lie in the neighbourhood or below the strange quark mass m_s , so that short-distance methods are inadequate. An appropriate approach is to look at virtual intermediate hadronic states (see figure 1.2). While in the K^0 case

²Please refer to appendix A for a brief reminder on the CKM matrix and charged current interactions in the Standard Model.

³Strong scattering between real intermediate states.

Figure 1.1: Box diagrams for the B_s^0 system.

the major states are reasonably limited to π , η , η' , and $\pi\pi$ there are many more to be considered for D^0 .

Figure 1.2: Example of a long distance effect diagram in the D^0 system through intermediate $K\pi$ states.

We are in particular concerned with $\Delta\Gamma$. In terms of simple power-counting, $\Delta\Gamma$ behaves like $\lambda^2 m_s^2$, $\lambda^2 m_c^2$, $\lambda^6 m_b^2$, $\lambda^4 m_b^2$ respectively for the systems previously discussed. The K^0 system is very peculiar in the sense that $\Gamma_S \gg \Gamma_L$ comes from a dynamical accident: the very different phase space available for the two possible final states $\pi\pi$ ($CP = +1$) and $\pi\pi\pi$ ($CP = -1$). For the D^0 we expect $\Delta\Gamma/\Gamma$ on the percent order due to the absorptive part of the same long distance diagrams contributing to Δm . We also note that the values of $\Delta\Gamma/\Gamma$ for the B_d^0 and B_s^0 systems are in a ratio of $\lambda^2 : 1$. Although it seems hopeless to measure $\Delta\Gamma/\Gamma$ for the B_d^0 system, the value of a few percent for the B_s^0 system is within experimental reach.

Mixing phenomena in neutral B meson systems provide an important testing ground for standard model flavour dynamics. The mass difference Δm_d between the B_d^0 eigenstates gave the first evidence for a large top quark mass and provides a valuable constraint on V_{td} in the CKM unitarity triangle. A measurement of Δm_s , the corresponding quantity for the B_s^0 system, would yield more information and help to reduce hadronic uncertainties in the extraction of CKM parameters [14]. Complementary knowledge can be obtained from measuring the width difference $\Delta\Gamma_s$ of the B_s^0 [15]. It is expected to be the largest rate difference in the B hadron sector and roughly of the order of 10% in the Standard Model, which brings it within current experimental reach. The width difference for the B_d^0 mesons, on the other hand, is CKM suppressed and experimentally much harder to determine.

A large value of the width difference would open up the possibility for observation of CP violation and the extraction of CKM phases from *untagged* B_s^0 data samples [16], [17], [18]. There are two interesting aspects of untagged data samples. First, tagging always costs in statistics and purity. Second, the rapid oscillations dependent on $\Delta m_s t$ all cancel in time evolutions of untagged B_s^0 samples, which are governed by the two exponentials $\exp(-\Gamma^{Short}t)$ and $\exp(-\Gamma^{Long}t)$ alone.

Another interesting point is, that new physics can only lead to a decrease of the width difference compared to the Standard Model value [19]. An experimental number which is considerably smaller than the theoretical lower bound, would thus be a hint for new physics that affects B_s^0 mixing. The large error on the theory prediction of $\Delta\Gamma_s/\Gamma_s$ though makes it improbable for new physics to be first observed in the width difference, since new physics contributions to the B_s^0 mixing phase large enough to cause an effect larger than the theoretical uncertainty would rather be seen elsewhere, for instance as a time-dependent CP asymmetry.

For this thesis we use fully reconstructed decays $B_s^0 \rightarrow J/\psi \phi$, reconstructed from 135 pb^{-1} of $p\bar{p}$ collisions taken with the CDF detector at Fermilab between January 2002 and August 2003. We have reconstructed about 125 signal events above background, with high spatial resolution from silicon information. Although this is already world-wide the biggest sample of fully reconstructed B_s^0 decays, the sample size is much smaller than initially expected (2 fb^{-1}). This is mainly due to luminosity problems with the Tevatron machine. Studies show (see appendix D) that 125 events are insufficient to measure $\Delta\Gamma_s/\Gamma_s$ with reasonable precision. Therefore we focus on the average B_s^0 lifetime in this decay. We also use similar, but higher statistics decay modes as control samples to gain confidence in the analysis method. These modes are $B_u^\pm \rightarrow J/\psi K^\pm$ and $B_d^0 \rightarrow J/\psi K^{*0}$, which have decay topologies and kinematics similar to $B_s^0 \rightarrow J/\psi \phi$. In the following sections we explain briefly the phenomenology of B meson lifetimes in general, and the B_s^0 mixing parameters and their connection to the average B_s^0 lifetime in particular.

1.2 Operator Product Expansion

In this section we give an overview over some theoretical methods used to estimate lifetime ratios and width differences in the Standard Model. We closely follow [20]. Weak decays of B mesons involve a large range of different mass scales: first there is the W boson mass M_W , which appears for instance in the weak $b \rightarrow c\bar{c}s$ decay amplitude. The second scale in the problem is the mass m_b of the decaying b quark. Finally there is the QCD scale parameter Λ_{QCD} , which sets the scale for the strong binding forces in the B mesons. QCD corrections associated with these scales must be treated in different ways.

The common tool for factorising amplitudes into short- and long-distance parts is the **Operator Product Expansion (OPE)**, which we will apply two times between our three scales. The OPE allows short (μ_{SD}) and long distance (μ_{LD}) contributions to be separated. Using renormalisation group methods we can sum large logarithms $\ln \frac{\mu_{SD}}{\mu_{LD}}$ to all orders in perturbation theory.

This can be done very efficiently by renormalisation group methods. The resulting renormalisation group improved perturbative expansion for $C_i(\mu)$ in terms of the effective coupling constant $\alpha_s(\mu)$ does not involve these logarithms.

In order to calculate the amplitude $A(M \rightarrow F)$ the matrix elements $\langle Q_i(\mu) \rangle$ have to be evaluated. Since they involve long-distance contributions one is forced in this case to use non-perturbative methods such as lattice QCD, QCD sum rules, and so on. One of the outstanding issues in the calculation of $\langle Q_i(\mu) \rangle$ is the compatibility (matching) of $\langle Q_i(\mu) \rangle$ with the Wilson coefficients $C_i(\mu)$. $\langle Q_i(\mu) \rangle$ must have the correct μ and renormalisation scheme dependence to ensure that the physical quantities are μ - and scheme-independent. Non-perturbative methods often struggle with this problem, but lattice calculations using non-perturbative matching techniques can meet this requirement.

In the decay of B hadrons the binding energy, which is of order Λ_{QCD} , is small compared to the b quark mass. In most of the phase space of the decay the energy release, which can be as large as $\mathcal{O}(m_b)$, is much larger than the typical scale of hadronic interactions. The large energy release implies a short distance, and we can use the same tools as before – an OPE (though not the same OPE as between the scales M_W and m_b as described before) – to separate short and long distances. This OPE is called **Heavy Quark Expansion (HQE)**. In this way, inclusive decay rates can be described with a double series in m_b and $\alpha_s(m_b)$.

1.3 B Meson Lifetimes

We now give a review of the main techniques in calculating lifetime ratios, closely following [20], [21], and [22].

In a naive spectator model, where the spectator quark has no effect on the decay of the heavy b quark in the B meson, the lifetimes of all B mesons are expected to be equal. Measurements of individual lifetimes yield information about spectator effects. These effects involve the participation of the light constituents in the decay and thus contribute to the differences in the decay widths and lifetimes of different species of beauty hadrons.

One can calculate inclusive decay widths in QCD proper through an expansion in inverse powers of the heavy flavour quark mass without recourse to phenomenological assumptions (HQE). The non-perturbative contributions are treated systematically in this way; they are found to produce corrections of order a few percent in beauty decays, i.e. typically somewhat smaller than the perturbative corrections. Extensive theoretical work has gone into calculating lifetime ratios such as $\tau(B_u)/\tau(B_d)$ or $\tau(B_s)/\tau(B_d)$, where many theoretical uncertainties related to the values of the b-quark mass and CKM elements cancel to large extent. The lifetime ratios are computed in this framework as series in inverse powers of the mass of the b quark. The leading term of this expansion corresponds to the decay of a free b quark. This term is universal, contributing equally to the lifetimes of all beauty hadrons. Remarkably, the first correction to this result is of order $1/m_b^2$.

Spectator effects also participate in the weak process. Though for decays of heavy particles, these effects are strongly suppressed due to the need for the b quark and a light quark in the heavy hadron to be close together.

Inclusive decay rates, which determine the probability of the decay of a particle into the sum of all possible final states with a given set of quantum numbers f , have two advantages from the theoretical point of view: first, bound-state effects related to the initial state can be accounted for in a systematic way using the HQE; secondly, the fact that the final state consists of a sum over many hadronic channels eliminates bound-state effects related to the properties of individual final state hadrons. This feature is based on the hypothesis of **parton-hadron duality** [23], i.e. the assumption that sufficiently many exclusive hadronic channels contribute to the inclusive rate, so that the accidentals of the low-energy resonance structure do not affect the total rates of inclusive processes.

Unitarity relates inclusive decay rates to the imaginary part of certain forward “scattering” amplitudes (optical theorem [1]), which opened the way for the consistent use of the OPE. The starting object is the transition operator $T(b \rightarrow f \rightarrow b)$ describing the forward scattering of b quarks via an intermediate state f (see e.g. [1]). To second order in the weak interactions the transition operator is given by

$$T(b \rightarrow f \rightarrow b) = i \int d^4x \mathcal{T}(\mathcal{H}_{eff}(x), \mathcal{H}_{eff}(0)) , \quad (1.3)$$

where \mathcal{H}_{eff} denotes the relevant effective weak Hamiltonian and \mathcal{T} is the time-ordered product.

Using the optical theorem, the inclusive decay width of a hadron H_b containing a b quark can be written as the forward matrix element of the imaginary part of the transition operator T

$$\Gamma(H_b \rightarrow X) = \frac{1}{2M_{H_b}} \text{Im} \langle H_b | T | H_b \rangle . \quad (1.4)$$

Notice that the factor $1/(2M_{H_b})$ (where M_{H_b} is the mass of H_b) reflects the relativistic normalisation of the state $|H_b\rangle$. Because $m_b \ll M_W$, we can write the effective Hamiltonian for the case of semileptonic and non-leptonic decays, renormalised at the scale $\mu = m_b$, as

$$\begin{aligned} \mathcal{H}_{eff} = & \frac{G_F}{\sqrt{2}} V_{cb} \{ c_1(m_b) [\bar{d}'_L \gamma_\mu u_L \bar{c}_L \gamma^\mu b_L + \bar{s}'_L \gamma_\mu c_L \bar{c}_L \gamma^\mu b_L] \\ & + c_2(m_b) [\bar{c}_L \gamma_\mu u_L \bar{d}'_L \gamma^\mu b_L + \bar{c}_L \gamma_\mu c_L \bar{s}'_L \gamma^\mu b_L +] \\ & + \sum_{l=e,\mu,\tau} \bar{l}_L \gamma_\mu \nu_l \bar{c}_L \gamma^\mu b_L \} + \text{h.c.} , \end{aligned} \quad (1.5)$$

where G_F is the effective Fermi coupling, $q_L = \frac{1}{2}(1 - \gamma_5)q$ denotes the left-handed quark fields, $d' = d \cos \theta_C + s \sin \theta_C$ and $s' = s \cos \theta_C - d \sin \theta_C$ are the Cabibbo-rotated down- and strange-quark fields, and we have neglected $b \rightarrow u$ transitions. The Wilson coefficients c_1 and c_2 take into account the QCD corrections arising from

the fact that the effective Hamiltonian is written at a renormalisation scale $\mu = m_b$ rather than M_W , the mass of the W boson. They can be calculated in perturbation theory. The combinations $c_{\pm} = c_1 \pm c_2$ have a multiplicative evolution under the renormalisation scale. To leading order, they are given by

$$c_{\pm}(m_b) = \left(\frac{\alpha_s(M_W)}{\alpha_s(m_b)} \right)^{a_{\pm}}, \quad a_- = -2a_+ = -\frac{12}{33 - 2n_f}, \quad (1.6)$$

where n_f is the number of active flavours ($n_f = 5$ in our case). This corresponds to $c_+ = 0.86$ and $c_- = 1.35$, with $\alpha_s(M_Z) = 0.117$.

Since the energy release in the decay of a b quark is large, it is possible to construct an OPE for the transition operator \mathbb{T} , in which it is expanded as a series of local operators with increasing dimension,⁴ whose coefficients contain inverse powers of the b quark mass (HQE).

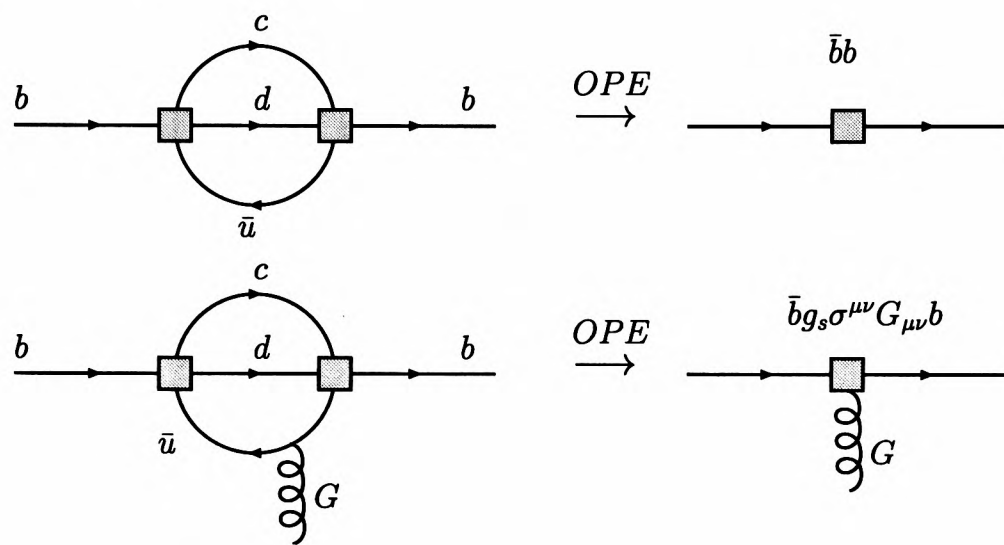


Figure 1.3: Top: Example of a contribution to the local dimension 3 operator $\bar{b}b$ (parton term). Bottom: Example of a contribution to the local dimension 5 operator $\bar{b}g_s\sigma^{\mu\nu}G_{\mu\nu}b$ (chromo-magnetic term).

The lowest dimensional term in this expansion will dominate in the limit $m_b \rightarrow \infty$; for beauty decays that is the dimension three operator $\bar{b}b$, which is shown in figure 1.3. The width for the decay of a beauty hadron H_b into an inclusive final state f is obtained by taking the expectation value of \mathbb{T} between the state H_b . Through order $1/m_b^2$ one finds:

$$\Gamma(H_b \rightarrow f) = \frac{G_F^2 m_b^5}{192\pi^3} \left\{ c_3(f) \frac{\langle H_b | \bar{b}b | H_b \rangle}{2M_{H_b}} + \frac{c_5(f)}{m_b^2} \frac{\langle H_b | \bar{b}g_s\sigma^{\mu\nu}G_{\mu\nu}b | H_b \rangle}{2M_{H_b}} + \mathcal{O}\left(\frac{1}{m_b^3}\right) \right\}, \quad (1.7)$$

⁴The term dimension of an operator is short for *mass dimension*. In natural units $\hbar = c = 1$ all dimensions are equivalent to the dimension of mass or inverse mass. The mass dimension of Dirac spinors is $\frac{3}{2}$, and 1 for scalar and vector fields. The dimension of a local operator built from these fields is simply the sum of the dimension of all its fields and coefficients.

where the dimensionless coefficients $c_i(f)$ include the relevant CKM matrix elements and the parton level characteristics of f (such as the ratios of the final state quark masses to m_b). The term containing the gluonic field strength tensor $G_{\mu\nu}$ describes the chromo-magnetic contribution. For semileptonic and non-leptonic decays, the coefficients $c_3(f)$ have been calculated at one-loop order, and the coefficients $c_5(f)$ at tree level [21].

In the next step, the forward matrix elements of the local operators in the OPE are systematically expanded in inverse powers of the b quark mass, using the heavy-quark effective theory (HQET) [24]. One finds

$$\frac{1}{2M_{H_b}} \langle H_b | \bar{b}b | H_b \rangle = 1 - \frac{\mu_\pi^2(H_b) - \mu_G^2(H_b)}{2m_b^2} + \mathcal{O}(1/m_b^3), \quad (1.8)$$

$$\frac{1}{2M_{H_b}} \langle H_b | \bar{b}g_s \sigma^{\mu\nu} G_{\mu\nu} b | H_b \rangle = 2\mu_G^2(H_b) + \mathcal{O}(1/m_b), \quad (1.9)$$

where $\mu_\pi^2(H_b)$ and $\mu_G^2(H_b)$ are parameters that describe the matrix elements of the kinetic-energy (π) and the chromo-magnetic (G) operators, respectively. The second term on the right hand side of equation 1.8 exactly reproduces the parton model spectator result in the limit $m_b \rightarrow \infty$, which attributes equal lifetimes to all hadrons of a given heavy flavour. The parameters $\mu_\pi^2(H_b)$ and $\mu_G^2(H_b)$ can be determined to some extent from the spectrum of heavy hadron states, from lattice QCD calculations or from QCD sum rules. For example for the B mesons μ_G^2 is given by the hyperfine splitting of the B^* and B masses [22]:

$$\mu_G^2(B) = \frac{3}{4}(M_{B^*}^2 - M_B^2) \approx 0.36 \text{ GeV}^2. \quad (1.10)$$

An existing analysis based on QCD sum rules [25] yields a numerical value for $\mu_\pi^2(B)$

$$\mu_\pi^2(B) \approx (0.54 \pm 0.12) \text{ GeV}^2, \quad (1.11)$$

whereas from an analysis of spectroscopy of heavy hadrons [26] gives

$$\mu_\pi^2(B) \approx (0.3 \pm 0.2) \text{ GeV}^2. \quad (1.12)$$

The most important aspect – qualitatively as well as conceptually – of the expression in equation 1.7 is contained in the element that is missing here: there are no non-perturbative contributions of order $1/m_b$ to fully integrated rates. The numerical impact of this fact is obvious: since the leading non-perturbative corrections arise only on the $1/m_b^2$ level, they fade away quickly with increasing heavy flavour quark mass. For beauty decays this is on the several percent level. Conceptually the absence of $1/m_b$ terms is more subtle and can be traced back to the conservation of colour flow [27].

To order $1/m_b^2$, the lifetime ratio for two beauty hadrons is given by

$$\begin{aligned} \frac{\tau(H_b^{(1)})}{\tau(H_b^{(2)})} &= 1 + \frac{\mu_\pi^2(H_b^{(1)}) - \mu_\pi^2(H_b^{(2)})}{2m_b^2} \\ &\quad + c_G \frac{\mu_G^2(H_b^{(1)}) - \mu_G^2(H_b^{(2)})}{2m_b^2} + \mathcal{O}\left(\frac{1}{m_b^3}\right), \end{aligned} \quad (1.13)$$

where $c_G \approx 1.2$ can be obtained using the results of [27], [28]. Assuming that in the case of the B_s^0 meson SU(3)-breaking effects in the values of the matrix elements are of order 20 %, we arrive at the following predictions:

$$\frac{\tau(B_u)}{\tau(B_d)} = 1 + \mathcal{O}(1/m_b^3) , \quad (1.14)$$

$$\frac{\tau(B_s)}{\tau(B_d)} = (1.00 \pm 0.01) + \mathcal{O}(1/m_b^3) . \quad (1.15)$$

Hard spectator effects manifest themselves first in the matrix elements of four-quark operators of dimension six, such as $(\bar{b}\Gamma_i q)(\bar{q}\Gamma_i b)$, where Γ_i denotes some combination of Dirac and colour matrices. Some examples of the corresponding contributions to the transition operator \mathbb{T} are shown in figure 1.4, before and after the heavy-quark expansion. Since these contributions arise from one-loop rather than two-loop diagrams, they receive a phase-space enhancement of order $16\pi^2$ relative to the other terms in the heavy-quark expansion.

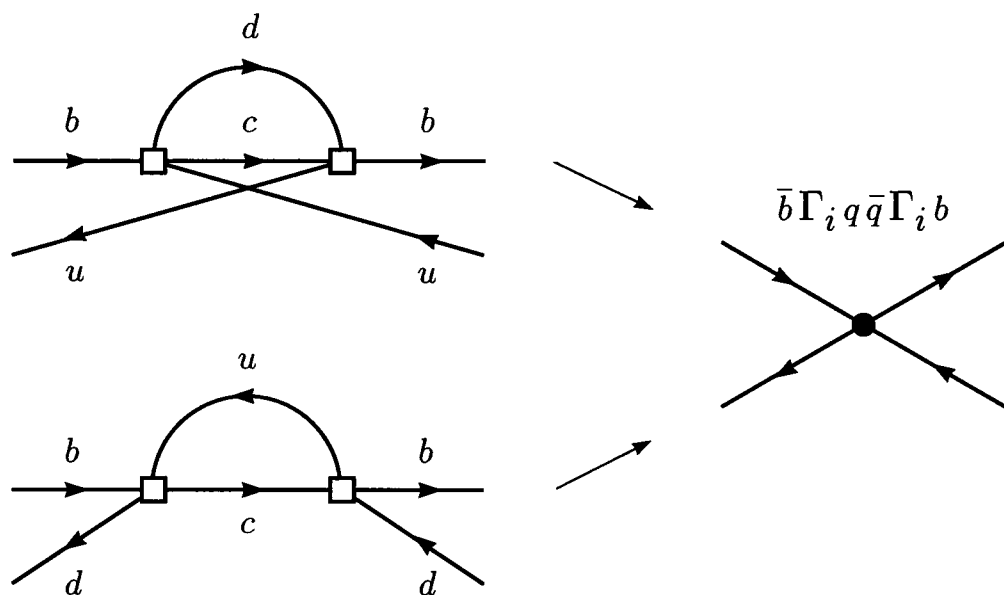


Figure 1.4: Spectator contributions to the transition operator \mathbb{T} (left), and the corresponding operator in the OPE (right). Here, Γ_i denotes some combination of Dirac and colour matrices.

These local four-quark operators of dimension six have been calculated in [21] and produce differences between the B meson lifetimes. For $\tau(B_u)/\tau(B_d)$ the spectator quark contribution is of order

$$\frac{\Gamma_{non-spect}(B)}{\Gamma_{spect}(B)} \approx \frac{2\pi f_{B_d}^2}{m_b^2} \approx 5 \% , \quad (1.16)$$

where f_B is the B_d meson decay constant. In the limit where SU(3)-breaking effects are neglected, the spectator contributions to the decay widths of B_s^0 and B_d^0 mesons are too similar to produce an observable lifetime difference. Therefore the ratio $\tau(B_s)/\tau(B_d)$ is dominated by SU(3)-breaking effects, which, even being on the order

of 20 % give an uncertainty on the ratio of only about 1 %. The current theoretical predictions therefore are

$$\begin{aligned}\tau(B_u)/\tau(B_d) &= 1 + 0.05 (f_{B_d}/(200 \text{ MeV}))^2 = 1.05 \pm 0.02, \\ \tau(B_s)/\tau(B_d) &= 1 \pm 0.01,\end{aligned}\tag{1.17}$$

with f_{B_d} being the B_d decay constant with current value of (200 ± 30) MeV from recent lattice calculations [29].

1.4 Mixing of Neutral Mesons

The standard formalism described here is model independent and general [20]. In general it can be applied to all four neutral meson states, although we make some assumptions driven by experimental facts, only typical for the B_s^0 system. In the following we drop the index s for convenience and just use B^0 to indicate the B_s^0 meson.

Mixing refers to transitions between the two flavour eigenstates $|B^0\rangle$ and $|\bar{B}^0\rangle$. They are produced by the strong interaction as **flavour eigenstates**.⁵ If only the strong and electromagnetic interactions existed, $|B^0\rangle$ and $|\bar{B}^0\rangle$ would be stable and form a particle-antiparticle pair with common mass m . But they live as eigenstates of the “whole theory”, including weak interaction and potential interactions beyond the Standard Model. These eigenstates are the **physical eigenstates** and have a definite mass and lifetime.⁶ Because of the weak interaction, $|B^0\rangle$ and $|\bar{B}^0\rangle$ decay. Moreover, neither electric-charge conservation nor any other conservation law respected by the weak interaction prevent $|B^0\rangle$ and $|\bar{B}^0\rangle$ from having both real and virtual transitions to common states. Such transitions allow $|\Delta B| = 2$ transitions $|B^0\rangle \leftrightarrow |\bar{B}^0\rangle$, where B is the bottom quantum number. As a consequence, $|B^0\rangle$ and $|\bar{B}^0\rangle$ oscillate between themselves before decaying.

A study of the time evolution of the B^0 states has to treat both physical eigenstates as one entity, because both eigenstates do not lose coherence [30] over the short distance they travel and will interfere. This can be done by using 2-dimensional vectors in either the physical eigenstate or the flavour eigenstate basis. In the Wigner-Weisskopf formalism⁷ [31], [32] the time evolution can be described with an effective Schrödinger equation:

$$i \frac{d}{dt} \begin{pmatrix} |B^0(t)\rangle \\ |\bar{B}^0(t)\rangle \end{pmatrix} = \begin{pmatrix} M_{11} - \frac{i}{2}\Gamma_{11} & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{12}^* - \frac{i}{2}\Gamma_{12}^* & M_{22} - \frac{i}{2}\Gamma_{22} \end{pmatrix} \begin{pmatrix} |B^0(t)\rangle \\ |\bar{B}^0(t)\rangle \end{pmatrix}.\tag{1.18}$$

⁵Sometimes also called **eigenstates of the strong interaction**. Since the strong interaction conserves flavour, they are flavour eigenstates as well.

⁶Thus justifying the naming **mass eigenstates**, or equivalently **lifetime eigenstates**. While lifetime eigenstates are the natural choice for the K system due to the huge lifetime difference between K_S and K_L , mass eigenstates are preferred in the B_d^0 system. In the B_s^0 system both types are equally common. In our lifetime analysis though we put more emphasis on the lifetime eigenstates.

⁷The Wigner-Weisskopf formalism is not exact, and there are tiny corrections to the exponential decay law at very short and very large times. These corrections though are too small to be noticeable in this analysis.

We see that the off-diagonal elements, induced by $|\Delta B| = 2$ transitions, mix the time evolution of the flavour eigenstates.

The *mass matrix* M and the *decay matrix* Γ are time-independent Hermitian matrices. The Hamiltonian $H = M - \frac{i}{2}\Gamma$ is thus not Hermitian. The resulting imaginary part of the energy eigenvalue leads to the expected exponential decay behaviour (equation 1.24). CPT invariance requires that

$$M_{11} = M_{22} \equiv m, \quad \Gamma_{11} = \Gamma_{22} \equiv \Gamma. \quad (1.19)$$

The matrices M and Γ are given, in second-order perturbation theory, by sums over intermediate states n :

$$\begin{aligned} M_{ij} &= m \delta_{ij} + \langle i | \mathcal{H}_W | j \rangle + \sum_n \mathcal{P} \frac{\langle i | \mathcal{H}_W | n \rangle \langle n | \mathcal{H}_W | j \rangle}{m - E_n}, \\ \Gamma_{ij} &= 2\pi \sum_n \delta(m - E_n) \langle i | \mathcal{H}_W | n \rangle \langle n | \mathcal{H}_W | j \rangle. \end{aligned} \quad (1.20)$$

The indices i, j can be B^0 or \bar{B}^0 , and the operator \mathcal{P} projects out the principal part.⁸ Note that the intermediate states contributing to M are virtual, while the ones contributing to Γ are physical common states. For this reason M_{12} is called the **dispersive part** and Γ_{12} the **absorptive part** of the transition amplitude $|B^0\rangle \rightarrow |\bar{B}^0\rangle$.

The mass eigenstates are defined as the eigenvectors of $M - \frac{i}{2}\Gamma$. They can be expressed as linear combination of the flavour eigenstates as

$$\begin{aligned} |B_L\rangle &= p \cdot |B^0\rangle + q \cdot |\bar{B}^0\rangle, \\ |B_H\rangle &= p \cdot |B^0\rangle - q \cdot |\bar{B}^0\rangle, \end{aligned} \quad (1.21)$$

with complex coefficients p, q , satisfying $|p|^2 + |q|^2 = 1$. $|B_L\rangle$ denotes by definition the lighter eigenstate and is in general not orthogonal to the heavier eigenstate $|B_H\rangle$. The corresponding eigenvalues are

$$\lambda_{L,H} = (M - \frac{i}{2}\Gamma) \pm \frac{q}{p} (M_{12} - \frac{i}{2}\Gamma_{12}), \quad (1.22)$$

where

$$\frac{q}{p} = \sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}}. \quad (1.23)$$

We choose a convention where $\text{Re}(q/p) > 0$ and $\text{CP}|B^0\rangle = -|\bar{B}^0\rangle$.

The time evolution in this new basis is governed by the eigenvalues $M_{L,H} - \frac{i}{2}\Gamma_{L,H}$ and simply reads

$$\begin{aligned} |B_L(t)\rangle &= e^{-(iM_L + \frac{\Gamma_L}{2})t} |B_L\rangle, \\ |B_H(t)\rangle &= e^{-(iM_H + \frac{\Gamma_H}{2})t} |B_H\rangle, \end{aligned} \quad (1.24)$$

⁸See standard textbooks on Complex Analysis [33] for details.

where the notation without the time argument denotes the mass eigenstates at time $t = 0$. The following definitions for the average mass and width and the mass and width differences are common:

$$\begin{aligned} m &\equiv \frac{M_H + M_L}{2} = M_{11} , & \Gamma &\equiv \frac{\Gamma_L + \Gamma_H}{2} = \Gamma_{11} \\ \Delta m &\equiv M_H - M_L , & \Delta\Gamma &\equiv \Gamma_L - \Gamma_H . \end{aligned} \quad (1.25)$$

Δm is positive by definition. The sign of $\Delta\Gamma$ has to be measured from experiment. Nevertheless, with our convention the Standard Model prediction turns out to be positive as we will see later.

1.5 Calculation of $\Delta\Gamma_s$

In a given theory one can calculate the off-diagonal elements M_{12} and Γ_{12} from diagrams which change that bottomness quantum number B by 2 ($|\Delta B| = 2$). We can then solve for the eigenvalues and eigenvectors of $M - \frac{i}{2}\Gamma$ and express Δm , $\Delta\Gamma$ and q/p in terms of M_{12} and Γ_{12} . The relative phase ϕ between M_{12} and Γ_{12} appears in many observables related to B mixing:

$$\phi \equiv \arg\left(-\frac{M_{12}}{\Gamma_{12}}\right) . \quad (1.26)$$

In the Standard Model for the B_s^0 system $\phi = 2\beta_s \approx 2\lambda^2\eta \approx 0.03$. It is possible to express Δm , $\Delta\Gamma$ and q/p in terms of $|M_{12}|$, $|\Gamma_{12}|$ and ϕ [20].

A simple, approximate form of the solution can be derived when using the approximations

$$|\Gamma_{12}| \ll |M_{12}| , \text{ and } \Delta\Gamma \ll \Delta m , \quad (1.27)$$

which hold for both the B_d^0 and the B_s^0 systems. First we note that $|\Gamma_{12}| \leq \Gamma$, because Γ_{12} comes from the decays into final states common to B^0 and \bar{B}^0 ; a small number compared to the total number of decay modes. For the B_s^0 meson the experimental lower bound on Δm_s guarantees $\Gamma_s \ll \Delta m_s$. Hence $\Gamma_{12}^s \ll \Delta m_s$, which implies equations 1.27. For the B_d^0 meson, experiment gives $\Delta m_d \approx 0.75 \Gamma_d$. The Standard Model predicts $|\Gamma_{12}^d|/\Gamma_d = \mathcal{O}(1\%)$, but Γ_{12}^d stems only from CKM-suppressed decay channels common to B_d^0 and \bar{B}_d^0 and could therefore be affected by new physics. New decay channels would, however, also increase Γ_d and potentially conflict with the precisely measured semileptonic branching ratio. A conservative estimate is $|\Gamma_{12}^d|/\Gamma_d < 10\%$ and justifies the expansion in Γ_{12}/M_{12} and $\Delta\Gamma/\Delta m$ as well. Neglecting terms of $\mathcal{O}(|\Gamma_{12}/M_{12}|^2)$ we find

$$\Delta m \approx 2|M_{12}| , \quad (1.28)$$

$$\Delta\Gamma \approx 2|\Gamma_{12}| \cos \phi , \quad (1.29)$$

$$\frac{q}{p} \approx -e^{-i\phi_M} \left\{ 1 - \frac{a}{2} \right\} , \quad (1.30)$$

where parameter a is defined as

$$a = \text{Im} \frac{\Gamma_{12}}{M_{12}} = \left| \frac{\Gamma_{12}}{M_{12}} \right| \sin \phi \quad (1.31)$$

and ϕ_M is the phase of M_{12} ,

$$M_{12} = |M_{12}| e^{i\phi_M} . \quad (1.32)$$

The deviation of $|q/p|^2$ from 1 – namely parameter a – describes CP violation in B mixing. It can only occur if $M_{12} \neq 0$, $\Gamma_{12} \neq 0$ and if the so-called CP violating phase ϕ between M_{12} and Γ_{12} is different from 0 or π . In the Standard Model ϕ for the B_s^0 system turns out to be small.

1.6 Standard Model Expectation for $\Delta\Gamma_s$

CP violation in mixing for the B_s^0 system is small in the Standard Model. There, M_{12} arises from the two-top-quark box diagram, while Γ_{12} is dominated by the decays $b \rightarrow \bar{c}cs$. Both M_{12} and Γ_{12} involve mainly the last two families, hence they cannot exhibit CP violation, since no CP violating phase comes in. In order to get a phase one must introduce the first family, via the contributions to Γ_{12} of the suppressed decays $b \rightarrow \bar{u}cs, \bar{c}us, \bar{u}us$. For this reason we neglect from now on CP violation in mixing in the B_s^0 system.

In calculating $\Delta\Gamma_s$ there are two different approaches: a **parton model calculation** [34] and a **sum over exclusive decay modes** [35]. It is important to realise for the parton model calculation that in the integral over final states, quark states rather than hadron states are used. The underlying assumption is called **parton-hadron duality** [23] which connects quantities evaluated on the quark-gluon level to the (observable) world of hadrons. For us it means that the effect of hadronisation cancels out in the sum over a sufficient number of hadronic final states. A parton model estimate of the total width for B mesons is justified since the b quark is heavy enough and there are many exclusive final states. At present duality is tested in various inclusive observables in B decays and no experimental evidence for duality violation in B meson widths has yet been found [23]. On the other hand, duality violations in the width difference $\Delta\Gamma_s$ may be larger than in the average width Γ_s , because we are dealing with a relatively small number of **common decay modes** to B^0 and \bar{B}^0 . For this reason the second approach, the sum over exclusive final states was performed [35] and the numerical result for $\Delta\Gamma_s$ is found to be consistent with the value from parton model approaches [34].

We will first briefly discuss the sum over exclusive final states, because it is more illustrating than the parton model calculation.

Summing Over Exclusive Final States

There are two main types of spectator diagrams contributing to the decays into common decay modes: **Colour-allowed** (figure 1.5) and **colour-suppressed** diagrams

(figure 1.6). In the colour-suppressed case, the colours of the quarks emerging from the W boson have to match the colours of the other quarks to form colourless hadrons. For that reason the colour-suppressed modes are suppressed by a factor of 0.20 relative to the colour-allowed ones. There is also a contribution from exchange diagrams

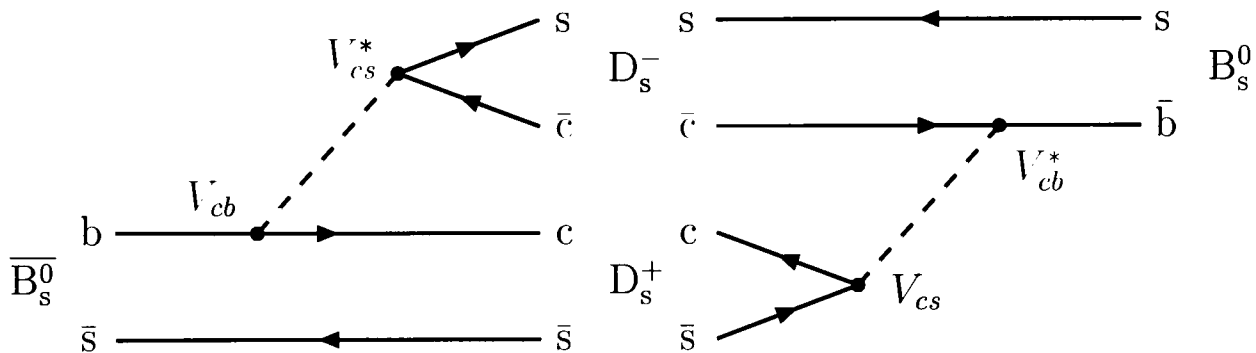


Figure 1.5: Colour-allowed spectator diagram. The left diagram should be read from left to right, the right diagram from right to left.

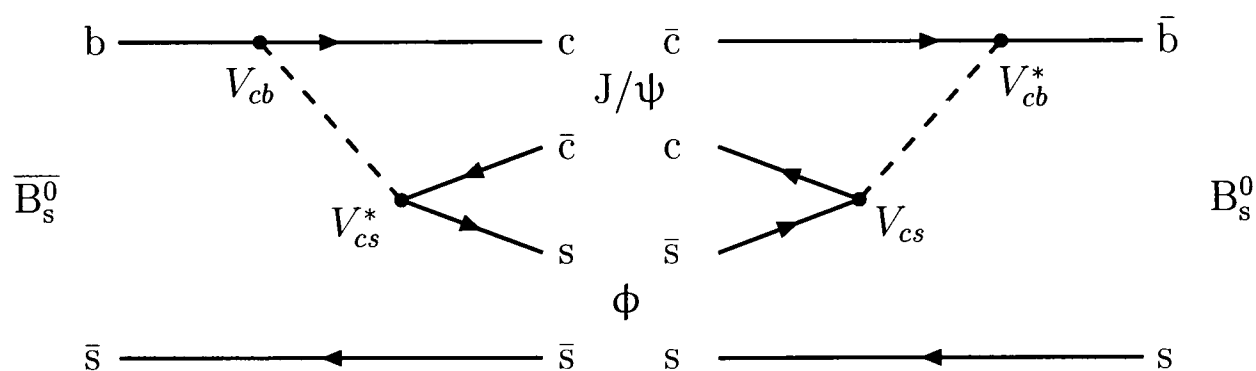


Figure 1.6: Colour-suppressed spectator diagram. The left diagram should be read from left to right, the right diagram from right to left.

(Figure 1.7), which are very much suppressed by colour and form factor effects.

As we neglect CP violation in the mixing of B_s^0 , the CP eigenstates B_+ , B_- with CP eigenvalues $+1$, -1 , are equal to the mass eigenstates B_H and B_L .⁹

$$\begin{aligned} |B_+\rangle &= \frac{1}{\sqrt{2}}(|B_s^0\rangle - |\overline{B}_s^0\rangle) = |B_H\rangle \\ |B_-\rangle &= \frac{1}{\sqrt{2}}(|B_s^0\rangle + |\overline{B}_s^0\rangle) = |B_L\rangle. \end{aligned} \quad (1.33)$$

From the definition $\Delta\Gamma_s \equiv \Gamma_L - \Gamma_H = \Gamma(B_+) - \Gamma(B_-)$ we see that CP even decay modes increase $\Delta\Gamma_s$, whereas CP odd decay modes decrease $\Delta\Gamma_s$. It is instructive to classify the main decay modes according to their CP value. Table 1.3 lists the main contributing decay modes and their contribution to $\Delta\Gamma_s/\Gamma_s$ according to theoretical calculations by [35]. The colour-allowed modes clearly dominate and all the

⁹We use the convention $CP|B_s^0\rangle = -|\overline{B}_s^0\rangle$.

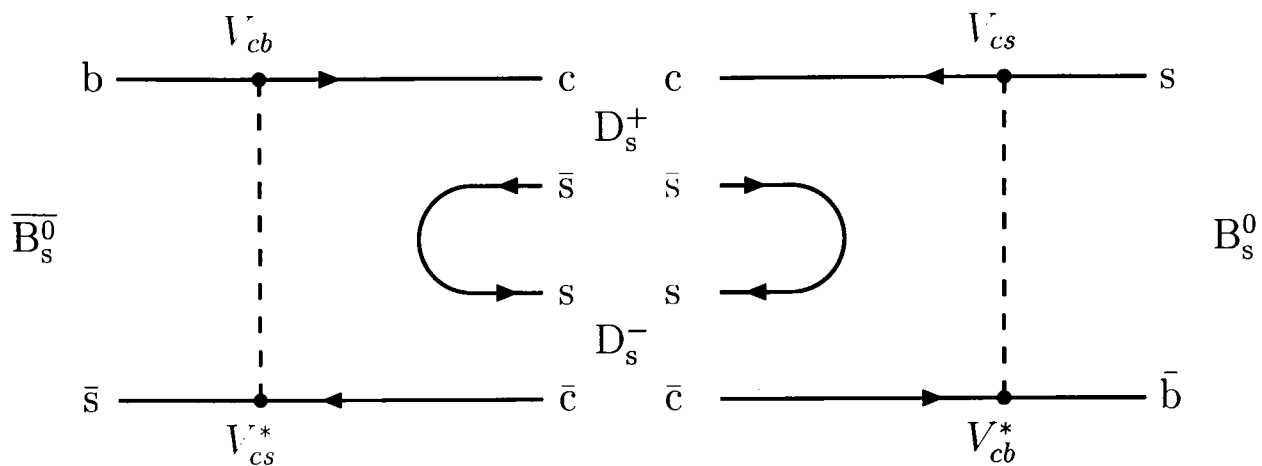


Figure 1.7: Exchange diagrams. The left diagram should be read from left to right, the right diagram from right to left.

contributions add up to a value of

$$\left. \frac{\Delta\Gamma_s}{\Gamma_s} \right|_{\text{Exclusive}} \approx 0.15 . \quad (1.34)$$

We note, however, that a significant fraction of hadronic modes, such as the baryonic mode $B_s \rightarrow \Xi_c \bar{\Xi}_c$, as well as penguin diagrams were not considered in order to get this number. Since in the Standard Model $\Delta\Gamma_s/\Gamma_s$ turns out to be positive, we have the following connection between mass, lifetime and CP eigenstates of the B_s^0 :

$$\begin{aligned} |B_H\rangle &= |B_+\rangle = |B_{\text{Short}}\rangle , \\ |B_L\rangle &= |B_-\rangle = |B_{\text{Long}}\rangle . \end{aligned} \quad (1.35)$$

$\Delta\Gamma_s$ in the Parton Model

Neglecting CP violation in mixing, the prediction of the width difference $\Delta\Gamma \approx 2|\Gamma_{12}|$ requires the calculation of $|\Gamma_{12}|$. Γ_{12} is determined from the absorptive part of the $|\Delta B| = 2$ transition amplitude. It receives contributions from all final states which are common to B_s^0 and \overline{B}_s^0 . The most dominant contributor to Γ_{12} is the CKM-favoured $b \rightarrow c\bar{c}s$ transition, with the CKM-suppressed $b \rightarrow c\bar{u}s$, $u\bar{c}s$, $u\bar{u}s$ processes playing a minor role.

The mass and width difference is determined by the familiar box diagram that give rise to an effective $|\Delta B| = 2$ Hamiltonian. On distance scales larger than $1/M_W$, but still smaller than $1/m_b$, this effective Hamiltonian contains a local $|\Delta B| = 2$ interaction as well as a bi-local part constructed from two local $|\Delta B| = 1$ transitions.

The *mass difference* is given by the real part of the box diagrams and is dominated by the top quark contribution. For this reason, M_{12} is generated by an interaction that is local already on scales greater than $1/M_W$ and theoretically well under control. The long-distance contribution is parameterised by the matrix element of a single four-quark operator between B_s^0 and \overline{B}_s^0 states. Corrections to this result are suppressed by powers of m_b^2/M_W^2 and completely irrelevant for all practical purposes.

Decay mode	Partial wave	Contribution to $\Delta\Gamma_s/\Gamma_s$ in %
$B_{s,+} \rightarrow D_s \bar{D}_s$	S	+3.13
$B_{s,+} \rightarrow D_s^* \bar{D}_s^*$	S+D	+7.04
$B_{s,+} \rightarrow D_s \bar{D}_s^* + \bar{D}_s D_s^*$	P	+4.40
$B_{s,-} \rightarrow D_s \bar{D}_s^* + \bar{D}_s D_s^*$	P	-0.02
$B_{s,-} \rightarrow D_s^* \bar{D}_s^*$	P	-0.19
$B_{s,+} \rightarrow \eta_c \eta$	S	+0.13
$B_{s,+} \rightarrow \eta_c \eta'$	S	+0.01
$B_{s,+} \rightarrow J/\psi \eta$	P	+0.06
$B_{s,+} \rightarrow J/\psi \eta'$	P	+0.00
$B_{s,+} \rightarrow \eta_c \phi$	P	+0.05
$B_{s,+} \rightarrow J/\psi \phi$	S+D	+0.31
$B_{s,-} \rightarrow J/\psi \phi$	P	-0.01
$B_{s,+} \rightarrow \psi' \eta$	P	+0.02
$B_{s,+} \rightarrow \psi' \eta'$	P	+0.00
$B_{s,+} \rightarrow \psi' \phi$	S+D	+0.21
$B_{s,-} \rightarrow \psi' \phi$	P	-0.01
Total		+0.15

Table 1.3: Contributions to the different ground state decay modes to $\Delta\Gamma_s/\Gamma_s$ (in %).

The *width difference* is given by the imaginary part of the box diagram and determined by real intermediate states, which correspond to common decay products of B_s^0 and \bar{B}_s^0 , so that only the bi-local part of the $|\Delta B| = 2$ Hamiltonian can contribute. The presence of long-lived (on hadronic scales) intermediate states would normally preclude a short-distance treatment of the width difference as indeed it does for neutral kaons. But for bottom mesons, the b quark mass m_b provides an additional short-distance scale that leads to a large energy release (compared to Λ_{QCD}) into the intermediate states. Thus, at typical hadronic distances greater than $1/m_b$, the decay is again a local process.

In a first step we apply the OPE to the W-mediated $b \rightarrow c\bar{c}s$ decay amplitude:

$$\begin{array}{c}
 b_i \xrightarrow{\quad} \bullet \xrightarrow{\quad} c_i \\
 \bar{c}_j \xleftarrow{\quad} \bullet \xleftarrow{\quad} \bar{s}_j
 \end{array}
 +
 \begin{array}{c}
 b_i \xrightarrow{\quad} \bullet \xrightarrow{\quad} c_j \\
 \bar{c}_i \xleftarrow{\quad} \bullet \xleftarrow{\quad} \bar{s}_j
 \end{array}
 \quad (1.36)$$

The diagram on the right side stands for the four different diagrams with gluon exchange between any two quark lines. An OPE then results in a sum of two local $|\Delta B| = 1$ operators Q_1 and Q_2 , with different colour structure, multiplied by short-distance Wilson coefficients C_1 and C_2 :

$$\begin{aligned}
 \mathcal{H}_{eff} &= \frac{G_F}{\sqrt{2}} V_{cb} V_{cs}^* \left\{ C_2 \cdot \begin{array}{c} b_i \quad c_i \\ \nearrow \quad \searrow \\ \text{---} Q_2 \text{---} \\ \nwarrow \quad \swarrow \\ \bar{c}_j \quad \bar{s}_j \end{array} + C_1 \cdot \begin{array}{c} b_i \quad c_j \\ \nearrow \quad \searrow \\ \text{---} Q_1 \text{---} \\ \nwarrow \quad \swarrow \\ \bar{c}_i \quad \bar{s}_j \end{array} \right\} \\
 &= \frac{G_F}{\sqrt{2}} V_{cb} V_{cs}^* \left\{ C_2 \cdot (\bar{c}_i b_i)_{V-A} (\bar{s}_j c_j)_{V-A} + C_1 \cdot (\bar{c}_j b_i)_{V-A} (\bar{s}_j c_i)_{V-A} \right\}.
 \end{aligned} \tag{1.37}$$

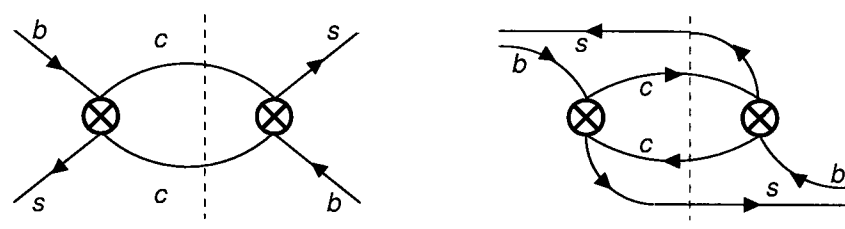
The Wilson coefficients are determined in such a way that the Standard Model amplitude is reproduced by $\langle c\bar{c}s | \mathcal{H}_{eff} | b \rangle$ up to terms of order m_b^2/M_W^2 . They contain the short-distance physics associated with the scale M_W , and QCD corrections can be computed in perturbation theory. The Wilson coefficients are then evolved to the next scale $\mu = \mathcal{O}(m_b)$ using the renormalisation group evolution. In the leading log approximation we get an infinite series of logarithms of the form $\alpha_s^n \ln^n(\mu/M_W)$ with $n = 0, 1, \dots$. The next-to-leading order (NLO) corrections to the coefficients comprise terms of order $\alpha_s^{n+1} \ln^n(\mu/M_W)$ and have been calculated in [36]. There are also penguin operators in the effective Hamiltonian, whose coefficients are very small [37],[38].

The width difference is related to \mathcal{H}_{eff} by the optical theorem:

$$\Delta\Gamma_s = 2|\Gamma_{12}| = \left| -\frac{1}{M_{B_s}} \text{Im} \langle \bar{B}_s | i \int d^4x \mathcal{T}(\mathcal{H}_{eff}(x), \mathcal{H}_{eff}(0)) | B_s \rangle \right|, \tag{1.38}$$

where the time-ordering operator is denoted by \mathcal{T} .

The main leading-order diagrams at the m_b scale are



$$\tag{1.39}$$

where the vertices correspond to the $|\Delta B| = 1$ operators Q_1, Q_2 , given in equation 1.37. The dashed lines show the two ways of cutting through a box diagram to form final states. The dominant contribution comes from the **spectator diagram** on the right, whereas the **weak annihilation diagram**¹⁰ on the left is highly suppressed due to colour and form factor effects.

Now we apply the HQE, an expansion in $1/m_b$ to describe Γ_{12} in terms of matrix elements of local $|\Delta B| = 2$ operators. Pictorially this means that we shrink the internal $c\bar{c}$ loop to a point. We get to leading order in α_s

$$\begin{aligned}
 &\left| \text{Im} \langle \bar{B}_s | i \int d^4x \mathcal{T}(\mathcal{H}_{eff}(x), \mathcal{H}_{eff}(0)) | B_s \rangle \right| \\
 &= -\frac{G_F^2 m_b^2}{12\pi} |V_{cb}^* V_{cs}|^2
 \end{aligned}$$

¹⁰Sometimes called **exchange diagram**.

$$\cdot \left\{ F \left(\frac{m_c^2}{m_b^2} \right) \langle \bar{B}_s | Q | B_s \rangle + F_S \left(\frac{m_c^2}{m_b^2} \right) \langle \bar{B}_s | Q_S | B_s \rangle \right\} \cdot \left[1 + \mathcal{O} \left(\frac{1}{m_b} \right) \right] \quad (1.40)$$

$$= -\frac{G_F^2 m_b^2}{12\pi} |V_{cb}^* V_{cs}|^2$$

$$\cdot \left\{ F \cdot \begin{array}{c} b_i \quad \bar{s}_j \\ \nearrow \quad \searrow \\ \boxed{Q} \\ \nwarrow \quad \swarrow \\ \bar{s}_i \quad b_j \end{array} + F_S \cdot \begin{array}{c} b_i \quad \bar{s}_j \\ \nearrow \quad \searrow \\ \boxed{Q_S} \\ \nwarrow \quad \swarrow \\ \bar{s}_i \quad b_j \end{array} \right\} . \quad (1.41)$$

The local dimension-6 operators Q and Q_S are obtained, which turn out to have a vector-minus-axial (V–A) and a scalar-minus-pseudoscalar (P–S) structure:

$$Q = (\bar{s}_i b_i)_{V-A} (\bar{s}_j b_j)_{V-A} , \quad Q_S = (\bar{s}_i b_i)_{S-P} (\bar{s}_j b_j)_{S-P} . \quad (1.42)$$

The new Wilson coefficients F and F_S also depend on the charm quark mass m_c , which is formally treated as a hard scale of order m_b , since $m_c \gg \Lambda_{QCD}$. They are independent of the QCD binding forces in the external B_s^0 states and can be calculated in perturbation theory at the parton level. The non-perturbative long-distance QCD effects completely reside in the hadronic matrix elements of Q and Q_S , which are usually parameterised as

$$\begin{aligned} \langle \bar{B}_s | Q | B_s \rangle &= \frac{8}{3} f_{B_s}^2 M_{B_s}^2 B , \\ \langle \bar{B}_s | Q_S | B_s \rangle &= -\frac{5}{3} f_{B_s}^2 \frac{M_{B_s}^2}{m_b + m_s} B_S . \end{aligned} \quad (1.43)$$

Here M_{B_s} and f_{B_s} are the mass and decay constant of the B_s^0 meson, and B and B_S are the bag parameters of the B_d^0 and B_s^0 . The quark masses m_b and m_s are defined in the \overline{MS} scheme [39].

The Wilson coefficients F and F_S have been calculated to leading-order (LO) in [34], and to next-to-leading order (NLO) in [38]:

	LO	NLO
$F(0.085)$	0.057	0.045
$F_S(0.085)$	−1.513	−1.045

We note that the NLO contribution to F_S is large (about 30% of the LO result), and it decreases the Standard Model value of the width difference considerably. The $\mathcal{O}(1/m_b)$ corrections have been performed in [37] and are of order -8% . They are large and decrease the width difference as well. Following [40] the NLO prediction becomes

$$\left. \frac{\Delta\Gamma_s}{\Gamma_s} \right|_{Parton} = \left(\frac{f_{B_s}}{245 \text{ MeV}} \right)^2 \{ (0.234 \pm 0.035) B_S(m_b) - 0.080 \pm 0.020 \} . \quad (1.44)$$

Table 1.4 shows the used input values as well as current theoretical predictions for the bag parameters and decay constant. As F is small the uncertainty on B is irrelevant.

Quantity	Value	Comment
$m_b(m_b) + m_s(m_b)$	4.3 GeV	$\overline{\text{MS}}$ scheme
m_c^2/m_b^2	0.085	
f_{B_s}	(245 ± 30) MeV	unquenched lattice QCD [41]
$B_S(m_b)$	0.87 ± 0.09	quenched lattice QCD [42]
$B(m_b)$	0.84 ± 0.08	quenched lattice QCD [42]

Table 1.4: Theoretical input values used for evaluating $\Delta\Gamma_s/\Gamma_s$ in the Standard Model.

With these numbers the value becomes

$$\left. \frac{\Delta\Gamma_s}{\Gamma_s} \right|_{\text{Parton}} = 0.12 \pm 0.06, \quad (1.45)$$

where the errors on B_S and f_{B_s} have conservatively been added linearly.

The main theoretical uncertainty comes from the *residual scale dependence*, which comes from effects in the matching of the NLO Wilson coefficients to the bag parameters computed on the lattice. The second major source of uncertainty comes from the $\mathcal{O}(1/m_b)$ corrections. For a concise discussion of the different sources of theoretical uncertainties see [39]. The authors state that for the future it appears that “Despite (or because of?) extensive work on radiative and $1/m_b$ corrections the theoretical prediction of $\Delta\Gamma_s/\Gamma_s$ remains rather uncertain. This is due to an unfortunate conspiracy of negative corrections at next-to-leading order in α_s and in the heavy quark expansion.”

1.7 New Physics Contributions to $\Delta\Gamma_s$

Various theories of flavour physics suggest that new physics effects are more likely to appear in processes involving the heavy generations. The B_s^0 system is unique in that it is the only neutral meson that is expected to exhibit mixing and does not involve first generation quarks.

The existence of new physics may modify the low-energy effective Hamiltonian governing B physics in several ways: (i) via contributions to the Wilson coefficients of the Standard Model operators, (ii) by generating new operators, or (iii) through the presence of new CP violating phases. These effects may originate from new interactions in tree-level decays or from virtual exchange of new physics in loop-mediated processes. The scale of new physics is expected to be large compared to M_W , and hence it is generally anticipated that additional tree-level contributions to B decays are suppressed, and to a very good approximation all decay amplitudes are given by the Standard Model. However, large new contributions may be present in loop processes, making mixing a fertile ground to reveal the influence of new interactions.

When new physics contributes comparably to or dominates over the Standard Model contribution to the B_s^0 mixing, CP may be significantly violated. New CP

violating contributions to mixing always reduce $\Delta\Gamma_s$ relative to the Standard Model prediction, as we can see from equation 1.29. Under the reasonable assumption that new physics does not significantly affect the leading decay process, Γ_{12} arises from $\Gamma(b \rightarrow c\bar{c}s)$. Consequently, the angle ϕ is the phase difference between the total mixing amplitude and the $b \rightarrow c\bar{c}s$ decay amplitude. In the Standard Model $\phi \approx 0$, and then $\cos\phi = 1$ to high accuracy. With new physics, however, new phases could be present, leading to $\cos\phi < 1$, and therefore to a reduced value of $\Delta\Gamma_s$.

The reduction of $\Delta\Gamma_s$ can be understood intuitively as follows. In the absence of CP violation, the two mass eigenstates are also CP eigenstates. The large $\Delta\Gamma_s$ is an indication that most of the $b \rightarrow c\bar{c}s$ decays are into CP even final states. With CP violation, in the basis where the $b \rightarrow c\bar{c}s$ amplitude is real, the mass eigenstates are no longer approximate CP eigenstates. Then, both mass eigenstates decay into CP even final states. Consequently, $\Delta\Gamma_s$ is reduced.

An experimental value of $\Delta\Gamma_s$ which is significantly lower than the Standard Model prediction would be a hint to new physics. The large theoretical uncertainty on the prediction of $\Delta\Gamma_s$ though makes it unlikely that new physics will be observed first in the width difference, for new physics contributions to the B_s^0 mixing phase are likely to have even stronger effects on time-dependent CP asymmetries, such as for example in the decay $B_s^0 \rightarrow J/\psi \phi$, and consequently will be observed there first.

1.8 Time Evolution of Untagged B Mesons

As time evolution for the flavour eigenstates we get from equations 1.18 and 1.24:

$$\begin{aligned} |B^0(t)\rangle &= g_+(t)|B^0\rangle + \frac{q}{p}g_-(t)|\bar{B}^0\rangle, \\ |\bar{B}^0(t)\rangle &= \frac{p}{q}g_-(t)|B^0\rangle + g_+(t)|\bar{B}^0\rangle, \end{aligned} \quad (1.46)$$

with

$$\begin{aligned} g_+(t) &= e^{-imt}e^{-\Gamma t/2} \left\{ \cosh\frac{\Delta\Gamma t}{4}\cos\frac{\Delta m t}{2} - i \sinh\frac{\Delta\Gamma t}{4}\sin\frac{\Delta m t}{2} \right\}, \\ g_-(t) &= e^{-imt}e^{-\Gamma t/2} \left\{ -\sinh\frac{\Delta\Gamma t}{4}\cos\frac{\Delta m t}{2} + i \cosh\frac{\Delta\Gamma t}{4}\sin\frac{\Delta m t}{2} \right\}. \end{aligned} \quad (1.47)$$

These equations describe the flavour oscillation in B mixing, with *oscillation frequency* Δm and a *modulation term* containing $\Delta\Gamma$. A non-zero value for $\Delta\Gamma$ will ensure that the “*same flavour coefficient*” $g_+(t)$ will never vanish, and the “*opposite flavour coefficient*” $g_-(t)$ will only be nought at $t = 0$. Therefore an initially produced B^0 will never turn into pure \bar{B}^0 or back into pure B^0 .

The time-dependent decay rates of an initially flavour tagged B meson into some final state is defined as follows:

$$\begin{aligned} \Gamma(B^0(t) \rightarrow f) &= \mathcal{N}_f |\langle f|B^0(t)\rangle|^2 \\ \Gamma(\bar{B}^0(t) \rightarrow f) &= \mathcal{N}_f |\langle f|\bar{B}^0(t)\rangle|^2, \end{aligned} \quad (1.48)$$

with \mathcal{N}_f being a time-independent normalisation factor. To calculate $\Gamma(B^0(t) \rightarrow f)$ we introduce the two decay amplitudes

$$A_f = \langle f | B^0 \rangle, \quad \bar{A}_f = \langle f | \bar{B}^0 \rangle. \quad (1.49)$$

The interference term λ_f defined as

$$\lambda_f = \frac{q}{p} \frac{\bar{A}_f}{A_f} \approx -e^{-i\phi_M} \frac{\bar{A}_f}{A_f} \left[1 - \frac{a}{2} \right] \quad (1.50)$$

plays an important role in CP asymmetries and other observables in B mixing.

Untagged data samples do not require the knowledge of the flavour at production. Because the two eigenstates have different lifetimes, suitably long times can be chosen where the longer lived state is highly enriched. Time is the tag here, in analogy to the neutral kaons.

