Search for the rare decay $B^0 \to J/\psi\,\gamma$ in the *BABAR* experiment

Sheila McLachlin
Department of Physics
McGill University
Montréal, Québec
CANADA

February 16, 2004

A thesis submitted to the
Faculty of Graduate Studies and Research
in partial fulfillment of the requirements
for the degree of Master of Science

Your file Votre référence
ISBN: 0-612-98700-0
Our file Notre référence
ISBN: 0-612-98700-0

Abstract

This thesis presents the results of a search for the decay $B^0 \to J/\psi\gamma$ at the *BABAR* experiment. This is expected to be a *very* rare decay, with a branching fraction of the order of 10^{-8}, too low to be measured at *BABAR*; so the aim is instead to determine a 90% CL upper limit for $\mathcal{B}(B^0 \to J/\psi\gamma)$. No prior limit has been set in this channel. The analysis uses the complete data set from Runs 1, 2, and 3, comprising a total integrated onpeak (offpeak) luminosity of 113.1 (12.0) fb^{-1}; and both SP4 and SP5 Monte Carlo samples for optimization and background studies. We obtain a 90% CL upper limit on the branching fraction for $B^0 \to J/\psi\gamma$ of 1.2×10^{-6}.

Résumé

Cette thèse présente les résultats d'une recherche de la décomposition $B^0 \to J/\psi\gamma$ à l'experience *BABAR*. On s'attend à ce que cela soit une décomposition *très* rare, avec une fraction de branchement de l'ordre de 10^{-8}, ce qui est trop petite pour être mesurée à *BABAR*; donc, le but de l'analyse est plutôt de déterminer une limite supérieur de confiance 90% sur la fraction de branchement. Aucune limite précédente n'a été déterminée pour ce mode. L'analyse utilise les données des "Runs" 1, 2, et 3, comprenant une luminosité intégrée de 113.1 (12.0) fb^{-1}; et les échantillons Monte Carlo SP4 et SP5 pour l'optimisation et pour des études de bruit de fond. On obtient une limite supérieur de confiance 90% sur la fraction de branchement de 1.2×10^{-6}.

Contributions of Authors

Chapter 5 is adapted from an (unpublished) supporting document that I wrote for the $B^0 \to J/\psi\gamma$ analysis presented in this thesis. I did the research under the supervision of Chris Hearty, who is listed as a co-author of the document. However, it was me who actually wrote the document, and I am listed as the primary editor.

The other four chapters provide background information about particle physics in general, and the *BABAR* experiment in particular. I researched and wrote them by myself.

Acknowledgments

By far the biggest thank-you goes to Chris Hearty of UBC, who guided me through this analysis and through *BABAR* in general. This guy is amazingly smart and patient and went out of his way to help me all the time. So a thank-you with a cherry on top to him.

Now for the other thank-yous.

Thanks to Ryan Beaton and Dominique Mangeol, who gave me my first introduction to a *BABAR* analysis.

Thanks to my officemates: Teresa Spreitzer, Carsten Mueller, Pascal Fortin, Graham Gauthier, and Thomas Linder; for all of their help.

Thanks to Paul Mercure and Doug Maas, for their help with computer stuff.

Thanks to the whole *BABAR* collaboration, and in particular to the friendly *BABAR*ians who answered all my Hypernews questions.

Thanks to Prof. Patel, for his supervision and help preparing this thesis for submission.

And finally, thanks to Kate and Sarah, who didn't help with this thesis but are my favorite person and best friend, respectively.

Contents

List of Figures

List of Tables

Introduction

A B decay is generally considered to be "rare" if it has a branching fraction of about 10^{-6} or less. This thesis presents a search for the *very* rare radiative decay $B^0 \to J/\psi\gamma$.[1] No prior limits have been set in this channel. The leading-order contribution to $B^0 \to J/\psi\gamma$ is shown in Fig. 1. As this is an annihilation process it is suppressed by a power of Λ_{QCD}/m_b. A recent calculation by Lu, Wang and Yang [1] in the QCD factorization framework based on the heavy quark limit $m_b \gg \Lambda_{QCD}$ gives an expected branching fraction of

$$\mathcal{B}(B^0 \to J/\psi\gamma) = 7.65 \times 10^{-9}. \tag{1}$$

If this calculation is correct, then the decay $B^0 \to J/\psi\gamma$ is far too small to be measured at *BABAR*, so the aim of the analysis is to set a 90% confidence level (CL) upper limit on the branching fraction.

Rare decays even when they are not observed can provide valuable clues to fundamental physics [2]. If an allowed process is not observed, this could mean that a conservation law or other suppression mechanism is preventing it from occurring. To take a classic example, it was the non-observation of the flavor-changing neutral transition $b \to s$ that led Glashow, Illiopoulos, and Maiani to predict in 1970 the existence of the charm quark, whose existence was confirmed three years later [3].[2] Conversely, an unexpectedly large signal for a decay expected to be tiny could be a signal of New Physics beyond the Standard Model. So it is worthwhile to search for $B^0 \to J/\psi\gamma$ just to confirm that it is not observed. However, if a signal is detected that would be interesting as well, signalling that something is enhancing the decay rate.

New Physics effects that could enhance the $B^0 \to J/\psi\gamma$ branching fraction include intrinsic charm in the B^0 wavefunction [4] or an admixture of $(V + A)$ current to the standard $(V - A)$ current [1]. Unfortunately, the very smallness that makes $B^0 \to J/\psi\gamma$ such a sensitive New Physics probe also makes it too small to be measured by the current generation of B physics experiments. The use of this channel to investigate New Physics

[1] Charge conjugation is implied throughout the thesis.

[2] The charm quark allows for the suppression of this transition via the *GIM mechanism*, which introduces new contributions to the weak current which cancel the unobserved transitions.

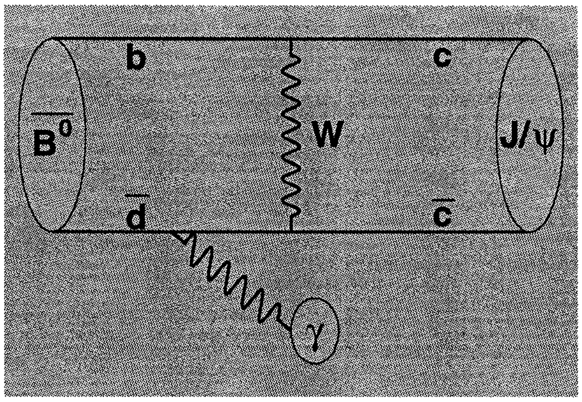

Figure 1: Leading-order contribution to $\overline{B}^0 \to J/\psi\,\gamma$. As this is an annihilation diagram the decay is suppressed by a power of Λ_{QCD}/m_b. Note that it matters that the photon is emitted from the light $\overline{d}$ quark — the other diagrams are suppressed by an additional power of Λ_{QCD}/m_b.

will therefore have to await the higher statistics of the planned Super*BABAR* and SuperBelle experiments. But in the meantime placing a 90% CL upper limit is a good start, and will at least allow us to confirm that the decay rate is tiny, as expected with or without New Physics beyond the Standard Model.

Thesis overview

The thesis is organized as follows:

- Chapter 1 describes the Standard Model of particle physics.
- Chapter 2 presents more advanced theoretical material required to understand how to obtain a prediction for $B^0 \to J/\psi\,\gamma$.
- Chapter 3 describes the PEP-II collider and the *BABAR* detector of the *BABAR* experiment.
- Chapter 4 describes the stages of processing that the data collected at *BABAR* pass through before they are used for an independent analysis like this one.
- Chapter 5 describes the analysis and presents the results of the search for $B^0 \to J/\psi\,\gamma$.

Chapter 1

The Standard Model of Particle Physics

1.1 Particles and forces

The Standard Model [5] is the present theory of particle physics. It explains all of the phenomena of particle physics in terms of the properties and interactions of a small number of fundamental particles. These are presented in Tables 1.1 and 1.2. The components of the Standard Model are:

- 2 forces, or types of interaction: the strong force, and the electroweak force. Electroweak interactions come in 3 types: electromagnetic, charged-current and neutral-current interactions.
- Quarks and leptons: The spin-1/2 particles (fermions) that make up matter. Quarks are the fermions that interact via the strong force; leptons are the ones that don't.
- Gauge bosons, or mediators: spin-1 particles that *mediate* interparticle interactions.
- The Higgs boson: a spin-0 particle that makes it possible for particles to have masses via a mechanism known as *spontaneous symmetry breaking.*

There is also a third force, gravity, but its effects at the particle level are too small to be measured, which makes it both impossible to study, and negligible in particle physics. Therefore gravity is ignored in the Standard Model.

Each type of interaction has its own gauge bosons to mediate that interaction. For the strong force, the mediator is the *gluon*; for the electromagnetic force, it is the *photon*; for the charged-current interactions it is the W^+ or W^- bosons; and for the neutral-current interactions it is the Z^0 boson.

Boson	Type of interaction	Charge	Mass (GeV/c^2)	Fermions affected
Photon (γ)	electromagnetic	0	0	charged particles (ie, all but neutrinos)
Gluon (g)	strong	0	0	colored particles (ie, all quarks)
Z^0 boson	neutral weak	0	91.187	all particles
W^+ boson	charged weak	+1	80.423	all particles
W^- boson	charged weak	−1	80.423	all particles

Table 1.1: The bosons (spin-1 particles). Each boson mediates a different type of interaction, or force.

	Flavor (abbrev.)/(mass in GeV/c^2)			
Generation	1	2	3	Charge
leptons	electron (e) (0.000511)	muon (μ) (0.106)	tau (τ) (1.777)	−1
	electron neutrino (ν_e) (0)	muon neutrino (ν_μ) (0)	tau neutrino (ν_τ) (0)	0
quarks	up (u) (0.003)	charm (c) (1.3)	top (t) (175)	+2/3 ("up-type")
	down (d) (0.006)	strange (s) (0.1)	bottom (b) (4.3)	−1/3 ("down-type")

Table 1.2: The fermions (spin-1/2 particles). Each quark also comes in one of three colors: red, blue, or green.

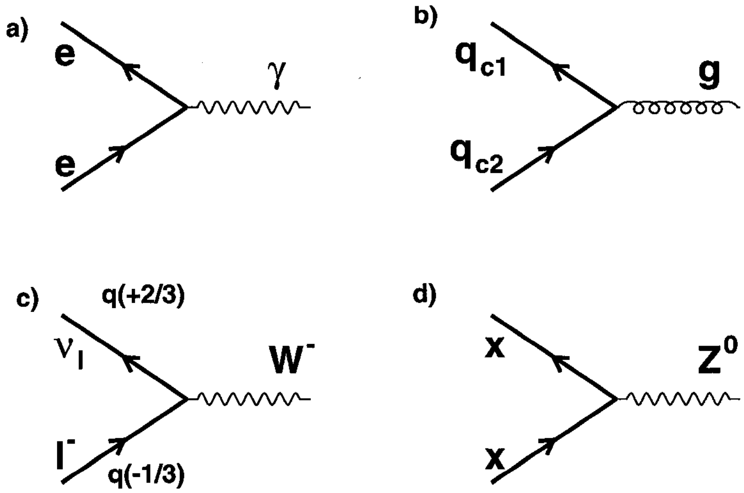

Figure 1.1: Feynman vertices for the 4 types of interaction. (a) Electromagnetic interaction. (b) Strong interaction. (c) Charged current interaction. (d) Neutral current interaction. "e" denotes a charged particle, "q" denotes a quark (with different colors c1, c2), and "x" denotes any quark or lepton.

Particles interact via a given force by exchanging the gauge bosons that mediate that force. A useful way to illustrate the mechanism of an interaction is with *Feynman diagrams.* The *Feynman vertices* of the 4 types of interaction are shown in Fig. 1.1. To represent a given interaction, you just piece together the relevant diagrams. More about this in Section 1.2.

In addition to those listed in Tables 1.2 and 1.1, every particle has an antiparticle — a particle with the same mass and spin but opposite quantum numbers. This is true of the fundamental particles and of particles consisting of more than one fundamental particle. The antiparticle of some particle a is denoted $\bar{a}$ and called "anti-a" or "a-bar." Some neutral particles — for example, the photon — can be their own antiparticles, but particles with nonzero quantum numbers are always distinct from their antiparticles.

The organization of the particles in Tables 1.1 and 1.2 warrants some explanation. The classification of the gauge bosons is pretty straightforward — they're just listed alongside the forces that they mediate. However, the spin-1/2 particles are classified according to several different properties: color, flavor, charge, and generation. The first three are called *quantum numbers*; a particle's quantum numbers are what gives it its own personal identity. They also determine whether and how a particle interacts via a given force. The rest of this section is devoted to elaborating this point.

The most important distinction is between spin-1/2 particles that carry the quantum number *color* and interact via the strong force, and those that do not. The colored spin-1/2 particles are called *quarks*, and the colorless ones are called *leptons*.

The strongly-interacting particles are quarks and gluons. A quark or a gluon can be one of three colors: red, green, or blue. These names were chosen in analogy to ordinary colors, to explain a strange experimental fact: no one has ever seen an isolated quark. Instead, quarks always show up as bound states of 3 quarks, or of a quark and an antiquark. This is explained by saying that quarks can exist only in colorless bound states. The 3-quark states, called *baryons*, consist of one red quark, one blue quark, and one green quark (red + green + blue = white, ie "colorless"). The familiar proton and neutron are both baryons. The 2-quark states, called *mesons*, are automatically colorless since they consist of a quark and an antiquark (red + antired = colorless, blue + antiblue = colorless, green + antigreen = colorless). Quark bound states in general are called hadrons; this includes both baryons and mesons.

There are six types, or *flavors* of quark: down, up, strange, charm, bottom, and top; these are often denoted by their first initials: u, d, s, c, b, t. There are also six flavors of lepton: electron (e), electron neutrino (ν_e), muon (μ), muon neutrino (ν_μ), tau (τ), and tau neutrino (ν_τ). Different flavors of particle are considered to be different "types" of particle because, as you can see in Table 1.2, a particle's flavor completely specifies the mass and quantum numbers of the particle — except for color, which must be specified separately.[1] Only the charged weak interactions can change a particle's flavor — the others can form new particles only by rearranging the quark flavors of the initial state.

Another way to classify particles is by the quantum number *electric charge*, or just *charge*, which specifies how they are affected by the electromagnetic force. From Table 1.2 you can see that there are three leptons with charge -1, and three leptons with charge 0. The charge 0, or neutral, leptons are called neutrinos, and they are massless,[2] whereas the charge -1 leptons have mass. As for the quarks, up, charm and top have charge $+2/3$, and down, strange, and bottom have charge $-1/3$. It is often convenient to refer to the charge $+2/3$ quarks as "up-type" quarks, and to the charge $-1/3$ quarks as "down-type" quarks. Because neutrinos are neutral, they cannot participate in electromagnetic interactions. All

[1] The alternative would be to say that there are 18 "types" of quark, from 6 colors $\times$ 3 flavors. But this is not a very useful distinction, since all observed particles are colorless anyway, and one rarely needs to specify the color of an individual quark. Plus, different flavors of particle have different masses, charges, and so on; whereas the color of a quark is unrelated to any other properties and so can always be specified independently.

[2] Recent evidence suggests that the neutrinos might have a very tiny mass after all (see Section 1.5), but for simplicity I will confine this discussion to the "classic" Standard Model with massless neutrinos and conserved lepton numbers, which is more than adequate for all of the physics in this thesis.

of the other quarks and leptons are charged, so they do interact via the electromagnetic force.

The charged-current interactions are the only interactions that can change a particle's flavor; however, there are restrictions on the way flavor is changed, and these are reflected in the grouping of quarks and leptons into doublets, or *generations.* Each lepton generation has its own conserved quantum number, called a *lepton number*; these are the *electron number*, *muon number*, and *tau number*, respectively. The electron and electron neutrino have the same electron number, +1; all other particles have zero electron number. The situation is analogous for the other two generations. So although the weak force can change lepton flavor, it can do this only within a generation — the only allowed transitions are $(e \leftrightarrow \nu_e)$, $(\mu \leftrightarrow \nu_\mu)$, and $(\tau \leftrightarrow \nu_\tau)$.

The situation for quarks as far as generations are concerned is a bit more complicated. The six different flavors of quark listed in Table 1.2 are by convention the *mass eigenstates*, so that each flavor of quark has a definite mass. However, the mass eigenstates are not the same as the weak eigenstates. What this means is that for quarks, the charged current weak interactions do conserve a generation number, but only if you choose new quark eigenstates. The convention is to transform the three down-like quarks from their mass/flavor eigenstates (q) to their weak eigenstates (q'):

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.1}$$

The matrix V_{ij} is called the *CKM matrix.* With this definition, you can define a conserved generation number: only the transitions $(u \leftrightarrow d')$, $(c \leftrightarrow s')$, and $(t \leftrightarrow b')$ are permitted, so in principle you could assign one conserved number to each of these "primed" generations. However, it is the mass eigenstates that are detected experimentally, and by convention these are well-established as the "real quarks." So the convention is to stick with the mass eigenstates, and say that any (down-type $\leftrightarrow$ up-type) transition is allowed, with the *quark mixing* described by the CKM matrix. More about this in Section 1.4.

The spin-0 Higgs boson is neither one of the matter particles, nor the mediator of a force. It was introduced into the Standard Model in order to account for particle masses. The mechanism by which the Higgs boson allows particles to have mass is called *spontaneous symmetry breaking*, and it is central to our understanding of electroweak unification, which will be explained in the next chapter. The Higgs boson is the only particle discussed so far that has not been found — apparently it is very heavy — but most particle physicists expect that it will be discovered soon.

Thus, particles are organized by spin, color, electric charge, flavor, and generation. All interactions conserve the quantum numbers color, charge, number of quarks, number of leptons, electron number, muon number, and tau number. Flavor is conserved in all interactions except for the charged-current interactions. These conservation laws suggests an underlying symmetry governing the particle interactions. Indeed, Noether's theorem [6] guarantees that every conserved quantum number must correspond to some symmetry. This principle is crucial in *quantum field theory*, the subject of the next section.

1.2 Quantum field theory

As the fundamental theory of particle interactions, the Standard Model must first of all follow the laws of quantum mechanics and relativity — it must be a *relativistic quantum field theory*. In quantum field theory particle behavior is described by a Lagrangian; and like in classical physics, the Lagrangian *is* the theory — the Standard Model Lagrangian plus the rules for what to do with it (from the Principle of Least Action) gives a complete description of absolutely any process in particle physics. But of course, the Lagrangian is not a "given" — like any physical law, it must be deduced from experimental observations. It was only after trying out many different models that theorists found the theory — the Standard Model — that reproduced the observed particle spectrum and interactions. The Standard Model is the union of two theories: the Glashow-Weinberg-Salam (GWS) theory of the electroweak interactions and Quantum Chromodynamics (QCD). Both QCD and the GWS theory are examples of a particular class of quantum field theory called a *gauge theory*, in which forces and conservation laws arise from a type of symmetry called *gauge invariance*. This section gives a brief overview of how the Standard Model Lagrangian is "derived" from basic physical principals plus the assumption of gauge invariance [5].

In relativistic quantum field theory, particles are described by fields, and the Lagrangian is a function of these fields and their covariant (spacetime) derivatives. A free particle A is described by terms containing only A-fields and derivatives, and the interaction of particles A and B is described by products of A-fields and B-fields. The "rules" for writing the Lagrangian for a set of spin-0, spin-1/2, and spin-1 particles come from basic physical requirements. First, a good relativistic theory must be Lorentz-invariant. Second, in order to be useful the theory should be *renormalizable* — that is, it should not depend on high energies outside of the energy range in which it applies. Additional requirements include probability conservation (probabilities should always add to one), locality (independence of physically separated events), and stability (the existence of a ground state). Among the most useful "rules" that come from these requirements are that each term must be a Lorentz scalar, and can contain no more than four fields or two covariant derivatives.

Having dealt with the basic physical requirements of relativistic quantum field theory, you can now proceed to search for the specific theory that describes our world. Looking back at Section 1.1, one of the most intriguing features of the observed particle spectrum is the existence of many conserved quantum numbers — charge, color, lepton family, and (often) flavor among them. Faced with conservation laws the next step is of course to turn to Noether's theorem for insight. According to Noether's theorem [6], conservation laws arise from symmetries in the Lagrangian. So to find the Lagrangian, you need to look for the symmetries that give rise to observed conservation laws. It it turns out that quantum number conservation comes from the demand that the Lagrangian be invariant under a certain class of symmetry transformations called *global gauge transformations.* There is a global gauge transformation for each conserved quantum number.

Demanding that the Lagrangian be invariant under the relevant global gauge transformation takes care of your conservation law, so you could just stop there. But in *gauge theory* you go a step further. Instead of enforcing the conservation law by demanding *global* gauge invariance, you demand *local* gauge invariance under the same gauge transformation. This has an amazing consequence. Not only does it impose conservation of the quantum number, but it actually gives rise to the mediator of the force associated with that quantum number! To see how this works, it helps to look at the example of quantum electrodynamics (QED), the theory that describes electromagnetic interactions. In QED, charge conservation comes from a very simple gauge transformation known as $U_{em}(1)$, where the "*em*" denotes electromagnetism.[3] Under a $U(1)$ transformation the electron field $\phi(x)$ transforms as follows:

$$\phi(x) \to e^{i\theta}\phi(x) \tag{1.2}$$

This is a *global* gauge transformation because the arbitrary parameter, θ, is a constant. Requiring the Lagrangian to be invariant under this global $U_{em}(1)$ transformation gives rise to electromagnetic charge conservation. Now comes the key step: suppose that you take the *global* $U_{em}(1)$ symmetry that gives charge conservation, and demand instead that it be *local* — that is, you allow θ to depend on x:

$$\phi(x) \to e^{i\theta(x)}\phi(x) \tag{1.3}$$

Now it is a bit harder to keep the Lagrangian invariant. But it can be done — provided you introduce a massless, spin-1 boson. For U(1), the form of the resulting Lagrangian

[3]Symmetry is described by a branch of mathematics called group theory. A *group* is a collection of elements whose relationships are defined by *group transformations.* Here, the group transformation is $U(1)$, and the group is made up of fields. For the Lagrangian, a symmetry exists if the Lagrangian is *invariant* (ie, does not change) under the group transformations. The corresponding conserved quantity — here, electromagnetic charge — is the *generator* of the group.

shows that this boson has zero charge, and mediates interactions between charged particles. You have just generated the photon! So while *global* gauge invariance gave you just charge conservation, *local* gauge invariance gave you charge conservation *and* the mediator of interactions between charged particles. This is the magic of gauge theory.

The Standard Model is the union of two gauge theories: quantum chromodynamics (QCD) which describes strong interactions, and the GWS theory of electroweak interactions. The procedure of generating massless bosons and conservation laws by demanding gauge invariance is analogous to that described above for the photon. For QCD, you group the colored particles (ie, the quarks) into $SU_c(3)$ *color triplets*:

$$\begin{pmatrix} r \\ b \\ g \end{pmatrix}$$

Here, r, b, and g denote the three quark colors: red, blue, and green. You then demand that the Lagrangian be invariant under local $SU_c(3)$ transformations. The result is color conservation and a spin-1 mediator with 8 different color states. This is, of course, the gluon.

Similarly, for electroweak theory, you group the left-handed[4] leptons and quarks into $SU_L(2)$ *weak isospin doublets*, and the right-handed ones into $SU_L(2)$ *weak isospin singlets*:

$$\begin{array}{cccccc}
\begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix} & e_R & \qquad & \begin{pmatrix} u_L \\ d_L \end{pmatrix} & u_R & d_R \\
\begin{pmatrix} \nu_{\mu L} \\ \mu_L \end{pmatrix} & \mu_R & & \begin{pmatrix} c_L \\ s_L \end{pmatrix} & c_R & s_R \\
\begin{pmatrix} \nu_{\tau L} \\ \tau_L \end{pmatrix} & \tau_R & & \begin{pmatrix} t_L \\ b_L \end{pmatrix} & t_R & b_R
\end{array}$$

(There are no right-handed neutrinos in the Standard Model.)

Finally, you group all the leptons and quarks into $U_Y(1)$ *weak hypercharge singlets*:[5]

$$e \quad \nu_e \quad \mu \quad \nu_\mu \quad \tau \quad \nu_\tau \qquad\qquad u \quad d \quad c \quad s \quad t \quad d$$

You then demand that the Lagrangian be invariant under $SU_L(2) \times U_Y(1)$ transformations. The result is weak-isospin and weak-hypercharge conservation, and 4 spin-1 mediators of

[4] Handedness is related to a particle's helicity, a measure of the relative directions of the particle's momentum and its spin.

[5] Well, actually you don't have to group anything, since all the quarks and lepton fields are *already* $U_Y(1)$ singlets; but I put it that way to show the parallel with the $SU_c(3)$ and $SU_L(2)$ cases.

the electroweak force, which can be identified by their interactions as the photon, the Z^0 boson, the W^+ boson, and the W^- boson.[6]

Thus, the requirement of $SU_c(3) \times SU_L(2) \times U_Y(1)$ invariance generates a gluon to mediate the strong interactions, and 4 bosons to mediate the electroweak interactions; and the resulting Lagrangian correctly describes all of the interactions and quantum numbers observed in experiments. Unless this is just an amazing coincidence, it appears that the underlying symmetry of the Lagrangian is indeed $SU_c(3) \times SU_L(2) \times U_Y(1)$.

There is one problem, however: all of the particles in this theory are massless! Local gauge invariance generates only *massless* spin-1 bosons, so all 4 weak mediators are massless. And there are no mass terms in the Lagrangian for the quarks and leptons, either. This is certainly not consistent with experiment. The solution to this problem is not obvious, but some very clever physicists finally found it in a phenomenon known as *spontaneous symmetry breaking.* In spontaneous symmetry breaking, the symmetry of a physical system is broken, not by an external agent, but by the ground state (from which things like masses are measured). In a local gauge theory, this makes it possible for the generated spin-1 bosons to be massive instead of massless. To break the symmetry, you have to introduce a new particle in the theory, the spin-0 *Higgs boson.* When the Higgs boson is included, the $SU_L(2) \times U_Y(1)$ symmetry is broken and weak isospin and weak hypercharge are no longer conserved. However, the $U_{em}(1)$ gauge symmetry of QED is *not* broken, and electromagnetic charge is still conserved. Therefore the W^+ , W^- and Z^0 bosons become massive, while the photon remains massless. Furthermore, the new Lagrangian contains mass terms for the quarks and leptons. In other words, the *Higgs mechanism* [7] has solved the mass problem, and the resulting *Standard Model Lagrangian* gives a perfect correspondence between theory and experiment.[7] Of course, there remains the slight problem of actually *finding* the Higgs boson, but there is good reason to believe that the reason it hasn't been found yet is just because it is very, very heavy [10]. There is hope that the high energies at the upcoming Large Hadronic Collider (LHC) at CERN [11] will be high enough to finally find the Higgs boson.

[6] In the unified electroweak theory, invariance under the QED U(1) gauge transformation is not imposed directly, but arises as a consequence of $SU_L(2) \times U_Y(1)$ invariance. This is because although QED does not need the weak interaction theory to be a gauge theory, the weak interaction theory *does* need to be unified with QED to be a gauge theory, as there is no separate gauge theory for weak interactions only.

[7] Actually, the original motivation for GWS theory was not to explain the W and Z^0 bosons, but rather to unify the electromagnetic and weak theories. GWS theory was proposed independently by Steven Weinberg in 1967 , and Abdus Salam in 1968, based on previous work by Glashow [8]. At that time the W and Z^0 bosons had not even been discovered yet, so GWS theory did not *explain* but rather *predicted* the existence of the heavy weak bosons. It was a major triumph for the theory when these were found at CERN in 1983 [9].

1.3 The Standard Model Lagrangian

The Standard Model Lagrangian can be separated into two parts, a free Lagrangian $\mathcal{L}_{free}$ and an interaction Lagrangian $\mathcal{L}_{int}$:

$$\mathcal{L} = \mathcal{L}_{free} + \mathcal{L}_{int} \tag{1.4}$$

The *free Lagrangian* $\mathcal{L}_{free}$ just contains all the basic terms describing the free propagation of each particle in the theory. That is, for each lepton and quark you have the standard Lagrangian for a free spin-1/2 particle; for each mediator, you have the standard Lagrangian for a free spin-1 (massive or massless) particle; and for the Higgs bosons, you have the standard Lagrangian for a free spin-zero particle. These would be present in any theory containing the same particles. The free Lagrangian is interesting from a theoretical point of view because it is one of the few Lagrangians in quantum field theory for which you can actually obtain a solution directly from a standard Lagrangian calculation — that is, using the Lagrangian to derive the equations of motion, which can then be solved to obtain the fields in terms of creation and annihilation operators. However, in particle physics you are usually more interested in more complicated processes than just free propagation, such as particle scattering, production, and decay — that is, you want to know about particle *interactions.*

The *interaction Lagrangian* $\mathcal{L}_{int}$ consists of all the terms that aren't part of the free Lagrangian, and each term in $\mathcal{L}_{int}$ describes a type of interaction. Unlike the free Lagrangian, the form of $\mathcal{L}_{int}$ is very much constrained by the $SU_c(3) \times SU_L(2) \times U_Y(1)$ invariance of the theory. A typical interaction term is just the product of the interacting fields, multiplied by a dimensionless *coupling factor* g_i describing the strength of the interactions, or *coupling.* Of particular interest are the terms $\mathcal{L}_{b-f}$ describing boson-fermion coupling ("h.c." = "hermitian conjugate"):

$$\mathcal{L}_{b-f} = \mathcal{L}_{gluon-fermion} + \mathcal{L}_{photon-fermion} + \mathcal{L}_{W^{\pm}-fermion} + \mathcal{L}_{Z-fermion}$$

$$\equiv \mathcal{L}_{gl-f} + \mathcal{L}_{em} + \mathcal{L}_{cc} + \mathcal{L}_{nc}$$

$$\mathcal{L}_{gl-f} = \tfrac{ig_3}{2} \textstyle\sum_q G^{\alpha}_{\mu} \bar{q} \gamma^{\mu} \lambda_a q + \text{h.c.}$$

$$\mathcal{L}_{em} = \textstyle\sum_f ieA_{\mu} \bar{f} \gamma^{\mu} Q f + \text{h.c.}$$

$$\mathcal{L}_{cc} = \tfrac{ig_2}{2\sqrt{2}} [W^+_{\mu} \overline{\nu}_m (1+\gamma_5) e_m + V_{mn} \overline{u}_m \gamma^{\mu} (1+\gamma_5) d_n] + h.c.$$

$\mathcal{L}_{nc} = \frac{ie}{\sin\theta_W \cos\theta_W} \sum_f Z_\mu^0 \bar{f}\gamma^\mu(g_V + \gamma_5 g_A)f$

The bosons can also couple with each other in various ways:

$\mathcal{L}_{b-b} = \mathcal{L}_{gl-gl} + \mathcal{L}_{WW\gamma} + \mathcal{L}_{WWZ} + \mathcal{L}_{WWWW} + \mathcal{L}_{WWZZ} + \mathcal{L}_{WW\gamma\gamma} + \mathcal{L}_{WWZ\gamma}$

(I will not write these out explicitly since they are not used nearly as often as the fermion-boson coupling terms.)