The untagged decay rate for the decay $B^0/\bar{B}^0 \rightarrow f$ into a final state f is given by

$$\begin{aligned} \Gamma_f(t) &= \Gamma(B^0(t) \rightarrow f) + \Gamma(\bar{B}^0(t) \rightarrow f) \\ &= \mathcal{N}_f |A_f|^2 \left\{ (1 + |\lambda_f|^2) (e^{-\Gamma_L t} + e^{-\Gamma_H t}) + 2\text{Re}\lambda_f (e^{-\Gamma_L t} - e^{-\Gamma_H t}) \right\} \end{aligned} \quad (1.51)$$

$$\Gamma_{\bar{f}}(t) = \mathcal{N}_f |\bar{A}_f|^2 \left\{ (1 + |\lambda_f|^2) (e^{-\Gamma_L t} + e^{-\Gamma_H t}) + 2\text{Re}\bar{\lambda}_f (e^{-\Gamma_L t} - e^{-\Gamma_H t}) \right\} \quad (1.52)$$

where we use $|\lambda_f| = |\bar{\lambda}_f|$ and $\Gamma(B^0 \rightarrow f) = \Gamma(\bar{B}^0 \rightarrow \bar{f})$. We see that the time-dependent oscillations with $\Delta m t$ cancel in untagged data samples. Also the equation reveals its usefulness for the determination of $\Delta\Gamma$. Two limits are of particular interest. For the decay modes which are flavour specific, $\lambda_f = \bar{\lambda}_f = 0$. This applies to semileptonic decays, and to a very good approximation to $b \rightarrow c\bar{u}d$. We get

$$\Gamma_f(t) = \Gamma_{\bar{f}}(t) = \mathcal{N}_f |A_f|^2 \{ e^{-\Gamma_L t} + e^{-\Gamma_H t} \}. \quad (1.53)$$

For a final state that is a CP eigenstate¹¹ $\lambda_f^* = \bar{\lambda}_f$ with $|\lambda_f| = 1$, and

$$\Gamma_f(t) = \mathcal{N}_f |A_f|^2 2 \left\{ (e^{-\Gamma_L t} + e^{-\Gamma_H t}) + \text{Re}\lambda_f (e^{-\Gamma_L t} - e^{-\Gamma_H t}) \right\}. \quad (1.54)$$

1.9 The Decay Mode $B_s^0 \rightarrow J/\psi \phi$

The decay of the B_s^0 into the two vector particles J/ψ and ϕ has a very small branching ratio of $(9.3 \pm 3.3) \times 10^{-4} \%$ [12]. In this thesis we use this decay mode with the subsequent decays $J/\psi \rightarrow \mu^+ \mu^-$ (branching ratio $(5.88 \pm 0.10) \%$) and $\phi \rightarrow K^+ K^-$ (branching ratio $(49.2 \pm 0.6) \%$). The two muons from the J/ψ can be used to trigger on the event in a hadron collider environment, where the $b\bar{b}$ cross-section is high enough to compensate for the small branching ratio of $B_s^0 \rightarrow J/\psi \phi$. Also, since its final state particles are all charged particles, the momentum of the B_s^0 can be fully reconstructed with standard spectrometer techniques, as we will see in chapter 4.

¹¹Or an incoherent mixture of CP even and odd eigenstates, like $B_s^0 \rightarrow J/\psi \phi$, as it will turn out in section 1.9.

To leading order the decay $B_s^0 \rightarrow J/\psi \phi$ is described by the colour-suppressed diagram shown in figure 1.6. Since both decay products, J/ψ and ϕ , are CP eigenstates, the final state is a CP eigenstate as well, in general an incoherent mixture of eigenstates with CP eigenvalue $+1$ and -1 . The CP even fraction f_{CP+} has been measured with rather poor precision to be $f_{CP+} = (0.79 \pm 0.19)$ [43].

The CP quantum numbers of the decay mode can be deduced from the orbital angular momentum quantum number L between its decay products J/ψ and ϕ . Since both the J/ψ and the ϕ are vector particles with spin one, the total spin can be $S = |\mathbf{S}_1 \oplus \mathbf{S}_2| = 0, 1, 2$. The orbital angular momentum L has to compensate the total spin in order to reproduce the nought total angular momentum $J = |\mathbf{L} \oplus \mathbf{S}| = 0$ of the B_s^0 meson. Therefore L must have the same value as S . We get

$$\begin{aligned} C &= C_{J/\psi} \cdot C_\phi = 1, \\ P &= P_{J/\psi} \cdot P_\phi \cdot (-1)^L = (-1)^L, \\ CP &= (-1)^L. \end{aligned} \tag{1.55}$$

Therefore

$$\begin{aligned} CP = +1 &\iff L = 0, 2 \quad (\text{S- or D-wave}), \\ CP = -1 &\iff L = 1 \quad (\text{P-wave}). \end{aligned}$$

In order to extract the contribution of each CP-eigenstate, an angular analysis can be performed. Rather than decomposing the decay amplitude A into S-, P- and D-waves, we can alternatively decompose it into three independent components, corresponding to linear polarisation states of the vector mesons. The polarisation states are either longitudinal (0) to their directions of motion, or transverse and parallel (\parallel), or transverse and perpendicular (\perp) to one another. The states 0 and \parallel are P-even, while the state \perp is P-odd. Since J/ψ and ϕ are both C-odd states, the properties under P are the same as those under CP.

Let $\epsilon_{J/\psi}$ and ϵ_ϕ be polarisation vectors of the vector mesons in the J/ψ rest frame. The independent decay amplitudes are the rotationally invariant quantities linear in $\epsilon_{J/\psi}^*$ and ϵ_ϕ^* and involving possible powers of $\hat{\mathbf{p}}_\phi$, a unit vector in the direction of the momentum of ϕ in the J/ψ rest frame.

The two CP-even decay amplitudes are the combinations $\epsilon_{J/\psi}^* \cdot \epsilon_\phi^*$ (contributing to A_0 and A_\parallel) and $(\epsilon_{J/\psi}^* \cdot \hat{\mathbf{p}}_\phi)(\epsilon_\phi^* \cdot \hat{\mathbf{p}}_\phi) = \epsilon_{J/\psi}^{*L} \cdot \epsilon_\phi^{*L}$ (contributing only to A_0), where $\epsilon^L \equiv \epsilon \cdot \hat{\mathbf{p}}_\phi$. Equivalently, one can subtract off the longitudinal component of the polarisation vectors to replace $\epsilon_{J/\psi}^* \cdot \epsilon_\phi^*$ by $\epsilon_{J/\psi}^{*T} \cdot \epsilon_\phi^{*T}$, contributing only to A_\parallel , where the superscripts T refer to projections perpendicular to $\hat{\mathbf{p}}_\phi$. The CP-odd amplitude $[\epsilon_{J/\psi}^* \wedge \epsilon_\phi^*] \cdot \hat{\mathbf{p}}_\phi$ contributes only to A_\perp . The total amplitude A can be written as [44]:

$$\begin{aligned} A(B_s(t) \rightarrow J/\psi \phi) &= \frac{A_0(t)}{x} \epsilon_{J/\psi}^{*L} \epsilon_\phi^{*L} - \frac{A_\parallel(t)}{\sqrt{2}} \epsilon_{J/\psi}^{*T} \epsilon_\phi^{*T} \\ &\quad - i \frac{A_\perp(t)}{\sqrt{2}} [\epsilon_{J/\psi}^* \wedge \epsilon_\phi^*] \cdot \hat{\mathbf{p}}_\phi, \end{aligned} \tag{1.56}$$

where $x \equiv p_{J/\psi} p_\phi / (m_{J/\psi} m_\phi)$. Note that by counting powers of $\hat{\mathbf{p}}_\phi$ we can see directly that A_0 and $A_{||}$ are P-even while A_\perp is P-odd.

Vanishing CP violation in the decay amplitude of $B_s^0 \rightarrow J/\psi \phi$ implies:

$$\bar{A}_0(0) = A_0(0), \quad \bar{A}_{||}(0) = A_{||}(0), \quad \bar{A}_\perp(0) = -A_\perp(0). \quad (1.57)$$

Therefore the final state is an admixture of different CP eigenstates. A_0 and $A_{||}$ are CP-even amplitudes and A_\perp is CP-odd. The decay rate reads

$$\Gamma(t) \propto |A_0(t)|^2 + |A_{||}(t)|^2 + |A_\perp(t)|^2, \quad (1.58)$$

with individual time evolutions

$$\begin{aligned} |A_0(t)|^2 &= |A_0(0)|^2 [e^{-\Gamma_L t} - e^{\Gamma t} \sin(\Delta m t) \phi], \\ |A_{||}(t)|^2 &= |A_{||}(0)|^2 [e^{-\Gamma_L t} - e^{\Gamma t} \sin(\Delta m t) \phi], \\ |A_\perp(t)|^2 &= |A_\perp(0)|^2 [e^{-\Gamma_H t} + e^{\Gamma t} \sin(\Delta m t) \phi]. \end{aligned} \quad (1.59)$$

The value of $\phi \equiv 2\beta_s \approx 2\lambda^2 \eta \approx 0.03$ is small in the Standard Model, so that the oscillation terms can be dropped to first order. The time evolution is then simply a sum of two exponential decays with lifetimes $1/\Gamma_H$ and $1/\Gamma_L$:

$$\Gamma(t) \propto (|A_0(t)|^2 + |A_{||}(t)|^2) e^{-\Gamma_L t} + |A_\perp(t)|^2 e^{-\Gamma_H t}. \quad (1.60)$$

The expected small value of $\Delta\Gamma_s/\Gamma_s$ in the Standard Model being of order 10 % makes it hard to separate two closely spaced lifetimes. Additional angular information separating the two mass eigenstates will increase the accuracy in measuring $\Delta\Gamma_s$.

Since there are four particles in the final state, the directions of their momenta can define three independent physical angles. The common convention for the definitions of angles is shown in figure 1.8. The x axis is the direction of ϕ in the J/ψ rest frame, the z axis is perpendicular to the decay plane of $\phi \rightarrow K^+ K^-$, and $p_y(K^+) \geq 0$. The coordinates (θ, φ) describe the direction of the μ^+ in the J/ψ rest frame, and ψ is the angle made by $\mathbf{p}(K^+)$ with the x axis in the ϕ rest frame. With this convention,

$$\begin{aligned} \hat{\mathbf{x}} &= \hat{\mathbf{p}}_\phi, \quad \hat{\mathbf{y}} = \frac{\hat{\mathbf{p}}_{K^+} - \hat{\mathbf{p}}_\phi(\hat{\mathbf{p}}_\phi \cdot \hat{\mathbf{p}}_{K^+})}{|\hat{\mathbf{p}}_{K^+} - \hat{\mathbf{p}}_\phi(\hat{\mathbf{p}}_\phi \cdot \hat{\mathbf{p}}_{K^+})|}, \quad \hat{\mathbf{z}} = \hat{\mathbf{x}} \wedge \hat{\mathbf{y}} \\ \sin \theta \cos \varphi &= \hat{\mathbf{p}}_{\mu^+} \cdot \hat{\mathbf{x}}, \quad \sin \theta \sin \varphi = \hat{\mathbf{p}}_{\mu^+} \cdot \hat{\mathbf{y}}, \quad \cos \theta = \hat{\mathbf{p}}_{\mu^+} \cdot \hat{\mathbf{z}}. \end{aligned} \quad (1.61)$$

Here, the vectors with a hat represent unit 3-vectors measured in the rest frame of the J/ψ .

The angle θ here is the **transversity angle** [45], [44], which separates out the CP-even and CP-odd components. The angular distribution in terms of θ is given by:

$$\begin{aligned} \frac{d\Gamma(t)}{d\cos\theta} &\propto (|A_0(t)|^2 + |A_{||}(t)|^2) \cdot \frac{3}{8}(1 + \cos^2 \theta) \\ &+ |A_\perp(t)|^2 \cdot \frac{3}{4} \sin^2 \theta, \end{aligned} \quad (1.62)$$

where the time evolutions of the terms are given as in equation 1.59.

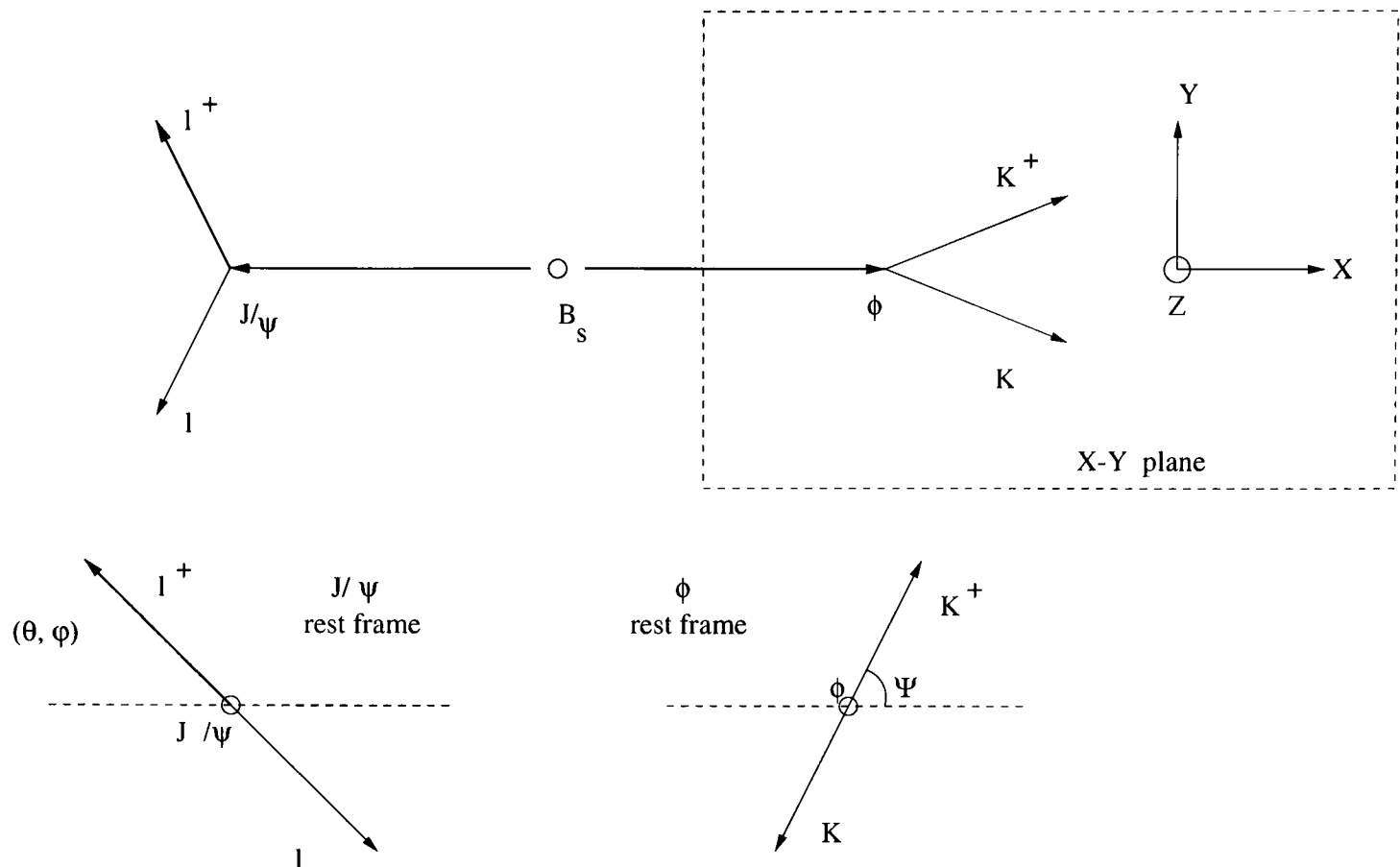


Figure 1.8: The definitions of the angles θ , φ , ψ of the transversity frame. The angle θ is the “transversity angle”.

The CP-even and CP-odd components are now separated not only by their different lifetimes (which are very close), but also by their decay angular distributions (which are significantly different). The width difference can be measured by a simultaneous fit to lifetime and transversity angle.

In appendix D we study the precision of such a fit, assuming yields and backgrounds we will obtain in chapter 4. Unfortunately yields are too low to measure $\Delta\Gamma_s/\Gamma_s$ with reasonable precision and the fit suffers from non-Gaussian tails. For this reason we only fit a single lifetime to the proper decay time distribution and ignore the transversity angle from now on.

1.10 Fitting the Average Lifetime

In this section we show what is meant by the term **average lifetime**, when fitting a sample, which contains two different exponential decays, with a single lifetime. This is important in order to interpret measured average lifetimes for example in the decay $B_s^0 \rightarrow J/\psi \phi$.

For a better understanding we begin with a simple case: we fit a single exponential function to a set of random data. If we have N data points x_i ($i = 0, 1, \dots, N$) we can

construct a likelihood function

$$\mathcal{L} = \prod_{i=1}^N \frac{e^{-x_i/\tau_{fit}}}{\tau_{fit}} . \quad (1.63)$$

Minimising $-\log \mathcal{L}$ by setting $\partial \log \mathcal{L} / \partial \tau_{fit} = 0$ we get

$$\tau_{fit} = \frac{1}{N} \sum_{i=1}^N x_i . \quad (1.64)$$

Thus the best fit lifetime τ_{fit} is simply the mean of the data sample. In case the data points are distributed according to an exponential $1/\tau_1 e^{-x_i/\tau_1}$, the fit would yield the expectation value of the exponential, which is the lifetime τ_1 , i.e. $\tau_{fit} = \tau_1$.

If the dataset consists of two subsets, x_1, \dots, x_{N_1} distributed according to $1/\tau_1 e^{-x_i/\tau_1}$ and $x_{N_1+1}, \dots, x_{N_1+N_2}$ distributed according to $1/\tau_2 e^{-x_i/\tau_2}$, then a single exponential fit with lifetime τ_{fit} would yield

$$\tau_{fit} = \frac{1}{N_1 + N_2} (N_1 \cdot \tau_1 + N_2 \cdot \tau_2) . \quad (1.65)$$

Thus, the average lifetime τ_{fit} is the weighted average of the individual lifetimes τ_1 and τ_2 . In the $B_s^0 \rightarrow J/\psi \phi$ decay, where the CP even fraction f_{CP+} decays with lifetime τ_{CP+} and the CP odd fraction with τ_{CP-} , we get

$$\tau_{average} = f_{CP+} \cdot \tau_{CP+} + (1 - f_{CP+}) \cdot \tau_{CP-} . \quad (1.66)$$

In terms of the variables Γ_s and $\Delta\Gamma_s$ we get

$$\tau_{average} = \frac{1}{\Gamma_s} \frac{1 - (1 - 2f_{CP+}) \frac{\Delta\Gamma_s}{2\Gamma_s}}{1 - \left(\frac{\Delta\Gamma_s}{2\Gamma_s}\right)^2} , \quad (1.67)$$

and $\tau_{average}$ contains $\Delta\Gamma_s/\Gamma_s$ to the first power.

For clarity we discuss two more examples: the inclusive B_s^0 lifetime and the flavour specific (\equiv semileptonic) B_s^0 lifetime. If we fit a single lifetime to the proper decay time distribution of an inclusive B_s sample we measure the **inclusive lifetime** τ_{incl} . An equal number of B_s^{Short} and B_s^{Long} mesons is selected, leading to a proper time distribution

$$P_{incl}(t) = \frac{1}{2} \left(\Gamma_s^{\text{Short}} e^{-\Gamma_s^{\text{Short}} t} + \Gamma_s^{\text{Long}} e^{-\Gamma_s^{\text{Long}} t} \right) , \quad (1.68)$$

which is sketched in figure 1.9 on the left. The inclusive lifetime then becomes

$$\tau_{incl} = \frac{1}{2} (\tau_{\text{Long}} + \tau_{\text{Short}}) = \frac{1}{\Gamma_s} \frac{1}{1 - \left(\frac{\Delta\Gamma_s}{2\Gamma_s}\right)^2} . \quad (1.69)$$

Note that $\Delta\Gamma_s$ appears only in second power.

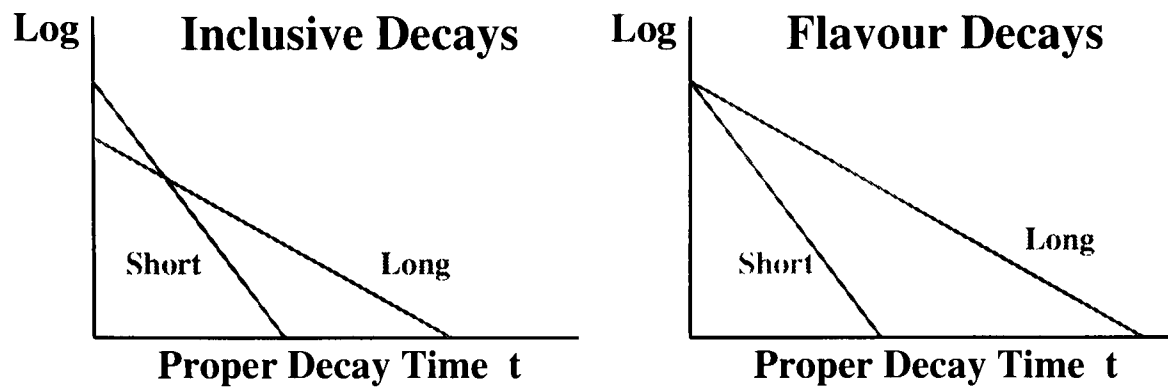


Figure 1.9: Sketch of the proper life distribution for an inclusive B_s sample (left) and a B_s sample decaying into flavour eigenstates (right). The vertical axis is on a logarithmic scale. In an inclusive B_s sample, the total number of B_s^{Short} and B_s^{Long} is the same. Since the B_s^{Short} has a shorter lifetime, in a small interval at $t = 0$ there will be more B_s^{Short} decays than B_s^{Long} decays, as can be seen in the left figure. The situation is different if we collect decays to flavour eigenstates (right). The mesons are produced as flavour eigenstates, live as a superposition of the lifetime eigenstates, before they decay as flavour eigenstates. At $t = 0$ there is an equal number of long- and short-lived mesons, and as the long-lived component lives longer, more of the selected flavour decays come from the long-lived component, as can be seen from the figure on the right.

In case the sample is collected through flavour specific decays, such as semileptonic decays, a single exponential fit will return the **flavour specific lifetime** τ_f . The mesons are produced as flavour eigenstates, live as a superposition of the lifetime eigenstates, before they decay as flavour eigenstates. At $t = 0$ there is an equal number of long- and short-lived mesons, and as the long-lived component lives longer, more of the selected flavour decays come from the long-lived component. Therefore the numbers of selected B_s^{Short} and B_s^{Long} mesons are proportional to $1/\Gamma^{\text{Short}}$ and $1/\Gamma^{\text{Long}}$, respectively, as is illustrated in the right picture of figure 1.9 (see also equation 1.53). Properly normalised, the proper time distribution reads

$$P_f(t) = \frac{\Gamma_s^{\text{Short}} \Gamma_s^{\text{Long}}}{\Gamma_s^{\text{Short}} + \Gamma_s^{\text{Long}}} \cdot \left(e^{-\Gamma_s^{\text{Short}} t} + e^{-\Gamma_s^{\text{Long}} t} \right), \quad (1.70)$$

and the flavour specific lifetime becomes

$$\tau_f = \frac{\tau_{\text{Short}}}{\tau_{\text{Short}} + \tau_{\text{Long}}} \tau_{\text{Short}} + \frac{\tau_{\text{Long}}}{\tau_{\text{Short}} + \tau_{\text{Long}}} \tau_{\text{Long}} = \frac{1}{\Gamma_s} \frac{1 + \left(\frac{\Delta \Gamma_s}{2\Gamma_s} \right)^2}{1 - \left(\frac{\Delta \Gamma_s}{2\Gamma_s} \right)^2}. \quad (1.71)$$

1.11 Current Experimental Status of B Lifetime Measurements

Measuring the lifetime of the B mesons in exclusive modes is a relatively new field. First precision measurements have been performed in the 1990's by the experiments at LEP and CDF at the Tevatron [12]. The current world average values for the B_d and B_u lifetimes are

$$\begin{aligned}\tau_{B_d} &= (1.537 \pm 0.015) \text{ ps} , \\ \tau_{B_u} &= (1.671 \pm 0.018) \text{ ps} , \\ \tau_{B_u}/\tau_{B_d} &= 1.085 \pm 0.017 ,\end{aligned}\tag{1.72}$$

and include a variety of measurements from the LEP experiments, CDF, and, with best precision, from the B factory experiments Belle and BaBar [12].

So far all B_s^0 lifetime analyses measure the average lifetime of the B_s^0 , which is evaluated under the assumption that the decays are governed by a single exponential function with a unique lifetime. The current combined experimental value is [12]

$$\tau(B_s) = (1.461 \pm 0.057) \text{ ps} .\tag{1.73}$$

Figure 1.10 shows the combined result together with its individual measurements. Most analyses are based on inclusive reconstruction of D_s mesons from B_s decays [46], [47], [48], [49], [50], [43], [52], [53], often in association with electrons and muons. One measurement from CDF [51] uses the fully reconstructed exclusive mode $B_s^0 \rightarrow J/\psi \phi$. All measurements agree well within their estimated errors. It is worth mentioning that the experiments running on the $\Upsilon(4S)$ resonance (Argus, Belle, BaBar) cannot produce B_s^0 mesons as their energy is below the threshold of producing B_s^0 pairs.

1.12 Current Experimental Status of $\Delta\Gamma_s$ Measurements

In the past five years there have been a few attempts to measure the width difference in the B_s^0 system. Unfortunately the small B_s^0 yields did not yet allow a good enough precision to measure a significantly non-zero value. Thus upper limits have been set. Many analyses use the constraint $\Gamma_s = \Gamma_d$. This is justified, since the ratio $\tau(B_s)/\tau(B_d)$ is expected to be very close to unity and corrections from SU(3) breaking effects are estimated to be smaller than 1 % (see section 1.3). The following summary of experimental constraints on $\Delta\Gamma_s/\Gamma_s$ is based on the report of a combined Working Group from the experiments ALEPH, CDF, DELPHI, L3, OPAL and SLD [10]:

- **Inclusive B sample (L3 [54]).** In an unbiased inclusive B sample, all B_s^0 decay modes are measured. Fitting the proper time distribution to a single exponential lifetime gives sensitivity to $\Delta\Gamma_s$ in second order

$$\tau_{B_s}^{incl} = \frac{1}{\Gamma_s} \frac{1}{1 - \left(\frac{\Delta\Gamma_s}{2\Gamma_s}\right)^2} .\tag{1.74}$$

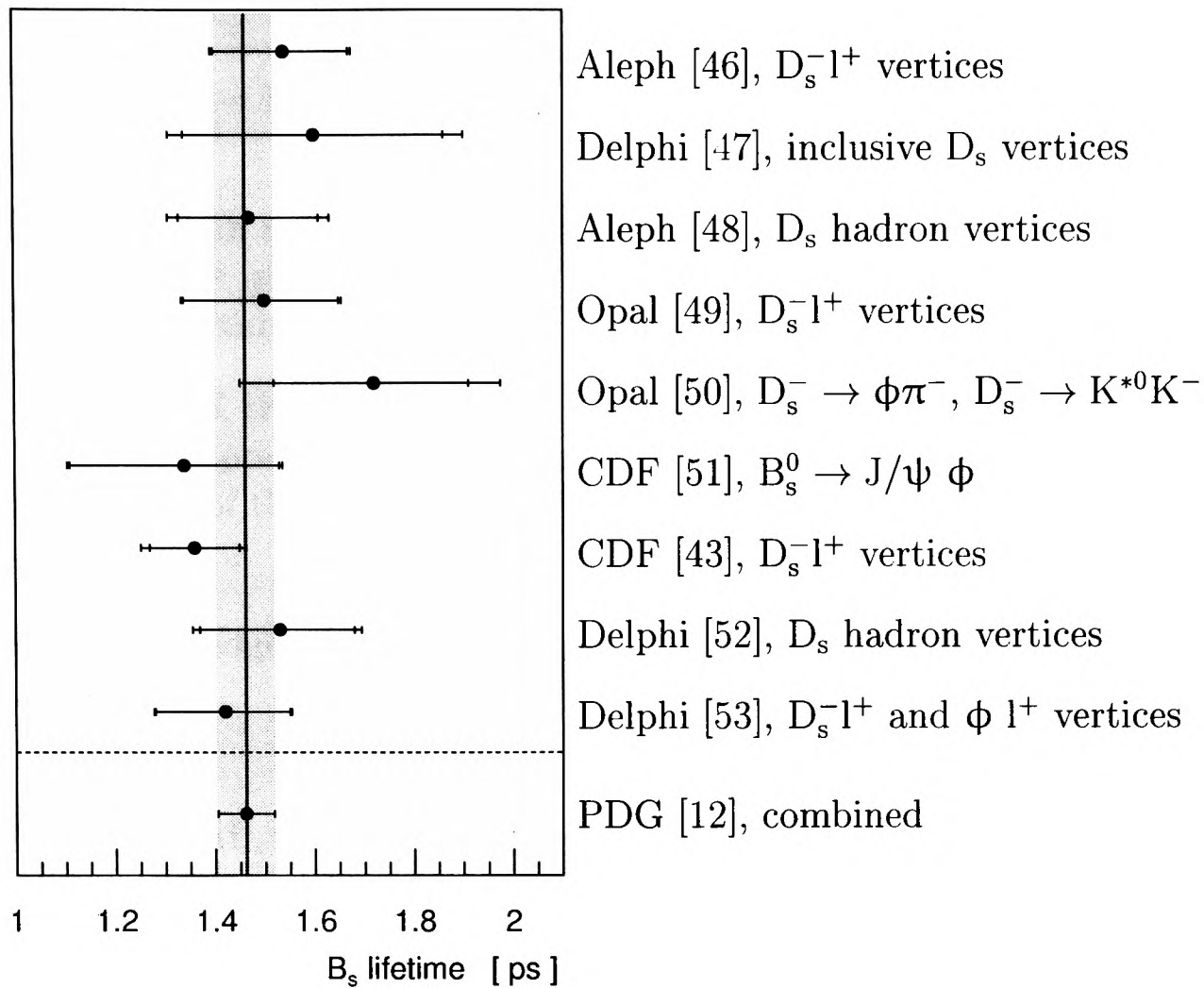


Figure 1.10: Compilation of recent measurements of the average lifetime of the B_s^0 meson and the combined PDG value. The error bars are statistical (inner error bar) and statistical \oplus systematic (total error bar) added in quadrature.

L3 incorporates $P_{incl}(t)$ into the proper lifetime fit and applies the constraint $1/\Gamma_s = (1.49 \pm 0.06)$.

- $\overline{B}_s^0 \rightarrow D_s^+ l^- \overline{\nu}_l X$ (Delphi [53]). In a sample of semileptonic B_s^0 decays the proper decay time distribution is given by equation 1.70. If this distribution is fitted assuming a single exponential lifetime $\tau_{B_s}^{semi}$, then the measured lifetime is given by

$$\tau_{B_s}^{semi} = \frac{1}{\Gamma_s} \cdot \frac{1 + \left(\frac{\Delta\Gamma_s}{2\Gamma_s}\right)^2}{1 - \left(\frac{\Delta\Gamma_s}{2\Gamma_s}\right)^2}. \quad (1.75)$$

A 95% CL limit $\Delta\Gamma_s/\Gamma_s < 0.46$ is obtained from the proper time distribution, assuming $\Gamma_s = \Gamma_d$.

- $B_s^0 \rightarrow \phi\phi X$ counting method (Aleph [55]). It is assumed that the width difference is due entirely to the $B_s^{Short} \rightarrow D_s^{(*)+} D_s^{(*)-}$ decay modes. Under this

assumption, $\Delta\Gamma_s \equiv \Gamma(B_s^{\text{Short}} \rightarrow D_s^{(*)+} D_s^{(*)-})$ where

$$\text{BR}(B_s^{\text{Short}} \rightarrow D_s^{(*)+} D_s^{(*)-}) = \frac{\Gamma(B_s^{\text{Short}} \rightarrow D_s^{(*)+} D_s^{(*)-})}{\Gamma_s^{\text{Short}}} = \frac{\Delta\Gamma_s}{\Gamma_s \left(1 + \frac{\Delta\Gamma_s}{2\Gamma_s}\right)}. \quad (1.76)$$

This results in a measured value of $\Delta\Gamma_s/\Gamma_s = 0.26_{-0.15}^{+0.30}$, which is so far the only measurement of $\Delta\Gamma_s/\Gamma_s$ which does not rely on a measurement of the average B_s^0 or B_d^0 lifetime.

- **$B_s^0 \rightarrow \phi\phi X$ lifetime method (Aleph [55]).** As the decay $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-} \rightarrow \phi\phi X$ is predominantly CP even, the proper time dependence of the B_s^0 component of the $\phi\phi$ component is therefore just a simple exponential

$$P(t) = \Gamma_s^{\text{Short}} e^{-\Gamma_s^{\text{Short}} t}. \quad (1.77)$$

With knowledge of Γ_s^{Short} it is easy to extract

$$\frac{\Delta\Gamma_s}{\Gamma_s} = 2 \left(\frac{\Gamma_s^{\text{Short}}}{\Gamma_s} - 1 \right). \quad (1.78)$$

A value of $\Delta\Gamma_s/\Gamma_s = 0.45_{-0.49}^{+0.80}$ is obtained assuming $\Gamma_s = \Gamma_d$.

- **$\overline{B}_s^0 \rightarrow D_s^+$ hadron (Delphi [52]).** A fully inclusive D_s^+ selection is expected to have an increased CP-even content, as the $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$ contribution is enhanced. $\Delta\Gamma_s/\Gamma_s$ is obtained by taking into account the relative abundance of B_s^{Short} and B_s^{Long} while fitting to the proper time distribution. Assuming $\Gamma_s = \Gamma_d$ the authors get $\Delta\Gamma_s/\Gamma_s < 0.69$ at 95% CL.
- **Other measurements of the B_s^0 semileptonic lifetime (Aleph, Opal, CDF)[56].** In certain B_s^0 lifetime analyses, a single exponential function was assumed in the fits to the data. The Working Group has taken the average of the measured lifetimes, $\tau_{B_s}^{\text{semi}} = (1.46 \pm 0.07)$ ps, to be equal to the expression in equation 1.75 with $\Gamma_s = \Gamma_d$. The lower bound $\Delta\Gamma_s/\Gamma_s < 0.30$ is obtained at 95% CL.
- **$B_s^0 \rightarrow J/\psi \phi$ (CDF [43]).** Studies of angular distributions by CDF favour the final state $J/\psi\phi$ to be predominantly CP even with a fraction of $f_{CP+} = 0.79 \pm 0.19$. The average B_s^0 lifetime is measured to be $(1.34_{-0.19}^{+0.23} \pm 0.05)$ ps in this final state. With the assumption $\Gamma_s = \Gamma_d$, this result corresponds to $\Delta\Gamma_s/\Gamma_s = 0.33_{-0.42}^{+0.45}$, where the uncertainty on the CP-even fraction is taken into account.

The Working Group on $\Delta\Gamma_s$ has combined the experimental results in terms of the likelihood over the $(1/\Gamma_s, \Delta\Gamma_s/\Gamma_s)$ plane. (The L3 result is not included in the combination because the likelihood function was not available over this space.) The combination is first performed without applying the constraint $\Gamma_s = \Gamma_d$, resulting in

$$\begin{aligned} \Delta\Gamma_s/\Gamma_s &= 0.24_{-0.12}^{+0.16} \\ \text{or } \Delta\Gamma_s/\Gamma_s &< 0.53 \text{ at 95\% CL.} \end{aligned} \quad (1.79)$$

When the constraint on Γ_s is applied – assuming no theoretical uncertainty on Γ_s/Γ_d – the results become

$$\begin{aligned} \Delta\Gamma_s/\Gamma_s &= 0.16^{+0.08}_{-0.09} \\ \text{or } \Delta\Gamma_s/\Gamma_s &< 0.31 \text{ at 95\% CL .} \end{aligned} \quad (1.80)$$

The likelihood distributions and probability density is shown in figure 1.11 and table 1.5 summarises the experimental status of $\Delta\Gamma_s/\Gamma_s$ measurements.

Experiment	Result	Selection	Comment
L3 [54]	< 0.67	Inclusive B sample	Uses $\tau_{B_s} = (1.55 \pm 0.05)\text{ps}$
Delphi [53]	< 0.46	$\overline{B}_s^0 \rightarrow D_s^+ l^- \bar{\nu}_l X$	
Others [56]	< 0.30	$\overline{B}_s^0 \rightarrow D_s^+ l^- \bar{\nu}_l X$	
Aleph [55]	$0.26^{+0.30}_{-0.15}$	$B_s^0 \rightarrow \phi\phi X$	Counting method, does not assume $\Gamma_s = \Gamma_d$
Aleph [55]	$0.45^{+0.80}_{-0.49}$	$B_s^0 \rightarrow \phi\phi X$	Lifetime method
Delphi [52]	< 0.69	$\overline{B}_s^0 \rightarrow D_s^+ \text{ hadron}$	
CDF [43]	$0.33^{+0.45}_{-0.42}$	$B_s^0 \rightarrow J/\psi \phi$	
[10]	< 0.31 (95% CL)	Combined	uses constraint $\Gamma_s = \Gamma_d$
[10]	< 0.53 (95% CL)	Combined	without constraint $\Gamma_s = \Gamma_d$

Table 1.5: Current experimental status of $\Delta\Gamma_s/\Gamma_s$ measurements. All upper limits are given at 95% CL.

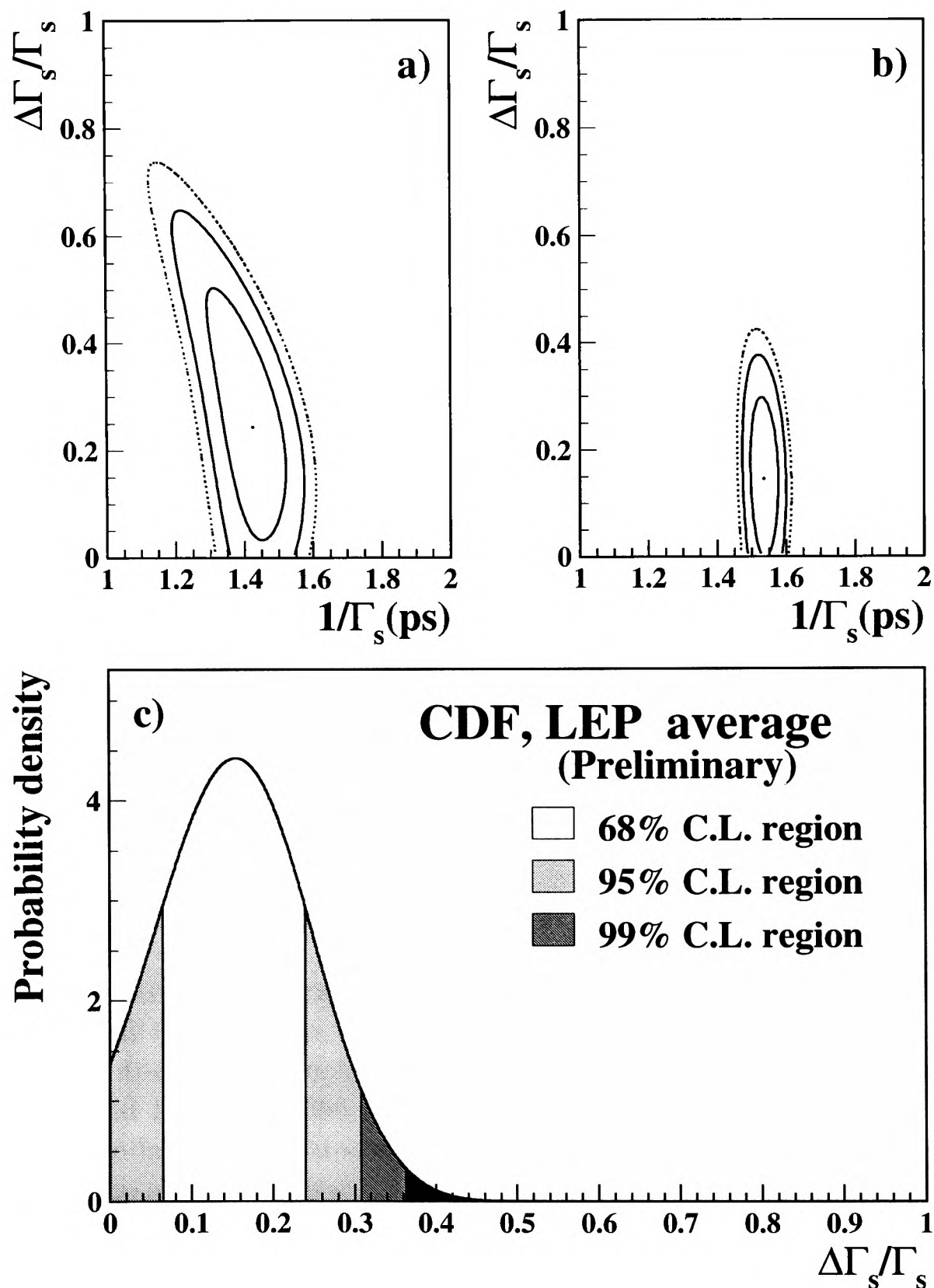


Figure 1.11: a) 68%, 95% and 99% C.L. contours of the negative log-likelihood distribution in the plane $(1/\Gamma_s, \Delta\Gamma_s/\Gamma_s)$. b) Same as a) but with the constraint $\Gamma_s = \Gamma_d$. c) Probability density distribution for $\Delta\Gamma_s/\Gamma_s$ after applying the constraint; the three shaded regions show the limits at the 68%, 95% and 99% CL respectively. Taken from [10].

Chapter 2

The CDF Experiment at Fermilab

The Collider Detector at Fermilab is a 5,000 ton multi-purpose particle physics experiment dedicated to the study of proton-antiproton collisions. It is located at the Fermilab Tevatron collider (near Chicago) with a centre-of-mass energy of 1.96 TeV. The experiment is performed by an international collaboration of about 790 physicists from 56 institutions of 11 countries to study a wide range particle physics program: electro-weak physics, top quark physics, QCD, B physics, and searches for particles beyond the Standard Model. It restarted data taking in autumn 2001 after a five-year shutdown with a major accelerator and detector upgrade. The current period of data taking is called Run IIa. We first give a brief overview over the Tevatron accelerator complex and then describe the components of the CDF detector relevant to understand this analysis.

2.1 The Tevatron Collider

The Tevatron proton-antiproton collider is the highest energy particle collider currently operational anywhere in the world and is described in detail elsewhere [57]. Currently 36 proton bunches collide with 36 antiproton bunches at two interaction regions (CDF and D0) every 396 ns at a centre-of-mass energy of 1.96 TeV. Figure 2.1 pictures Fermilab's accelerator chain which is briefly explained in the following.

The first stage of acceleration is achieved with the **Cockcroft-Walton accelerator**. It accelerates 25 keV H^- ions to 750 keV. The following 150 m long **linear accelerator (LINAC)** increases the energy to 400 MeV. The electrons of the H^- ions are then stripped off leaving protons, which are accelerated further to 8 GeV in the **Booster** – a synchrotron of about 75 m in radius. The **Main Injector (MI)**, circumference about 3 km) takes over and accelerates the proton up to 150 GeV, before they are injected into the **Tevatron**, the final stage in the chain. The Tevatron has a circumference of about 2π kilometres and brings the protons to their final energy of 980 GeV. It uses a lattice of superconducting dipoles and quadrupoles to keep the beam in position and conventional RF cavities for acceleration.

The Main Injector is also used for the production of antiprotons. Protons of 120 GeV are focused on a nickel target. Antiprotons of 8 GeV are selected, focused

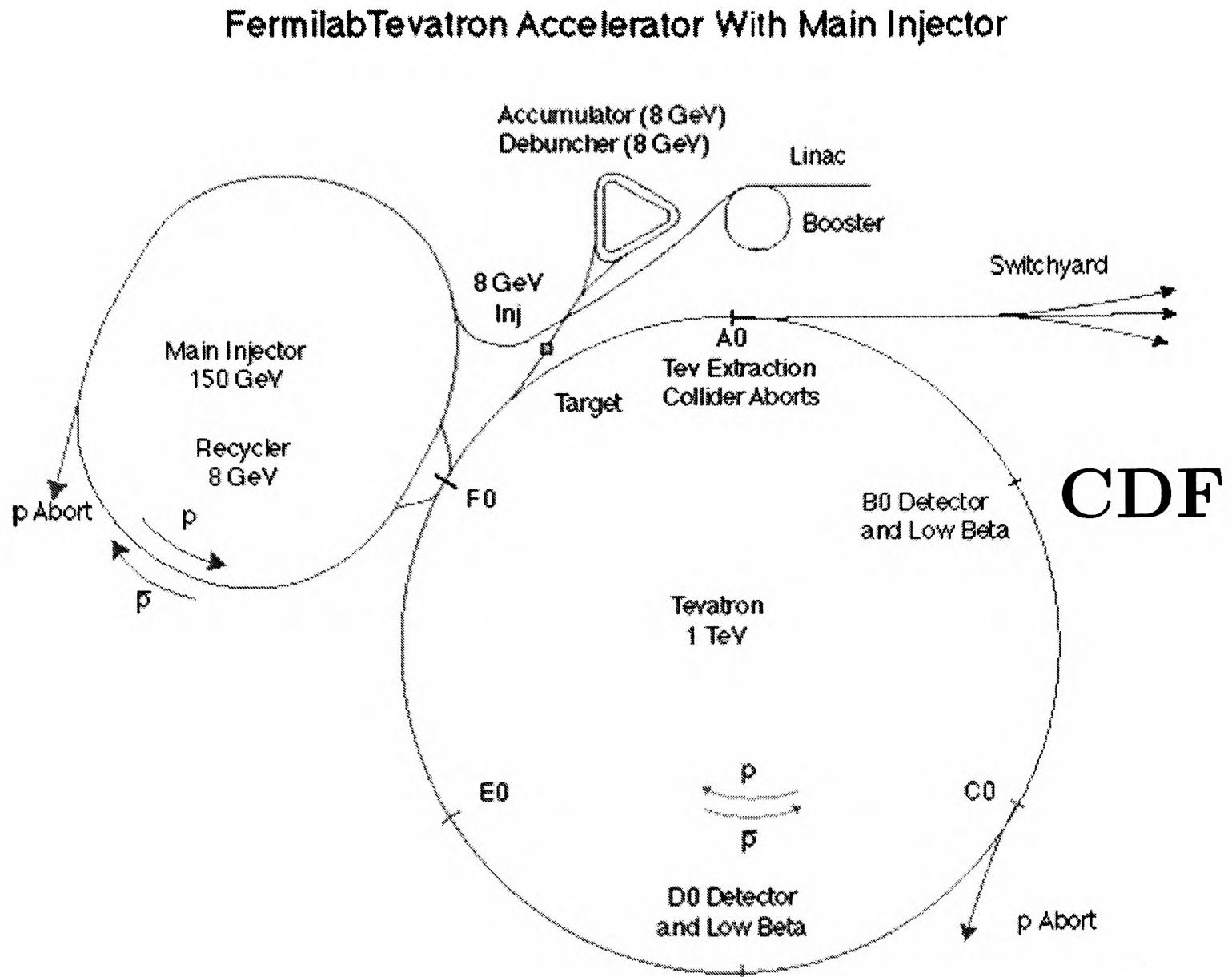


Figure 2.1: The Fermilab Tevatron Collider with Main Injector.

and delivered to the **Debuncher Ring** where they are debunched and stochastically cooled. These are then transferred to the **Accumulator Ring** where they are stacked and stochastically cooled again. Once a sufficient number of antiprotons has been produced, they are sent back to the Main Injector to be accelerated to 150 GeV before they enter the Tevatron for their final energy of 980 GeV.

With both beams at 980 GeV in the Tevatron a low beta squeeze is executed, and the beams are finally brought into collisions at the two interaction regions (CDF and D0). An entire filling procedure typically takes about 3 hours from the end of the previous store to the start of the next store. A collider store generally lasts about 14 hours which is approximately one luminosity lifetime.

The current performance [58] is at roughly 25 % of the Run IIa goal: initial luminosities average $2.2 \times 10^{31} \text{ cm}^{-2} \text{ s}^{-1}$ with a best luminosity to date of $4.22 \times 10^{31} \text{ cm}^{-2} \text{ s}^{-1}$. Current performance and Run IIa goals are listed in table 2.1. It is evident that current performance is characterised primarily by a 37% shortfall in proton intensity and a 55% shortfall in antiprotons relative to Run IIa goals. At present, the antiproton production rate is not an impediment to peak luminosity performance, because the

Parameter	Current (typical)	Run IIa Goal	Unit
Protons/bunch	1.70×10^{11}	2.70×10^{11}	
Antiprotons/bunch	1.39×10^{10}	3.0×10^{10}	
Total Antiprotons	5.00×10^{11}	1.08×10^{12}	
Peak \bar{p} Production Rate	1.10×10^{11}	2.00×10^{11}	\bar{p} /hour
Accumulator \rightarrow 150 GeV			
coalesced efficiency	0.80	0.90	
MI 150 GeV \rightarrow Low β efficiency	0.60	0.90	
Accumulator \rightarrow Low β efficiency	0.48	0.81	
Longitudinal Emittance (proton)	6	3	eV s
Longitudinal Emittance (\bar{p})	4.5	2	eV s
Typical Luminosity	2.2×10^{31}	8.1×10^{31}	$\text{cm}^{-2}\text{s}^{-1}$
Integrated Luminosity	4	16	$\text{pb}^{-1}/\text{week}$

Table 2.1: Current performance of the Tevatron compared with Run IIa goals (taken from [58]).

average store duration (about 14 hours) is sufficient to allow accumulation of the requisite number of antiprotons at the current stacking rate.

The low proton intensities in the Tevatron are due to beam instabilities, significant beam loss during the acceleration ramp and anomalous longitudinal emittance growth on the Main Injector acceleration cycle. On the other hand the Tevatron experiences low antiproton intensities from a transmission efficiency of only 48%. Transmission efficiency through the bunch coalescing process in the Main Injector is roughly 80% and efficiency from the Main Injector through the initiation of collisions in the Tevatron is roughly 60%. There are a number of issues limiting the performance, some of them the same as for the protons, but in addition emittance dilution and low efficiency (about 90%) in the Main Injector to Tevatron beam transfer.

All these issues are currently under study and a number of fixes and additions are or will be made to achieve the design luminosities. See [58] for further details.

2.2 The CDF II Detector

The CDF detector underwent a major upgrade for the Tevatron collider Run IIa to make it suitable for the order of magnitude increase in luminosity. Changes include simplification and improvement of the angular range of the calorimetry and muon coverage, improvement of the speed of electronics and trigger system, addition of a time-of-flight counter, and a complete replacement of the charged particle track detectors. We will refer to this version of the CDF detector by CDF II. Figures 2.2 and 2.3 show its inner structure: a tracking system is located in a super-conducting solenoid, which provides a magnetic field of 1.4 T for momentum measurement of charged tracks. It is surrounded by the electromagnetic and hadronic calorimeter, and

the muon system. A detailed description can be found in the Technical Design Report [59]. Here we want to give only a brief description of those detector components, which are relevant to this analysis: the tracking, muon and trigger system.

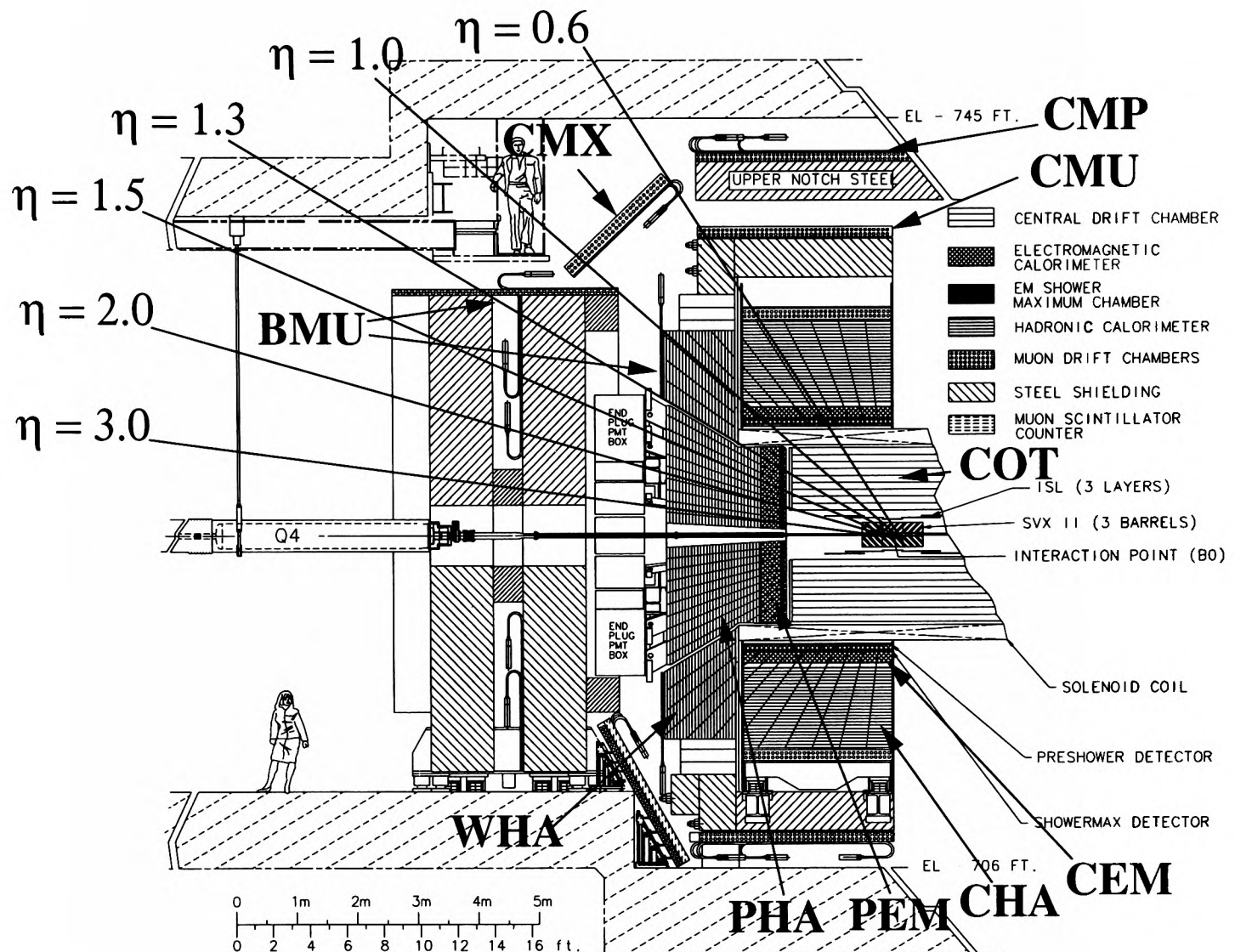


Figure 2.2: Elevation view of one half of the CDF II detector.

CDF uses a cylindrical coordinate system (r, ϕ, z) with the origin at the centre of the detector and the z axis along the nominal direction of the proton beam (from west to east). The x axis points away from the Tevatron centre, and the y axis is $\mathbf{y} = \mathbf{z} \wedge \mathbf{x}$. Sometimes different coordinates such as the polar angle θ , or **pseudorapidity** η are being used, where

$$\eta \equiv -\log \left(\tan \frac{\theta}{2} \right) . \quad (2.1)$$

Tracks are fit to helical trajectories. The plane perpendicular to the beam is referred to as the “transverse plane”, and the transverse momentum of the track is referred to as p_T .

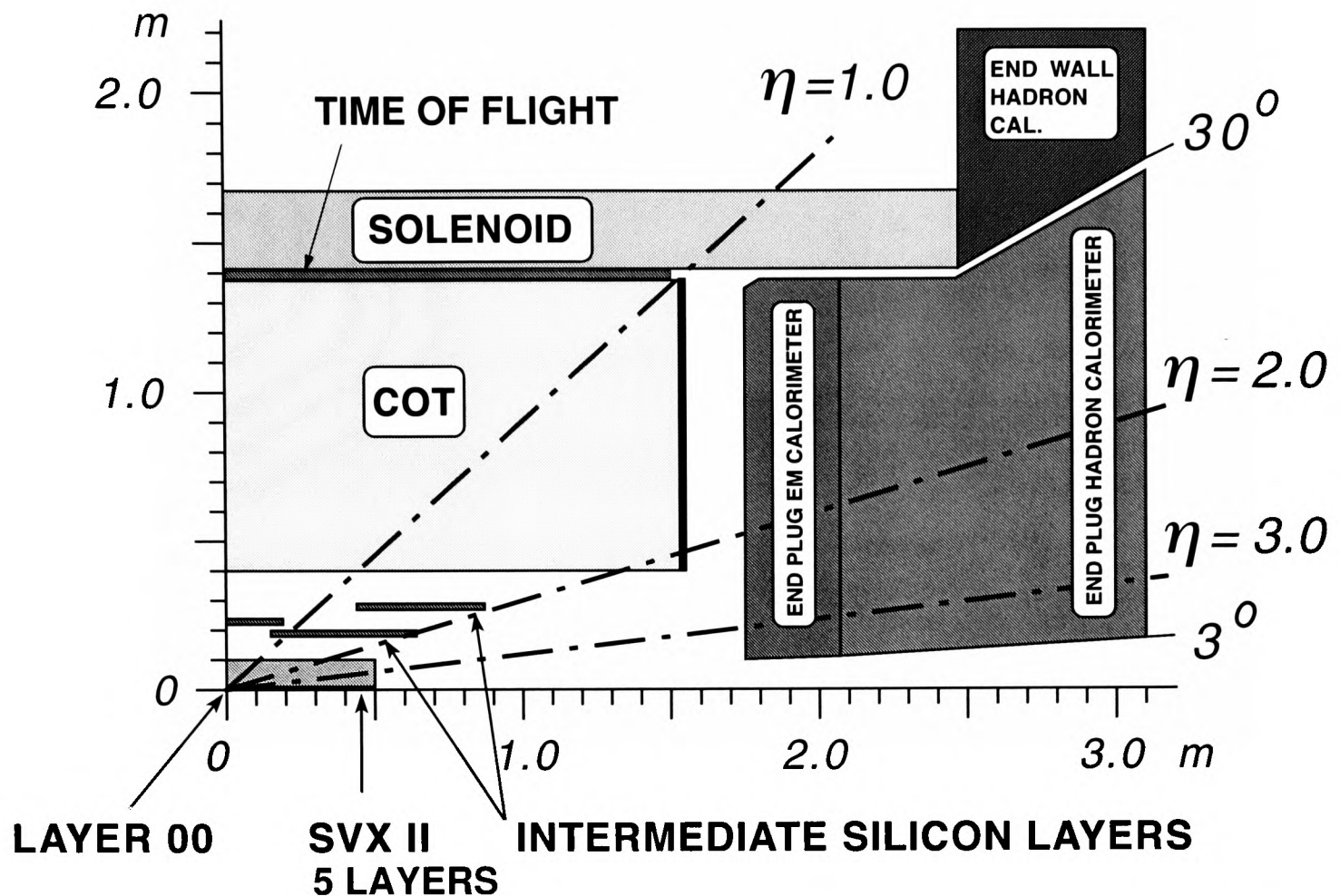


Figure 2.3: Longitudinal view of the CDF II tracking system, representing a quarter of the whole detector. The central calorimeters and the muon systems are not shown.

2.3 Silicon Vertex Detectors

At the heart of the CDF II detector are the silicon vertex detectors[60] : Layer 00 (L00), the Silicon Vertex Detector (SVX), and the intermediate Silicon Layers (ISL). Figure 2.4 shows in an end view at the top the five layers of the SVX and at the bottom its three-barrel structure. The side view shown in Figure 2.5 is a cross-section of one half of the silicon tracker, using a compressed z scale.

The silicon system provides precise three-dimensional track impact parameter determination over as wide an acceptance range as possible. Many different physics programs benefit: B physics, studies of top production, supersymmetry searches and the search for the Higgs boson. In addition they bridge more seamlessly between the vertex detectors and outer tracker, to improve purity and efficiency of tracking, and to increase the angular acceptance for well-reconstructed tracks.

The central vertexing portion of the detector is called **SVX II**. It is made up of three 29 cm long barrels, which extend 45 cm in z on either side of the detector origin, leading to an acceptance of $|\eta| < 2$. The barrels are separated by gaps of about 1.5 cm, which cause tracking inefficiencies in those regions. On every barrel there are two readout units, so-called half-ladders, which consist of two wafers in the SVX II – the smallest unit in a layer. So each barrel is four wafers in length. The five layers of ladders are arranged in twelve azimuthal wedges that alternate in radius

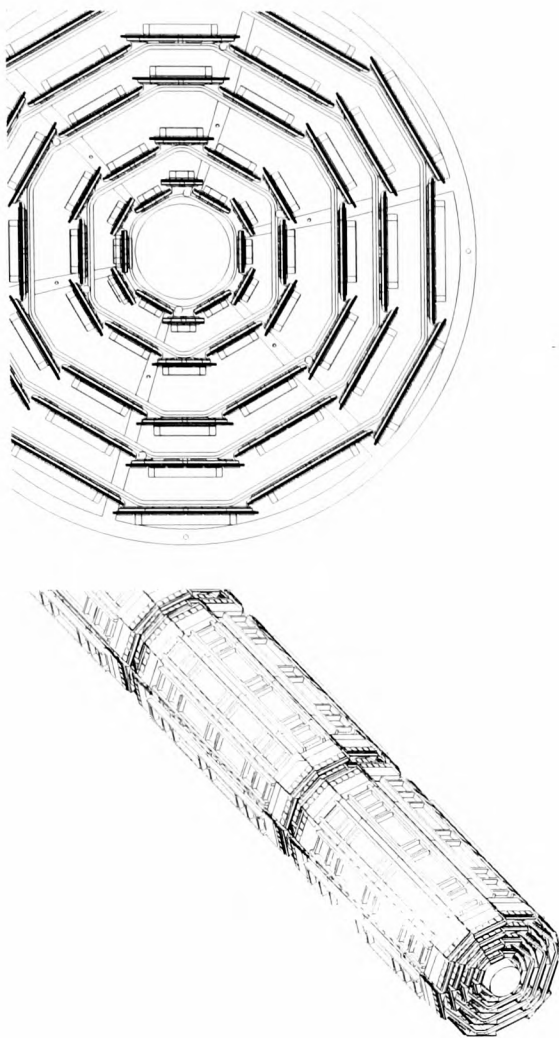


Figure 2.4: Top: the five layers of the SVX. Bottom: the three-barrel structure of the SVX.