Ultimately, the goal is to calculate decay and scattering *amplitudes.* But with the interaction terms included in the Lagrangian, it is no longer possible to obtain exact solutions for the equations of motion. So physicists used their favorite approximation method, perturbation theory, to develop a general approximation scheme for calculating amplitudes. In this method $\mathcal{L}_{int}$ is treated as a perturbation to $\mathcal{L}_{free}$, with the coupling factors as the small expansion parameters. The result is a formal expression for the amplitude as a perturbation series in the coupling factors g_k:

$$A = \sum_k g_k \langle f|T_k|i\rangle \tag{1.5}$$

The terms in the perturbation series are all matrix elements, $\langle f|T_k|i\rangle$, where $|i\rangle$ and $\langle f|$ are the initial and final states, respectively, and T_k consists of integrals over one or more factors of H_{int}, the interaction Hamiltonian, which usually is just the negative of $\mathcal{L}_{int}$. The first one or two terms usually give sufficient accuracy:

$$A = \langle f|H_{int}^0(x=0)|i\rangle + \frac{(-i)^2}{2!}\int d^4x \langle f|H_{int}^0(x)H_{int}^0(x=0)|i\rangle + \ldots \tag{1.6}$$

Once you have the amplitude, you can make physical predictions. The cross section σ (for a scattering event) or lifetime τ (for a decay) are both directly related the amplitude A.

The Feynman Rules

It is possible to calculate amplitudes straight from Eq. 1.6, but in practice nobody does that. Instead, it is much easier to use the *Feynman rules* derived from this equation [5]. With this method, instead of evaluating matrix elements of the Lagrangian, all you have to do is draw diagrams and look up the associated mathematical expressions.

To find the amplitude for some process $i \to f$, the first step is to draw the diagrams connecting the initial and final state in question. To do this, you use the "puzzle pieces" of diagrams provided in tables of Feynman rules, provided in most particle physics textbooks. You saw four of the most important "puzzle pieces," or *Feynman vertices*, in Section 1.1. Any diagram made by piecing together Feynman vertices represents a possible physical process. For example, by sticking together two of the basic Feynman vertices in Fig. 1.1,

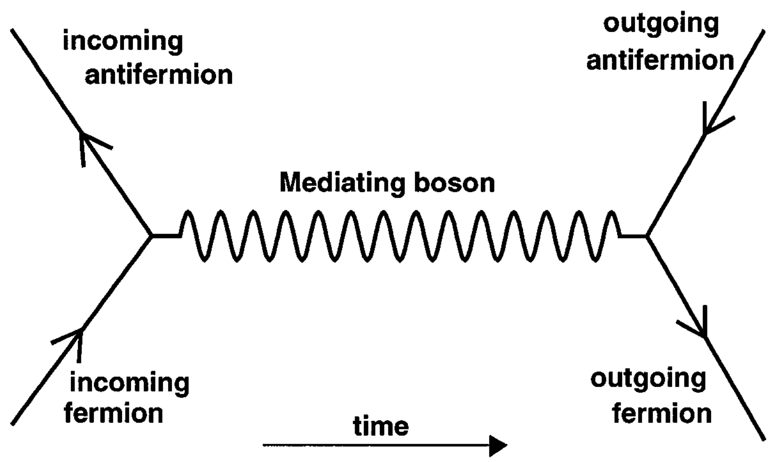

Figure 1.2: Interaction of two spin-1/2 particles via the exchange of a mediating boson. This is one of the most common processes in particle physics.

you obtain the diagram for one of the most common physical processes, the interaction of two spin-1/2 particles via the simple exchange of a mediating boson (Fig. 1.2).

There are actually an infinite number of diagrams for any given process, since you can always add more lines and loops between the initial and final states. These just correspond to the infinite number of terms in a perturbation series. However, in practice you need only the lowest-order diagrams, since even the best experiments cannot measure to an accuracy beyond about second order, anyway. One of the great things about Feynman diagrams is that as a general rule, the lowest-order diagrams are easy to find — they're just the simplest diagrams. The reason for this is that each vertex in a diagram contributes a factor of the very small coupling factor g_i, so that complicated diagrams with many vertices are suppressed by many powers of g_i, while simple diagrams with few vertices dominate.

The next step in a Feynman calculation is to write the mathematical term corresponding to each part of the diagram. Together, these terms make up the expression for the diagram's contribution to the amplitude. For example, for the electron-muon scattering diagram in Fig. 1.3 you get:

$$-i(2\pi)^4 A = \frac{(2\pi)^{-2}}{(16E_1E_2E_2E_4)^{1/2}}(-ie^2)[\bar{u}_{3i}\gamma^{\mu}_{ij}u_{1j}]\frac{-ig_{\mu\nu}}{(p_1-p_3)^2}[\bar{u}_{4k}\gamma^{\nu}_{kp}u_{2p}] \tag{1.7}$$

This is just a number. You still have to do a long, convoluted calculation to find out just what that number *is*, but in principle you're done.

So the Feynman rules provide a very nice shortcut to get from matrix elements of the

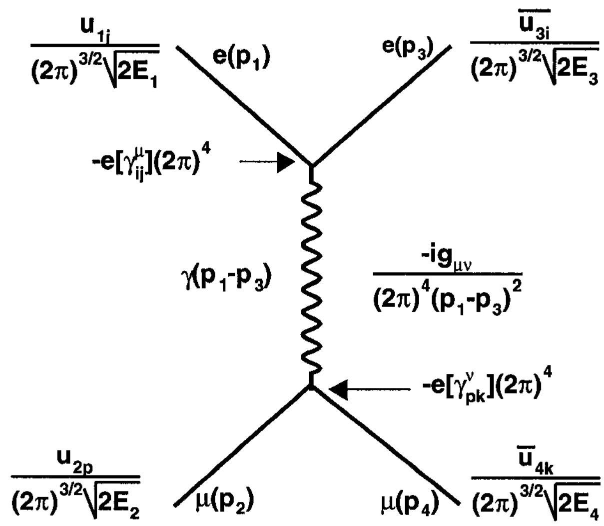

Figure 1.3: Using the Feynman rules to calculate the amplitude for electron-muon scattering.

Lagrangian, to the expression for an amplitude. But they're also extremely useful for other reasons. For one thing, just as drawings they provide an intuitive and "physical" picture of the mechanism of interactions. The isomorphism between processes and their diagrams is so complete in people's minds that it would be regarded as very odd if you were to talk about a decay or scattering event without providing its Feynman diagram, whether you plan to do the calculation or not. Another bonus of Feynman diagrams is that they provide an easy way to tell whether or not a physical process is even possible: if you can draw it, then it is; if you can't, then it's not. It's that simple. You could, of course, test if a decay is possible without Feynman diagrams by trying to do the calculation and seeing if the Lagrangian can accommodate it. But drawing pictures is much easier. A third advantage is that the Feynman rules enforce energy and momentum conservation, and the relevant quantum-number conservation laws, at each vertex. So the relevant conservation laws are automatically imposed for the overall process as well.

One interesting side effect of applying the kinematic conservation laws at each vertex is that *virtual* particles — particles represented by internal lines, which begin and end in the diagram — do not have to satisfy $E^2 = p^2 + m^2c^4$ as an ordinary particle does. Instead, they have whatever energy and momentum are required to satisfy the conservation laws. A common interpretation of this is to say that the virtual particles do not have to have the same mass as the real particles — they "do not lie on their mass shells." As a result, the virtual versions of heavy particles can sometimes be produced at much lower energies than the real versions would require. This provides an indirect method for studying heavy particles via their effects on low-energy processes.

1.4 The three forces

This section gives a brief overview of some of the most important properties of the electromagnetic, strong, and weak forces [5].

1.4.1 QED

- **coupling** The basic QED coupling is $e\bar{e}\gamma$, where e is a charged particle and γ is the mediating photon.

- **charge** The quantum number associated with the electromagnetic force is charge. The photon couples only with charged particles. In practice, this means that all of the fundamental particles except the neutrinos can interact electromagnetically. The type of interaction is determined by the sign of the charge — like charges attract, and unlike charges repel.

- **bound states** Oppositely-charged particles can and do form bound states with negative energy. The classic example is the hydrogen atom, a bound state of a positively-charged proton and a negatively-charged electron.

- **long-range** A force mediated by a boson of mass M has a range of about 1/M. So the electromagnetic force, mediated by the massless photon, has an infinite range ($1/M = 1/0 \to \infty$). However, in practice this effect is not too apparent because the strength of the force decreases with distance. Classically, the electromagnetic force between two particles is given by an inverse square law

$$F = -\alpha \frac{q_1 q_2}{r^2} \tag{1.8}$$

  where q_1 and q_2 are the particle charges, and r is the distance between the two particles.

- **coupling factor** The strength of the electromagnetic force is characterized by the *fine structure constant* $\alpha = 1/137$. The fact that this parameter is so small makes it possible to do calculations using a perturbative expansion in α. Both the QED Lagrangian and the Feynman calculus are derived from this expansion.

  Actually, α is not quite constant, as you'll see in Section 2.1 — it decreases with increasing distance (or equivalently, with decreasing energy). The variation is modest, but it is important in precision electroweak measurements.

1.4.2 The weak force

Both the W and Z^0 bosons are very heavy. Their masses $M_W = 80\,\text{GeV}$ and $M_Z = 91\,\text{GeV}$ are much higher than typical experimental energies. This has two important consequences:

- **short-range** The weak force has a range of about $1/M$ ($M = M_W$ or M_Z). This is nearly zero compared to typical distances in low-energy experiments, so W-mediated interactions resemble single-point interactions at low energies.

- **weak** In the Feynman calculus, mediating bosons of mass M generally contribute a factor of $1/(p^2 - M^2)$. At low energies $p^2 \ll M^2$, the mass M^2 dominates, and the weak force is effectively suppressed by a factor of $1/M^2$ compared to the electromagnetic interactions, which contribute just $1/p^2$. So the weak force is indeed weak at typical experimental energies. However, at high energies $p^2 \gg M^2$, the fact that the bosons are massive becomes less important, and the strength of the weak force becomes comparable to that of the electromagnetic force. The $SU_L(2) \times U_Y(1)$ symmetry is still broken, of course — but the theory is "less asymmetric" at high energies.

Other important properties of the weak force include:

- **coupling factor** The GWS theory revealed that the weak and electromagnetic forces come from the same $SU_L(2) \times U_Y(1)$ symmetry, and therefore share a common coupling factor. The weak coupling factor is just (a constant multiple of) the electromagnetic coupling factor α. So the weak force is weak, not because of an unusually small coupling factor as was once thought, but because it is suppressed by a power of $1/M_W^2$ at typical lab energies.

- **no bound states** The weak force is too weak to form bound states.

- **the Z^0 boson**

  The Z^0 boson can couple with any fermion f ($\bar{f}fZ^0$). The interactions of the Z^0 boson are called *neutral currents*, because like the photon the Z^0 boson does not carry charge (or any other quantum numbers). So any particle can emit or absorb a Z^0 boson, with no effect on the particle's charge (or other quantum numbers).

  You would be unlikely to notice a Z^0-mediated interaction unless you were looking for it. In any process in which a photon is exchanged ($q\bar{q}\gamma$), a Z^0 boson can be exchanged as well ($q\bar{q}\,Z^0$). However, at typical experimental energies the Z^0-mediated contributions to such a process are suppressed by a factor of $1/M_Z^2$ compared to the dominant electromagnetic contribution. So it is difficult to observe the effects of Z^0 bosons coupling with charged particles, because the analogous photon coupling always dominates at low energies. One option is to try to observe the one thing the Z^0 boson can do that photons can't do: couple with the *neutral* particles, the neutrinos ($\overline{\nu}_\ell\ \nu_\ell\ Z^0$). However, this is not much easier, because neutrinos are notoriously difficult to observe. A better way to observe Z^0-mediated processes is to run experiments at the Z^0 resonance energy — obviously $\frac{1}{p^2-M_Z^2}$ is very big *there*, and Z^0-mediated processes dominate over electromagnetic ones. This was the strategy at experiments like CERN in Geneva.

- **the W boson and the CKM matrix**

  W-mediated interactions are rather special because they are the only interactions that can change a particle's flavor. The interactions of the $W^\pm$ boson are called *charged current* interactions, because the $W^\pm$ boson carries one unit of charge. This allows it to couple with a lepton and its neutrino, or with a down-type quark and an up-type quark.

The CKM matrix As mentioned in Section 1.1, the quark weak eigenstates — the ones that couple with the W boson — are not the same as the quark mass eigenstates. The transformation between the two bases is described by the Cabbibo-Kobayashi-Maskawa (CKM) matrix V_{ij}, which rotates the down-type mass quarks (q) into down-type weak quarks (q') [10] [12]. When this matrix is written out explicitly, the Lagrangian for quarks becomes:

$$\mathcal{L}_{cc} = \begin{pmatrix} u & c & t \end{pmatrix} \frac{\gamma^\mu(1+\gamma_5)}{2} \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} W_\mu^- + \text{h.c.} \tag{1.9}$$

$$= \begin{pmatrix} u & c & t \end{pmatrix} \frac{\gamma^\mu(1+\gamma_5)}{2} \begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} W_\mu^- + \text{h.c.}$$

The latest PDG values of the CKM matrix elements, at the 90% confidence level, are[8] [10]:

$$V = \begin{pmatrix} 0.9741 \text{ to } 0.9756 & 0.219 \text{ to } 0.226 & 0.0025 \text{ to } 0.0048 \\ 0.219 \text{ to } 0.226 & 0.9732 \text{ to } 0.9748 & 0.038 \text{ to } 0.044 \\ 0.004 \text{ to } 0.014 & 0.037 \text{ to } 0.044 & 0.9990 \text{ to } 0.9993 \end{pmatrix} \tag{1.10}$$

The CKM matrix elements V_{ij} scale the rates of the various quark transitions, enhancing some and suppressing others. For example, the fact that $V_{cb} \gg V_{ub}$ means that the rate of $b \to c$ transitions is greater than the rate of $b \to u$ transitions.

[8] These values assume unitarity and only three quark generations.

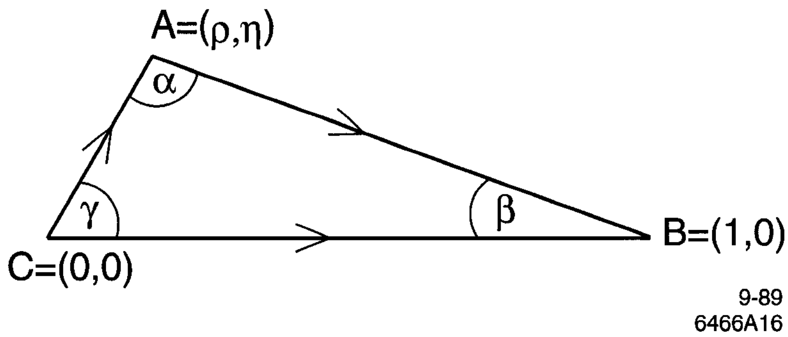

Figure 1.4: The unitarity triangle. The angles of the triangle are related to *CP* violation.

The Standard Model makes only one prediction for the CKM matrix: it should be unitary. This imposes 9 constraints on the 18 matrix elements:

$$V^\dagger V = VV^\dagger = 1 \Longrightarrow \sum_j V_{ji}^* V_{jk} = \sum_j V_{ij}^* V_{kj} = \delta_{ik} \tag{1.11}$$

An additional 5 degrees of freedom are eliminated by phase redefinitions of the left-handed fields, so that the CKM matrix actually depends on just 4 independent parameters — 3 real parameters and a 1 complex phase. In the Standard Model, this single phase is the source of *CP* violation.

Six of the unitarity conditions [1.11] require that three complex numbers add to zero. This can be depicted geometrically as the requirement that the corresponding three vectors in the complex plane form a closed triangle (Fig. 1.4). The most important of the six "unitarity triangles," often called "the" unitarity triangle, is the one whose angles are all of the same order of magnitude:

$$V_{td}V_{tb}^* + V_{cd}V_{cb}^* + V_{ub}V_{ud}^* = 0 \tag{1.12}$$

with the side $V_{cd}V_{cb}^*$ chosen real and rescaled to unit length. It is shown in Fig. 1.4. The sides of the triangle correspond to magnitudes of CKM matrix elements, and can be measured via the rates of processes involving these elements. The angles of this triangle, α, β, and γ, are related to the CKM phase and can be determined by measuring *CP*-violating asymmetries. One of the main goals of B physics today is to overdetermine the parameters of the CKM matrix, and thus test the CKM picture of *CP* violation.

1.4.3 QCD and the strong interactions

- **color** The quantum number associated with the strong force is color. Both the strongly-interacting fermions — the quarks — and their spin-1 mediator — the gluon

— are colored. Color is conserved in the strong interactions (and in all other interactions, for that matter).

- **bound states and quark confinement** One of the strange properties of the strong force is that quarks not only *can* exist in bound states, but they *have to*. Quarks exist only in bound states called hadrons — no one has ever succeeded in observing an isolated quark

  In QED, bound states consist of oppositely-charged particles — the net charge is zero. Similarly, in QCD bound states consist of colorless combinations of the quarks — the net color is zero. There are two possibilities — the quarks can form *mesons*, consisting of a quark and an antiquark with, say, red + antired = 0 (or blue, or green); or they can form *baryons*, consisting of three quarks or three antiquarks, with red + blue + green = 0.

- **asymptotic freedom** The strong force keeps the quarks confined in hadrons. But strangely enough, *within* the hadrons themselves the strong force suddenly becomes much weaker, so that the quarks are virtually free within a certain *confinement radius* r_c. This property is called *asymptotic freedom.*

- **coupling factor and perturbation theory** Both quark confinement and asymptotic freedom can be understood in terms of the behavior of the strong coupling factor α_S. The strong coupling factor varies dramatically with interquark distance, becoming much stronger at long distances and much weaker at short distances. This is in marked contrast to the electromagnetic coupling factor, which becomes weaker at longer distances.

  The large variation of the strength of the coupling factor α_S has profound implications for QCD. As for electromagnetism, both the QCD Lagrangian rules and the Feynman calculus are based on a perturbative expansion in the coupling factor, in this case α_S. At short distances α_S is small, so you can use perturbation theory. However, at longer distances α_S becomes large, and a perturbative expansion in α_S is no longer valid.

- **flavor $\mathbf{SU_f(3)}$** The quark bound states — hadrons and mesons — belong to different *multiplets* of a symmetry group called *flavor* $SU(3)$, or $SU_f(3)$ with the f denoting flavor. What this means is that there is an underlying symmetry, $SU_f(3)$, determining which quark combinations form bound states. It would be a perfect symmetry if the quarks all had the same masses; but they don't, so it's not. However, the symmetry is pretty good for the three lightest quarks, u, d, and s, since within the confinement radius they are highly relativistic and so their masses are negligible.[9]

[9]The limit $m_u, m_d, m_s \to 0$ is called the *chiral limit*, and the symmetries that arise in this limit are called

- **evidence for quarks** Although no one has ever observed a quark on its own, there is still experimental evidence for the quark model. First, of course, there is the success of the $SU_f(3)$ symmetry scheme in predicting hadron properties. But in addition, there is dynamical evidence for the quark model from deep inelastic scattering experiments, in which electrons are used to probe the internal structure of the nucleons [13]. The results show that the electrons are scattering off smaller pointlike fermions with fractional charges within the nucleons [14]. Additional evidence for quarks comes from the *jets* of hadrons produced in many high-energy processes, which arise as a result of quark and gluon interactions [15].

- **evidence for color** As you cannot observe a quark on its own, neither can you observe an object with nonzero color on its own. Nevertheless, the evidence for color is very strong. First, without color the hadronic wavefunctions violate the Pauli exclusion principle, and the very successful $SU_f(3)$ symmetry scheme wouldn't work. Second, experimental results for the ratio R, given by

$$R = \frac{\sigma(e^+e^- \to \text{hadrons})}{\sigma(e^+e^- \to \mu^+\mu^-)} \tag{1.13}$$

give clear evidence that there are three quark colors [16].

1.5 New Physics?

So far, all experimental observations in particle physics are consistent with the Standard Model, in spite of stringent testing of its conservation laws and other constraints [10]. Nevertheless, many physicists believe that the Standard Model does not tell the "whole story" in particle physics. The search for New Physics effects is motivated by a number of unresolved issues related to the Standard Model:

- **arbitrary input parameters** The Standard Model contains 18 arbitrary input parameters that must be determined from experiment: 6 quark masses, 3 charged lepton masses, 3 coupling factors, 4 CKM parameters, one vacuum expectation energy, and the mass of the Higgs boson. The Standard Model *allows* for all observed values of these parameters, but it does not *explain* them. That is, the Standard Model permits particles to have masses, and permits *CP* violation, and has an underlying $SU_c(3) \times SU_L(2) \times U_Y(1)$ symmetry that gives rise to all known interactions. But the model makes no predictions about the *values* of the masses or of the *CP*-violating phase in the CKM matrix; nor does it explain *why* the underlying symmetry should

chiral symmetries.

be $SU_c(3) \times SU_L(2) \times U_Y(1)$. Now, there is nothing actually *wrong* with this — after all, there is no law that says the universe can't be arbitrary. However, many particle physicists find it disturbing, and hope to find a "deeper" theory that would provide a better explanation of why things are the way they are.

- **matter/antimatter asymmetry** There is way more matter than antimatter in the universe, and physicists want to know why. One possibility is simply that the universe has always had more matter than antimatter. But once again, to many particle physicists this explanation seems "too arbitrary." Another possibility is that the universe started out with perfect matter/antimatter symmetry, and then something happened to change that. *CP* violation could do the job, provided there is thermal nonequilibrium and baryon number nonconservation [17]. The trouble with this model is that there is not enough *CP* violation in the Standard Model to account for the observed matter/antimatter asymmetry of the universe. This is one of the reasons that particle physicists are so interested in possible alternative (non-Standard Model) mechanisms of *CP* violation.

- **neutrino mixing** The Standard Model predicts the separate conservation of the three lepton generation numbers, and neutrinos that are exactly massless and stable, and never change flavor. Until very recently, all experimental observations were consistent with this prediction. However, recent results from the Sudbury Neutrino Observatory (SNO) [18] found that there must be ν_μ and ν_τ neutrinos coming from the sun, even though the solar neutrinos produced in the sun's nuclear reactions are pure ν_e. This is strong evidence for neutrino flavor-mixing. If this is confirmed, then the Standard Model will have to be modified to incorporate this effect. In particular, mixing would imply that the neutrinos may have masses after all, in the mV range.

- **general relativity** Quantum field theory gives extremely accurate predictions for all non-gravitational phenomena. Meanwhile general relativity give accurate predictions for the behavior of the large-scale universe. However, it is difficult to put them together because the resulting theory is nonrenormalizable. Theorists are on the job, but so far the problem has not been solved.

- **grand unification** In the same way that the electromagnetic and weak interaction theories have been "unified" into the GWS theory of electroweak interactions, most particle physicists expect that the GWS theory and QCD can be unified into a single *grand unified theory.* Although the strong and electroweak forces behave very differently at the energies accessible in the lab, there would be some very high energy scale — say, $M_X^2c^4$, with $M_X \sim 10^{15}\,\mathrm{GeV}/c^2$ — at which the coupling factors α_S and α

merge to one and the same value. (It would have to be a very high energy, to explain why we haven't *already* seen the effects of this heavier degree of freedom.)

There is no shortage of GUT models on the market today, though of course the huge unification energy $M_X^2 c^4$ makes it difficult to test them. As an example, one of the simplest and most popular models groups the quark and lepton generations together, so that each of the three generations consists of three down-type quarks, a positive lepton, and its antineutrino; or three up-type quarks, a negative lepton, and its neutrino. The charges in each generation add to zero, and this can be shown to be a requirement of the theory, thus explaining why the quark charges are fractions of the lepton charges. It also makes a pretty good prediction of the weak mixing angle, and predicts a proton lifetime of $\sim 10^{32} - 10^{33}$ years, which would explain why we have never seen proton decay (the universe is only about 10^{10} years old) despite its needing to be at least *permitted* to account for the baryon asymmetry of the universe (see above). But that's just one example of how a GUT can be constructed. Today the most promising modern GUT models are based on supersymmetry and string theory.

1.6 B physics

The search for $B^0 \to J/\psi\gamma$ at *BABAR* described in this thesis takes place in the wider context of rare B decay searches, which in turn is a branch of B physics, the study of B mesons. So before ending this chapter it is worthwhile to take a quick look at this exciting field of study.

B Physics is the study of B mesons, mesons that contain bottom (b) quarks. B physics is one of the most important branches of particle physics today. B mesons decay to a very wide variety of final states, so their decays can be used to study the properties of many different particles and processes. The major goals of B physics today are [19]:

- To measure the values of the CKM matrix elements V_{ij} with high accuracy and precision. As parameters in the Lagrangian, the CKM matrix elements are as important to know as the values of particle masses and coupling factors. According to the Standard Model, the CKM matrix is a 3×3 unitary matrix, so technically you would need only four independent measurements to completely determine the CKM matrix. However, researchers aim to *overdetermine* the parameters of the CKM matrix — that is, to make as many independent measurements of the paramters as possible. Then, assuming the Standard Model is correct, the "extra" measurements can be used as additional constraints. Conversely, if measurements reveal that it is impossible to satisfy all of the Standard Model constraints, this would be a sign of New Physics.

B physics makes several contributions to CKM measurements. $|V_{ub}|$ and $|V_{cb}|$ are measured in semileptonic B decays. FCNCs (see the discussion of rare B decays, below) have great potential for measuring the "top-flavored" CKM matrix elements, $|V_{td}|$, $|V_{ts}|$ and $|V_{tb}|$. The CKM angle β is measured in *CP* violation studies (see the next section), and B physicists are working on measuring the other two CKM angles α and γ as well.

- To study *CP* violation in B meson systems. In the Standard Model, *CP* violation arises as a result of a single, nonzero phase in the CKM matrix. *CP*-violating B decays are used to measure the angles of the CKM triangle, which is related to this phase. However, *CP* asymmetry measurements are also interesting in and of their own right, as they are related to such puzzles as why there is more matter than antimatter in the universe. Thanks to large B mixing and lifetimes, *CP* asymmetries in B meson systems are much larger than for kaons or D mesons, making B decays ideal for the study of this phenomenon.

  The recent precise measurements of sin2β from *BABAR* and Belle [21] [22], combined with measurements of the *CP*-violating parameter ε from kaon decays, lead Nir [23] to conclude that the CKM mechanism is very likely the dominant source of *CP* violation in flavor changing processes. So far, there is no evidence of New Physics; however, this possibility is not excluded at higher order or in other types of *CP* violation. In particular, *CP* violation has not yet been observed in B decays besides $B \to \psi K$, and direct *CP* violation has not been observed in B decays. Experiments attempting to measure the second-easiest CKM angle, sin2α, via $B \to \pi\pi$ decays are still ongoing. So much work remains to be done.

- To study heavy-flavor spectroscopy. Although *BABAR* and Belle produce and study only the lightest B^0 ($d\bar{b}$) and B^+ ($u\bar{b}$) B mesons, there are other types of heavier B mesons as well, such as B_s^0 ($\bar{b}s$) and the as-yet unobserved B_c^+ ($\bar{b}c$) mesons. These can be studied at higher-energy hadronic colliders like the Tevatron at Fermilab. Also of interest are the Λ_b and D particles, and the $c\bar{c}$ and $b\bar{b}$ onium states. And then there are the excited versions of all of these states — B^* and so on. All of these need to be studied further.

- To search for rare B decays. Rare B decays are those with very small branching fractions, typically of the order of 10^{-6} or less. In general, a rare B decay is one that does not proceed via the dominant $b \to c$ transition. These come in two types: there are the tree-level $b \to u$ decays, which are rare because they are CKM-suppressed, and the loop $b \to s$ and $b \to d$ decays, which are rare because the leading contribution comes from a loop diagram. The smallness of these decays make them sensitive probes of possible higher-order New Physics effects. For example, in the Standard Model, $b \to s$ and $b \to d$ flavor-changing neutral currents (FCNCs) are forbidden at tree level but can proceed via higher order loop diagrams. This allows FCNCs to serve as *indirect* probes of high-energy physics, because heavy particles such as the top quark or the as-yet undiscovered Higgs boson can appear in the loop. Also, as mentioned above, FCNC B decays play an important role in the measurement of the "top-flavored" CKM parameters V_{td}, V_{ts}, and V_{tb}. Combined with other CKM measurements these serve as constraints on the Standard Model and probes of possible New Physics effects, as explained above.

The study of rare B decays owes much to the pioneering work of the CLEO collaboration; notably the observation of the decay $B \to K^*(892)\gamma$, which provided the first evidence for one-loop penguin diagrams [24]. Today, interest in rare B decays is increasing as ever-higher samples of $B\overline{B}$ events become available and many rare B decays become measurable for the first time. Since *BABAR* is designed to make precision measurements of the CKM matrix, it is well-suited to rare decay studies. Indeed, most of the (potentially) *CP*-violating channels of interest at *BABAR* are themselves rare B decays. *BABAR* has obtained measurements or set tight limits in a wide variety of rare hadronic, leptonic and radiative B decays [25] [26].

The decay $B^0 \to J/\psi\,\gamma$ is not an FCNC; however, as described in Chapter 2, it too has the potential to serve as a sensitive probe for New Physics effects such as right-handed coupling and intrinsic charm. Unfortunately, even *BABAR* does not have high enough statistics for a measurement of the tiny $B^0 \to J/\psi\,\gamma$ branching fraction. However, the large number of $B\overline{B}$ events collected by *BABAR* should make it possible to set a reasonable upper limit for this decay mode.

1.7 Summary

This chapter reviewed the Standard Model of particle physics. The Standard Model is a union of two gauge theories, Quantum Chromodynamics (QCD) and and Glashow-Weinberg-Salaam (GWS) theory of electroweak interactions. The $SU_c(3) \times SU_L(2) \times U_Y(1)$ invariance of the Lagrangian gives rise to the four gauge bosons which serve as the mediators of lepton and quark interactions. The introduction of the Higgs boson in the theory breaks the $SU_L(2) \times U_Y(1)$ invariance of the electroweak interactions, allowing particles to have their experimentally observed masses. Calculations in the Standard Model are generally done using the Feynman calculus, which is based on a perturbative expansion in terms of the small coupling factors in the Lagrangian. The Standard Model has never yet failed an experimental test, though searches for New Physics effects are ongoing. B physics experiments are important in the ongoing investigation of the Standard Model, providing measurements of the CKM matrix and CP violation, and contributing to the search for New Physics. Rare decays like $B^0 \to J/\psi\,\gamma$ are useful tools in that search.

Chapter 2

Calculations for $B^0 \to J/\psi\,\gamma$

This chapter describes the calculation of the predicted branching fraction for $B^0 \to J/\psi\,\gamma$. So far, the only prediction of the branching fraction for $B^0 \to J/\psi\,\gamma$ comes from Lu, Wang, and Yang [1]. As for any B decay, the starting point is the effective Hamiltonian for B decays, derived from an operator product expansion that separates the long- and short- distance contributions into matrix elements and Wilson coefficients, respectively. This chapter begins with a review of the problem of renormalization in general, and works its way to Lu, Wang, and Yang's calculation. It also briefly describes some New Physics phenomena that could enhance the $B^0 \to J/\psi\,\gamma$ decay rate.

2.1 Renormalization and the Renormalization Group

As described in Section 1.2, to calculate amplitudes in quantum field theory you use the Feynman rules to compute the contributions from diagrams at successively higher orders in perturbation theory. The leading order contributions generally come from tree diagrams. However, the higher-order loop diagrams are often divergent. Unlike in tree diagrams, in loop diagrams the momentum of the virtual loop particles can range from zero to infinity. The trouble with this is that there is a high-energy scale beyond which quantum field theory is not meant to apply. This reflects our ignorance of physics at very short distance (high-energy) scales, beyond the range of experiments. It may be that at such low resolutions there are other heavy degrees of freedom that quantum field theory does not take into account. But in integrating loops all the way up to $p = \infty$, you are acting as if the theory describes physics at all energy scales — hence the divergences. In order to get rid of the divergences, then, you need to take the high-energy limit into account. This section explains how this is done [27] [5].