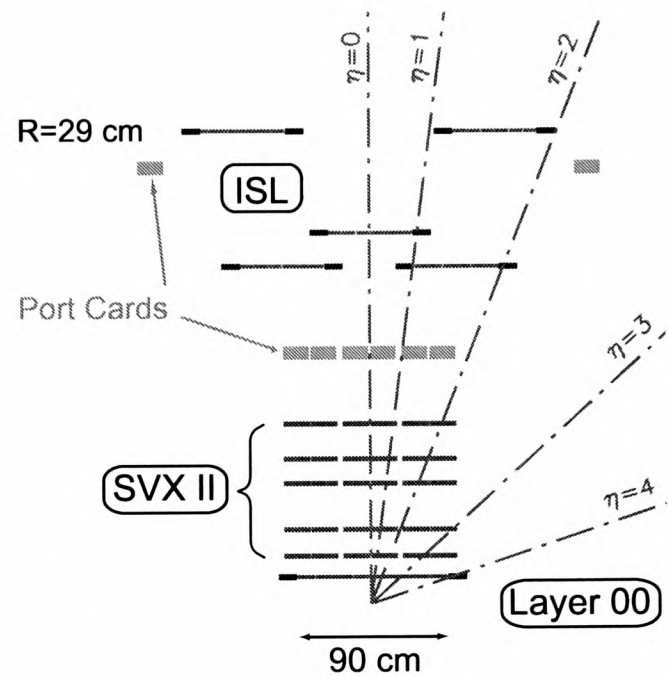


Figure 2.5: A side view of half of the CDF Run II silicon system on a scale in which the z coordinate is compressed.

within each layer (sublayers a/b). Each wedge is instrumented with 5 double-sided silicon layers with a combination of both $90^\circ z$ and small-angle stereo (SAS) layers. Both types have axial p-strips parallel to the z axis for measuring the azimuthal ϕ coordinate of a hit. These strips are spaced in $r\phi$ by approximately 60 to 65 μm , depending on layer, and have implant widths of 14 to 15 μm .

The n-strips are stereo strips, rotated with respect to the non-stereo strips by $(90, 90, -1.2, 90, +1.2)^\circ$, respectively from the innermost to outermost SVX II layers. The stereo strips are spaced by (141, 125.5, 60, 141, 65) μm , and have implant widths of 20 μm for the $90^\circ z$ strips and 15 μm for the SAS layers. The $90^\circ z$ layers measure in addition the z position via strips perpendicular to the z axis, with a pitch that ranges from 58 to 60 μm depending on layer. The SAS layers measure a stereo coordinate via strips twisted by a small angle of $\pm 1.2^\circ$ versus the ϕ strips. While the z resolution of a measurement with a $90^\circ z$ layer is about 50 ($\approx 1/\sin 1.2^\circ$) times better than with a SAS layer, there are more hit ambiguities with the $90^\circ z$ layers. Figure 2.6 shows three real hits in a $90^\circ z$ layer, and six so-called *ghost hits* arising from ambiguities. The same situation is better for the SAS layer – as shown in figure 2.7 – where such ambiguities do not arise at the cost of a less accurate z estimation. The alternating

pattern of 90° and SAS layers is designed to permit good resolution in locating the z position of secondary vertices and to enhance the 3D pattern recognition capability of the silicon tracker.

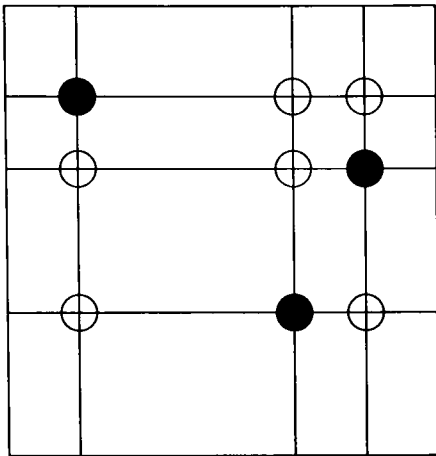


Figure 2.6: Three real hits (dark) and six ghost hits (light) in a $90^\circ z$ layer.

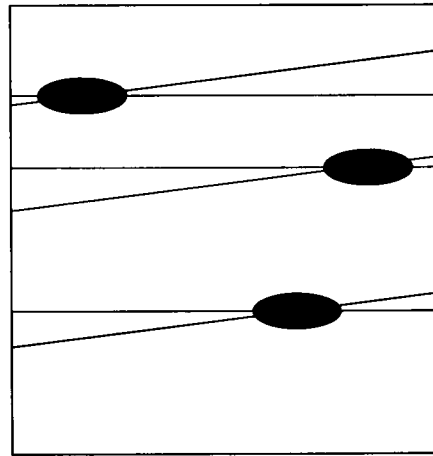


Figure 2.7: Three hits in a SAS layer without ambiguities.

Water/glycol coolant for the silicon and electronics flows within internal channels that are machined into the beryllium bulkheads at each barrel end. The readout electronics for the SVX II is mounted as close as possible to the sensors, due to readout speed and capacitance reasons. The large instrumented length of silicon along the beam pipe requires these electronics to be located within the active sensitive volume, with resulting negative consequences on impact parameter resolution.

A layer of silicon called **Layer 00** is mounted directly onto the beam pipe at very small radius to mitigate these effects. This layer is single-sided, for capacitance and space reasons, to minimise material, and to allow large bias voltages to be used to ensure depletion even after extensive radiation damage.

Outside SVX II there are three additional layers of double-sided silicon (SAS) – the **Intermediate Silicon Layers (ISL)** – positioned at a radius of 20, 23 and 29 cm from the collision point. As can be seen in figure 2.9 it consists of three barrels and accepts the wider region $|\eta| < 2$. The ISL ladders are composed of six wafers, arranged as half-ladders of three wafers each. A fixed strip pitch of $112\ \mu\text{m}$ is used on both the axial and 1.2° stereo sides. The ISL space frame, shown in figure 2.9, also supports the SVX II and all associated readout and utility components. The central barrel of the ISL consists of silicon ladders that are staggered alternately at either 22.6 or 23.1 cm in radius from the nominal beam line. This layer is extended by an additional set of ladders mounted at 19.7 or 20.2 cm in radius that are contained in the barrels at each end of the space frame. The outermost layer (L6) of the ISL is populated only in these end barrels, and consists of ladders mounted alternately at radii of either 28.6 or 29.0 cm. Temperature control is provided for the ISL electronics by a water/ethylene glycol coolant mixture flowing in aluminum tubes attached to beryllium ledges mounted on the space frame. Blockage of some of these tubes delayed the commissioning of the ISL. Although the cooling tubes have been cleared from these

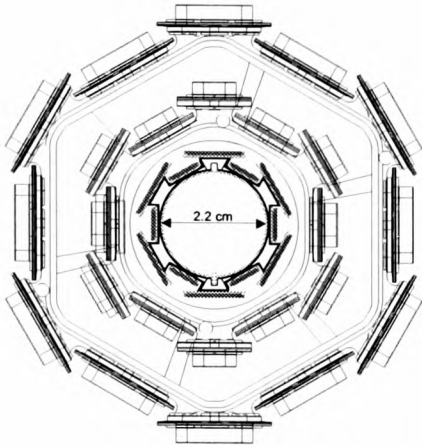


Figure 2.8: End view of the innermost three layers of the CDF Run II silicon system, showing Layer 00 along with the first two layers of the SVX II region.

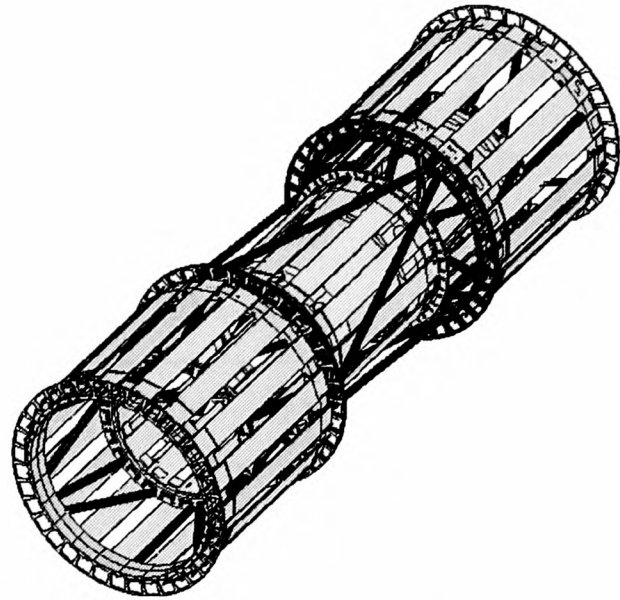


Figure 2.9: A perspective view of the ISL space frame and silicon placement. For clarity, only every other silicon ladder assembly is shown.

blockages, it was too late to use ISL for this analysis.

Table 2.2 summarises geometric information for each layer.

All SVX II sensors are about $300\text{ }\mu\text{m}$ thick. The total amount of material in the silicon system averaged over azimuthal angle and z varies roughly as 10% of a radiation length divided by the sine of the polar angle in the region of pseudorapidity $|\eta| < 1$. The average material traversed by particles increases to roughly twice this value for $1 < |\eta| < 2$ due to increased likelihood to encounter cables, cooling bulkheads, and portions of the support frame.

The complete CDF II silicon detectors have 211,968 readout channels in $r\phi$, and 193,536 readout channels in rz . The data acquisition system is a fully pipelined and buffered DAQ and trigger architecture that can operate without deadtime losses at machine bunch crossing intervals as low as 132 ns between cycles. An online trigger identifies displaced tracks at Level 2 and will operate with about $20\text{ }\mu\text{s}$ latency at rates up to 300 Hz. This trigger opens up a lot of new exciting hadronic decay channels of beauty and charmed hadrons. It will not play a rôle for this analysis though as we rely on the so-called J/ψ dimuon trigger (see section 3.6).

Present estimates indicate that the innermost SVX II sensors will degrade beyond usable levels after the first 2 to 3 years of operation, corresponding to several 10^{13} n/cm^2 . Layer 00, which is single-sided and thus can operate acceptably even when not fully depleted, should be able to withstand the higher radiation doses that it will encounter at its small inner radius. Along with the remaining layers of the SVX II and ISL, this should preserve functional $r\phi$ tracking and at least some stereo capability to as much as 5 fb^{-1} of accumulated Tevatron data.

System	Layer Name	Radius (cm)	Readout	Wedges	Strips		Strip Pitch (μm)	
					ϕ	z	ϕ	z
L00	L00a	1.35	ϕ	12	128	—	50	—
	L00b	1.62			256	—	50	—
SVX II	L0a	2.5325	ϕ , 90°	12	256	256	60	141
	L0b	2.9825						
	L1a	4.1075	ϕ , 90°	12	384	384	62	125.5
	L1b	4.5575						
	L2a	6.5075	ϕ , −1.2°	12	640	640	60	60
	L2b	6.9075						
	L3a	8.2075	ϕ , 90°	12	768	512	60	141
	L3b	8.7075						
	L4a	10.1325	ϕ , +1.2°	12	896	896	65	65
	L4b	10.6825						
ISL	L5a	19.71	ϕ , SAS	24	512	512	112	112
	L5b	20.215						
	L5c	22.585	ϕ , SAS	28				
	L5d	23.09						
	L6a	28.585	ϕ , SAS	36				
	L6b	28.99						

Table 2.2: Overview of silicon layers.

Without misalignments, the estimated intrinsic impact parameter resolution should improve from an average of $\sigma_{d_0} = (9 \oplus 50/p_T) \mu\text{m}$ for tracks at normal incidence without Layer 00 to an average of $\sigma_{d_0} = (6 \oplus 25/p_T) \mu\text{m}$ with the addition of this layer[10], for p_T in GeV/c.

The large geometrical acceptance ($|\eta| < 2$) of the silicon system does not only provide additional information for the tracks reconstructed in the COT ($|\eta| < 1$), but also opens up the challenging chance of standalone silicon track reconstruction.

2.4 Central Outer Tracking Chamber

The silicon system is surrounded by the **Central Outer Tracker (COT)**[61], a 310 cm long open-cell drift chamber with a radial range of $40 \text{ cm} < r < 137 \text{ cm}$ and a corresponding full acceptance of $|\eta| < 1$. It provides central tracking in the high luminosity environment by measuring charged tracks with transverse momentum as low as 400 MeV/c in the 1.4 T field of the solenoid with a resolution of $\delta p_T/p_T^2 \leq 0.1\%/ \text{GeV}/c$. Due to its larger radial extent its momentum resolution and purity is much better than the silicon tracker's. The expected large number of multiple-vertex events at high luminosity requires a good rz (stereo) reconstruction as well. The chosen gas mixture is Argon-Ethane-CF₄ (50:45:5) has a sufficiently high drift

velocity to resolve the beam crossings at 396 bunch spacing.

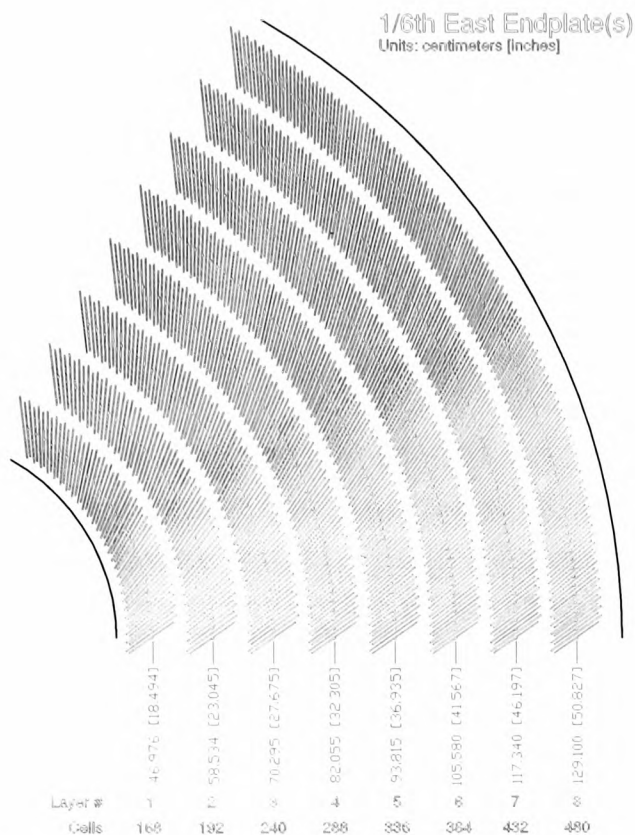


Figure 2.10: Layout of the endplate of the COT. The long slots are for field sheets, the short ones are for shaper panels and sense wires.

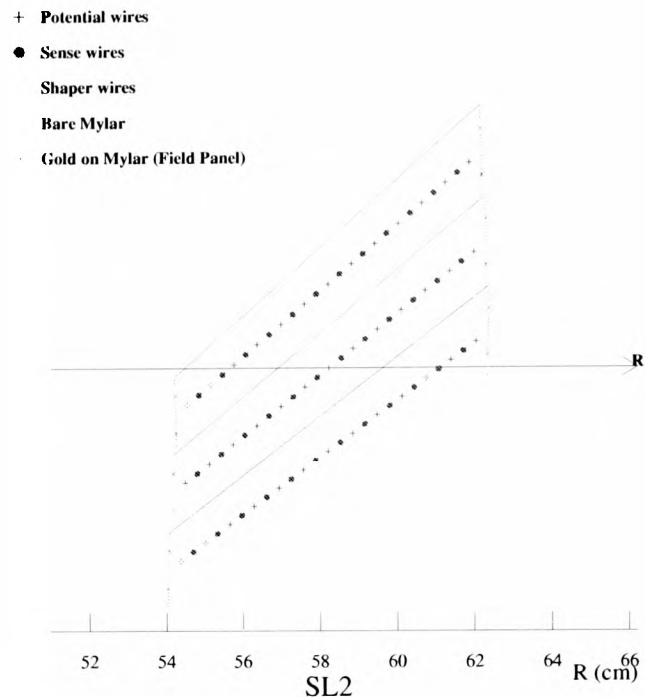


Figure 2.11: Nominal cell layout for SL-2. Other super-layers – including stereo – are similar.

Figure 2.10 reveals the design of eight superlayers (SL) of cells. Each of the eight superlayers consists of twelve layers of sense wires, alternating with potential wires. They are comprised in cells which are tilted by the 35° Lorentz angle. Figure 2.11 illustrates the cell geometry using superlayer 2 as an example. The cathode is gold on 0.25 mm Mylar, referred to as a “field panel”. The taper of the cell requires a varying potential between the sense wire plane and the field plane. To achieve this effect the sense and potential wire voltages are varied. The sides of each cell are closed, both mechanically and electrostatically, by Mylar strips with field-shaping wires attached, called the “shaper panels”. For both sense and potential wires 2 mm gold-plated copper-beryllium wire is used. The electric field gradient is about 2.6 kV/cm. Four of the superlayers are axial, alternating with stereo superlayers in the pattern $(+3, 0, -3, 0, +3, 0, -3, 0)^\circ$. Thus the COT provides in the ideal case 48 axial and 48 stereo measurements per charged particle traversing.

The trailing edge of the signal of a sense wire is logarithmically related to the total charge deposited. Converting this information into energy loss dE/dx allows particle identification to a certain degree. Studies [59] show that the K- π separation power for $p_T > 2$ GeV/ c is on the order of about one σ . Because of this weak separation power and not yet studied possible biases on the lifetime of reconstructed particles,

Number of Layers	96							
Super-Layer Number	1	2	3	4	5	6	7	8
Stereo Angle (°)	+3	0	-3	0	+3	0	-3	0
Cells per Layer	168	192	240	288	336	384	432	480
Radius at Centre of SL (cm)	46	58	70	82	94	106	119	131
Number of Sense Wires	30,240							
Number of Wires	63,000							

Table 2.3: Mechanical summary of the COT.

we do not use particle identification via dE/dx in this analysis.

The average total material of the central part of the COT corresponds to about 0.01 radiation lengths [62]. The single hit resolution for each layer is measured from data to be on the order of 150 μm [63].

2.5 Time-of-Flight (TOF)

The Time-of-Flight detector is new in Run II and was built to provide π -K-p separation for low p_T tracks, which enhances the tagging performance for the B physics program. In addition it allows rejection of cosmic rays. It consists of 216 scintillator bars ($2.8 \text{ m} \times 4 \text{ cm} \times 4 \text{ cm}$) and is located between the COT and the cryostat of the superconducting solenoid. Photomultiplier tubes, attached to both ends of each bar, provide time and pulse height measurements. The difference between the two time signals estimates the z coordinate of the particle crossing the bar and can be measured with a resolution of about 100 ps. The individual interaction time t_0 is determined offline to a precision of 30 ps by a fit to all PMT hits, assuming the tracks come from pions, kaons and protons with known production fractions. For a given track matched to a TOF hit the path length can be determined. Using the time difference $t_{hit} - t_0$ gives the speed β of the particle and can be translated into invariant mass using the known particle momentum.

The TOF system has a timing resolution of 100 ps and provides a K- π separation of about 2σ for tracks with $p_T < 1.5 \text{ GeV}/c$. For this analysis time-of-flight information could be useful to reduce background for low p_T $\phi \rightarrow K^+K^-$. To date the TOF reconstruction suffers from a very low efficiency about 55% of fiducial tracks due to high occupancy. Also possible biases on the proper lifetime have not been studied yet. For this reason we do not use time-of-flight information in this analysis.

2.6 Solenoid Magnet

A superconducting solenoid 3 m in length surrounds the tracking systems and provides a magnetic field of 1.4 T for measuring momenta of charged particles through their curvature. The field is uniform to 0.1 % in the region $|z| < 250 \text{ cm}$ and $|r| < 150 \text{ cm}$.

The solenoid is built of an Al-stabilised NbTi superconductor and operates at liquid helium temperature. Its return yoke guides the field lines over and underneath the detector and is also used as shielding for the upper and lower part of the CMP muon system (see section 2.8). The solenoid current is regulated by a feedback loop driven by a Nuclear Magnetic Resonance (NMR) probe, which also provides a measurement of the magnetic field. The typical time interval for the feedback is on the order of a minute and ensures a stable magnetic field in time. The solenoid and cryogenic equipment corresponds to 0.85 radiation lengths (X_0) of material at 90° incident angle.

2.7 Calorimeters

The solenoid and tracking volume are surrounded by the calorimeters, designed to measure the energy of photons, electrons and jets, and to determine the missing transverse energy as well. The calorimeter is divided into two physical sections: central ($|\eta| < 1.1$) and plug ($1.1 < |\eta| < 3.64$). Each section is subdivided into an electromagnetic (CEM, PEM) and hadronic part (CHA, PHA). The end-wall hadronic calorimeter (WHA) covers a gap between the central and plug sections, as shown in figures 2.2 and 2.3.

The calorimeter system is a scintillator sampling system with tower segmentation. The electromagnetic section of lead-scintillator is backed by the steel-scintillator hadron calorimeter. The signal is carried via wavelength shifters embedded in the scintillator to the photo-multiplier read-out. Table 2.4 summarises properties of the calorimeters.

	CEM	CHA/WHA	PEM	PHA
Coverage in η	< 1.1	< 1.3	$1.1 - 3.6$	$1.2 - 3.6$
Tower $\Delta\eta \times \Delta\phi$	$0.1 \times 15^\circ$		varies $0.1 - 0.16 \times 7.5 - 15^\circ$	
Thickness	$19 X_0, 1 \lambda$	4.5λ	$21 X_0, 1 \lambda$	7λ
Energy Resolution	$14\%/\sqrt{E}$	$75\%/\sqrt{E}$	$16\%/\sqrt{E} \oplus 1\%$	$80\%/\sqrt{E} \oplus 5\%$

Table 2.4: Characteristics of the CDF II calorimeters. X_0 is the radiation length and λ is the hadronic interaction length. The energy E is in GeV.

Proportional chambers (CES and PES) are embedded near the shower maximum at about 6 radiation lengths (X_0) within the electromagnetic calorimeters. These chambers have wires in $r\phi$ and cathode strips in z to record the 3D position of the shower, which helps to match showers to tracks. A second set of proportional chambers – the central or plug preradiator (CPR or PPR) – is placed between the front face of the electromagnetic calorimeter and the magnetic coil. Acting as shower pre-sampler, this chamber becomes useful in π^0 - γ separation and electron identification.

The calorimeters are used for this analysis only as passive absorbers for the muon system. Momenta of tracks from B mesons are only a few GeV/ c , thus much more

precisely measured from the COT than from the calorimeters.

2.8 Muon Systems

The outermost component of the CDF detector is a set of scintillators, drift tubes and steel absorbers, used for the detection of muons above $1.4 \text{ GeV}/c$. Muons are identified by their penetrating ability, and detected by chambers located outside substantial material. There are four muon detectors: the Central Muon Detector (CMU), the Central Muon Uppgrade Detector (CMP), Central Muon Extension Detector (CMX), and the Barrel Muon Detector (BMU). Their positions are labeled in figure 2.2. The four detectors are composed of layers of single wire drift chambers, which operate in proportional mode to survive the high luminosity and event rates.

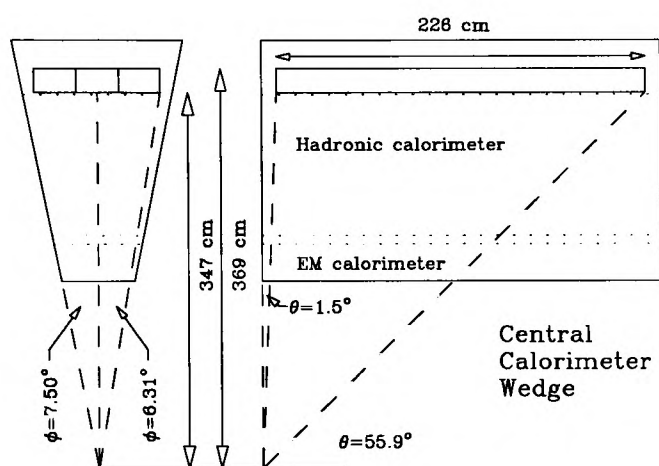


Figure 2.12: Three CMU sections embedded in a central calorimeter wedge at its outer radius.

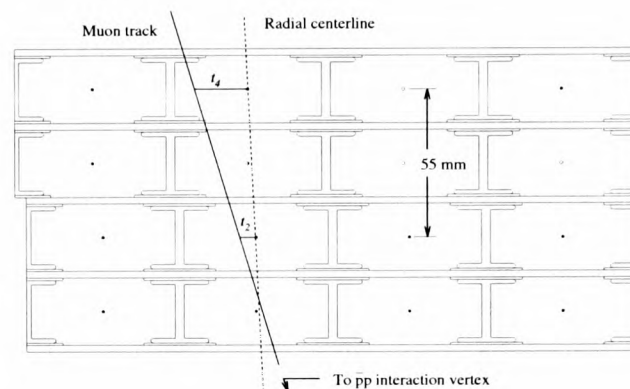


Figure 2.13: One of the three CMU towers in a wedge.

The **Central Muon Detector (CMU)** covers $|\eta| < 0.6$, and is embedded in the central calorimeter wedges at their outer radius (see figure 2.12), behind about 5.5λ of absorber (the CHA). A gap of 18 cm between east and west chambers (at $\eta = 0$) is due to the central calorimeter arch support structure and the high voltage fan out mounted on this end of the chambers. There are three separate sections (towers) of 16 cells (4 wide by 4 high) in each wedge, which makes 144 cells in total. Figure 2.13 shows one of the towers. Each cell is $(6.35 \times 2.68 \times 226) \text{ cm}$ in size and has a $50 \mu\text{m}$ stainless steel wire at the centre. The 16 cells in a module are stacked four deep in radial direction, with a small ϕ offset between the first pair and second pair of layers. Two adjacent wires in each layer are ganged at $z = 0$ to reduce the channel count as chamber occupancy is low enough to resolve the ambiguity. Each wire pair is instrumented with a TDC to measure the muon's location in ϕ , and an ADC on each end to measure the muon's location in z via charge division.

The maximum drift time of the drift chamber is on the order of 800 ns. This means that for bunch crossing times of 396 ns the muon systems will integrate over

several bunches. This is resolved by associating a muon signal with several bunches for the Level 1 trigger decision. The correct bunch is then determined by associating the muon signal with an XFT track and appropriate hadron calorimeter TDC time, which both are capable of resolving a single bunch crossing.

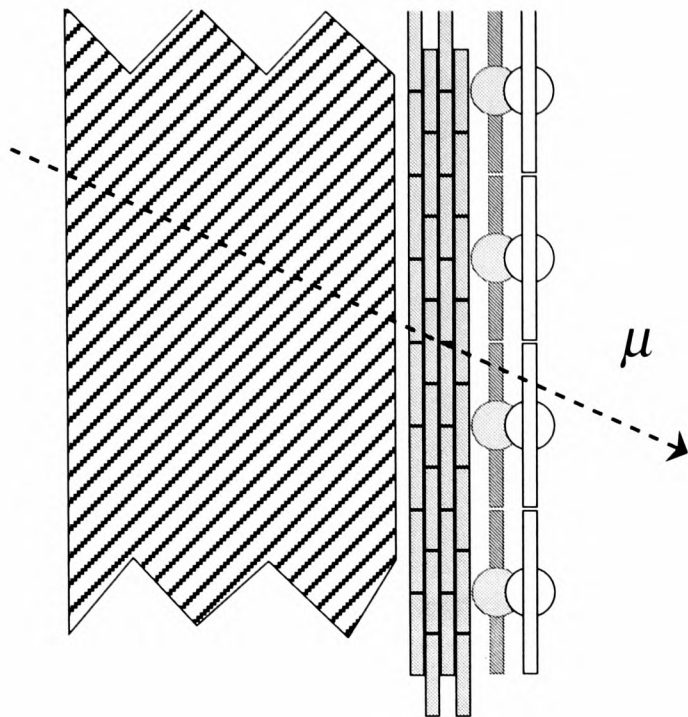


Figure 2.14: Detail showing the configuration of steel, chambers and counters for the Central Muon Upgrade walls.

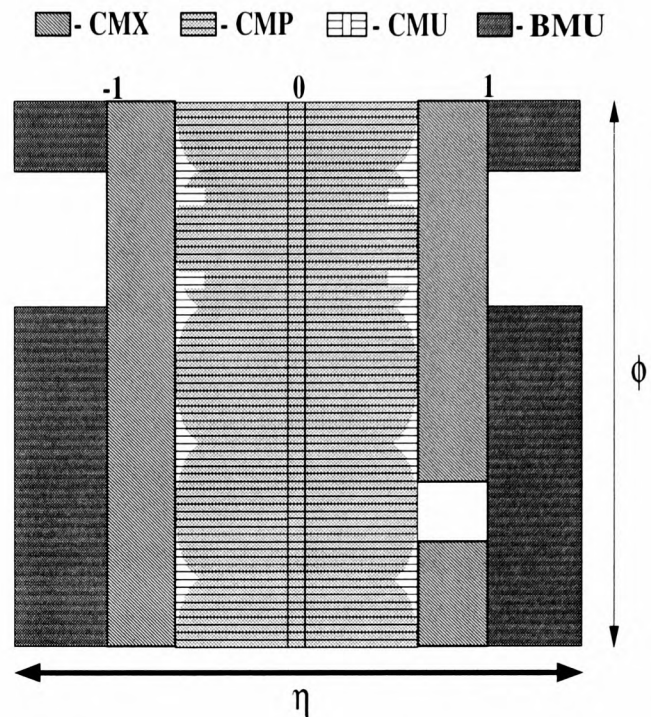


Figure 2.15: Coverage of the Run II muon systems in azimuth ϕ and pseudorapidity η .

The **Central Muon Upgrade Detector (CMP)** consists of a second set of muon chambers behind an additional 60 cm of steel in the region $55^\circ < \theta < 90^\circ$. It provides confirmation for CMU tracks but with reduced background. Its rectangular geometry (see figure 2.14) gives a pseudo-rapidity coverage which varies with azimuth as shown in figure 2.15. The return yoke of the CDF solenoid provides the necessary steel above and below the central detector. Steel has been added on the two sides with two non-magnetised walls.

On the outer surface of the CMP, a layer of scintillation counters are installed, which cover the upgrade chambers in width and half in length. They provide timing information which is used to associate the CMP muon signals to the appropriate event (since the drift time is longer than the bunch crossing time). Also, muon triggers which require CMP confirmation make use of the scintillation counters in conjunction with trigger information from the CMU chambers. Such triggers are not used for the dataset used for this analysis though. More details of the trigger system and muon triggers are given in the following section 2.9.

The **Central Muon Extension (CMX)** extends muon coverage to the pseudo-rapidity region $0.6 < |\eta| < 1.0$. It consists of free-standing conical arrays of 180 cm

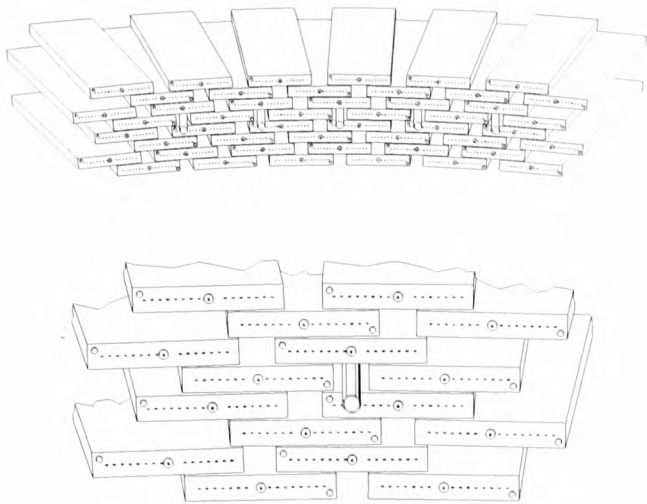


Figure 2.16: Overlapping geometry of CMX drift tubes.

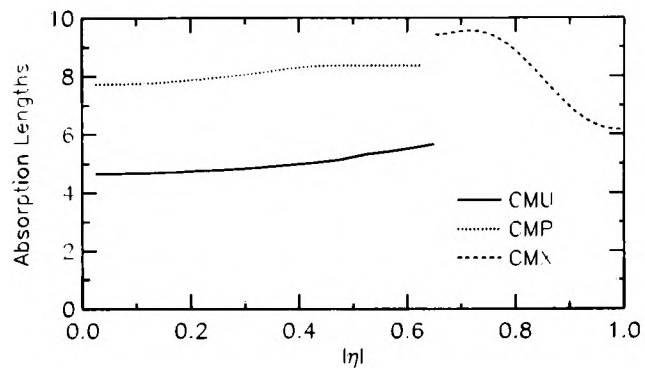


Figure 2.17: Number of absorption lengths as a function of pseudorapidity averaged over azimuthal acceptance of the CMU, CMP and CMX systems.

long drift tubes, with scintillators on both sides. There is azimuthal angular coverage from -45° to 75° and 105° to 225° . The two top wedges interfere with the CDF cryogenics penetration on the east side and with the main ring steel shielding on both sides. The bottom 90° part penetrates into the nominal floor of the collision hall and requires a different geometrical configuration than the upper 270° . During the data-taking for this analysis this part of the detector was turned off due to noise problems.

The drift tubes partially overlap each other with greatest overlap at the inner edge of the detector to be able to form a conical surface with rectangular cells (see figure 2.16). This gives a stereo angle of 3.6 mrad between adjacent cells allowing the measurement of the polar angle of the muon. On both the outside and inside a layer of scintillation counters is installed.

The **Barrel Muon Chambers (BMU)** is a horizontal cylindrical barrel of CMP-like drift tubes and scintillators, shielded by the steel of the Run I toroid magnets. In addition there is a vertical pinwheel of scintillation counters on the endwall ($1 < |\eta| < 1.3$) and two pinwheels ($1.3 < |\eta| < 1.5$, $1.5 < |\eta| < 2$) between the toroids for triggering. The BMU system will make it possible to trigger on forward muons, up to a pseudorapidity of 1.5, and to identify muons up to a pseudorapidity of 2. At the time the data used in this analysis was collected, there were noise problems in the chambers. BMU trigger have not been implemented yet and the BMU reconstruction software was still being developed. Therefore we do not use the BMU detector in this analysis.

Table 2.5 summarises essential properties of the CDF II muons systems. Figure 2.17 shows the number of absorption lengths as a function of pseudorapidity averaged over azimuthal acceptance of the CMU, CMP and CMX systems.

	CMU	CMP	CMX	BMU
η Coverage	0 – 0.6	0 – 0.6	0.6 – 1	1 – 1.5
Drift Tube Thickness (cm)	2.68	2.5	2.5	2.5
Drift Tube Width (cm)	6.35	15	15	8.4
Drift Tube Length (cm)	226	640	180	363
Maximum Drift Time (μ s)	0.8	1.4	1.4	0.8
Number of Drift Tubes	2,304	1,074	2,208	1,728
Scintillator Thickness (cm)	–	2.5	1.5	2.5
Scintillator Width (cm)	–	30	30 – 40	17
Scintillator Length (cm)	–	320	180	180
Number of Counters	–	269	324	864
π Interaction Lengths	5.5	7.8	6.2	6.2 – 20
Minimum Detectable p_T (GeV/ c)	1.4	2.2	1.4	1.4 – 2
Multiple Scattering Resolution (cm GeV/ p c)	12	15	13	13 – 25

Table 2.5: Summary of properties of the CDF II muon systems. Pion interaction lengths and multiple scattering are computed at a reference angle of $\theta = 90^\circ$ in CMU and CMP, at $\theta = 55^\circ$ in CMX, and show the range of values for the BMU.

2.9 The Trigger System

With the current 36 bunches the Tevatron has a beam crossing rate of 2.5 MHz. At luminosities of order $10^{31} \text{ cm}^{-2}\text{s}^{-1}$ there will be on average one interaction per beam crossing. The total cross section of proton-antiproton collisions is 75 mb, whereas interesting physics processes have cross sections of many orders of magnitude smaller. For example the cross section for $b\bar{b}$ is 100 μb . For $B_s^0 \rightarrow J/\psi \phi$ with $J/\psi \rightarrow \mu^+\mu^-$, $\phi \rightarrow K^+K^-$ it is 3.4 μb , for top production it is 6 pb, and for Higgs production it is of order 1 pb. This translates into signal to background ratios of 1.3×10^{-3} , 4.5×10^{-5} , 8×10^{-11} , and 1×10^{-11} , respectively. The trigger has the important task to select efficiently the most interesting events, while keeping the accept rate lower than the rate data can be written to permanent storage. At CDF the maximum data storage rate is about 50 Hz.

The CDF trigger system has a three level architecture with each level providing a rate reduction sufficient to allow for processing in the next level with minimal dead-time. Figure 2.18 shows the three level pipelined and buffered trigger system, for the design goal of 132 ns bunch spacing, although the machine is currently running at 396 ns. Level 1 uses programmable custom designed hardware to find physics objects based on a subset of the detector information and makes a decision based on simple counting of these objects. The latency of Level 1 is 5.5 μs to allow time for transmission and processing of the trigger signals to make the trigger decision. This requires each detector element to have local data buffering for the 14 beam crossings (at 396 ns separation) during the latency period. This is sufficient to average out the

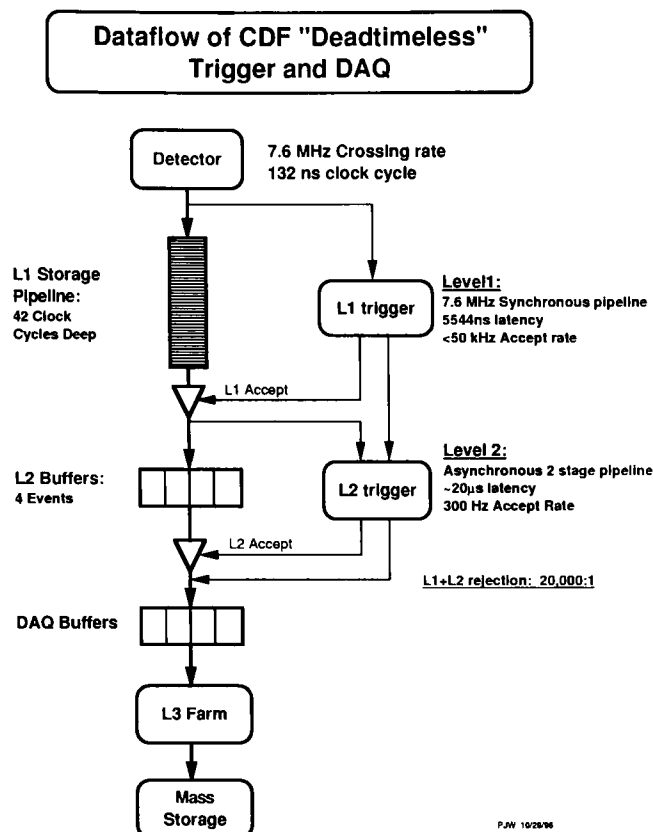


Figure 2.18: Block diagram of the three level pipelined and buffered trigger system. Each level provides a sufficient rate reduction to permit processing of the next level with minimal deadtime.

RUN II TRIGGER SYSTEM

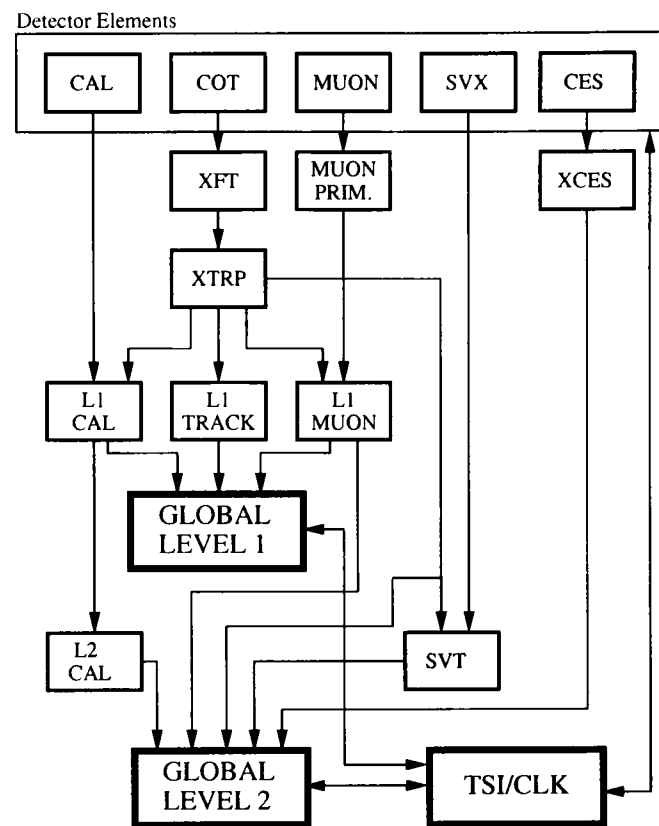


Figure 2.19: Block diagram of the Run II trigger system.

rate fluctuations and allow a 40 kHz Level 1 accept rate with $\leq 10\%$ deadtime for the anticipated 20 μ s Level 2 processing time. If an event is accepted by the Level 1 trigger, the front-end electronics move the data to one of four on-board Level 2 buffers. The Level 2 trigger uses custom hardware to do a limited event reconstruction which can be processed in programmable processors. All of the information used in the Level 1 decision is available to the Level 2 system, but with higher precision. In addition, data from the central calorimeter shower-maximum detector allows improved identification of electrons and photons. Jet reconstruction is provided by the Level 2 cluster finder; secondary-vertex information is produced by the Silicon Vertex Tracker (SVT). The data acquisition system will allow the Level 2 trigger to accept as many as 300 events per second. A Level 2 accept initiates full detector readout for the event. The Trigger System Interface (TSI) and Clock systems synchronise the trigger and data acquisition systems. The Level 3 trigger uses the full detector resolutions to fully reconstruct events on a farm of about 250 Dual Pentium Linux Personal Computers. The Level 3 reconstruction is very similar to the offline reconstruction code, and includes full three-dimensional track reconstruction and tight matching of tracks to calorimeter and muon-system information. The Level 3 accept rate is currently

smaller than 50 Hz.

A trigger path is composed of a specific Level 1, Level 2, and Level 3 trigger. An event passes a specific trigger path if the specific requirements are met at each level. The dataset used for this analysis is taken with two different trigger paths: the so-called JPSI_CMU1.5_CMx2 and JPSI_CMUCMU1.5, both low-momentum J/ψ dimuon triggers. They are described later. In the following section we describe in more detail the Level 1 trigger, with emphasis on the requirements for the J/ψ dimuon trigger paths.

2.10 The Level 1 Trigger

Figure 2.19 shows a block diagram for the the trigger system. The input to the Level 1 hardware comes from the calorimeters, tracking chamber and muon detectors. The decision to retain an event for further processing is based on the number and energies of electron, muon, and jet candidates, as well as missing transverse energy in the event. A Level 1 accept can also be generated based on the kinematic properties of observed track pairs.

The Level 1 hardware consists of three parallel synchronous processing streams which feed inputs into the single Global Level 1 decision unit. One stream finds calorimeter based objects, another finds muon objects, while the third finds tracks in the COT. Since the muon and electron triggers require the presence of a track pointing at the corresponding outer detector element, the tracks must be sent to the calorimeter and muon streams as well as the track-only stream. Up to 64 different streams can be formed using logical ANDs and ORs of objects from these streams. All elements of the Level 1 trigger are synchronised with the same 132 ns clock with a decision made every 396 ns by Global Level 1.

The goal of the **Level 1 calorimeter (L1 CAL)** trigger is to trigger on electrons, photons, jets, total event transverse energy and missing transverse energy. It is described in detail elsewhere [59] and is irrelevant for our analysis.

The **eXtremely Fast Tracker (XFT)** is a trigger track processor that identifies charged tracks in the COT. This trigger is at the heart of much of the physics that CDF hopes to do in Run II. It is needed to identify high momentum leptons for top, electro-weak, and exotic physics. It is also needed to identify low-momentum charged tracks for B physics analyses, such as this thesis. The minimum track p_T is 1.5 GeV/ c , and simply set by the fact that muons are stopped in the calorimeter for track momenta below this.

The XFT processor uses hit data from the four axial superlayers of the COT. A charged track passing through an axial superlayer will generate 12 hits of prompt and/or delayed data. The definition of a prompt or a delayed hit depends upon the maximum drift in the COT: 0 – 44 ns for a prompt, and 44 – 121 ns for a delayed hit. Track identification is accomplished in two processes by the Finder and the Linker.

The Finder searches for high- p_T track segments in each of the four axial superlayers. To do this quickly, each COT superlayer is divided into groups of four cells

each, with each group processed by a single Finder PLD¹, which compares hit data to a predefined set of patterns. A track segment is defined by the cell and whether a prompt or delayed hit was generated in each of the 12 wire planes within a superlayer. A collection of the cell numbers and hit types for the 12 wires in an axial layer is called a mask. The mask will change depending on the ϕ of the track, and its angle through the cell. The Finder works by comparing masks with all possible masks for tracks with $p_T > 1.5$ GeV/c, which are stored in a hard-wired look-up table.

The Linker searches for a four-out-of-four match among segments in the four layers, consistent with a prompt high- p_T track. If no track is found, the Linker searches for a three-out-of-three match among segments in the innermost three layers. Figure 2.20 illustrates this process. Each Linker module covers a 15° region of ϕ . There are 12 Linker PLDs that search for the best track in a region of 1.25° in ϕ . The XFT therefore logically divides the COT into 288 segments, each covering 1.25° . One track is allowed per segment. Every 132 ns, data from all 288 segments are sent from the XFT to the XTRP.

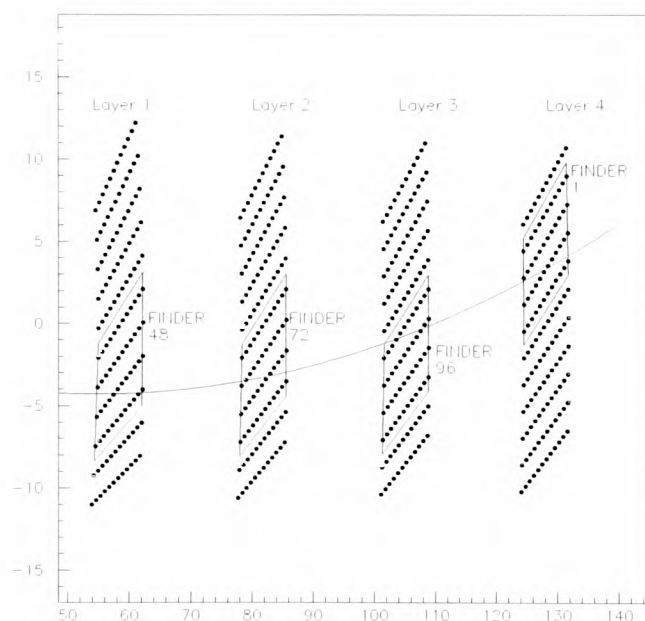


Figure 2.20: An XFT track in the COT. All cells in all four axial layers are shown. The relevant Finder in each layer is highlighted.

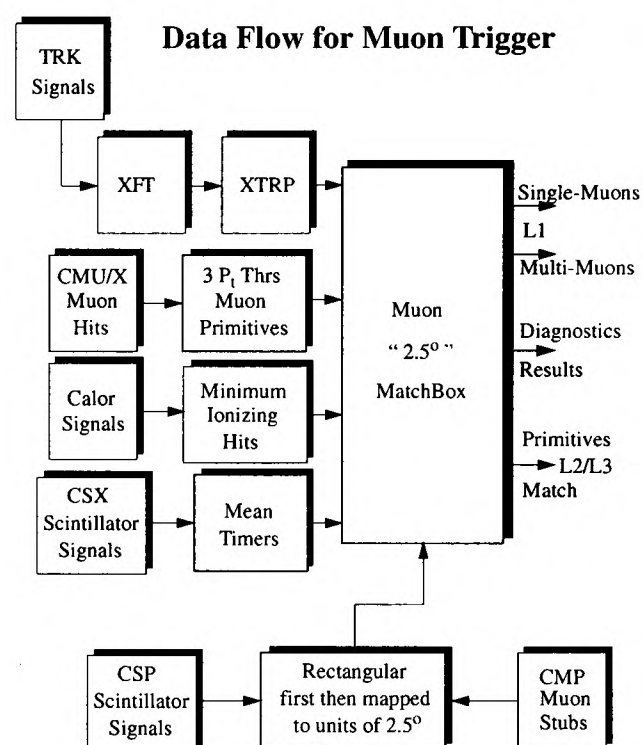


Figure 2.21: Data flow for the muon trigger.

The performance of XFT tracking is described in [64]. Table 2.6 lists some properties of the XFT, measured by matching XFT tracks with offline tracks from 10 GeV Jet trigger data at Level 1.

The purpose of the **Extrapolation Unit XTRP** is to receive tracks from the XFT and distribute the tracks or information derived from the tracks to the Level 1 and Level 2 trigger subsystems. After receiving the tracks from the XFT, signals

¹PLD = Programmable Logic Device

Efficiency	96%
Fakes	3 %
Momentum Threshold	1.5 GeV/ c
Momentum Resolution	1.7 $\frac{\%}{\text{GeV}/c}$
Angular Resolution	5.5 mrad

Table 2.6: XFT Performance measured from data [64].

are sent to the Level 1 Muon system (L1MUON), the Level 1 Calorimeter trigger (L1CAL), and the Level 1 Track Trigger (L1TRACK), as shown in figure 2.19. The tracks are also put into a pipeline and stored in Level 2 buffers upon receiving a Level 1 accept. They are then sent to the Level 2 processor and the silicon vertex trigger (SVT). In the following we only describe the part of the XTRP which is relevant for this thesis: the extrapolation to the muon chambers.

The XFT track is extrapolated to the radii of the CMU and CMX chambers by means of look-up tables. Bits are set corresponding to 2.5° segmentation (matching the segmentation of the CMU), are set according to p_T , ϕ , and amount of multiple scattering. These bits are sent to the L1MUON systems. Three separate p_T thresholds are available. For this thesis, the lowest p_T threshold of 1.5 GeV/ c is used.

The **Level 1 Muon Trigger (L1 MUON)** provides single and dimuon objects for the Level 1 trigger decision. There are 24 muon wedges each covering 15° in ϕ . In each CMU wedge, there are 12 muon stacks with four radially aligned layers per stack, and each stack covers about 1.25° in ϕ . In each CMX wedge, there are six muon stacks with eight radially aligned layers per stack, and each stack covers about 2.5° in ϕ . The sense wires in alternating layers are offset by known amounts. Therefore the drift velocity can be determined by comparing differences in drift time for alternating layers. In the CMU, the timing information read out from each adjacent pair of muon stacks is ORed, thus forming a 2.5° tower.

The Level 1 Muon Trigger starts from muon towers with at least one pair of hits from alternating layers in the muon chamber. These hits are combined to a trigger object called a **muon stub**. For the four layer CMU chambers, the pair of hits are required to be either from layer 0 and 2, or layer 1 and 3. For the eight layer CMX chambers, the combination of layers for hit pairs are (0, 4), (1, 5), (2, 6), and (3, 7). The stub timing window ($\Delta T < 396$ ns), where ΔT is the calculated time difference between hits in two radially aligned wires in alternating layers, corresponds to a rough p_T threshold cut of 1.5 GeV/ c at the stub level.

Using look-up tables the XTRP projects XFT tracks with $p_T > 1.5$ GeV/ c (CMU) or $p_T > 2$ GeV/ c (CMX) into the muon chambers. In more detail, given p_T , ϕ_{SL6}^2 and the charge of an XFT track and assuming that the track originates from the centre of the detector, the XTRP decides into which muon towers this track is likely to enter, multiple scattering and misalignment taken into account. The matching window is

² ϕ as measured in superlayer 6 of the COT

$\sqrt{(3 \cdot 2.5/p_T \text{ GeV}/c)^2 + 5^2}$ degrees in ϕ .

A Level 1 muon trigger decision is formed when a match is found between a fired Level 1 muon tower and an XTRP projection of an XFT track into the same tower. There are 12 of these muon match boxes collecting primitives from the muon chambers and XTRP to make the L1MUON decision. Each match box covers two wedges or 30° in ϕ . Figure 2.21 pictures the data flow for the muon trigger. The trigger efficiencies are very high and dominated by the XFT efficiency. The plateau value of the single CMU Level 1 trigger efficiency was measured by [65] to be 0.977 ± 0.002 , and 0.962 ± 0.003 for the CMX.

If the event has two XFT tracks matched to non-adjacent muon stubs, it is accepted by the **dimuon trigger**. There are two Level 1 dimuon triggers relevant for the dataset used in this thesis: The CMU-CMU with a momentum threshold of $1.5 \text{ GeV}/c$, and CMU-CMX triggers with $1.5 \text{ GeV}/c$ and $2 \text{ GeV}/c$, respectively. The trigger cross sections for those triggers are $(1,000 \pm 500) \text{ nb}$ for CMU-CMU, and $(300 \pm 70) \text{ nb}$. The trigger efficiencies are flat over the J/ψ p_T range of 2 - 20 GeV/c with a value of about 90% (see [66]).

The Level 2 muon trigger boards are yet to be commissioned. Therefore the events passing Level 1 dimuon triggers are automatically passed through to Level 3. Level 3 uses full event information, and uses reconstruction software similar to offline reconstruction. For this reason we describe the Level 3 trigger requirements for the J/ψ triggers in chapter 3.6, after offline reconstruction has been explained.

Chapter 3

Reconstruction

This chapter gives a brief overview over the offline reconstruction of the physics objects relevant for this analysis: tracks with silicon information, muons in the CMU and CMX, and the beam line.

3.1 Track Parameterisation

Charged particles in a homogeneous magnetic field travel on a helix whose axis is parallel to the magnetic field. At CDF the direction of the magnetic field is in negative z direction, following the antiproton beam's direction from east to west. A helix can be described with six parameters: three to parameterise the circle in the xy plane and two for the straight line in the rz projection. We use the following track parameterisation [67]:

- **Curvature C :** The projection of the track helix onto the transverse plane is a circle of radius r . The signed track curvature is defined as $C = q/(2r)$, where $q = \pm 1$ is the charge of the particle.
- **Impact Parameter d_0 :** It is defined as $[\hat{\mathbf{p}} \wedge \mathbf{d}] \cdot \hat{\mathbf{z}}$, where $\hat{\mathbf{p}}$ is the unit vector in the direction of the particle, \mathbf{d} the vector pointing from the primary interaction point to the point of closest approach. The unit vector in z direction is denoted by $\hat{\mathbf{z}}$. The absolute value $|d_0|$ is simply the distance of closest approach to the interaction point.
- **Azimuthal Angle ϕ_0 :** Azimuthal angle of the particle trajectory at the point of closest approach to the interaction point.
- **$\cot \theta$:** Cotangent of the angle θ between the z axis and the momentum of the particle. Sometimes $\cot \theta$ is called λ .
- **z_0 :** Position along the z axis at the point of closest approach.

3.2 COT Track Reconstruction

The COT hardware provides the drift time for each detected hit. In order to perform track reconstruction, this time measurement needs to be converted into a position measurement. Without corrections, a simple relation holds to relate the drift time t (time difference between wire signal and beam crossing) to the particle position:

$$\mathbf{x} = \mathbf{x}_{wire} \pm \mathbf{v} \cdot t, \quad (3.1)$$

with \mathbf{x}_{wire} being the wire position and \mathbf{x} the position in the local COT coordinate system. The drift velocity is denoted with \mathbf{v} . There is a \pm ambiguity as to which side of the wire the real particle passed. This ambiguity is resolved during track finding. In reality however, corrections are applied to account for field disuniformities and nonlinearities.

Segment Fitting

Once the position of COT hit candidates is known, the tracking algorithm scans each of the eight superlayers looking for line segments. These segments are used in subsequent steps, either as basic building blocks (in the Segment Linking algorithm) or as reconstruction seeds (in the Histogram Tracking algorithm).

Segment finding begins by looking for triplets of still unused, aligned hits belonging to consecutive layers.¹ Both possible choices of drift signs are taken into consideration for each hit. A list of candidate segments is formed by selecting the cases in which the central hit lies close enough to the midpoint of the external ones, and the overall slope α of the segment with respect to the radial direction is not too high.

Candidate segments are arranged in increasing order of $|\alpha|$ (so that high momentum tracks will be given precedence), and fitted to a straight line which is then extrapolated to the other layers of the superlayer. Then hits of adjacent layers which are close enough to the candidate segments are added, and the straight line fit recalculated. If for a given segment candidate, the number of wires without hit becomes bigger than the number of wires with a hit, the candidate is discarded.

Each COT hit only belongs to a single segment. In case a hit is shared between two different segments, it is assigned to the segment with the greater number of hits. The only exception to this rule applies to segments which lie entirely within a single cell. In this case, another valid segment (“ghost” segment) is obtained by flipping the drift sign of all the hits. Figure 3.1 illustrates the process of segment fitting. At the level of segment finding, the single hit resolution is about 240 μm .

Once segments are available, two different algorithms try to assemble them into tracks: the *segment linking* and *histogram linking* algorithm. Aiming for most efficient COT tracking, the track lists resulting from both algorithms are merged.

¹The terms *cell*, *superlayer* and *layer* give easily rise to confusion. A *cell* is the physical unit of 12 sense wires, tilted by the Lorentz angle. The cells are arranged in 8 superlayers. A *layer* is the collection of wires located at a fixed radius. Each layer contains exactly one sense wire from each cell of the superlayer in which the layer is in. A straight track going through the fiducial volume of the detector sees $12 \times 8 = 96$ layers.

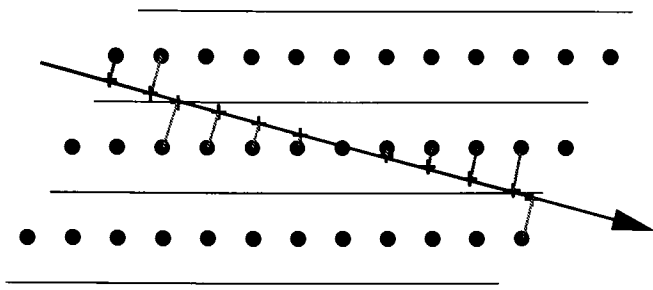


Figure 3.1: Segment with hits from three adjacent cells. The wires are indicated by circles, hit positions by crosses, and drift path by a short line.

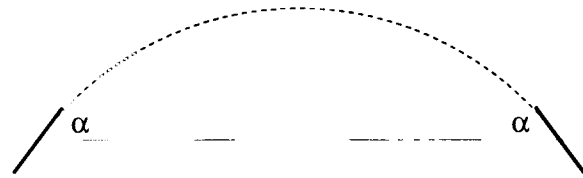


Figure 3.2: Geometrical principle of segment linking.

Segment Linking

The segment linking algorithm is done in two steps: at first, axial segments are joined in a 2D track; then, stereo segments and individual stereo hits are attached to each axial track. Axial segment linking is based on the simple geometry of segments belonging to the same track. Figure 3.2 shows that two segments are tangent to the same circumference when forming equal angles α with the line that connects their central points.

The algorithm starts taking the outermost axial superlayer's segments as seeds. Each outer segment is extrapolated toward the beam pipe, assuming that the track had a null impact parameter. This defines a region in the inner superlayers, in which other axial segments are looked for. All axial segments of the inner superlayers within this region are then compared to the current outer segments, as shown in figure 3.2. After segment linking for axial segments, the found axial hits are fitted to circles in the transverse plane. Then, stereo segments are linked to axial tracks. In case stereo segment linking fails to find a match on some superlayer, single stereo hits are matched to the track.

Histogram Linking

The histogram linking process starts from a given so-called “telescope”, in our case a single axial segment. A histogram is constructed for each of the layers that could contain hits belonging to the track. Each histogram is filled with the log-likelihood (as a function of track curvature) that the telescope track, shifted by an appropriate distance, might pass through a hit belonging to the current layer. All the histograms are then summed together. Ideally, all the contributions due to hits belonging to the same telescope track should have the same curvature. This results in a sharp peak in the sum histogram. By looking for this peak, it is possible to determine the set of hits that are best aligned to the telescope. It is worth noting that the choice of hits assigned to a track only depends on the sum of all contributions, not on the order in which these contributions are provided.

3.3 Outside-In Silicon Tracking

The outside-in tracking algorithms [68] take advantage of the excellent COT pattern recognition, by using COT tracks as seeds. The high spatial resolution of the silicon is deployed by progressively adding silicon hits to the track. Since the z resolution of the COT is about an order of magnitude worse than the $r\phi$ resolution, stereo linking of COT to silicon is much more challenging and needs special care. For this reason it is necessary to separate the z from the $r\phi$ coordinates. Two outside-in tracking strategies are currently used at CDF: **outside-in tracking** in the $r\phi$ plane and **outside-in z tracking** in all three dimensions. Outside-in tracking requires an excellent alignment between COT and silicon detectors and the COT error matrices properly calculated.

Given a COT track the algorithms progressively search inwards for silicon hits. At each sensor they add silicon hits that lie in a 4σ road around the track-sensor intersection point. The track error matrix is propagated to the intersection point and takes energy loss due to multiple scattering in the silicon layers into account. A new track candidate is generated for each hit in the road. The track itself is also included in case the true hit is missing. Only the two candidates with the greatest number of silicon hits found so far are extrapolated to the next layer of silicon, where the procedure is repeated. At the end of this process the candidate with the largest number of silicon layers is chosen to be the best candidate. For the outside-in z tracking stereo hits are added in a similar fashion.

Silicon Standalone Tracking

The basic idea of picking up silicon hits in the Standalone tracking [69] is similar to the outside-in tracking. The significant difference is that the seed tracks have to be obtained first. In case of outside-in tracking there are the COT tracks. For the Standalone tracking they have to be constructed from the hits in the silicon. The drift chamber is at larger radii than the silicon tracker, resulting in a lower relative track density. Tracks in the COT are more isolated which leads to less combinatorics for the track reconstruction. Thus tracking in the COT is much purer and faster than in the silicon. For the outside-in tracking there are up to 200 COT seed tracks depending on event type, and an outside-in track is expected for most of them. For the Standalone tracking there are up to 50 000 seed tracks and only a few of them are real tracks. The major problem of the Silicon Standalone strategy is to distinguish between real and fake seed tracks. For this reason the Standalone tracking is only run on the remaining unused hits after outside-in tracking has been performed. Thus the main purpose of the algorithm is to find tracks in the forward region $1 < |\eta| < 2$, where four SAS layers are located.

Seed tracks are required to have two out of four stereo hits. In addition the two stereo hits have to be consistent in ϕ within $\Delta\phi = \pm\pi/20$. This corresponds to a weak cut on transverse momentum $p_T > 100 \text{ MeV}/c$. The straight line in the rz plane has to be consistent with one of the primary vertices found by the pre-tracking primary vertex finder [70] within $\pm 5 \text{ cm}$. The pattern recognition step is similar to

Track Parameter	Error Scale Factor
$\cot\theta$	$\sqrt{1 + 0.580/(p_T^2 \cdot \sin^3 \theta)}$
z_0	$\sqrt{1 + 0.653/(p_T^2 \cdot \sin^3 \theta)}$
C	$\sqrt{1 + 5.33/p_T^2}$
d_0	$\sqrt{1 + 3.01/p_T^2}$
ϕ_0	$\sqrt{1 + 3.70/p_T^2}$

Table 3.1: COT track parameter error scaling factors. Transverse momentum p_T is assumed to be in unit of GeV/ c .

the outside-in tracking and is performed for one track candidate after another. Once a track is accepted, its hits are flagged as already used. The list of track candidates is sorted according to a ranking criterion which combines p_T and hit pattern criteria in a way that the candidates which are trusted most are being processed first.

3.4 Tracking Corrections

There are two levels of different offline track quality. The first level is the standard CDF reconstruction, which uses a fixed version of alignment, and calculates the energy loss corrections assuming a pion hypothesis using the GEANT geometry description. These tracks are used for a rather loose pre-selection of the decay modes under study, and efficiency losses due to the track quality of that level cannot be recovered by subsequent corrections.

It was found [71] that the track covariance matrix of the COT reconstruction is underestimated and shows a p_T and θ dependence. It is therefore scaled to the correct value, implementing parameterised effects of multiple scattering in the COT as measured in a GEANT simulation study. The used error scale factors are listed in table 3.1.

After the pre-selection, for every candidate the particle hypothesis is known, i. e. whether we assume the track comes from a π , K, or μ . The silicon tracking algorithm is performed again under the new particle hypothesis to take into account the energy loss, which is particle type specific. This time two additional corrections are applied: missing material in the geometry description, and a corrected value for the magnetic field. In addition the latest silicon alignment constants are used. These two corrections come from a calibration to the observed $J/\psi \rightarrow \mu^+\mu^-$ signal [72], [73]. Figure 3.3 shows how this calibration was performed. Each data point is the fitted J/ψ mass for different bins in p_T and different stages of corrections. In the mass fit special care is taken to include the small but significant radiative tail in order to determine the J/ψ mass correctly. The lowest curve is the J/ψ mass without energy loss corrections. A significant shift from the PDG value [12] $m_{J/\psi} = (3096.87 \pm 0.04)\text{MeV}/c^2$ and a linear p_T dependence is visible. The second curve includes energy loss corrections under μ hypothesis and is therefore shifted to higher masses and is less steep.

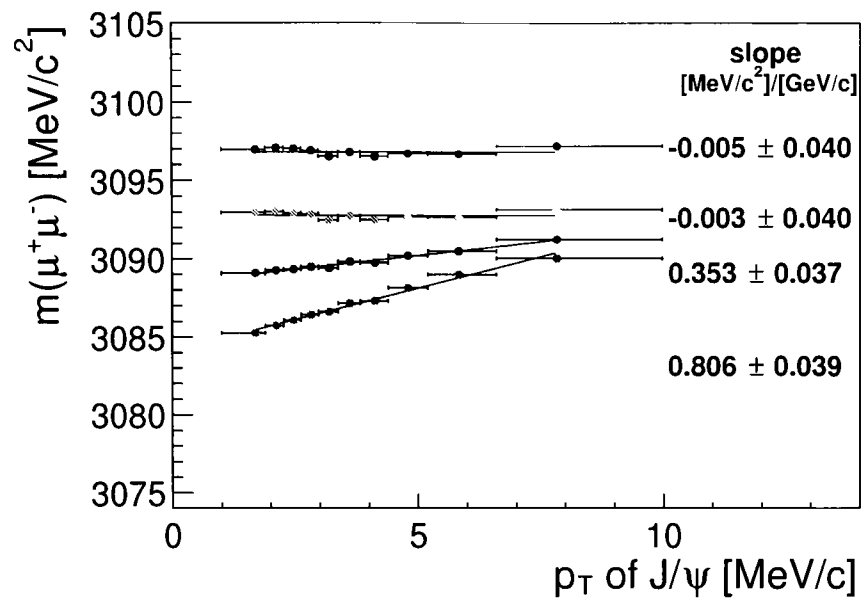


Figure 3.3: The J/ψ mass as a function of its p_T . Four versions of the data are shown: The raw mass with only the radiative bias correction is the lowest and steepest set of points, above which the mass after nominal energy loss corrections have been included. The flat points above are the result after energy loss corrections, where a small additional layer has been added to the nominal material model. The top set of points is the final results after magnetic field corrections and energy loss corrections are applied.

A particle's energy loss is almost constant at high momenta. This implies that the relative effect becomes smaller as the momentum of the particle increases – therefore the J/ψ mass rises with p_T when energy loss has not been fully accounted for. The remaining p_T dependence is attributed to missing material in the geometry description of the detector, which is known to be accurate up to about 20%. An additional phantom layer of silicon is added ad hoc to the geometry description at a radius of 34 cm to account for the missing material in an average manner. At the somewhat arbitrarily chosen radius of 34 cm the geometry description allows the inclusion of the additional layer. Its thickness is tuned to make the J/ψ mass flat in p_T . This gives a thickness of 0.25 cm, corresponding to an 11 % under-estimation of present material. The third curve shows the J/ψ mass after this correction has been performed. Figure 3.4 shows the energy loss of all muons in the J/ψ sample, before and after including the additional material layer. The remaining discrepancy between the measured J/ψ mass and the PDG value questions the correct knowledge of the overall magnetic field scale. The magnetic field was found to be stable in time and field non-uniformities are taken care of in track reconstruction. Increasing the nominal magnetic field from 1.41160 T to 1.41394 T shifts the J/ψ mass to the PDG value. The credibility of this correction is backed up by mass measurements of other particles, such as K_S and D mesons on the lower mass scale, and Υ on the higher mass scale. All these masses agree with their PDG value after the magnetic field correction has been applied.

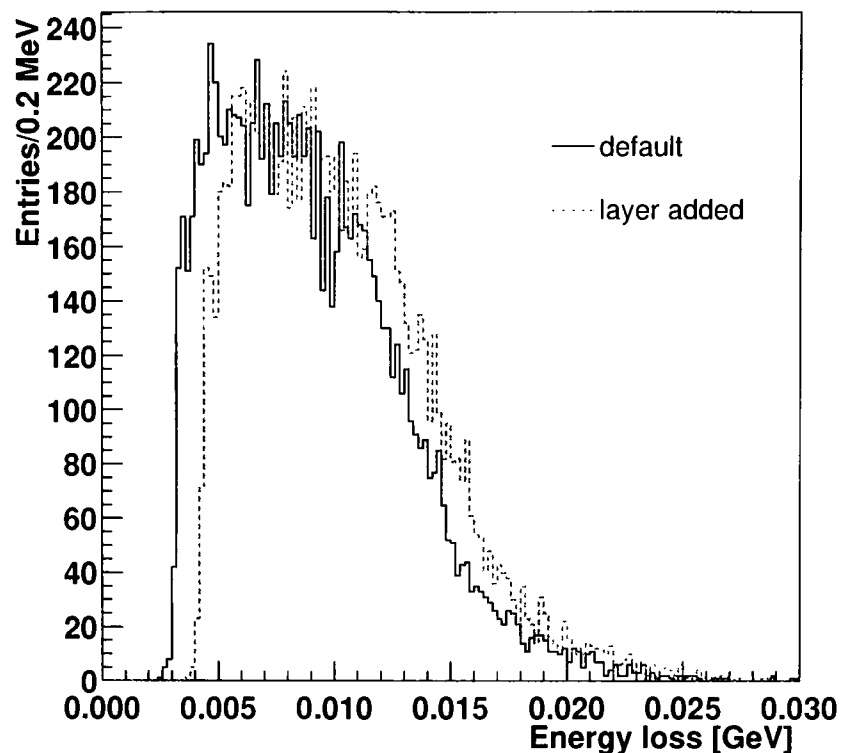


Figure 3.4: Energy loss of all muons in the J/ψ sample. The average energy loss is 9 MeV without additional corrections and 10.2 MeV after including an additional layer of missing material of thickness 0.25 cm.

Tracking Performance

Figure 3.5 shows the calculated errors on p_T , ϕ_0 , z_0 and d_0 for silicon outside-in tracks, where silicon stereo hits, Layer 00 and ISL hits have been dropped. This was done to obtain tracks similar to the ones used in the subsequent analysis. For the same purpose we require more than two silicon $r\phi$ hits, $p_T > 1$ GeV/ c , $|z_0| < 50$ cm, and $|\eta| < 1$. The errors include all the corrections mentioned above. The green (grey) solid histogram is for the outside-in tracks and the black non-solid histogram is for its parent COT track. It can be seen that the p_T and z_0 resolution are due to the performance of the COT and not much improved by silicon information (remember that silicon stereo information is not considered). The ϕ_0 resolution improves with silicon information by a factor of roughly four, whereas the d_0 resolution improves dramatically. The impact parameter resolution of high momentum tracks has been independently measured to be (23 ± 3) μm [63] from back-to-back tracks from $Z^0 \rightarrow \mu^+\mu^-$. This is in agreement with the d_0 error distribution, where the high transverse momentum tracks indeed populate the high resolution tail.

3.5 Muon Reconstruction

A muon passing through one of the muon detectors will leave hits in the drift chamber layers. A short reconstructed track in the muon detectors is called a **muon stub**. It is required to have hits in at least 3 chambers. A muon candidate is formed by linking a stub to a track from the tracking detectors.

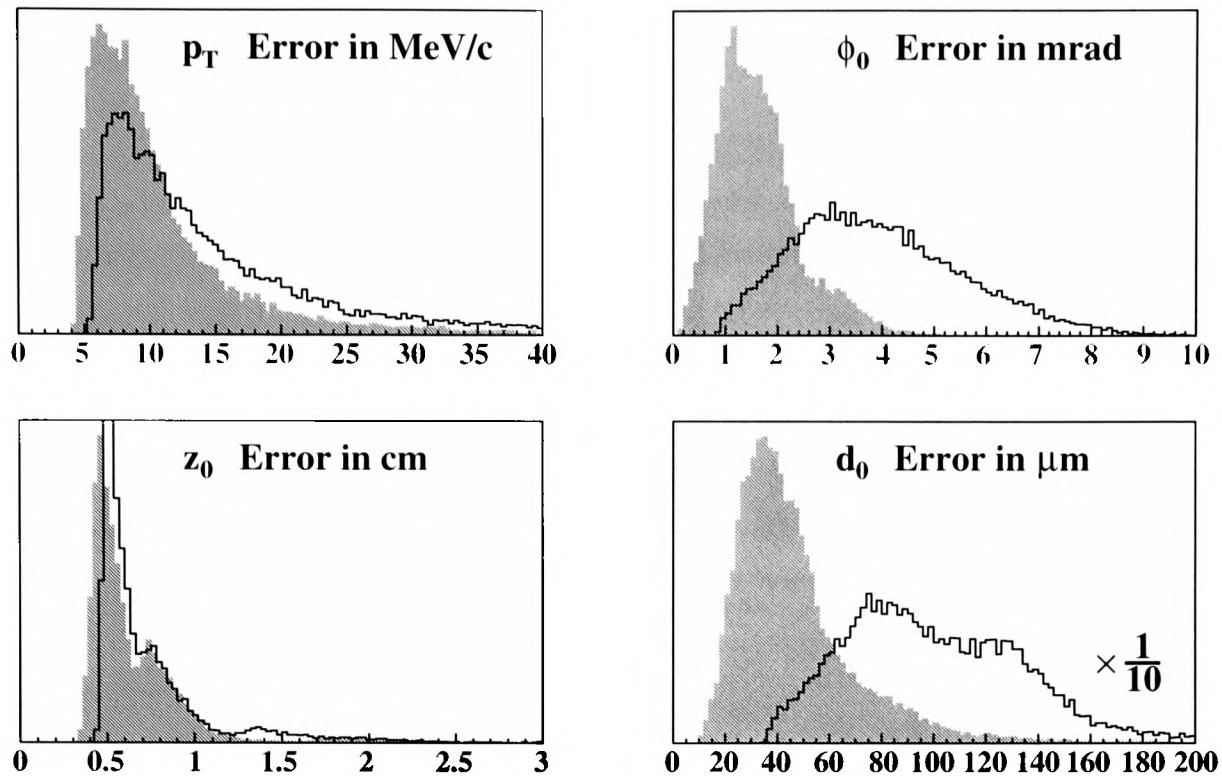


Figure 3.5: Errors of track parameters (green/grey solid histogram) for outside-in silicon tracks with at least three silicon hits in $r\phi$, stereo hits dropped, $p_T > 1\text{GeV}/c$, $|z_0| < 50\text{ cm}$, and $|\eta| < 1$. The black non-solid histogram is for the parent COT track. The d_0 error for the COT parent track was scaled by a factor of $1/10$.

The algorithms for **stub finding** are the same for CMU and CMP. The geometry of the CMX detector requires a more refined algorithm, which is based on the same basic ideas. For this reason we only describe the CMU/CMP stub finding algorithm and refer to [74] for more information on the CMX algorithm.

The stub finding algorithm loops over appropriate-sized regions (wedges for the CMU, regions for the CMP). In each region, the finder counts the number of hits, and the number of layers that have hits on them. Regions with fewer than three hits are not considered. Once the finder has chosen a region to search for stubs, it forms four vectors of hits, one for each detector layer, containing all the hits in that region and layer. The finder examines all pairs of hits taken from layers 0 and 2 that are within a specified distance from each other in x (7.5 cm for CMU, 12 cm for CMP). After that it looks for hits in layers 1 and 3 that are in a specified road (0.5 cm for CMU, 1.5 cm for CMP) around the assumed position, assuming a linear muon stub without any drift model applied. All possible combinations of hits are considered.

For each of these stub candidates a fit is performed, assuming a specific drift model. Figure 3.6 pictures the simple geometry of the drift model used. The stubs are then ranked by their number of hits and χ^2 of residuals – the distances between hit position and reconstructed muon stub, along the wire planes. Figure 3.7 shows a simplified example of a muon stub, with residuals drawn as dashed lines. Lower-ranking duplicates are removed.

The **track-stub linking** algorithm attempts to use the reconstructed stubs to

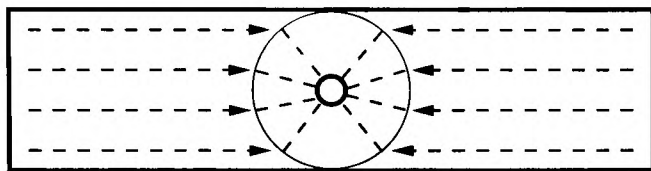


Figure 3.6: Simple linear and radial geometry of the muon drift model for one muon drift chamber.

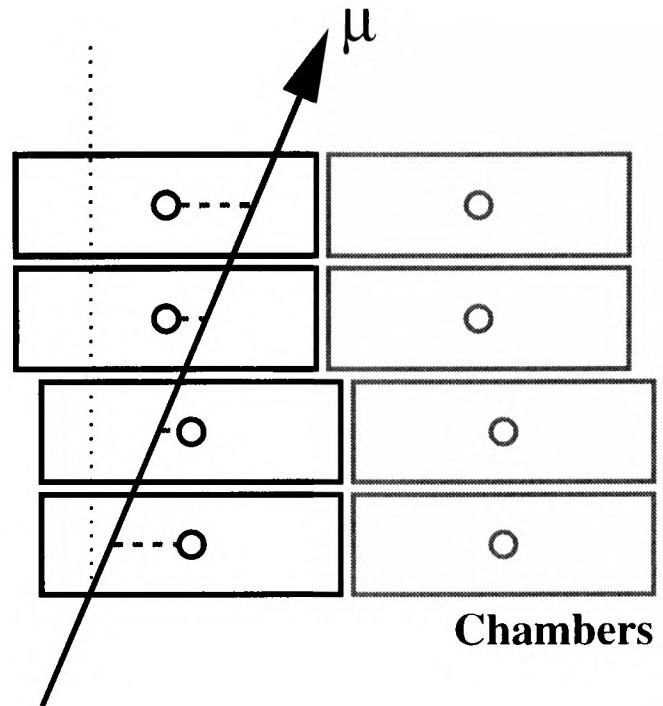


Figure 3.7: Muon stub reconstructed in the CMU chambers. The drift distances are drawn as dashed lines.

identify which tracks (reconstructed in the SVX, COT, or both) are most likely to belong to muons. This is done by extrapolating the tracks to the stubs, through the inhomogeneous magnetic field and dense material, and seeing whether they are close to each other. For this two matching variables are used: Δx and z_{track} . Δx is the distance in $r\phi$ between the stub and extrapolated track, whereas z_{track} is the z coordinate of the extrapolated track at the muon detector. There is the option of either cutting directly on Δx or on the χ^2 of the fit used in the matching procedure. Table 3.5 lists the imposed track quality cuts and track-stub matching criteria. No minimum ionising selection criteria are imposed to form a muon candidate.

Track Quality			
p_T	$> 1.3 \text{ GeV}/c$		
Number of axial COT hits	≥ 10		
$ d_0 $	$< 6 \text{ cm}$		
$ z_0 $	$< 200 \text{ cm}$		
Track Stub Matching			
	CMU	CMP	CMX
$\Delta x \text{ (cm)}$	< 30	< 60	< 50
$ z_{track} \text{ (cm)}$	$z < 250$	$z < 330$	$325 < z < 550$
Track and stub in same detector half	yes, unless $z < 20 \text{ cm}$	yes	yes

Table 3.2: Track-stub matching criteria for CMU, CMP and CMX.

There are a lot of hadrons in the events, that can mimic a muon, either by so-called “punch-through” or “decay-in-flight”. Punch-through occurs when a hadron does not get stopped in the hadron calorimeter or particles from the hadron shower reach the muon chambers and fake a muon. On the other hand there are real muons from decays of pions and kaons. Since those muons do not come from the main process under study, they are considered background muons as well. The muon fake-rate of single low- p_T muons is on the order of 50% [75], but a detailed study has yet to be done. In our case we require two muons to be consistent with a J/ψ , which reduces the fake-rate by a large factor (see chapter 4.1). The offline muon reconstruction efficiency for fiducial low p_T muons is measured from data (J/ψ events) to be well above 95% [76].

3.6 The J/ψ Dimuon Triggers

We use a dataset that passes one of two different J/ψ dimuon trigger paths:

JPSI_CMUCMU1.5 and JPSI_CMU1.5_CMX2.

They both require two muons: two muons in the CMU for JPSI_CMUCMU1.5, and one in the CMU and one in the CMX for JPSI_CMU1.5_CMX2. The momentum threshold for CMU muons is 1.5 GeV/ c and 2 GeV/ c for CMX muons. At Level 1 two XFT tracks with p_T above the momentum threshold are required, matching low p_T muon stubs in the CMU and/or CMX. Tracks which pass within 1.5 cm of the centre of the COT wire planes in any of the four axial layers are excluded. Also, due to a problem in a specific muon matchbox, muon stubs that fall in the range 240° to 270° are excluded. Since the Level 2 muon trigger boards are yet to be commissioned, Level 2 passes events through to Level 3, where fully reconstructed tracks and muons are available in offline quality as described in the previous chapter. Table 3.3 lists all cuts applied at Level 3 and table 3.4 shows the trigger cross sections measured from data.

Variable	CMU1.5	CMX2.0
Charge Product $q(\mu_1) \cdot q(\mu_2)$	-1	
$ z_0(\mu_1) - z_0(\mu_2) $	< 5 cm	
$\Delta\phi(\mu_1\mu_2)$	< 130°	
Mass Window (GeV/ c^2)	$2.7 < M(\mu_1\mu_2) < 4.0$	
$p_T(\mu)_{COT}$ Threshold	> 1.5 GeV/ c	> 2 GeV/ c
Stub-Track Match Δx	< 30 cm	< 50 cm

Table 3.3: Level 3 cuts for the J/ψ triggers.

	JPSI_CMUCMU1.5	JPSI_CMU1.5_CMx2
Level 1 (nb)	$1,000 \pm 500$	300 ± 70
Level 3 (nb)	32.2 ± 6.4	8.0 ± 1.6

Table 3.4: Measured J/ψ trigger cross sections.

3.7 The Luminous Region

The beam profile is determined by finding the interaction vertex (vertices) for each event of a given run [77]. The iterative algorithm performs a fit of all silicon tracks to a common vertex. For silicon inside-out tracks at least 4 silicon $r\phi$ hits, and for silicon stand-alone tracks at least 5 $r\phi$ and 2 stereo hits are required. In each iteration cycle the track with the highest χ^2 contribution greater than 16 is removed. The vertex fit is repeated until all the remaining tracks give a χ^2 contribution of less than 16.

The transverse profile of the luminous region at CDF can be described by Gaussian functions in x and y . The means of these Gaussians define the beam position. The widths vary along the length of the interaction region due to the beam focusing. Typically the beam position is around $(x, y) = (-2, 5)$ mm and tilted by about 0.6 mrad in x and 0.2 mrad in y . The typical width is about 30 μm in the central region and tends to grow with $|z|$. In our analysis we treat the beam width as constant in z and study variations of z as systematic effects. In z the luminous region can be described by a Gaussian with mean at nought and a width of 25 cm.

The beam positions and profile are calculated and stored on a run by run basis. For most runs the beam position does not change during a run. There are a few long runs identified to have a beam position changing slowly in time by a few μm . For them the beam position is calculated for small enough sections, where the beam position is quasi-stable in time.

The option of using an event-by-event interaction vertex rather than a run-averaged beam position is not considered. Events from $b\bar{b}$ are of fairly low multiplicity. After removing displaced tracks from b events it is not clear that enough prompt tracks remain in the event to ensure interaction vertex finding with high efficiency and better spatial resolution than from the run-averaged approach. Also tracks coming from the opposite B hadron might bias the interaction vertex away from the B meson of study.

Figure 3.8 shows for a particular run the impact parameter of all tracks with $p_T > 1.5\text{GeV}/c$ and at least three $r\phi$ silicon hits. The impact parameter is calculated with respect to the beam line. The left plot shows no ϕ_0 dependence, which would be expected in case of an offset beam line. On the right the impact parameter distribution is overlaid with a fit of two Gaussians plus flat background ($\chi^2/n_{dof} = 1.03$). A fraction of about 0.65 is contained in a Gaussian with $\sigma = 42.1 \mu\text{m}$. This resolution includes a 30 μm beam width uncertainty. Therefore the intrinsic impact parameter resolution is about 33 μm .

In this chapter we have shown the principles and performance of track and muon

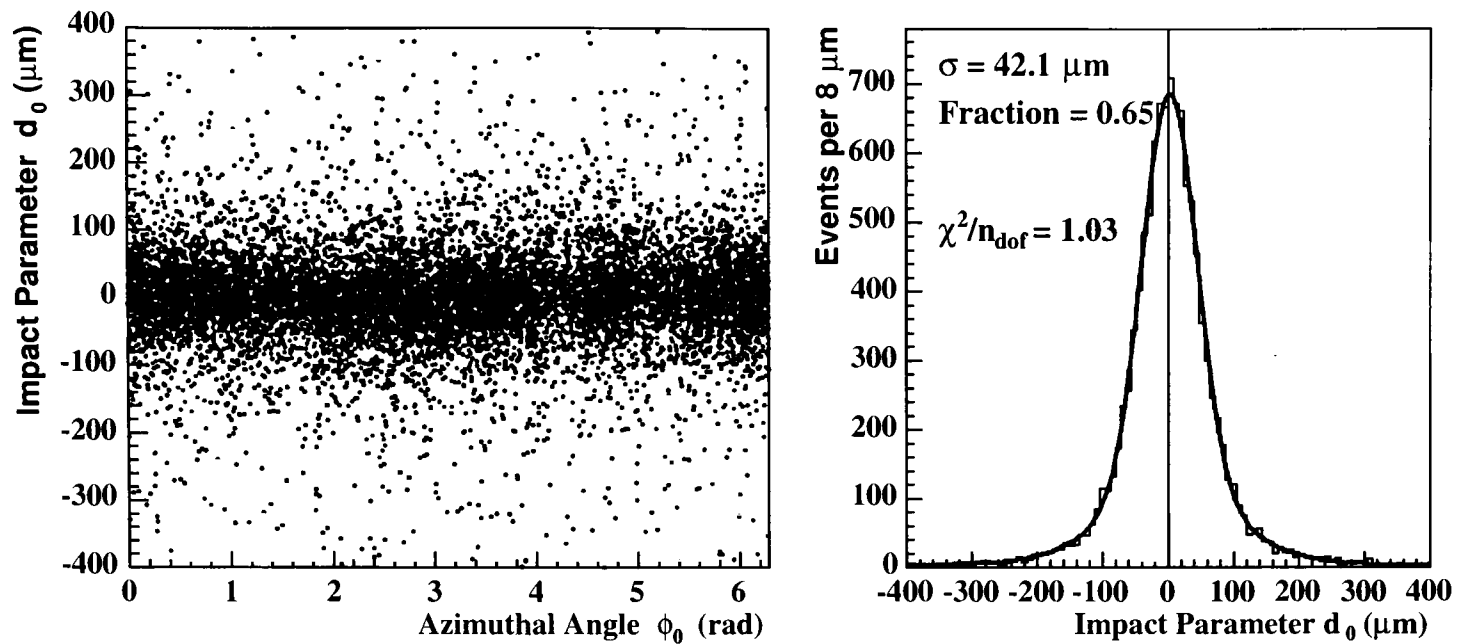


Figure 3.8: Impact parameter distributions with respect to the measured beam line.

reconstruction. We can reconstruct tracks with $p_T > 0.4 \text{ GeV}/c$ and muons with $p_T > 1.5 \text{ GeV}/c$. The silicon systems provides an excellent impact resolution of $42 \mu\text{m}$, which will be necessary for a lifetime measurement of B mesons. The J/ψ dimuon trigger performs well, and in chapter 4 we will see that they collect a sample which contains B_u , B_d and B_s mesons.

Chapter 4

Selection of $B_s^0 \rightarrow J/\psi \phi$

In this chapter we describe the selection of the decay mode $B_s^0 \rightarrow J/\psi \phi$, as well as of the similar higher statistics modes $B_u^\pm \rightarrow J/\psi K^\pm$ and $B_d^0 \rightarrow J/\psi K^{*0}$. The selection procedure should neither select candidates of different proper decay times with varying efficiency, nor should it prefer statistical upward fluctuations in the signal-to-background ratio. For this reason it is important to ensure that the selection procedure is done in an unbiased way.

We first describe the basic track and muon selection cuts, which guarantee a sample of reasonably well and controlled measured candidates. For each track the 4-momentum p for the given particle hypothesis is inferred from the measured 3-momentum \mathbf{p} by assuming its mass to be its current world average m [12]:

$$p = (\sqrt{m^2 + \mathbf{p}^2}, \mathbf{p}) \quad (4.1)$$

In search for B candidates we perform nested loops (3-fold for B_u , 4-fold for B_d and B_s) over all tracks, apply the appropriate charge requirements and check for double-counting of tracks. Invariant masses of intermediate resonances or the B meson are calculated from the sum of all track 4-momenta. A vertex fit is applied and required to converge for each intermediate resonance stage and for the B meson candidate.

The actual physics cuts to extract the signal from background are described in later sections. They are kept as simple as possible and are transverse momentum cuts on the B meson and non- J/ψ component, as well as a cut on the vertex fit quality.

4.1 Common Selection Cuts

Track Selection

To ensure minimal track quality without big efficiency loss we only consider tracks which have at least 20 axial, 16 stereo COT hits and a transverse momentum $p_T > 0.4 \text{ GeV}/c$. For a lifetime measurement precise vertex reconstruction is crucial. Thus we require all tracks to have at least 3 SVX $r\phi$ hits. After this first pre-selection we reapply the silicon fit algorithm including the tracking corrections mentioned in the previous chapter. In the refitting process we do not consider hits in the ISL and

Layer 00, because corresponding sub-detectors are not yet understood to a degree needed for this analysis. Also dropped are SVX 90° and small-angle-stereo hits. Keeping these hits does not directly contribute to the quality of the lifetime analysis, which is performed in the $r\phi$ plane only.

Muon Selection

For this analysis muon candidates only with stubs in the CMU or CMX are used. Muons involving CMU stubs are required to have a matching $\chi^2 < 9$ between the extrapolated track and the stub. No matching requirements are made for muon candidates with CMX stubs. The p_T of the track associated with the muon must be larger than 1.5 GeV/ c . This cut re-implements the CMU1.5 trigger cut for the dimuon trigger.

J/ ψ Selection

The J/ ψ selection is based on the muon selection just described. A J/ ψ candidate is made up of two oppositely charged muons. The invariant mass of the two muons is calculated by fitting their tracks to a common vertex as described in appendix C. The vertex fit must converge for the candidate to be considered further. No χ^2 cut is applied. Only candidates in a mass window of ± 80 MeV/ c^2 around the nominal J/ ψ mass of 3.09687 GeV/ c^2 are kept.

Figure 4.1 shows the dimuon mass spectrum around the J/ ψ mass. A single Gaussian with a first order polynomial is used as a crude estimate of yield and mass resolution, although the fit is obviously not satisfactory. At masses below the peak we identify a radiative tail which shifts the fit mass to too small values. In addition we see that one Gaussian is insufficient in describing the average mass resolution.

To first order we get 1,040,000 J/ ψ candidates, with an average mass resolution of $\sigma_m = 15.5$ MeV/ c^2 . The signal-to-background-ratio is 5.4 in $\pm 3\sigma_m$ around the world average J/ ψ mass [12], which is indicated by a vertical line at mass 3.09687 GeV/ c^2 .

Transverse Decay Length L_{xy}

In an ideal situation the decay length in the xy plane would be the distance in the laboratory frame between the primary interaction point and the decay point of the B meson. The existence of measurement uncertainties and mis-reconstruction introduces unphysical situations where the decay vertex vector and the reconstructed B meson momentum are not parallel and sometimes even of opposite direction. Introducing a signed transverse decay length¹ L_{xy} as defined in figure 4.2 – the projection of the decay vertex vector onto the B meson direction – incorporates these unphysical cases in a natural way.

Figure 4.3 shows the measured decay length L_{xy} of the J/ ψ candidates. We expect the peak at nought to come from prompt background and prompt J/ ψ . In order to get a rough estimate on the L_{xy} resolution, a Gaussian fit in the range of ± 100 μm

¹We will refer to it often as just “decay length”

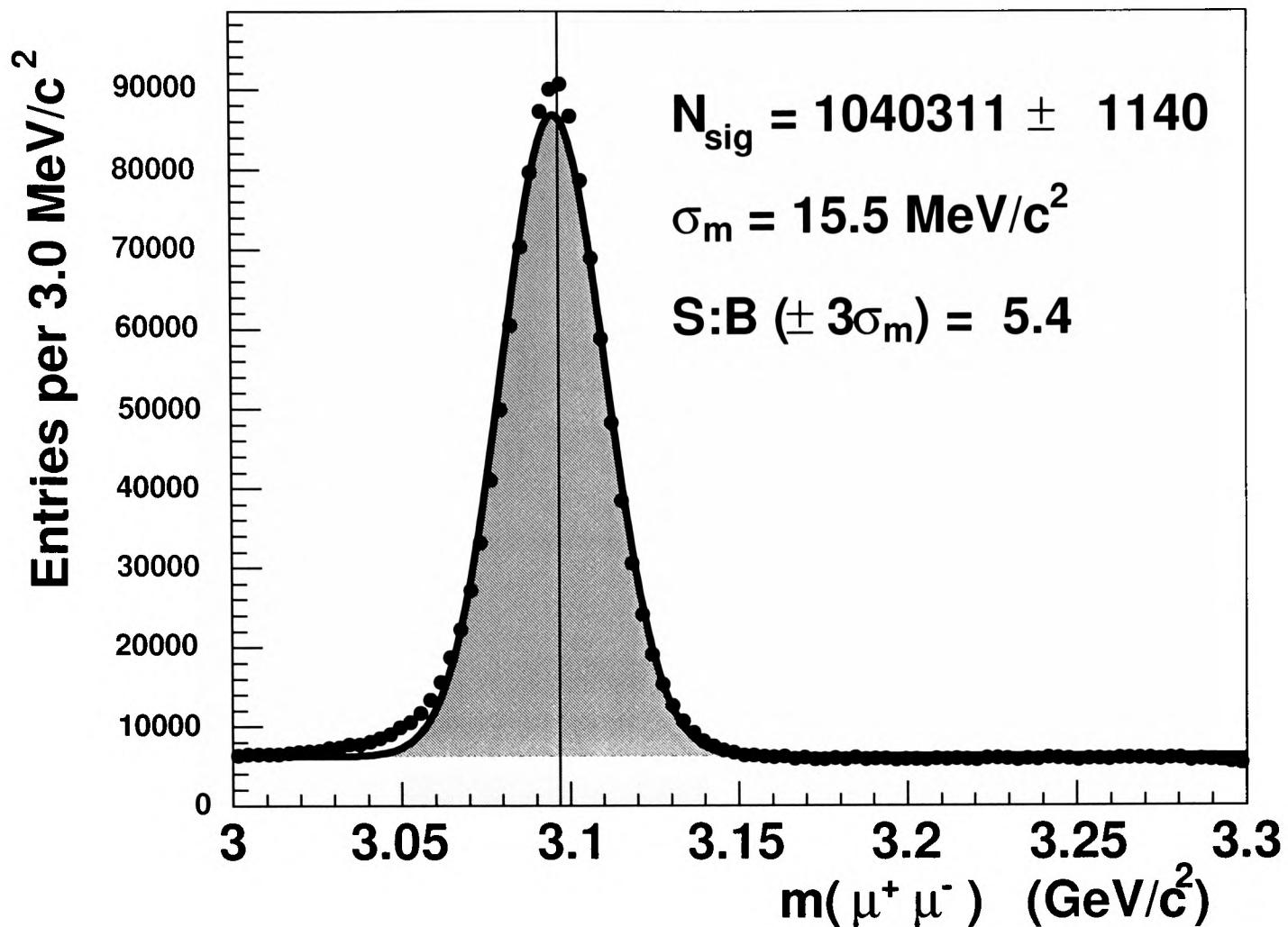


Figure 4.1: J/ψ mass spectrum fitted with a single Gaussian and first order polynomial.

around nought is overlaid, resulting in an average resolution of $58 \mu\text{m}$ (beam width included). As every event contributes with an individual resolution rather than a fixed resolution we do not expect our simple model to describe the data perfectly. We also see negative and positive tails as expected from mis-reconstruction and a long positive tail as expected from secondary J/ψ coming from B decays.

For a lifetime measurement we would like the decay length of B candidates to be reasonably well measured. Therefore we require a maximum value of the error $\sigma_{L_{xy}}$ on the decay length. For the 4-prong vertices ($B_d^0 \rightarrow J/\psi K^{*0}$ and $B_s^0 \rightarrow J/\psi \phi$) we require $\sigma_{L_{xy}} < 100 \mu\text{m}$. The 3-prong $B_u^\pm \rightarrow J/\psi K^\pm$ vertex is on average of lower quality than the 4-prong vertices. It turns out that we reject quite some signal with a cut value of $100 \mu\text{m}$. For that reason we loosen the cut to $150 \mu\text{m}$ in order to be more efficient.

Transverse Momentum Cuts

The most simple and effective quality cuts to suppress background are cuts on the total B meson transverse momentum and the non- J/ψ transverse momentum. The

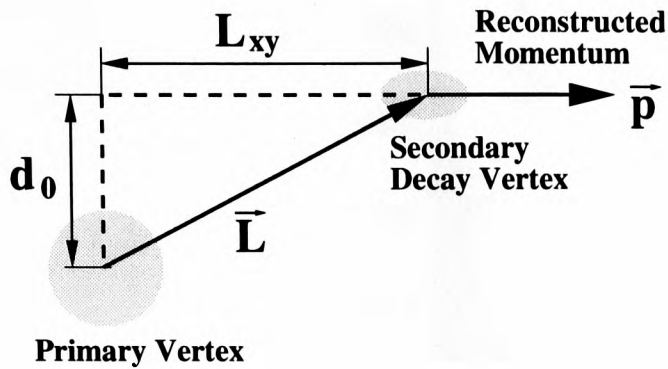


Figure 4.2: Definition of the transverse decay length L_{xy} and impact parameter d_0 of a reconstructed B candidate.

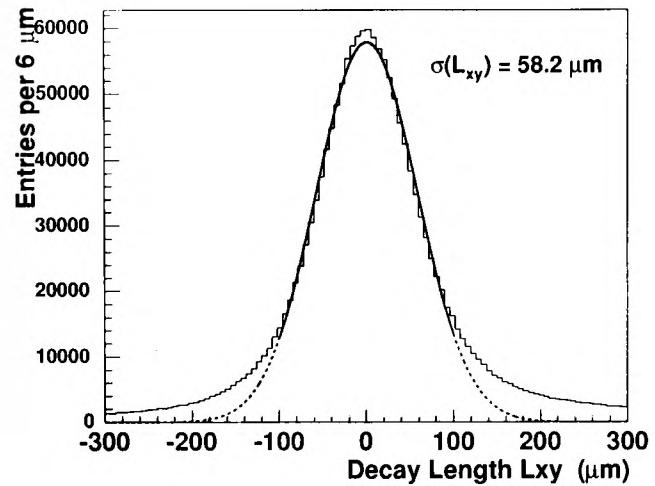


Figure 4.3: L_{xy} distribution of J/ψ candidates. A Gaussian fit in the range $-100 \dots 100 \mu\text{m}$ is overlaid. The dotted line shows the extrapolation of the fit.

cut values are chosen in an unbiased way to minimise the statistical error on the background subtracted signal. For our lifetime measurement we consider long-lived candidates as signal, i.e. candidates with a proper decay length ct larger than our resolution (about $60 \mu\text{m}$). In appendix B we describe in more detail the optimisation procedure and the choice of the actual cut values: $p_T(\text{non-}J/\psi) > 2 \text{ GeV}/c$ and $p_T(B) > 6.5 \text{ GeV}/c$. We use the same cuts for all three decay modes, because kinematics and backgrounds are similar and a comparison can be done more easily when kinematic cuts are kept the same.

Vertex Fit Quality

The three decay modes $B_u^\pm \rightarrow J/\psi K^\pm$, $B_d^0 \rightarrow J/\psi K^{*0}$ and $B_s^0 \rightarrow J/\psi \phi$ all have in common a one-vertex decay topology with 3, 4, and 4 tracks, respectively. We fit the candidate tracks to a common vertex with an algorithm described in Appendix C. Because the natural width ($87 \text{ keV}/c^2$) of the J/ψ is much smaller than our J/ψ mass resolution ($15 \text{ MeV}/c^2$), we improve our B meson mass resolution by constraining the mass of the muon pair to $3.09687 \text{ GeV}/c^2$. We do not require the resulting B meson momentum to point back to the primary vertex, because it may bias the decay time distribution.

Since silicon z hits are not considered in this analysis, we vertex effectively only in the $r\phi$ plane. Therefore the B vertex quality is controlled by a cut not on the total χ^2 but only on the $\chi_{r\phi}^2$ contribution returned by the vertex fit. The corresponding number of degrees of freedom for an effective vertex fit in the $r\phi$ plane is $n_{trks} - 2 + 1$.

Figure 4.4 shows the $\chi_{r\phi}^2$ distribution for $B_u^\pm \rightarrow J/\psi K^\pm$ for sideband subtracted data and simulation. See appendix B for details on the simulation and background subtraction. We notice an excess of simulation at low $\chi_{r\phi}^2$ values, but overall rather

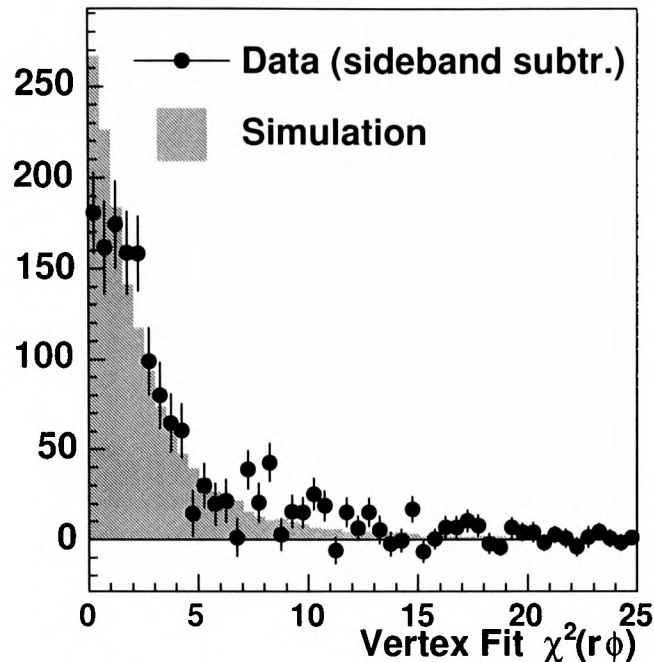


Figure 4.4: Distribution of $\chi^2_{r\phi}$ for $B_u^\pm \rightarrow J/\psi K^\pm$ data (sideband subtracted) and simulation.

good agreement. To see how this distribution corresponds to a χ^2 distribution with four degrees of freedom, we translate the χ^2 into upper tail probability²(or **vertex fit χ^2 probability**), which we expect to be flat.

Figure 4.5 shows the vertex fit probability in a linear and logarithmic scale. The distribution (left) with a linear x scale is rather flat as it should be, though a slight negative slope is visible, which could indicate that the track errors are overestimated. Also a big peak is visible at χ^2 probabilities very close to nought. This feature is not understood. Simulation follows the same trends as data, only the peak at nought is lower. It is interesting to see that there is quite a significant fraction of signal at very low χ^2 probabilities.

The distribution on the right shows the logarithm of the vertex fit χ^2 probability. The first bin artificially acts as underflow bin, i.e. contains the entries less than minus four as well. We see that there is a significant fraction of the signal at extremely low values of the vertex fit χ^2 probability.

A cut on the $\chi^2_{r\phi}$ (or vertex fit χ^2 probability) should be subject to a signal optimisation procedure based on a formal optimisation of $\mathcal{S}^2/(\mathcal{S}+\mathcal{B})$ using simulation. Such an optimisation leads to extremely small χ^2 probabilities (below 10^{-5}) as cut values [93]. While it may be true that one achieves a net gain in effective statistics of the signal by cutting so loosely, we are uncomfortable about relying upon an L_{xy} distribution derived from such unreasonably improbable vertex fits. Most earlier CDF analyses [43], [78], [51], [79] use either 0.1% or 1% cuts, which seem to be conservatively sensible choices. We choose the lower value of 0.1% that allows us

² $= \int_{\chi^2}^{\infty} dt f(t)$ with $f(t)$ being the χ^2 distribution for n_{dof} degrees of freedom.

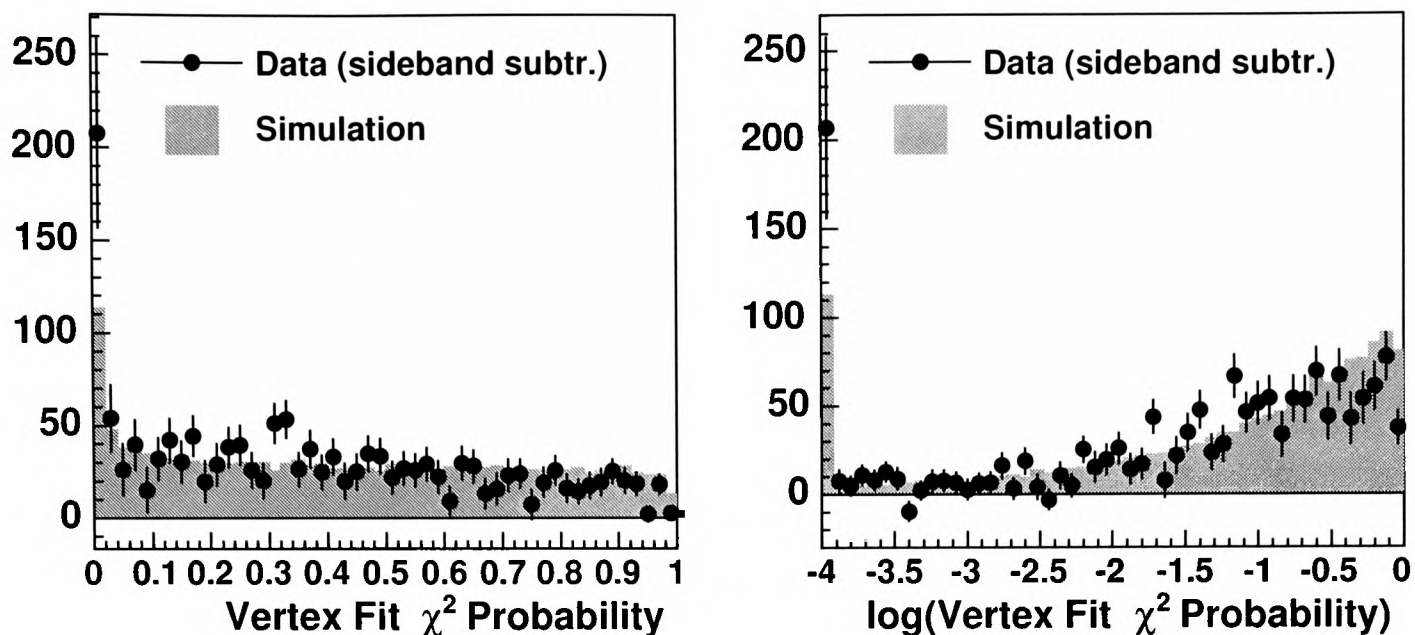


Figure 4.5: Distributions of the vertex fit χ^2 probability in a linear and logarithmic scale for $B_u^\pm \rightarrow J/\psi K^\pm$ data (sideband subtracted) and simulation.

Decay	n_{dof}	Maximum $\chi_{r\phi}^2$
$B_u^\pm \rightarrow J/\psi K^\pm$	2	13.82
$B_d^0 \rightarrow J/\psi K^{*0}$	3	16.27
$B_s^0 \rightarrow J/\psi \phi$	3	16.27

Table 4.1: Cut values on $\chi_{r\phi}^2$ corresponding to a vertex fit χ^2 probability of 0.1 % for different decay topologies.

to be more efficient with the knowledge that the vertex fit χ^2 scale cannot be taken literally in the first place. It is important to notice that the choice of the cut is unbiased. An additional discussion of the vertex quality cut is done in the section on systematic uncertainties. Table 4.1 shows for each decay mode the number of degrees of freedom along with the $\chi_{r\phi}^2$ cut values that correspond to a vertex fit χ^2 probability cut of 0.1%.

Multiple B Candidates

In the case of multiple B candidates in the same event all of them are kept so that the true one is not discarded. Keeping multiple candidates should not bias the lifetime result, because the background model of the lifetime fit will parameterise these candidates in positive exponential tails (see 5). The size of this contribution is found to be 2.2 %, 20.3 %, and 1.3 %, respectively for $B_u^\pm \rightarrow J/\psi K^\pm$, $B_d^0 \rightarrow J/\psi K^{*0}$, and $B_s^0 \rightarrow J/\psi \phi$. It is a small effect for $B_u^\pm \rightarrow J/\psi K^\pm$ and $B_s^0 \rightarrow J/\psi \phi$. The big value for $B_d^0 \rightarrow J/\psi K^{*0}$ comes from the ambiguity of particle mass assignment to the decay $K^{*0} \rightarrow K^\pm \pi^\mp$. In case two $B_d^0 \rightarrow J/\psi K^{*0}$ candidates have the opposite $K\pi$ mass

Description	Requirement
Number of COT axial hits	≥ 20
Number of COT stereo hits	≥ 16
Number of SI $r\phi$ hits	≥ 3
Muon momentum	$p_T(\mu^\pm) > 1.5 \text{ GeV}/c$
Track-to-stub matching for muons with CMU stub	$\chi^2 < 9.0$
Two oppositely charged muons	
J/ψ mass window	$(3096.87 \pm 80) \text{ MeV}/c^2$
Non- J/ψ momentum	$p_T(\text{non-}J/\psi) > 2.0 \text{ GeV}/c$
B momentum	$p_T(B) > 6.5 \text{ GeV}/c$
B vertex quality	$\text{Prob}(\chi_{r\phi}^2) > 0.1\%$
Beamline	Measured with SVX tracks

Table 4.2: Summary of quality cuts applied at different stages of the analyses. These cuts are common to all three decay modes considered.

assignment, the candidate whose $K\pi$ mass is closer to the world average K^{*0} mass is chosen, the other one discarded. This measure reduces the number of duplicates to 3.0%. Effects of this procedure on the lifetime measurement will be studied in more detail later.

Beam Position Measurement

We only take candidates from runs for which the SVX beam position has been measured. SVX beamlines that are known to be incorrect are not used. The beam position in the xy plane is calculated as the point on the beamline with the z coordinate given by the mean z of the B vertex. This is a good enough approximation since beam slopes are very small (typically of the order of a few tenths mrad) and the contribution to the uncertainty of L_{xy} is negligible compared to the transverse size of the beam (typically $\sigma_x = \sigma_y = 30 \text{ }\mu\text{m}$ at small z and $\sigma_x = \sigma_y = 35 \text{ }\mu\text{m}$ at $z = 100 \text{ cm}$). In addition no lifetime bias is introduced by this procedure due to (approximate) cylindrical symmetry and measuring in the $r\phi$ plane only. If the beam spot size is not available,³ a size of $30 \text{ }\mu\text{m}$ in x and y is assumed.

Summary

Table 4.2 summarises all common cuts to the three decay modes.

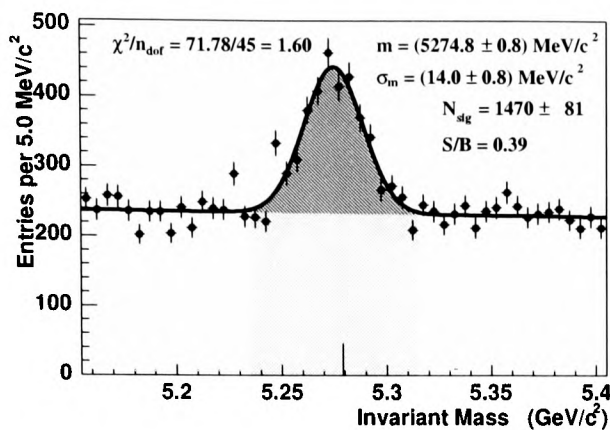


Figure 4.6: Invariant mass spectrum of the $B_u^\pm \rightarrow J/\psi K^\pm$ candidates, fitted with a single Gaussian plus first order polynomial.

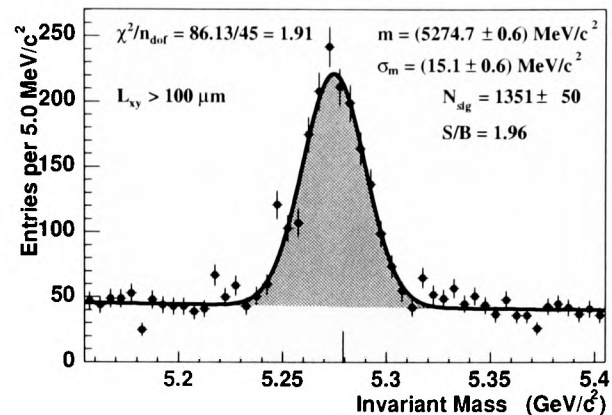


Figure 4.7: Same for candidates with $L_{xy} > 100 \mu\text{m}$.

4.2 Selection of $B_u^\pm \rightarrow J/\psi K^\pm$

Figure 4.6 shows the invariant mass spectrum of candidates surviving the cuts mentioned above. Overlaid is a simple fit of a Gaussian plus a linear function to get an estimate of the peak position, experimental width and signal yield. We see about 1,470 B_u candidates above background. The χ^2 per degree of freedom is 1.60. This low fit quality comes from deviations from mainly two bins. Studying these two bins in more details has not revealed anything peculiar. Neither do the events have an excess of a certain run number nor direction. Statistical fluctuations of this order are very unlikely. Also there is no reason to believe that the background description is inadequate and we do not see any broad bumps that would arise from reflections.

In $\pm 40 \text{ MeV}/c^2$ around the peak position we have a signal-to-background ratio of 0.39. The regions in which the signal-to-background ratio is estimated is indicated dark grey (green) for the signal, and light grey (yellow) for the background. The width of the peak is about $14 \text{ MeV}/c^2$ and is purely due to experimental resolution. The short line on the x -axis indicates the current world average value of the B_u mass of $(5.2790 \pm 0.0005) \text{ GeV}/c^2$ [12].

Figure 4.7 shows the same with an additional cut of $L_{xy} > 100 \mu\text{m}$. This cut improves the signal-to-background ratio to about 2 without losing much signal. It demonstrates that most of the B_u candidates are long-lived.

We have chosen the mass window to be from $5,155 \text{ MeV}/c^2$ to $5,405 \text{ MeV}/c^2$. The lower bound is to remove partially reconstructed B decays like $B_u^\pm \rightarrow J/\psi K^{*\mp}$ ($K^{*\mp} \rightarrow K^\mp \pi^0$) and $B_d^0 \rightarrow J/\psi K^{*0}$ ($K^{*0} \rightarrow K^\pm \pi^\mp$). Those decays have branching ratios larger or similar to $B_u^\pm \rightarrow J/\psi K^\pm$, and once the pion is missed they distort our mass background shape and will also contribute to the long-lived background component in the proper decay time distribution. Since they enter our mass spectrum below $m(B^{+0}) - m_\pi$, a cut of $m > 5.155 \text{ GeV}/c^2$ effectively removes them.

³Due to technical problems filling the database.

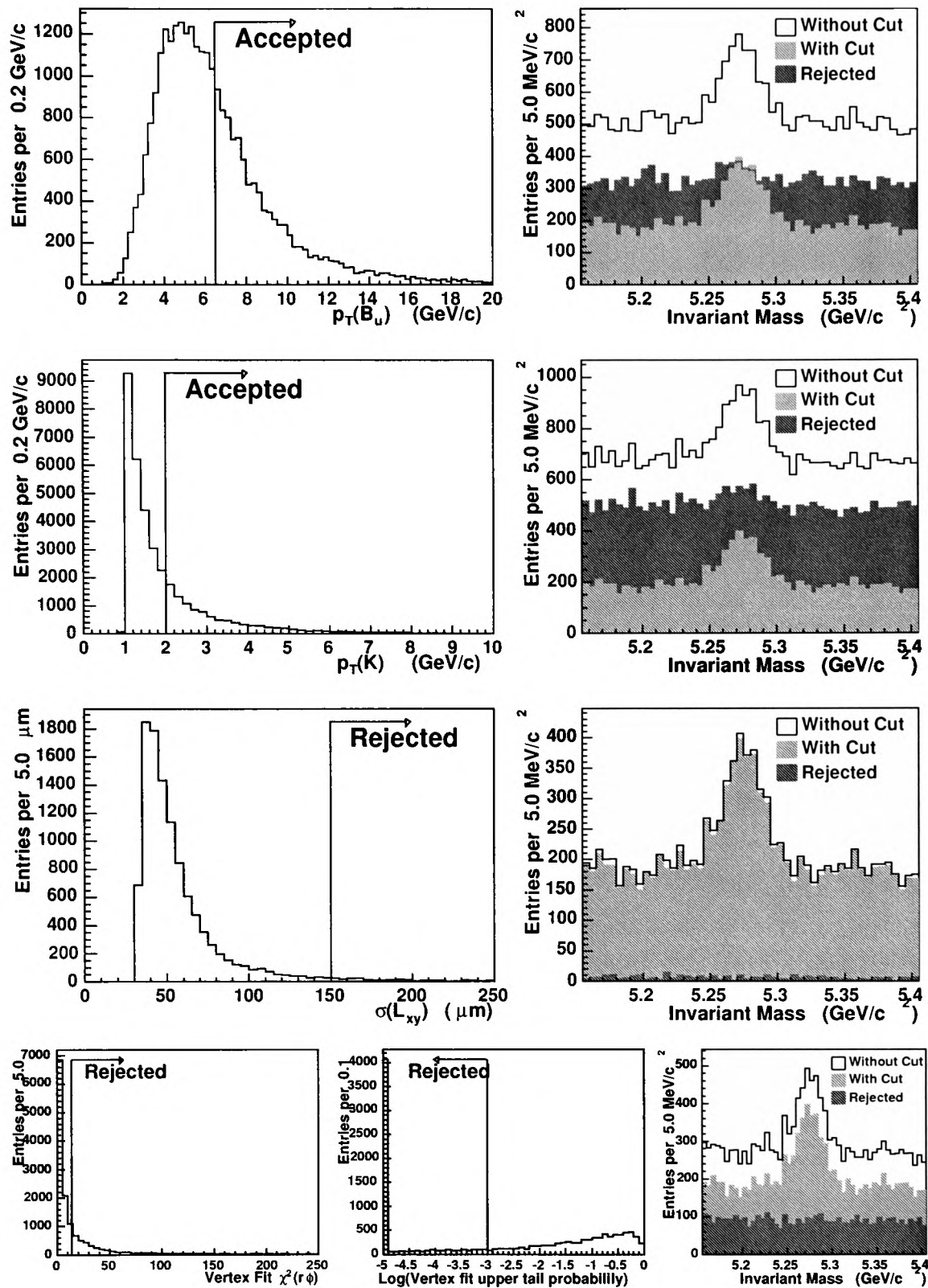


Figure 4.8: Each line shows the impact of each cut for $B_u^\pm \rightarrow J/\psi K^\pm$, when all the other cuts have already been applied. The left plot shows the distribution of the cut quantity, while the right picture contains invariant mass histograms of candidates without cut, with cut and those being rejected.

The impact of each cut can be studied by applying all other cuts first and then look at the mass distribution before and after the cut, as well as the mass distribution of rejected candidates. Figure 4.8 shows for each selection cut the distribution of the cut quantity on the left, and the B_d mass distribution without the cut, with the cut, and of the rejected candidates. We see that the transverse momentum and vertex quality cuts are very powerful. The cut on the decay length error does not have much impact. Also the mass distribution of rejected candidates is flat for the $\sigma(L_{xy})$ and vertex quality cuts, which indicates that very little signal is rejected by these cuts. For the transverse momentum cuts we see an indication that we reject some signal, but the large gain in purity justifies the inefficiencies.

4.3 Selection of $B_d^0 \rightarrow J/\psi K^{*0}$

We select K^{*0} mesons in their decay to $K^\pm \pi^\mp$. All oppositely charged track pairs passing our track quality requirements are fitted to a common vertex under the $K\pi$ particle hypothesis. We do not use particle identification, as dE/dx reconstruction is not yet in a state where we could use it, and the track momenta are too high for the particle identification capabilities of the time-of-flight system (see chapter 2.5).

We require the vertex fit to converge and the fitted mass to be in a window of $\pm 80 \text{ MeV}/c^2$ around the true mass of $896.10 \text{ MeV}/c^2$. Each track pair (t_1, t_2) is likely to pass these requirements for both mass assignments, (K, π) and (π, K) . In case two B_d candidates have opposite $K\pi$ mass assignment, we chose the one whose K^{*0} mass is closer to the current world average [12]. We discard the other candidate. Such a choice is known to be wrong in several percent of the cases, with some effect on the mass spectrum shape, but no significant effect on the lifetime. We shall discuss this issue in chapter 6 in more detail. The mass resolution of the K^{*0} will be smaller than the natural width of the K^{*0} , which is $(50.7 \pm 0.6) \text{ MeV}/c^2$ [12]. Therefore we do not constrain the $K\pi$ mass to the world average K^{*0} mass in our final vertex fit.

Figure 4.9 shows the invariant mass spectrum of all K^{*0} candidates and of K^{*0} candidates from B_d decays. The upper histogram shows all candidates with $p_T(K^{*0}) > 2 \text{ GeV}/c$ and converging vertex fit. No signal can be seen due to huge combinatorial background. The lower histogram on the other hand shows the K^{*0} candidates only for B_d candidates with the additional requirements $p_T(B^0) > 6.5 \text{ GeV}/c$ and $5.155 \text{ GeV}/c^2 < m(B_d) < 5.405 \text{ GeV}/c^2$. Figure 4.10 shows the same histogram again. We see a broad peak at $0.897 \text{ GeV}/c^2$ on a lot of background. We fit the signal with a Breit-Wigner,⁴ whose width we fix to the world average value of $50.7 \text{ MeV}/c^2$. For the background we use the following function:

$$f(m) = p_1 \cdot (m - m_{min})^{p_2} \cdot \exp(-p_3(m - m_{min})) , \quad (4.2)$$

with the free parameters m_{min} , p_1 , p_2 , and p_3 . The indicated mass window of

⁴We have tried to convolve it with a Gaussian of resolution σ to account for the experimental mass resolution. But since the mass resolution is expected to be on the order of just a few MeV/c^2 , it is of no significance compared to the width of $50.7 \text{ MeV}/c^2$.