The procedure for dealing with the divergences involves two steps. The first step is to *regularize* the theory, making the diagrams finite by taking the cutoff into account. There are many different *regulators*, or methods of regularization. For example, a simple method to make an integral finite is to impose a momentum cutoff. However, this method is seldom used except as an example because it spoils the gauge symmetry of the theory. Instead, most people use dimensional regularization (DR), as this regulator preserves all the symmetries of the theory. In any case, regularization gets rid of the infinities, but now you have another problem — your theory depends on an unknown and arbitrary cutoff energy Λ.

Now, this cutoff-dependence is not necessarily fatal. Of course, it it problematic, because we know that physical observables — cross sections, lifetimes, masses — give the same value in every measurement. So obviously they do not depend on an inherently arbitrary scale Λ. However, as long as the theory can be formulated in such a way that it still gives unambiguous (and correct!) predictions of measured quantities, then it should be fine. Strange as it sounds, there actually is a way to make sure that the cutoff-dependence introduced by the regulator does not extend into the theory's physical predictions. This process is called *renormalization.*

The Standard Model Lagrangian contains a number of parameters that must be input from experiments. The ones relevant to renormalization are the masses m (one per particle) and the coupling factors g (for a given type of interaction). The key to renormalization lies in the difference between the *physical* or *renormalized* parameters of the theory — the ones that are actually measured — and the unphysical, *bare* parameters in the Lagrangian. Although they have the same names, the bare masses and coupling factors are actually very different from the physical masses and coupling factors. Experimental quantities like cross sections and lifetimes are calculated in terms of the bare parameters, but the bare parameters themselves have no physical meaning — they're just calculation tools. To get values for the bare parameters, you have to relate them to the measured physical parameters.

What does this have to do with the problem of diverging loops? To relate the bare and physical parameters, you have to do a Feynman calculation for some physics process. For example, to calculate the physical mass in terms of the bare mass, you do a Feynman calculation for a propagating particle; and to calculate the physical coupling factor in terms of the bare coupling factor, you do a Feynman calculation for a two-body interaction. However, these calculations include a whole series of divergent higher-order loop diagrams (Fig. 2.1), so you have to use a regulator. This gives relations of the form

$$\begin{aligned} m_r &= m_0 + \delta m(m_r, g_0, \Lambda) \\ g_r &= g_0 + \delta g(m_r, g_0, \Lambda) \end{aligned} \tag{2.1}$$

between the bare (g_0, m_0) and physical or renormalized (m_r, g_r) masses and coupling factors.

Having related the bare and physical quantities, the next step is to make m_r and g_r

independent of Λ. You can see from Eq. 2.1 that in order to enforce this condition, you will have to let m_0 and g_0 depend on Λ. But that's okay, because m_0 and g_0 are never measured, so you can set them to be *whatever they have to be* to make the physical parameters cutoff-independent. This is called "hiding" or "absorbing" the divergences in the bare parameters. So the renormalization procedure is just to apply this "hiding" recursively, moving up one order in perturbation theory each time, to whatever order you want. Once you have found the relation between the bare and physical parameters, you can take the bare parameters out of the theory, and rewrite the theory in terms of the physical parameters. This is much better, because (a) you can get values for the physical parameters from experiment, and (b) the divergences are gone.

The amazing thing about renormalization is that it eliminates the divergences from the theory by fixing just these two parameters — the masses and the coupling factors. A theory is called *renormalizable* if (a) the divergence-hiding procedure works at all orders in perturbation theory, and (b) it needs to be applied to only a finite number of parameters to remove all of the divergences. In the 1970s 't Hooft showed that all gauge theories — including the Standard Model's component QCD and GWS theories — are renormalizable [28]. This is fortunate, because although a nonrenormalizable theory may be correct, such a theory is not very useful for making predictions because it depends on unknown physics at high-energy scales.

Renormalization gives you m_r and g_r in terms of the bare parameters. But it turns out that there is more than one solution: the possible values of m_r and g_r form a set $[m_r(\mu), g_r(\mu)]$. The *renormalization scale* μ is an arbitrary parameter with energy dimension. Different values of μ define different (but equally valid) parameter sets $[m_r(\mu), g_r(\mu)]$. To apply the theory, you have to:

- Choose the scale μ_0 at which you wish to use the theory.
- Measure the parameters in an experiment at some scale E.
- Obtain values for the parameters in the set $[m_r(\mu_0), g_r(\mu_0)]$ by comparison to the measured set $[m_r(\mathrm{E}), g_r(\mathrm{E})]$.

The simplest case occurs when the scale of interest μ_0 is the same as the measurement scale E. The choice $\mu = E$ lets you use the simple relations $m_r(\mu) = m(E)$ and $g_r(\mu) = g(E)$. This gives a simple relation between the two parameter sets because with this choice, logarithmic terms like $\log(\mu/E)$ vanish.

If the scale of interest μ_0 is *not* the same as the measurement scale E, then you have two choices:

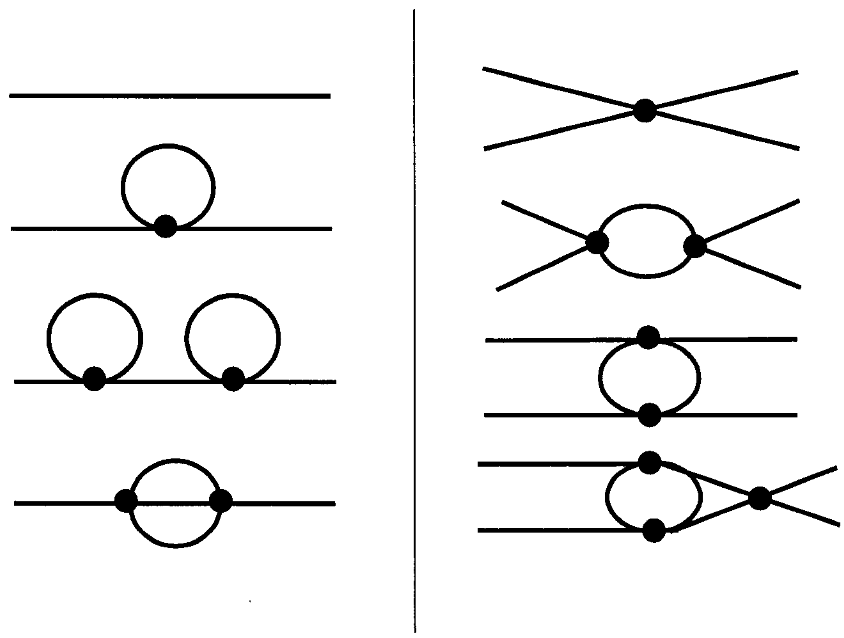

Figure 2.1: Typical Feynman diagrams used to calculate masses and coupling factors. Right: The "two-point" function for a propagating particle, used for mass calculations; and a few higher-order corrections. Left: The "four-point" function for a two-particle interaction, used for coupling factor calculations; and a few higher-order loop corrections.

- Relate the sets $[m_r(\mu_0), g_r(\mu_0)]$ and $[m_r(\mathrm{E}), g_r(\mathrm{E})]$ directly. The trouble with this method is that now the logarithmic terms like $\log(\mu_0/E)$ do not vanish. Furthermore, in the common case $\mu_0 \gg E$ these and similar logarithmic terms can become very large and may even spoil the perturbation series on which the theory is based.

- An alternative method is to start by setting $\mu = E$ as before, and then to use equations for $g_r(\mu)$ and $m_r(\mu)$ in terms of μ to obtain the set $[m_r(\mu_0), g_r(\mu_0)]$ from the set $[m_r(\mathrm{E}), g_r(\mathrm{E})]$. This is called "running" or "evolving" the masses and coupling factors from $\mu = E$ to $\mu = \mu_0$. This method is better because it avoids the aforementioned large logarithms.

To use the second method, you need to determine *how* m_r and g_r depend on μ. To do this you impose the requirement that the bare parameters be μ-independent. This gives the *renormalization group equations* (RGEs), first-order differential equations for $m_r(\mu)$ and $g_r(\mu)$:

$$\left.\begin{array}{l} \frac{\partial}{\partial\mu} m_0[m_r(\mu), \mu] = 0 \\ \frac{\partial}{\partial\mu} g_0[g_r(\mu), \mu] = 0 \end{array}\right\} \Longrightarrow \text{RGEs for } m_r(\mu) \text{ and } g_r(\mu) \tag{2.2}$$

Solving the RGEs gives the explicit μ-dependence of $m_r(\mu)$ and $g_r(\mu)$. $m_r(\mu)$ and $g_r(\mu)$ are often called the *running mass* and *running coupling factors* because they "run with μ." Of course, although this second method avoids the large logarithms, no information has been lost. It's just that the sum of large logarithms is taken into account when you solve the RGEs. We say that solving the RGEs automatically sums the large logarithms.

The expressions for the renormalized QED and QCD coupling factors, α and α_S (which are related to g_i), are of particular interest.[1] The one-loop calculations for α and α_S give the following expressions for the running of the electromagnetic and strong coupling factors [29]:

$$\frac{1}{\alpha_i(\mu)} = \frac{1}{\alpha(\mu_0)} + b_i \log\frac{\mu^2}{\mu_0^2} \tag{2.3}$$

These equations relate the theory at scale μ_0 to the theory at scale μ. The coefficients b_i are given by

$$b_{em} = \frac{1}{3\pi}\left[\frac{1}{9}n_{1/3} + \frac{4}{9}n_{2/3} + n_{-1}\right] \tag{2.4}$$

$$b_s = \frac{1}{12\pi}\left[2n_q - 11N\right] \tag{2.5}$$

Here, n_q is the number of fermions, and $n_{2/3}$, $n_{1/3}$, and n_{-1} are the number of (u, c, t), (d, s, b), and (e, μ, τ), respectively, with masses below the cutoff energy μ. The important

[1] It turns out that the variation in the masses is extremely tiny, so to a good approximation they can be treated as energy-independent constants. And after all that work!

point is that b_{em} is positive, so the electromagnetic coupling increases with energy — or equivalently, decreases with distance. So in principal, for QED perturbation theory becomes *more* accurate at long distances. This is in marked contrast to QCD, in which b_s is negative and the strong coupling increases with distance. When α_S gets too big, perturbation theory becomes invalid. This point will be crucial in the the following sections. For the sake of reference, it is conventional to define the *QCD scale*, Λ_{QCD}, as the energy scale at which the strong coupling (in theory) becomes infinite. Experimental measurements give $\Lambda_{QCD}=$ $\mathcal{O}(100$ MeV$)$.

2.2 Calculations for B decays

In QCD calculations, physics at different scales must be treated very differently. QCD actually gives fairly good predictions for quark *bound* states, because for short interquark distances the coupling factor α_s is small, the quarks are relatively free (*asymptotic freedom*) and the theory is perturbative, just like QED. Short-distance effects in QCD can therefore be calculated using the usual quantum field theory method: perturbation theory, with the coupling factor as the small expansion parameter. However, for hadron *decays* the situation is quite different — when the quarks begin to separate, long-distance effects come into play, the coupling factor α_s blows up and the theory is no longer perturbative. So the usual perturbative quantum field theory methods do not apply, and the calculation becomes very difficult. Theorists have worked very hard to develop nonperturbative methods to deal with long-distance effects in QCD, such as lattice gauge calculations, QCD sum rules, HQET and chiral perturbation theory. The following is so important that I will put it in bold print: **To calculate things in QCD, you need to separate short-distance from long-distance effects. The short-distance effects are calculated in perturbative QCD, in powers of α_S. The long-distance effects must be calculated using some nonperturbative method.**

Although the decay of a B meson is governed by weak interactions, QCD effects are unavoidable in any physics process involving quarks, because quarks exist only in strongly-bound states (hadrons). So all B decays receive QCD corrections. Therefore, the first step in any B decay calculation must be to separate the long- and short- distance contributions. For decays of B and other heavy mesons, the standard procedure is to use an operator product expansion (OPE) to derive an effective Hamiltonian in which the long-distance and short-distance contributions are separated. Once the crucial separation of scales is achieved, the short-distance contributions can be evaluated using standard perturbative and renormalization methods in quantum field theory, and the long-distance contributions can be tackled using some non-perturbative method.

2.2.1 Operator product expansions

There are many contexts in quantum field theory in which it is a good idea to replace a product of operators with a single local operator. This is known as an operator product expansion (OPE) [30] [31] [32], and it is a powerful theoretical tool for separating the long- and short-distance contributions in a problem.

The physical idea behind the operator product expansion technique is that at low energy scales, a short-range interaction mediated by a heavy boson corresponds approximately to a point interaction. Most B decays are mediated by a W boson, which travels only a distance of $\mathcal{O}(1/M_W)$. As far as the initial and final quarks (characteristic length $\mathcal{O}(1/m_q)$) are concerned, the distance travelled by the W boson is so small that it might as well be zero — they "see" only a point interaction. Accordingly, in the OPE, the interaction is "shrunk to a point" and replaced by an effective vertex.[2] This removal or "integrating out" of heavy mediating bosons as explicit, dynamical degrees of freedom is a common strategy for constructing a low-energy effective theory. The new theory, having fewer degrees of freedom, is simpler than the old theory; and yet the procedure ensures that the physical effects of the original operator are retained in the effective theory.

Using an operator product expansion you get the effective Hamiltonian for B decays, which serves as the basic starting point for any B decay calculation [31]:

$$\mathcal{H}_{eff} = \frac{G_F}{\sqrt{2}} \sum_i V^i_{CKM} C_i(\mu) Q_i(\mu) \tag{2.6}$$

valid up to corrections of order (m_b^2/m_W^2). The amplitude for the decay of a B meson from an initial state i into some final state f is then given by:

$$A = \frac{G_F}{\sqrt{2}} \sum_i V^i_{CKM} C_i(\mu) \langle f|Q_i(\mu)|i\rangle \tag{2.7}$$

Here, G_F is the Fermi constant, V^i_{CKM} denote the relevant CKM factors, and μ is the renormalization scale, usually chosen to be $\mu = m_b$. With this choice, the short-distance contributions ($E > m_b$) go into the *Wilson coefficients* $C_i(\mu)$, and the long-distance contributions ($E < m_b$) go into the matrix elements $\langle f|Q_i(\mu)|i\rangle$.[3] The Wilson coefficients are independent of the particular decay under consideration, so they can be tabulated, just like "real" coupling factors. Meanwhile, all of the process-dependence is contained in the matrix elements.

[2] In terms of Feynman calculations, this corresponds to the familiar approximation $\frac{1}{p^2-m_W^2} \to \frac{-1}{m_W^2}$ in the propagator.

[3] In this context μ is often called the *factorization scale*, where "factorization" refers to the separation of short- and long-distance physics.

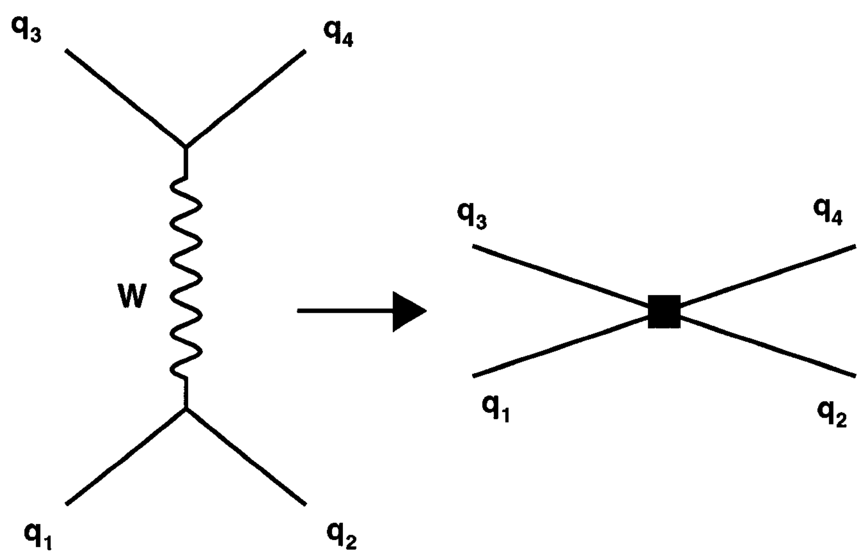

Figure 2.2: Integrating out the W boson. The effective four-quark operators that replace the W boson and the other, more complicated interactions in Fig. 2.3 are typically represented by black boxes.

The local operators Q_i that make up the basis for $B^0 \to J/\psi\gamma$ are given by [1]:

$$\begin{array}{ll} Q_1 = (\bar{c}_\alpha b_\beta)_{V-A}(\bar{d}_\beta c_\alpha)_{V-A} & Q_2 = (\bar{c}_\alpha b_\alpha)_{V-A}(\bar{d}_\beta c_\beta)_{V-A} \\ Q_3 = (\bar{d}_\alpha b_\alpha)_{V-A}(\bar{c}_\beta c_\beta)_{V-A} & Q_4 = (\bar{d}_\alpha b_\beta)_{V-A}(\bar{c}_\beta c_\alpha)_{V-A} \\ Q_5 = (\bar{d}_\alpha b_\alpha)_{V-A}(\bar{c}_\beta c_\beta)_{V-A} & Q_6 = (\bar{d}_\alpha b_\beta)_{V-A}(\bar{c}_\beta c_\alpha)_{V-A} \\ Q_7 = \frac{3}{2}(\bar{d}_\alpha b_\alpha)_{V-A} e_c (\bar{c}_\beta c_\beta)_{V-A} & Q_8 = \frac{3}{2}(\bar{d}_\alpha b_\beta)_{V-A} e_c (\bar{c}_\beta c_\alpha)_{V-A} \\ Q_9 = \frac{3}{2}(\bar{d}_\alpha b_\alpha)_{V-A} e_c (\bar{c}_\beta c_\beta)_{V-A} & Q_{10} = \frac{3}{2}(\bar{d}_\alpha b_\beta)_{V-A} e_c (\bar{c}_\beta c_\alpha)_{V-A} \end{array} \tag{2.8}$$

For other decays, the form of the operators is very similar, although the quark content might differ (eg., s instead of d quarks). Fig. 2.3 shows the original diagrams in the full theory, from which the Q_i are derived. The operator product expansion takes each diagram and replaces it with an *effective* vertex, represented by a black box (Fig. 2.2).[4] From the original diagrams in Fig. 2.3 you can see how the Wilson coefficients and the matrix come to depend on μ — you need renormalization to deal with all those QCD loops.

Both the Wilson coefficients $C_i(\mu)$ and the matrix elements $\langle f|Q_i(\mu)|i\rangle$ depend individually on the renormalization scale μ and on the choice of renormalization scheme for the local operators. In principal these dependencies should cancel in Eq. 2.7, as decay amplitudes are physical quantities and cannot depend on μ or the renormalization scheme. So

[4]Just to warn you, people often refer to the diagrams in Fig. 2.3 as the diagrams for the Q_i, although this is not technically correct. The Q_i are the *effective* operators, so they're just black boxes.

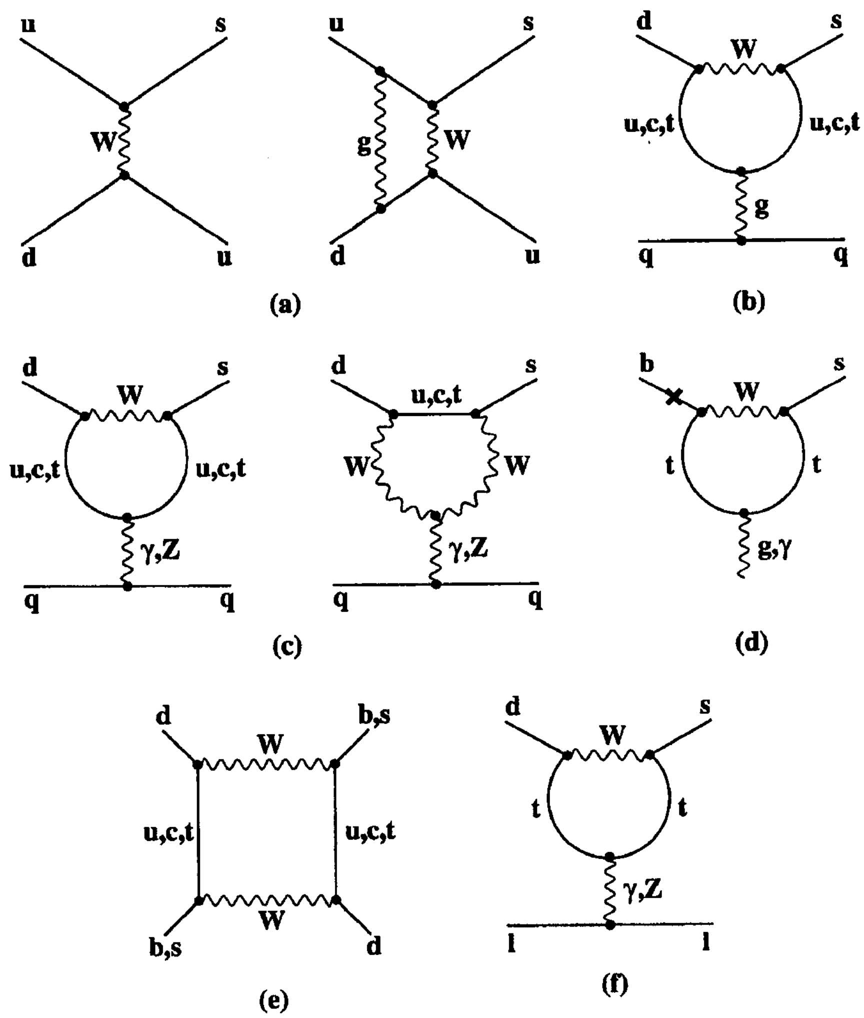

Figure 2.3: Typical diagrams in the full theory, from which the effective operators in Eq. 2.8 are derived.

if every term in the perturbation expansion were calculated, the result would be $\mu-$ and scheme-independent. However, in practice $C_i(\mu)$ and $\langle f|Q_i(\mu)|i\rangle$ can be calculated only to leading or next-to-leading order in perturbation theory, and the resulting *truncated* series for the decay amplitude may well be μ- and scheme- dependent. Unfortunately there are no good solutions to this problem; all you can do is keep it in mind, and possibly try to find ways to calculate to higher orders.

2.2.2 Calculating the Wilson coefficients

At short distances α_S is very small, so the short-distance Wilson coefficients C_i can be calculated perturbatively in α_S. The first step is to obtain an initial value for the Wilson coefficients at some μ to the desired order in α_S by *matching* the full theory on to the effective theory, as follows:[5]

- Calculate the amplitude A in the full theory (with the W propagator) to the desired order in α_S, for arbitrary μ and external states.

- Compute the matrix elements $\langle f|Q_i|i\rangle$ to the same order in α_S, with the same μ and same treatment of external states.

- Extract C_i from $A = C_i\langle Q_i\rangle$.

In practice $\mu = \mathcal{O}(M_W)$ is the most convenient choice for the matching procedure. However, ultimately you want to know the Wilson coefficients at a scale more appropriate to B decays, $\mu = \mathcal{O}(m_b)$. Plus you are also faced with the usual problem of large logarithms ($\ln M_W/\mu$) which spoil the perturbation series in $\alpha_s(\mu)$. To deal with both of these issues you treat the coupling factors C_i just like the coupling factors in Section 2.1. There, the requirement that the bare coupling factors be independent of μ gave a set of equations, the *renormalization group equations* (RGEs), for the value of the running coupling factors at any μ. Solving the RGEs automatically took care of the large logarithms. Similar RGEs can be derived for the Wilson coefficients, and once you solve them you can evolve $C_i(\mu = m_W)$ down to the more appropriate scale $\mu = m_b$. That gives you $C_i(m_b)$ and no large logs — exactly what you want.

Thus, the calculation of the Wilson coefficients is done using ordinary perturbation theory plus the renormalization group. This is possible thanks to the operator product expansion, which ensures that the Wilson coefficients contain only perturbation-friendly short-distance elements.

[5]This matching is analogous to the part of renormalization where you input some values from experiment.

2.2.3 Calculating the matrix elements

At long distances the strong coupling factor α_S blows up, so for the long-distance matrix elements $\langle f|Q_i|i\rangle$ a perturbative expansion in α_S is out of the question. Therefore it is necessary to resort to a nonperturbative method to determine the matrix elements. Some of the most common methods include:

- **lattice gauge theory [33]** Lattice QCD is a "brute-force" approach to QCD calculations. It bypasses the problem of finding a good approximation scheme, instead trying to solve QCD directly via a numerical simulation over a discrete lattice of space-time points.

- **$1/N_c$ expansion [34]** The symmetry of QCD is $SU_c(N_c)$, with the number of colors $N_c = 3$. But the $1/N_c$ expansion treats N_c as a variable in order to explore the symmetries that arise in the large-N_c limit, when $1/N_c$ becomes a small parameter.

- **HQET and HQE [32]** Heavy-quark effective theory (HQET) and the Heavy Quark Expansion (HQE) are both based on the heavy-quark limit $m_b \gg \Lambda_{QCD}$. Instead of α_S, they rely on expansions in inverse powers of the b-quark mass, $1/m_b$ or Λ_{QCD}/m_b, which are small because the b quark is heavy.

- **chiral perturbation theory [35]** This method exploits the symmetries that appear in the limit that the masses of the light quarks (u, d, s) go to zero.

- **QCD sum rules [32]** QCD sum rules are useful for calculating spectra, decay constants, and form factors. The primary assumption is *quark-hadron duality* — the idea that the B meson and b quark can sometimes be treated interchangeably in theoretical calculations. QCD sum rules often use HQET as well.

In the next section, you will see how a relatively new technique called *QCD factorization* is used to apply HQE/HQET-like techniques to *exclusive* B meson decays like $B^0 \to J/\psi\gamma$.

2.2.4 Summary and practical considerations

Thus, to do calculations for B decays you use the effective Hamiltonian derived from an operator product expansion, which divides the problem into two tasks: the calculation of the short-distance Wilson coefficients C_i using ordinary perturbative methods, and the calculation of the long-distance matrix elements $\langle f|Q_i|i\rangle$ using some nonperturbative technique.

In practice, you can usually skip the Wilson coefficient calculation and just look up the tabulated values instead, for example in Ref. [31]. But the matrix elements for different decays are all different, so unless someone has done the calculation for the decay that you are interested in, you may have to do the calculation yourself, using one of the long-distance techniques like QCD sum rules or lattice gauge theory. You will have to choose a renormalization scheme and renormalization scale μ, and make sure you use the same ones for the Wilson coefficients and the matrix elements. In principle, that should give you your amplitude. It is the nonperturbative calculation of the long-distance matrix elements that is the biggest challenge.

2.3 QCD factorization

To calculate the long-distance contribution to $B^0 \to J/\psi\gamma$ Lu, Wang and Yang [1] use a relatively new method called *QCD factorization* [36]. In the context of hadronic B decays, the word "factorization" actually describes two completely different procedures:

- *Factorization of short- and long- distance physics.* This is the meaning of the word factorization in the context of techniques like operator product expansions, HQE, or HQET. In particular, when μ from the previous section is called the "factorization scale", this refers to μ's role as the "dividing line" between short- and long-distance contributions.

- *Factorization of matrix elements into form factors, decay constants, or other useful quantities.* In this context, the word factorization refers to a model-based approach to the calculation of two-body decays, in which the matrix elements of a four-quark operator are expressed as the product of two matrix elements of color-singlet currents.

QCD factorization is aptly named; for it implements *both* types of factorization. This section explains why nonleptonic B decays are so hard to calculate, and describes the traditional *naive factorization* approach to these decays. It then describes how QCD factorization is used to obtain a more reliable result.

2.3.1 Naive factorization

The previous section described how an operator product expansion is used to separate physics at short- and long-distance scales. With the b-quark mass m_b as the "dividing line," the short distance contributions ($\mu > m_b$) go into the Wilson coefficients and the long-distance contributions ($\mu < m_b$) go into the matrix elements. However, the scale at which QCD becomes nonperturbative, $\Lambda_{QCD} \sim 0.1\,\text{GeV}$, is actually quite a bit lower than $m_b \sim 5\,\text{GeV}$; $m_b \gg \Lambda_{QCD}$. So the "long-distance" matrix elements actually still include some "short-distance" physics in the range $\Lambda_{QCD} < \mu < m_b$. So a good first step in tackling the calculation of the matrix elements is yet another separation of scales, this time with

Λ_{QCD} as the dividing line, and the short-distance contributions entering at $\mathcal{O}(m_b)$ instead of $\mathcal{O}(M_W)$. For *inclusive* B decays, this additional factorization of short-distance effects at $\mathcal{O}(m_b)$ from long distance effects at $\mathcal{O}(\Lambda_{QCD})$ is achieved using the Heavy Quark Expansion (HQE). However, HQE does not work for *exclusive* B decays. In particular, *nonleptonic* exclusive B decays have generally proven to be the class of B decays most resistant to any systematic theoretical treatment.

For a long time two-body B decays have been calculated in the "factorization" or "naive factorization" model-based approach. Models are generally used only as a last resort, because they are not controlled approximations to QCD, but arbitrary *replacements* of the original theory with another which — hopefully — is motivated by some physical picture and at least behaves itself in the relevant limits [37]. But because theorists could not find a way to treat exclusive nonleptonic decays systematically, they were forced to resort to naive factorization. In this method, the 4-quark operator separates, or *factorizes*, into two factors of matrix elements of bilinear currents. For example, for B decays to two mesons $B \to M_1 M_2$ (the most common application), factorization gives

$$\begin{aligned}
&\langle M_1 M_2 | Q_i | B \rangle = \langle M_1 M_2 | (\bar{q}_1 q_2)_{V-A} (\bar{q}_3 q_4)_{V-A} | B \rangle \\
&\longrightarrow \langle M_1 | (\bar{q}_1 q_2)_{V-A} | B \rangle \langle M_2 | (\bar{q}_3 q_4)_{V-A} | 0 \rangle \propto f_{M_2} F^{B \to M_1}.
\end{aligned} \tag{2.9}$$

In other words, you take the operator Q_i, which has the generic form $(\bar{q}_1 q_2)_{V-A}(\bar{q}_3 q_4)_{V-A}$, and put half of it in one matrix element and half in another. In terms of Feynman diagrams, you are splitting the Feynman diagram in half, and calculating each half separately. Essentially, you are treating the decay $B \to M_1 M_2$ as if it can be broken down into two processes — the decay $B \to M_1$ and the appearance of M_2 in the vacuum — without too much error. The two-quark matrix elements are proportional to physical quantities: the *form factor* $F^{B \to M_1}$ describing the overlap of the B and M_1 during the weak decay, and the *decay constant* f_{M_2} describing the single-meson state M_2. These physical quantities can often be obtained from experiments, and in any case are easier to calculate than the original matrix elements.

The main justification for this procedure comes from the *color transparency* argument [38], which says that factorization should work provided M_2 is a light color singlet. The quarks from weak B decay are produced in an environment of gluons and other quark-antiquark pairs, with which they would normally interact strongly. However, if one of the products is a light meson ($m < m_b$), it gets a high momentum (since it comes from a heavy b quark) and its quarks stay close together and speed away from all the other quarks and gluons. As long as they can remain in the color singlet state until they escape the system, the quark pair will interact with the rest of the system as a colorless meson, rather than as two colored quarks. This makes it okay, as an approximation, to neglect final-state

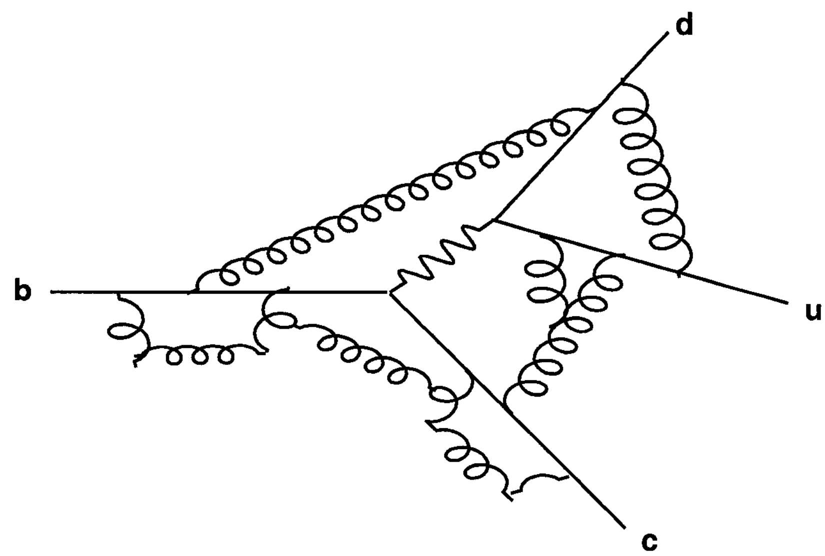

Figure 2.4: A typical nonleptonic B decay. All nonleptonic B decays receive $\mathcal{O}(\alpha_S)$ radiative corrections from gluon loops connecting the external quark lines.

interactions. That's the argument. But usually it is the desire to use factorization that motivates this argument, rather than the other way around.

However, there is a more fundamental problem with naive factorization (and other models). All nonleptonic B decays receive $\mathcal{O}(\alpha_S)$ radiative corrections from gluon loops connecting the external quark lines, as in Fig. 2.4. In QCD, these short-distance contributions to $\langle f|Q_i(\mu)|i\rangle$ would be calculated perturbatively, and the result would depend on μ, as it should. But because models are not derived from QCD, they miss the radiative corrections altogether. Now, if it were simply a question of missing some higher-order corrections, this might not be such a big deal — you'd just have to reconcile yourself to having only a leading order result. But losing the μ-dependence of the original matrix elements $\langle f|Q_i(\mu)|i\rangle$ *is* a big problem, because it means that the crucial *cancellation* of this μ-dependence in the product $C_i(\mu)\langle f|Q_i(\mu)|i\rangle$ is lost as well. (The same is true for the renormalization scheme dependence, which is an issue when the Wilson coefficients are calculated to next-to-leading order.) Since physical quantities cannot depend on the arbitrary renormalization scale μ, this is a very bad thing.

So although naive factorization is better than nothing, what is really needed is a systematic method derived from QCD, that takes into account the radiative corrections and the μ-dependence of the matrix elements. That is where QCD factorization comes in.

2.3.2 QCD factorization

QCD factorization [36] is a systematic, model-independent method for applying the heavy-quark limit $m_b \gg \Lambda_{QCD}$ to nonleptonic decays. The fact that $m_b \gg \Lambda_{QCD}$ allows for power-counting in the small expansion parameter Λ_{QCD}/m_b, making it possible to identify leading and sub-leading contributions. For example, up to power corrections in Λ_{QCD}/m_b, QCD factorization gives the following formula for a B decay to a heavy meson M_1 and a light meson M_2, $B \to M_1 M_2$ [39]:

$$\langle f|Q_i|i\rangle = \sum_j F_j^{B\to M_1}(m_2^2) \int_0^1 du\, T_{ij}^I(u)\, \Phi_{M_2}(u). \tag{2.10}$$

Here, the long-distance contributions go into the form factor $F^{B\to M_1}(m_2^2)$ for $B \to M_1$ and the *light-cone distribution amplitude* (LCDA) Φ_{M_2} of the light meson M_2. Both can be calculated using some nonperturbative technique, like lattice QCD or QCD sum rules, or they can be obtained from experiment. Meanwhile, the short-distance contributions go into the *hard-scattering kernel* T_{ij}^I, which can be calculated in perturbation theory.[6] In particular, radiative corrections from gluon loops enter as $\mathcal{O}(\alpha_S)$ corrections to T_{ij}^I. So as in naive factorization, the matrix elements $\langle f|Q_i|i\rangle$ have been separated into well-understood quantities like form factors and LCDAs. However, this time the separation is done *systematically*, and in such a way as to achieve the crucial separation of physics at long and short distances.

In the power-counting scheme, the QCD factorization formula [2.10] is valid to leading order in Λ_{QCD}/m_b. When you do the calculation, the leading term turns out to be none other than the naive factorization result. However, this time the result includes the radiative corrections as well; they appear as $\mathcal{O}(\alpha_S)$ corrections to the leading term. So QCD factorization makes it possible to systematically compute $\mathcal{O}(\alpha_S)$ corrections to naive factorization, and promotes the latter from an *ad hoc* hypothesis to a leading term in a well-defined limit. And this time, the radiative corrections are included, so the renormalization scale- and scheme- dependencies of the original matrix elements are not lost. In addition, the strong rescattering phases that plague most B decay calculations are either of $\mathcal{O}(\alpha_S)$ and calculable, or power suppressed.

QCD factorization has now replaced naive factorization as the method of choice for the calculation of exclusive two-body B decays. It has been used to study B decays to a wide variety of two-meson final states containing pions, kaons, η, or ρ mesons of all charges and types, as well as calculations for radiative B decays.

[6] Not that this time you are using $\alpha_S\ (m_b) \sim 0.20$ for the short-distance hard-scattering kernels, rather than the $\alpha_S\ (M_W) \sim 0.12$ used for the $\mathcal{O}(M_W)$ short-distance Wilson coefficients. It all depends on how short your short distance is.

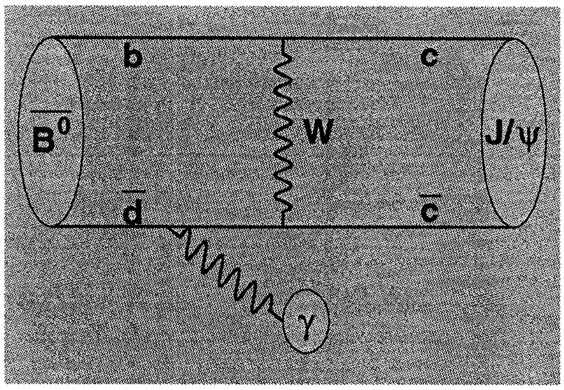

Figure 2.5: Leading order contribution to $\overline{B}^0 \to J/\psi\gamma$. As this is an annihilation diagram the decay is suppressed by a power of Λ_{QCD}/m_b.

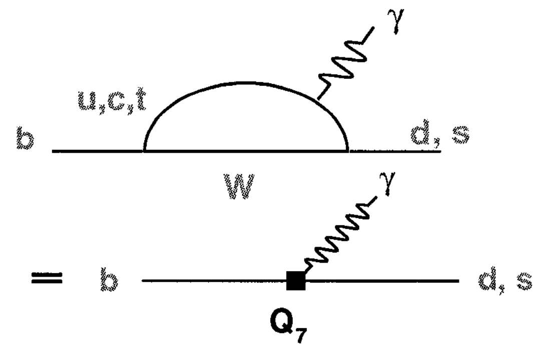

Figure 2.6: The magnetic penguin operator Q_7 for $B \to V\gamma$ decays like $B \to \rho\gamma$ and $B \to \omega\gamma$.

2.4 Predictions for $B^0 \to J/\psi\gamma$

2.4.1 About $B^0 \to J/\psi\gamma$

The leading order diagram for $\overline{B}^0 \to J/\psi\gamma$ is shown in Fig. 2.5. This is an *annihilation* diagram — that is, both the b and $\overline{d}$ quarks participate in the decay, "annihilating" to produce the final $c\overline{c}$ (J/ψ) state. In contrast, other $B \to V\gamma$ (V = vector meson) decays like $B \to K^*\gamma$, and the rarer $B \to \rho\gamma$ and $B \to \omega\gamma$ modes, proceed via the magnetic penguin operator Q_7 (Fig. 2.6) [40]. The annihilation mechanism requires that the wave functions of the two initial quarks overlap; as a result, $B^0 \to J/\psi\gamma$ is power suppressed compared to those proceeding via Q_7. Since only upper limits for $B \to \rho\gamma$ and $B \to \omega\gamma$ have been measured [41], you can see that an upper limit is probably all that it will be possible to measure for the even *smaller* $B^0 \to J/\psi\gamma$.

2.4.2 Calculation of the branching fraction for $B^0 \to J/\psi\gamma$

Recently, Lu, Wang and Yang [1] used the QCD factorization method to calculate the first-ever prediction of the branching ratio for $B^0 \to J/\psi\gamma$, with and without the radiative corrections. I will briefly review their procedure here. The leading order diagram for $B^0 \to J/\psi\gamma$ in the low-energy effective theory is shown in Fig. 2.7 (top). The star represents an effective operator Q_i. Even this leading diagram is suppressed by one order of (Λ_{QCD}/m_b) — this has to do with the fact that the J/ψ meson must be transversely polarized. Only the diagram with the photon coming from the light quark contributes at leading order — the others are suppressed by another factor of (Λ_{QCD}/m_b) and therefore are neglected. The

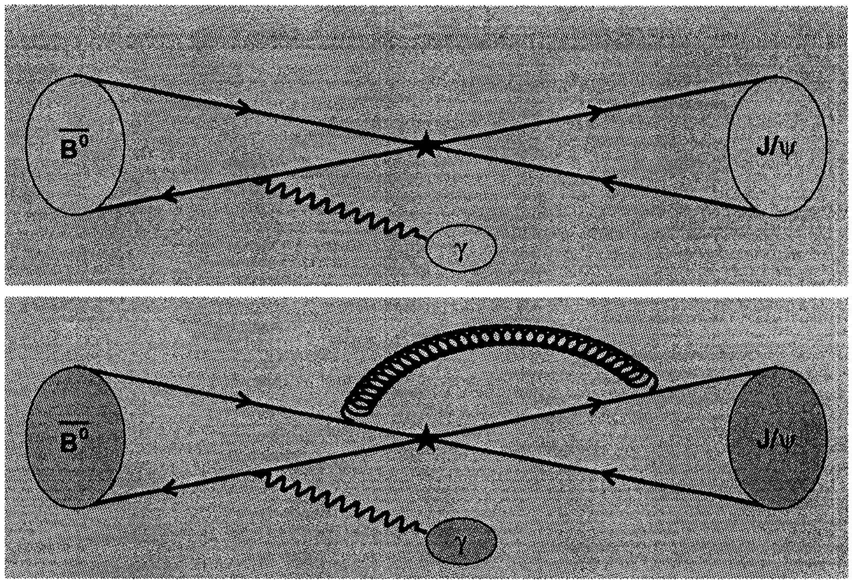

Figure 2.7: The Feynman diagrams for $B^0 \to J/\psi\gamma$. The star represents an effective operator Q_i. Top: The leading order diagram. Bottom: Typical radiative correction to the leading order process.

diagrams for the radiative corrections are the similar, but they have an added gluon loop connecting initial and final quark lines (Fig. 2.7, bottom).

Using the methods described in the previous two sections, Lu, Wang and Yang obtained an expression for the branching ratio of $B^0 \to J/\psi\gamma$ in terms of quantities that you can look up: meson masses and lifetimes, quark masses, CKM matrix elements, decay constants, the B meson's distribution amplitude (needed to calculate form factors), and Wilson coefficients at $\mu = m_b$. The resulting predictions for $B^0 \to J/\psi\gamma$ are:

$$\text{naive factorization} : \mathcal{B}(B^0 \to J/\psi\gamma) = 5.29 \times 10^{-8}$$
$$\text{QCD factorization} : \mathcal{B}(B^0 \to J/\psi\gamma) = 7.65 \times 10^{-9}$$

2.4.3 Intrinsic charm and $B^0 \to J/\psi\gamma$

The original motivation for this study of $B^0 \to J/\psi\gamma$ actually came from a comment in a paper by Susan Gardner and Stan Brodsky [4] about a phenomenon called *intrinsic charm.* In this paper, they briefly mention $B^0 \to J/\psi\gamma$ as one of the decays whose branching ratio might be enhanced by intrinsic charm.

According to the paper, the idea behind intrinsic charm is the following. The wave functions of a bound state in QCD contain *Fock states* of arbitrarily high particle number. For example, the Fock state decomposition of a B^0 is given by:

$$|B^0\rangle = \psi_{\bar{b}d}|\bar{b}d\rangle + \psi_{\bar{b}dg}|\bar{b}dg\rangle\psi_{\bar{b}dd\bar{d}}|\bar{b}dd\bar{d}\rangle + \psi_{\bar{b}dc\bar{c}}|\bar{b}dc\bar{c}\rangle + \ldots \quad (2.11)$$

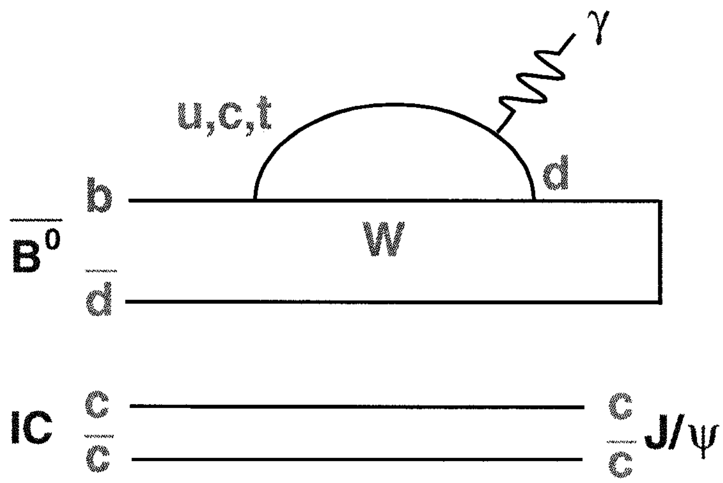

Figure 2.8: How intrinsic charm ("IC") could enhance the rate of $\overline{B}^0 \to J/\psi\,\gamma$. (The many, many gluons connecting the intrinsic $c\overline{c}$ pair to the b and $\overline{d}$ quarks are not shown.)

These Fock states are *intrinsic* to the hadron's structure and included in its nonperturbative bound state; they are to be distinguished from perturbative corrections which are *extrinsic* to the hadron's structure. In terms of Feynman diagrams, the *intrinsic* quarks are entangled by many, many gluons, whereas perturbative corrections arise from things like gluon splitting and are not multiply connected to the quarks of the bound state. Evidence for intrinsic charm has been found in the proton, and it could explain some other physics puzzles as well.

In spite of this, calculations for most B meson decays generally consider only the valence quarks, $|\bar{b}d\rangle$. More specifically, the other quarks might be included, but it is assumed that the matrix elements $\langle f|Q_i|i\rangle$ involve only the valence quarks, while any intrinsic extras appear only as spectators. If the intrinsic quarks actually participate directly in the weak decay, then this could enhance hadronic decay rates. The decays most sensitive to such an enhancement would be those for which the usual, valence-quark-only decay mechanism is suppressed. The decay $B^0 \to J/\psi\,\gamma$ certainly fits that criterion. Fig. 2.8 shows how intrinsic charm could allow $B^0 \to J/\psi\,\gamma$ to proceed via the magnetic penguin operator Q_7 as well, like other $B \to V\gamma$ decays (see the previous section). This alternative decay mechanism could enhance the $B^0 \to J/\psi\,\gamma$ decay rate (though probably not enough to make it measurable at *BABAR*).

2.4.4 Right-handed coupling factors and $B^0 \to J/\psi\,\gamma$

The coupling of quarks and leptons to the W boson contains a factor of $(1-\gamma^5)$ from the Feynman vertex. $(1-\gamma^5)$ or $(V-A)$ ("vector – axial") currents are called "left-handed," while $(1+\gamma^5)$ or $(V+A)$ currents are called "right-handed." In the Standard Model, charged currents are always pure left-handed or $(V-A)$. A right-handed charged current would therefore be a New Physics phenomenon. Such an admixture has not yet been ruled out experimentally and can therefore be sought as a possible sign of New Physics. Lu, Wang, and Yang calculate the effects of an admixture of $(V+A)$ charged current to the standard $(V-A)$ current; they find that this would result in an enhancement of the $B^0 \to J/\psi\,\gamma$ decay rate [1]. Again, though, it is highly unlikely that it would be enhanced enough to be measured at *BABAR*.

2.5 Summary

This chapter described the many theoretical tools required to calculate the $B^0 \to J/\psi\,\gamma$ decay rate. In any QCD calculation, the crucial step is to separate the perturbative short-distance contributions from the nonperturbative long-distance contributions. For B decays this is achieved using an operator product expansion to write an effective Hamiltonian in which the short-distance effects are contained in the Wilson coefficients, while the long-distance effects are contained in the matrix elements of the effective operators. The calculation of the Wilson coefficients uses renormalization group improved perturbation theory. The long-distance matrix elements require a nonpertubative technique, such as QCD factorization. The resulting prediction for the $B^0 \to J/\psi\,\gamma$ decay rate is extremely tiny, of the order of 10^{-8}. The decay rate could be enhanced somewhat by possible physics effects such as intrinsic charm or a right-handed charged current. However, even with the enhancement the decay rate is expected to be far too low to measure at *BABAR* — the best I can expect to be able to do is set an upper on the branching fraction.

Chapter 3

The *BABAR* Experiment

This chapter describes the *BABAR* experiment. It begins with an overview of B physics experiments, and explains the motivation for an asymmetric B factory. This is followed by a description of the PEP-II collider, and then the *BABAR* detector. Each subdetector of the *BABAR* detector is described in detail.

3.1 Experiments in B physics

B physics experiments are distinguished by two main features: whether they are e^+e^- or hadronic (usually $p\overline{p}$) experiments; and whether they run at the *threshold energy* for $B\overline{B}$ production or far above it. At the turn of the century B physics was dominated by three main types of experiment:

- **high-energy $p\overline{p}$ colliders** Thanks to a huge cross section for $b\overline{b}$ production ($\sigma_{b\overline{b}} \sim$ 100 nb, compared to $\sim$ 1 nb for e^+e^- colliders), hadronic colliders are able to accumulate extremely large samples of B mesons at very high rates, giving them a huge statistical advantage. In addition, the B mesons are produced at high momenta, facilitating lifetime measurements; and the final states include not only B_d but also the heavier B_s mesons and b-flavored Λ_b particles, which are inaccessible to threshold machines. Unfortunately, these advantages are offset by extremely high backgrounds from other hadronic events. $B\overline{B}$ events actually make up only a small fraction of the total signal, and the B mesons are typically surrounded by jets of other hadrons, making them difficult to reconstruct. To date the only hadronic collider to contribute much to B physics has been Fermilab's Tevatron [42], so-named because it runs at high energies of several TeV. The Tevatron's CDF and D0 experiments both have B meson working groups. To reduce backgrounds, the experiments have generally focused on B events that are relatively easy to trigger on, like B decays to charmonium, and dilepton events. The Tevatron has made important contributions to lifetime and mixing measurements, and in particular to heavy-B spectroscopy.

- **e^+e^- colliders at the $\Upsilon(4S)$ resonance** At the opposite end of the spectrum are the threshold e^+e^- colliders. Low-energy e^+e^- colliders generally provide the "cleanest" environment for $B\overline{B}$ production, with a low signal-to-background ratio and relatively isolated, easy-to-reconstruct B mesons. The main source of background is from continuum events ($e^+e^- \to q\overline{q}$), but the cross section for continuum is only about 3 times the $B\overline{B}$ cross section, and the relatively spherical topologies of B events makes them easy to distinguish from jetlike continuum events. So although hadronic experiments produce much larger overall samples of B mesons, low-energy e^+e^- experiments provide larger *useful* samples of B mesons, because theirs are not as swamped by background. The ideal choice of low energy is the $\Upsilon(4S)$ resonance energy, which is just above the threshold for $B\overline{B}$ production. A major advantage of running at this energy is that the initial and final states are completely specified. The B meson pair from $e^+e^- \to \Upsilon(4S) \to B\overline{B}$ is produced in a coherent state with $L = 1$; and each B gets half the beam energy in the center-of-mass frame. These tight constraints makes it easier to relate experimental measurements to theoretical predictions. For the first two decades following the discovery of the b quark, the CLEO experiment [43] at CESR was the world leader in B physics, studying B decays at the $\Upsilon(4S)$ resonance. All known decay modes of the B meson down to the level of 10^{-5} branching fractions were first seen by CLEO. The ARGUS experiment [44] at DESY was also very successful, and is credited with the discovery of B mixing [45]. The main disadvantage of threshold production is that the B mesons are produced with low momenta and therefore have very short decay lengths. This makes it difficult or impossible to measure B lifetimes, and to distinguish a B's production vertex from its decay vertex. Another disadvantage is that low-energy e^+e^- colliders have a low cross section for $b\overline{b}$ production ($\sim 1\,\text{nb}$). To obtain the high statistics needed for good results, e^+e^- experiments must strive for the highest possible luminosities.

- **e^+e^- colliders at the Z^0 resonance** Some of the disadvantages of threshold e^+e^- colliders can be overcome by operating at higher energies — the Z^0 resonance being the best choice. The main advantage is that unlike B mesons produced at threshold, B mesons from Z^0 decay are in motion in the center-of-mass frame, and move faster in the lab frame, giving them measurable decay lengths. Because of this, experiments like LEP and SLC provided some of the best B lifetime measurements during the first two decades of B physics. In addition, Z^0 machines provide a relatively clean environment and a relatively large cross section for $b\overline{b}$ production ($\sigma_{b\overline{b}} \sim 7\,\text{nb}$).

However, Z^0 experiments also lack many of the advantages of threshold experiments. The kinematic constraints used in reconstruction at threshold don't apply at higher energies. Another issue (also an issue at hadronic colliders) is *fragmentation* — at higher-energy colliders not only B_d are produced, but also B_s and B_c mesons and other b-flavored particles. While this is useful if you want to study B mesons other than B_d, it also means that the initial state is not as well-known as at threshold machines, so theoretical predictions require knowledge of branching ratios of the Z^0 to the many different possible final states.

B physics began with the discovery of the b quark at Fermilab in 1977 [46]. The first 20 years of B physics focused on measuring the properties of B mesons, such as lifetimes, mixing, and branching ratios for many different decays. However, measurements of both large mixing and a surprisingly long B lifetime introduced the exciting possibility of measuring *CP* violation in B meson systems. Unfortunately, none of the existing B physics experiments met the stringent requirements for *CP* violation studies, so work began immediately to find the best B experiment design for this purpose. The two main requirements for a B physics experiment to study *CP* violation were determined to be:

- **an extremely large sample of B mesons** This is needed because the branching ratios of *CP*-violating B decays are very tiny, of the order of 10^{-4} or less. For hadronic experiments, which already have a high rate of $B\overline{B}$ production, this requires improved background-fighting techniques, such as improved triggering and reconstruction methods. For e^+e^- colliders, it is the production rate itself that is the obstacle, so to improve statistics an e^+e^- experiment needs to operate at high luminosities. At the time of the studies, the largest $B\overline{B}$ samples and best performance had come from the CLEO e^+e^- experiment, operating at the $\Upsilon(4S)$ resonance. Record-high luminosities gave CLEO very high statistics; at the turn of the century CLEO had a cumulative data sample of over 17 million B meson pairs [47]. Since the hadronic experiments had not yet managed to solve their background problem, it seemed that an even higher luminosity e^+e^- collider at the $\Upsilon(4S)$ resonance, with its clean environment for B studies, would be the best option to provide the large samples of B mesons needed for *CP* violation measurements.

- **the ability to resolve the difference in decay lengths of the B and $\overline{B}$ mesons of a $B\overline{B}$ pair** This is needed because the decay length difference is a measure of *CP* asymmetry. At higher-energy $p\overline{p}$ or e^+e^- experiments, this is not a problem, because the B mesons are produced at high momentum and have measurable decay length ($\Delta z = \gamma\beta c\Delta t \sim 2300\,\mu$m at the Tevatron, $\sim 730\,\mu$m at LEP). But at threshold e^+e^-

colliders the B mesons are produced almost at rest, and so have almost zero decay length.

So the designers faced a dilemma. On the one hand, the ideal experimental design for B physics was an e^+e^- collider at the $\Upsilon(4S)$ resonance. But these experiments are unable to measure the crucial decay-length difference Δz. A solution was proposed by P. Oddone [48]: use an e^+e^- experiment at the $\Upsilon(4S)$ resonance, but with *asymmetric* beam energies. The B mesons would still be at rest in the center-of-mass frame, but they would be boosted in the lab frame; and with a sufficient boost, the decay length difference would be resolvable. This sounds like a little thing, but it was a brand new concept for high-energy physics colliders — previously, all colliding-beam experiments had been symmetric, both because this allows for higher energies (for a given beam energy) and because the collider design is much simpler. Oddone's model was adopted as the best design for a "B factory", and the *BABAR* and Belle experiments were born — their PEP-II and KEKB colliders are both very high-luminosity asymmetric e^+e^- colliders operating at the $\Upsilon(4S)$ resonance.

Today, the B factories dominate B physics. *BABAR* and Belle began taking data in 1999, and so far have been very successful. Both have far surpassed previous luminosity records and routinely log data samples of 100-200 pb^{-1} per day. They have measured the *CP*-violation parameter $\sin 2\beta$ [21] [22] and improved measurements of branching ratios for many B decays, and contributed to mixing and lifetime measurements as well. So dramatic has been their success that CLEO, the previous B physics leader, decided that there was no longer any useful role for a symmetric B collider and shifted its focus to charm physics at the charm threshold ($3-5\,\mathrm{GeV}$). Meanwhile, the Z^0 colliders, LEP and SLD, have finished their runs and been dismantled, and there are no plans for a new Z^0 collider.

So today, the undisputed leaders in B physics are *BABAR* and Belle. However, this is not the end of the story. After so many years of being hindered by huge backgrounds, hadronic colliders are finally set to take their place as world leaders in B physics. The Large Hadron Collider (LHC) at CERN is scheduled to begin running in 2007, and it is expected to be the best particle physics experiment in the world, operating at $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ luminosities and TeV energies. The main focus at LHC will be on searches for the Higgs boson and other New Physics phenomena like sypersymmetry and additional space-time dimensions. However, it will also include a B physics experiment, LHCb. And although CDF and D0 plan to bow out once LHC gets going, Fermilab has plans for its own new Tevatron B physics experiment, BTeV. Using experience gained from CDF, LHCb and BTeV will deal with the small signal-to-background ratio at hadronic B experiments with sophisticated and efficient multilevel triggering systems, and detectors with good particle identification and good momentum resolution. With $b\bar{b}$ cross sections of several hundred nb and TeV energies, a single year of data-taking (2 fb^{-1})[1] should be enough for LHCb and BTeV to obtain unprecedented 1% accuracy measurements of sin2β and of B_s mixing, as well as a few % accuracy measurements of $\sin 2\alpha$ and the as-yet unmeasured angle γ [49]. If all goes as planned, LHCb and BTeV will make *BABAR* and Belle obsolete.[2] But there are about four years before LHC and BTeV are due to begin running, and in the meantime it is certainly worthwhile to get as much out of the *BABAR* and Belle experiments as possible.

3.2 The *BABAR* experiment

The *BABAR* experiment is one of many ongoing experiments at the Stanford Linear Accelerator Center (SLAC) in Stanford, California. The primary purpose of the *BABAR* experiment is the study of *CP* asymmetries in the B meson system, particularly those that can be cleanly related to angles of the CKM triangle. However, the large data samples of relatively clean events make the *BABAR* experiment ideal for many other studies of B and non-B physics as well, including precision measurements of the CKM triangle sides, studies of rare B decays, charm physics, tau physics, and two-photon physics. The experiment began in 1999, and was designed and is expected to last until 2010, by which time the detector will have aged too much to continue.

BABAR shares the SLAC linear accelerator (linac) with other experiments, but has its own two-ring storage ring facility called PEP-II, and its own detector called the *BABAR* detector. The SLAC linac supplies the e^+e^- beams; the PEP-II B factory produces B mesons at a high rate by sending the e^+e^- beams into head-on collision; and the *BABAR* detector surrounds the collision point and records the data for the experiment.

The remainder of this chapter describes in detail the PEP-II collider and the *BABAR* detector. It is based primarily on Refs. [50].

3.3 PEP-II

The PEP-II collider is an asymmetric e^+e^- collider operating at the $\Upsilon(4S)$ resonance. Unlike a symmetric collider, in which the electron and positron beams can use the same storage ring, an asymmetric collider requires two storage rings. PEP-II consists of a High Energy Ring (HER) for the 9.0 GeV electron beam, and a Low Energy Ring (LER) for the 3.1 GeV positron beam. These beam energies result in a center-of-mass energy of 10.58 GeV (the rest mass of the $\Upsilon(4S)$); and give the lab frame a boost of $\gamma\beta = 0.56$ with respect to

[1] At LHCb, the LHC beam will be "detuned" from 10^{34} to $10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ luminosity, so that the level-1 trigger can cope. Otherwise the yearly integrated luminosity would be higher.

[2] Partly for this reason, both *BABAR* and Belle have plans for "Super*BABAR*" and "SuperBelle" experiments at even higher luminosities [$\mathcal{O}$ ($10^{36}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$)] once their current runs are over.

the lab frame. The result is a resolvable vertex separation of $\sim \gamma\beta c\Delta t = 250\,\mu$m. PEP-II stores and maintains the energies of the circulating beams until they are ready to be sent into head-on collision in the *interaction region* (IR), where the *BABAR* detector is located.

The two PEP-II storage rings are located at the end of SLAC's 3 km, high-rate linear accelerator (linac) which supplies both the electron and positron beams for the experiment. To accommodate both beams in the linac, the electron and positron beams are produced out of phase. Once the beams reach the required energies they are extracted into bypass lines and redirected by dipole magnets to the PEP-II storage rings. This process is called "injection" into the storage rings. It takes about 10-15 minutes to fill the empty storage rings; once filled, they get a 3-minute "top-off" from the linac about once per hour to compensate for beam losses (due to finite beam lifetime).

PEP-II was designed to operate at an unprecedented (at the time) luminosity of $3 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, a value determined by the need to measure *CP* asymmetries with errors at the 10% level or better. At PEP-II, high luminosity is achieved via high beam currents and a large number of bunches. This places significant demands on the interaction region. In order to avoid secondary collisions (collisions not at the interaction point) it is necessary to reorient, focus and bring the beams close together only *just* before they collide, and separate them immediately afterwards. This is a lot more difficult when the bunches are coming in at such a high rate (they are separated by about 8.4 ns on average). Nevertheless, it was accomplished, thanks to an elaborate arrangement of dipole (B1) and quadrupole (Q1) magnets placed very close to the interaction point. Having dipole magnets located only ± 21 cm from the interaction point limits the acceptance of the detector, but this was deemed a necessary evil.

Since it began running in 1999 PEP-II has reached and exceeded the $10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ luminosity mark. Thanks to PEP-II's high luminosity *BABAR* has managed to accumulate 113.1 fb^{-1} of on-resonance data as of June 2003. Overall, the PEP-II B factory has performed much better than expected, and has proven an ideal collider for the study of *CP* violation.