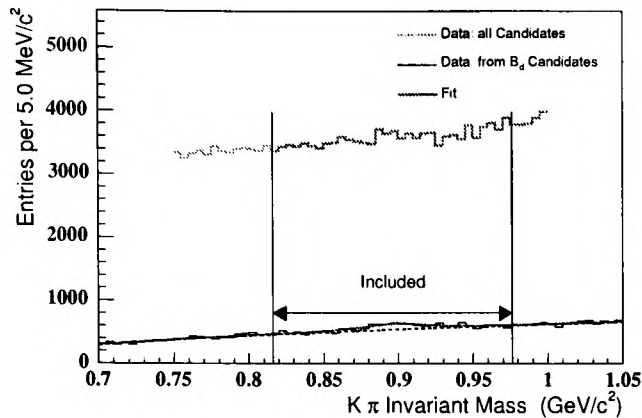


Figure 4.9: Invariant mass spectrum of all K^{*0} candidates (upper histogram) and K^{*0} candidates from B_d decays (lower histogram). A fit for background (dotted line) and background + Breit-Wigner (solid line) is overlaid.

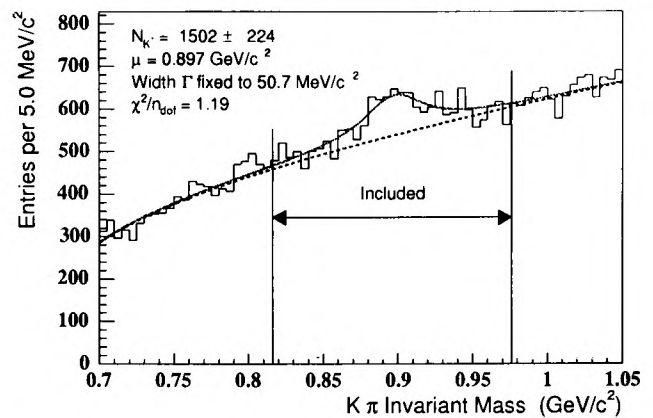


Figure 4.10: Invariant mass spectrum of K^{*0} candidates from B_d decays. A fit for background (dotted line) and background + Breit-Wigner (solid line) is overlaid.

$\pm 80 \text{ MeV}/c^2$ around the world average K^{*0} mass seems to be a reasonable (unbiased) choice.

In addition we require transverse momentum cuts on the K^{*0} and B^0 : $p_T(K^{*0}) > 2.0 \text{ GeV}/c$ and $p_T(B^0) > 6.5 \text{ GeV}/c$. These cuts are chosen to be the same as for $B_u^\pm \rightarrow J/\psi K^\pm$ to facilitate comparisons. The mass window was chosen under similar considerations as made in the $B_u^\pm \rightarrow J/\psi K^\pm$ section, with an additional reason being uniformity between the two decay modes.

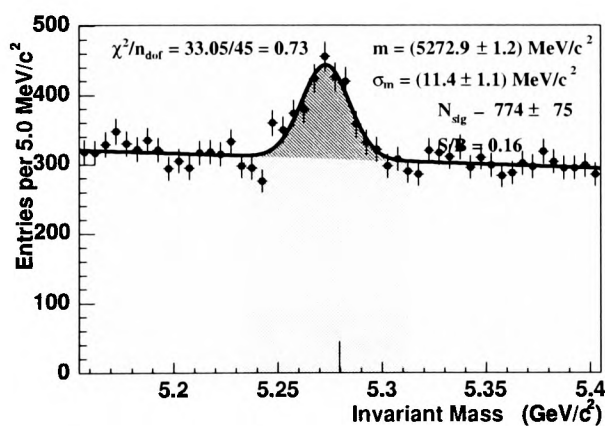


Figure 4.11: Invariant mass spectrum of the $B_d^0 \rightarrow J/\psi K^{*0}$ candidates, fitted with a single Gaussian plus first order polynomial.

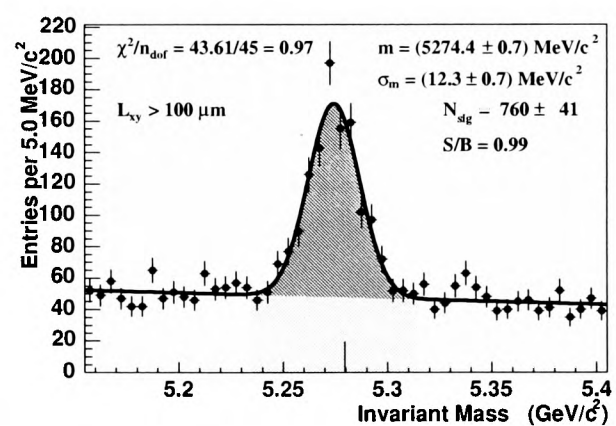


Figure 4.12: Same for candidates with $L_{xy} > 100 \mu\text{m}$.

Figure 4.11 shows the invariant mass spectrum of candidates surviving the cuts mentioned above. Overlaid is a simple fit of a Gaussian plus a linear function to get an estimate on the peak position, experimental width and signal yield. We see about 774

B_d candidates above background. In $\pm 40 \text{ MeV}/c^2$ around the peak position we have a signal-to-background ratio of 0.16. The regions in which the signal-to-background ratio is estimated is indicated dark grey (green) for the signal, and light grey (yellow) for the background. The width of the peak is about $11.4 \text{ MeV}/c^2$ and is purely due to experimental resolution. The short line on the x -axis indicates the current world average value of the B_d mass of $(5.2794 \pm 0.0005) \text{ GeV}/c^2$ [12].

Figure 4.12 shows the same with an additional cut of $L_{xy} > 100 \mu\text{m}$. This cut improves the signal-to-background ratio to about 0.99 without losing much signal, and demonstrates that most B_d signal candidates are long-lived.

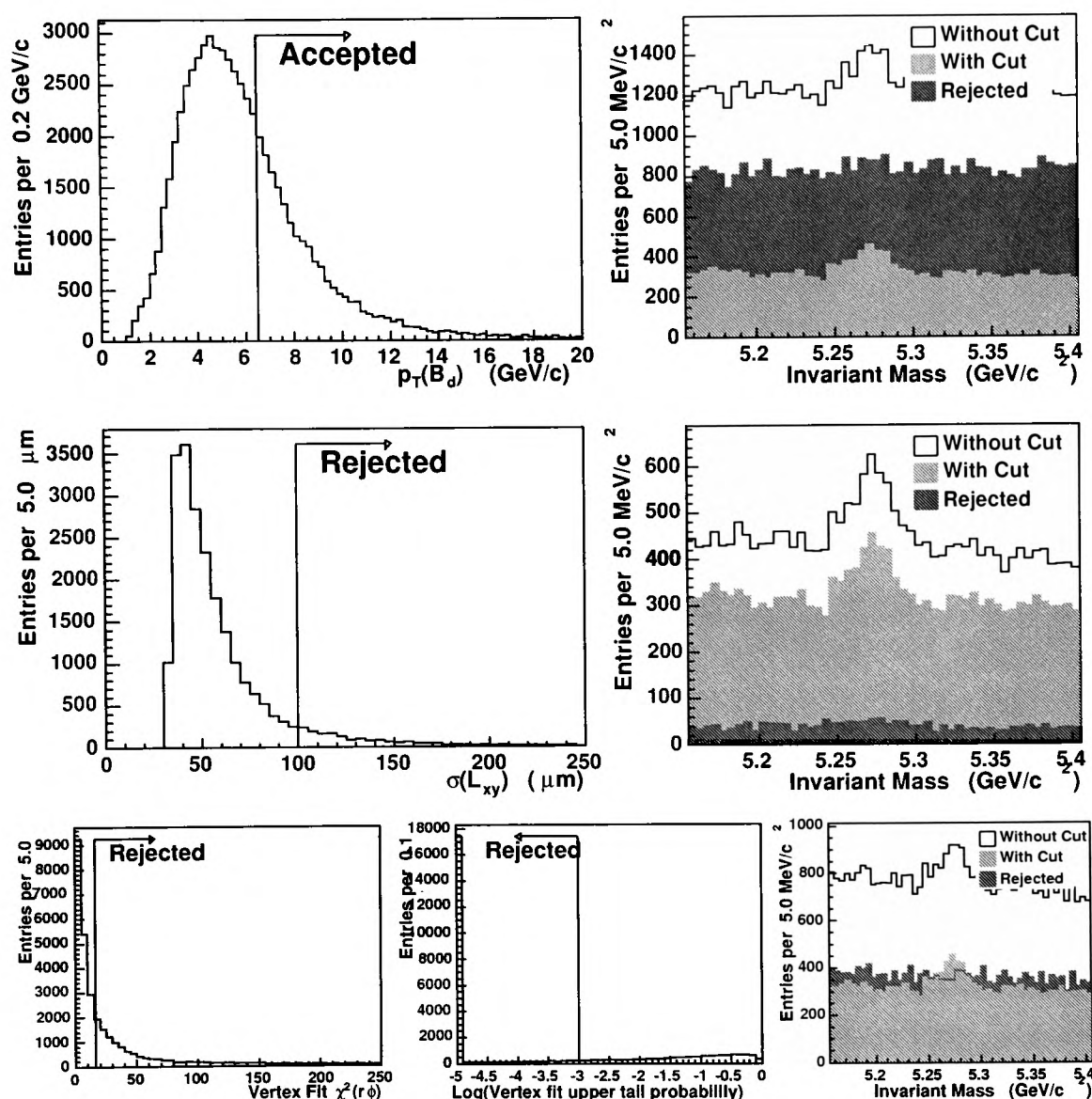


Figure 4.13: Each line shows the impact of each cut for $B_d^0 \rightarrow J/\psi K^{*0}$, when all the other cuts have already been applied. The left plot shows the distribution of the cut quantity, while the right picture contains invariant mass histograms of candidates without cut, with cut and those being rejected.

Figure 4.13 shows again on the left for each selection cut the distribution of the cut quantity, and on the right the B_d mass distribution without the cut, with the cut, and of the rejected candidates. The figure for the $p_T(K^{*0})$ cut is not shown. Due to the large number of combinatorial background the $p_T(K^{*0}) > 2 \text{ GeV}/c$ cut was

implemented at a very early stage of the analysis procedure, not allowing a histogram for the $p_T(K^{*0})$ cut to be made within an appropriate time. The same is true with the cut on the K^{*0} mass window. Again we demonstrate the power and efficiency of the cuts.

4.4 Selection of $B_s^0 \rightarrow J/\psi \phi$

The intermediate ϕ resonance is reconstructed in its decay K^+K^- . All oppositely charged track pairs passing our track quality requirements are fitted to a common vertex under kaon hypothesis. We require the vertex fit to converge and the fitted mass being in a window of $\pm 10 \text{ MeV}/c^2$ around the world average mass of $1.019456 \text{ GeV}/c^2$ [12]. The mass resolution of our ϕ reconstruction is on the same order of magnitude as the natural width of the ϕ ($(4.26 \pm 0.05) \text{ MeV}/c^2$ [12]). Therefore we do not constrain the KK mass to the world average ϕ mass in our final vertex fit.

In addition we require transverse momentum cuts on the ϕ and B_s : $p_T(\phi) > 2.0 \text{ GeV}/c$ and $p_T(B_s) > 6.5 \text{ GeV}/c$, much the same way we do for the other two decay modes. We only consider B_s candidates that fall in the mass window set from $5220 \text{ MeV}/c^2$ to $5520 \text{ MeV}/c^2$. This window is slightly wider than the one we use for the lighter B mesons to allow better determination of the background shape.

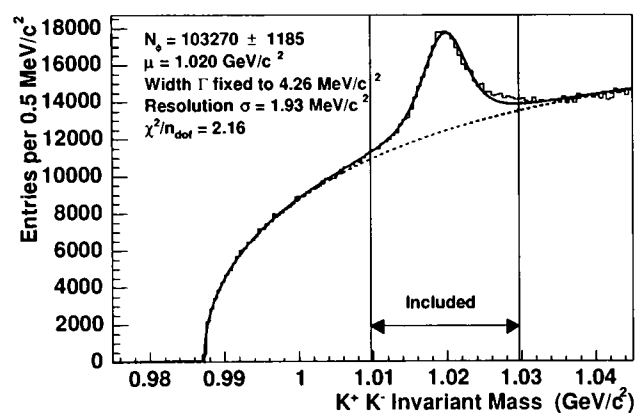


Figure 4.14: Invariant mass spectrum of all ϕ candidates. A fit for background (dotted line) and background + Breit-Wigner (convolved with Gaussian resolution) (solid line) is overlaid.

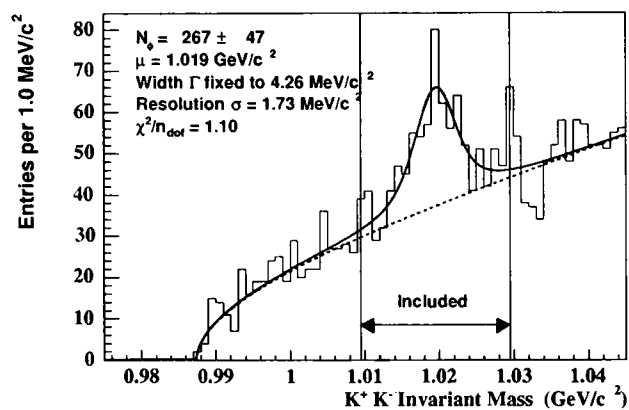


Figure 4.15: Invariant mass spectrum of ϕ candidates from B_s decays. A fit for background (dotted line) and background + Breit-Wigner (convolved with Gaussian resolution) (solid line) is overlaid.

Figure 4.14 shows the invariant mass spectrum of all ϕ candidates with $p_T(\phi) > 2 \text{ GeV}/c$ and a converging vertex fit. We fit the signal with a Breit-Wigner convolved with a Gaussian of resolution σ . The width of the Breit-Wigner is fixed to the world average value of $4.26 \text{ MeV}/c^2$. As background function we use again equation 4.2. About 100,000 candidates can be seen above a huge background. The fit quality is not perfect because the error on the mass is different for each event. Nevertheless our

simple fit yields an average mass resolution of about $2 \text{ MeV}/c^2$. The indicated mass window of $\pm 10 \text{ MeV}/c^2$ around the current world average ϕ mass seems reasonable. Figure 4.15 shows the ϕ candidates for B_s candidates satisfying all cuts except the ϕ mass window cut.

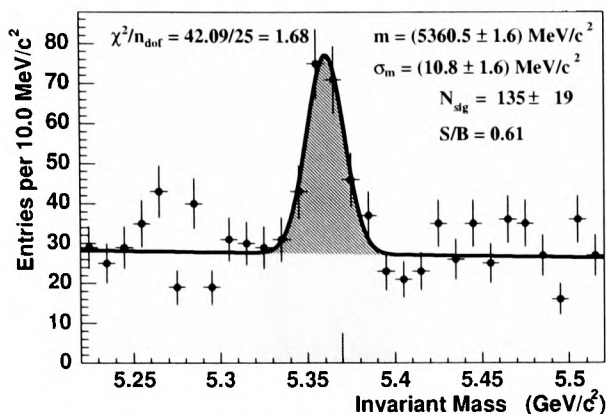


Figure 4.16: Invariant mass spectrum of the $B_s^0 \rightarrow J/\psi \phi$ candidates, fitted with a single Gaussian plus first order polynomial.

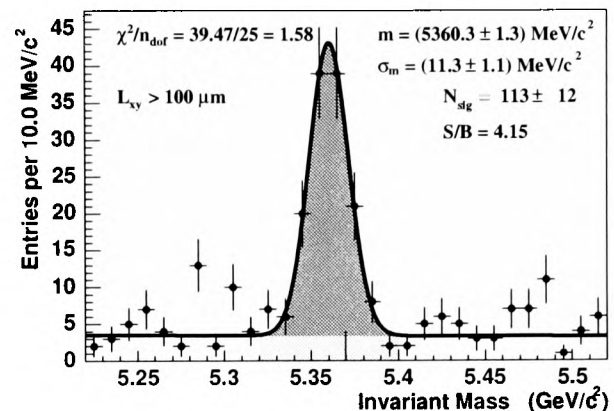


Figure 4.17: Same for candidates with $L_{xy} > 100 \mu\text{m}$.

Figure 4.16 shows the invariant mass spectrum of candidates surviving the cuts mentioned above. Overlaid is a simple fit of a Gaussian plus a linear function to get an estimate on the peak position, experimental width and signal yield.

We see about 135 B_s candidates above background. In $\pm 40 \text{ MeV}/c^2$ around the peak position we have a signal-to-background ratio of 0.61. The regions in which the signal-to-background ratio is estimated is indicated dark grey (green) for the signal, and light grey (yellow) for the background. The width of the peak is about $10.8 \text{ MeV}/c^2$ and is purely due to experimental resolution. The short line on the x -axis indicates the current world average value of the B_s mass of $(5.3696 \pm 0.0024) \text{ GeV}/c^2$ [12].

Figure 4.17 shows the same with an additional cut of $L_{xy} > 100 \mu\text{m}$. This cut improves the signal-to-background ratio to about 4 without losing much signal, demonstrating once that most of the B_s candidates are long-lived.

As with the other decay modes we study the impact of each cut by applying all other cuts first and then look at the mass distribution before and after the cut, and also at the mass distribution of rejected candidates. Figure 4.18 shows for each selection cut the distribution of the cut quantity on the left, and the B_s mass distribution without the cut, with the cut, and of the rejected candidates. We see that the transverse momentum, ϕ mass window and vertex quality cuts are very powerful. The cut on the decay length error does not have much impact. Also the mass distribution of rejected candidates is flat, which indicates that very little signal is rejected by these cuts.

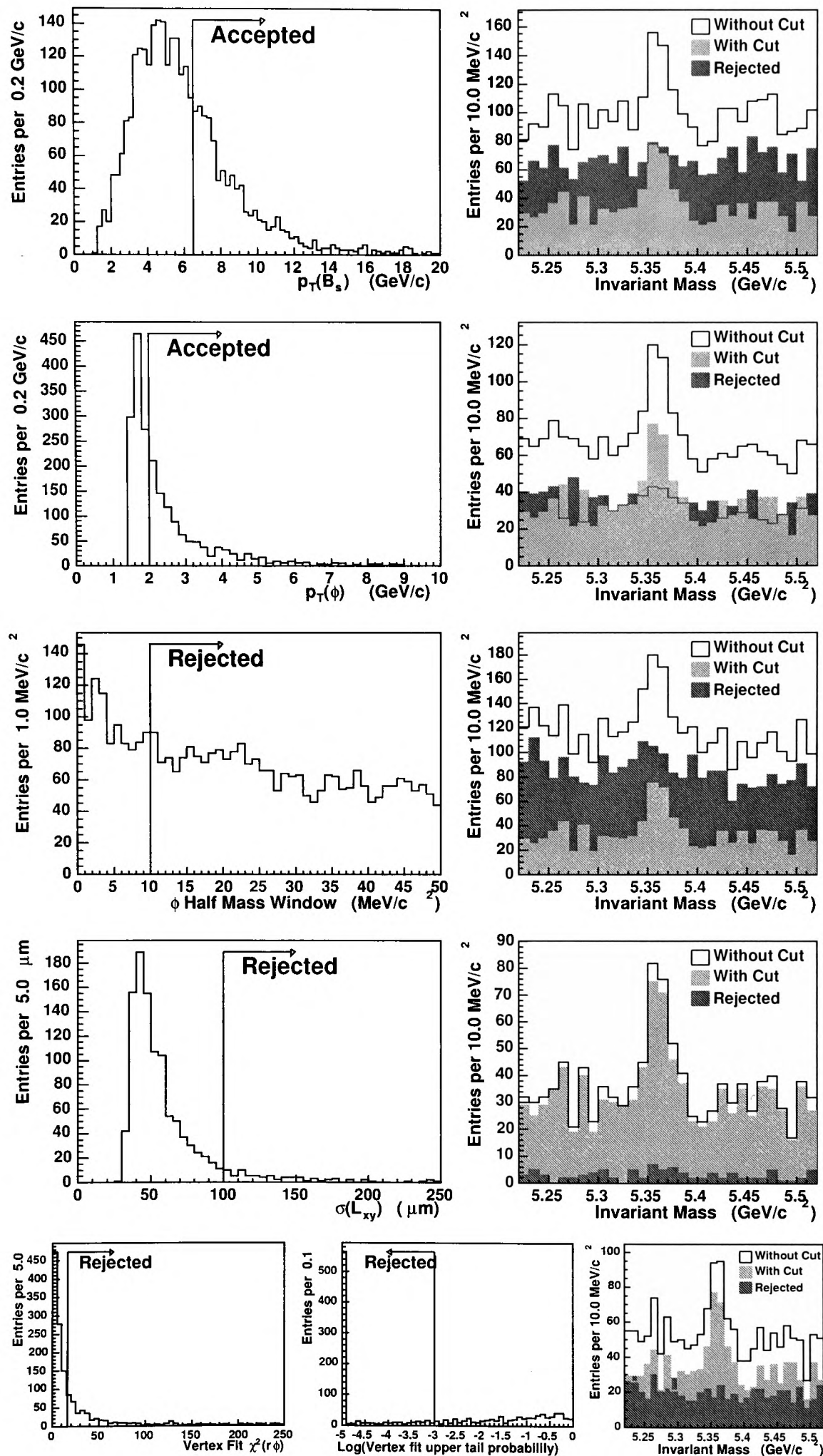


Figure 4.18: Each line shows the impact of each cut for $B_s^0 \rightarrow J/\psi \phi$, when all the other cuts have already been applied. The left plot shows the distribution of the cut quantity, while the right picture contains invariant mass histograms of candidates without cut, with cut and those being rejected.

4.5 Proper Decay Time Calculation

A lifetime measurement requires the estimation of the proper decay time from measurable quantities for each B meson candidate. The decay time in the laboratory frame can be obtained from the measured decay length by dividing by the velocity βc of the B meson. Transforming into the B rest frame with a factor $1/\gamma$ yields the proper decay time. The unknown factor $\beta\gamma c$ can easily be expressed in terms of the measurable quantities 3-momentum \mathbf{p} and invariant mass m by the relation $\beta\gamma c = |\mathbf{p}|/m$. This is still true if we only use transverse components L_{xy} , β_T , γ_T and p_T . Therefore we have

$$\begin{aligned} t &= \frac{L}{\gamma\beta c} = L \frac{m}{|\mathbf{p}|} \\ &= \frac{L_{xy}}{\gamma_T\beta_T c} = L_{xy} \frac{m}{p_T} \end{aligned} \quad (4.3)$$

and the proper decay time is fully expressed through the measurable quantities L_{xy} , m and p_T . We choose to use the current world average value for the B meson mass m , rather than the reconstructed value.

Ignoring possible correlations we can write the error σ_t as

$$\sigma_t^2 = \left| \frac{m}{p_T} \sigma_{L_{xy}} \right|^2 + \left| \frac{t}{p_T} \sigma_{p_T} \right|^2 + \left| \frac{L_{xy}}{p_T} \sigma_m \right|^2. \quad (4.4)$$

Note that the contribution from the momentum measurement scales with t , whereas the vertexing contribution only adds an uncertainty constant with t . Figure 4.19 shows the distribution of the three contributions (unsquared) for the $B_u^\pm \rightarrow J/\psi K^\pm$ candidates. The mass contribution is of order 10^{-4} ps and occupies only the very first bin. The p_T contribution just overlaps with the dominating L_{xy} contribution at values of about 0.03 ps. The $B_d^0 \rightarrow J/\psi K^{*0}$ and $B_s^0 \rightarrow J/\psi \phi$ case are very similar, though the centre of the L_{xy} distribution varies slightly.

Figures 4.20, 4.21 and 4.22 show distributions of transverse decay length L_{xy} (left), proper decay time t (middle) and proper decay time error σ_t (right) for the $B_u^\pm \rightarrow J/\psi K^\pm$, $B_d^0 \rightarrow J/\psi K^{*0}$ and $B_s^0 \rightarrow J/\psi \phi$ candidates, respectively. In order to get an estimate on the average resolution, we fit the L_{xy} and t distributions to a single Gaussian centred at nought, in the range indicated by the solid line. The resulting resolutions are listed in table 4.3. The resolutions are statistically consistent to be the same for all three decay modes. Comparing the t resolutions with the peak of the appropriate σ_t distribution reveals the fact that the calculated σ_t is on average underestimated by a factor of about 1.3.

Apart from negative non-Gaussian tails in the t distributions we see large positive tails, which indicate the long-lived heavy flavour content. In the following chapter we will see how to infer the signal lifetime from the positive tail of the t distribution.

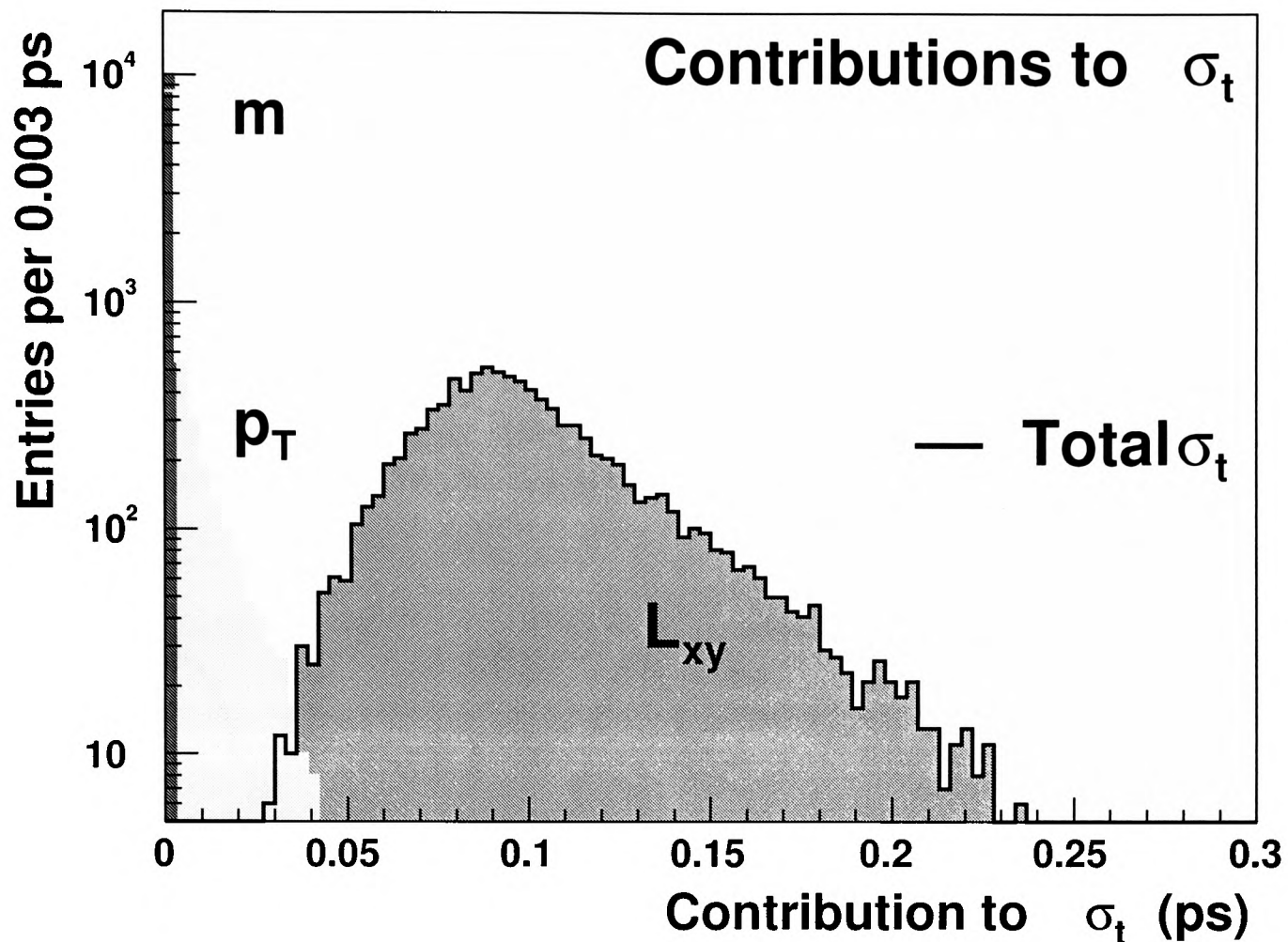


Figure 4.19: Contributions to the total proper decay time error from the uncertainty on decay length L_{xy} , momentum p_T and meson mass m . The mass contribution is of order 10^{-4} ps and occupies the very first bin, whereas the L_{xy} completely dominates. The total error is the sum of errors squared from all three contributions and is nearly identical to the L_{xy} contribution. This figure is made from the $B_u^\pm \rightarrow J/\psi K^\pm$ candidates, but represents qualitatively the $B_d^0 \rightarrow J/\psi K^{*0}$ and $B_s^0 \rightarrow J/\psi \phi$ case as well.

Decay Mode	L_{xy} Resolution	t Resolution
$B_u^\pm \rightarrow J/\psi K^\pm$	$(58.5 \pm 1.0) \mu\text{m}$	$(0.121 \pm 0.002) \text{ ps}$
$B_d^0 \rightarrow J/\psi K^{*0}$	$(60.2 \pm 0.8) \mu\text{m}$	$(0.122 \pm 0.002) \text{ ps}$
$B_s^0 \rightarrow J/\psi \phi$	$(64.3 \pm 4.2) \mu\text{m}$	$(0.137 \pm 0.010) \text{ ps}$

Table 4.3: Summary of average L_{xy} and t resolutions obtained from a single Gaussian fit to the data

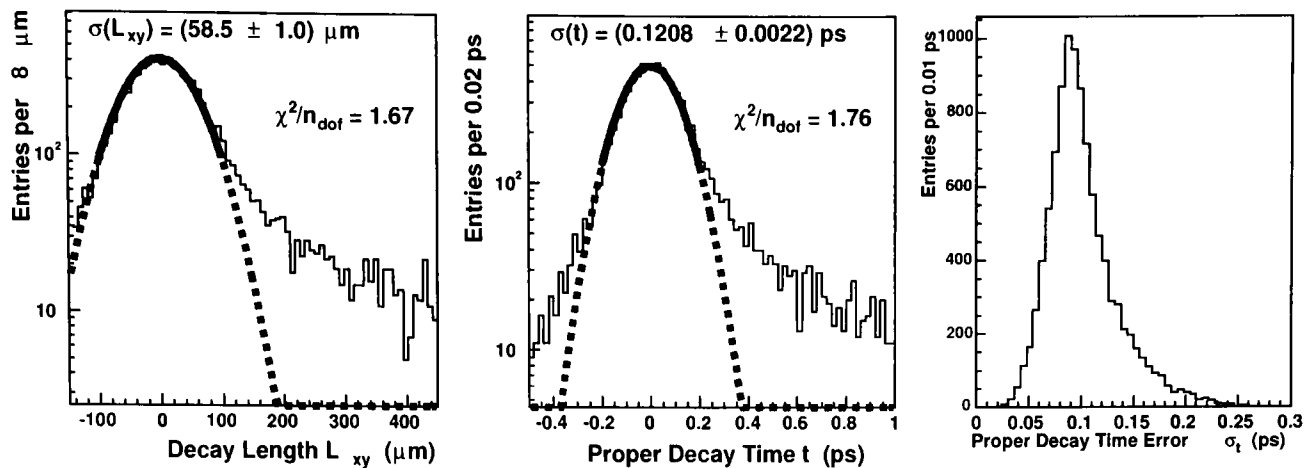


Figure 4.20: Distributions for $B_u^\pm \rightarrow J/\psi K^\pm$ of decay length L_{xy} (left), proper decay time t (middle) and proper decay time error σ_t (right). Overlaid in the left and middle picture is a Gaussian, with the fit range indicated by the solid line.

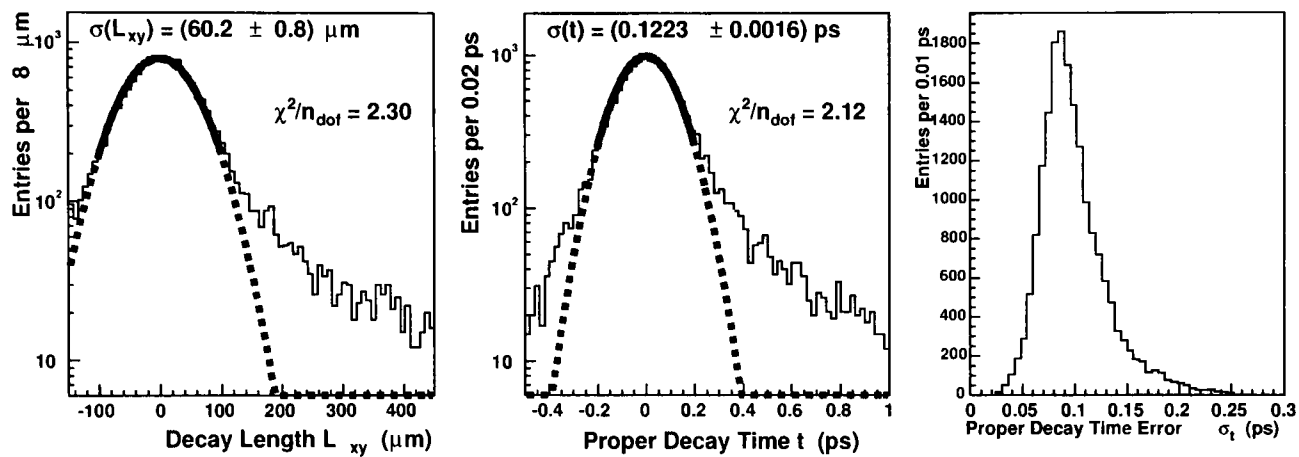


Figure 4.21: Distributions for $B_d^0 \rightarrow J/\psi K^{*0}$ of decay length L_{xy} (left), proper decay time t (middle) and proper decay time error σ_t (right). Overlaid in the left and middle picture is a Gaussian, with the fit range indicated by the solid line.

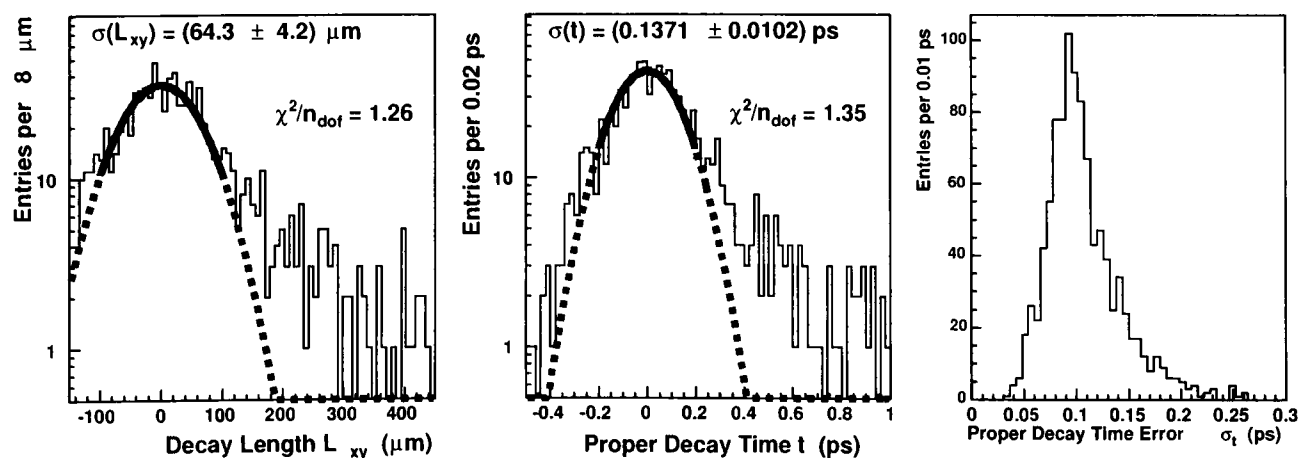


Figure 4.22: Distributions for $B_s^0 \rightarrow J/\psi \phi$ of decay length L_{xy} (left), proper decay time t (middle) and proper decay time error σ_t (right). Overlaid in the left and middle picture is a Gaussian, with the fit range indicated by the solid line.

Chapter 5

Lifetime Analysis

5.1 Fit Model

In this section we describe the fit model we use to extract B lifetimes. We fit simultaneously invariant mass and proper decay time for all candidates in a region about the B meson mass. The lifetime is extracted by minimising an unbinned negative log-likelihood function, which uses proper decay time t , its error σ_t and invariant mass information m from each B meson candidate. The minimisation is performed with the MIGRAD algorithm in the MINUIT package [80].

The likelihood function is given by

$$\mathcal{L} = \prod_{i=1}^N P(m_i, t_i, \sigma_{t,i}; \mathbf{A}), \quad (5.1)$$

where $P(m_i, t_i, \sigma_{t,i}; \mathbf{A})$ is the probability density function for each candidate, \mathbf{A} comprises all fit parameters, and i indicates an individual B meson candidate with measured m_i , t_i and $\sigma_{t,i}$.

The probability function P is split into signal part and background part:

$$\begin{aligned} P(m_i, t_i, \sigma_{t,i}; \mathbf{A}) &= f_{sig} \cdot g_{sig}(m_i) \cdot F_{sig}(t_i, \sigma_{t,i}) \\ &+ (1 - f_{sig}) \cdot g_{bk}(m_i) \cdot F_{bk}(t_i, \sigma_{t,i}), \end{aligned} \quad (5.2)$$

with f_{sig} being *signal fraction*, g_{sig} and F_{sig} modelling mass and proper decay time for the signal, and g_{bk} and F_{bk} for the background.

The proper decay time error $\sigma_{t,i}$ results from fitting the candidate's tracks to a common vertex. The error estimate is therefore only correct if the tracks come from a real physical vertex. Our philosophy is to use the errors in the contributions that are believed to come from a physical vertex (signal and prompt background), through convolving t_i with a resolution function $r(t_i - t', \sigma_{t,i}; S_t)$. The remaining contributions that are not believed to come from a physical vertex have a distorted error estimate and are parameterised without the error σ_t and a resolution function.

For simplicity we choose the proper decay time **resolution function** \mathbf{r} as being Gaussian, allowing an overall scale factor S_t for the Gaussian error $\sigma_{t,i}$, thus correcting

for under- or overestimated errors:

$$r(t_i - t', \sigma_{t,i}; S_t) \equiv \frac{1}{\sqrt{2\pi}S_t\sigma_{t,i}} \exp \left\{ -\frac{1}{2} \left(\frac{t_i - t'}{S_t\sigma_{t,i}} \right)^2 \right\}. \quad (5.3)$$

The background model should *effectively* account for background proper decay time and decouple background from signal. There are many different possible choices for a background model. A separate discussion about our particular choice of background model as well as cross-checks can be found in chapter 6. Our chosen background model distinguishes four different background contributions to the t distribution. A δ -function at $t = 0$ convolved with the resolution function describes prompt background. A negative exponential describes negative tails and two positive background components account for positive tails:

$$\begin{aligned} F_{bk}(t_i, \sigma_{t,i}; S_t, f_-, f_+, f_{++}, \tau_-, \tau_+, \tau_{++}) &= (1 - f_- - f_+) \cdot \delta(0) \otimes r(t_i, \sigma_{t,i}; S_t) \\ &+ f_- \cdot \begin{cases} \frac{1}{\tau_-} e^{t_i/\tau_-} & (t_i < 0) \\ 0 & (t_i \geq 0) \end{cases} \\ &+ f_+ \cdot \begin{cases} 0 & (t_i < 0) \\ \frac{1}{\tau_+} e^{t_i/\tau_+} & (t_i \geq 0) \end{cases} \\ &+ f_{++} \cdot \begin{cases} 0 & (t_i < 0) \\ \frac{1}{\tau_{++}} e^{-t_i/\tau_{++}} & (t_i \geq 0) \end{cases}. \end{aligned} \quad (5.4)$$

The negative tail has fraction f_- and lifetime τ_- , and the short(long)-lived positive tail has fraction f_+ (f_{++}) with lifetime τ_+ (τ_{++}). The negative and short-lived positive exponential backgrounds are due to prompt vertices mixed with mis-reconstructed tracks that appear to be displaced. In addition, the following cases contribute to the positive tails as well:

- Displaced J/ψ s (from other B decays) paired with random tracks.
- Sequential semileptonic decays $b \rightarrow c \rightarrow s$ faking J/ψ .
- Partially reconstructed B decays.
- Reflections ($B_s^0 \rightarrow J/\psi \phi$ reconstructed as $B_d^0 \rightarrow J/\psi K^{*0}$ or, with a missed K, as $B_u^\pm \rightarrow J/\psi K^\pm$).
- In case of $B_d^0 \rightarrow J/\psi K^{*0}$, candidates from $K\pi$ mis-assignment when reconstructing K^{*0} .

In principle each background contribution should be parameterised independently. The limited statistics and their common exponential behaviour allow us to combine and effectively describe them on average by two exponential functions. In the case of $B_s^0 \rightarrow J/\psi \phi$ we will see later that statistics is so low that even one positive exponential is enough.

As signal model for the proper decay time we obviously choose an exponential convolved with the resolution function:

$$F_{sig}(t_i, \sigma_{t,i}; \tau, S_t) = \left\{ \begin{array}{ll} 0 & (t_i < 0) \\ \frac{1}{\tau} e^{-t_i/\tau} & (t_i \geq 0) \end{array} \right\} \otimes r(t_i, \sigma_{t,i}; S_t) \quad (5.5)$$

$$= \int_0^\infty dt' \frac{e^{-t'/\tau}}{\tau} \cdot \frac{1}{\sqrt{2\pi} S_t \sigma_{t,i}} \exp\left(-\frac{(t_i - t')^2}{2 S_t^2 \sigma_{t,i}^2}\right) \quad (5.6)$$

$$= \frac{1}{2\tau} \exp\left(\frac{S_t^2 \sigma_{t,i}^2}{2\tau^2} - \frac{t_i}{\tau}\right) \cdot \left\{ 1 - \text{Erf}\left(\frac{S_t \sigma_{t,i}}{\sqrt{2}\tau} - \frac{t_i}{\sqrt{2} S_t \sigma_{t,i}}\right) \right\} \quad (5.7)$$

where the parameter τ is the lifetime of the B meson, and $\text{Erf}()$ denotes the error function.¹

For simplicity we choose a Gaussian with peak position M and width W for the signal mass function and a first order polynomial describes well the background mass distribution:

$$g_{sig}(m_i; M, W) = \frac{1}{\sqrt{2\pi}W} e^{-\frac{(m_i - M)^2}{2W^2}} \quad (5.8)$$

$$g_{bk}(m_i; A) = \left\{ \begin{array}{ll} 0 & (m_i < m_{min} \text{ or } m_i > m_{max}) \\ A \cdot m_i + b & (m_{min} < m_i < m_{max}) \end{array} \right\} \quad (5.9)$$

The offset b normalises g_{bk} to 1 in the chosen mass range $[m_{min}, m_{max}]$, which gives

$$b = \frac{1}{m_{max} - m_{min}} - \frac{A}{2}(m_{max} + m_{min}) . \quad (5.10)$$

The resulting 12 parameters are summarised in table 5.1. It would be helpful to test our fit model on a simulated sample, and to check the composition and origins of the various backgrounds. Unfortunately we yet lack a reliable simulation of an inclusive bottom and charm sample, as well as a correct description of the underlying event and hard scattering process. But as we will see in chapter 6, our background model gives a reasonable description of the backgrounds.

5.2 Fitting $B_u^\pm \rightarrow J/\psi K^\pm$

Figure 5.1 shows the result of the fit to the 10,516 $B_u^\pm \rightarrow J/\psi K^\pm$ candidates. The range of the proper decay time histogram was chosen quite tight for more visibility, but it prevents 5 events on the negative side and 12 events on the far positive side not to enter the histogram. Nevertheless those events are used in the fit of course. The invariant mass histogram on the other hand contains all B_u candidates. We obtain a lifetime

$$\tau(B_u) = (1.623 \pm 0.048(\text{stat.})) \text{ ps} . \quad \text{corresponding to} \quad (5.11)$$

$$c\tau(B_u) = (486.7 \pm 14.5(\text{stat.})) \mu\text{m} . \quad (5.12)$$

¹ $\text{Erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x dt \exp(-t^2)$

1	M	Mass	GeV/ c^2
2	W	Width	MeV/ c^2
3	A	Background slope	c^2/GeV
4	f_{sig}	Signal fraction	–
5	τ	Signal lifetime	ps
6	S_t	t error scale factor	–
7	f_-	Background fraction of negative tail	–
8	f_+	Background fraction of positive short-lived tail	–
9	f_{++}	Background fraction of positive long-lived tail	–
10	τ_-	Lifetime of negative background tail	ps
11	τ_+	Lifetime of positive short-lived background tail	ps
12	τ_{++}	Lifetime of positive long-lived background tail	ps

Table 5.1: Summary of the 12 parameters used in the combined mass and proper decay time fit.

Estimating the goodness of fit for an unbinned likelihood fit is still an area in statistics where more effort is needed, though some success has already been achieved [81]. We use two different, more intuitive approaches: On the one hand we calculate a χ^2 by binning the data. On the other hand we compare the value of the likelihood function with the distribution of likelihoods from many generated experiments.

In the first method we bin the data and calculate the χ^2 between data and fit projection:²

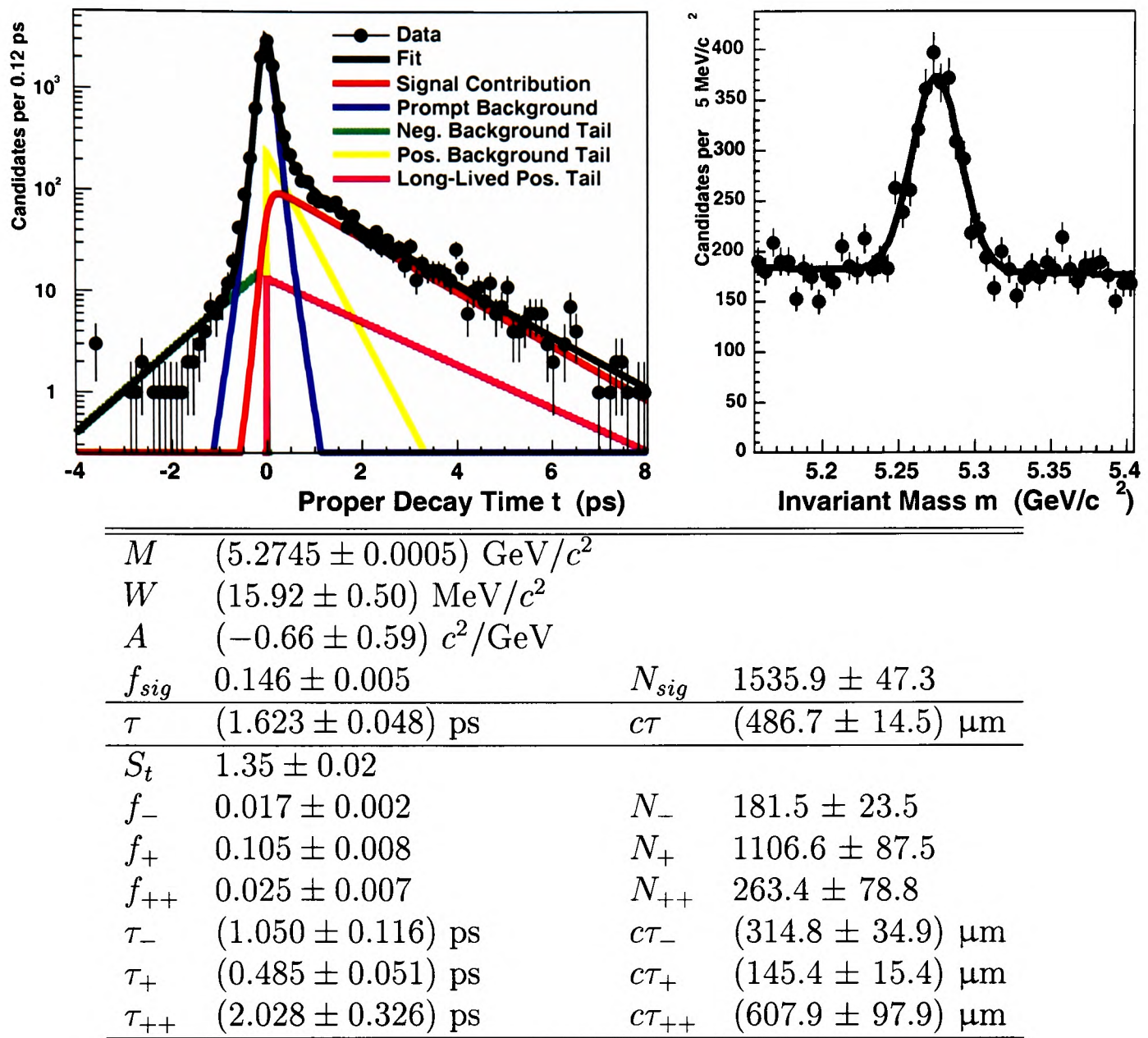
$$\chi^2 = \sum_{1 \leq i \leq N_{bins}} \frac{(y_i^{data} - y_i^{proj})^2}{\sigma_{y_i^{proj}}^2} = \sum_{1 \leq i \leq N_{bins}} \frac{(y_i^{data} - y_i^{proj})^2}{y_i^{proj}}, \quad (5.13)$$

where y_i^{data} is the number of entries in the i^{th} bin of the data histogram, and y_i^{proj} the integral of the projection over the range of bin i . It is important to have enough data entries in each bin to ensure Gaussian errors. This is the case in the chosen mass histogram in figure 5.1. The proper decay time histogram on the other hand has many low statistics bins. In that case we combine consecutive low statistics bins to a “combined bin”, until its number of entries becomes 20 or larger.

For the mass histogram we get a χ^2 value of 69.7 for 50 bins. The number of degrees of freedom n_{dof} is the number of bins minus the number of parameters. For the mass histogram 4 parameters are important: M , W , A and f_{sig} . So $\chi^2/n_{dof} = 69.7/46 = 1.51$. The probability of getting a bigger χ^2 is only 1.3 %, which indicates a rather poor fit.

For the proper decay time histogram we get a χ^2 of 67.7 for 43 bins. The parameters that define the proper decay time projection are f_{sig} , τ , S_t , f_- , f_+ , f_{++} , τ_- , τ_+ , τ_{++} . Therefore $n_{dof} = n_{bins} - 9$, where n_{bins} is the number of bins. So $\chi^2/n_{dof} = 67.7/34 = 1.99$, which translates into a probability of only 0.05 %.

²i.e. the curve resulting from the maximum likelihood fit as in figure 5.1.

Figure 5.1: Result of the combined mass and decay time fit for $B_u^\pm \rightarrow J/\psi K^\pm$.

The very low probabilities prompt an investigation into the underlying cause. The low fit probability in the mass distribution is not driven by any kind of problem in the maximum likelihood fit; the χ^2 fit of the mass histogram shown in figure 4.6 shows a similar χ^2/n_{dof} . Inspecting the mass histogram in figure 5.1 reveals that the poor fit quality comes only from a few bins. Changing the binning of the mass histogram can “smoothen” a few bins and increase the fit probability to up to about 5 %, but that does not change the fact that the fit is rather poor. It is also unlikely that our description of the mass distribution (single Gaussian for the signal and linear function for the background) is inadequate. All in all, we could not find anything wrong with our analysis and conclude that the occasional “outliers” seem to be a characteristic of the data themselves.

The poor quality in the proper decay time fit is due to at least two factors:

- In the region $t < -0.4 \text{ ps}$ the background is inadequately described by our background model. Not using this region does not change the resulting lifetime,

but enhances the fit quality to a fit probability of 8.4 %. We will discuss different background models and their impact on the lifetime more thoroughly in chapter 6.

- There is a clear excess of events in the interval $[3.80, 4.04]$ ps. The fit expects about 24 events whereas there are 39 in the data. The probability of one of 43 bins to fluctuate 2.95σ high is about 14 %. Therefore we could see a statistical fluctuation. All attempts of studying the candidates from this region have not revealed any peculiarity:
 - No excess of a certain run number or azimuthal direction.
 - The vast majority of them are from the signal region in the mass spectrum.
 - The B_d and B_s data does not show an excess in this interval.

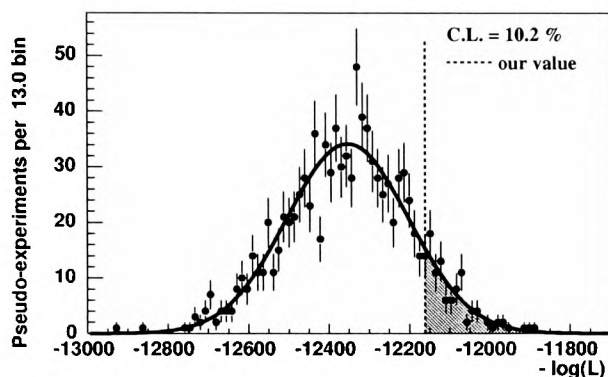


Figure 5.2: Distribution of the negative log-likelihood of many generated experiments with 10,516 $B_u^\pm \rightarrow J/\psi K^\pm$ candidates each. A Gaussian fit is overlaid. The dashed line represents the likelihood value of the data fit.

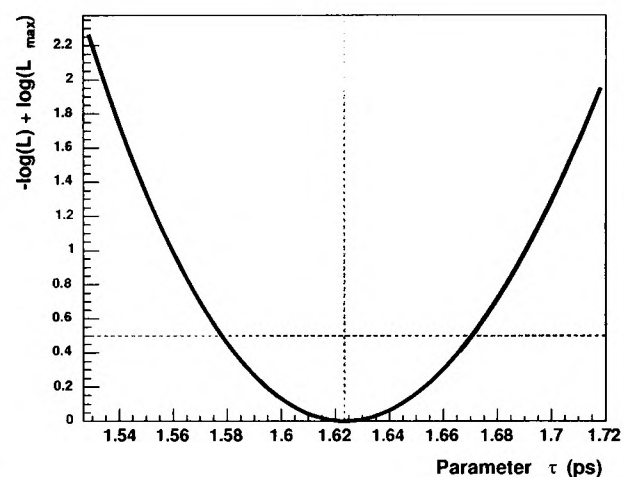


Figure 5.3: Likelihood $-\log(\mathcal{L}) + \log(\mathcal{L}_{max})$ as a function of lifetime fit parameter τ in the region about the minimum for $B_u^\pm \rightarrow J/\psi K^\pm$.

Another way of estimating the goodness of fit is to compare the likelihood of the actual experiment against the likelihood distribution of generated experiments. A confidence level can be obtained by computing the number of experiments with a higher value of $-\log(\mathcal{L})$ than the actual experiment.

Figure 5.2 shows the distribution of $-\log(\mathcal{L})$ for 2,000 generated experiments, that were simulated according to the parameters of the result of the data fit. The fraction of experiments with a higher $-\log(\mathcal{L})$ value than the data value is 10.2 % – an acceptable value.

In figure 5.3 we see $-\log(\mathcal{L}) + \log(\mathcal{L}_{max})$ as a function of the lifetime fit parameter τ about the fit minimum. It is approximately parabolic which justifies symmetric 1σ errors.

Table 5.2 lists the correlation coefficients for the fit. The correlation matrix ρ is defined as

$$\rho_{ij} \equiv \frac{\text{cov}(i, j)}{\sigma_i \sigma_j}, \quad (5.14)$$

where cov is the covariance matrix from the fit, and σ_i the error of parameter i . Coefficients related to the lifetime parameter τ are marked bold. Notice that τ has no bigger correlations other than with the signal fraction. Therefore our background model effectively separates signal and background.

	M	W	A	f_{sig}	τ	S_t	f_-	f_+	f_{++}	τ_-	τ_+
W	3										
A	-5	1									
f_{sig}	1	30	1								
τ	-0	-5	-1	-22							
S_t	-0	0	0	-0	0						
f_-	0	1	0	2	-1	-36					
f_+	-1	-2	-0	-10	3	-20	8				
f_{++}	0	-16	-0	-11	-5	-4	3	-58			
τ_-	-0	0	0	-0	0	29	-54	-8	-3		
τ_+	-1	-0	0	2	1	13	-8	20	-75	7	
τ_{++}	-1	8	0	6	-4	3	-2	54	-83	2	61

Table 5.2: Correlation coefficients for the $B_u^\pm \rightarrow J/\psi K^\pm$ fit in percent. The upper half and the diagonal are not shown for clarity.

The main correlations can be understood qualitatively:

- *Fraction-Lifetime-Anti-correlations:* The fit compensates a lowering of the fraction of a certain contribution by raising the appropriate lifetime.
- *Scale-Factor-Tails-Correlations:* If the scale-factor S_t increases, the prompt background includes more of the background tails. Therefore the tail fractions f_- , f_+ and f_{++} decrease, and the tail lifetimes increase.
- *Positive-Tails:* An increase of the long-lived positive tail fraction f_{++} decreases the short-lived tail fraction f_+ and steepens the short-lived tail's lifetime distributions, i.e. decreases τ_+ .

5.3 Fitting $B_d^0 \rightarrow J/\psi K^{*0}$

We repeat the fit for our 16,057 $B_d^0 \rightarrow J/\psi K^{*0}$ candidates. The result is displayed in figure 5.4. There are 6 events hiding in the underflow bin and 6 events in the overflow

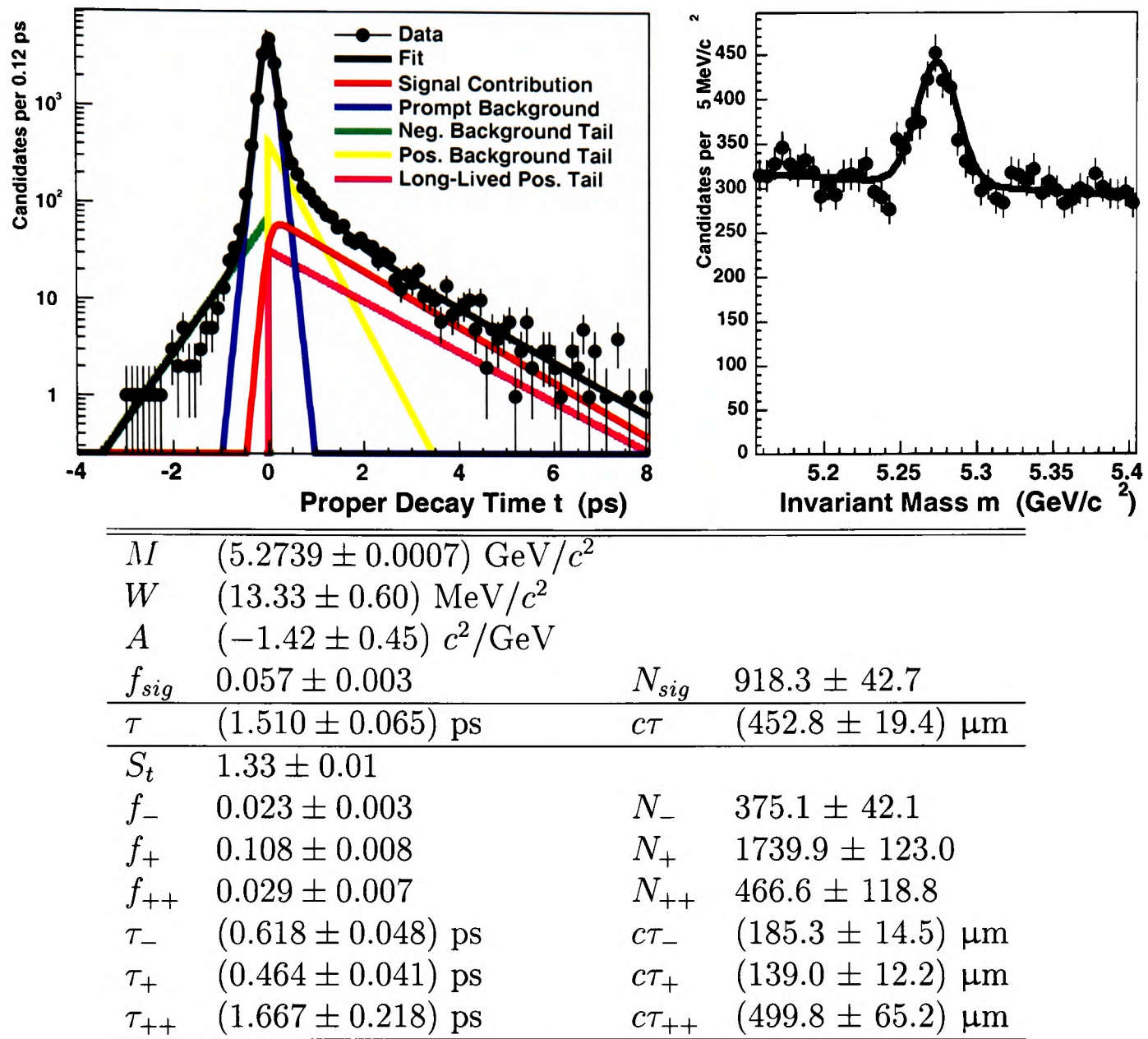


Figure 5.4: Result of the combined mass and decay time fit for $B_d^0 \rightarrow J/\psi K^{*0}$.

bin. Nonetheless they are included in the fit. As lifetime we get

$$\tau(B_d) = (1.510 \pm 0.065(\text{stat.})) \text{ ps} , \quad \text{corresponding to} \quad (5.15)$$

$$c\tau(B_d) = (452.8 \pm 19.4(\text{stat.})) \mu\text{m} . \quad (5.16)$$

Binning the data as described in the previous chapter gives $\chi^2/n_{dof} = 37.1/46 = 0.81$ for the mass distribution, which corresponds to a probability of 82 %. For the decay time distribution it is $\chi^2/n_{dof} = 47.4/32 = 1.48$ corresponding to 3.9 % probability. The poor fit quality of the proper decay time part is again due to an inadequate description of the negative tail in the region $t < -1$ ps. Not using this part of the data does not change the lifetime τ , but gives a fit probability of 20.1 %.