PEP-II backgrounds Since the $\Upsilon(4S)$ always decays to $B\overline{B}$ pairs, there is no background from resonant decay to something that isn't B mesons. The event backgrounds come mainly from *continuum* production of $u\overline{u}$, $d\overline{d}$, $s\overline{s}$ and $c\overline{c}$ directly from the e^+e^- collisions. The ratio of $\Upsilon(4S)$ to continuum $q\overline{q}$ production is about 1:3. About 10% of the data collected by *BABAR* is taken *off-resonance*, 40 MeV below the $\Upsilon(4S)$ energy, for background studies. This generally gives a better picture of continuum background than Monte Carlo simulations. Machine backgrounds are a problem, both because they contaminate the data and because they cause radiation damage to the detectors. The main source of background at PEP-II is synchrotron radiation from the dipole and quadrupole magnets, but fortunately it was possible to design the interaction region so that most of the synchrotron photons are channeled away from the detector. The Silicon Vertex Tracker, being closest to the interaction point, must be especially resilient. It receives a radiation dose of about 240 krad/year, and will have to survive a total of about 2 Mrad over the experiment's 10 year lifetime. Other sources of machine backgrounds include the interaction of beam particles with "beam gas", residual gas around the rings; and radiative bhabha scattering (e^+e^- scattering).

3.4 Requirements for the *BABAR* detector

The primary purpose of the *BABAR* detector is to measure time-dependent asymmetries in B meson decays. The constraints from this requirement come in two forms: those related directly to the *CP* asymmetry measurement, and more general constraints from the need to take full advantage of — and survive — the asymmetric, high-luminosity environment provided by PEP-II.

CP violation studies at *BABAR* center on time-dependent *CP* asymmetries that can be cleanly related to angles in the CKM triangle. As explained in Appendix A, to achieve this *BABAR* uses a particular type of B decay, in which one B meson decays to a *CP* eigenstate, and the other decays to a *tagging mode*. To achieve the required sensitivity for *CP* measurements, the objectives for the *BABAR* detector are:

- To reconstruct the relevant exclusive decays to *CP* eigenstates with high efficiency and low backgrounds,

- To tag the flavor of the other B meson with high efficiency and purity, and

- To measure the relative decay time of the two B mesons.

These requirements make the following tasks particularly important for the *BABAR* detector:

- **Pion and kaon PID** Most of the *CP* eigenstates to be reconstructed at *BABAR* have kaons and/or pions as primary or secondary decay products. Therefore hadron particle identification (PID) is a priority at *BABAR*, more so than in most particle physics experiments.

- **Lepton and kaon PID** Flavor-tagging of the B meson that decays to the tagging mode is done using charged leptons and kaons. Electrons, muons and kaons must be well identified in order to provide the required tagging efficiency and purity.

- **Excellent vertex resolution** The most important measurement at *BABAR* is the measurement of the vertex separation of the two B mesons. This must be measured with the best possible resolution.

- **Excellent tracking** Good tracking is needed to reconstruct the *CP* eigenstate decays.

In addition, to take full advantage of — and survive — PEP-II's asymmetric, high-luminosity environment, the following considerations are important for the *BABAR* detector:

- **asymmetry** As PEP-II is asymmetric, so is the *BABAR* detector. PEP-II's energy-asymmetric beams provide the boost needed to make the vertex separation of the two B mesons resolvable. A side-effect of this is that the decay products are Lorentz boosted in the forward direction — more than half of them end up in the region $\cos\theta_{lab} > 0.5$. Therefore it is more important that the detector have good coverage in the forward direction than in the backward direction. As a first step, the whole detector is offset from the interaction point by 0.37 m. In addition, as you will see, electronics and readout elements are located at the backs of their subdetectors whenever possible, so as not to get in the way.

- **acceptance** The *BABAR* experiment needs very large data samples, so it is important to avoid "losing" events to regions that are not covered by the detector. Therefore the detector should have the maximum possible acceptance (coverage), particularly in the forward direction, as just explained.

- **high rates** The event rate at a B factory is much higher than for previous e^+e^- colliders (although not unprecedented for hadronic colliders). This is necessary to provide large data samples, but it will not do any good unless the detector is able to collect the data at this rate. So the subdetectors must be able to take data at high rates with negligible deadtime.

- **machine components in the interaction region** As explained in the previous section, in order to provide high luminosity it was necessary to include the B1 dipole magnets and other machine components very close to the collision point. This limits the detector's angular acceptance to $17° < \theta_{lab} < 150°$.

- **radiation damage** Although radiation levels at the B factories are not nearly as high as levels at hadronic colliders, they are unprecedentedly high for e^+e^- experiments. Therefore possible radiation damage to the subdetectors was a concern for the designers.

- **material in the active volume of the detector** Despite the Lorentz boost, most of the decay products of the B mesons have relatively low momenta (less than 1 GeV/c). Because of this, charged track measurements are limited mainly by multiple scattering, and photon energy measurements are severely limited by material in front of the calorimeter. This makes it particularly important to minimize the amount of material in the active volume of the detector.

3.5 The *BABAR* detector

The *BABAR* detector is shown in Fig. 3.1. Like most particle detectors, it is made up of subdetectors, arranged in layers outward from the interaction point. These are listed here; from inside to outside:

1. The Silicon Vertex Tracker (SVT) provides position measurements, and is the only tracking device for those low-energy particles that don't make it to the Drift Chamber. It also provides $\mathrm{d}E/\mathrm{d}x$ measurements for particle identification (PID).

2. The Drift Chamber (DCH) provides the main tracking and momentum measurements for charged particles. It also provides $\mathrm{d}E/\mathrm{d}x$ measurements for PID.

3. The Detector of Internally Reflected Cerenkov light (DIRC) is used for PID, especially for K/π separation.

4. The Electromagnetic Calorimeter (EMC) measures the energy of electromagnetic particles — mostly photons and electrons — for PID.

5. A superconducting coil provides an axial magnetic field of 1.5 T.

6. An Instrumented Flux Return (IFR) serves both as the flux return for the magnet, and as a muon and hadron detector.

The *BABAR* coordinate system is a right-handed coordinate system such that:

- The +z axis is parallel to the magnetic field of the solenoid and points in the direction of the high energy electron beam.
- The +y axis points vertically upward.
- The +x axis points horizontally, away from the center of the PEP-II ring.
- The origin (0,0,0) is defined as the nominal interaction point.

The rest of this chapter describes in detail the subdetectors of the *BABAR* detector.

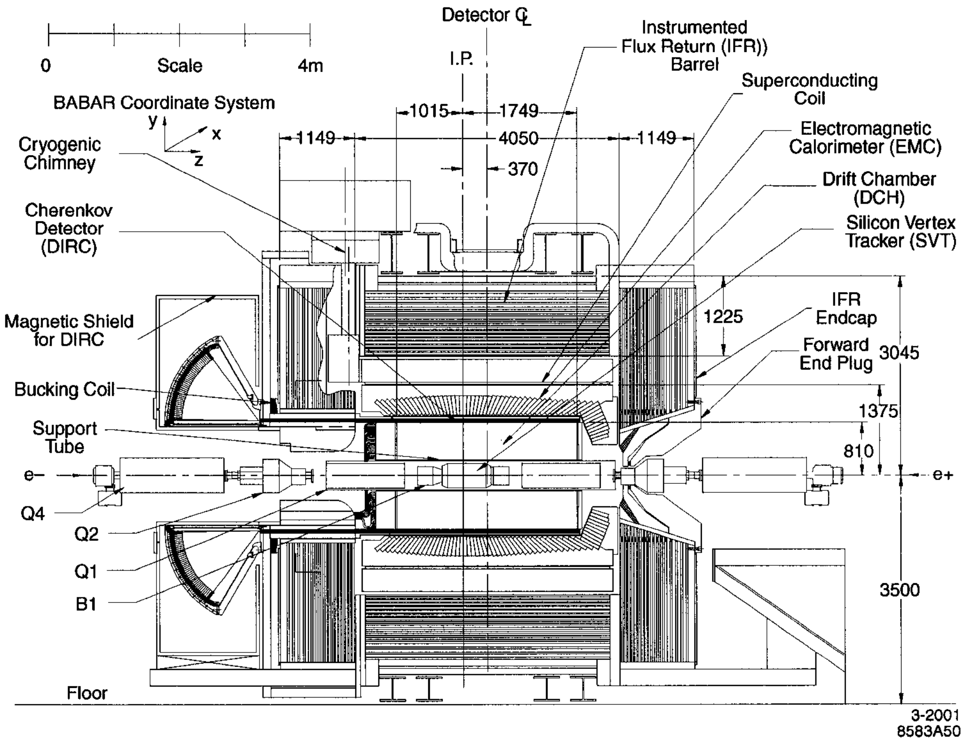

Figure 3.1: The *BABAR* Detector.

Tracking Devices

Tracking devices are devices that measure the positions of charged particles. The basic idea is that a charged particle passing through the detector can ionize the atoms of the detector material. Therefore it leaves a *charged track*, a trail of individual ionization events which can be joined in a "connect-the-dots" fashion to reconstruct the particle's trajectory. Furthermore, if the detector is enclosed in a magnetic field, then the resulting curvature of the charged tracks can be used to determine the particle's momentum and charge. Thus, tracking devices are crucial in that they provide two fundamental measurements: position and momentum. In addition, charged particles lose energy as they ionize, and measurements of the *ionization energy loss* $\mathrm{d}E/\mathrm{d}x$ can be used for particle identification.

Most modern particle detectors have two main tracking devices as their inner layers. First, there is an small, expensive and extremely precise *vertex detector*, which provides position measurements very close to the interaction point. Next, there is a much larger and more general-purpose tracking chamber, which is used to reconstruct the trajectories of the many charged particles in a physics event. The *BABAR* detector has a Silicon Vertex Tracker (SVT) as its innermost detector, and a Drift Chamber (DCH) as its main tracking chamber. These are described in more detail in the following sections.

3.6 The Silicon Vertex Tracker (SVT)

Vertex detectors

Vertex detectors — and in particular, silicon vertex detectors — are becoming more and more common in particle physics experiments because they are capable of extremely precise position measurements. A vertex detector is a detector used to measure particle tracks very close to the interaction point. It is usually a semiconducting device. A silicon vertex detector is actually an array of many smaller detectors called silicon microstrip detectors. These are semiconducting devices, asymmetric p-n junctions reverse-biased to work as detectors. In a silicon detector used for tracking, the particle "tracks" are made up of electron-positron pairs produced when a charged particle passes through the semiconducting material. The number of pairs produced is proportional to the energy loss of the particle. An applied electric field causes the pairs to separate and drift to their respective electrodes, where the charge collected is translated into a signal. The microstrips are arranged in layers, and the detector records the precise position at which particles traverse each layer. The advantages of a silicon vertex detector are those of any semiconducting device: extremely good position resolution and two-track separation, a low response time, and low bias voltages.

The Silicon Vertex Tracker

The main task of *BABAR*'s Silicon Vertex Tracker (SVT) is to reconstruct the decay vertices of the B mesons in a $B\overline{B}$ pair, in order to determine the distance (and thus the time) between the two decays. Therefore the SVT needs good vertex resolution. To keep the reduction in precision of *CP* asymmetries below 10%, the SVT must be able to measure the vertex separation of the two B mesons with a minimum precision better than half the mean separation of about $250\,\mu$m, corresponding to a single vertex resolution of about 80 μm. Lower resolution is even better, as it helps with pattern recognition, vertex reconstruction, and background rejection. Modern vertex detectors have resolutions much lower than 80 μm — so the minimum achievable vertex resolution is set, not by the intrinsic detector resolution, but by multiple scattering: 10-15 μm for the inner layers, and 30-40 μm for the outer layers. The SVT was designed to have a resolution at least as low as these limits from multiple scattering.

In addition to vertexing, the SVT also acts as a tracking device, providing precise information on the position and direction of charged particle tracks near the interaction point. To cope with the high backgrounds from PEP-II the SVT must have extremely high tracking efficiency. The best tracking measurements are the combined measurements from the SVT and the DCH. However, the SVT also provides the sole tracking measurements for particles with insufficient transverse[3] momentum p_t to reach and traverse the drift chamber ($p_t \leq 100\,\mathrm{MeV}/c$), so it must be capable of stand-alone tracking for these particles. In addition, the SVT dominates the determination of track angles, and provides $\mathrm{d}E/\mathrm{d}x$ measurements for PID.

The basic technology is double-sided microstrip detectors made of 300 μm thick, high-resistivity silicon. Microstrip detectors are ideal due to their high precision for position measurements and ability to tolerate high background levels. Double-sided detectors were chosen because they provide precise vertexing in z and ϕ while minimizing the material between measurements. Each side of the silicon wafers is divided into strips, and the strips on one side are perpendicular to the strips on the other side, for z and ϕ position measurements. All of the strips are biased, but some of the strips are not connected to the readout electronics. The purpose of these "floating strips" (as in floating voltage) is to improve the resolution by increasing capacitive charge sharing — that is, they allow the charge to spread out more. This is especially important for particles coming in at large incident angles, for which the resolution is lower.

The SVT consists of 340 of these double-sided microstrip detectors arranged in five concentric cylindrical layers (Fig. 3.2). It provides five (z, ϕ) measurements for each particle, one measurement per layer. The purpose of the inner 2 layers is to measure the track angle, which is best done as close as possible to the interaction point to minimize the effects of multiple scattering on the measurement. The outer 2 layers are mainly for linking tracks in the SVT to tracks in the DCH, and layer 3 helps with pattern recognition. 5 layers provide redundant measurements, for reliability — if one of the layers were to fail, the SVT could still work (albeit not as well). The SVT extends from 3 cm (near the interaction point) to 20 cm (the inner radius of PEP-II's support tube).

The SVT has performed very well so far, and shows relatively little radiation damage. It has a hit-finding efficiency of about 97%, which is very good. Single point resolutions for tracks originating from the interaction point are about 20 μm in both z and ϕ for the inner 3 layers, and 40 μm in z and 20 μm in ϕ for the outer 2 layers. This is more than adequate to resolve the vertex separation of the two B mesons of a B meson pair.

3.7 The Drift Chamber (DCH)

Drift chambers

The main tracking chamber of a particle detector is very often a *drift chamber*. A drift chamber is a type of *wire chamber*, which in turn is a type of *gaseous ionization detector*.

[3]In colliders, "transverse" means in the plane perpendicular to the beam direction.

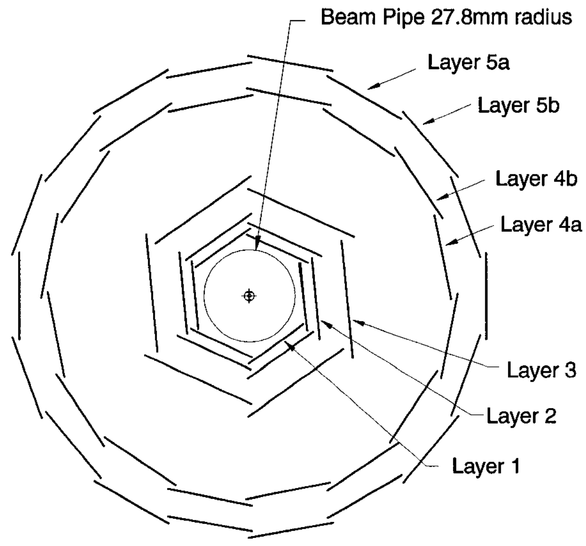

Figure 3.2: The Silicon Vertex Tracker's 5-layer arrangement.

Let's take these words one at a time:

- Gaseous ionization detectors are gas chambers under an applied voltage. Charged particles passing through the chamber ionize the gas and therefore leave ionization tracks. At lower voltages, the signal comes only from this *primary ionization*; however, at higher voltages ion collisions cause *secondary ionizations*, and the resulting *avalanches* of charge amplify the signal. Drift chambers operate in *proportional mode*, meaning that there are avalanches, but the signal is still proportional to the primary ionization.

- A wire chamber is a type of gaseous ionization detector in which the signal is picked up by anode wires called *sense wires*, each of which acts as an individual detector. Cathode wires called *field wires* maintain the near-constant electric field required to make the ionization electrons drift to the anode wires. It is only right next to the anodes that the electric field is very high; so it is near the anodes that the arriving electrons set off avalanches.

- A drift chamber is a special type of wire chamber in which track positions are deduced from the time it takes the ionization electrons to drift to the anode wire. This allows for wider wire spacing and actually obtains better spatial resolution than traditional multiwire proportional chambers (MWPCs), which require many closely-spaced wires for good spatial resolution. Since fewer channels have to be equipped with electronics, drift chambers are less expensive than traditional MWPCs. In order to translate good time resolution into good spatial resolution, you need a predictable electron drift velocity. The simplest and most reliable *time-to-distance relations* are for an electron in a uniform electric field, so a drift chamber needs many field wires to keep the field uniform. A typical drift cell includes one sense wire and several field wires.

The Drift Chamber (DCH)

The Drift Chamber (DCH) is *BABAR*'s main tracking device. Tracking is a fundamental task for any particle detector, and it goes without saying that the main tracking device must provide excellent position and momentum resolution. At *BABAR*, good tracking is important for the efficient reconstruction of exclusive B decays, which tend to have many multiply charged decay products. A spatial resolution of 140 μm, and a momentum resolution of $\sigma_{p_t}/p_t \simeq 0.3\%$ were deemed adequate for this important task, giving an average tracking efficiency of better than 98% (which is very good).

In addition to basic tracking, the DCH has several other important functions. Vertices from B daughter decays often occur in the DCH rather than the SVT, so the DCH must be able to reconstruct these secondary vertices. The DCH also provides a charged track trigger, one of the principal triggers for the experiment (the other is from the EMC). Finally, like the SVT the DCH supplies $\mathrm{d}E/\mathrm{d}x$ measurements to help with PID in the regimes not covered by the DIRC: low momenta $p_t < 700\,\mathrm{MeV}/c$, and the backward direction.

The DCH designers took particular care to maximize the detector acceptance. Acceptance was a concern because of studies that showed that even small decreases in acceptance would lead to unacceptably low efficiencies for B reconstruction, due to "missing tracks" — tracks outside of the acceptance. Because of the B1 dipole magnets, the DCH unfortunately cannot have the full 180° acceptance, but the designers made sure that the DCH acceptance would not be any lower than the limit set by the magnets, $17^\circ < \theta_{lab} < 150^\circ$.

Another major concern for the designers was to reduce the material in the active detector volume, in order to minimize multiple scattering and to avoid degrading the performance of the DIRC and the EMC. Multiple scattering is a concern because for low-momentum particles — like most B decay products — the limits on the momentum resolution come from multiple scattering. The material in the drift chamber also interferes with particles trying to reach the DIRC and EMC, thus degrading the performance of these detectors. To minimize these effects the DCH was built using low-mass, low-Z materials, including aluminum wires and a helium-based gas. Also, the inner and outer walls and endplates were made thin so that particles can make it relatively unhindered to the next subdetector, the DIRC.

The DCH is a 280 cm long cylindrical detector, extending from the support tube at a radius of 23.6 cm to the DIRC at 80.9 cm. It is made of 7104 small hexagonal drift cells with typical dimensions $1.2 \times 1.8\,\mathrm{cm}^2$. Each cell has a gold-plated tungsten sense wire surrounded

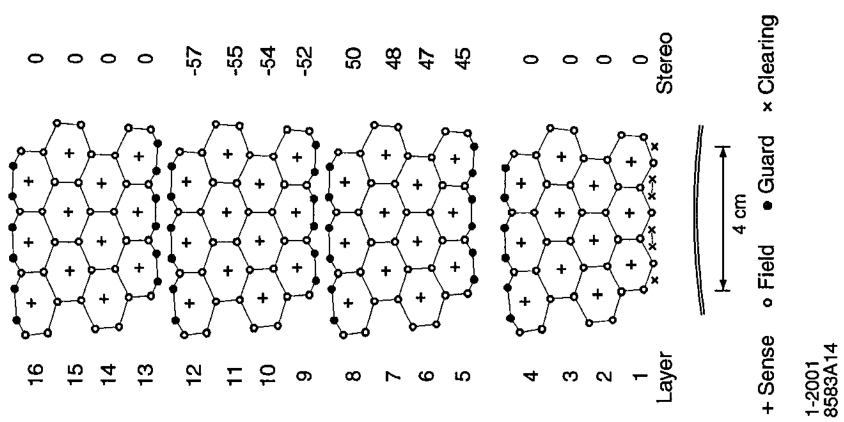

Figure 3.3: The first four superlayers in the Drift Chamber. The 4-layer superlayers allow for 3-out-of-4 majority logic for the trigger and for track segment finding.

by 6 gold-plated aluminum field wires. Both tungsten and aluminum have relatively large radiation lengths, and the gold plating prevents oxidation of the wires. The sense wires provide a voltage (1900-1960 V), giving an avalanche gain of order $\sim 10^4$, while the field wires maintain and shape the electric field.

The cells are arranged into 10 superlayers of 4 layers each, for a total of 40 layers. So the DCH provides 40 spatial and $\mathrm{d}E/\mathrm{d}x$ measurements, one per layer (unless, of course, the particle decays before reaching the end of the DCH). Figure 3.3 shows the first four superlayers. The grouping of the layers into superlayers allows for 3-out-of-4 majority logic for the trigger and for track segment finding (see Section 4.5.2). The superlayers alternate in orientation: first axial (A), then a small positive stereo angle (U), and then a small negative stereo angle (V); the pattern from inside to outside is "AUVAUVAUVA." The stereo superlayers are used to obtain z measurements.

The choice of gas was driven by the need to reduce the total amount of material, minimize multiple scattering, and operate efficiently in a 1.5 T magnetic field. The DCH gas is an 80:20 mixture of helium and hydrocarbon gas. It has a long radiation length ($X_0 = 807\,\mathrm{m}$) and provides good spatial resolution. It also provides a $\mathrm{d}E/\mathrm{d}x$ resolution of 7% and good K/π separation up to $700\,\mathrm{MeV}/c$, above which the DIRC takes over as the main PID device.

Precise and accurate timing information is crucial for good performance of the DCH. The time-to-distance relation, needed for position measurements, is determined from bhabha scattering and dimuon events. The relationship is monitored and calibrated offline. The DCH also provides event time information. The drift time is measured by a TDC; the charge deposit in each cell is measured separately by an ADC.

Calorimeters

A calorimeter is a detector that measures the energy and position of a particle by absorbing it. The particle's interactions with the absorber material generate secondary particles which in turn generate further particles, and so on, so that a *shower* develops. The shower is mostly in the direction of the particle, but has some transverse component due to multiple scattering. The energy of the original particle is distributed among the secondaries making up the shower, so that as the number of particles increases, the energy per particle decreases, until the shower stops. The shower particles constitute a signal, and the characteristics of the shower are used to locate and identify the original particles. Detection by total absorption is an example of a *destructive* measurement, in which the nature of the particle is changed by the detector. This is why calorimeters are the outermost subdetectors — that way, a particle is not destroyed until the other detectors are finished with it, since it reaches the calorimeters last.

Calorimeters are very important for several reasons. Neutral particles do not leave charged tracks, but they can shower, so calorimeters are able to detect neutral particles. Also, the signal produced is very fast ($\sim 10 - 100$ ns), which is useful in a high-rate experiment. Finally, unlike tracking devices which work better at low energies, calorimeters have better resolution at higher energies — certainly an advantage when studying high-energy collisions.

Calorimeters are classified as either *homogeneous* or *sampling* calorimeters. In a homogeneous calorimeter, the absorber and detector are one and the same material. In a sampling calorimeter, there are separate layers of absorber and detector. So for sampling calorimeters the designer has greater freedom to choose the absorber and detector, as each material need only be able to perform one function. However, the showers in sampling calorimeters are discontinuous (they stop in the detector layers, then restart in the absorber layers), so sampling calorimeters typically have poor energy resolution and are used primarily for position measurements.

In most particle detectors, the outer two layers are an *electromagnetic* calorimeter, and then (moving outwards) a *hadronic* calorimeter. Incident electromagnetic particles — mostly electrons and photons — are totally absorbed in the electromagnetic calorimeter. In electromagnetic showers, electrons radiate photons, which pair-produce electrons, which radiate photons, and so on; so that bremsstrahlung and pair production are the dominant energy-loss mechanisms. The showering stops once the shower particles reach the *critical energy*, at which the dominant energy-loss mechanism for the electrons is no longer bremsstrahlung but ionization. The electromagnetic shower length scales with the radiation length X_0, while its width scales with the Molière radius r_M. Incident hadrons, on the other hand, may start showering in the electromagnetic calorimeter, but are usually fully absorbed only in the outer hadronic calorimeter. Hadronic showers consist mostly of pions and nucleons, and in general hadronic shower shapes are highly variable. Hadronic calorimeters also tend to have worse energy resolution than electromagnetic calorimeters, because many of the shower secondaries are neutral pions which are "lost" to photon decay ($\pi^0 \to \gamma\gamma$), and because a considerable fraction of the energy is converted into breakup of nuclei and therefore is not detectable.

Calorimeters are used to detect and identify electrons, photons, and hadrons (and particles that decay to them) based on the properties of their showers. In addition, calorimeters can provide signatures for particles that are not absorbed: muons and neutrinos. Muons are identified as particles which leave ionization tracks in the tracking devices (ie, are charged), but do not shower. Neutrinos leave no signal in a calorimeter, but their existence can be inferred: if the detectable particles are missing total energy and momentum that ought to be there, then a neutrino has probably carried it off.

The outermost layers of the *BABAR* detector are the homogeneous Electromagnetic Calorimeter (EMC), and a sampling hadronic calorimeter called the Instrumented Flux Return (IFR). These are discussed in the following sections.

3.8 The Electromagnetic Calorimeter (EMC)

BABAR's Electromagnetic Calorimeter (EMC) is used to identify electrons, photons, and to a lesser extent muons; as well as neutral pions and η particles from $\pi^0 \to \gamma\gamma$ and $\eta \to \gamma\gamma$. Very good energy and angular resolutions are required to provide the high signal-to-background ratio needed to reconstruct *CP* eigenstates with their tiny branching fractions.

The focus is on detection of low-energy photons from π^0 decays, as this is the dominant source of photons produced in $B\overline{B}$ events. However, higher energies are also important. In general, B meson reconstruction requires efficient photon detection and good angular and energy resolutions in the energy range 0.02-5.0 GeV; while for other physics processes such as bhabha scattering, τ decays, or two-photon events, the EMC must be able to operate at energies as high as the beam energies. Therefore the EMC was designed to operate over the full energy range, from 20 MeV up to the kinematic limit at 9 GeV.

In addition to photon detection, the EMC is also the main source of information for electron identification. Furthermore, the EMC has to be able to operate in a high radiation, high magnetic field environment.

Low-energy e^+e^- experiments like *BABAR* typically use cesium iodide (CsI) calorimeters, as these are sensitive to photons in the 10-20 MeV energy range. A CsI calorimeter is a scintillating crystal calorimeter — a homogeneous calorimeter made of high-Z inorganic scintillating material. *BABAR*'s EMC is made of thallium-doped cesium iodide, CsI(Tl), chosen for its high light yield and small Molière radius, which give excellent energy and angular resolutions, respectively. Doping with thallium improves the light yield (and thus the energy resolution): CsI(Tl) has 50-60 thousand photons/ MeV, compared to 2-10 thousand photons/ MeV for undoped CsI. Furthermore, the photon spectrum for CsI(Tl) peaks at 560 nm, making it possible to use silicon photodiodes for readout. The other useful properties of CsI(Tl) are also those of CsI: a low Molière radius of $r_M = 3.8\,\mathrm{cm}$ makes for good angular resolution, and a low radiation length of $X_0 = 1.86\,\mathrm{cm}$ gives complete shower containment within a compact design. CsI has good mechanical and thermal stability, and is *radiation hard*, so it can cope with PEP-II's ~ 1.5 krad/year radiation dose. The design for *BABAR*'s EMC was based on CLEO-II's very good CsI(Tl) calorimeter, but aimed for even better resolution.

The EMC design is asymmetric, consisting of a barrel and a cone-shaped forward endcap. As there is no backward endcap, rear leakage is a problem, particularly at high energies where it is the main degrader of the calorimeter resolution. However, the calorimeter is very expensive, and the financial cost of including a backward endcap outweighed the benefits. Another way in which the EMC is asymmetric is that like the SVT, DCH, and DIRC, the EMC has its cooling, cables and services located in the backward region, so as not to get in the way of forward particles.

The barrel and forward endcap are made up of a total of 6580 CsI(Tl) crystals. The majority (5760) of the crystals are located in the barrel, which has 48 rows in θ, each containing 120 crystals in ϕ. The barrel compartments are organized into modules of 3 crystals in ϕ and 7 crystals in θ. The remaining 820 crystals form the forward endcap, with 8 rings in θ, and a 9th ring close to the beam which is made of lead in order to reduce the beam background effects on the endcap crystals. The endcap modules contain 41 crystals each.

The geometry of the EMC is shown in Fig. 3.4. All the crystals are arranged in a projective geometry, meaning that the inner crystal faces (dimensions $\sim 4.7 \times 4.7\,\mathrm{cm}$) have smaller area than the outer faces ($\sim 6.0 \times 6.0\,\mathrm{cm}$). Choosing dimensions of the same order as the Molière radius ($r_M = 3.8\,\mathrm{cm}$) improves the angular resolution. The average energy of the photons is different in different regions of the detector (forward particles are more energetic), so to keep the volume down and thus reduce cost the radial crystal lengths in the barrel vary with the polar angle to match the average energy of the photons at that angle, ranging from 29.76 cm (16.1 X_0) in the backward region to 32.55 cm (17.6 X_0) at the front. In the endcaps, all the crystals have length 32.55 cm, except for the two innermost rings which are 1 X_0 shorter due to space limitations.

To avoid degrading the performance of the EMC, material in front of the calorimeter

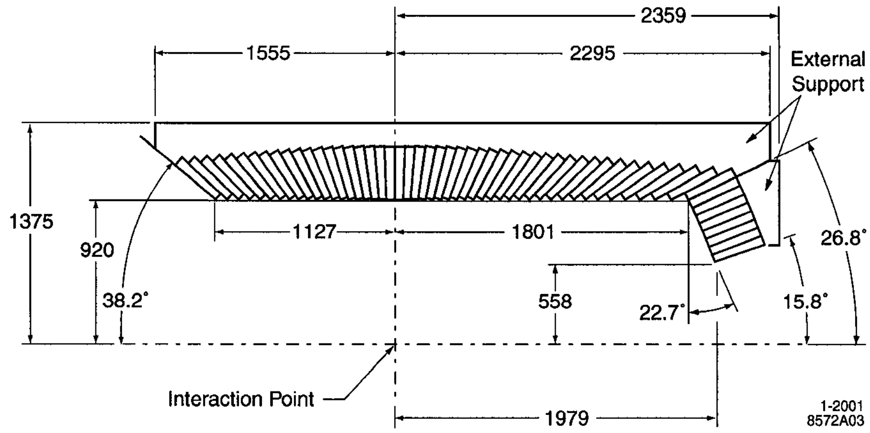

Figure 3.4: The geometry of the Electromagnetic Calorimeter.

was kept to a minimum. The total amount of material at normal incidence is about 0.25 X_0 in the barrel, and 0.20 X_0 in the endcap. The main contribution in the barrel comes from the DIRC ($\sim .14X_0$), while in the forward endcap, the endplate of the drift chamber is the main contributor ($\sim 0.065X_0$). Photons that interact in the material in front of the EMC usually convert and are lost entirely (as opposed to just losing energy and/or changing direction), so material in front of the EMC mostly degrades photon efficiency, rather than resolution.