Figure 5.5 shows the distribution of $-\log(\mathcal{L})$ for 2,000 simulated experiments, that were generated using the result of the data fit. We get a value of 26.5 % for the fraction of experiments with a higher $-\log(\mathcal{L})$ value than the data value.

In figure 5.6 we display $-\log(\mathcal{L}) + \log(\mathcal{L}_{max})$ as a function of lifetime parameter

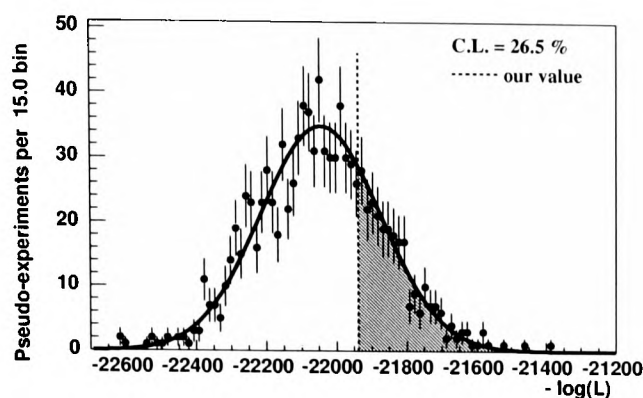


Figure 5.5: Distribution of the negative log-likelihood of many generated experiments with 16,057 $B_d^0 \rightarrow J/\psi K^{*0}$ candidates each. A Gaussian fit is overlaid. The dashed line represents the likelihood value of the data fit.

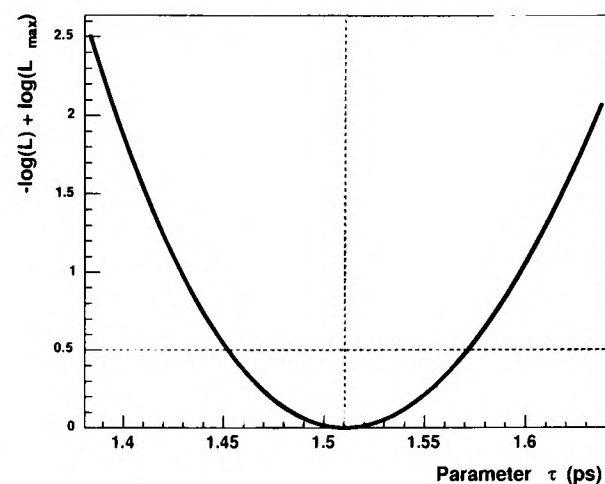


Figure 5.6: Likelihood $-\log(\mathcal{L}) + \log(\mathcal{L}_{max})$ as a function of lifetime fit parameter τ in the region about the minimum for $B_d^0 \rightarrow J/\psi K^{*0}$.

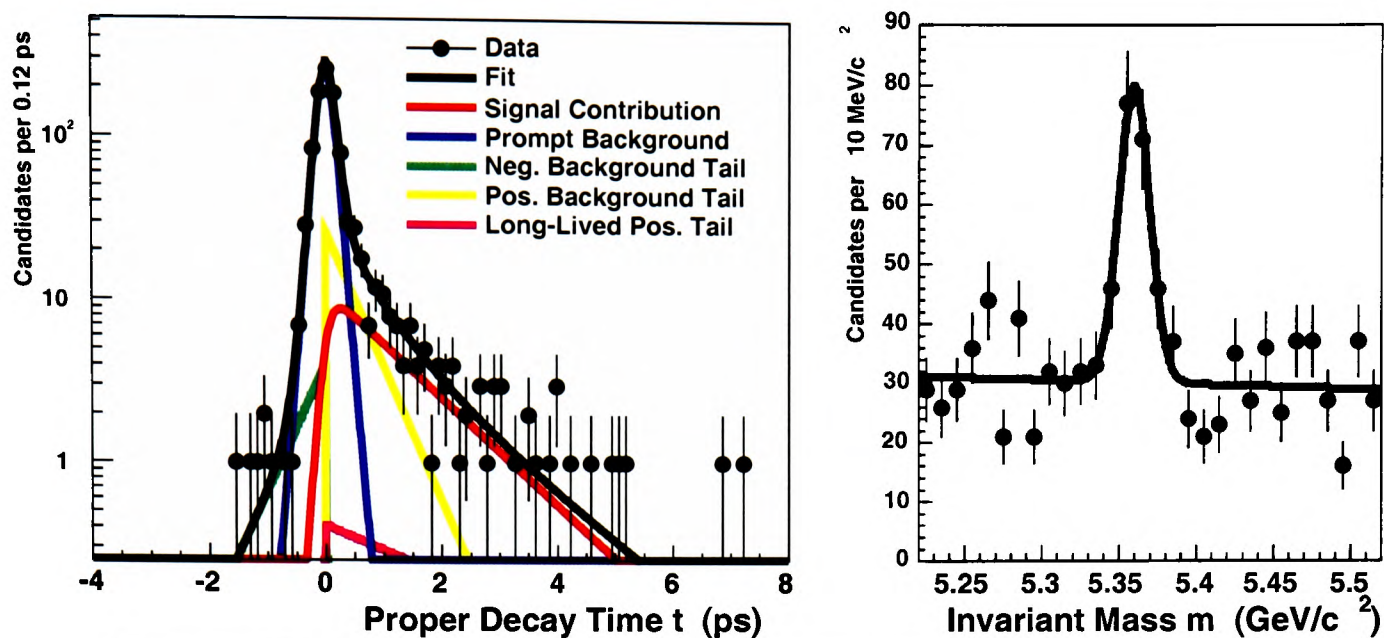
τ about the minimum. The parabolic shape justifies symmetric errors for the lifetime parameter τ . The correlation coefficients of our fit can be found in table 5.3.

	M	W	A	f_{sig}	τ	S_t	f_-	f_+	f_{++}	τ_-	τ_+
W	-4										
A	-4	1									
f_{sig}	-3	31	2								
τ	2	-3	-1	-28							
S_t	-0	-2	-0	-1	-0						
f_-	0	5	0	2	-1	-44					
f_+	-3	-0	-0	-9	7	-19	8				
f_{++}	3	-7	-0	-6	3	-6	8	-72			
τ_-	-0	-4	-0	-2	0	34	-67	-8	-6		
τ_+	-0	-6	-0	-4	-2	15	-15	41	-81	13	
τ_{++}	-3	-0	0	4	-13	4	-6	65	-86	5	69

Table 5.3: Correlation coefficients for the $B_d^0 \rightarrow J/\psi K^{*0}$ fit in percent. The upper half and the diagonal are not shown for clarity.

5.4 Fitting $B_s^0 \rightarrow J/\psi \phi$

The same procedure is now applied to our main decay mode. Figure 5.7 shows the result of the fit to the 1,023 $B_s^0 \rightarrow J/\psi \phi$ candidates. There is no underflow and only one event ($t = 9.9$ ps) overflow not shown in the histogram but used in the fit.



M	$(5.3597 \pm 0.0011) \text{ GeV}/c^2$		
W	$(10.06 \pm 0.96) \text{ MeV}/c^2$		
A	$(-0.89 \pm 1.28) c^2/\text{GeV}$		
f_{sig}	0.122 ± 0.013	N_{sig}	125.2 ± 13.0
τ	$(1.314 \pm 0.138) \text{ ps}$	$c\tau$	$(394.0 \pm 41.4) \mu\text{m}$
S_t	1.45 ± 0.05		
f_-	0.021 ± 0.011	N_-	21.4 ± 11.2
f_+	0.128 ± 0.022	N_+	130.8 ± 22.7
f_{++}	0.011 ± 0.013	N_{++}	11.1 ± 12.8
τ_-	$(0.550 \pm 0.217) \text{ ps}$	$c\tau_-$	$(165.0 \pm 65.1) \mu\text{m}$
τ_+	$(0.526 \pm 0.119) \text{ ps}$	$c\tau_+$	$(157.7 \pm 35.8) \mu\text{m}$
τ_{++}	$(2.850 \pm 1.822) \text{ ps}$	$c\tau_{++}$	$(854.3 \pm 546.1) \mu\text{m}$

Figure 5.7: Result of the combined mass and decay time fit for $B_s^0 \rightarrow J/\psi \phi$.

Comparing the fit result to binned data gives a good fit quality. For the proper decay time distribution we get $\chi^2 = 5.59/4 = 1.40$, corresponding to a probability of 23.2 %, and for the mass distribution we obtain $\chi^2 = 32.4/25 = 1.27$ and 14.8 % probability.

Figure 5.9 shows the distribution of $-\log(\mathcal{L})$ for 2,000 simulated experiments, that were generated using the result of the data fit. We obtain a confidence level of 58.7 %. Table 5.4 lists the correlation coefficient of our fit, and figure 5.9 shows $-\log(\mathcal{L}) + \log(\mathcal{L}_{max})$ about the minimum. As the curve is non-parabolic we calculate the asymmetric error on the lifetime τ using the MINOS algorithm of the MINUIT package [80]. Although the other parameters have asymmetric errors too, we are content with their symmetric approximation as the error on those parameters is not of prime interest. We get for the average B_s lifetime:

$$\tau(B_s) = (1.314^{+0.148}_{-0.129}(\text{stat.})) \text{ ps}, \text{ corresponding to} \quad (5.17)$$

$$c\tau(B_s) = (394.0^{+44.3}_{-38.8}(\text{stat.})) \mu\text{m}. \quad (5.18)$$

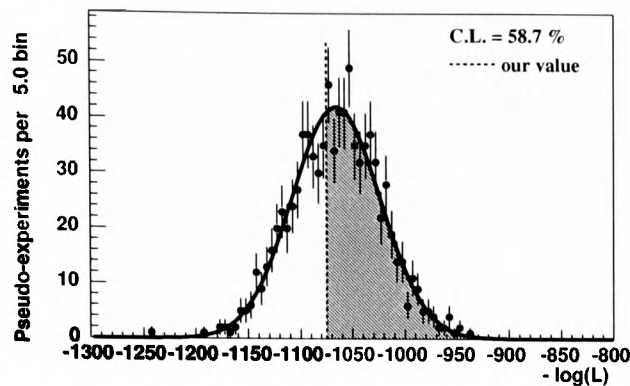


Figure 5.8: Distribution of the negative log-likelihood of many generated experiments with 1,023 $B_s^0 \rightarrow J/\psi \phi$ candidates each. A Gaussian fit is overlaid. The dashed line represents the likelihood value of the data fit.

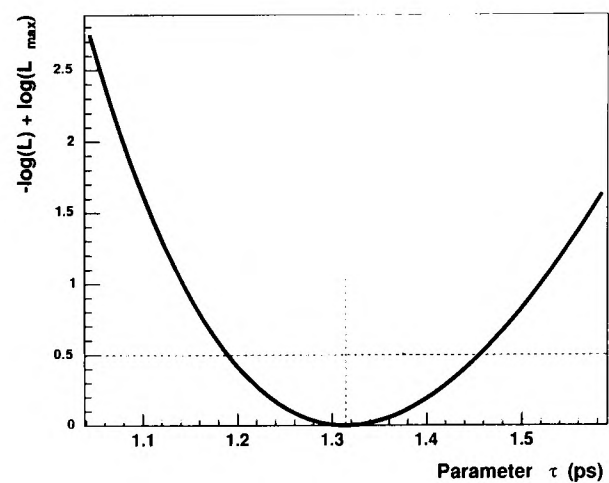


Figure 5.9: Likelihood $-\log(\mathcal{L}) + \log(\mathcal{L}_{max})$ as a function of lifetime fit parameter τ in the region about the minimum for $B_s^0 \rightarrow J/\psi \phi$.

	M	W	A	f_{sig}	τ	S_t	f_-	f_+	f_{++}	τ_-	τ_+
W	9										
A	-2	1									
f_{sig}	6	20	1								
τ	-2	-4	-1	-22							
S_t	-1	0	0	0	-0						
f_-	-0	-1	0	1	-1	-33					
f_+	1	-4	-1	-12	12	-28	17				
f_{++}	-8	-19	0	-6	-11	-7	4	-22			
τ_-	0	1	0	-0	0	27	-73	-16	-4		
τ_+	5	10	0	6	1	20	-14	-22	-69	13	
τ_{++}	8	15	-0	4	5	6	-3	22	-83	3	59

Table 5.4: Correlation coefficients for the $B_s^0 \rightarrow J/\psi \phi$ fit in percent. The upper half and the diagonal are not shown for clarity.

Systematic uncertainties as well as cross-checks are studied in chapter 6.

Chapter 6

Systematics and Cross-Checks

In this chapter we study possible systematic effects and their influence on the measured lifetime τ . We also perform cross-checks with different methods to gain trust and confidence in our main analysis procedure.

6.1 Residual Silicon Mis-Alignment

We determine the position of the B decay vertex from intersecting tracks to a common point. The silicon hits assigned to the tracks provide the accurate spatial information. The position of a hit is known to about 10 μm relative to the individual ladder position which detects the hit [82].

The knowledge about the position of a given ladder within the silicon system is roughly subdivided into two parts. First, there is the internal position of a given ladder within the silicon system, which is dealt with by the internal alignment. Then, there is the positioning of the silicon system in the global reference frame, which is addressed as global alignment.

Corrections are made on the individual nominal wafer position based on a certain alignment procedure, which performs track fits on part of the data, with free parameters coming from a misalignment model [82]. The resulting set of alignment parameters is then used to correct misalignments in all runs.

The internal alignment takes bow effects into account. When two ends of a ladder are pinned to the SVX frame, it has been observed that the ladder in most cases is bent outwards, so that the two central wafers are at a radius larger than nominal. Only in a few cases the ladders are bowed inwards. The radial shift is on average +70 μm , but there are some ladders which bend out by as much as 200 μm . Using the overlap regions for the various ladders the uncertainty of bowing effects was reduced to ± 20 μm on the outer layer of the SVX. This propagates to an 0.02 % error on the measurement of L_{xy} . The uncertainty on the internal alignment between the outer SVX layer and the other layers was found to be negligible [82].

In order to probe the sensitivity of misalignments of the SVX on the lifetime measurements, we conduct a variety of checks using our highest statistics sample $B_u^\pm \rightarrow J/\psi K^\pm$. For these checks we use different versions of alignment, which are

Name	Description
A	Best version of alignment.
B	Alignment constants from mechanical survey.
C ₁	Best with all ladders shifted radially by +50 μm .
C ₂	Best with all ladders shifted radially by -50 μm .
C ₃	Best with all ladders bowed outward by +50 μm .
C ₄	Best with all ladders bowed outward by -50 μm .

Table 6.1: Different versions of alignment constants used for systematic studies.

defined in table 6.1. Versions C_1 , C_2 , C_3 , and C_4 have been created for systematic studies by experts [82]. The values of $\pm 50 \mu\text{m}$ are considered to be reasonable estimates of our uncertainties on the alignment. In the future we hopefully get more different versions to study the effect of misalignments more thoroughly.

We make the following comparisons:

- Difference in reconstructed lifetime between alignment version A and B in data for common events.
- Difference in reconstructed lifetime for realistically simulated events between A and C_1, C_2, C_3, C_4 , respectively.
- Same for data.

Case	Alignment	Type of Data	$c\tau$ (μm)
1	A	Data	487 ± 14
2	A	Data common with B	485 ± 15
3	A	Simulation	476 ± 14
4	B	Data	477 ± 15
5	C ₁	Data	488 ± 14
6	C ₂	Data	490 ± 14
7	C ₃	Data	487 ± 14
8	C ₄	Data	489 ± 14
9	C ₁	Simulation	477 ± 14
10	C ₂	Simulation	478 ± 14
11	C ₃	Simulation	477 ± 14
11	C ₄	Simulation	478 ± 14

Table 6.2: Summary of $B_u^\pm \rightarrow J/\psi K^\pm$ $c\tau$ measurements using different versions of alignment.

The result of the different $B_u^\pm \rightarrow J/\psi K^\pm$ lifetime measurements is summarised in table 6.2. The statistics are too small to infer significant alignment effects. We

use the change of the central value to estimate such effects. Comparing case 4 to 1 gives a systematic shift of $-10\text{ }\mu\text{m}$. Such a shift comes from an alignment-dependent selection efficiency, and also from the alignment affecting the B vertex position. A separation of these two effects is possible by measuring the lifetime with alignment version A, but only for data which is common to the dataset obtained with version B. This measurement was done in case 2 and leads to a shift of only $-2\text{ }\mu\text{m}$. Alignment versions C_i ($i = 1, 2, 3, 4$) provide the best guess on the uncertainty on the individual ladder alignments, which is $50\text{ }\mu\text{m}$ [82]. From data we get an increase of $0.3\text{ }\mu\text{m}$. It is interesting to note that all cases lead to a small increase. This makes us conclude that lifetime variations are mostly statistical (bad alignment lowers the number of candidates slightly) and individual ladder alignments have no significant impact on the lifetime. The fitted value of the simulated sample using alignment version A gives $(476 \pm 14)\text{ }\mu\text{m}$, which is consistent within statistics with the value of $492\text{ }\mu\text{m}$ used for generating the sample. Versions C_i ($i = 1, 2, 3, 4$) give a systematic increase of $1.2\text{ }\mu\text{m}$.

We conclude that $10\text{ }\mu\text{m}$ forms a reasonable upper estimate on alignment effects, and $2\text{ }\mu\text{m}$ a reasonable lower estimate. We consider the value of $5\text{ }\mu\text{m}$ as best estimate for the systematic uncertainty due to residual mid-alignment, which we regard as common across all decay modes.

6.2 Beam Spot Size

Under normal circumstances the beam spot error matrix is determined for each run along with the beam spot position. This on average results in a Gaussian beam profile with $\sigma_x = \sigma_y = 30\text{ }\mu\text{m}$. Unfortunately, at present time this mechanism is not well tested and the size of the beam spot not entirely understood. Values between 25 and $35\text{ }\mu\text{m}$ have been measured. There are also indications of a z dependence of the beam spot size. In addition, for a number of runs technical problems¹ force us to use a default value of $30\text{ }\mu\text{m}$.

In order to study the influence on the lifetime measurement, we repeat the analysis for $B_u^\pm \rightarrow J/\psi K^\pm$, but fix the beam spot size to 25 , 30 and $35\text{ }\mu\text{m}$. No variation of the lifetime is observed. The effect, as expected, is completely absorbed by a change of the scale factor S_t , which changes from 1.42 for $25\text{ }\mu\text{m}$ to 1.28 for $35\text{ }\mu\text{m}$.

No systematic error is assigned due to the uncertainty on the beam spot size.

6.3 Proper Decay Time Resolution Function

High statistics studies using the inclusive B decays $B \rightarrow J/\psi X$ firmly determine the ct resolution function for L_{xy} , where the vertex is only partially reconstructed by the J/ψ vertex [83]. The study points out that a single Gaussian seems to be an inadequate description of the resolution function, while a double Gaussian or even a double Gaussian with exponential tails do a better job.

¹empty database entries

For our low statistics exclusive B decays we have chosen a single Gaussian. This was done to keep the model as simple as possible, although we have already seen indications in the negative tails of the $B_u^\pm \rightarrow J/\psi K^\pm$ and $B_d^0 \rightarrow J/\psi K^{*0}$ data of the need for a better description of the background.

In this section we study systematic uncertainties on the measured lifetime arising from the choice of the resolution function. We study the following two resolution functions:

- Double Gaussian:

$$r(t - t', \sigma_t; S_t, f_2, S_2) \equiv (1 - f_2) \cdot G(t - t', S_t \sigma_t) + f_2 \cdot G(t - t', S_2 \sigma_t), \quad (6.1)$$

where $G(\mu, \sigma)$ is a standard Gaussian with mean μ and variance σ , f_2 the fraction of the second Gaussian, and S_2 its error scale factor.

- Single Gaussian with symmetric exponential tails:

$$r(t - t', \sigma_t; S_t, f_e, \tau_e) \equiv (1 - f_e) \cdot G(t - t', S_t \sigma_t) + \frac{f_e}{2\tau_e} \exp\left(-\frac{|t - t'|}{\tau_e}\right), \quad (6.2)$$

with f_e being the fraction of the exponential tails and τ_e their lifetime.²

A change of resolution function will have the biggest direct influence on the prompt background. The direct influence on the signal contribution is small, due to low statistics. The big influence on the prompt background will however have impact on the tails and therefore influence indirectly all other contributions.

Table 6.3 shows the result of the $B_u^\pm \rightarrow J/\psi K^\pm$ fit for the three different resolution functions. In the double Gaussian case the second Gaussian helps the scale factor S_t , and the negative tail can be flatter. The symmetric exponential tails of the other resolution function have a very small fraction of 0.5 %. The resulting very long lifetime helps very much the regular negative exponential tail (f_- , τ_-) as well as the long-lived positive tail. The impact on the lifetime however is small in both cases. Table 6.4 summarises the obtained lifetimes using different resolution functions; not only for $B_u^\pm \rightarrow J/\psi K^\pm$, but also for the other two decay modes. As a result of this study we assign a systematic uncertainty of ± 0.01 ps (or ± 3 μm for ct) due to the choice of resolution function.

²The convolution of the exponential with this resolution function breaks up into the sum of a term like in equation 5.7 and the convolution of an exponential with a symmetric exponential. The latter gives

$$\int_{-\infty}^{+\infty} dt' \frac{1}{\tau} e^{-t'/\tau} \cdot \frac{1}{2\tau_e} e^{-|t-t'|/\tau_e} = \begin{cases} \frac{1}{2(\tau+\tau_e)} \exp\left(-\frac{t}{\tau_e}\right) & , \quad t < 0 \\ \frac{1}{2(\tau-\tau_e)} \left[\exp\left(-\frac{t}{\tau}\right) - \exp\left(-\frac{t}{\tau_e}\right) \right] & , \quad t \geq 0, \tau \neq \tau_e \\ \left[\frac{t}{2\tau^2} + \frac{1}{4\tau} \right] \exp\left(-\frac{t}{\tau}\right) & , \quad t \geq 0, \tau = \tau_e \end{cases} \quad (6.3)$$

	Single Gaussian	Double Gaussian	Gaussian + Exp. Tails	Unit
M	5.2745 ± 0.0005	5.2745 ± 0.0005	5.2745 ± 0.0005	MeV/c^2
W	15.92 ± 0.50	15.98 ± 0.50	15.87 ± 0.50	MeV/c^2
A	-0.66 ± 0.59	-0.66 ± 0.59	-0.67 ± 0.59	c^2/GeV
f_{sig}	0.146 ± 0.005	0.147 ± 0.005	0.146 ± 0.005	
τ	1.623 ± 0.048	1.617 ± 0.048	1.621 ± 0.048	ps
S_t	1.35 ± 0.02	1.11 ± 0.04	1.30 ± 0.02	
f_-	0.017 ± 0.002	0.009 ± 0.002	0.038 ± 0.008	
f_+	0.105 ± 0.008	0.090 ± 0.009	0.104 ± 0.011	
f_{++}	0.025 ± 0.007	0.023 ± 0.008	0.042 ± 0.011	
τ_-	1.050 ± 0.116	1.585 ± 0.226	0.348 ± 0.063	ps
τ_+	0.485 ± 0.051	0.543 ± 0.071	0.370 ± 0.051	ps
τ_{++}	2.028 ± 0.326	2.110 ± 0.374	1.382 ± 0.209	ps
		f_- 0.281 ± 0.067	f_e 0.005 ± 0.002	
		S_t 2.10 ± 0.16	τ_e $(4.171 \pm 1.012) \text{ ps}$	

Table 6.3: Results of the fit to $B_u^\pm \rightarrow J/\psi K^\pm$ data using different parameterisations of the resolution function.

Decay Mode	Single Gaussian	Double Gaussian	Gaussian + Exp. Tails
$B_u^\pm \rightarrow J/\psi K^\pm$	1.623 ± 0.048	1.617 ± 0.048	1.621 ± 0.048
$B_d^0 \rightarrow J/\psi K^{*0}$	1.510 ± 0.065	1.501 ± 0.064	1.509 ± 0.065
$B_s^0 \rightarrow J/\psi \phi$	1.314 ± 0.138	1.314 ± 0.138	1.317 ± 0.138

Table 6.4: Effect of different resolution functions on the lifetime parameter τ (in ps).

6.4 Proper Decay Time Background Parameterisation

In this section we study effects of the proper decay time background parameterisation on the fitted lifetime. The prompt contribution, described by a δ -function convolved with the resolution function, is not controversial, bar the exact shape of the latter as we have seen in the previous section.

As a first case study we make the two short-lived tails symmetric in their lifetime and their fraction. The four parameters f_- , f_+ , τ_- , and τ_+ become the two parameters f_\pm and τ_\pm . The result of applying this fit model to the $B_u^\pm \rightarrow J/\psi K^\pm$ is shown in the middle column of table 6.5. The correlation coefficients can be found in table 6.6. The long-lived positive tail now has to take on the task of the former short-lived positive tail. Therefore the background model is not able to account for long-lived background events; the signal lifetime parameter gets correlated with the parameters describing the background proper decay time distribution. Also the increase in width W supports this statement, as the mass Gaussian has to accommodate long-lived

Our model		Sym. tails	No long-lived tail	Unit
M	5.2745 ± 0.0005	5.2745 ± 0.0005	5.2745 ± 0.0007	GeV/c^2
W	15.92 ± 0.50	16.22 ± 0.51	16.43 ± 0.52	MeV/c^2
A	-0.66 ± 0.59	-0.66 ± 0.58	-0.66 ± 0.59	c^2/GeV
f_{sig}	0.146 ± 0.005	0.147 ± 0.005	0.148 ± 0.005	
τ	1.623 ± 0.048	1.645 ± 0.048	1.660 ± 0.048	ps
S_t	1.35 ± 0.02	1.36 ± 0.01	1.36 ± 0.02	
f_-	0.017 ± 0.002	$f_{\pm} \quad 0.031 \pm 0.004$	$f_- \quad 0.016 \pm 0.002$	
f_+	0.105 ± 0.008	$f_{++} \quad 0.105 \pm 0.007$	$f_+ \quad 0.110 \pm 0.006$	
τ_-	1.050 ± 0.116	$\tau_{\pm} \quad 1.691 \pm 0.151$	$\tau_- \quad 1.101 \pm 0.146$	ps
τ_+	0.485 ± 0.051	$\tau_{++} \quad 0.564 \pm 0.044$	$\tau_+ \quad 0.818 \pm 0.045$	ps
f_{++}	0.025 ± 0.007			
τ_{++}	2.028 ± 0.326			ps

Table 6.5: Fit results on the $B_u^{\pm} \rightarrow J/\psi K^{\pm}$ data with different alternative background fit models.

background which blows up the width. In addition we see an increase of the signal lifetime by 0.022 ps, which is compatible with the signal lifetime accounting for the long-lived background.

	M	W	A	f_{sig}	τ	S_t	f_{\pm}	f_+	τ_{\pm}	τ_{++}
W	4									
A	-2	0								
f_{sig}	2	30	0							
τ	-2	-6	0	-24						
S_t	-3	-1	1	-1	-0					
f_{\pm}	0	-1	-0	3	-3	-31				
f_{++}	-1	-14	-0	-22	15	-19	-16			
τ_{\pm}	-1	-10	1	-4	-6	20	-42	13		
τ_{++}	-1	-20	-0	-10	-6	21	-31	-34	4	

Table 6.6: Correlation coefficients for the fit on $B_u^{\pm} \rightarrow J/\psi K^{\pm}$ data with the model with symmetric short-lived exponential tails.

A similar scenario arises when we drop the long-lived positive background contribution from our fit model. The fit result is displayed in the third column of table 6.5 and the correlation coefficients can be found in table 6.7. Again the long-lived positive background has to be accounted for on average by the short-lived positive background contribution. This hypothesis is backed by an increase in signal lifetime (+0.037 ps) and correlations of the lifetime with the short-lived positive background parameters.

	M	W	A	f_{sig}	τ	S_t	f_-	f_+	τ_-
W	7								
A	-7	0							
f_{sig}	-4	29	2						
τ	-1	-4	-1	-23					
S_t	-3	3	0	-2	-1				
f_-	2	1	-0	2	-1	-30			
f_+	13	-17	-2	-28	15	-26	11		
τ_-	8	3	-1	0	-1	23	-49	-14	
τ_+	-7	-25	1	-6	-17	21	-10	-36	5

Table 6.7: Correlation coefficients for the fit on $B_u^\pm \rightarrow J/\psi K^\pm$ data with the model without a long-lived positive tail.

We see the same effects when applying the different background model version to the $B_d^0 \rightarrow J/\psi K^{*0}$ data. In the $B_s^0 \rightarrow J/\psi \phi$ case (table 6.8) it seems that the low statistics of the sample allows us to drop the long-lived positive background contribution without a big effect. Note that f_{++} is compatible with nought. We see a decrease in lifetime rather than the expected increase, which is likely to come from low statistics: both, the errors on the signal lifetime τ and the long-lived lifetime τ_{++} are big.

	Our model		Sym. tails		No long-lived tail	Unit
M	5.3597 ± 0.0011		5.3598 ± 0.0011		5.3598 ± 0.0012	GeV/ c^2
W	10.06 ± 0.96		10.13 ± 0.91		10.14 ± 0.98	MeV/ c^2
A	-0.89 ± 1.28		-0.88 ± 1.28		-0.88 ± 1.28	c^2/GeV
f_{sig}	0.122 ± 0.013		0.124 ± 0.013		0.124 ± 0.013	
τ	1.314 ± 0.138		1.300 ± 0.135		1.306 ± 0.128	ps
S_t	1.45 ± 0.05		1.46 ± 0.05		1.47 ± 0.05	
f_-	0.021 ± 0.011	f_\pm	0.058 ± 0.034	f_-	0.018 ± 0.011	
f_+	0.128 ± 0.022	f_{++}	0.088 ± 0.019	f_+	0.114 ± 0.018	
τ_-	0.550 ± 0.217	τ_\pm	0.438 ± 0.147	τ_-	0.599 ± 0.230	ps
τ_+	0.526 ± 0.119	τ_{++}	0.931 ± 0.161	τ_+	0.820 ± 0.110	ps
f_{++}	0.011 ± 0.013					
τ_{++}	2.850 ± 1.822					ps

Table 6.8: Fit results on the $B_s^0 \rightarrow J/\psi \phi$ data with different alternative background fit models.

As the changes in signal lifetime can be traced back to an inappropriate description of the background proper decay time distribution, which provides a worse decoupling of the background from signal, we can not treat them as systematic uncertainties.

They rather provide a cross-check and additional indications that our chosen fit model is better than the others.

On the other hand we feel uncomfortable by claiming that our choice of background model is perfect and would not introduce a systematic uncertainty. One possibility would be to refine the background fit model rather than simplify it. The effect of for example an additional positive background contribution on the signal lifetime would give us an estimate of this systematic uncertainty. Such refinements essentially have been performed already in the last section with the change of the resolution function. The effect was on the order 0.01 ps. As this systematic uncertainty is already denoted as the systematic uncertainty due to the choice of the resolution function, we do not assign a systematic uncertainty on the choice of the proper decay time background model.

6.5 Biases from the Fitting Procedure

This section will demonstrate that the fitting procedure is properly implemented and reveal possible biases from the fitting procedure.

First we look at the most simple situation: we set the signal fraction to one and fix all parameters in the fit but the signal lifetime τ . We then generate pseudo-experiments with 5,000 events each, perform the fit and calculate the pull variable p for each experiment, which is defined as

$$p \equiv \frac{\tau_{fit} - \tau_{true}}{\sigma_{fit}}, \quad (6.4)$$

where τ_{true} is the true value of the parameter, τ_{fit} its fitted value, and σ_{fit} the error on it as determined from the fit. Figure 6.1 shows distributions of $\tau_{fit} - \tau_{true}$ (left) and the pull variable (right) for 2,000 such experiments. A Gaussian fit is overlaid and shows that there is no significant deviation from nought for $\tau_{fit} - \tau_{true}$ and that the pull distribution is standard Gaussian.

The situation changes when small datasets are being used. Figure 6.2 shows the same distributions for 2,000 pseudo-experiments with 125 events each. The $\tau_{fit} - \tau_{true}$ distribution is well centred around nought, but the mean of the pull distribution is shifted down to negative values by about three σ due to the negative tail. This indicates that σ_{fit} tends to come out too small for fits that yield a value smaller than the true value. This negative shift of the pull distribution for lifetime fits is expected and consistent with [84]. The origin is the correlation, in a lifetime fit, of the estimation of the lifetime τ and the estimation of the error σ_τ , and is similar to the set of circumstances that produces a student-T distribution rather than a normal distribution when the data itself is used to estimate the error. We cannot precisely estimate the expected downward shift due to the complicated nature of the fit we perform. Note that this does not represent a bias in the fitting procedure and has not to be corrected for, unless there is a shift in the $\tau_{fit} - \tau_{true}$ distribution.

We now come to our complete fit model. First we check the performance of our fitting procedure in the limit of infinite events, i.e. with a big number of events. This

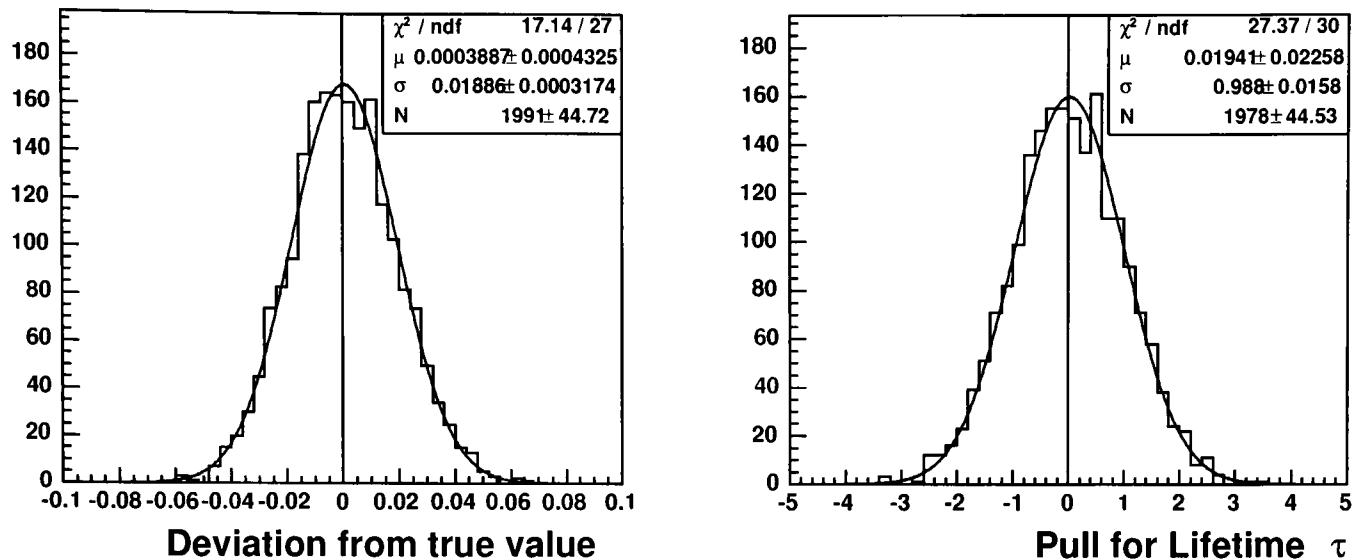


Figure 6.1: Left: $\tau_{fit} - \tau_{true}$ distribution for 2,000 pseudo-experiments with 5,000 events each. Right: pull distribution for the same experiments.

is done to spot any major flaws in the technical implementation of the fit. Table 6.9 shows the mean μ and width σ of pull distributions from 5,000 pseudo-experiments with 200,000 events each. The parameters used for generating the samples have been set to the resulting values from the fit to $B_s^0 \rightarrow J/\psi \phi$ data. The deviations $p_{fit} - p_{true}$ for parameter p , displayed in the first row, are very close to nought, although few small deviations are visible. All pull distributions are also close to standard Gaussian, with small deviations. The deviations are likely to come from correlations between fit parameters and still too low statistics. For the pulls the big negative shift of the lifetime is mainly due to the reason described in the last paragraph. For deviations of the pull width, effects from having used symmetric rather than asymmetric errors may contribute as well. Summarising, we conclude that the fitting performs well enough, without obvious flaws. Small biases in the lifetime will be dealt with in separate studies for each decay mode.

In order to check for possible biases due to the small number of events in the actual data, we generate many pseudo-experiments with the same number of candidates as in data. Table 6.10 shows the result for 5,000 experiments with 1,023 events, generated using the parameters from the fit to $B_s^0 \rightarrow J/\psi \phi$ data. Gaussian fits to the deviation distributions $p_{fit} - p_{true}$ and the pull distributions have been performed and the resulting parameters listed in the table.

Although many distributions have long tails and should not be fitted by a Gaussian, we use the Gaussian as a tool to get an approximate answer. We see that there are huge discrepancies for the various background contributions. But we note that small background fractions of the percent order lead to very small event numbers of order 10. For small statistics biases can occur because the distributions do not converge to a standard Gaussian fast enough. In the signal lifetime we see a deviation of -1.6σ and -5.8σ for the mean of the pull distribution. Before drawing any conclusions we take a look at the actual distributions, shown in figure 6.3. Rather

Par	μ_{dev}	$\frac{\mu_{dev}}{\sigma(\mu_{dev})}$	μ_{pull}	$\frac{\mu_{pull}}{\sigma(\mu_{pull})}$	σ_{pull}	$\frac{1-\sigma_{pull}}{\sigma(\sigma_{pull})}$
M	-0.000002	-2.2	-0.032 ± 0.014	-2.3	0.989 ± 0.011	1.0
W	0.000003	2.6	0.040 ± 0.015	2.8	1.033 ± 0.011	-3.0
A	-0.000620	-0.5	-0.008 ± 0.013	-0.6	0.946 ± 0.010	5.5
f_{sig}	-0.000010	-0.8	-0.019 ± 0.014	-1.4	0.950 ± 0.009	5.3
τ	-0.000165	-1.2	-0.114 ± 0.014	-8.0	1.004 ± 0.010	-0.4
S_t	0.000119	2.5	0.038 ± 0.014	2.8	0.961 ± 0.010	3.7
f_-	-0.000001	-0.1	-0.008 ± 0.014	-0.6	0.964 ± 0.010	3.5
f_+	-0.000014	-0.7	-0.014 ± 0.013	-1.0	0.944 ± 0.009	6.1
τ_-	-0.000235	-1.3	-0.034 ± 0.014	-2.5	0.967 ± 0.010	3.4
τ_+	-0.000446	-4.2	-0.057 ± 0.014	-4.1	0.980 ± 0.010	2.0
f_{++}	0.000041	3.8	0.023 ± 0.014	1.6	0.998 ± 0.011	0.1
τ_{++}	-0.000497	-0.3	-0.038 ± 0.014	-2.7	1.003 ± 0.010	-0.3

Table 6.9: Test of the fitting procedure with 5,000 pseudo-experiments with 200,000 events each. Two histograms are filled for each parameter: the deviation of the fitted value from the true value $p_{fit} - p_{true}$, and the pull variable. The table lists results from a Gaussian fit to these histograms with parameters μ for the mean and σ for the variance. We denote with $\sigma(X)$ the error on parameter X.

Par	μ_{dev}	$\frac{\mu_{dev}}{\sigma(\mu_{dev})}$	μ_{pull}	$\frac{\mu_{pull}}{\sigma(\mu_{pull})}$	σ_{pull}	$\frac{1-\sigma_{pull}}{\sigma(\sigma_{pull})}$
M	0.000009	0.6	0.007 ± 0.015	0.5	1.028 ± 0.010	-2.7
W	-0.000047	-3.5	-0.121 ± 0.014	-8.3	0.971 ± 0.010	2.9
A	0.015228	0.8	0.014 ± 0.014	1.0	1.022 ± 0.010	-2.1
f_{sig}	0.000393	2.2	-0.003 ± 0.014	-0.2	1.006 ± 0.010	-0.6
τ	-0.003013	-1.6	-0.084 ± 0.014	-5.8	0.981 ± 0.011	1.8
S_t	-0.002705	-3.8	-0.078 ± 0.014	-5.5	1.011 ± 0.011	-1.1
f_-	0.002719	17.0	0.113 ± 0.016	7.3	0.837 ± 0.011	15.3
f_+	-0.002225	-6.5	-0.076 ± 0.013	-5.8	0.935 ± 0.010	6.8
τ_-	-0.007018	-2.5	-0.183 ± 0.016	-11.2	0.847 ± 0.011	14.1
τ_+	-0.035678	-24.2	-0.372 ± 0.015	-24.4	1.006 ± 0.013	-0.5
f_{++}	0.003322	20.2	0.229 ± 0.012	19.6	0.804 ± 0.012	16.1
τ_{++}	-0.037094	-1.7	-0.124 ± 0.014	-8.9	0.698 ± 0.008	36.0

Table 6.10: Test of the fitting procedure with 5,000 pseudo-experiments with 1,023 events each and parameter values according to $B_s^0 \rightarrow J/\psi \phi$ data. Two histograms are filled for each parameter: the deviation of the fitted value from the true value $p_{fit} - p_{true}$, and the pull variable. The table lists results from a Gaussian fit to these histograms with parameters μ for the mean and σ for the variance. We denote with $\sigma(X)$ the error on parameter X.

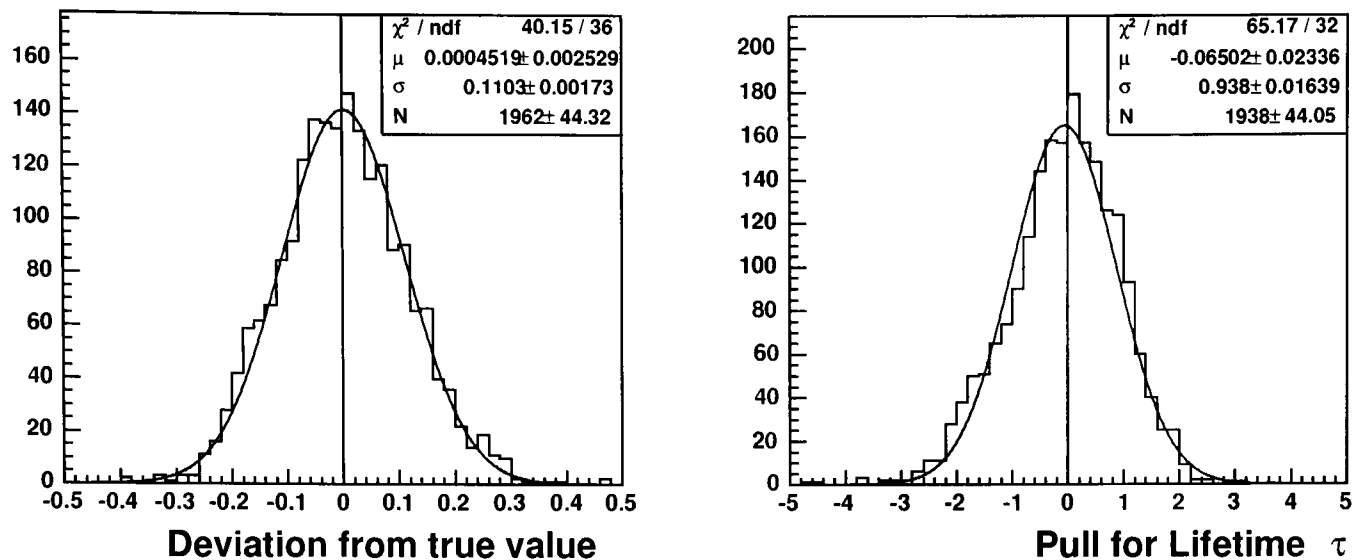


Figure 6.2: Left: $\tau_{fit} - \tau_{true}$ distribution for 2,000 pseudo-experiments with 125 events each. Right: pull distribution for the same experiments.

than fitting a Gaussian we can use $\text{RMS}/\sqrt{5,000}$ as an unbiased estimator of the standard deviation, which gives 0.0020. Therefore there is no significant bias for the lifetime and we do not assign any systematic uncertainty or apply any correction to the measured B_s lifetime.

We repeat the test for biases for our 10,516 $B_u^\pm \rightarrow J/\psi K^\pm$ candidates, for which we generate 5,000 pseudo-experiments with 10,516 events according to the results of the data fit. Table 6.11 shows the result of this test. There is no significant bias in the lifetime parameter τ . Table 6.12 shows the same result for $B_d^0 \rightarrow J/\psi K^{*0}$ from 5,000 pseudo-experiments with 16,057 events and parameters according to the data fit. Again no bias is seen.

We have studied possible lifetime biases from the fitting procedure using many pseudo-experiments. No biases or significant systematic uncertainties are present.

6.6 Selection Cuts

In this section we study the dependence of the fitted lifetime on the cuts applied in the selection procedure. We analyse the following cuts:

- $p_T(B)$ cut
- $p_T(\text{non-}J/\psi)$ cut
- Mass window of the ϕ for $B_s^0 \rightarrow J/\psi \phi$
- Mass window of the K^{*0} for $B_d^0 \rightarrow J/\psi K^{*0}$
- Fit quality cut
- σ_{Lxy} cut

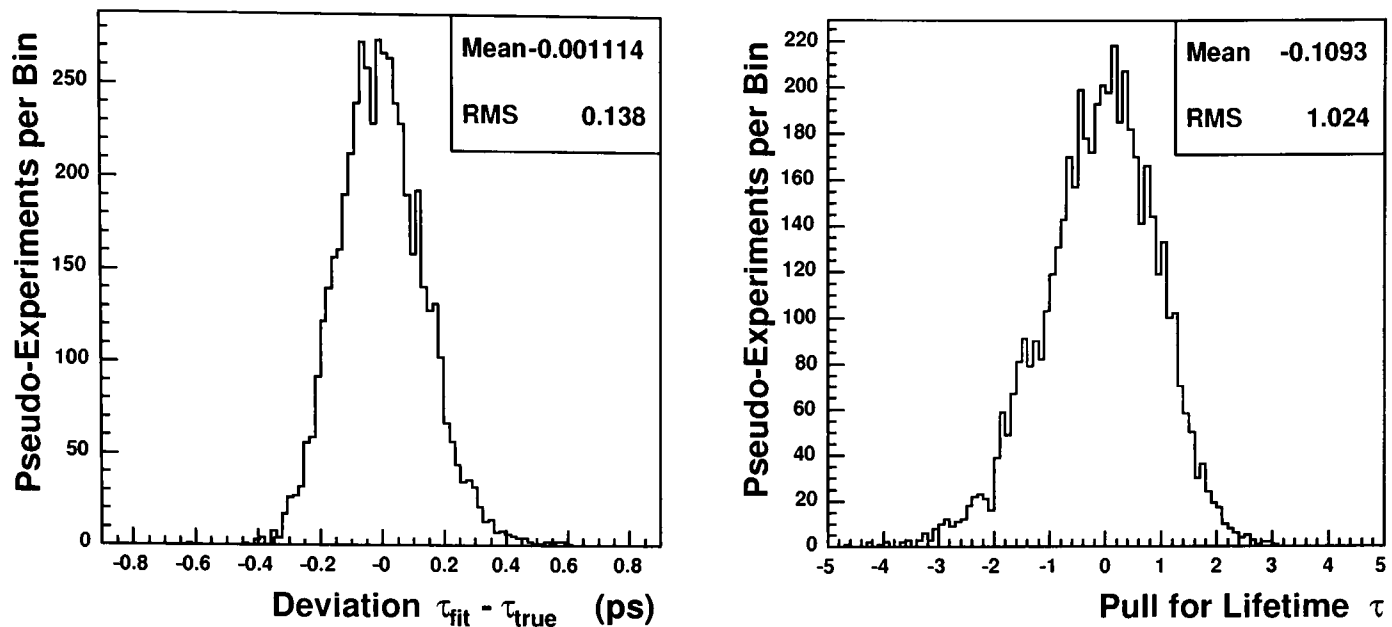


Figure 6.3: Deviation (left) and pull (right) for the lifetime parameter τ from 5,000 pseudo-experiments with 1,023 events, generated according to the result of the $B_s^0 \rightarrow J/\psi \phi$ fit from data.

Par	μ_{dev}	$\frac{\mu_{dev}}{\sigma(\mu_{dev})}$	μ_{pull}	$\frac{\mu_{pull}}{\sigma(\mu_{pull})}$	σ_{pull}	$\frac{1-\sigma_{pull}}{\sigma(\sigma_{pull})}$
M	0.000003	0.4	0.008 ± 0.014	0.5	0.992 ± 0.010	0.8
W	-0.000010	-1.6	-0.023 ± 0.014	-1.6	1.007 ± 0.011	-0.6
A	-0.000597	-0.1	0.001 ± 0.014	0.1	0.969 ± 0.010	3.2
f_{sig}	-0.000042	-0.7	-0.011 ± 0.014	-0.8	0.978 ± 0.010	2.3
τ	0.000178	0.3	-0.010 ± 0.014	-0.7	1.007 ± 0.011	-0.7
S_t	-0.000493	-2.6	-0.036 ± 0.014	-2.5	1.006 ± 0.010	-0.6
f_-	0.000064	2.3	-0.004 ± 0.015	-0.3	0.996 ± 0.010	0.4
f_+	-0.000281	-2.4	0.016 ± 0.014	1.2	0.974 ± 0.010	2.7
τ_-	0.001786	1.2	-0.057 ± 0.015	-3.8	0.977 ± 0.010	2.2
τ_+	-0.005040	-7.2	-0.124 ± 0.014	-8.6	1.010 ± 0.011	-1.0
f_{++}	0.000749	7.0	0.006 ± 0.015	0.4	0.953 ± 0.010	4.5
τ_{++}	0.014685	3.2	-0.062 ± 0.015	-4.1	0.895 ± 0.010	10.2

Table 6.11: Test of the fitting procedure with 5,000 simulated experiments with 10,516 events each and parameter values according to $B_u^\pm \rightarrow J/\psi K^\pm$ data. Two histograms are filled for each parameter: the deviation of the fitted value from the true value $p_{fit} - p_{true}$, and the pull variable. The table lists results from a Gaussian fit to these histograms with parameters μ for the mean and σ for the variance. We denote with $\sigma(X)$ the error on parameter X .

Par	μ_{dev}	$\frac{\mu_{dev}}{\sigma(\mu_{dev})}$	μ_{pull}	$\frac{\mu_{pull}}{\sigma(\mu_{pull})}$	σ_{pull}	$\frac{1-\sigma_{pull}}{\sigma(\sigma_{pull})}$
M	-0.000009	-0.9	-0.011 ± 0.014	-0.8	0.983 ± 0.010	1.6
W	-0.000003	-0.3	-0.022 ± 0.014	-1.6	1.003 ± 0.011	-0.3
A	0.001874	0.3	-0.000 ± 0.014	-0.0	0.979 ± 0.010	2.2
f_{sig}	0.000033	0.9	-0.002 ± 0.014	-0.1	0.972 ± 0.010	2.9
τ	0.000441	0.5	-0.004 ± 0.014	-0.3	0.966 ± 0.010	3.3
S_t	-0.000228	-1.5	-0.022 ± 0.014	-1.6	0.986 ± 0.010	1.4
f_-	0.000069	2.2	-0.007 ± 0.014	-0.5	1.000 ± 0.010	-0.0
f_+	-0.000353	-3.3	0.026 ± 0.014	1.8	0.972 ± 0.010	2.9
τ_-	0.001054	1.7	-0.056 ± 0.015	-3.7	0.943 ± 0.010	5.9
τ_+	-0.003004	-5.6	-0.093 ± 0.014	-6.5	1.011 ± 0.011	-1.0
f_{++}	0.000616	5.7	0.000 ± 0.015	0.0	0.998 ± 0.011	0.2
τ_{++}	0.006932	2.5	-0.089 ± 0.016	-5.7	0.951 ± 0.010	4.6

Table 6.12: Test of the fitting procedure with 5,000 simulated experiments with 16,057 events each and parameter values according to $B_d^0 \rightarrow J/\psi K^{*0}$ data. Two histograms are filled for each parameter: the deviation of the fitted value from the true value $p_{fit} - p_{true}$, and the pull variable. The table lists results from a Gaussian fit to these histograms with parameters μ for the mean and σ for the variance. We denote with $\sigma(X)$ the error on parameter X .

The transverse momentum cuts for the B meson and the non- J/ψ component are correlated and are studied simultaneously. Table 6.13 shows the fitted lifetime for each pair of cut values ($p_T(B), p_T(\text{non-}J/\psi)$) in an appropriate range around the nominal cut values. Horizontally the $p_T(\text{non-}J/\psi)$ cut varies from 1.5 GeV/ c to 3.0 GeV/ c , and vertically the cut value of $p_T(B)$ changes from 4.5 GeV/ c to 7.5 GeV/ c . All other cuts are fixed to their nominal value. The samples defined by different cut pairs largely overlap, which correlates the fitted lifetime results. We see that the mean value of the fitted lifetime remains acceptably stable under cut variations.

To determine the other systematic effects we fix all cuts to their nominal values and vary only the cut in question within a reasonable range. The nominal lifetime is compared to the lifetime at different cut values. Because the sample size and composition change when adjusting the cuts, there is a statistical component in the lifetime variation which should be separated from the systematic effect being analysed. This separation task is not well-defined. With caution we use the following prescription:

Let us assume that the studied cut value is a tighter cut than the nominal cut. Then the sample we get with the nominal cut, which we denote with \mathcal{N} , comprises completely the sample we get with the studied cut, which we will call \mathcal{C} . We then perform the lifetime fit on both samples and get as lifetime $\tau_N \pm \sigma_N$ for sample \mathcal{N} and $\tau_C \pm \sigma_C$ respectively for sample \mathcal{C} . We are interested in the change of the lifetime due to the sample composition, i.e. $\Delta\tau \equiv \tau_C - \tau_N$. For its statistical significance we have to calculate the statistical error on $\Delta\tau$. This is a non-straightforward task as

	1.5	1.75	2.0	2.25	2.5	2.75	3.0
4.5	1.628 ± 0.045	1.631 ± 0.045	1.634 ± 0.046	1.640 ± 0.049	1.648 ± 0.051	1.633 ± 0.053	1.640 ± 0.055
5.0	1.617 ± 0.046	1.618 ± 0.045	1.626 ± 0.047	1.634 ± 0.049	1.640 ± 0.051	1.636 ± 0.053	1.648 ± 0.056
5.5	1.583 ± 0.045	1.605 ± 0.044	1.610 ± 0.046	1.617 ± 0.049	1.627 ± 0.050	1.623 ± 0.052	1.646 ± 0.055
6.0	1.599 ± 0.047	1.614 ± 0.046	1.609 ± 0.047	1.616 ± 0.049	1.621 ± 0.051	1.618 ± 0.053	1.632 ± 0.056
6.5	1.618 ± 0.046	1.632 ± 0.047	1.623 ± 0.048	1.625 ± 0.050	1.633 ± 0.052	1.630 ± 0.054	1.649 ± 0.057
7.0	1.618 ± 0.047	1.639 ± 0.048	1.630 ± 0.050	1.643 ± 0.051	1.650 ± 0.053	1.645 ± 0.055	1.667 ± 0.057
7.5	1.637 ± 0.048	1.652 ± 0.051	1.640 ± 0.053	1.649 ± 0.054	1.657 ± 0.055	1.639 ± 0.056	1.659 ± 0.058

Table 6.13: Cut variation of $p_T(\text{B})$ (horizontal) and $p_T(\text{non-J}/\psi)$ (vertical) for the $\text{B}_u^\pm \rightarrow \text{J}/\psi \text{K}^\pm$ decay mode. All other cuts are kept to their nominal value. In bold face the nominal cut value pair.

sample \mathcal{C} is fully contained in sample \mathcal{N} , which correlates the errors $\sigma_{\mathcal{C}}$ and $\sigma_{\mathcal{N}}$.

Suppose we have the two statistically independent samples \mathcal{C} and $\mathcal{N} \setminus \mathcal{C}$,³ with fitted lifetime results $\tau_{\mathcal{C}} \pm \sigma_{\mathcal{C}}$ and $\tau_{\mathcal{N} \setminus \mathcal{C}} \pm \sigma_{\mathcal{N} \setminus \mathcal{C}}$. Combining those two independent measurement then yields

$$\tau_{\mathcal{N}} = \left(\frac{\tau_{\mathcal{C}}}{\sigma_{\mathcal{C}}^2} + \frac{\tau_{\mathcal{N} \setminus \mathcal{C}}}{\sigma_{\mathcal{N} \setminus \mathcal{C}}^2} \right) / \left(\frac{1}{\sigma_{\mathcal{C}}^2} + \frac{1}{\sigma_{\mathcal{N} \setminus \mathcal{C}}^2} \right) \text{ and } \frac{1}{\sigma_{\mathcal{N}}^2} = \frac{1}{\sigma_{\mathcal{C}}^2} + \frac{1}{\sigma_{\mathcal{N} \setminus \mathcal{C}}^2}. \quad (6.5)$$

Then we can write

$$\Delta\tau = \frac{\sigma_{\mathcal{C}}^2}{\sigma_{\mathcal{C}}^2 + \sigma_{\mathcal{N} \setminus \mathcal{C}}^2} (\tau_{\mathcal{C}} - \tau_{\mathcal{N} \setminus \mathcal{C}}). \quad (6.6)$$

The error $\sigma_{\Delta\tau}$ can be obtained by propagating the errors through, and expressing afterwards $\sigma_{\mathcal{N} \setminus \mathcal{C}}$ with $\sigma_{\mathcal{C}}$ and $\sigma_{\mathcal{N}}$:

$$\sigma_{\Delta\tau} = \sigma_{\mathcal{C}} \sqrt{1 - (\sigma_{\mathcal{N}}/\sigma_{\mathcal{C}})^2}. \quad (6.7)$$

This expression has the sensible behaviour that if sample \mathcal{N} and \mathcal{C} are identical, then $\sigma_{\Delta\tau} = 0$. It was assumed that we use a cut tighter than the nominal cut. This prevents the radicand of the square root to be negative, as $\sigma_{\mathcal{C}} > \sigma_{\mathcal{N}}$. However, in our fit we have the interplay between signal and background, and fluctuations in the signal fraction parameter f_{sig} can in principle yield to $\sigma_{\mathcal{C}} < \sigma_{\mathcal{N}}$. Although as it turns out that this extreme case never happens in our study, fluctuations can sometimes lead to too small error bars. For this reason we have to cautiously interpret the result. Rather than taking a single deviation too serious we look for significant trends under cut variations.

In case the cut is looser than the nominal cut, we have the reverse situation and a similar calculation yields:

$$\sigma_{\Delta\tau} = \sigma_{\mathcal{N}} \sqrt{1 - (\sigma_{\mathcal{C}}/\sigma_{\mathcal{N}})^2}. \quad (6.8)$$

Figure 6.4 shows the variation of cuts $p_T(K)$, $p_T(\text{B}_u)$, vertex fit quality, and $\sigma(L_{xy})$. The error bars are determined with the procedure just described and only make sense when referred to the value obtained with nominal cuts, which is indicated by a vertical

³i.e. \mathcal{N} without \mathcal{C}

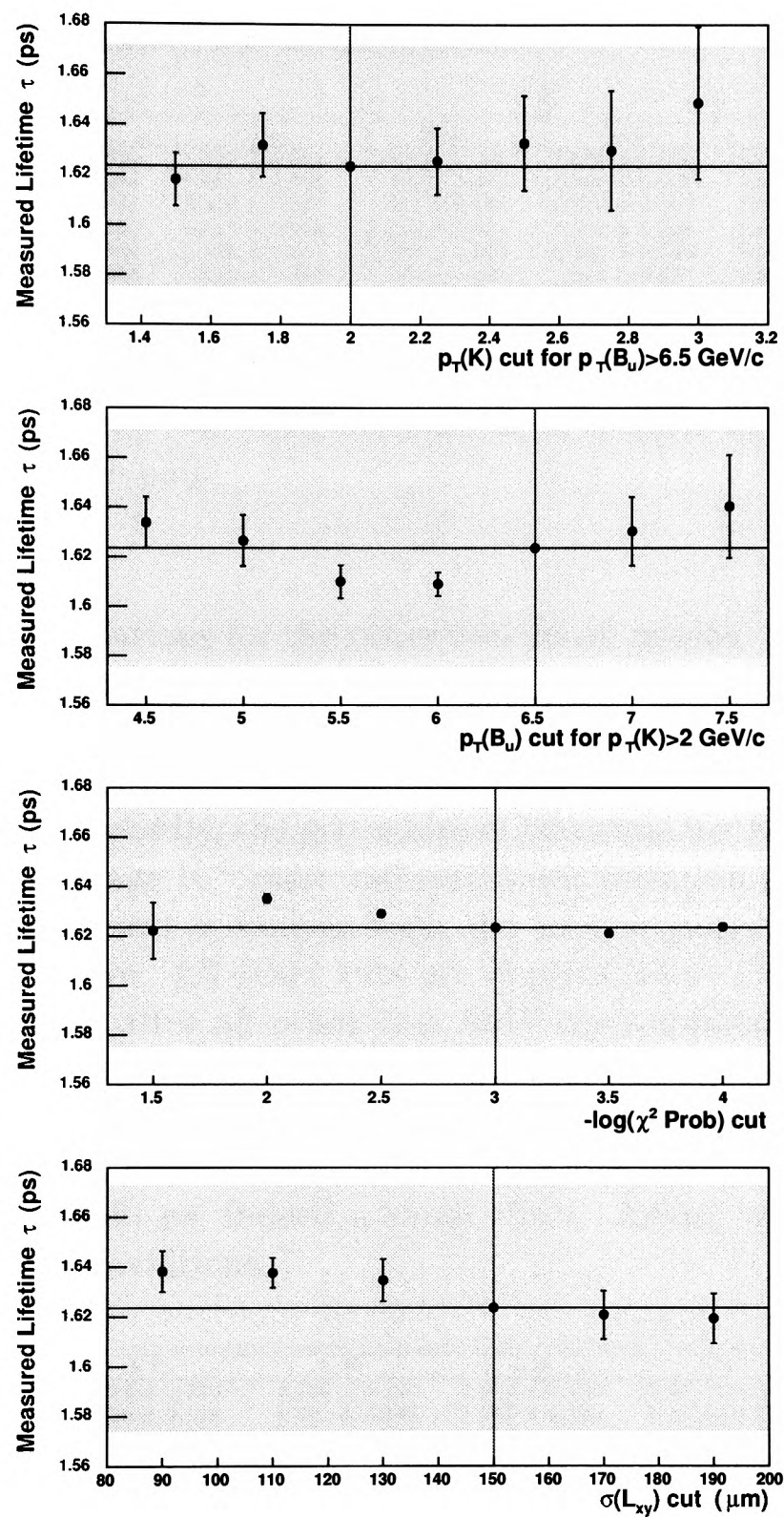


Figure 6.4: Lifetime result for $B_u^\pm \rightarrow J/\psi K^\pm$ under variation of different cut values, while keeping all other cuts on their nominal value. The error bars are with respect to the nominal cut, which is indicated by a vertical line. The error of the measurement under nominal cuts is shown as a grey (yellow) horizontal band.

line. The original error of the fit with nominal cut values is shown as a grey (yellow) band. We do not see significant deviations or trends, apart from few deviations which could be due to statistical fluctuations of the signal fraction f_{sig} . We do not assign a systematic uncertainty due to variations of the cut values studied, but rather treat these studies as cross-checks of the measurement.