The EMC has good resolution thanks to thanks to CsI(Tl)'s high photon yield, and also thanks to good photodiodes. However, there are many different factors that degrade the resolution. At high energies, there are fluctuations in energy loss due to leakage, particularly rear leakage. This could be reduced by including a backward endcap, but as mentioned, this was deemed too expensive. At intermediate energies, fluctuations in the transverse shower spread become important. These can never be entirely eliminated, but the effect is minimized by keeping the crystal size between 1 and 2 Molière radii ($r_M = 3.8\,\mathrm{cm}$). At low energies, noise and beam-related backgrounds have the most severe effect on EMC resolution.

The standard parametrization of energy resolution in calorimeters reflects the fact that the energy resolution of a calorimeter is determined both by statistical fluctuations inherent in the development of showers, and by instrumental and calibration limits. Since showers are statistical processes, their contribution to the resolution goes as $\sigma_E/E \propto 1/\sqrt{E}$. On the other hand, the instrumental and calibration limits are generally energy-independent and therefore contribute a constant term. Therefore, the resolution is expressed as

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt[4]{E(\mathrm{GeV})}} \oplus b. \tag{3.1}$$

Here, a is the (energy-dependent) term which comes mostly from fluctuations in photon statistics; this is the dominant contribution at low energies. b is the constant term from the instrumental limits. At *BABAR*, the energy resolution for photons is measured at low energies from radioactive source calibrations, and at high energies from bhabha scattering, to be

$$\frac{\sigma_E}{E} = \frac{(2.32 \pm 0.30)\%}{\sqrt[4]{E(\,\mathrm{GeV})}} \oplus (1.85 \pm 0.12)\%. \tag{3.2}$$

Similarly, the angular resolution is parametrized as

$$\sigma_\theta = \sigma_\phi = \frac{c}{\sqrt{E(\,\mathrm{GeV})}} \oplus d \tag{3.3}$$

where again, c is an energy-dependent term and d is a constant. The angular resolution is determined mostly by the transverse crystal size and distance from the interaction point. At *BABAR*, it is measured from π^0 and η decays to be

$$\sigma_\theta = \sigma_\phi = \frac{(3.9 \pm 0.1)\,\mathrm{mrad}}{\sqrt{E(\,\mathrm{GeV})}} + (0.00 \pm 0.04)\,\mathrm{mrad}. \tag{3.4}$$

The design goals were a = 1%, b = 1.2%, c = 3 mrad, and d = 2 mrad, so the energy resolution is slightly above the design goals, but the angular resolution is within the design goals.

The EMC is more than 96% efficient for photons with $E > 20\,\mathrm{MeV}$, where this lower limit comes from beam backgrounds and material in front of the calorimeter. The EMC has an acceptance of $15.8° < \theta_{lab} < 141.8°$ (asymmetric, as usual) and full azimuthal coverage, so that about 90% of the total solid angle is covered. (The angular acceptance is also limited by the magnets, however.)

3.9 The Instrumented Flux Return (IFR)

BABAR's outermost subdetector is the Instrumented Flux Return (IFR), a hadronic sampling calorimeter used to detect muons and neutral hadrons, especially K_L^0. About 18% of all B decays contain at least one muon in the region covered by the *BABAR* detector. Muons provide the cleanest tag of B flavor in semileptonic B decays, so to provide high B-tagging efficiency the IFR must have high efficiency for muon detection. The muons used for tagging vary widely in momentum; there are high-momentum muons ($p \geq 1.2\,\mathrm{GeV}/c$) from direct ($B \to \mu$) decays, and low-momentum muons ($\sim 0.5\,\mathrm{GeV}/c$) from indirect ($B \to D \to \mu$) decays. Therefore the IFR needs to be able to detect muons over a wide momentum range. The exception is for momenta so low that the muons do not reach the IFR in the first place; this sets the lower limits at $\sim 450\,\mathrm{MeV}/c$ in the barrel region, and $\sim 250\,\mathrm{MeV}/c$ in the endcaps.

As for K_L^0, kaons in general are important because they show up in so many *CP* eigenstates of interest. In particular, K_L^0 mesons are products in the *CP* conjugate decay of the

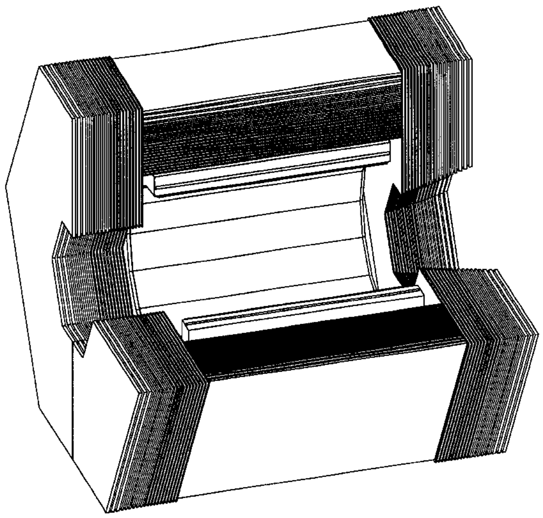

Figure 3.5: The Instrumented Flux Return. The IFR consists of alternating layers of iron absorber plates and Resistive Plate Chamber (RPC) detectors.

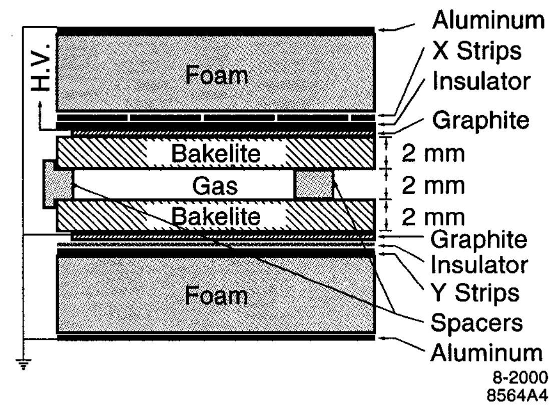

Figure 3.6: A Resistive Plate Chamber (RPC) in *BABAR*'s IFR.

most important decay in a B factory, the "golden mode" $B^0 \to J/\psi\ K_S^0$, which provides the cleanest and best measure of sin2β. The decay $B^0 \to J/\psi\ K_L^0$ contributes to the event sample for sin2β, and also serves as a systematic check on $B^0 \to J/\psi\ K_S^0$. A hadronic sampling calorimeter like the IFR has very poor energy resolution; therefore, K_L^0 detection is useful only in channels for which knowledge of the K_L^0 direction alone is sufficient to reconstruct the decay. Fortunately, $B^0 \to J/\psi\ K_L^0$ is in this category. To permit reconstruction of $B^0 \to J/\psi\ K_L^0$ with a good signal-to-background ratio, the IFR must provide good resolution of the K_L^0 direction (aka good angular resolution).

BABAR's IFR does double duty — as the flux return for the magnetic solenoid, and as a muon and neutral hadron detector. A flux return by itself is not a particle detector. But at *BABAR* the flux return is segmented (divided into layers) and instrumented (equipped with active detectors) so it can act as a particle detector as well. Specifically, the IFR is a sampling hadronic calorimeter, which as you saw means that it consists of alternating layers of absorber and detector (Fig. 3.5). The iron layers of flux return serve as the absorber, and these alternate with layers of Resistive Plate Chambers (RPCs) which serve as the active detectors. Particles shower (or not) in the iron plates, and they or their showers are detected in the RPCs.

The IFR is shown in Fig. 3.5. The basic structure of the IFR reflects its role as a flux return and as the support for the whole *BABAR* detector. A flux return is a device to contain the magnetic field; as such, it must surround the solenoid — and by extension, the entire

BABAR detector. The IFR is a large, barrel-shaped iron structure with two endcaps. It supports the weight of the entire *BABAR* detector and its stable against seismic loads. The IFR is 4.05 m long and extends radially from 1.78 m to 3.01 m. The iron barrel structure is made out of large iron plates arranged into a hexagonal structure about the beam axis. The barrel plates are rectangular, while the endcap plates are hexagonal donuts (the hole is for the beam pipe, of course). There is a gap of 15 cm between the endcaps and the barrel, resulting in a loss of solid angle of about 7%. To allow access to the detector, the endcaps are split vertically into two halves, which can be rolled apart like wheeled doors to reveal the detector.

A novel feature of *BABAR*'s IFR is the graded segmentation of the iron — the outer layers are fatter than the inner layers. Finer segmentation gives more precise measurements, improving both muon detection identification and K_L^0 detection, but it is more expensive. So the designers decided to compensate by having finer segmentation only in the region that is most important — the inner layers. The inner measurements are more important because in general, only muons get far beyond the inner layers. Therefore PID in the outer layers is practically automatic — if the particle is there, it's probably a muon. But a particle that doesn't get beyond the first layers could be one of several species, such as a low-energy muon, a pion, or a K_L^0. Monte Carlo studies showed that for a given length of iron, the efficiencies for K_L^0 identification, B-tagging with muons, and general muon identification were almost as good when the segmentation was graded as when all of the layers were very thin. So the IFR was built with an 18 layer configuration[4]; from inside to outside, the segmentation is $9 \times 2\,\mathrm{cm} + 4 \times 3\,\mathrm{cm} + 2 \times 10\,\mathrm{cm}$. This gives the same efficiency as a 30-layer, $30 \times 2\,\mathrm{cm}$ configuration, but is a lot cheaper.

The Resistive Plate Chambers (RPCs) are tracking devices. Like a drift chamber, a RPC is a gaseous ionization detector — a gas chamber under an applied voltage, in which passing charged particles leave ionization tracks. But unlike a drift chamber, which operates in proportional mode, RPCs operate in *limited streamer mode*; that is, close to the gas breakdown point. This results in visible streamers along the tracks of charged particles. RPCs differ from traditional wire chambers also in that the voltage is maintained by two high-resistivity electrodes instead of a wire. The wire's function of localizing the electric field, and so containing the avalanche, is taken over by the plates, whose high resistivity results in the prompt disappearance of the electric field around the discharge point. That way, the detector is unaffected so it can make other measurements right away. At *BABAR*, the area affected by a discharge is only about $3\,\mathrm{mm}^2$, and the deadtime is only about 10 ms — both quite low. This enables the RPCs to operate at very high rates. The streamers

[4]There have been a few upgrades since then, with some absorber layers added and some RPCs removed and/or replaced.

in streamer chambers are kept from developing in to full-blown, uncontrolled avalanches by a *quenching gas*. Streamers are fainter than avalanches, but they give better position measurements and shorter deadtimes. RPCs were chosen because they were expected to have high efficiencies and good performance at high rates.

Figure 3.6 shows a diagram of one of *BABAR*'s RPCs. The electrodes in *BABAR*'s RPCs are made of Bakelite, a phenolic resin with high resistivity ($\sim 10^{11} - 10^{12}\,\Omega\,\mathrm{cm}$). Bakelite is the original and most common material for RPC electrodes; the only alternative that has been studied much is resistive glass (that's what Belle uses). The plates are prevented from sliding, and kept separated and parallel, by a rectangular frame of PVC G-10 and a square grid of cylindrical Lectern spacers ("buttons") which are glued to the plates at 10 cm from each other. Both Lectern and G10 also have high resistivities ($\sim 10^{13}\,\Omega\,\mathrm{cm}$).

The gas in *BABAR*'s original RPCs was a mixture of the inert gas argon and two quenching gases, Freon and isobutane, in the ratio 45/50.2/4.8. The advantage of argon over other inert gases is that argon is cheaper and has a higher specific ionization. As for the quenching gases, isobutane absorbs stray UV photons, while Freon captures the outer electrons of the avalanches. The mix at *BABAR* was chosen to have only a small amount of isobutane because isobutane is flammable. The particular Freon 134a was chosen because the traditional RPC Freon, Freon 13B1, is destructive to the ozone layer and has been banned. Since *BABAR* started running, the gas mixture has twice been slightly modified to increase the concentration of argon to make the gas less reactive.

The inner surfaces of the electrodes and spacers have a thin coating of linseed oil, which smooths the surface. Otherwise, even tiny bumps would cause sparking and breakdowns in the gas. The outer surfaces of the electrodes are painted with a layer of graphite with a high surface resistivity ($\sim 10k\Omega/\,\mathrm{cm}^2$). It is this layer that is actually connected to the high voltage and supplies the voltage to the electrodes. The graphite paint covers the entire surface of the Bakelite to ensure that the voltage is evenly distributed. The high surface resistivity of the paint is important because it makes the graphite layer transparent to the electric signal sent from the chamber to the aluminum readout strips. The aluminum readout strips are glued onto the outside of the detector, and they run in two orthogonal directions to provide z and ϕ measurements. The aluminum strips are separated and insulated from the high voltage graphite strips by two insulating layers — first a thin $100\,\mu$m Mylar strip, and then a 4 mm thick foam layer. The signal comes from the movement of ions in the streamers, which generates an electric pulse by induction. The induced charge is of the order of 100 pC, and the pulse has a rise time of 2 ns and lasts for 10 ns.

For several years the performance of the IFR was severely compromised by high inefficiency rates in the RPCs. Things started off well, with 75% of the RPCs having muon efficiencies higher than 90% at the beginning of the experiment; but the efficiency began immediately to decline. It appears that in many of the RPCs, the linseed oil did not cure properly and leaked into the active region, creating nonuniformities and even short circuits which messed up the voltage across the RPCs. The problem was partially resolved during the 2002 detector upgrade, when RPCs in the endcaps were replaced with new ones with better quality control and less linseed oil. These new RPCs appear to be working with little loss of efficiency. However, overall the RPC technology has proven very troublesome, and the plan is to replace the barrel RPCs with Limited Streamer Tubes (LSTs) (another type of active detector) this summer (2004). Hopefully the new LSTs will work without any problems all the way to the end of the *BABAR* experiment in 2010.

Dedicated particle identification (PID) devices

The primary distinguishing feature of a particle is its mass. The mass of a particle can be determined from two quantities: its momentum and energy ($E^2 = p^2c^2 + m^2c^4$), or its momentum and speed ($m = p/\gamma\beta$). For charged particles, you have the momentum from the main tracking chamber; so to measure a charged particle's mass, it remains to measure either energy or speed. Energy measurements can be obtained in the calorimeters and in the tracking devices via the ionization energy loss $\mathrm{d}E/\mathrm{d}x$. However, detectors may also include a *dedicated particle identification device* designed specifically to provide energy or speed measurements for charged particles. If they are included, dedicated PID devices are usually located outside the tracking devices (so as not to interfere with the crucial tracking measurements) but inside the calorimeters (because the particles are absorbed in the calorimeters).

A Cerenkov detector is one of the most practical devices for charged hadron PID over a wide momentum range, typically 1-150 GeV/c. The basic principal behind a Cerenkov detector is that when a charged particle travels through a medium at a speed greater than the speed of light in that medium ($\beta > 1/n$), radiation is emitted at the *Cerenkov angle* θ_c, where $\cos\theta_c = 1/\beta n$. Thus, the Cerenkov angle is a measure of the particle speed. With momentum measurements from the tracking chambers, the mass can then be calculated: $m = p/\gamma\beta = p\sqrt{n_{quartz}^2\cos^2\theta_c - 1}$.

Cerenkov detectors can provide very good PID for charged particles; however, they compromise the operation of other detector elements, taking up space that could otherwise have been used for a larger tracking chamber, and putting extra material in front of the calorimeter. Because of these disadvantages, most general-purpose experiments get their PID information from the trackers and calorimeters, and do not include dedicated PID devices. Cerenkov counters are used only in experiments in which hadron identification is one of the top priorities. However, this is definitely the case for B physics experiments, all

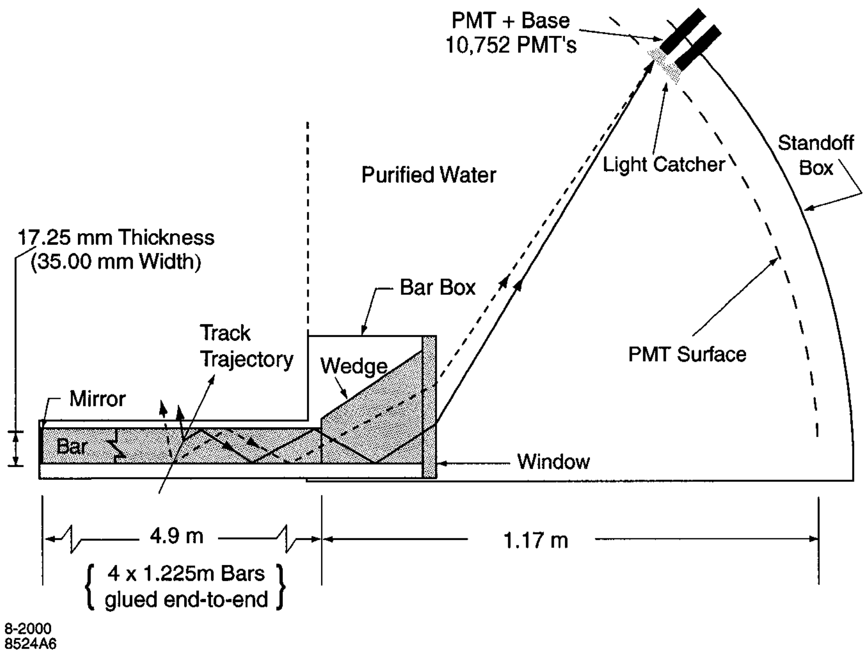

Figure 3.7: The Detector of Internally Reflected Cerenkov Radiation (DIRC).

of which use Cerenkov detectors of some sort.

3.10 The Detector of Internally Reflected Cerenkov Radiation (DIRC)

At *BABAR*, the ability to accurately identify and distinguish between kaons and pions is crucial both for B-flavor tagging and for the reconstruction of *CP*-violating B decays, which usually include pions and/or kaons as decay products. Therefore *BABAR* includes a dedicated PID device, a ring-imaging Cerenkov counter called the Detector of Internally Reflected Cerenkov radiation (DIRC). The required momentum coverage of the DIRC is set on the low end by kaon tagging for time-dependent asymmetry measurements, where the typical momentum is less than $1\,\mathrm{GeV}/c$; and on the high end by the need for good K/π separation to distinguish $B^0 \to \pi^+\pi^-$ from $B^0 \to K^+\pi^-$, for which the momentum can be as high as $4\,\mathrm{GeV}/c$. Therefore the DIRC must provide very good K/π separation over a wide momentum range, all the way up to $4\,\mathrm{GeV}/c$. At the same time, it must have a minimum of material so as not to degrade the operation of the calorimeter, and it must have fast timing to cope with PEP-II's high rates.

In *BABAR*'s DIRC, long, thin quartz bars serve both as radiators and as light guides. Cerenkov light is produced when charged particles pass through the quartz bars. The bars direct the Cerenkov light by internal reflection to photomultiplier tubes (PMTs) located at the back end of the detector. Before detection the image is expanded in a water-filled tank called the Standoff Box. Like in the SVT and DCH, in the DIRC all of the readout takes place at the back of the detector in order to minimize the material in the forward direction — yet another example of subdetector asymmetry.

The DIRC uses 144 rectangular quartz bars arranged into 12 "bar boxes," each of which consists of 12 bars placed side by side. Each bar box is one side of a 12-sided polygon about the beam line, a design which maximizes the acceptance, minimizes edge effects, and is simple to construct. The bars have rectangular cross section and dimensions 1.7 cm thick by 3.5 cm wide, and 4.9 m long. They have to be long in order to bring the Cerenkov light outside the tracking and magnetic field regions where the PMTs are located. Light travels through the bars via successive total internal reflections, but the Cerenkov angle is well-preserved thanks to the high optical quality and careful polishing of the quartz. Light travelling in the forward direction is redirected back toward the PMTs by mirrors located at the front end of each bar. The DIRC actually benefits from PEP-II's asymmetry — since most particles are boosted in the forward direction, they have steeper incident angles and therefore pass through more DIRC material. This means more photons are produced and trapped in the quartz bars, giving a higher light yield and thus better energy resolution.

Upon reaching the backward end of the bars, the light passes through a donut-shaped (because it surrounds the DCH) water tank called the Standoff Box, which allows the image to expand. Water is an ideal choice because it is inexpensive and transmits light very well, but most of all because it has an index of refraction similar to that of quartz, and therefore minimizes total internal reflection at the quartz/water interface. The Standoff Box extends radially from 1.2 m to 3 m.

The image is detected by an array of 10572 PMTs, which cover the back surface of the Standoff Box at an average distance of 1.2 m from the bar ends. They are operated directly in the water. Images of the Cerenkov rings are reconstructed from the location and time of the signals in the PMTs. Measurements of the diameter of the ring, combined with tracking information from the DCH, are used to determine the Cerenkov angle. The PMTs are located outside of the flux return in order to keep them out of the magnetic field. Furthermore, the Standoff Box has steel walls to keep the field out. The PMTs must and do provide high quantum efficiency, adequate gain, low noise, and low cost. They also provide good timing measurements ($\sim$ 1 ns), for use in background hit rejection, resolving ambiguities and separating hits from different tracks; and independent measurements of the Cerenkov angle based on timing information about photon angles.

The DIRC was designed to have minimal impact on EMC performance. The bars were made thin so that they represent only 0.14 X_0 of material between the DCH and the EMC, which isn't too bad. It occupies only 8 cm of radial space, so it does not steal too much space from the DCH or force a too-large radius on the (very expensive) EMC.

One of the nicest things about *BABAR*'s PID system is that the PID information from $\mathrm{d}E/\mathrm{d}x$ measurements neatly complements that from the DIRC. The DIRC covers the momentum range from $4\,\mathrm{GeV}/c$ down to $700\,\mathrm{MeV}/c$, and the SVT and DCH cover the range below $700\,\mathrm{MeV}/c$.

The DIRC has performed very well so far. It has achieved a resolution of 2.4 mrad, corresponding to a 4σ K/π separation at $3.0\,\mathrm{GeV}/c$, and a 3σ K/π separation at $4.0\,\mathrm{GeV}/c$, which is very good. The DIRC also provides muon identification below $750\,\mathrm{MeV}/c$, where the IFR is inefficient; and proton identification above $1.3\,\mathrm{GeV}/c$.

3.11 The superconducting solenoid

The superconducting solenoid is not a subdetector, but it is still important, for it provides the 1.5 T magnetic field required for momentum measurements in the DCH. It surrounds all of the subdetectors except the IFR, which serves as its flux return. A solenoid is a popular choice of magnet shape for colliders, because the field lines are axial, parallel to the beam direction, and therefore the best momentum resolution is for particles travelling in the radial direction.

The solenoid has to be superconducting to provide the 1.5 T field. If the designers had chosen a 1 T field, the magnet could have been a cheaper ordinary magnet, but this would have made the momentum resolution unacceptably worse, so they stuck with a 1.5 T field. Unfortunately, a higher field makes low-momentum particles harder to detect, as it causes them to spiral more, so that they are more likely to decay before reaching the DCH.

For good momentum resolution, the field must be kept uniform, especially in the DCH. For this reason, the current density at the two ends of the solenoid is twice that in the center. To avoid degrading the performance of the IFR, the solenoid is thin. The solenoid's backward shield is designed to accommodate the DIRC, whose quartz bars penetrate through the back of the detector to its external PMTs.

3.12 Summary

This chapter described the *BABAR* experiment. The primary aim of the *BABAR* experiment is the study of *CP* violation in B decays. This goal led to the construction of PEP-II, a high-luminosity, asymmetric e^+e^- collider at the $\Upsilon(4S)$ resonance. The asymmetric beam energies give the Lorentz boost needed for *CP* violation measurements, while the high luminosity is needed to get enough statistics. The *BABAR* detector was designed to take advantage of and accommodate PEP-II's high luminosity and asymmetry. Otherwise, it's just a typical particle detector, with the usual layers of subdetectors: tracking devices and calorimeters, as well as a Cerenkov detector for dedicated particle identification. Although CP violation is the main focus, Babar's clean, high-luminosity environment makes it ideal for many other types of studies as well, including searches for rare B decays like $B^0 \to J/\psi\,\gamma$.

Chapter 4

Data processing at *BABAR*

The previous chapter described the PEP-II collider used to produce $B\overline{B}$ events, and the *BABAR* detector used to record them. However, the raw data recorded by the detector goes through many stages of processing before it shows up as histograms in a published report or thesis. The initial filtering and reconstruction of the data are done centrally at *BABAR*. The reconstructed data is stored in a *BABAR*'s central database, the Event Store, where users can access and use it to produce their own ntuples for offline analysis. This chapter describes in more detail how data is processed at *BABAR*.

4.1 Data acquisition (DAQ)

The data acquisition (DAQ) [53] system's job is to get data from the detector to the online prompt reconstruction (OPR) system. There are two main stages of processing that the data must undergo before it can be passed to OPR. First, it must be digitized — that is, translated from raw detector signals to "computer language". Second, it must be filtered by the *trigger* system to get rid of as much background as possible, enough so that OPR can cope with the rate of incoming data. The main steps in DAQ are the following:

- **front-end electronics (FEEs): initial processing** Each subdetector has its own *front-end electronics* (FEEs), which are located right on the subdetector and receive the raw detector signals. The FEEs amplify and digitize the signals, and perform some other subdetector-specific processing. Then they send the data to the subdetector's readout modules.

- **readout modules (ROMs): primitive construction** Each subdetector also has its own *readout modules* (ROMs), located off the detector. The ROMs use the FEE data from the EMC and DCH to construct basic data objects called *primitives* and send them to the Level-1 trigger.

- **Level-1 trigger** The Level-1 (L1) trigger is a quick hardware trigger that filters out the most "obvious" background. It uses the primitives from the ROMs and other simple selection criteria to make a quick decision whether to accept or reject events.
- **feature extraction** If an event passes the L1 trigger, a signal is sent to the ROMs telling them to collect the rest of the FEE data. The readout modules then perform *feature extraction*, transforming the raw data into useful information such as particle hit time and energy. The output of feature extraction is called an *event fragment.*
- **event assembly** The DAQ system collects the event fragments from the ROMs of all of the subdetectors, and assembles them into an event for the Level-3 trigger.
- **Level-3 trigger** The Level-3 trigger performs the Trigger's final event selection. It runs in software, and uses more complex algorithms than the Level-1 trigger to decide whether to accept or reject events. Events passing the L3 trigger are sent to the online prompt reconstruction (OPR) system, which performs the main event analysis.

4.2 The Trigger

A trigger is a very important part of high-energy particle physics experiments, particularly those with very high backgrounds. The trigger's job is to act as a filter, quickly selecting interesting physics events for further analysis while rejecting background. Since it must keep up with the event rate, the trigger's selection criteria are typically very simple, based on things like event topology or track multiplicity. Many experiments are limited by *deadtime* — time when data is not being recorded because the data-taking electronics can't keep up with the event rate. A good trigger should minimize or eliminate deadtime. For experiments with particularly high rates, this usually requires a *multilevel* trigger, with each level receiving data at a lower rate and using more complex selection criteria than the previous level.

The complete reconstruction of an event by the online prompt reconstruction (OPR) [54] system is a complicated process, far more complex than the loose pseudo-reconstruction performed by the trigger system. The OPR system can accept events at a maximum rate of about 120 Hz. This is fast enough to allow for the complete reconstruction of all *interesting* events at *BABAR* — $B\bar{B}$ physics, charm physics, τ physics, and two-photon physics events occur at rates of several Hz each, for an overall rate of less than 100 Hz. However, the detector records not only the events of interest but also background events — bhabha scattering, beam-induced backgrounds from the interactions of "lost" beam particles, and cosmic muon events. These events are "not interesting" in the sense that they are already well understood, and there are no *BABAR* projects that study them. Due to its high luminosity, PEP-II has very high background rates for an e^+e^- experiment — at the design luminosity of $3 \times 10^{33}\ \mathrm{cm}^{-2}\mathrm{s}^{-1}$ (which *BABAR* has exceeded), the background rate is about 1200 Hz. Even if *BABAR* wanted to log all of these background events (which it doesn't), it is beyond OPR's capacity. *BABAR*'s trigger system must filter out most of this background in order to bring the incoming data rate down to a level that OPR can cope with.

BABAR's Trigger [55] has two levels, the Level 1 (L1) hardware trigger and the Level 3 (L3) software trigger. The purpose of the L1 trigger is solely to reduce backgrounds while remaining as "open" as possible to events of interest. The L3 trigger then uses the complete event information to select the physics events of interest to be sent for reconstruction. The maximum permissible rates for the L1 and L3 triggers are set by the maximum rate of the next system downstream. For the L1 trigger, this is the data acquisition (DAQ) system, which has a maximum rate of 2 kHz. For the L3 trigger, it is the online prompt reconstruction (OPR) system, which can reconstruct events at a maximum rate of 120 Hz.

The Level 1 (L1) trigger

The Level-1 trigger receives the raw data straight from the detector, before any processing. It consists of a drift chamber trigger (DCT), an electromagnetic calorimeter trigger (EMT), and a global trigger (GLT).[1] The EMT and DCT construct basic data objects called *primitives* from the raw hits in the subdetectors. The idea is that a primitive corresponds to a (possible) particle. The DCT primitives are long and short tracks in the DCH. The EMT primitives are clusters of crystals with energy above a certain threshold. The results from both the EMT and DCT are sent to the GLT, which decides — mostly on the basis of track and cluster multiplicities — whether to reject the event or allow it to be sent to Level 3.

The Level 3 (L3) trigger

Once the events have been assembled, there is plenty of information that can be used to filter out background events. The Level 3 trigger runs in software, and uses more complex algorithms to analyze the complete event data, combining DCT tracks and EMT clusters from L1 with the full DCH and EMC information, and using information from the other subdetectors as well. Like L1, the L3 trigger looks at event topologies and track multiplicities, and matches DCH and SVT tracks. And like L1, the L3 trigger has separate and orthogonal DCH and EMC triggers. However, it uses many other selection criteria as well. Track impact parameters from the SVT are particularly useful for rejecting beam backgrounds, by rejecting events that did not originate from the primary vertex (the re-

[1] There is also an IFR trigger (IFT), but it is used only for diagnostic purposes.

constructed collision point). Timing information is used to reject background events from other beam crossings. Backgrounds can also be reduced by matching DCH tracks to SVT tracks. The only type of background which is not reduced via topological cuts is bhabha events; these must be vetoed instead.

The L3 trigger passes various sets of *trigger lines* [55]. The physics events of interest to most users are in the physics line. However, L3 also passes diagnostic and calibration lines for diagnostic and calibration studies.

Events passing the L3 trigger are written to temporary files, called XTC (eXtended Tagged Container) files, which serve as the input to OPR. Each XTC file contains all of the events from a single run[2] of the collider, typically about 300,000 events.

4.3 *BABAR* software

BABAR software is written in a computer language called C++, which is designed to support object-oriented programming [56]. Object-oriented programming makes it easy to model real-life objects and concepts by designing *classes* to represent different types of objects. This is useful for *BABAR* because it allows for the association of things like particles and detectors with software objects. For example, the class `ChargedTracks` was created to represent charged tracks in the DCH and SVT. Furthermore, C++ allows for a hierarchy of classes, via the concept of *inheritance* — for instance, `ChargedTracks` *inherits from* the generic particle candidate class `BtaMicroCandidate`.

BABAR software for reconstruction, simulation, and ntuple production is organized in terms of *packages*, self-contained sets of code intended to perform a specific task (for example, to find calorimeter clusters). *BABAR* is constantly improving and updating its software, with regular *releases* of the most recent stable version of each package, as well as the libraries and binaries needed for particular machine architectures. Most researchers use the release designated as the *current* release; this is the most recent release to have reached an acceptable level of quality.

Releases are numbered chronologically, so that the highest numbers represent the most recent releases. Releases with the same first number X are very similar, and are often referred to as "the release X series," or just "release X." For example, the release 10 series includes releases 10.2.3b, 10.3.0f, 10.4.1a, and far too many more to list here. The differences between the releases in a series are just minor bug fixes and the like. On the other hand, releases with different first numbers are very different — data processed using a release in the release 8 series has a different format than data processed using release 10. Since the most recent releases represent the latest and greatest *BABAR* software, *BABAR* implements

[2]A run is the interval between two beam injections at PEP-II.

a system of *reprocessing* the data — completely redoing the reconstruction for all old, already-reconstructed data, but this time using the new release. Thus, you have the option of analyzing data initially processed using release 8, with the new and improved release 10 software; but not vice versa.[3]

A given analysis should use *compatible* releases of the different types of software. In practice what this means is that the releases used for reconstruction, simulation, and ntuple production should have the same first number. For example, the data set from Runs 1 and 2 was reconstructed using release 10 reconstruction software, so for Runs 1 and 2 I used release 10 ntuple production software to make my real data ntuples, and I used SP4 ("Simulation Production 4") software, which is also release 10, to make my simulated ("Monte Carlo") ntuples. Similarly, the Run 3 data set was reconstructed using release 12 reconstruction software, so for Run 3 I used release 12 ntuple production software to make my real data ntuples, and I used SP5 software, which is also release 12, to make my Monte Carlo ntuples. Using compatible releases ensures that all of the code works together properly, and also that the background conditions are the same in the real data and Monte Carlo simulated sets.

4.4 Online Prompt Reconstruction (OPR)

Reconstruction is the last stage of data processing to be performed centrally at *BABAR*. The aim is to reconstruct particle candidates from the raw hits from the detector. You have already seen the first few steps of this process, with the detector hits being used to form primitives for the trigger, and assembled into events for the L3 trigger. But it is only after an event is selected by the L3 trigger and passed to the *online prompt reconstruction* (OPR) system [54], that the true reconstruction begins.

The OPR system aims to completely reconstruct all physics events passed by the Level-3 trigger within several hours of acquisition (hence "online" and "prompt"). The need to process events quickly comes not only from the desire to obtain the latest physics results as quickly as possible, but also from the need to provide feedback for the detector operations staff so that they can fix problems as they arise. This was particularly important in the first stages of the experiment. The rate of interesting physics is around 100 Hz, most of which is passed by the L3 trigger, so the OPR system was designed to be able to accept data at rates of up to 120 Hz (with zero *deadtime*), and to finish processing it within several hours (minimizing *latency*).

Calibrating the detector is a continuous and important part of detector operations.

[3]So when people refer to "release X data," they mean data taken when X was the current release, even if this data has since been reprocessed.

BABAR uses a system of *rolling calibrations*, in which calibration information in the *Conditions Database* is continually updated and used in reconstructing events from the same time period. These conditions are also used in the production and reconstruction of simulated Monte Carlo data.

4.5 Reconstruction algorithms

To translate the raw detector hits into a description of particles and their decays, the OPR system uses reconstruction algorithms implemented in software. Reconstruction takes place in three steps. First, the hits are reconstructed into the basic objects corresponding to individual particle candidates: tracks in the tracking devices, and clusters in the calorimeters. Second, particle identification (PID) algorithms are used to assign an identity hypothesis to each particle candidate. Finally, tagging creates a database of tag bits, simple boolean or boolean-like flags for quick data skims. This section describes these steps in more detail.

4.5.1 OPR filters

There are two stages of filtering at the OPR level [55]. The `DigiFilter` is based only on the L1 and L3 trigger output, and is run before the full reconstruction. The main purpose is to select the physics line of physics events, and reject the diagnostic and calibration lines. The effect is to require that events pass either the DCH or EMC trigger.

After some initial reconstruction another filter, `BGFilter` (background filter) [55], is applied. `BGFilter` consists of selectors that tag physics events as multihadron events, τ events, two-photon events, and so on. An event must have at least one of these tags set to true in order to be written to the Event Store database; otherwise it will be available only in the XTC file. Events passing `BGFilter` are collectively called `isPhysicsEvents` events.

4.5.2 Track and cluster finding

The aim of the *BABAR* detector is to detect and identify particles; but the raw data obtained by the detector is in the form of hits in the various subdetectors, not particles or particle candidates. The first step in reconstruction is therefore to run algorithms which use the hits to form *basic objects* corresponding to particle candidates. In the tracking devices, the basic object is a *charged track*, formed by "connecting the dots" between the hits in the many layers of the DCH and SVT. In the calorimeters, where particles are identified based on their showers, the basic object is a *cluster*, a bunch of hits in the same general region of the calorimeter.

Because of the magnetic field, the charged tracks in the SVT and DCH are helical.

They would be exact helices if not for multiple scattering, energy loss in material, and inhomogeneities in the magnetic field. So the tracking algorithms are designed to look for ways to join the detector hits into near-helical tracks, taking these three effects into account. Track reconstruction begins in the DCH with the tracks and the event time (T0) estimate from the L3 trigger. The algorithm looks for track segments in the individual DCH superlayers, which consist of 4 layers each to allow for 3-out-4 majority logic in segment finding and also in triggering decisions. The algorithm then tries to piece the segments together to form tracks. The tracks are fit to the expected near-helix, and then the tracks from the DCH are extrapolated to the SVT. Finally, additional algorithms are run in the SVT to check for tracks that don't extend through to the DCH (corresponding to charged particles that decayed before they reached the DCH).

A particle showering in the EMC will typically deposit energy in several crystals. Groups of crystals containing the energy deposit from a single particle are called *clusters* or *bumps.* More specifically, a cluster is a set of adjacent crystals with the sum of their energies above a certain minimum energy. A cluster represents a single particle candidate, except in cases where two or more particles deposit energy in the same region. In these cases the clusters typically contain bumps — local energy maxima in the clusters — and it is the bumps that represent single particle candidates. So the EMC reconstruction algorithms look for clusters, and then check the clusters for bumps. The particle candidates in the IFR are also called clusters, but as the IFR does not provide energy measurements an IFR cluster is just a group of adjacent hits.

The last step is track-cluster matching. This is a search for tracks and clusters that correspond to the same particle candidate. This is determined by extrapolating the charged tracks to the EMC and DCH, and seeing if there are any clusters along these extrapolated trajectories. Clusters that are associated with charged tracks correspond to charged particles, and so are often called charged clusters. Clusters that have no associated track represent neutral particles and are called neutral clusters. Thus, track and cluster finding outputs lists of charged and neutral particle candidates, where charged candidates are track-cluster associations or just tracks, and neutral candidates are clusters that don't have associated tracks.

4.5.3 Particle identification (PID)

Once the track and cluster particle candidates have been found, the next step is to determine their identities. The only particles that are observed directly in the detector — that is, that form tracks and/or clusters — are electrons, muons, pions, kaons, photons, and protons.[4]

[4]Neutrons are seldom produced in B decays, so they are not very important at *BABAR*.

Other particles decay too quickly to be observed, so their existence and properties must be inferred from their track or cluster decay products. So the aim of particle identification is to identify each charged track and cluster as one of these six species of observable particle.

A particle is identified based on its mass and its quantum numbers; or equivalently, its mass and how it interacts. Charged particles are much easier to identify than neutral particles, since both their charges and momenta can be determined from the curvature of their tracks, and their masses can be calculated using the momentum plus either a speed or energy measurement. At *BABAR*, charged particle energy (specifically, the ionization energy loss $\mathrm{d}E/\mathrm{d}x$) is measured in the SVT and DCH, and charged particle speed can be obtained from a Cerenkov angle (θ_c) measurement in the DIRC. Neutral particles, on the other hand, must be identified based solely on their showers in the calorimeters.

At *BABAR*, particle identification is implemented via *particle identification selectors*, sets of cuts developed and maintained by the *BABAR* Particle Identification Group. There is a selector for each type of observed particle, and each has different levels of selection. These are usually called "very loose," "loose," "tight," and "very tight," with looser cuts having higher efficiency but also higher mis-identification rates. The selector algorithms run over the tracks and clusters, and assign to each one tags that indicate which selectors the candidate passed. For example, a candidate which very much resembles a muon would pass the "tight" or "very tight" muon selector and would get a `muMicroTight` or `muMicroVTight` tag.

Here is a brief overview of how particles are identified at *BABAR* [37]:

- *Electrons* are charged clusters that shower in the EMC with a characteristic electromagnetic shower shape, as parametrized by shower-shape variables like E/p and LAT.

- *Muons* are charged tracks that do not shower, and penetrate further in the IFR than any other particle.

- *Photons* are neutral clusters that shower in the EMC with a characteristic electromagnetic shower shape, as parametrized by shower-shape variables like LAT.

- *Charged pions, kaons, and protons* are identified and distinguished from each other (and from electrons and muons) using *likelihood selectors.* The discriminating variables are the ionization energy loss $\mathrm{d}E/\mathrm{d}x$ in the SVT and DCH; and at high enough momentum, the Cerenkov angle θ_c and the number of photons N_γ from the DIRC.

- *Neutral pions* π^0 are neutral clusters that decay to two photons, $\pi^0 \to \gamma\,\gamma$.

- *Neutral long-lived kaons* K_L^0 are neutral clusters that are not neutral pions, that reach the IFR but don't penetrate very far, and whose EMC and IFR showers are characteristic of a hadron. (Neutral short-lived kaons K_S^0 are too short-lived to be directly observed.)

The details for each selector can be obtained from the Particle Identification Group. As well, there will be more details about the selectors for particles relevant to the $B^0 \to J/\psi\gamma$ analysis (photons and leptons) in Chapter 5.

The output from reconstruction goes to the *BABAR*'s central database, the Event Store. The output of track and cluster finding and PID is a list of particle candidates, class `BtaMicroCandidate` in *BABAR*'s C++ notation. The `BtaMicroCandidates` are sorted into more specific lists such as `ChargedTracks`, `CalorNeutral`, and `GoodTracksLoose`. These fundamental lists are used by all *BABAR* collaborators for their analyses, and are are stored at the Micro level of the Event Store.

4.5.4 Event tagging

The last step in reconstruction is called tagging. Events are tagged with *tag bits*, which contain global information about events such as event parameters, which triggers and filters the event passed, and interesting-physics flags. The tag bits are stored at the Nano level of the Event Store. Tag bits are used to create subsets of data corresponding to specific physics processes. These subsets are called *skims*, *streams*, or *collections*.

Thus, OPR fills *BABAR*'s Event Store database with particle candidates. The lists of `BtaMicroCandidates` are stored at the Micro level, and the *tag bits* are stored at the Nano level. These two databases contain all of the information needed for a typical physics analysis.[5]

4.6 Skims, streams, and collections

The Micro database contains many different types of interesting physics events. However, the typical researcher is interested in only one or a few types of event. For example, a researcher in *BABAR*'s τ physics group will generally not need to study charmonium events. For this reason, each Analysis Working Group (AWG) defines its own selection criteria to produce its own *skims* or *streams* — subsets of the data that contain the events of interest to that AWG. The selection criteria typically involve one or several tag bits in the Nano database. For example, the Charmonium AWG defines a skim called `Jpsitoll`, which picks

[5]The Event Store also includes two other levels of storage, Raw and Reco, but these are too detailed to be useful, and are being phased out.

out events with the JpsiELoose or Psi2SELoose or JpsiMuLoose or Psi2SMuLoose tag bits. (In case you can't tell from the names, the purpose of this skim is to select $B \to J/\psi\ X$, $J/\psi \to \ell^+\ell^-$ and $B \to \psi(2S)\ X$, $\psi(2S) \to \ell^+\ell^-$ events.) For releases 10 and 12 the skim selections were run as part of the reconstruction. In release 10 there are 21 streams; for release 12 there are just 4 big streams.

The data subsets produced in skims and streams are called *collections*. Each collection is assigned a name which specifies all of the information about the collection: the stream or skim name, the software release version used to produce it, its run number, and which version it is (there is more than one version if the collection has been reprocessed).

4.7 Ntuple production

Ntuple production involves running *BABAR* code to access collections in *BABAR*'s Event Store, and storing them into a very convenient and portable data format called an *ntuple*.[6] The ntuples serve as the user's own private copy of the data, so there is no longer any need to interact with *BABAR* Central. The user can continue to work online with *BABAR* software if he desires; or he can copy the ntuples to his own computer and analyze them offline. Either way, once the ntuples are produced the user can begin independent analysis.

To produce ntuples, the user needs two things: data from the Event Store, and an executable to turn it into ntuples. To make the executable, the user needs *BABAR* ntuple production software. First the user checks out a release of ntuple production code and adds the appropriate packages. Once he has a package-filled release, the user sets the parameters for the *preselection*, the loose selection applied during the production of the ntuples. This involves making minor modifications to the code, such as the selection of particular tag bits, and the definition of the reconstruction criteria for the user's decay of interest. The preselection is to be distinguished from the final selection, which the user imposes at the ntuple level during his independent, offline analysis.

Once the code is all ready, the user links and compiles it to make the executable.[7] The executable will make ntuples, but to do this it needs data, so the user must communicate with the Event Store using a computer language called tcl ("tool command language"). This step is made somewhat easier by a *BABAR* tool called skimData, which can automatically generate tcl files to access particular collections. With the executable, and tcl files with

[6] "Ntuple production" is my term. The general term at *BABAR* seems to be "an analysis job", or just "a job", since the code is run as a "job" on a queue at SLAC or other *BABAR* servers. But "ntuple production" is a bit more descriptive, I think.

[7] This is just standard programming procedure — linking puts all of the code together to make one big code, and compiling turns it into a program or *executable* that can actually do something.

which to access the data, the user is ready to run the job. The executable takes the data and produces ntuples.[8]

The ntuple production code does two main things: first, it performs additional reconstruction; and second, it stores the results in the ntuple format. As you've seen, Micro data consists of a list of `BtaMicroCandidates`, the C++ version of tracks and clusters. The ntuple production takes these lists and determines the most probable identities and decay trees of the particle candidates. Some of the most important reconstruction tasks in a typical ntuple production job are the following:

- **nano tag filters** Ntuple production usually begins with a pass through the Nano database of tag bits. As explained in Section 4.5.4, tag bits are used for quick filtering to find the events of most interest to the user.

- **PID decisions** In Micro, particle candidates are stored in lists of `BtaMicroCandidates`, and for each `BtaMicroCandidate` there are PID tags indicating which PID selectors the candidate passed. The ntuple production code uses this information to determine the most probable identity for each particle. It takes the lists of `BtaMicroCandidates` and separates them into blocks of electron candidates, muon candidates, kaon candidates, and so on.

- **composition** Composition involves recreating the "decay trees" in the event — determining the mother and daughters of each particle.

- **vertexing** Vertexing takes decay trees and determines the most likely position of the vertex, and the most likely momenta of the particles at the vertex.

- **kinematic fitting** Kinematic fitting uses kinematic constraints to determine the best values for the momenta, mass and energies of a particle. (This is as opposed to just automatically using the measurements from the detector.)

- **calculations of useful variables** Ntuple production code also performs calculations of many quantities useful for physics analysis. This is helpful not only because it saves the user the trouble of doing it himself, but also because in many cases the information required to do the calculation is not available at the ntuple level — for example, to calculate *LAT* you need to know which crystals make up a given cluster, and this is listed in the Micro database but not in an ntuple.

The ntuple production code then stores the data in ntuples. It would be hard to overstate the usefulness of the ntuple format. Figure 4.1 shows how data is organized in a typical

[8] The executable does not affect data in the Event Store, of course — it just makes a copy of the parts needed for the ntuples.

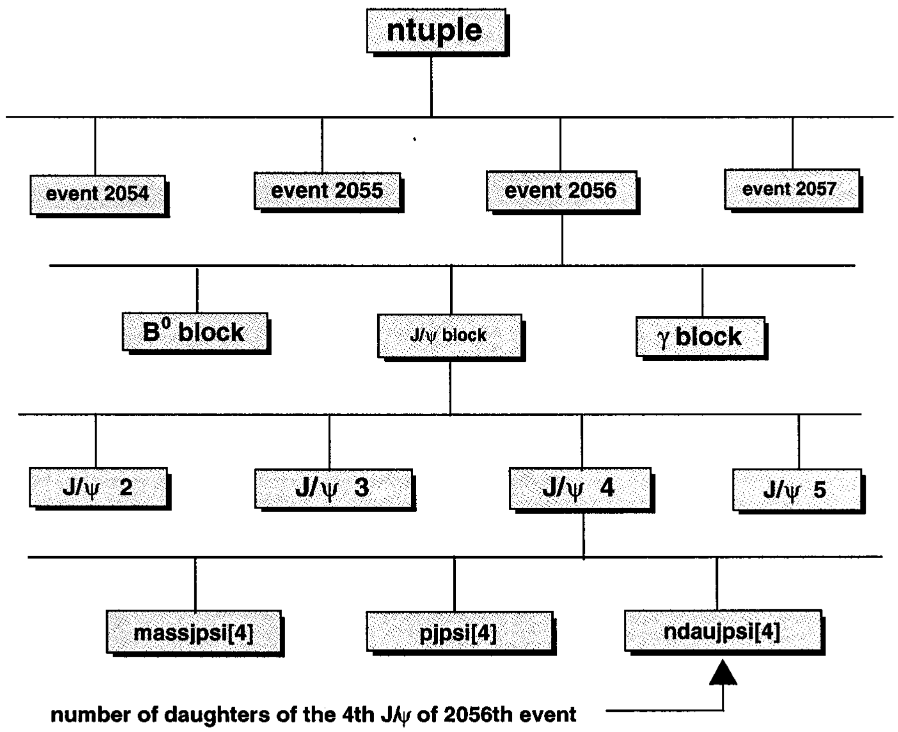

Figure 4.1: The structure of a *BABAR* ntuple. This particular example shows how you would access properties of the 4th J/ψ candidate in event # 2056.

BABAR ntuple. An ntuple typically contains several thousand events. For each event, there is a list of all of the particle candidates in the event, organized into blocks according to their hypothesized identities: a block of B^0 candidates, a block of J/ψ candidates, a block of π^0 candidates, and so on. For each particle candidate, the relevant quantities from the reconstruction are listed, including measured kinematic quantities like momentum and mass, and the ID numbers and (probable) identities of its daughters. Figure 4.1 shows how these properties would be accessed for a specific particle candidate, the 4th J/ψ candidate in event #2056.[9] Thus, an ntuple provides a complete and intuitive description of the most likely decay tree of every event, as determined by *BABAR* reconstruction and ntuple production algorithms.

4.8 Independent analysis

It is once the ntuples are produced that the user really takes over the analysis. The ntuple production job creates ntuples as .hbook files for analysis in PAW, or as .root files for analysis in ROOT.[10] Once the data is in ntuple form, the user has his very own copy of the data, and can use PAW or ROOT on any computer, anywhere, to do offline analysis. Individual researchers typically focus on a single decay or other physics process of interest; and a typical analysis involves using tight final selection criteria to isolate and study the decay.

4.9 Simulated data (Monte Carlo)

Simulated or "Monte Carlo" (MC) data sets are a very important tool in particle physics experiments, providing a way to test whether experimental results are consistent with theoretical predictions. "Monte Carlo" refers to the standard method for producing simulated data in particle physics (and many other disciplines). The Monte Carlo method generates a random set of events distributed according to input probability density functions, which reflect our current knowledge of particle physics. The probability density functions are based on world-average values for the properties of the relevant particles (such as masses, lifetimes, branching ratios); and on the laws of particle physics (such as conservation laws, quantum field theory). If the theory input into the probability density functions is correct, then the real and simulated data sets should agree. If they differ, the first thing to check

[9] Of course, in practice a researcher is rarely interested in only a single particle; in this example you would probably have a code that loops over all of the J/ψ candidates in an event, and would be part of a greater loop over all events.

[10] PAW and ROOT are are physics analysis programs.

is that there are no mistakes in the simulation. If there are no mistakes, then differences between real and simulated data could indicate that the theory is wrong or incomplete.

Simulations are used to model the signal and background distributions for the decay of interest.[11] A typical approach is to to test and optimize an analysis strategy on Monte Carlo data. This works because the simulated data set comes with "truth" information, so that in addition to OPR's interpretation of an event you also know what really happened. In particular, you can look up whether a given event is really a signal event or a background event, so you can use Monte Carlo data sets to find ways to select signal events while rejecting background.

In a *BABAR* simulation, there are three basic things to be modeled: the detector, the particles and their interactions in the detector, and the detector response. *BABAR* uses a GEANT-based [57] simulation with four stages:

1. Generation of the underlying physics event. The output is the four-vectors of the intermediate and final-state particles, with some "smearing" of the beam energies and collision coordinates to make things more realistic.

2. Simulation of the particle interactions and the detector response. The output consists of the signals ("`GHits`") that the imaginary particles produce in the imaginary sub-detectors, as well as truth information about the particles that produced the signals. Effects like multiple scattering and energy loss are taken into account.

3. Overlaying of backgrounds and digitization of the energy deposits. This stage models the detector response, taking the `GHits` and translating them into imaginary FEE signals. It also mixes in backgrounds measured from real data.

4. Reconstruction. Reconstruction is done in essentially the same way as for real data.

The simulations are set up so that real and simulated data looks exactly the same as far as the OPR is concerned. The only difference between a real data ntuple and a Monte Carlo ntuple is that a Monte Carlo ntuple has an additional, separate "truth block" which can be used to access the truth information about the particles.

4.10 Review

In this chapter, you have seen the many stages of processing undergone by the data set as it goes from detector to desktop. The trigger filters out background, reducing PEP-II's high rates to a level that the Online Prompt Reconstruction system (OPR) can handle.

[11] The "signal" is the decay you are interested in; "background" is everything else. So for my analysis, the signal is $B^0 \to J/\psi\,\gamma$.

OPR runs tracking and particle identification algorithms and stores the results in the Event Store database, which includes the list of `BtaMicroCandidates` in the Micro database, and tag bits in the Nano database. Individual Analysis Working Groups (AWGs) define and use the tag bits to sort the data into *collections* of events of interest to their group, and individual researchers use the collections to produce their very own ntuples. At each stage, increasingly tight selection is applied, so that the ntuples contain only events that are (probably) of interest to the particular researcher. The researchers use the ntuples for independent analysis, generally involving much tighter and carefully optimized selection criteria and statistical analysis. An example of an independent analysis is my study of the decay $B^0 \to J/\psi\,\gamma$; this is the subject of Chapter 5.

Chapter 5

Analysis

5.1 Preface

The chapter is adapted from an (unpublished) supporting document that I wrote for *BABAR* for the $B^0 \to J/\psi\,\gamma$ analysis. It includes some *BABAR*-specific terms, particularly in Sections 5.3 and 5.4, as these sections describe parts of the analysis that are standardized at *BABAR*. I have retained these terms as a reference for other *BABAR*ians, but they are not required to understand the chapter. In addition, for the reader unfamiliar with the *BABAR* experiment, Chapter 4 and the glossary at the end of this thesis should help to fill in some of the blanks. Alternatively, the reader could just ignore the *BABAR*-specific terms, as they are not required to understand this chapter.

5.2 Overview

This section gives a brief overview of the analysis.

When possible, analyses at *BABAR* are *blind*, meaning that the physics result is kept hidden until the analysis is essentially complete. This helps to reduce the possibility of bias on the part of the experimenter. A blind analysis is considered essential for rare decay searches in which the decay mode has not yet been conclusively detected [58].

The recommended analysis method for rare B decay searches is the *hidden signal region* method [58]. In this method we define a *signal region* in which the signal is expected to be concentrated. For exclusive B decays the signal region is usually defined in terms of the standard kinematic variables ΔE and $m_{\rm ES}$, whose ranges are tightly restricted for signal events. A blind analysis then just means that when studying onpeak data[1], the

[1] *Onpeak* data is data collected at the $\Upsilon(4S)$ resonance energy. The $\Upsilon(4S)$ resonance energy is the usual run energy at *BABAR*; however, *BABAR* sometimes also runs below the $\Upsilon(4S)$ resonance energy, and studies of this *offpeak* data are used to learn about continuum background.

experimenter does not look inside the signal region (or possibly a larger hidden region) in any way. However, he is still permitted to study onpeak data *outside* of the hidden region. In addition, he can study simulated and/or offpeak data in any region, including the hidden region.

The analysis searches the full data set recorded by the *BABAR* detector from its startup in 1999 to 2003. It uses simulated data for optimization and background studies. Simulated data samples are conventionally referred to as "Monte Carlo" or "MC" samples; we will use this convention throughout the text. The Monte Carlo data include *signal* Monte Carlo samples, consisting only of real $B^0 \to J/\psi\,\gamma$ events; and *background* Monte Carlo samples, consisting only of fake $B^0 \to J/\psi\,\gamma$ events. The background Monte Carlo samples are further categorized as $B\overline{B}$, continuum, or inclusive J/ψ background, and weighted to the onpeak luminosity in order to model the background in the real onpeak data.

Rather than study the whole set of real and simulated *BABAR* data, we select for analysis only candidate $B \to J/\psi\ X$ events in which the J/ψ decays to two leptons. ("X" denotes some arbitrary final state.)

We define the signal and hidden regions for the blind analysis in the ΔE-$m_{\rm ES}$ (see above) kinematic plane. The kinematic constraints for $e^+e^- \to B\overline{B}$ events at the $\Upsilon(4S)$ resonance cause exclusive B decays to be concentrated in a small region of this plane. We define specific boundaries for this *signal region* and use it to select signal events with high efficiency. In addition, we surround the signal region with a larger *hidden region*, the region that is kept blind throughout the analysis. The researcher must not look at onpeak data inside the hidden region.

There are two main steps in the analysis. The first step is to choose, optimize and apply selection criteria to accept signal events and reject background events. The second step is to estimate and subtract the remaining background to obtain the number of signal events, n_s.

To pass the final selection, a candidate must (a) fall in the signal region, and (b) pass a set of tight selection cuts. The primary purpose of the signal region is to isolate the signal. The primary purpose of the final cuts is to reduce background. The final cuts include cuts for J/ψ and photon selection, as well as cuts to suppress continuum background.

We use signal and background Monte Carlo samples to optimize the final selection. First we optimize the final cuts with the signal region held constant; then we optimize the signal region with the final cuts held constant.

The branching fraction for $B^0 \to J/\psi\,\gamma$ is given by

$$\mathcal{B} = \frac{n_s}{N_{B\overline{B}}\,\varepsilon_{signal}\,\mathcal{B}(J/\psi \to \ell^+\ell^-)}$$

where n_s is the number of signal events. The number of $B\overline{B}$ events $N_{B\overline{B}}$ and the branching fraction $\mathcal{B}(J/\psi \to \ell^+\ell^-)$ are known constants; and the signal efficiency ε_{signal} for events to pass the final selection (signal region + final cuts) is determined from signal Monte Carlo.

Even with the final cuts applied the number of onpeak data events in the signal region, $n_0 = n_s + n_b$, still includes a contribution from background n_b in addition to signal n_s. The number of background events in the signal region n_b is estimated from onpeak data outside of the hidden region, and subtracted from n_0 to obtain the signal n_s. The scale factor needed for the background estimate is determined from background Monte Carlo.

We perform two cross-checks of the Monte Carlo background modeling. We compare the background estimate from onpeak data to that from background Monte Carlo; the results are consistent. We also compare the continuum background estimate from offpeak data to that from continuum Monte Carlo; the results are consistent. This increases our confidence in the final background estimate.

To correct for differences between data and Monte Carlo we apply standard *BABAR* efficiency corrections and determine the associated systematic errors.

The final result is just one number n_0 — the number of onpeak events that pass the final selection (signal region + final cuts). We use this result to calculate the likelihood function for the branching fraction. We then use the likelihood function to obtain a 90% CL upper limit on the branching fraction.

We obtain a final result of $n_0 = 0$ onpeak events that pass the final selection. The corresponding 90% CL upper limit on the branching fraction for $B^0 \to J/\psi\gamma$ is 1.2×10^{-6}. In this chapter we describe in detail how we arrive at this conclusion.

The remainder of this chapter is organized as follows:

- Section 5.3 lists the data and simulated (Monte Carlo) sets used in this analysis.
- Section 5.4 describes the initial selection ("preselection") for the subset of *BABAR* events used in this analysis.
- Section 5.5 presents the final cuts.
- Section 5.6 describes the kinematic variables ΔE and $m_{\rm ES}$, and the regions of the ΔE-$m_{\rm ES}$ plane relevant to the blind analysis.
- Section 5.7 gives the formula for the branching fraction.
- Section 5.8 describes the optimization of the final cuts (but not of the signal region).
- Section 5.9 describes the optimization of the signal region.
- Section 5.10 describes the method for the final background estimate.

- Section 5.11 describes the cross-checks on the Monte Carlo modeling of the background.
- Section 5.12 presents the efficiency corrections and their associated systematic errors.
- Section 5.13 presents the quantities used in the final calculation of the 90% CL upper limit for $\mathcal{B}(B^0 \to J/\psi\gamma)$.
- Section 5.14 calculates the expected number of onpeak events in the signal region, n_0.
- Section 5.15 describes how the 90% CL upper limit on the branching fraction is calculated from n_0.
- Section 5.16 presents and discusses the final results.
- Section 5.17 presents our conclusions and suggestions for future studies.

5.3 Data Set

This analysis uses the complete data sample recorded by the *BABAR* detector from 1999 to 2003. The data consists of two data sets:

- The "Summer 2002" data set includes data collected between *BABAR*'s startup back in 1999 and the summer shutdown in 2002. It includes both Run 1 and Run 2 data, and was reconstructed using the 10-series release of the *BABAR* software. It contains 81.9 fb^{-1} of onpeak data, as well as 9.6 fb^{-1} of offpeak data collected at about 40 MeV below the $\Upsilon(4S)$ resonance for background studies.
- The Run 3 data set contains data collected between December and June 2003. It was reconstructed using the 12-series release of the *BABAR* software. We use all of the Run 3 data: 31.2 fb^{-1} of onpeak data and 2.4 fb^{-1} of offpeak data.

The combined data set corresponds to an integrated onpeak (offpeak) luminosity of 113.1 (12.0) fb^{-1}.

The analysis also uses simulated or Monte Carlo data for background and optimization studies. The names of the *BABAR* production cycles for our samples are SP4 and SP5. The SP4 Monte Carlo was produced for the Summer 2002 data set using the release 10 software. So to produce the SP4 samples we used release 10.4.4 (aka "analysis-13b"). The SP5 Monte Carlo was produced during Run 3 using the release 12 software. So to produce the SP5 samples we used release 12.5.2-physics-1a (aka "analysis-14a"). (*BABAR* software nomenclature is explained in Section 4.3.)

The data and Monte Carlo samples are produced in the form of *ntuples*, the convenient and portable data storage format described in Section 4.7. The ntuples serve as the user's own private copy of the data.

The different Monte Carlo sets used in this analysis are (Table 5.1):

- signal $B^0 \to J/\psi\gamma$, 39000 events
- generic $B^0\overline{B}^0$, 661.7 fb^{-1}
- generic B^+B^-, 680.2 fb^{-1}
- generic uds,[2] 171.4 fb^{-1}
- generic c, 206.2 fb^{-1}
- generic τ, 146.3 fb^{-1}
- inclusive J/ψ,[3] where the J/ψ meson is reconstructed only in the low-background, high-efficiency $J/\psi \to \ell^+\ell^-$ mode, 375.9 fb^{-1}

See Table 5.1 for more detail. The values provided are the *equivalent time-integrated luminosities* of the samples. Luminosity is the number of events per unit area, per unit time, so the time-integrated luminosity is just the number of events per unit area for a given time period of data-taking. It is a measure of the size of the data sample. The units of time-integrated luminosity for *BABAR*-sized samples are inverse femtobarns (fb^{-1}). The *equivalent* luminosity is defined as follows: if a Monte Carlo sample contains (for example) N *uds* events, then the equivalent luminosity of the MC sample is defined as the luminosity of a real data sample that would contain N *uds* events.