	2.00	2.25	2.50	2.75	3.00
4.5	1.468 ± 0.060	1.469 ± 0.061	1.455 ± 0.066	1.488 ± 0.066	1.492 ± 0.069
5.0	1.465 ± 0.062	1.481 ± 0.062	1.469 ± 0.063	1.500 ± 0.067	1.500 ± 0.069
5.5	1.467 ± 0.061	1.478 ± 0.064	1.472 ± 0.064	1.510 ± 0.069	1.506 ± 0.070
6.0	1.506 ± 0.062	1.517 ± 0.063	1.519 ± 0.064	1.552 ± 0.066	1.547 ± 0.069
6.5	1.511 ± 0.066	1.526 ± 0.064	1.522 ± 0.065	1.549 ± 0.068	1.560 ± 0.070
7.0	1.530 ± 0.065	1.535 ± 0.069	1.529 ± 0.068	1.515 ± 0.071	1.521 ± 0.074
7.5	1.519 ± 0.071	1.520 ± 0.071	1.512 ± 0.070	1.527 ± 0.075	1.529 ± 0.076

Table 6.14: Cut variation of $p_T(B)$ (horizontal) and $p_T(\text{non-}J/\psi)$ (vertical) for the $B_d^0 \rightarrow J/\psi K^{*0}$ decay mode. All other cuts are kept to their nominal value. In bold face the nominal cut value pair.

We repeat the cut variations for the other two decay modes. Table 6.14 and figure 6.5 contain the result of the study for $B_d^0 \rightarrow J/\psi K^{*0}$. We note two technical differences to the $B_u^\pm \rightarrow J/\psi K^\pm$ case. Firstly, cut values for $p_T(K^{*0})$ smaller than 2 GeV/ c are technically challenging due to big combinatorial background. We therefore only increase the cut value. Secondly, the procedure of choosing the $K\pi$ assignment closest to the current world average K^{*0} mass influences our selection procedure. Technical limitations force us to select a sample with the loosest possible cuts, followed by applying the $K\pi$ procedure. All other cuts are applied later. Correctly, the $K\pi$ procedure should be applied after all other cuts have been applied. The complexity of the technical implementation of this order makes us prefer the wrong order, which we expect to have only a small effect. Using nominal cut values we see 103 candidates less than with the correct order, and the lifetime comes out to be (1.511 ± 0.066) ps rather than (1.510 ± 0.065) ps; indeed a small effect. Again, we are not able to see any significant trend or deviations.

	1.50	1.75	2.00	2.25	2.50	2.75
4.5	1.358 ± 0.124	1.313 ± 0.123	1.290 ± 0.126	1.342 ± 0.146	1.400 ± 0.160	1.494 ± 0.182
5.0	1.371 ± 0.130	1.320 ± 0.128	1.291 ± 0.130	1.349 ± 0.162	1.412 ± 0.171	1.485 ± 0.180
5.5	1.326 ± 0.130	1.327 ± 0.130	1.288 ± 0.137	1.337 ± 0.152	1.394 ± 0.154	1.450 ± 0.169
6.0	1.334 ± 0.131	1.334 ± 0.135	1.292 ± 0.130	1.354 ± 0.142	1.409 ± 0.150	1.457 ± 0.174
6.5	1.331 ± 0.136	1.344 ± 0.135	1.314 ± 0.138	1.360 ± 0.149	1.417 ± 0.152	1.461 ± 0.171
7.0	1.329 ± 0.140	1.339 ± 0.140	1.338 ± 0.142	1.368 ± 0.152	1.444 ± 0.169	1.487 ± 0.179
7.5	1.347 ± 0.149	1.343 ± 0.146	1.339 ± 0.148	1.366 ± 0.157	1.435 ± 0.163	1.491 ± 0.184

Table 6.15: Cut variation of $p_T(B)$ (horizontal) and $p_T(\text{non-}J/\psi)$ (vertical) for the $B_s^0 \rightarrow J/\psi \phi$ decay mode. All other cuts are kept to their nominal value. In bold face the nominal cut value pair.

We repeat the cut variations for the $B_s^0 \rightarrow J/\psi \phi$ data, shown in table 6.15 and figure 6.6. Again we do not assign any systematic uncertainty due to the studied cut variations.

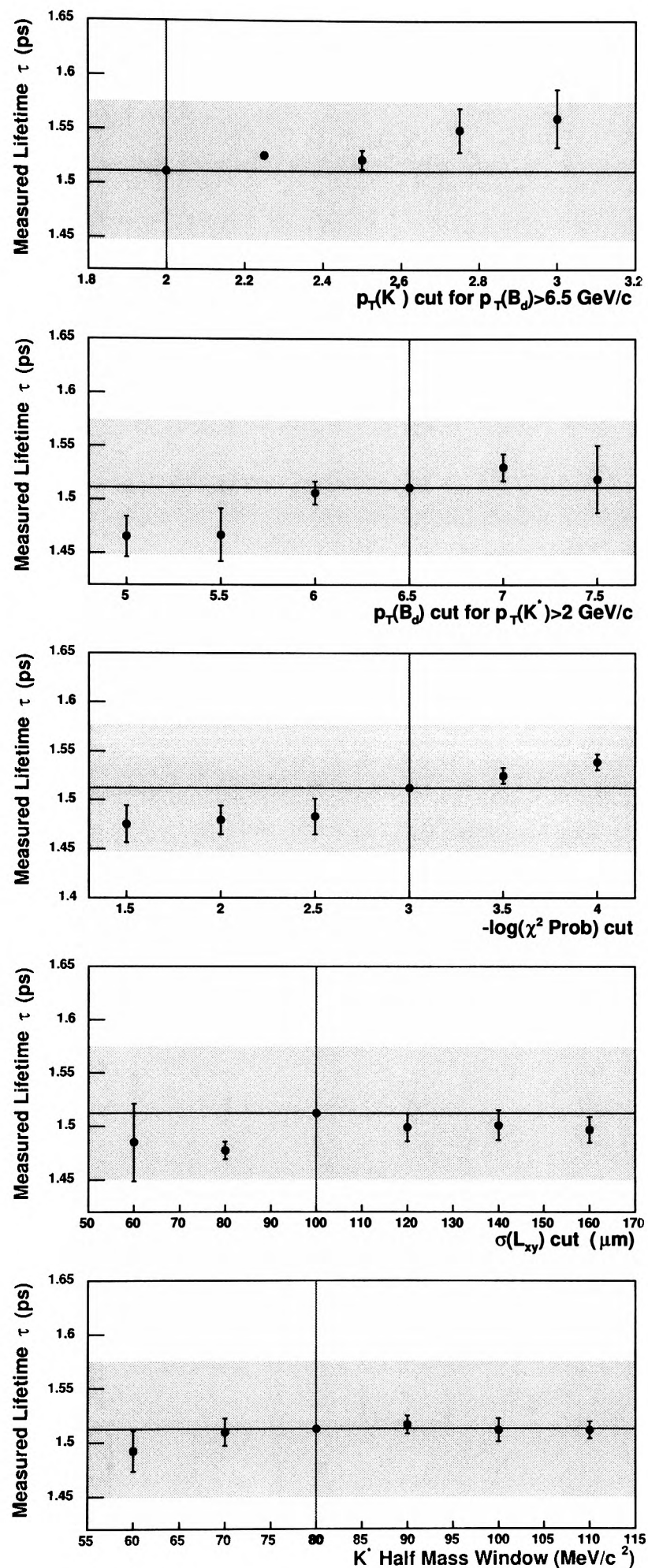


Figure 6.5: Lifetime result for $B_d^0 \rightarrow J/\psi K^{*0}$ under variation of different cut values, while keeping all other cuts on their nominal value. The error bars are with respect to the nominal cut, which is indicated by a vertical line. The error of the measurement under nominal cuts is shown as a grey (yellow) horizontal band.

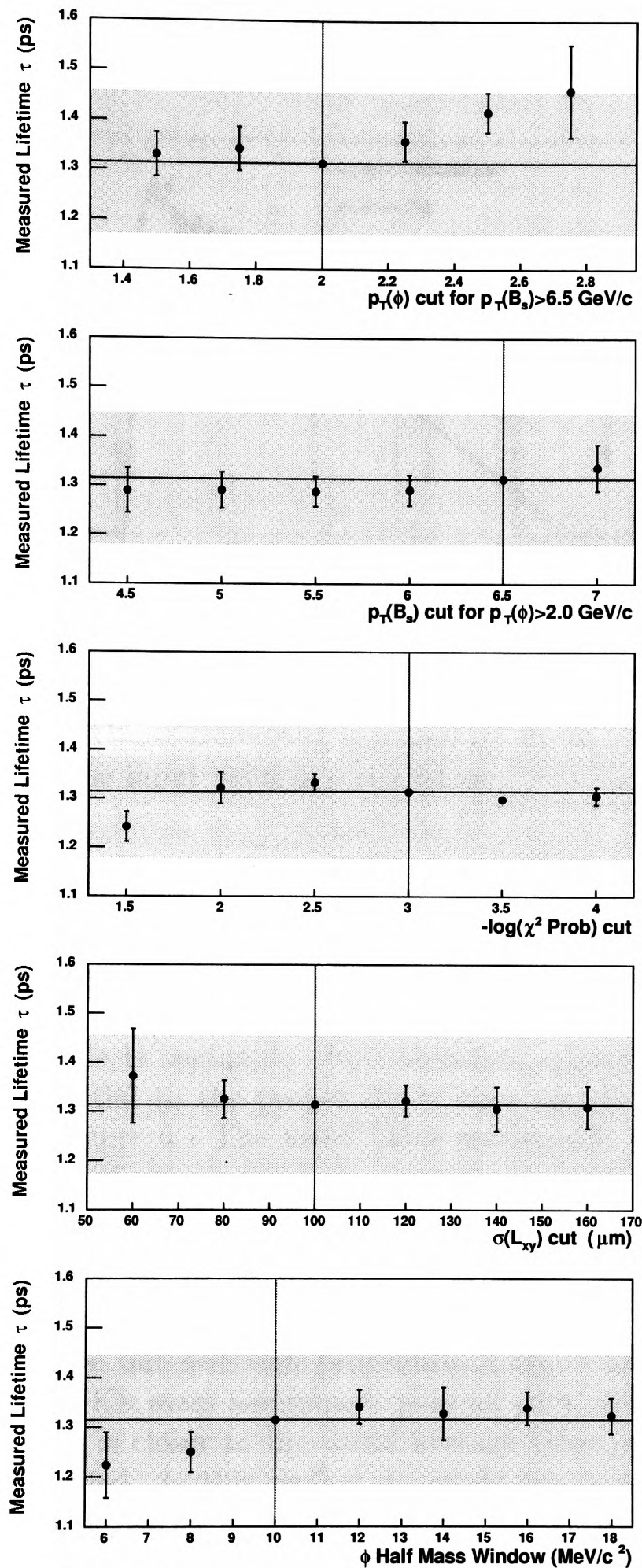


Figure 6.6: Lifetime result for $B_s^0 \rightarrow J/\psi \phi$ under variation of different cut values, while keeping all other cuts on their nominal value. The error bars are with respect to the nominal cut, which is indicated by a vertical line. The error of the measurement under nominal cuts is shown as a grey (yellow) horizontal band.

6.7 Cross-Check with Realistic Simulation

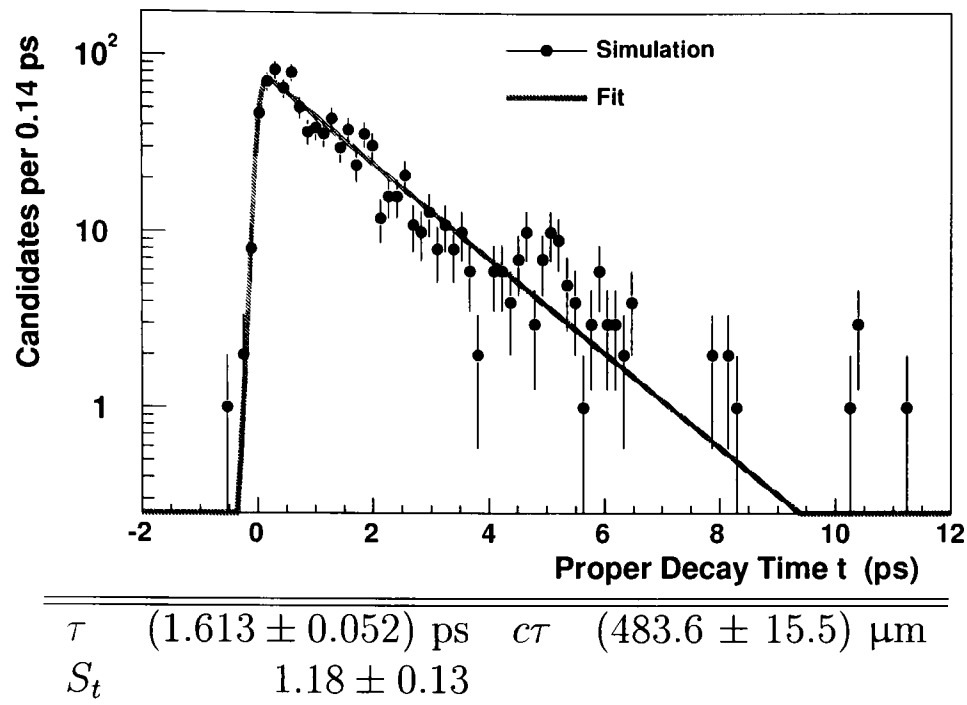


Figure 6.7: Result of a Gauss-convolved exponential fit to realistically simulated $B_u^\pm \rightarrow J/\psi K^\pm$ events. The input value is $\tau = 1.64$ ps.

In this section we verify that the reconstruction procedure, trigger, or detector acceptance do not introduce any bias in the extracted lifetime. For this reason we perform a lifetime measurement on a sample of realistically simulated $B_u^\pm \rightarrow J/\psi K^\pm$ decays. The sample is described in more detail in appendix B. The combinatorial background of this sample is negligible. It is therefore appropriate to fit a simple Gauss-convolved exponential to the proper decay time distribution of the sample. The result is shown in figure 6.7. The fitted value statistically agrees well with the true input value of $\tau = 1.64$ ps.

6.8 $K\pi$ Assignment Ambiguity

In section 4.3 we describe our selection procedure of $B_d^0 \rightarrow J/\psi K^{*0}$. In case two candidates with opposite $K\pi$ mass assignment pass all cuts, we select the candidate whose $K\pi$ invariant mass is closer to the world average value of the K^{*0} mass. The other candidate is discarded. In this section we study the procedure's influence on the fitted lifetime by comparison with a result obtained without discarding K^{*0} candidates. We keep all $K\pi$ mass assignments and change our fit model to accommodate the wrong mass assignments. We use information obtained from a sample of realistically simulated $B_d^0 \rightarrow J/\psi K^{*0}$ data. By matching reconstructed tracks to the generated quantities on the particle level we have a means to distinguish between correct and incorrect $K\pi$ mass assignments. Figure 6.8 shows the invariant mass distributions for correct (left) and incorrect (right) $K\pi$ assignments, with a Gaussian plus first

order polynomial fit. Of all candidates that pass in addition the mass window for the B_d mass, $772/(2421 + 772) = 24.2\%$ are assigned incorrectly. Both distributions peak at the same mass, but their widths differ: the width of the incorrectly assigned candidates is bigger by a factor of 2.7.

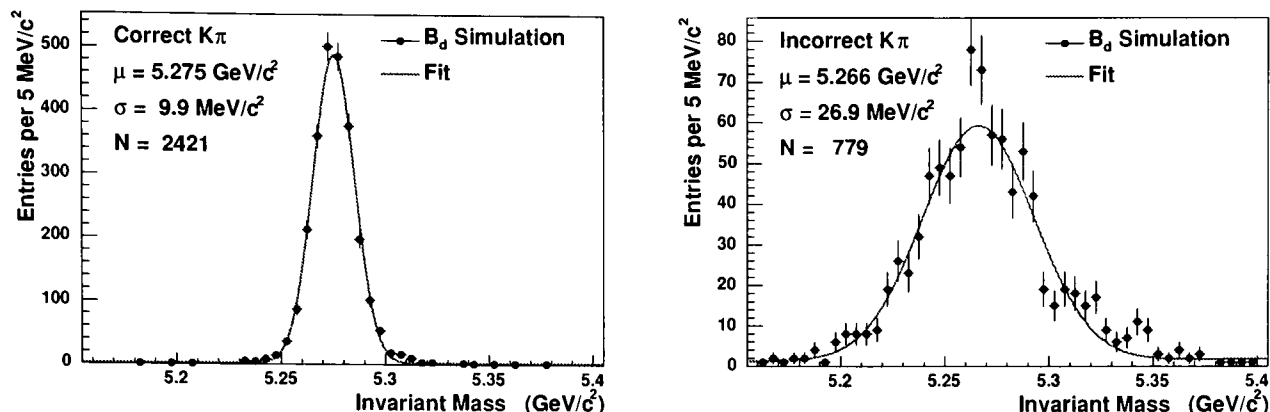


Figure 6.8: $B_d^0 \rightarrow J/\psi K^{*0}$ Invariant mass distributions for correct (left) and incorrect (right) $K\pi$ assignments, with a Gaussian plus first order polynomial fit.

Figure 6.9 shows the fit result (Gauss-convolved exponential) to the correct candidates, and figure 6.10 to the incorrect candidates. Both obtained lifetimes are statistically in agreement with each other. This is expected as both $K\pi$ assignments share the same J/ψ vertex and have the same transverse momentum (up to effects from energy corrections, which are only of the order $10 \text{ MeV}/c$). Also, the lifetimes agree well with the true input value of 1.54 ps .

To test the effect of our selection procedure on the signal lifetime, we abandon the procedure of selecting a particular mass assignment. Instead all mass assignments are selected. A simple change in our fit model can use all information obtained from the realistic simulation: Instead of a single Gaussian for the signal mass distribution we choose a weighted sum of two Gaussians:

$$G(M, W) \rightarrow 0.76 \cdot G(M, W) + 0.24 \cdot G(M, 2.7 \cdot W) . \quad (6.9)$$

Fitting the lifetime with such a model gives us a sense on how large the effect of the $K\pi$ mass assignment might be on the lifetime. Table 6.16 shows the result of this fit. We see that the central value of the lifetime has increased by 0.026 ps . Since the different selection procedure has increased the sample size from 920 to 1240 candidates, it is not clear whether this is a real systematic effect or a statistical effect. We take half the value of the deviation as systematic uncertainty: 0.013 ps ($4 \mu\text{m}$).

6.9 Cross-Check: Background from Sidebands

As additional cross-check of our simultaneous mass and decay time fit, we perform the following procedure:

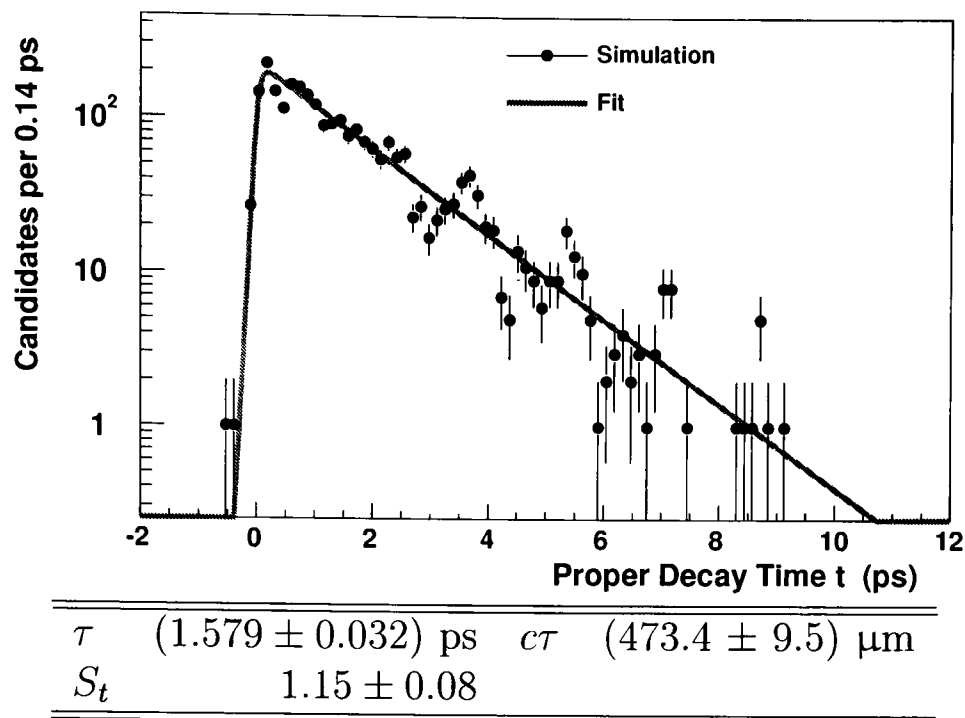


Figure 6.9: Result of a Gauss-convolved exponential fit to realistically simulated $B_d^0 \rightarrow J/\psi K^{*0}$ events, reconstructed with the *correct* $K\pi$ assignment. The input value is $\tau = 1.54 \text{ ps}$.

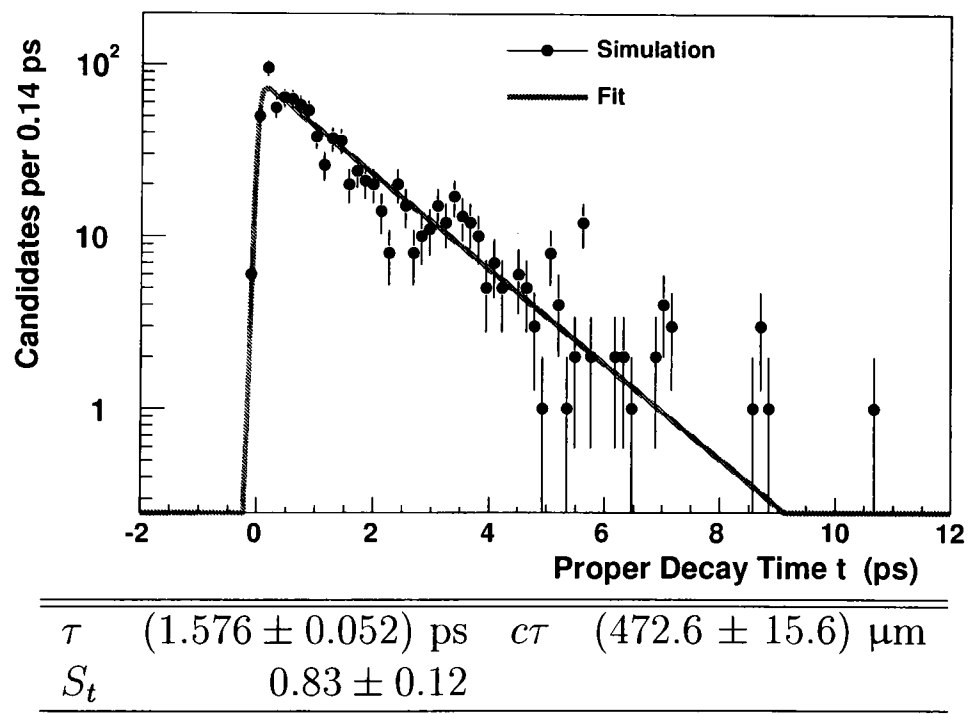


Figure 6.10: Result of a Gauss-convolved exponential fit to realistically simulated $B_d^0 \rightarrow J/\psi K^{*0}$ events, reconstructed with the *incorrect* $K\pi$ assignment. The input value is $\tau = 1.54 \text{ ps}$.

M	$(5.2740 \pm 0.0006) \text{ GeV}/c^2$		
W	$(12.64 \pm 0.61) \text{ MeV}/c^2$		
A	$(-1.92 \pm 0.40) c^2/\text{GeV}$		
f_{sig}	0.061 ± 0.003	N_{sig}	1236.7 ± 54.9
τ	$(1.536 \pm 0.058) \text{ ps}$	$c\tau$	$(460.3 \pm 17.4) \mu\text{m}$
S_t	1.33 ± 0.01		
f_-	0.023 ± 0.002	N_-	466.2 ± 47.4
f_+	0.106 ± 0.007	N_+	2123.7 ± 135.5
f_{++}	0.026 ± 0.007	N_{++}	531.0 ± 134.7
τ_-	$(0.584 \pm 0.044) \text{ ps}$	$c\tau_-$	$(175.2 \pm 13.1) \mu\text{m}$
τ_+	$(0.429 \pm 0.036) \text{ ps}$	$c\tau_+$	$(128.7 \pm 10.8) \mu\text{m}$
τ_{++}	$(1.574 \pm 0.210) \text{ ps}$	$c\tau_{++}$	$(471.9 \pm 63.1) \mu\text{m}$

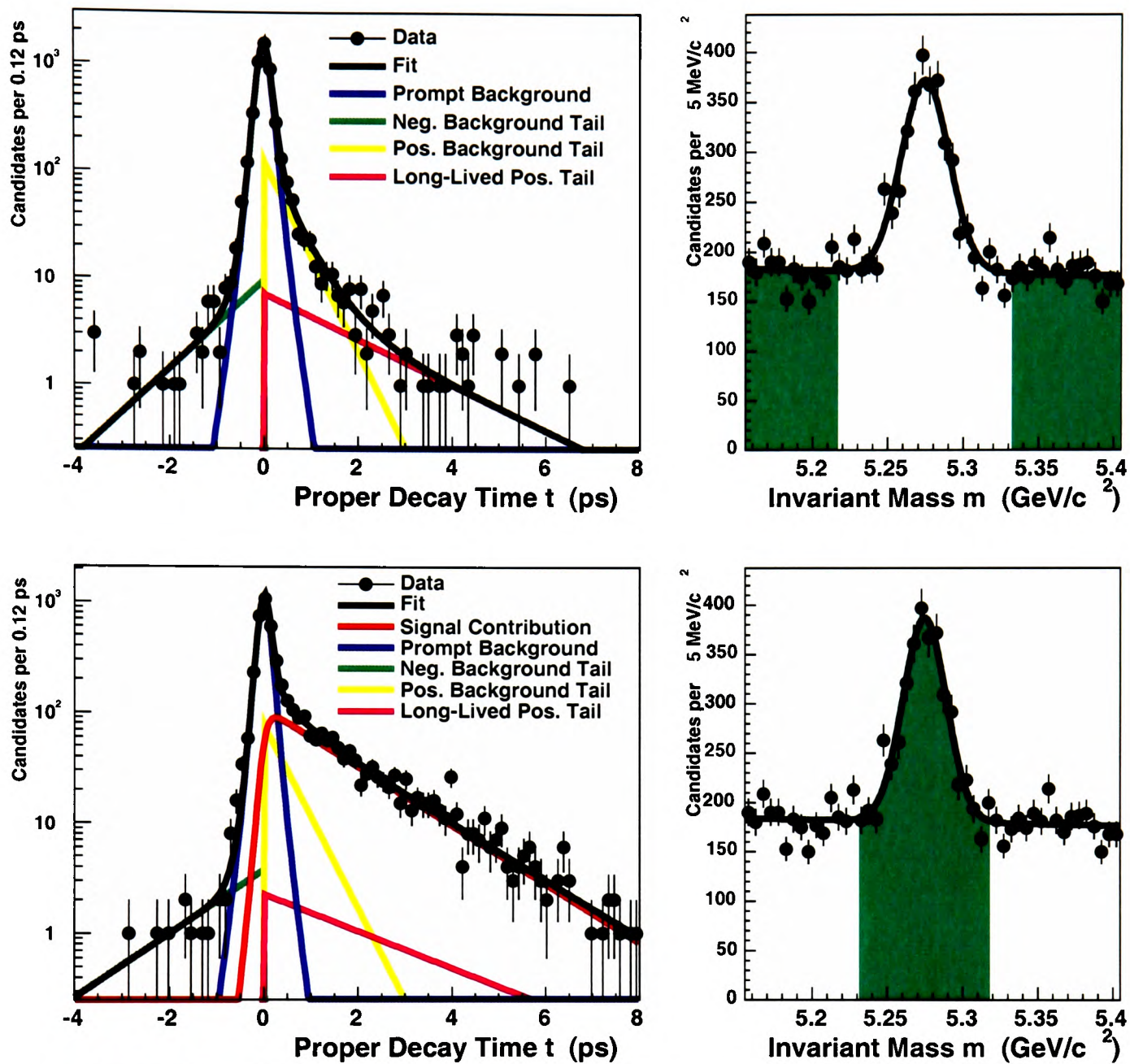
Table 6.16: Result of the combined mass and decay time fit for $B_d^0 \rightarrow J/\psi K^{*0}$, where the fit model is modified in order to include incorrectly assigned $K\pi$ candidates. Compare with our nominal result of $\tau = 1.510 \text{ ps}$.

- Binned likelihood fit of the mass distribution, with free parameters M , W , A and f_{sig} . We define the signal region as $[M - 3W, M + 3W]$ and the sidebands as $m < M - 4W$ and $m > M + 4W$.
- Unbinned likelihood fit to the sideband candidates, with free parameters S_t , f_- , f_+ , f_{++} , τ_- , τ_+ , τ_{++} .
- Unbinned likelihood fit to the signal region, with free parameter τ . All other parameters are fixed to the values determined by the previous two fits.

Figures 6.17, 6.18, 6.19 show the results of this procedure applied to the decay modes $B_u^\pm \rightarrow J/\psi K^\pm$, $B_d^0 \rightarrow J/\psi K^{*0}$, and $B_s^0 \rightarrow J/\psi \phi$, respectively. In the upper plots the proper decay time distribution is shown for the sidebands, which are indicated as grey (green) area in the mass distribution on the right. The lower plots show the distribution for the signal region, where the parameters describing the background are set to the values obtained from the fit to the sidebands. The table summarises the resulting fit parameters for each step. We can reproduce roughly the results from the simultaneous mass and decay time fit (tables 5.1, 5.4, and 5.7), with small differences. In the following we discuss their significance.

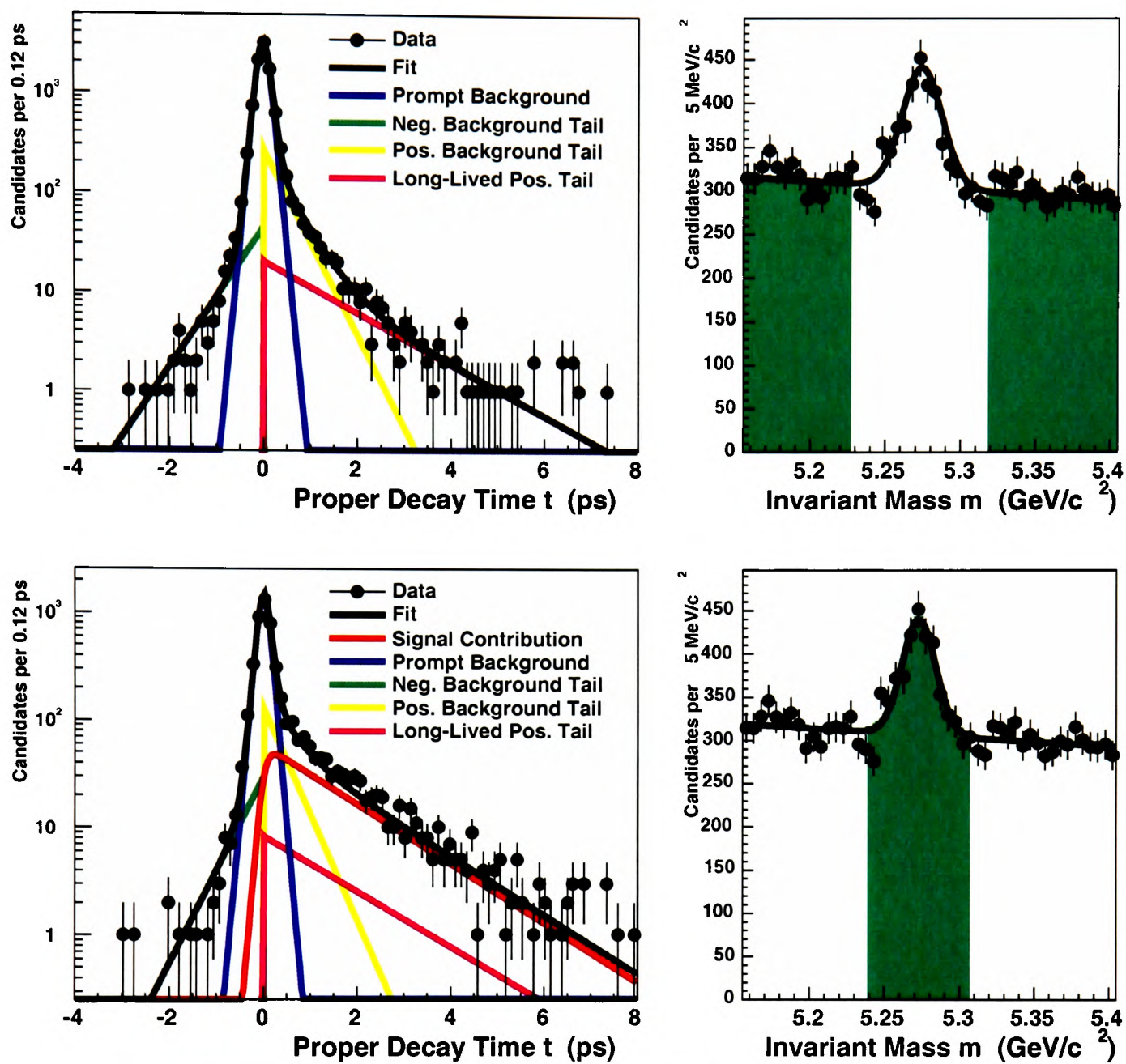
The difference between this cross-check fit and our nominal fit is, that the cross-check fit estimates the background only with the sideband events, which are roughly $\frac{2}{3}$ of the total background. The nominal fit on the other hand uses the non-sideband region events as well to determine the background.

For $B_u^\pm \rightarrow J/\psi K^\pm$ ($B_s^0 \rightarrow J/\psi \phi$) the cross-check result is about 0.3σ (-0.3σ) away from the nominal fit value, which we do not consider significant given that we are not using about a third of the events for background estimation. In the $B_d^0 \rightarrow J/\psi K^{*0}$ case we see a 1σ deviation. Although 1σ seems significant under the circumstance



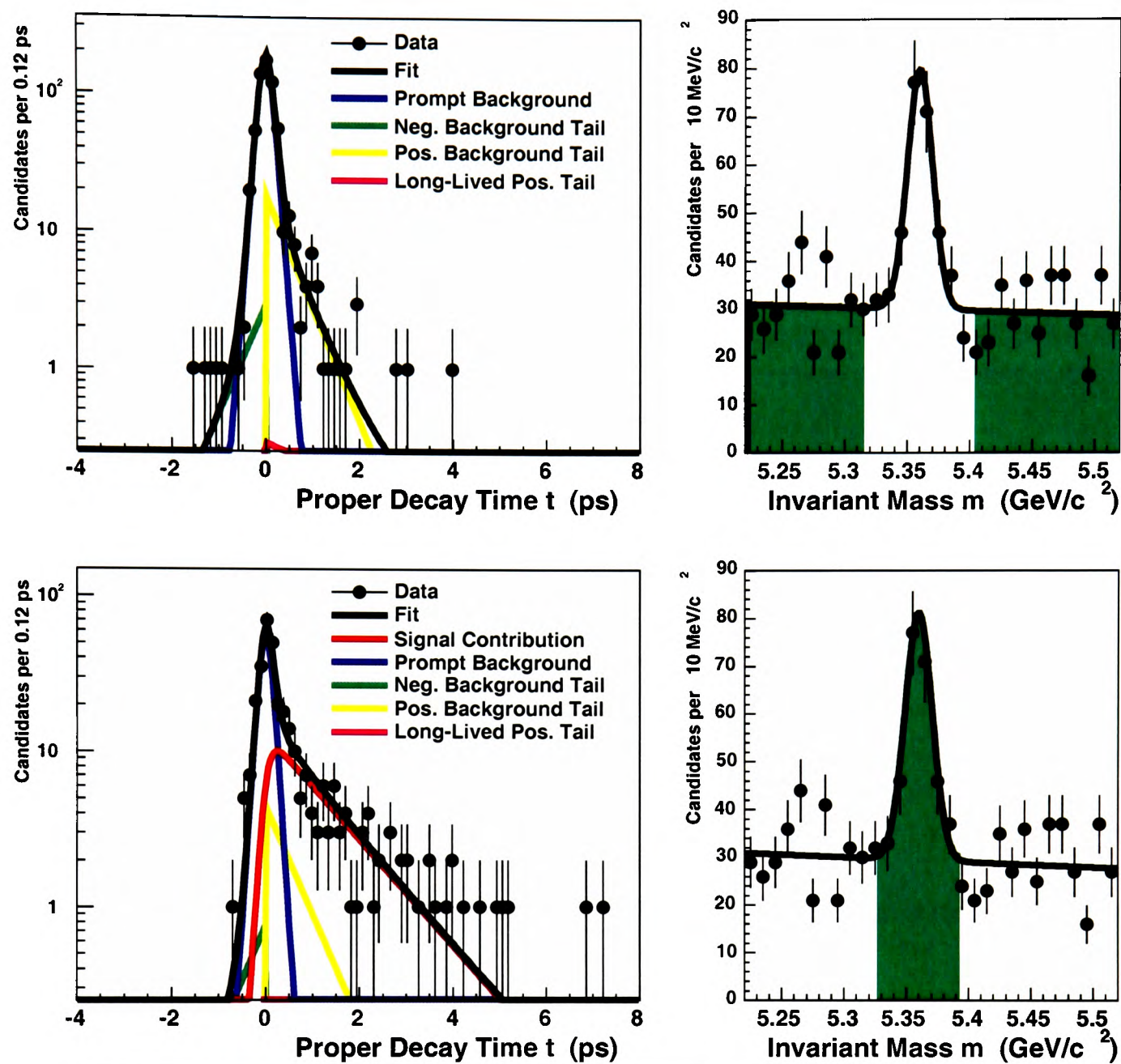
M	$(5.2747 \pm 0.0007) \text{ GeV}/c^2$		
W	$(14.45 \pm 0.77) \text{ MeV}/c^2$		
A	$(-0.65 \pm 0.59) c^2/\text{GeV}$		
f_{sig}	(0.142 ± 0.007)	N_{sig}	1495.9 ± 74.4
τ	$(1.633 \pm 0.047) \text{ ps}$	$c\tau$	$(489.7 \pm 14.0) \mu\text{m}$
S_t	1.36 ± 0.02		
f_-	0.015 ± 0.003	N_-	158.6 ± 27.5
f_+	0.107 ± 0.008	N_+	1124.6 ± 87.7
f_{++}	0.016 ± 0.006	N_{++}	164.5 ± 63.3
τ_-	$(1.508 \pm 0.217) \text{ ps}$	$c\tau_-$	$(452.0 \pm 65.2) \mu\text{m}$
τ_+	$(0.523 \pm 0.061) \text{ ps}$	$c\tau_+$	$(156.9 \pm 18.4) \mu\text{m}$
τ_{++}	$(2.590 \pm 0.592) \text{ ps}$	$c\tau_{++}$	$(776.6 \pm 177.4) \mu\text{m}$

Table 6.17: Result of the separate fit to mass distribution, sidebands and signal region for $B_u^\pm \rightarrow J/\psi K^\pm$ data. Compare with our nominal result of $\tau = 1.623$ ps.



M	$(5.2729 \pm 0.0012) \text{ GeV}/c^2$		
W	$(11.41 \pm 1.15) \text{ MeV}/c^2$		
A	$(-1.38 \pm 0.45) c^2/\text{GeV}$		
f_{sig}	(0.047 ± 0.005)	N_{sig}	762.7 ± 74.6
τ	$(1.571 \pm 0.064) \text{ ps}$	$c\tau$	$(471.1 \pm 19.3) \mu\text{m}$
S_t	1.31 ± 0.02		
f_-	0.029 ± 0.004	N_-	465.6 ± 60.8
f_+	0.113 ± 0.008	N_+	1808.3 ± 133.6
f_{++}	0.028 ± 0.007	N_{++}	444.2 ± 113.5
τ_-	$(0.503 \pm 0.047) \text{ ps}$	$c\tau_-$	$(150.8 \pm 14.1) \mu\text{m}$
τ_+	$(0.439 \pm 0.040) \text{ ps}$	$c\tau_+$	$(131.6 \pm 12.1) \mu\text{m}$
τ_{++}	$(1.684 \pm 0.226) \text{ ps}$	$c\tau_{++}$	$(504.9 \pm 67.8) \mu\text{m}$

Table 6.18: Result of the separate fit to mass distribution, sidebands and signal region for $B_d^0 \rightarrow J/\psi K^{*0}$ data. Compare with our nominal result of $\tau = 1.510$ ps.



M	$(5.3598 \pm 0.0016) \text{ GeV}/c^2$		
W	$(11.02 \pm 1.65) \text{ MeV}/c^2$		
A	$(-1.22 \pm 1.28) c^2/\text{GeV}$		
f_{sig}	(0.140 ± 0.018)	N_{sig}	142.9 ± 18.5
τ	$(1.268 \pm 0.124) \text{ ps}$	$c\tau$	$(380.2 \pm 37.0) \mu\text{m}$
S_t	1.46 ± 0.06		
f_-	0.020 ± 0.011	N_-	20.8 ± 11.7
f_+	0.123 ± 0.022	N_+	125.4 ± 22.1
f_{++}	0.006 ± 0.007	N_{++}	5.7 ± 7.4
τ_-	$(0.606 \pm 0.267) \text{ ps}$	$c\tau_-$	$(181.8 \pm 80.0) \mu\text{m}$
τ_+	$(0.625 \pm 0.127) \text{ ps}$	$c\tau_+$	$(187.3 \pm 38.0) \mu\text{m}$
τ_{++}	$(4.311 \pm 3.479) \text{ ps}$	$c\tau_{++}$	$(1292.5 \pm 1043.0) \mu\text{m}$

Table 6.19: Result of the separate fit to mass distribution, sidebands and signal region for $B_s^0 \rightarrow J/\psi \phi$ data. Compare with our nominal result of $\tau = 1.314 \text{ ps}$.

of largely correlated samples, we have to examine whether we want to attribute a systematic uncertainty due to this effect.

For $B_d^0 \rightarrow J/\psi K^{*0}$ we expect the wrong $K\pi$ assignments to slightly distort the background shape in the signal region, which then would affect the cross-check fit and the nominal fit in a different way. It is clear that the interplay between long-lived background and signal in the signal region influences the value of the fitted lifetime and could easily contribute the 1σ effect we see. On the other hand we believe that our nominal fit model estimates the background better than the cross-check model. Keeping in mind that we have already assigned a systematic uncertainty due to the $K\pi$ assignment ambiguity in section 6.8, we do not assign a systematic uncertainty due to the difference we see between the cross-check and our nominal fit result.

6.10 Simultaneous Fit to Sideband and Signal Regions

In order to further cross-check our fit, we perform a decay time fit simultaneously for the sideband and signal regions. This technique can be regarded as a modification of our main fit by replacing the continuous mass variable by a discrete variable taking three values: signal, sideband, or neither. This is of interest because possible correlations between measured mass and decay time have less impact on the fit, because sideband and signal regions are well-defined, and the event-by-event mass information is only used to classify events according to those regions.

As in section 6.9, we perform a binned likelihood fit to the mass distribution, with free parameters M , W , A and f_{sig} , to be able to define the signal region as $[M - 3W, M + 3W]$ and the sidebands as $m < M - 4W$ and $m > M + 4W$.

We then fit to proper decay time simultaneously for the sideband and signal regions. We use a modification of equation 5.2 to define the probability function per event:

$$P(m, t, \sigma_t) = \theta(\text{sideband}) \cdot \frac{f_{sig}}{6W} \cdot F_{sig}(t, \sigma_t) + \theta(\text{signal or sideband}) \cdot \frac{1 - f_{sig}}{m_{max} - m_{min}} \cdot F_{bk}(t, \sigma_t), \quad (6.10)$$

where θ is 1 if the event belongs to the region specified in its argument, and 0 otherwise.

Tables 6.20, 6.21, and 6.22 display the results of this fit applied to the three decay modes. They differ only marginally from our main results (tables 5.1, 5.4, and 5.7), which leads us to declare the cross-check successful.

6.11 Cross-check: Stability Over Time

In order to check the time stability of our results we have divided the data into two approximately equal sized parts: a set of early data and a set of later data. Each

M	$(5.2747 \pm 0.0007) \text{ GeV}/c^2$		
W	$(14.45 \pm 0.77) \text{ MeV}/c^2$		
A	$(-0.65 \pm 0.59) c^2/\text{GeV}$		
f_{sig}	0.144 ± 0.004	N_{sig}	1518.6 ± 47.0
τ	$(1.613 \pm 0.050) \text{ ps}$	$c\tau$	$(483.5 \pm 14.9) \mu\text{m}$
S_t	1.35 ± 0.02		
f_-	0.017 ± 0.002	N_-	181.8 ± 23.5
f_+	0.103 ± 0.009	N_+	1084.0 ± 90.0
f_{++}	0.030 ± 0.008	N_{++}	313.4 ± 83.3
τ_-	$(1.047 \pm 0.116) \text{ ps}$	$c\tau_-$	$(314.0 \pm 34.9) \mu\text{m}$
τ_+	$(0.466 \pm 0.055) \text{ ps}$	$c\tau_+$	$(139.8 \pm 16.4) \mu\text{m}$
τ_{++}	$(1.968 \pm 0.286) \text{ ps}$	$c\tau_{++}$	$(589.9 \pm 85.7) \mu\text{m}$

Table 6.20: Result of the simultaneous lifetime fit in signal and sideband regions for $B_u^\pm \rightarrow J/\psi K^\pm$ data. Compare with our nominal result of $\tau = 1.623 \text{ ps}$.

M	$(5.2729 \pm 0.0012) \text{ GeV}/c^2$		
W	$(11.41 \pm 1.15) \text{ MeV}/c^2$		
A	$(-1.38 \pm 0.45) c^2/\text{GeV}$		
f_{sig}	0.056 ± 0.003	N_{sig}	900.3 ± 42.6
τ	$(1.537 \pm 0.067) \text{ ps}$	$c\tau$	$(460.9 \pm 20.0) \mu\text{m}$
S_t	1.32 ± 0.01		
f_-	0.024 ± 0.003	N_-	378.5 ± 41.2
f_+	0.109 ± 0.008	N_+	1750.6 ± 123.0
f_{++}	0.031 ± 0.008	N_{++}	496.3 ± 123.5
τ_-	$(0.614 \pm 0.048) \text{ ps}$	$c\tau_-$	$(184.1 \pm 14.4) \mu\text{m}$
τ_+	$(0.454 \pm 0.041) \text{ ps}$	$c\tau_+$	$(136.0 \pm 12.2) \mu\text{m}$
τ_{++}	$(1.604 \pm 0.196) \text{ ps}$	$c\tau_{++}$	$(481.0 \pm 58.8) \mu\text{m}$

Table 6.21: Result of the simultaneous lifetime fit in signal and sideband regions for $B_d^0 \rightarrow J/\psi K^{*0}$ data. Compare with our nominal result of $\tau = 1.510 \text{ ps}$.

M	$(5.3598 \pm 0.0016) \text{ GeV}/c^2$		
W	$(11.02 \pm 1.65) \text{ MeV}/c^2$		
A	$(-1.22 \pm 1.28) c^2/\text{GeV}$		
f_{sig}	0.122 ± 0.013	N_{sig}	124.3 ± 13.5
τ	$(1.311 \pm 0.143) \text{ ps}$	$c\tau$	$(392.9 \pm 43.0) \mu\text{m}$
S_t	1.45 ± 0.05		
f_-	0.021 ± 0.011	N_-	21.4 ± 11.3
f_+	0.128 ± 0.022	N_+	131.0 ± 22.9
f_{++}	0.011 ± 0.013	N_{++}	10.8 ± 13.3
τ_-	$(0.550 \pm 0.218) \text{ ps}$	$c\tau_-$	$(165.0 \pm 65.3) \mu\text{m}$
τ_+	$(0.544 \pm 0.136) \text{ ps}$	$c\tau_+$	$(163.0 \pm 40.8) \mu\text{m}$
τ_{++}	$(2.873 \pm 1.866) \text{ ps}$	$c\tau_{++}$	$(861.3 \pm 559.4) \mu\text{m}$

Table 6.22: Result of the simultaneous lifetime fit in signal and sideband regions for $B_s^0 \rightarrow J/\psi \phi$ data. Compare with our nominal result of $\tau = 1.314 \text{ ps}$.

dataset corresponds to an integrated luminosity of about 67 pb^{-1} . Since the data was taken during a period when the luminosity was gradually increasing, the early dataset contains many runs with on average smaller luminosity, and the later one with higher luminosity. Figure 6.11 compares the fitted lifetime for the two datasets, for all three decay modes. The results for the first half of the data are statistically compatible with the second half.

6.12 Summary

Systematic Effect	$B_u^\pm \rightarrow J/\psi K^\pm$	$B_d^0 \rightarrow J/\psi K^{*0}$	$B_s^0 \rightarrow J/\psi \phi$
Alignment	± 0.017	\leftarrow same	\leftarrow same
Beam Spot Size	negligible	\leftarrow same	\leftarrow same
Resolution Function	± 0.010	\leftarrow same	\leftarrow same
Background Model	negligible	\leftarrow same	\leftarrow same
Fitting Procedure	negligible	\leftarrow same	\leftarrow same
Event Selection	negligible	\leftarrow same	\leftarrow same
$K\pi$ assignment	n/a	± 0.013	n/a
Total	± 0.020	± 0.024	± 0.020

Table 6.23: Summary of systematic uncertainties on the measured lifetime τ (ps).

Table 6.23 summarises the studied systematic effects and their resulting uncertainty on the measured lifetime τ . The biggest effects come from residual silicon misalignments and from the choice of resolution function. In the decay $B_d^0 \rightarrow J/\psi K^{*0}$

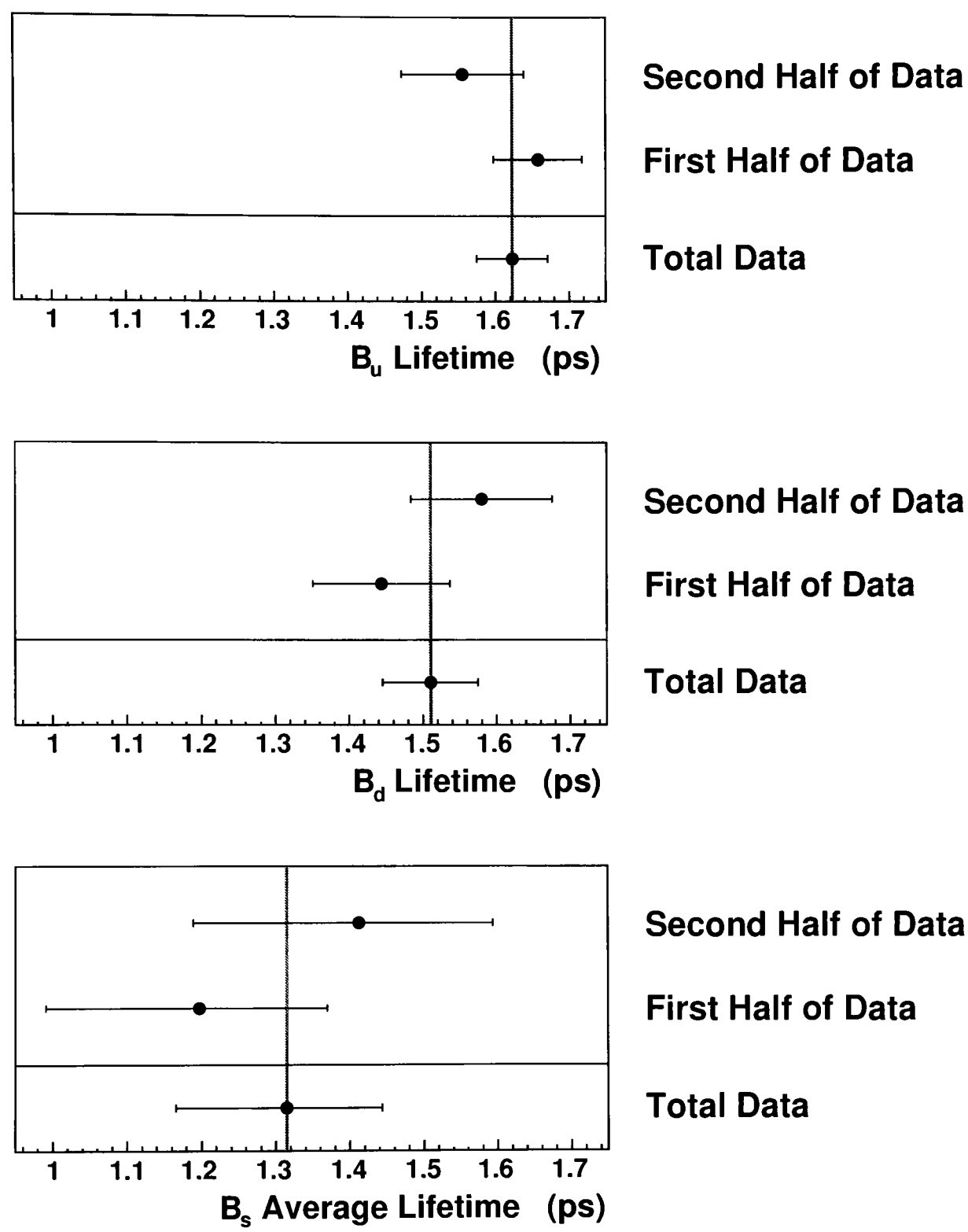


Figure 6.11: Fitted lifetime of B_u , B_d and B_s (from top to bottom) for two different parts of the data.

also the $K\pi$ ambiguity contributes a big systematic uncertainty. The last line of the table adds up all the contributions in quadrature to the total systematic uncertainty.

Chapter 7

Summary and Conclusions

We have performed a measurement of the average B_s lifetime using the exclusive decay mode $B_s^0 \rightarrow J/\psi \phi$. Our result is

$$\tau(B_s) = (1.31_{-0.13}^{+0.15}(\text{stat.}) \pm 0.02(\text{syst.})) \text{ ps} \quad (7.1)$$

As cross-checks we measured the lifetime of similar, but higher statistics B meson decay modes, such as $B_u^\pm \rightarrow J/\psi K^\pm$ and $B_d^0 \rightarrow J/\psi K^{*0}$. We get

$$\begin{aligned} \tau(B_u) &= (1.623 \pm 0.048(\text{stat.}) \pm 0.020(\text{syst.})) \text{ ps} \\ \tau(B_d) &= (1.510 \pm 0.065(\text{stat.}) \pm 0.024(\text{syst.})) \text{ ps} \end{aligned} \quad (7.2)$$

We note that all three measurements are dominated by their statistical error. Due to the small amount of data most of the systematic studies we have performed could not reveal significant systematic effects. This might change in future measurements with more data. The statistical error will decrease and the systematic uncertainties will become more important. In addition, the same systematic studies might then point out significant systematic effects.

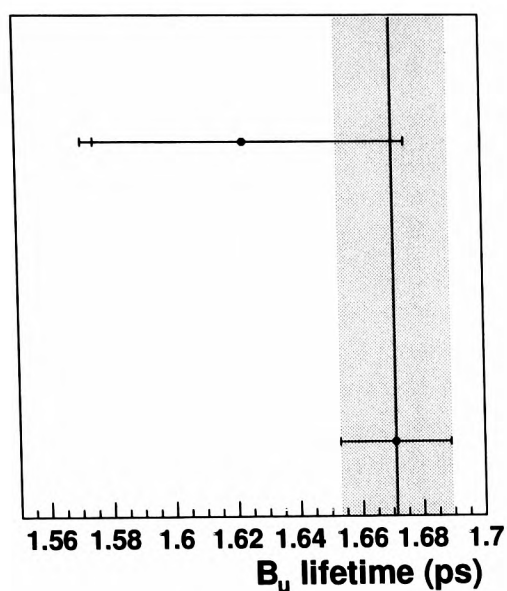
We can calculate the lifetime ratio $\tau(B_u)/\tau(B_d)$, where some of the common systematic uncertainties cancel. Our measurements result in

$$\tau(B_u)/\tau(B_d) = 1.075 \pm 0.056(\text{stat.}) \pm 0.017(\text{syst.}) , \quad (7.3)$$

where we propagated the uncertainties taking into account correlations. Since the systematic errors are largely correlated, we split the systematic uncertainty on $\tau(B_d)$ into a part common to both lifetimes (0.020 ps), and a part unique to $\tau(B_d)$ (0.013 ps). The result is consistent with the current theory prediction of $\tau(B_u)/\tau(B_d) = 1.05 \pm 0.02$ (see equation 1.17). A comparison with the theory prediction of the lifetime ratio $\tau(B_s)/\tau(B_d)$ is not meaningful, as our measured average B_s lifetime $\tau_{\text{average}} = f_{CP+} \cdot \tau_{CP+} + (1 - f_{CP+}) \cdot \tau_{CP-}$ is in general not the same as the mean lifetime $\tau = \hbar/\Gamma_s$.

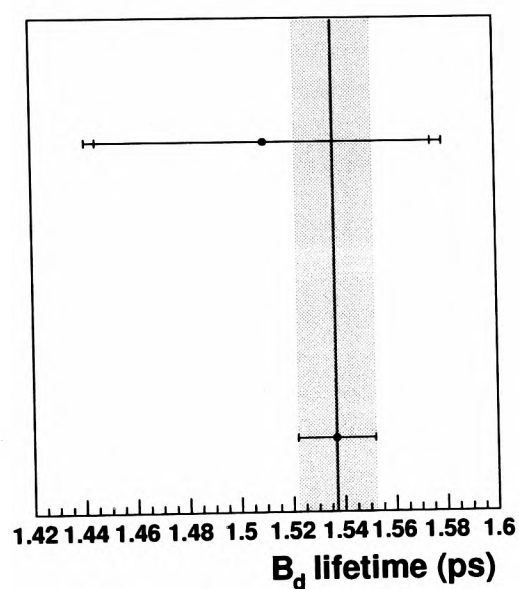
Figure 7.1 puts the lifetime measurements into perspective with the current world average values [12], and in the B_s case, with a former measurement of CDF [51] using the same decay mode. There is good agreement between all measurements within

their uncertainties. It has been noted that all lifetimes of this analysis come out somewhat smaller than the current world averages. With the current small datasets we cannot make any conclusions whether this is a systematic effect or pure accident. We have studied, to our knowledge, all sources of systematic effects and believe that our assigned systematic errors provide the best estimate on the uncertainties to possible systematic effects. In the B_s case the small measured lifetime could be an indication that the decay $B_s^0 \rightarrow J/\psi \phi$ has a large CP even component and therefore predominantly decays via its B_s^{Short} component. An angular analysis similar to [43] could be performed to support this hypothesis.



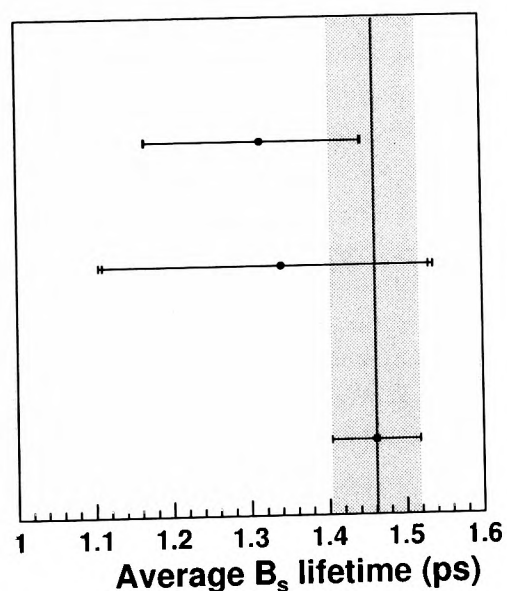
This Analysis

PDG [12], World Average



This Analysis

PDG [12], World Average



This Analysis

CDF [51], $B_s^0 \rightarrow J/\psi \phi$

PDG [12], World Average

Figure 7.1: Comparison between the lifetimes of B_u , B_d and B_s (from top to bottom) measured in this analysis and the current world average values [12]. In addition, for $B_s^0 \rightarrow J/\psi \phi$ a former CDF measurement on run 1 data [51] using the same decay mode is shown. The error bars are statistical (inner error bar) and statistical \oplus systematic (total error bar) added in quadrature.

Appendix A

Charged Current Interactions in the Standard Model

In the Standard Model flavour changing stems from charged current interactions in the quark and lepton sectors, involving the charged $SU(2)_L$ gauge bosons W^\pm [1]. In the mass basis the relevant parts of the Lagrangian have the form

$$-\mathcal{L}_{W^\pm}^q = \frac{g}{\sqrt{2}} \bar{u}_{Li} \gamma^\mu V_{ij}^{CKM} d_{Lj} W_\mu^+ + \text{h.c.} , \quad (\text{A.1})$$

$$-\mathcal{L}_{W^\pm}^l = \frac{g}{\sqrt{2}} \bar{\nu}_{Li} \gamma^\mu e_{Li} W_\mu^+ + \text{h.c.} , \quad (\text{A.2})$$

with g as the weak coupling, and u_{Li} , d_{Li} representing up- and down-type quarks, respectively, of generation $i = 1, 2, 3$, and e_{Li} , ν_{Li} electrons and neutrinos of generation i . The Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix [86] is a unitary 3×3 matrix with 4 independent physical parameters: three real angles and one complex phase. This phase allows CP violation in the Standard Model.

The experimentally found hierarchy of CKM matrix elements – diagonal elements being close to one and small off-diagonal elements – is reflected in the popular **Wolfenstein Parameterisation** [13], which uses 4 parameters A , λ , ρ , η and approximates V_{CKM} up to $\mathcal{O}(\lambda^3)$:

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

The current experimental values can be found in [14]. The parameter λ is the cosine of the Cabibbo mixing angle and is found to be $\lambda = \cos \theta_C = 0.2196 \pm 0.0026$.

Appendix B

Optimisation of Transverse Momentum Selection Cuts

We describe in this section the procedure to find the optimal minimal values for the total B meson transverse momentum, as well as that of the non-J/ ψ component. These cuts are optimised in an unbiased way to get the smallest statistical error on the background subtracted signals. The B candidates which are most important in determining the lifetime are those with a value of proper decay time t larger than our ct resolution (about 60 μm). We therefore choose the p_T cuts in a way that the error on the signals with $ct > 60 \mu\text{m}$ is minimised, or equivalently, the quantity $\mathcal{S}^2/(\mathcal{S} + \mathcal{B})$ is maximised. In this context \mathcal{S} (\mathcal{B}) denotes the number of signal (background) candidates. The motivation for maximising $\mathcal{S}^2/(\mathcal{S} + \mathcal{B})$ is, that the square-root of this quantity has a statistical meaning: it is the number of signal events (\mathcal{S}) compared to a 1σ fluctuation of the whole sample ($\sqrt{\mathcal{S} + \mathcal{B}}$).

In the optimisation procedure we use for the signal a simulated sample, which describes the data reasonably well. This ensures an unbiased selection. For the background it is preferred to use real data, since no simulated background sample exists that would describe the data sufficiently well. We consider $p_T(\text{B}) > 4 \text{ GeV}/c$ and $p_T(\text{non-J}/\psi) > 1 \text{ GeV}/c$ as reasonable starting point (loosest cuts) of the optimisation procedure. After requiring a proper lifetime of $ct > 60 \mu\text{m}$ we are left with a very small number of background events in the B_s case. For this reason we optimise the cuts for the $\text{B}_u^\pm \rightarrow \text{J}/\psi \text{ K}^\pm$ decay mode instead, and apply the same cuts to the $\text{B}_s^0 \rightarrow \text{J}/\psi \phi$ and $\text{B}_d^0 \rightarrow \text{J}/\psi \text{ K}^{*0}$ modes. Bearing in mind that the kinematics of the B meson and the non-J/ ψ component is similar for all three decay modes, and that the choice of cuts is unbiased, this seems like a reasonable approach. In addition using the same cuts for all three decay modes facilitates comparisons.

As signal region we define a window of $\pm 40 \text{ MeV}/c^2$ (approximately 3σ of the B peak observed in data) around the world average B^+ mass of $5,279.1 \text{ MeV}/c^2$. The number of signal events in this region, \mathcal{S} , as a function of p_T cuts is obtained from a simulated signal sample. The number of background events in the signal region, \mathcal{B} , is obtained from data by integrating the fitted background distribution over the region in question.

The signal sample is generated with the software **Bgenerator** [87], which generates single B^+ mesons according to the b quark transverse momentum - rapidity distribution calculated in [88], and a subsequent fragmentation according to Peterson fragmentation [89].¹

The B^+ mesons are subsequently forced to decay into $J/\psi K^+$, $J/\psi \rightarrow \mu^+\mu^-$. The resulting events are put through a parametric trigger simulation [90] and the full GEANT detector simulation [91], [92]. Obviously our procedure relies on the fact that the simulation describes the data well. Figure B.1 shows on the left the invariant mass of B candidates with $p_T(B) > 4 \text{ GeV}/c$ and $p_T(\text{non-}J/\psi) > 1 \text{ GeV}/c$, for data and simulated signal candidates. The simulated sample is normalised to the

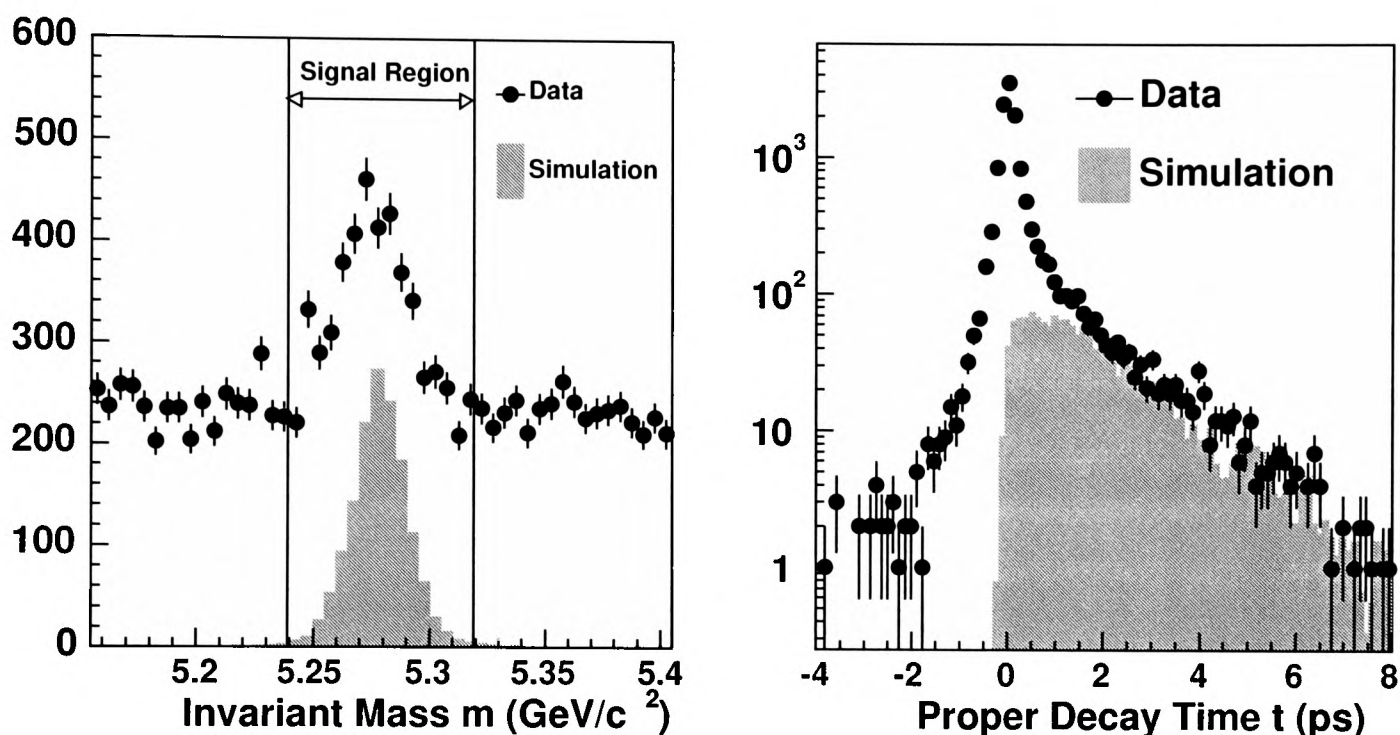


Figure B.1: Data-Simulation comparison of invariant mass (left) and proper decay time (right) for $B_u^\pm \rightarrow J/\psi K^\pm$ candidates.

number of signal events in data as extracted from a fit (single Gaussian + linear background). The simulated peak is slightly shifted by about $4 \text{ MeV}/c^2$ and has a width ($12 \text{ MeV}/c^2$) much narrower than the peak in data ($15 \text{ MeV}/c^2$). But since the signal region covers nearly the whole peak these differences should not influence our optimisation procedure. On the right we see the proper lifetime distribution for the same candidates. For long positive lifetimes we see the expected agreement between data and simulation. For shorter lifetimes we see the effect of background in the data, which does not exist in the simulated signal sample.

Figure B.2 proves that indeed the quantities of merit agree reasonably well between data and simulation. Transverse momenta $p_T(B)$ and $p_T(\text{non-}J/\psi)$ as well as proper lifetime t are shown, for sideband subtracted data and signal simulation. The sideband subtraction procedure works the following way: The sidebands are chosen to be the

¹The Peterson parameter $\epsilon_B = 0.006$ is used.

	1.00	1.25	1.50	1.75	2.00	2.25	2.50
4.0	677 ± 10	838 ± 13	950 ± 15	1019 ± 16	1051 ± 17	1033 ± 17	1002 ± 17
4.5	729 ± 11	882 ± 14	984 ± 16	1043 ± 17	1068 ± 17	1046 ± 17	1008 ± 17
5.0	775 ± 12	919 ± 15	1015 ± 16	1063 ± 17	1079 ± 18	1051 ± 17	1015 ± 17
5.5	820 ± 13	957 ± 16	1041 ± 17	1079 ± 18	1087 ± 18	1055 ± 17	1016 ± 17
6.0	847 ± 14	969 ± 16	1046 ± 17	1075 ± 18	1079 ± 18	1048 ± 17	1011 ± 17
6.5	865 ± 15	970 ± 16	1036 ± 17	1058 ± 18	1063 ± 17	1030 ± 17	993 ± 16
7.0	875 ± 15	967 ± 17	1020 ± 17	1032 ± 17	1029 ± 17	994 ± 16	957 ± 16
7.5	860 ± 16	938 ± 17	986 ± 17	987 ± 17	983 ± 16	947 ± 16	910 ± 15
8.0	835 ± 15	901 ± 16	934 ± 16	931 ± 16	931 ± 16	896 ± 15	870 ± 14

Table B.1: The quantity $\mathcal{S}^2/(\mathcal{S} + \mathcal{B})$ is optimised in the two cut variables: minimum p_T cut for the non-J/ ψ in horizontal direction (1.0 – 2.5 GeV/ c) and for the B in vertical direction (4.0 – 8.0 GeV/ c). The maximum value is highlighted in bold face.

interval 5,155 MeV/ c^2 to 5,405 MeV/ c^2 without a ± 60 MeV/ c^2 window around the world average B^+ mass of 5,279.1 MeV/ c^2 . We subtract from the $p_T(B)$ distribution of the signal region the $p_T(B)$ distribution of the sideband region, which we normalise to the number of background events expected in the signal region.² This gives the $p_T(B)$ distribution as expected for the signal only. We repeat the procedure for the $p_T(K)$ and proper decay time distribution.

Table B.1 shows the quantity $\mathcal{S}^2/(\mathcal{S} + \mathcal{B})$ for different transverse momentum cuts on the non-J/ ψ (horizontal direction, 1.0–2.5 GeV/ c) and on the B (vertical direction, 4.0–8.0 GeV/ c). The optimal value is high-lighted and corresponds to $p_T(\text{non-J}/\psi) > 2$ GeV/ c and $p_T(B) > 5.5$ GeV/ c . We observe negligible, non-significant variations on those values when varying the L_{xy} cut between nought and 100 μm . Similar cut optimisations have been performed at earlier stages of the analysis ([93], [94]), with less data for background and less accurate simulation. They resulted in higher central values for the $p_T(B)$ cut: 6.5 GeV/ c for [93], and 7.2 GeV/ c for [94], although statistically consistent with our result. The CDF collaboration internally has agreed on the values $p_T(\text{non-J}/\psi) > 2$ GeV/ c and $p_T(B) > 6.5$ GeV/ c based on those studies, which we will use for this analysis from now on.

²We get the expected number of background events in the signal region from a fit of a Gaussian plus first order polynomial to the mass distribution. Integrating the first order polynomial over the signal range gives the number of expected background events in the signal region.

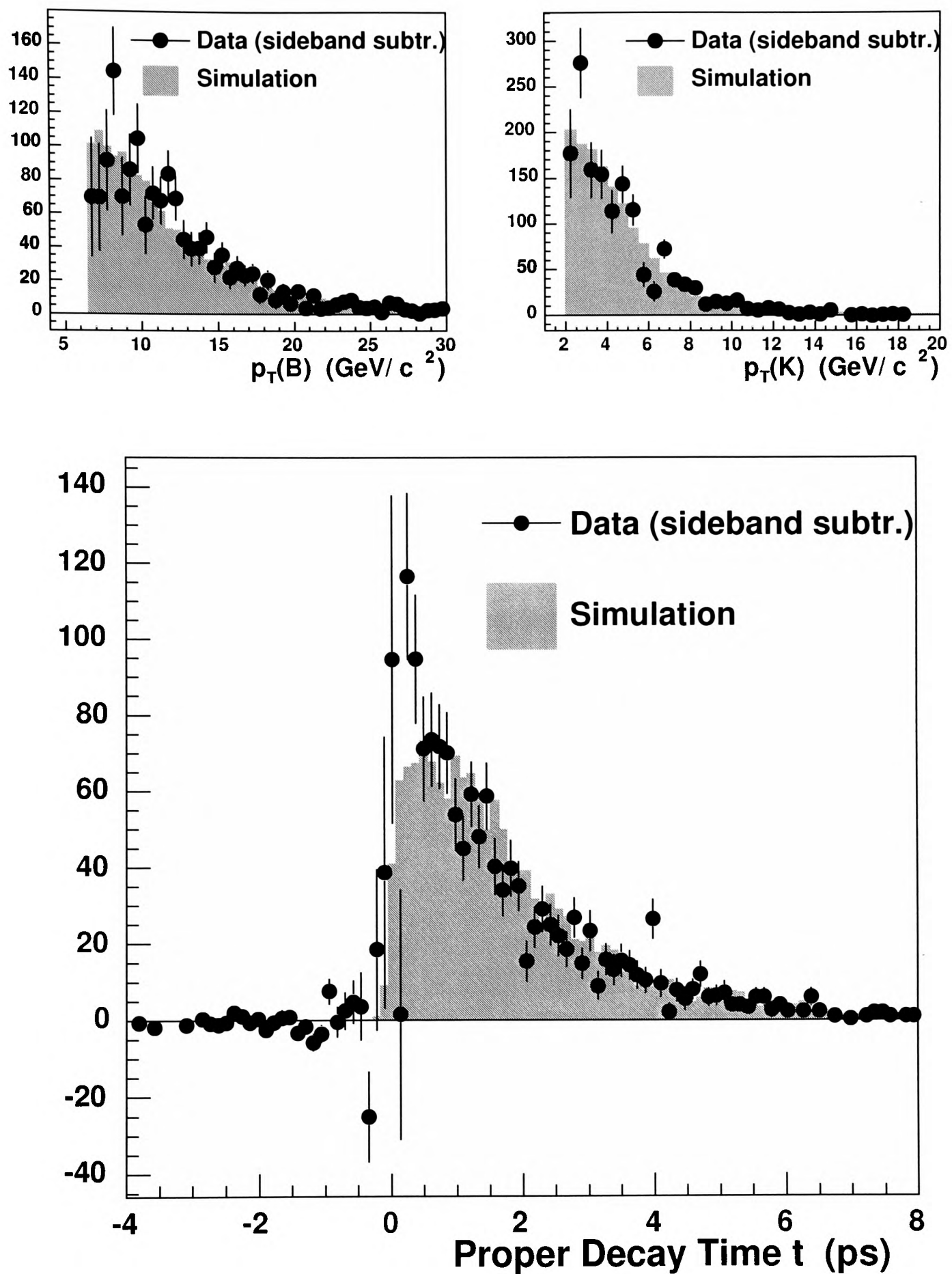


Figure B.2: Comparison between sideband subtracted data and signal simulation for the transverse momenta $p_T(B)$, $p_T(\text{non-}J/\psi)$, and the proper lifetime t .

Appendix C

Vertex Fit

In order to find 3-dimensional vertices for a given set of tracks and vertex topology, we perform a vertex fit with the ability of applying mass and pointing constraints [95]. For our analysis we do not use the pointing constraint option though. In the vertex fit the track parameters are varied in order to find the best set compatible with the given vertex topology. This procedure usually improves the mass resolution of the particle reconstruction.

In the case of a fit of n tracks to a common vertex \mathbf{x}_v the track parameters C_i , ϕ_{0i} , λ_i ($i = 1..n$) and the vertex coordinates \mathbf{x}_v are taken as free parameters. The common vertex requirement then determines already the values for the track parameters d_{0i} and z_{0i} . A χ^2 is formed of the residuals $\boldsymbol{\xi}$ between the measured and fitted track parameters. Each track i gives a contribution

$$\chi_i^2 = \boldsymbol{\xi}_i^\top \mathbf{G}_i^{-1} \boldsymbol{\xi}_i \quad \text{with} \quad \boldsymbol{\xi}_i = \begin{pmatrix} \Delta C_i \\ \Delta \phi_{0i} \\ \Delta \lambda_{0i} \\ \Delta d_{0i} \\ \Delta z_{0i} \end{pmatrix}, \quad (\text{C.1})$$

\mathbf{G}_i^{-1} being the track error matrix and d_{0i} , z_{0i} depending fully on C_i , ϕ_{0i} and λ_{0i} . The error matrix \mathbf{G}_i^{-1} should in principle be taken at the vertex \mathbf{x}_v , to incorporate multiple scattering correctly. For higher energetic tracks and for short decay distances, where the decay takes place inside the beam pipe, this effect is negligible and the error matrix can be taken at the point of closest approach to the origin.

The number of degrees of freedom n_{dof} for such a fit is the number of measured track parameters minus the number of free parameters. The number of measured track parameters is 5 for each track, and the number of free parameters is 3 per track and 3 per vertex. With n_{tracks} and $n_{vertices}$ being the number of tracks and vertices, we get:

$$n_{dof} = 5n_{tracks} - (3n_{tracks} + 3n_{vertices}) = 2n_{tracks} - 3n_{vertices}. \quad (\text{C.2})$$

In the case where several tracks originate from an intermediate resonance whose width is much smaller than its experimental mass resolution, a *mass constraint* helps

to improve the mass resolution. Mass constraints are included in the fit with Lagrange multipliers giving a χ^2 contribution

$$\chi_{\text{mass}}^2 = \mu \cdot (M^2 - M_0^2) , \quad (\text{C.3})$$

where μ is the Lagrange multiplier, M_0 is the exact mass of the resonance, and M is the invariant mass of the vector sum of all tracks originating from this resonance at the vertex \mathbf{x}_v :

$$M^2 = \left(\sum_i E_i\right)^2 - \left(\sum_i p_{xi}\right)^2 - \left(\sum_i p_{yi}\right)^2 - \left(\sum_i p_{zi}\right)^2 \quad (\text{C.4})$$

The energy E_i is calculated from the momentum of the track under a certain particle hypothesis. Since a mass constraint constrains the momenta on a vertex to a certain mass, it makes the free parameters dependent on each other and effectively decreases the number of free parameters by one. The number of degrees of freedom therefore goes up by one:

$$n_{\text{dof}} = 2n_{\text{tracks}} - 3n_{\text{vertices}} + 1n_{\text{mass}} , \quad (\text{C.5})$$

where n_{mass} is the number of mass constraints in the fit.

In our analysis we perform vertexing effectively only in the $r\phi$ plane. A track i in the $r\phi$ plane has 3 parameters: C_i , ϕ_{0i} and d_{0i} . The common vertex requirement then already determines the value for d_{0i} , so the number of free track parameters drops to 2. Therefore we get for the number of degrees of freedom for a 2-dimensional fit:

$$n_{\text{dof}} = n_{\text{tracks}} - 2n_{\text{vertices}} + 1n_{\text{mass}} . \quad (\text{C.6})$$

The χ^2 can be used to estimate the goodness of the fit. The probability P that the hypothesis (i. e. a true physical vertex) would lead to a χ^2 value greater than χ_0^2 is called **χ^2 probability** and is defined by

$$P(\chi^2) = \int_{\chi_0^2}^{\infty} dz f(z; n_{\text{dof}}) \quad (\text{C.7})$$

where $f(z; n_{\text{dof}})$ is the χ^2 distribution with n_{dof} degrees of freedom. Requiring χ^2 to be smaller than χ_0^2 corresponding to a fixed value of P therefore leads to an efficiency of vertex finding of $1 - P$. Thus small values (typical 0.01) of P are used to keep the efficiency high. Table C.1 shows values of χ^2 corresponding to certain values of P and n_{dof} .

n_{dof}	$P = 0.01$	$P = 0.005$	$P = 0.001$
1	6.63	7.88	10.83
2	9.21	10.60	13.82
3	11.34	12.84	16.27
4	13.28	14.86	18.47
5	15.09	16.75	20.52
6	16.81	18.55	22.46
7	18.47	20.28	24.32
8	20.09	21.95	26.12
9	21.67	23.59	27.88

Table C.1: Values of χ^2 corresponding to different values of P and n_{dof} .

Appendix D

Study of a $\Delta\Gamma/\Gamma$ Transversity Analysis

In this appendix we discuss the basic method of extracting the width difference with an analysis of the transversity angle θ_T . We describe the used fitting model, and study the performance of the fit under pseudo-real conditions with many generated pseudo-experiments. Unfortunately the outcome is that our current amount of data is not sufficient for a reliable transversity analysis.

The probability density function is tri-variate, by using per event proper decay time t_i , invariant mass m_i and transversity angle $\cos\theta_{T,i}$. The probability function P for each candidate is split into two signal parts describing the CP even and odd eigenstates of the B_s and a part describing the background:

$$\begin{aligned}
 P(m_i, t_i, \sigma_{t,i}, \cos\theta_{T,i}; \mathbf{A}) &= f_{sig} \cdot g_{sig}(m_i) \\
 &\cdot \left\{ f_{CP+} \cdot \frac{3}{8}(1 + \cos^2\theta_{T,i}) \cdot F_{CP+}(t_i, \sigma_{t,i}) \right. \\
 &\quad \left. + (1 - f_{CP+}) \cdot \frac{3}{4}(1 - \cos^2\theta_{T,i}) \cdot F_{CP-}(t_i, \sigma_{t,i}) \right\} \\
 &+ (1 - f_{sig}) \cdot \frac{1}{2} \cdot g_{bk}(m_i) \cdot F_{bk}(t_i, \sigma_{t,i}) .
 \end{aligned} \tag{D.1}$$

This is exactly the same probability function as in equation 5.2, only that the lifetime part of the signal is split into a CP-even part and a CP-odd part with different lifetimes (τ_{CP+} and τ_{CP-}) and transversity angle distribution. The parameter f_{CP+} denotes the CP-even fraction of the signal, the mass parts g_{sig} , g_{bk} are defined in 5.8 and the lifetime parts in 5.5, 5.4. In the following we drop for simplicity the long-lived positive background part (described by f_{++} and τ_{++}). The $\cos\theta_T$ distribution for the background is assumed to be flat, which explains the normalisation factor $\frac{1}{2}$ in the last term. For real data this of course cannot be assumed unless explicitly proven. In addition efficiency and acceptance corrections need to be applied to the transversity angle distribution.

Instead of fitting for the two lifetimes τ_{CP+} and τ_{CP-} we use an equivalent set of parameters: the flavour specific lifetime τ_f (see chapter 1.10) and the relative width

difference $\Delta\Gamma/\Gamma$. The transformation and its inverse reads:

$$\tau_f = \frac{\tau_{CP+}^2 + \tau_{CP-}^2}{\tau_{CP+} + \tau_{CP-}}, \quad (\text{D.2})$$

$$\frac{\Delta\Gamma}{\Gamma} = 2 \cdot \frac{\tau_{CP-} - \tau_{CP+}}{\tau_{CP-} + \tau_{CP+}}, \quad (\text{D.3})$$

and

$$\tau_{CP+} = \frac{1 - \frac{\Delta\Gamma}{2\Gamma}}{1 + \left(\frac{\Delta\Gamma}{2\Gamma}\right)^2} \cdot \tau_f, \quad (\text{D.4})$$

$$\tau_{CP-} = \frac{1 + \frac{\Delta\Gamma}{2\Gamma}}{1 + \left(\frac{\Delta\Gamma}{2\Gamma}\right)^2} \cdot \tau_f. \quad (\text{D.5})$$

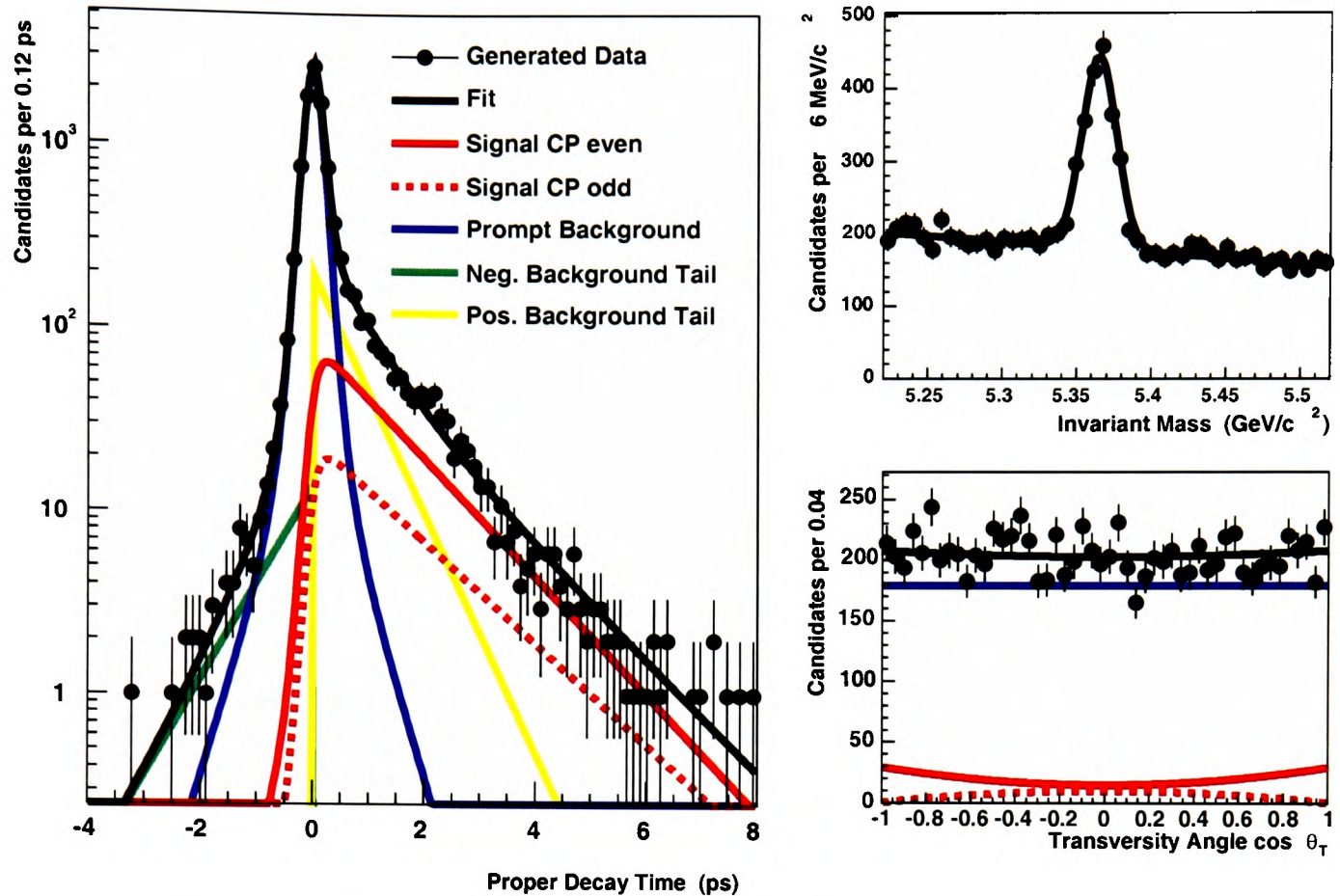
The resulting 12 parameters are summarised in table D.

1	M	Mass	GeV/c^2
2	W	Width	MeV/c^2
3	A	Background slope	c^2/GeV
4	f_{sig}	Signal fraction	–
5	f_{CP+}	Signal CP even fraction	–
6	τ_f	Signal flavour-specific lifetime	ps
7	$\Delta\Gamma/\Gamma$	Signal relative width difference	ps
8	S_t	t error scale factor	–
9	f_-	Background fraction of negative tail	–
10	f_+	Background fraction of positive tail	–
11	τ_-	Lifetime of negative background tail	ps
12	τ_+	Lifetime of positive background tail	ps

Table D.1: Summary of the 12 parameters used in the combined mass, transversity angle and lifetime fit.

Figure D.1 gives a graphical illustration. We have generated 10,230 events according to the fitting model, with parameters set to the values obtained from the average lifetime fit to the $B_s^0 \rightarrow J/\psi \phi$ data. In addition we use $f_{CP+} = 79\%$, $\tau_f = 1.461$ and $\Delta\Gamma/\Gamma = 15\%$. The fit result is shown in the table, and the three histograms display the generated data for t , m , and $\cos\theta_T$, overlaid with the fit result and the different contributions.

The correlation coefficients for the same generated data can be found in table D.2. We can classify the parameters into mass parameters (W, A, f_{sig}), signal parameters ($f_{CP+}, \tau_f, \Delta\Gamma/\Gamma$) and background parameters ($S_t, f_-, f_+, \tau_-, \tau_+$). From the table we see that the fit decouples these blocks well from one another. In the signal block there are large positive correlations between f_{CP+}, τ_f and $\Delta\Gamma/\Gamma$:



M	$(5.3648 \pm 0.0004) \text{ GeV}/c^2$		
W	$(11.62 \pm 0.34) \text{ MeV}/c^2$		
A	$(-2.86 \pm 0.40) c^2/\text{GeV}$		
f_{sig}	0.126 ± 0.004	N_{sig}	1289.8 ± 41.7
f_{CP+}	0.751 ± 0.049	N_{CP+}	968.9 ± 62.7
τ_f	$(1.490 \pm 0.114) \text{ ps}$	$c\tau_f$	$(446.8 \pm 34.2) \mu\text{m}$
$\Delta\Gamma/\Gamma$	$(0.256 \pm 0.167) \text{ ps}$		
S_t	1.31 ± 0.01		
f_-	0.011 ± 0.002	N_-	107.7 ± 17.4
f_+	0.125 ± 0.006	N_+	1273.8 ± 62.2
τ_-	$(0.824 \pm 0.112) \text{ ps}$	$c\tau_-$	$(247.0 \pm 33.7) \mu\text{m}$
τ_+	$(0.661 \pm 0.029) \text{ ps}$	$c\tau_+$	$(198.2 \pm 8.7) \mu\text{m}$

Figure D.1: Graphical illustration of the fit model used for the transversity analysis. The 10,230 candidate events were generated according to the fit parameters obtained from the average lifetime fit (see figure 5.7). In addition $f_{CP+} = 79 \%$, $\tau_f = 1.461 \text{ ps}$, and $\Delta\Gamma/\Gamma = 15 \%$ was used. The fitted parameters are given in the table below.

	M	W	A	f_{sig}	f_{CP+}	τ_f	$\Delta\Gamma/\Gamma$	S_t	f_-	f_+	τ_-
W	-1										
A	-2	1									
f_{sig}	-1	20	1								
f_{CP+}	-0	-1	0	1							
τ_f	-0	-3	-0	-3	43						
$\Delta\Gamma/\Gamma$	-1	-1	0	7	32	90					
S_t	0	-0	-0	-1	0	0	0				
f_-	-0	1	0	1	-0	-0	0	-17			
f_+	0	-11	-0	-15	1	3	-1	-23	5		
τ_-	0	-0	-0	-0	0	-0	0	12	-50	-5	
τ_+	0	-9	-0	-2	1	1	5	14	-5	-45	4

Table D.2: Correlation coefficients (in percent) for the fit to 10,230 generated events. The upper half and the diagonal are not shown for clarity.

- Increasing the flavour specific lifetime τ_f from its true value has to be compensated by an increase of the width difference $\Delta\Gamma/\Gamma$.
- A too big CP-even fraction f_{CP+} steepens the CP-even lifetime and therefore increases τ_f and $\Delta\Gamma/\Gamma$.

We use 5,000 pseudo-experiments with 1,023 events each to estimate the precision on $\Delta\Gamma/\Gamma$ with such a small number of events. As input value for $\Delta\Gamma/\Gamma$ we use the rather optimistic value of 15 %. Table D.3 shows the fitted $\Delta\Gamma/\Gamma$, f_{CP+} , τ_f for true values $f_{CP+} = 50\%$, 60% , 70% , 80% , 90% . The vertical line in each histogram indicates the true value of the quantity. The maximum of the $\Delta\Gamma/\Gamma$ can be found to good approximation at the true value of 15 %, though the negative tail gets bigger and bigger with increasing f_{CP+} . The width of the distribution gives an estimate of the 1σ error of an individual measurement. For $f_{CP+} = 50\%$ – the best scenario for the transversity analysis – the width is about 40 %, nearly three times as big as the actual true value.

The other columns show the same distribution for f_{CP} and τ_f . With increasing f_{CP} we see negative tails for f_{CP} and a positive tails for τ_f , and a bias of the maximum to smaller values than the true value.

This study indicates that with the current small number of events, even under good circumstances (high true value of $\Delta\Gamma/\Gamma$ and f_{CP} close to 50 %), a transversity analysis will measure $\Delta\Gamma/\Gamma$ with a precision of only 40 %. In addition one has to deal with tails and biases. The precision only gets worse for data as additional efficiency and acceptance corrections for the transversity angle have to be applied.

We can constrain the flavour specific lifetime τ_f to the current world average

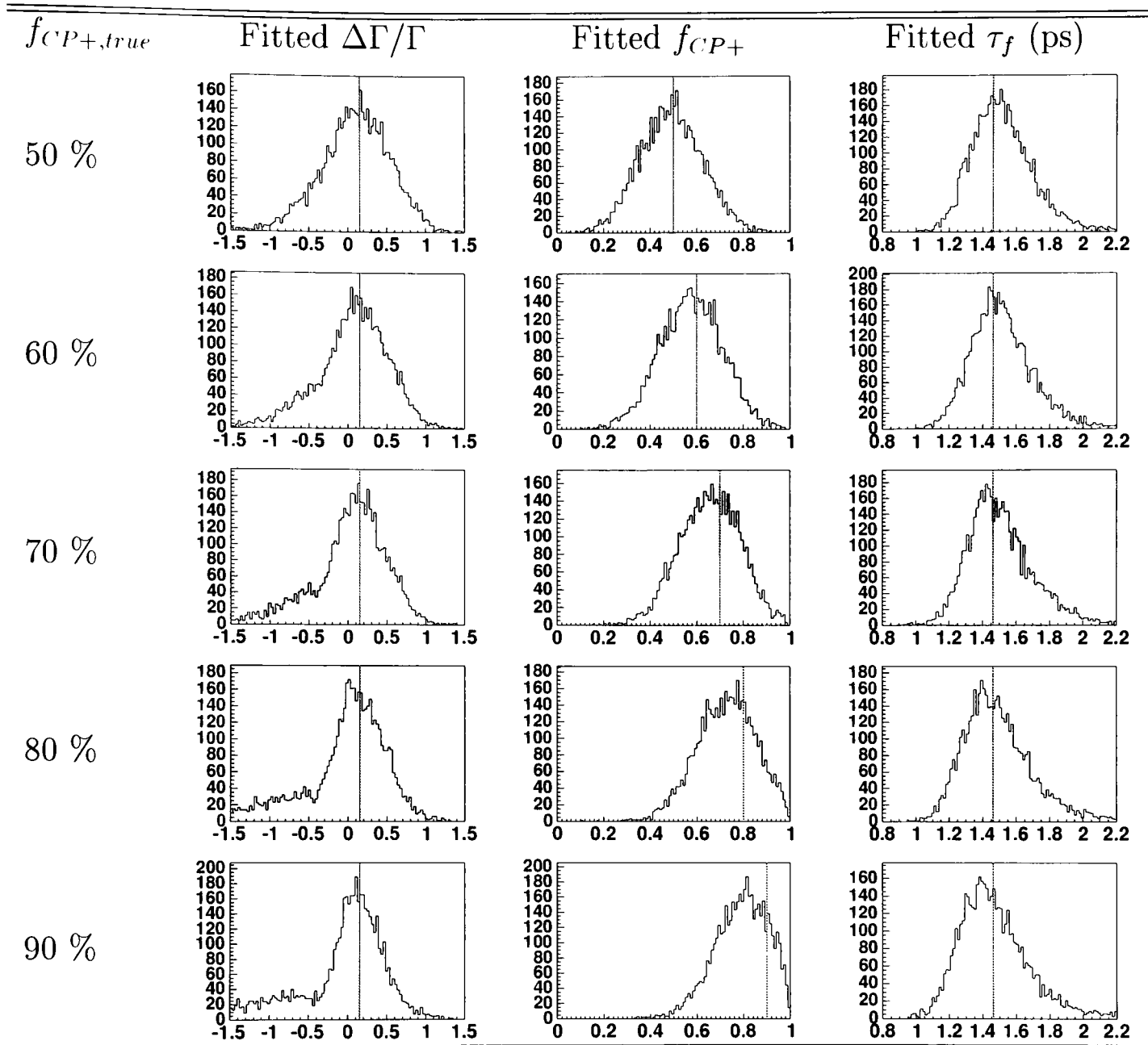


Table D.3: Fitted $\Delta\Gamma/\Gamma$, f_{CP+} , τ_f for 5,000 generated pseudo-experiments for true values $f_{CP+} = 50\%, 60\%, 70\%, 80\%, 90\%$. The vertical line in each histogram indicates the true value of the quantity.

semi-leptonic B_s lifetime $\tau = (1.461 \pm 0.057)$ ps [12], by multiplying a Gaussian term

$$\exp \left\{ \frac{1}{2} \left(\frac{\tau_f/\text{ps} - 1.461}{0.057} \right)^2 \right\} \quad (\text{D.6})$$

to the likelihood function \mathcal{L} . The resulting distributions are shown in table D.4. Note that the τ_f distribution is a Gaussian with $\mu = 1.461$ and $\sigma = 0.057$ by construction. There is small improvement in the $\Delta\Gamma/\Gamma$ precision, but biases arise now even in the $\Delta\Gamma/\Gamma$ distributions.

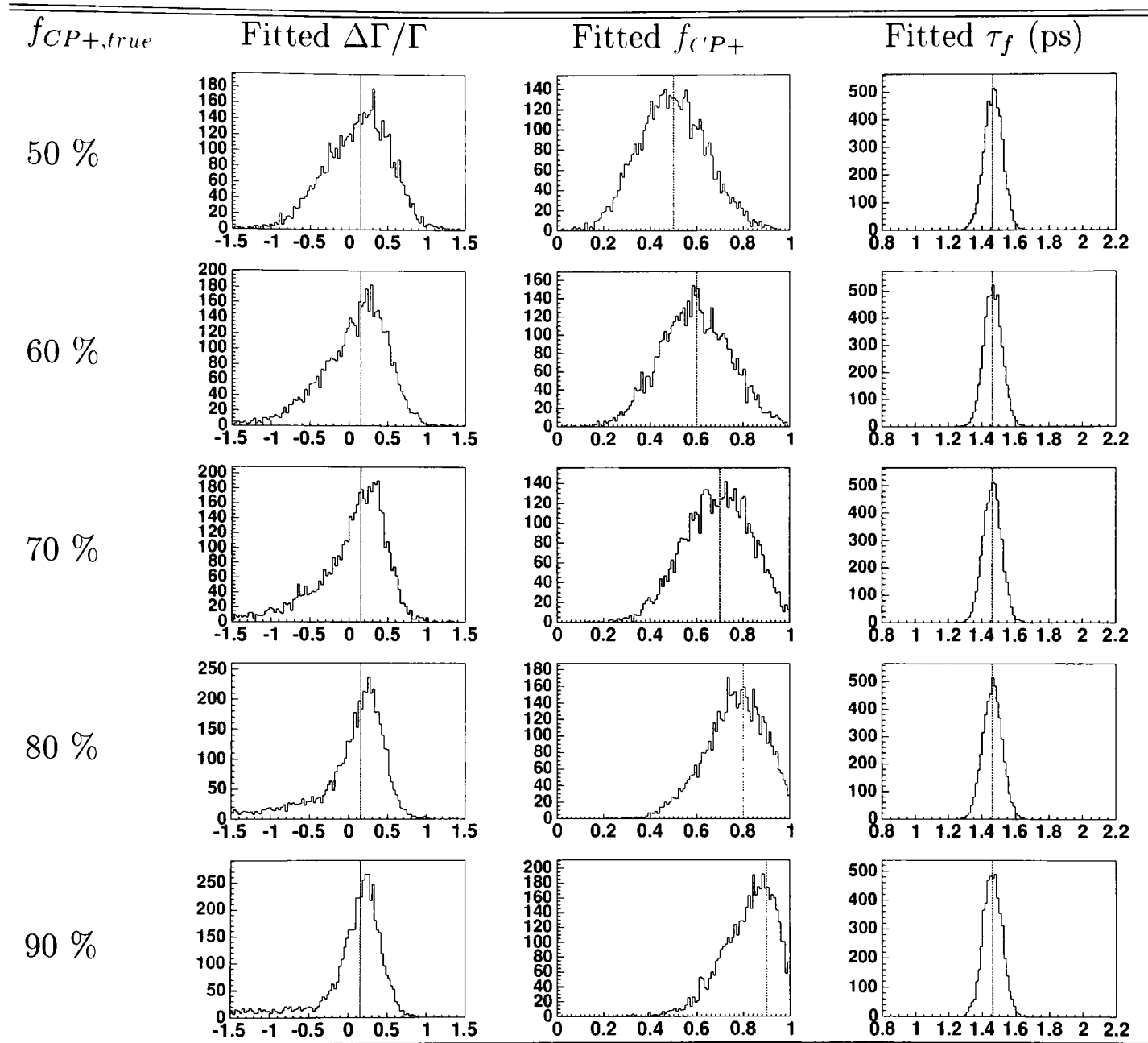


Table D.4: For a fit with constraint $\tau_f = (1.461 \pm 0.057)$ ps: Fitted $\Delta\Gamma/\Gamma$, f_{CP+} , τ_f for 5,000 generated pseudo-experiments for true values $f_{CP+} = 50\%, 60\%, 70\%, 80\%, 90\%$. The vertical line in each histogram indicates the true value of the quantity.

For all these reasons we abandon the aim of measuring $\Delta\Gamma/\Gamma$ with our data sample and be content with measuring the average lifetime of the B_s by fitting only a single lifetime to the proper decay time distribution.

Bibliography

- [1] M. E. Peskin and D. V. Schroeder, “*An Introduction To Quantum Field Theory*”, Addison-Wesley, 1993.
- [2] Donald H. Perkins, “*Introduction to High Energy Physics*”, Cambridge University Press, 2000.
- [3] M. Gell-Mann and A. Pais, “*Behavior Of Neutral Particles Under Charge Conjugation*”, Phys. Rev. **97** (1955) 1387.
- [4] K. Lande, E. T. Booth, J. Impeduglia, L. M. Lederman and W. Chinowsky, “*Observation Of Longlived Neutral V Particles*”, Phys. Rev. **103** (1956) 1901.
- [5] F. Muller *et al.*, “*Regeneration And Mass Difference Of Neutral K Mesons*”, Phys. Rev. Lett. **4** (1960) 418.
- [6] J. H. Christenson, J. W. Cronin, V. L. Fitch and R. Turlay, “*Evidence For The 2 π Decay Of The $K(2)0$ Meson*”, Phys. Rev. Lett. **13** (1964) 138.
- [7] C. Albajar *et al.* [UA1 Collaboration], “*Search For $B0$ Anti- $B0$ Oscillations At The Cern Proton - Anti-Proton Collider. (Paper 2.)*”, Phys. Lett. B **186** (1987) 247 [Erratum-ibid. **197B** (1987) 565].
- [8] H. Albrecht *et al.* [ARGUS COLLABORATION Collaboration], “*Observation Of $B0$ - Anti- $B0$ Mixing*”, Phys. Lett. B **192** (1987) 245.
- [9] M. Artuso *et al.*, “ *$B0$ Anti- $B0$ Mixing At The Upsilon ($4s$)*”, Phys. Rev. Lett. **62** (1989) 2233.
- [10] D. Abbaneo *et al.* [ALEPH, CDF, DELPHI, L3, OPAL, SLD Collaboration], “*Combined results on b -hadron production rates, lifetimes, oscillations and semileptonic decays*”, [arXiv:hep-ex/0009052].
- [11] F. Abe *et al.* [CDF Collaboration], “*Measurement of $B0$ anti- $B0$ mixing at the Fermilab Tevatron collider*”, Phys. Rev. Lett. **67** (1991) 3351.
- [12] K. Hagiwara *et al.*, Phys. Rev. **D66** (2002) 010001.
- [13] L. Wolfenstein, “*Parametrization Of The Kobayashi-Maskawa Matrix*”, Phys. Rev. Lett. **51** (1983) 1945.

- [14] A. Hocker, H. Lacker, S. Laplace and F. Le Diberder, “*A new approach to a global fit of the CKM matrix*”, Eur. Phys. J. C **21** (2001) 225 [arXiv:hep-ph/0104062].
- [15] T. E. Browder and S. Pakvasa, “*A Comment on the experimental determination of $|V_{ts}/V_{td}|^2$* ”, Phys. Rev. D **52** (1995) 3123 [arXiv:hep-ph/9501224].
- [16] I. Dunietz, “ *$B(s)$ - anti- $B(s)$ mixing, CP violation and extraction of CKM phases from untagged $B(s)$ data samples*”, Phys. Rev. D **52** (1995) 3048 [arXiv:hep-ph/9501287].
- [17] R. Fleischer and I. Dunietz, “*CP violation and CKM phases from angular distributions for B_s decays into admixtures of CP eigenstates*”, Phys. Rev. D **55** (1997) 259 [arXiv:hep-ph/9605220].
- [18] R. Fleischer and I. Dunietz, “*CP violation and the CKM angle γ from angular distributions of untagged B_s decays governed by $\bar{b} \rightarrow \bar{c}u\bar{s}$* ”, Phys. Lett. B **387** (1996) 361 [arXiv:hep-ph/9605221].
- [19] Y. Grossman, “*The B_s width difference beyond the standard model*”, Phys. Lett. B **380** (1996) 99 [arXiv:hep-ph/9603244].
- [20] K. Anikeev *et al.*, “*B physics at the Tevatron: Run II and beyond*”, arXiv:hep-ph/0201071.
- [21] M. Neubert and C. T. Sachrajda, “*Spectator effects in inclusive decays of beauty hadrons*”, Nucl. Phys. B **483** (1997) 339 [arXiv:hep-ph/9603202].
- [22] I. I. Bigi, B. Blok, M. A. Shifman, N. Uraltsev and A. I. Vainshtein, “*Nonleptonic decays of beauty hadrons: From phenomenology to theory*”, arXiv:hep-ph/9401298.
- [23] M. A. Shifman, “*Quark-hadron duality*”, arXiv:hep-ph/0009131.
- [24] M. Neubert, “*Heavy quark symmetry*”, Phys. Rept. **245** (1994) 259 [arXiv:hep-ph/9306320].
- [25] I. I. Bigi, B. Blok, M. A. Shifman and A. I. Vainshtein, “*The Baffling semileptonic branching ratio of B mesons*”, Phys. Lett. B **323** (1994) 408 [arXiv:hep-ph/9311339].
- [26] M. Neubert, “*Introduction to B physics*”, arXiv:hep-ph/0001334.
- [27] I. I. Bigi, N. G. Uraltsev and A. I. Vainshtein, “*Nonperturbative corrections to inclusive beauty and charm decays: QCD versus phenomenological models*”, Phys. Lett. B **293** (1992) 430 [Erratum-ibid. B **297** (1993) 477] [arXiv:hep-ph/9207214].
- [28] A. F. Falk, Z. Ligeti, M. Neubert and Y. Nir, “*Heavy Quark Expansion For The Inclusive Decay Anti- $B \rightarrow$ Tau Anti-Neutrino X*”, Phys. Lett. B **326** (1994) 145 [arXiv:hep-ph/9401226].

- [29] M. Battaglia *et al.*, “*The CKM matrix and the unitarity triangle*” [arXiv:hep-ph/0304132].
- [30] C. Gay, “*B mixing*”, Ann. Rev. Nucl. Part. Sci. **50** (2000) 577 [arXiv:hep-ex/0103016].
- [31] V. Weisskopf and E. P. Wigner, “*Calculation Of The Natural Brightness Of Spectral Lines On The Basis Of Dirac’s Theory*”, Z. Phys. **63** (1930) 54.
- [32] T. D. Lee, R. Oehme and C. N. Yang, “*Remarks On Possible Noninvariance Under Time Reversal And Charge Conjugation*”, Phys. Rev. **106** (1957) 340.
- [33] Serge Lang, “*Complex Analysis*”, Springer-Verlag New York Inc., 1999.
- [34] J. S. Hagelin, “*Mass Mixing And CP Violation In The B^0 Anti- B^0* ”, Nucl. Phys. B **193** (1981) 123.
- [35] R. Aleksan, A. Le Yaouanc, L. Oliver, O. Pene and J. C. Raynal, “*Estimation of Delta Gamma for the $B(s)$ - anti- $B(s)$ system: Exclusive decays and the parton model*”, Phys. Lett. B **316** (1993) 567.
- [36] A. J. Buras and P. H. Weisz, “*QCD Nonleading Corrections To Weak Decays In Dimensional Regularization And ’T Hooft-Veltman Schemes*”, Nucl. Phys. B **333** (1990) 66.
- [37] M. Beneke, G. Buchalla and I. Dunietz, “*Width Difference in the $B_s - \bar{B}_s$ System*”, Phys. Rev. D **54**, 4419 (1996) [arXiv:hep-ph/9605259].
- [38] M. Beneke, G. Buchalla, C. Greub, A. Lenz and U. Nierste, “*Next-to-leading order QCD corrections to the lifetime difference of B/s mesons*”, Phys. Lett. B **459** (1999) 631 [arXiv:hep-ph/9808385].
- [39] M. Beneke and A. Lenz, “*Lifetime difference of B/s mesons: Theory status*”, J. Phys. G **27** (2001) 1219 [arXiv:hep-ph/0012222].
- [40] U. Nierste, “*The width difference of B/s mesons*”, in *Proc. of the 5th International Symposium on Radiative Corrections (RADCOR 2000)* ed. Howard E. Haber, arXiv:hep-ph/0105215.
- [41] S. Hashimoto, “*B decays on the lattice*”, Nucl. Phys. Proc. Suppl. **83** (2000) 3 [arXiv:hep-lat/9909136].
- [42] N. Yamada *et al.* [JLQCD Collaboration], “ *B^0 anti- B^0 mixing with quenched lattice NRQCD*”, Nucl. Phys. Proc. Suppl. **94** (2001) 379 [arXiv:hep-lat/0010089].
- [43] F. Abe *et al.* [CDF Collaboration], “*Measurement of the B/s^0 meson lifetime using semileptonic decays*”, Phys. Rev. D **59** (1999) 032004 [arXiv:hep-ex/9808003].

-
- [44] A. S. Dighe, I. Dunietz, H. J. Lipkin and J. L. Rosner, “*Angular distributions and lifetime differences in $B_s \rightarrow J/\psi\phi$ decays*”, Phys. Lett. B **369** (1996) 144 [arXiv:hep-ph/9511363].
 - [45] J. L. Rosner, “*Determination Of Pseudoscalar Charmed Meson Decay Constants From B Meson Decays*”, Phys. Rev. D **42** (1990) 3732.
 - [46] D. Buskulic *et al.* [ALEPH Collaboration], “*Study of the $B(s)0$ anti- $B(s)0$ oscillation frequency using $D(s)$ - lepton+ combinations in Z decays*”, Phys. Lett. B **377** (1996) 205.
 - [47] P. Abreu *et al.* [DELPHI Collaboration], “*Mean lifetime of the $B(s)0$ meson*”, Z. Phys. C **71** (1996) 11.
 - [48] R. Barate *et al.* [ALEPH Collaboration], “*Study of $B/s0$ oscillations and lifetime using fully reconstructed D/s - decays*”, Eur. Phys. J. C **4** (1998) 367.
 - [49] K. Ackerstaff *et al.* [OPAL Collaboration], “*Measurements of the $B/s0$ and $\Lambda/b0$ lifetimes*”, Phys. Lett. B **426** (1998) 161 [arXiv:hep-ex/9802002].
 - [50] K. Ackerstaff *et al.* [OPAL Collaboration], “*A measurement of the $B/s0$ lifetime using reconstructed D/s - mesons*” Eur. Phys. J. C **2** (1998) 407 [arXiv:hep-ex/9708023].
 - [51] F. Abe *et al.* [CDF Collaboration], “*Measurement of B hadron lifetimes using J/ψ final states at CDF*”, Phys. Rev. D **57** (1998) 5382.
 - [52] P. Abreu *et al.* [DELPHI Collaboration], “*Study of $B0(S)$ - anti- $B0(S)$ oscillations and $B0(S)$ lifetimes using hadronic decays of $B0(S)$ mesons*”, Eur. Phys. J. C **18** (2000) 229 [arXiv:hep-ex/0105077].
 - [53] P. Abreu *et al.* [DELPHI Collaboration], “*Measurement of the $B/s0$ lifetime and study of $B/s0$ anti- $B/s0$ oscillations using D/s l events*”, Eur. Phys. J. C **16**, 555 (2000) [arXiv:hep-ex/0107077].
 - [54] M. Acciarri *et al.* [L3 Collaboration], “*Upper limit on the lifetime difference of short- and long-lived $B/s0$ mesons*”, Phys. Lett. B **438** (1998) 417.
 - [55] R. Barate *et al.* [ALEPH Collaboration], “*A study of the decay width difference in the $B/s0$ anti- $B/s0$ system using $\Phi\Phi$ correlations*”, Phys. Lett. B **486** (2000) 286.
 - [56] D. Buskulic *et al.* [ALEPH Collaboration], “*Study of the $B(s)0$ anti- $B(s)0$ oscillation frequency using $D(s)$ - lepton+ combinations in Z decays*”, Phys. Lett. B **377** (1996) 205.
- K. Ackerstaff *et al.* [OPAL Collaboration], “*Measurements of the $B/s0$ and $\Lambda/b0$ lifetimes*”, Phys. Lett. B **426** (1998) 161 [arXiv:hep-ex/9802002].

- [57] Fermilab Beams Division, “*Run 2 Handbook*”, March 12, 2001, unpublished.
<http://www-bd.fnal.gov/runII/index.html>.
M. Church, “*Tevatron Run II Performance and Plans*”, FERMILAB-Conf-02/142-E (2002).
- [58] “*Collider Run IIa: Plan for FY2003, Summary Report*”, September 30, 2002, unpublished.
<http://www-ap.fnal.gov/AP/SummaryPlan4DoE.pdf>.
- [59] The CDF II Collaboration, “*The CDF II Detector, Technical Design Report*”, FERMILAB-Pub-96/390-E (1996)
<http://www-cdf.fnal.gov/upgrades/tdr/tdr.html>.
The CDF Collaboration, “*Update to Proposal P-909: Physics Performance of the CDF II Detector with An Inner Silicon Layer and A Time of Flight Detector*”, CDF Note 5264 (1999), unpublished.
- [60] S. Nahn [CDF Collaboration], “*Status of the CDF Run II silicon detector*”, Nucl. Instrum. Meth. A **511** (2003) 20.
- [61] K. T. Pitts [CDF Collaboration], “*The Cdf Central Outer Tracker*,” Nucl. Phys. Proc. Suppl. **61B** (1998) 230.
- [62] Y.-K. Kim, *et. al*, “*Determination of the COT Inner Cylinder Materials*”, CDF Note 6168 (2002), unpublished.
- [63] M. Bishai, *et. al*, “*COT and SVXII Tracking Performance*”, CDF Note 5931 (2002), unpublished.
- [64] E. J. Thomson *et al.*, “*Online track processor for the CDF upgrade*”, IEEE Trans. Nucl. Sci. **49** (2002) 1063.
- [65] Y. Gotra, J. Kraus, T. Miao, “*Run II Level I Low p_T Muon Trigger Efficiency Measurement*”, CDF Note 6162 (2003), unpublished.
- [66] M. Bishai, *et. al*, “*Run II Dimuon Trigger Optimization and Efficiency Measurement*”, CDF Note 6004 (2002), unpublished.
- [67] H. Wenzel, “*Tracking in the SVX*”, CDF Note 1790 (1998), unpublished.
- [68] W.-M. Yao, K. Bloom, “*Outside-In Silicon Tracking at CDF*”, CDF Note 5991 (2002), unpublished.
- [69] M. Feindt, S. Menzemer, K. Rinnert, “*TrackingKal A Tracking and Alignment Software Package for the CDFII Silicon Detector*”, CDF Note 5968 (2002), unpublished.
- [70] M. Feindt, C. Lecci, S. Menzemer, K. Rinnert, “*Pre-Tracking PV z Finder*”, CDF Note 5988 (2002), unpublished.

- [71] Robyn Madrak, private communication.
- [72] K. Anikeev, *et. al*, “*Calibration of Energy Loss and Magnetic Field using J/ψ Events in Run II*”, CDF Note 5859 (2002), unpublished.
- [73] G. Bauer, *et. al*, “*Update on Calibration of Energy Loss and Magnetic Field using J/ψ Events in Run II*”, CDF Note 6355 (2003), unpublished.
- [74] J. Bellinger. *et. al*, “*A Guide to Muon Reconstruction and Software for Run 2*”, CDF Note 5870 (2002), unpublished.
- [75] Ting Miao, private communication.
- [76] K. Bloom, W. Dagenhart, “*Muon-Reconstruction Efficiency for Winter 2003 Conferences*”, CDF Note 6347 (2003), unpublished.
- [77] H. Stadie, *et. al*, “*The Beam Position in Run II*”, CDF Note 6327 (2003), unpublished.
- [78] F. Abe *et al.* [CDF Collaboration], “*Improved measurement of the B^- and anti- B^0 meson lifetimes using semileptonic decays*”, Phys. Rev. D **58** (1998) 092002 [arXiv:hep-ex/9806018].
- [79] F. Abe *et al.* [CDF Collaboration], “*Measurement of the lifetime of the B/s^0 meson using the exclusive decay mode $B/s^0 \rightarrow J/\psi \Phi$* ”, Phys. Rev. Lett. **77** (1996) 1945.
- [80] F. James and M. Roos, “*‘Minuit’ A System For Function Minimization And Analysis Of The Parameter Errors And Correlations*”, Comput. Phys. Commun. **10** (1975) 343.
- [81] R. Raja, “*A measure of the goodness of fit in unbinned likelihood fits*”, arXiv:physics/0207083.
- [82] Raymond Culbertson, private communication.
- [83] K. Anikeev, *et al.*, “*The inclusive B Lifetime using Run II Data at CDF*”, CDF Note 6023 (2002), unpublished.
- [84] L. Demortier, L. Lyons, “*Everything you always wanted to know about pulls*”, CDF Note 5776 (2001), unpublished.
- [85] N. Cabibbo, “*Unitary Symmetry And Leptonic Decays*”, Phys. Rev. Lett. **10** (1963) 531.
- [86] M. Kobayashi and T. Maskawa, “*CP Violation In The Renormalizable Theory Of Weak Interaction*”, Prog. Theor. Phys. **49** (1973) 652.
- [87] P. Sphicas, “*A $b\bar{b}$ Monte Carlo Generator*”, CDF Note 2655 (1994), unpublished.

-
- [88] P. Nason, S. Dawson and R. K. Ellis, “*The Total Cross-Section For The Production Of Heavy Quarks In Hadronic Collisions*”, Nucl. Phys. B **303** (1988) 607.
- [89] C. Peterson, D. Schlatter, I. Schmitt and P. M. Zerwas, “*Scaling Violations In Inclusive $E^+ E^-$ Annihilation Spectra*”, Phys. Rev. D **27** (1983) 105.
- [90] I. Furié, C. Paus, C. Seely, *Track Trigger Simulation at the Generator Level in Run II*, CDF Note 5548 (2000), unpublished.
- [91] E. Gerchtein and M. Paulini, “*CDF detector simulation framework and performance*”, arXiv:physics/0306031.
- [92] R. Brun, F. Carminati, “*GEANT Detector Description and Simulation Tool*”, CERN Programming Library Long Writeup **W5013** (1993), unpublished.
- [93] K. Anikeev, *et al.*, “*Exclusive B Lifetime Measurement with Run II Data at CDF*”, CDF Note 6024 (2002), unpublished.
- [94] J. Boudreau, *et al.*, “*Yields and cut optimization for $B \rightarrow J/\psi K$, $J/\psi \rightarrow \mu^+ \mu^-$* ”, CDF Note 6226 (2002), unpublished.
- [95] J. Marriner, “*Secondary Vertex Fit with Mass and Pointing Constraints (CTVMFT)*”, CDF Note 1996 (1993), unpublished.