5.3.1 Background Monte Carlo samples

In order to obtain a meaningful result in particle physics, we need to be able to distinguish signal from background. But this is difficult in real data, as signal and background contributions occur together. Furthermore, there often are not enough signal events in real data to adequately model the signal distributions of analysis variables. These are some of the reasons that we use simulations.

[2] *uds* is *BABAR* shorthand for the light quark continuum background, $u\overline{u}$, $d\overline{d}$, $s\overline{s}$. Similarly, c is short for the heavy quark continuum background, $c\overline{c}$.

[3] Note: In the table, $\sigma(\text{inclusive } J/\psi) = 2\sigma(e^+e^- \to B\overline{B})\mathcal{B}(B \to J/\psi\ X)\mathcal{B}(J/\psi \to \ell^+\ell^-) = 2(1.096\,\text{nb})(0.0115)(0.1181) = 0.00297\,\text{nb}$. Note also that the luminosity provided in Table 5.1 is that of the *total* inclusive J/ψ sample, including the contribution from $B\overline{B}$ Monte Carlo. See Section 5.3.1.

Table 5.1: Cross sections and weights for the various Monte Carlo and data samples. To make a Monte Carlo soup that accurately represents the data, samples are weighted to the onpeak luminosity. However, we also use the offpeak weights for cross-checks. See Section 5.11.

	Number of events			Cross-section	Luminosity (fb^{-1})			Onpeak	Offpeak
ource	Runs 1,2	Run 3	Total	(nb)	Runs 1,2	Run 3	Total	weight	weight
ds	150060000	208078000	358138000	2.09	71.8	99.6	171.4	0.659	0.070
	97481800	170536000	268017800	1.30	75.0	131.2	206.2	0.548	0.058
	72346300	7826000	80172300	0.94	132.0	14.3	146.3	0.772	0.082
$^0\overline{B}^0$	158300100	204284000	362584100	0.548	288.9	372.8	661.7	0.171	0.018
$^+B^-$	154915700	217844000	372759700	0.548	282.7	397.5	680.2	0.166	0.018
nclusive J/ψ	294000	824000	1118000	0.00297	98.8	277.0	375.9	0.108	0.011
ignal	39000	0	39000						
npeak data					81.9	31.2	113.1	1	0.11
ffpeak data					9.6	2.4	12.0	9.42	1

We distinguish between signal Monte Carlo, which includes only true $B^0 \to J/\psi\,\gamma$ events, and background Monte Carlo, which consists of events containing only fake $B^0 \to J/\psi\,\gamma$ events. The latter are events that pass the preselection for $B^0 \to J/\psi\,\gamma$, despite not really being $B^0 \to J/\psi\,\gamma$ events. In addition to the total background sample, we sort the background into three different samples:

- inclusive J/ψ = all $B \to J/\psi\,X$, $J/\psi \to \ell^+\ell^-$ events[4] (except $B^0 \to J/\psi\,\gamma$)
- generic $B\overline{B}$ = all B decays (except $B \to J/\psi\,X$ and $B^0 \to J/\psi\,\gamma$)
- continuum = all $q\overline{q}$ events ($q = u,\ d,\ s,\ c$) and τ events

These samples are almost just simple sums of the sets listed in the previous section, but not quite. The original inclusive J/ψ Monte Carlo sets of Section 5.3 may contain $B^0 \to J/\psi\,\gamma$ events, which must be moved to the signal category. Similarly, the B^0 and $B^\pm$ sets contain $B \to J/\psi\,X$ events and may contain $B^0 \to J/\psi\,\gamma$ events, which must be moved to the inclusive J/ψ and signal categories, respectively. This ensures that the correct weights are assigned to each sample.

Each histogram in this chapter comes with a statistics box. In this box, the number of `Entries` in the histogram is the raw, unweighted number of entries. However, the `Integral` in the statistics box gives the weighted number of events in the full histogram. To obtain

[4] "X" here denotes some arbitrary final state.

a background sample in which the different modes occur with the expected ratios and magnitudes, the background MC and offpeak data samples are weighted to the onpeak luminosity. The weights are given in Table 5.1.

5.4 Candidate Preselection

Preselection refers to the initial, loose selection applied during ntuple production to select events of physics interest to the user. Ntuple production is the process whereby data and Monte Carlo from the central *BABAR* database are converted to the useful, portable *ntuple* format for independent analysis by the user (see Section 4.7). During this process, the basic lists of observed particle candidates ("`BtaCandidates`") identified during the central *BABAR* reconstruction are used to form new lists of composite objects (e.g., the reconstruction of π^0 candidates from two photons from the `GoodPhotonsLoose` list).

5.4.1 Tag bits

A typical analysis job begins with a quick skim of the database to select only events that contain a certain set of *tag bits*, loose selectors that flag events of physics interest. Individual tag bits are defined by different *BABAR* Analysis Working Groups (AWGs).

We choose our tag bits to select candidate $B \to J/\psi X$ events in which the J/ψ candidate decays to two lepton candidates. We use only the $J/\psi \to \ell^+\ell^-$ mode because this mode can be reconstructed with low background and high efficiency. To impose the requirement, in our ntuples each event is required to contain the `isBCMultihadron` tag bit,[5] plus either the `JpsiELoose` or `JpsiMuLoose` tag bit.

5.4.2 The `isBCMultihadron` tag bit

Most studies at *BABAR* are on $B\overline{B}$ events. However, as explained in Section 4.5.1, the *BABAR* trigger passes not only $B\overline{B}$ and other multihadron events but other types as well, such as muon-muon, tau, two-prong, and two-photon events, for use in other types of studies. To select only the $B\overline{B}$ events of interest, we use the `isBCMultihadron` tag bit.

The `isBCMultihadron` tag bit [59] is often called the "B-counting" tag bit, because it is used to identify $B\overline{B}$ events for B-counting. It requires events to pass a background filter tag bit for multihadron events, and then imposes additional requirements to obtain

[5]During the production of the ntuples, we inadvertently did not apply the requirement to contain the `isBCMultihadron` tag bit to the signal ntuples. However, this has very little effect on the analysis, as the demand that an event be a signal event pretty much guarantees that it will pass the `isBCMultihadron` selection.

high efficiency for $B\overline{B}$ events ($\varepsilon_{BB} = 0.954$) and low efficiency for bhabha, two-photon and beam-gas events. Briefly, these requirements are:

- The event must satisfy one of the physics triggers — ie, the EMC or DCH trigger.
- Number of `GoodTracksLoose` in the fiducial volume (see below) ≥ 3.
- Ratio of the second to zeroth Fox Wolfram moment [60] of all the particles $R2 < 0.5$. $R2$ is an event-shape variable, described in Section 5.5.3.
- Total energy of all particles in the fiducial region (see below) $E_{total} > 4.5\,\text{GeV}$. (Here, only neutral candidates with $E > 30\,\text{MeV}$ are considered.)
- Distance of closest approach between the primary vertex and the measured beam spot must be $< 0.5\,\text{cm}$ in the xy plane, and $< 6\,\text{cm}$ in the z-direction.

The fiducial region is a *BABAR*-defined region of the detector with well-measured reconstruction efficiency, and good Monte Carlo modeling of the event. The fiducial region cuts are $0.41 < \theta < 2.54$ rad for charged candidates and $0.41 < \theta < 2.409$ rad for neutral candidates.

`GoodTracksLoose` are tracks[6] from the basic `ChargedTracks` list that pass the following selection for "loose but good" track reconstruction:

- Transverse momentum $p_t \geq 0.1\,\text{GeV}/c^2$.
- Momentum $p \leq 10.0\,\text{GeV}/c^2$.
- Number of hits in the DCH ≥ 12 (to ensure that p and $\mathrm{d}E/\mathrm{d}x$ are well-measured).
- Track fit χ^2 probability ≥ 0 (to select well-reconstructed tracks).
- Distance of closest approach between the primary vertex and the measured beam spot must be $< 1.5\,\text{cm}$ in the xy plane, and $< 10\,\text{cm}$ in the $\pm z$-direction.

5.4.3 The `JpsiELoose` and `JpsiMuLoose` tag bits

To select only events in which the J/ψ candidate decays to two leptons, we use the `JpsiELoose` and `JpsiMuLoose` tag bits. These tag bits impose the following requirements:

- $J/\psi \to e^+e^-$ candidates are formed from oppositely-charged electron pairs that have passed bremsstrahlung-recovery algorithms to account for energy loss due to radiated

[6]Tracks are the basic charged particle candidates, reconstructed in the SVT and the DCH. See Section 4.5.2.

photons. Electron candidates are required to pass the standard electron cut-based selector [61] in `Loose` mode. To form a J/ψ candidate, the electron pairs are required to have an invariant mass within the mass window 2.5-3.3 GeV/c^2.

- $J/\psi \to \mu^+\mu^-$ candidates are formed from oppositely-charged muon pairs. Muon candidates are required to pass the standard muon cut-based selector [62] in `Loose` mode. To form a J/ψ candidate, the muon pairs are required to have an invariant mass within the mass window 2.8-3.3 GeV/c^2.

A brief description of *BABAR*'s electron and muon selectors is provided in Appendix B.

5.4.4 $B^0 \to J/\psi\gamma$ reconstruction

Only events containing a $B^0 \to J/\psi\gamma$ candidate decay are written to the ntuple set. So in each event, either the B^0 or the $\overline{B}^0$ candidate is reconstructed from a J/ψ candidate and a photon candidate. We demand the following for a decay to be a $B^0 \to J/\psi\gamma$ candidate:

- J/ψ candidates are just objects that contain the above mentioned `JpsiELoose` and `JpsiMuLoose` tag bits. However a tighter mass constraint of $2.9 < m_{J/\psi} < 3.3$ GeV/c^2 is applied in both the $J/\psi \to e^+e^-$ and $J/\psi \to \mu^+\mu^-$channels. The J/ψ mass is subsequently constrained to the Particle Data Group (PDG) value [10].

- Photon candidates are members of the `GoodPhotonsLoose` list, which is derived from the basic Micro-level `CalorNeutral` list. The `CalorNeutral` list consists of single bumps[7] in the EMC that are not matched with any charged tracks. To be included in the `GoodPhotonsLoose` list, the bumps must also have a raw energy (the energy uncorrected for any types of leakage) of at least 30 MeV and a maximum lateral energy distribution (*LAT*) of 0.8. *LAT* is a calorimeter shower-shape variable, described in Section 5.5.2.

- B^0 candidates are formed from J/ψ and γ candidates for which vertexing and kinematic fitting are successful.

5.5 Final $B^0 \to J/\psi\gamma$ selection

In the final selection we impose a series of tighter cuts on the events in the ntuples. The aim of the cuts is to reject background without significantly compromising the signal yield. The values of the cuts are chosen via the optimization procedure described in Section 5.8. In this section we present the final cuts.

[7]Bumps are the basic EMC particle candidates. See Section 4.5.2.

We present plots of each selection variable, for onpeak data, signal Monte Carlo, and background Monte Carlo. The onpeak plots are the purple plots at the left;[8] the signal MC plots are the green plots in the middle; and the background MC plots are the red plots on the right. We also show in blue[9] the contribution of inclusive J/ψ MC events to the total background, since inclusive J/ψ events turn out to be the background that is most difficult to eliminate in this analysis (see Section 5.9). The background plots are weighted to the onpeak luminosity. The signal plots are not weighted.

In each plot, all final cuts are applied, except for the cut on the selection variable in the plot, and except for the requirement that events fall in the signal region. The values of the cuts are indicated by the vertical black lines.

5.5.1 Final J/ψ selection

The final J/ψ selection does not introduce any new selection variables, but rather just tightens the preselection windows for the J/ψ mass and the $J/\psi \to \ell^+\ell^-$ lepton identification (Section 5.4.3).

- J/ψ mass cuts

  The mass of a true J/ψ particle is $m_{J/\psi} = 3.09687 \pm 0.0004\,\mathrm{GeV}/c^2$ [10]. The final mass cuts are:

$$3.00 < m(J/\psi \to e^+e^-) < 3.14 \;\; \mathrm{GeV}/c^2$$
$$3.07 < m(J/\psi \to \mu^+\mu^-) < 3.13 \;\; \mathrm{GeV}/c^2$$

  Figure 5.1 shows the effect of the mass cuts. There is a long tail in the $J/\psi \to e^+e^-$ distribution because the e^+e^- pair loses energy due to the radiation of Bremsstrahlung photons. Therefore as in most analyses the optimal mass cuts in the $J/\psi \to e^+e^-$ channel are looser than those in the $J/\psi \to \mu^+\mu^-$ channel.

- Lepton identification

  - For $J/\psi \to e^+e^-$, both electrons are required to pass the `Likelihood` selector. The electron `Likelihood` selector has been found to be more efficient than the alternative standard cut-based electron selector. The decision to use the electron `Likelihood` selector is not optimized.

[8] These plots were kept blind throughout the analysis. They were *not* used in the optimization.

[9] For those reading the black-and-white version of this thesis, note that the histograms can be distinguished by the fact that the inclusive J/ψ histograms are are slightly darker than the background MC ones. And of course, the inclusive J/ψ MC contribution to the MC background is always smaller than the total MC background.

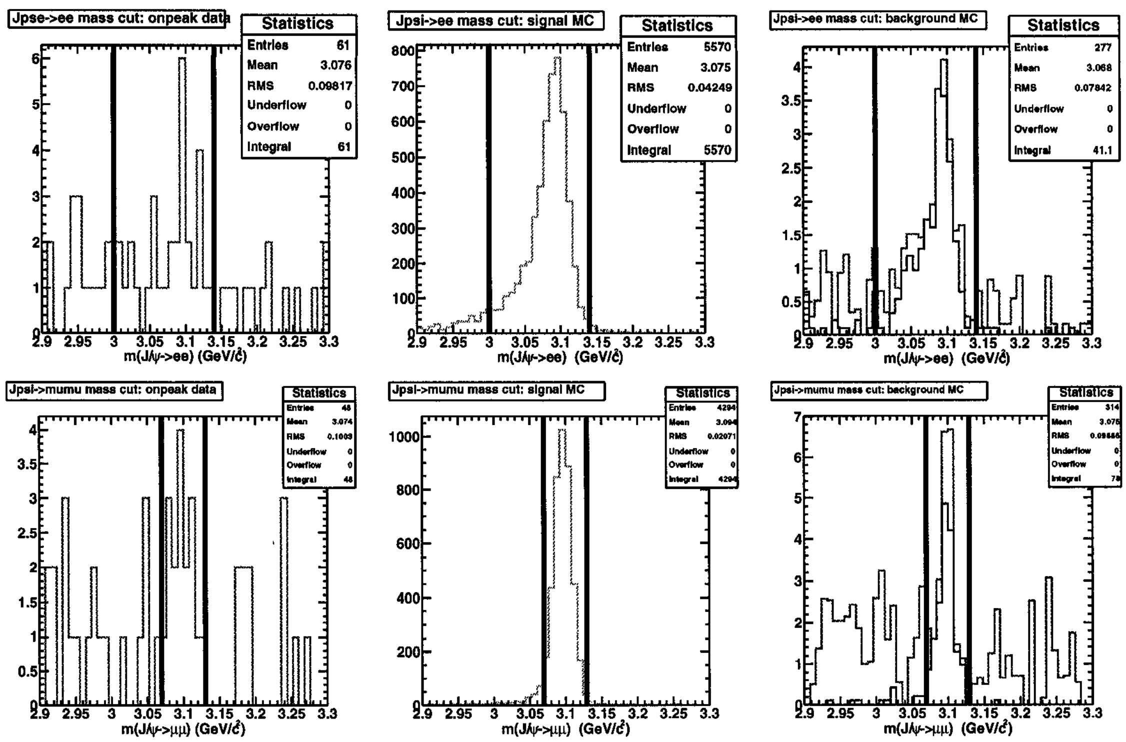

Figure 5.1: Final J/ψ mass cuts in the $J/\psi \to e^+e^-$ and $J/\psi \to \mu^+\mu^-$ channels, with all other cuts applied. From left to right, the plots are: onpeak data (purple), signal MC (green), and total background MC (red), with the contribution to background from inclusive J/ψ MC events shown in blue.

- For $J/\psi \to \mu^+\mu^-$, both muons are required to pass the `Tight` level of the standard muon cut-based selector [62]. This choice is optimized.

5.5.2 Final photon selection

For the final photon selection we impose two standard cuts: a cut on the shower-shape variable LAT, and a veto of photons from neutral pion and neutral eta decays.

- LAT cut

  One source of fake photons is hadrons that shower in the EMC. LAT, the lateral energy distribution of a candidate, is the quantity most often used to distinguish electromagnetic from hadronic showers. It is given by [63]

$$LAT = \frac{\sum_{i=3}^{N} E_i r_i^2}{\sum_{i=3}^{N} E_i r_i^2 + E_1 r_0^2 + E_2 r_0^2}. \tag{5.1}$$

  Here, N is the number of crystals in the shower, E_i is the energy deposited in the ith crystal, r_i is the polar radius in the plane perpendicular to the line pointing from

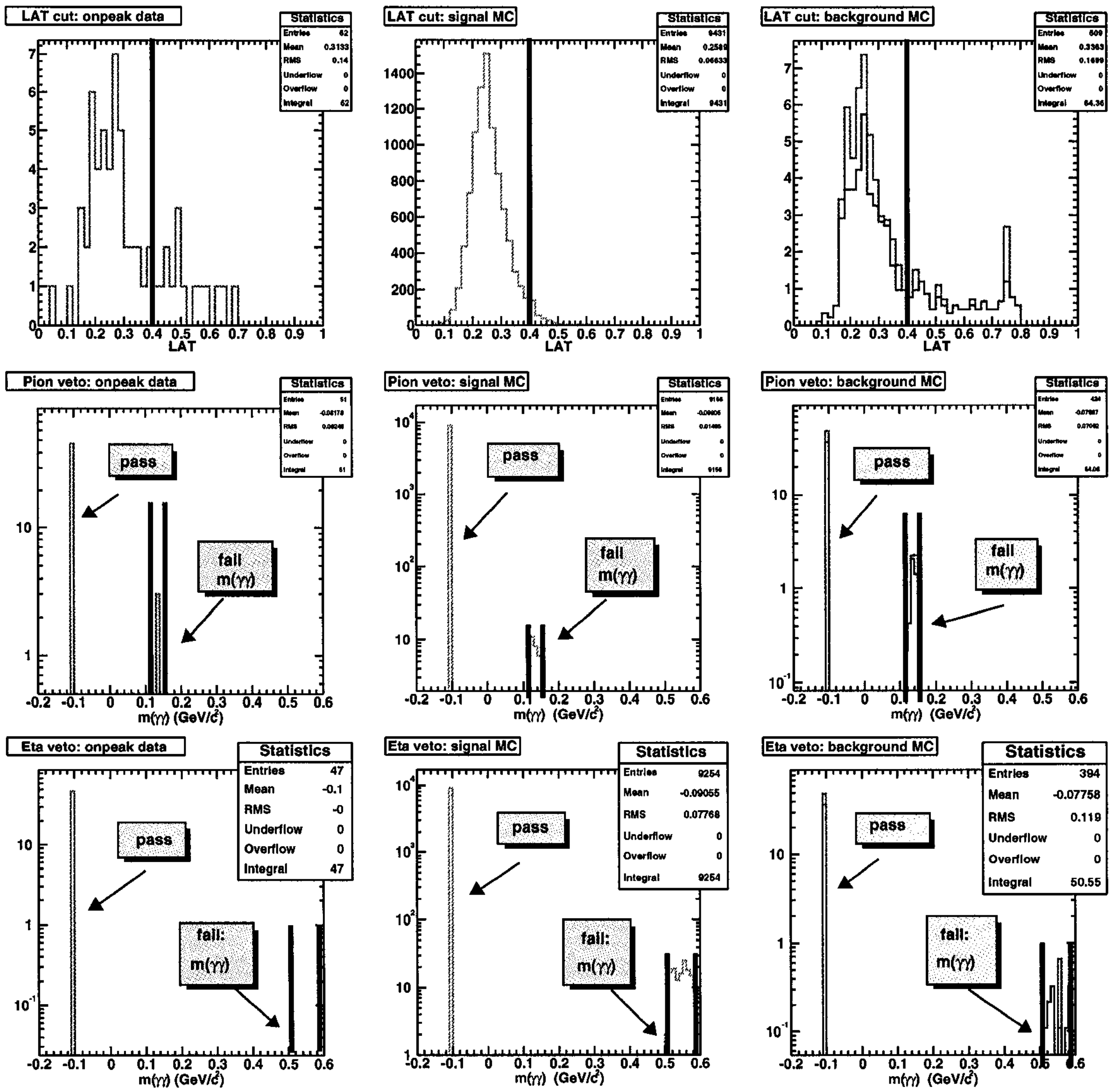

Figure 5.2: Final photon selection cuts. From left to right, the plots are: onpeak data (purple), signal MC (green), and total background (red), with the contribution to background MC from inclusive J/ψ MC events shown in blue. Each cut is shown with all other cuts applied. The π^0 (η) veto is applied to photons with energy $E > 50\,\mathrm{MeV}$ ($250\,\mathrm{MeV}$). The π^0 and η veto plots use a logarithmic scale, and the two vertical black lines indicate the mass range for the veto. For events that fail the veto, the invariant mass of the photon pair with the mass closest to the real π^0 (η) mass is plotted. Events in the negative bin are those that pass the veto.

the interaction point to the shower center, and r_0 is the average distance between two crystals. The energies are numbered such that $E_1 > E_2 > \dots E_N$, so that the sum in the numerator excludes the contribution from the two highest-energy crystals. Since electromagnetic showers typically deposit most of their energy in the first few crystals, they tend to have lower values of LAT.

As part of the preselection (Section 5.4.4), the photon candidate is required to be on the `GoodPhotonsLoose` list, which requires $LAT < 0.8$. In the final photon selection we tighten the LAT cut to $LAT < 0.40$. Figure 5.2 shows the effect of this cut.

- π^0 and η vetoes

  Another source of background to $B^0 \to J/\psi\gamma$ photons is photons from other physics processes. At PEP-II, the majority of photons come from the decay $\pi^0 \to \gamma\gamma$, so we apply a cut intended to veto these photons. We also veto photons from $\eta \to \gamma\gamma$ decays.

  We apply the neutral pion veto to photon candidates with energy greater than 50 MeV, and the eta veto to photon candidates with energy greater than 250 MeV. For a given photon candidate, we loop over all other photon candidates with $E > 50\,\mathrm{MeV}$ (250 MeV) and calculate the invariant mass of the combined four-momentum of the pair,

$$m_{\gamma pair} = \sqrt{(p_{1\gamma} + p_{2\gamma})^2}. \tag{5.2}$$

  To veto the neutral pions ($m_{\pi^0} = 0.135\,\mathrm{GeV}/c^2$), we reject all $B^0 \to J/\psi\gamma$ events in which the photon candidate forms a pair with another photon candidate such that

$$0.115 < m_{\gamma pair} < 0.155\,\mathrm{GeV}/c^2. \tag{5.3}$$

  To veto the etas ($m_\eta = 0.547\,\mathrm{GeV}/c^2$), we reject all $B^0 \to J/\psi\gamma$ events in which the photon candidate forms a pair with another photon candidate such that

$$0.507 < m_{\gamma pair} < 0.587\,\mathrm{GeV}/c^2. \tag{5.4}$$

  These mass ranges and minimum photon energy requirements are the same as those used in *BABAR* Analysis Document 665 [64] for the analysis of $B \to K^*\gamma$. They are not optimized but are deemed reasonable. (However, the decision of whether or not to apply the vetoes is optimized—the answer, obviously, is yes.) The reason for the $E > 50\,\mathrm{MeV}$ (250 MeV) restriction is to avoid photons from beam backgrounds, whose pairing with photons from real $B^0 \to J/\psi\gamma$ events could result in a four-vector with the pion mass, and thus the rejection of real $B^0 \to J/\psi\gamma$ events. The effects of the vetoes are shown in Fig. 5.2.

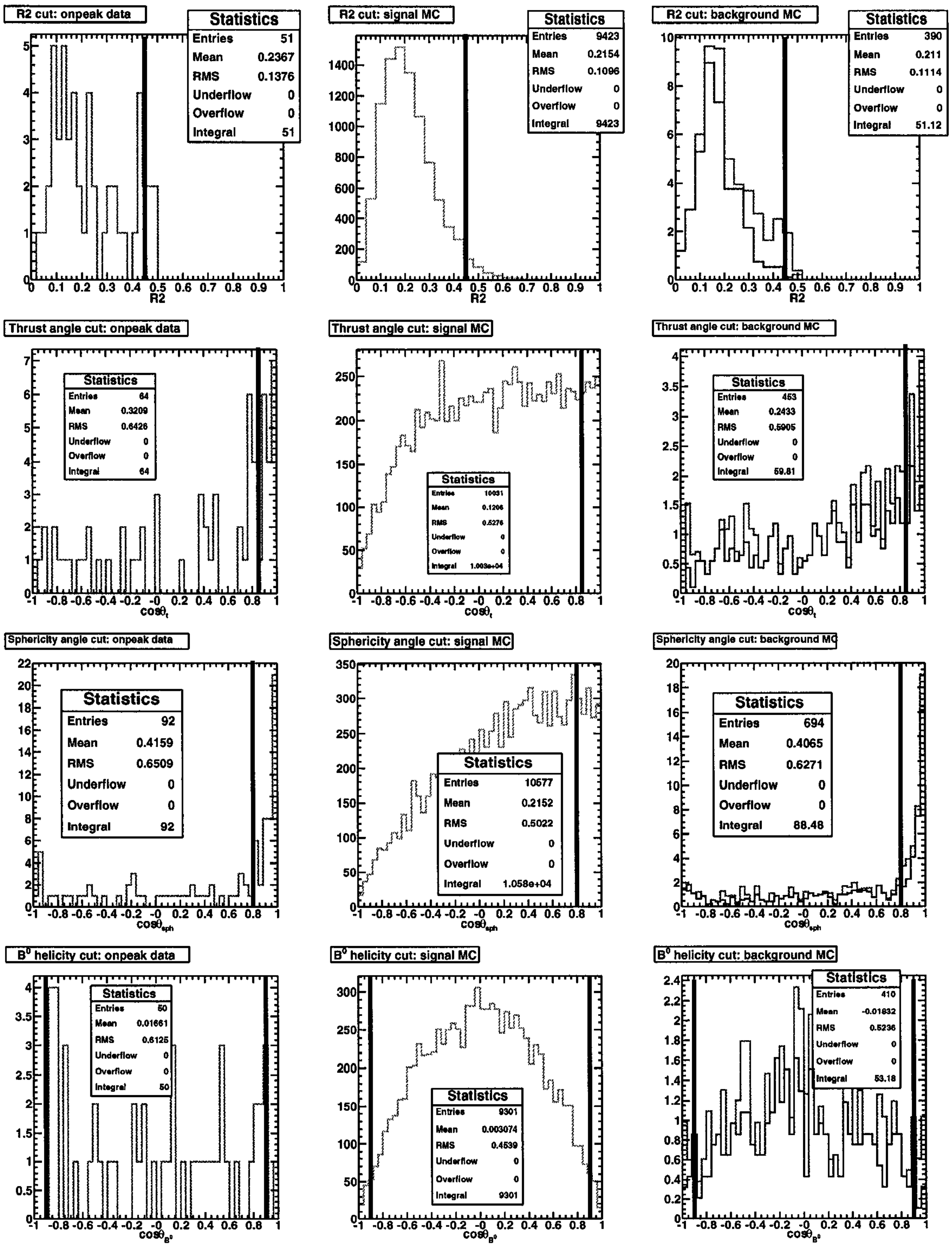

Figure 5.3: Continuum suppression cuts. From left to right, the plots are: onpeak data (purple), signal MC (green), and total background MC (red), with the contribution to background from inclusive J/ψ MC events shown in blue. Each cut is shown with all other cuts applied.

5.5.3 Continuum suppression (event shape) cuts

Event shape cuts provide a powerful method for separating signal from continuum events. In true signal events, the B mesons are produced nearly at rest in the center-of-mass frame, so the B decays are nearly isotropic. On the other hand, continuum events have a pronounced two-jet structure. *BABAR* [37] has identified many discriminating variables for continuum rejection. In this analysis we apply cuts on the following general event-shape variables:

- $R2$, the ratio of the second and zeroth Fox-Wolfram moments

  The Fox-Wolfram moments, H_ℓ, for an event are defined as [60]

$$H_\ell = \sum_{i,j} \frac{|\mathbf{p_i}| \cdot |\mathbf{p_j}|}{E^2_{vis}} P_\ell(\cos\theta_{ij}) \tag{5.5}$$

  where P_ℓ are the Legendre polynomials, $\mathbf{p_{i,j}}$ are the particle candidate momenta, θ_{ij} is the opening angle between particles i and j, and E_{vis} is the total visible energy of the event. $R2 \equiv H_2/H_0$ ranges from 0 to 1. The jetlike continuum events have high values of $R2$, while the spherical B events tend to have lower values of $R2$. The `isBCMultihadron` tag bit (Section 5.4.2) requires $R2 < 0.5$; in the final selection this cut is tightened to $R2 < 0.45$. The effect of the cut is shown in Fig. 5.3.

- θ_t and θ_{sph}, the B thrust and sphericity angles

  θ_t and θ_{sph} measure the correlation between the direction of the candidate B meson and the direction of the other particle candidates in the event, in the longitudinal and transverse directions, respectively. The thrust (sphericity) axis is defined as the direction that maximizes the thrust T (sphericity S). Thrust is related to this direction $\hat{\mathbf{T}}$ by [65]

$$T = \sum_i \frac{|\hat{\mathbf{T}} \cdot \mathbf{p}_i|}{\sum_i |\mathbf{p}_i|}. \tag{5.6}$$

  Sphericity is (3/2) times the sum of the two largest eigenvalues of the sphericity tensor S^{ab} [66]:

$$S^{ab} = \sum_i \frac{p_i^a p_i^b}{\sum_i |\mathbf{p_i}|^2} \quad a, b = x, y, z \tag{5.7}$$

$$S = \frac{3}{2}(\lambda_2 + \lambda_3). \tag{5.8}$$

  The thrust (sphericity) axis of the B candidate is calculated from its (alleged) J/ψ and γ daughters, while the axis of the rest of the event is calculated using all of the other charged and neutral objects. The thrust (sphericity) angle is just the angle between the thrust (sphericity) axis of the B meson, and that of the rest of the event.

For the highly directional continuum events, the distributions of $\cos\theta_t$ and $\cos\theta_{sph}$ are both sharply peaked near 1. In contrast, for the isotropic B decays the distributions are more uniform.

The final cuts on these variables are:

$$\cos\theta_t < 0.85 \tag{5.9}$$

$$\cos\theta_{sph} < 0.80 \tag{5.10}$$

The effects of these cuts are shown in Fig. 5.3.

- θ_B, the B helicity angle

  B mesons from exclusive B decays are generally subject to tight angular momentum constraints which do not apply to fake B candidates from background. The spin structure of a decay is characterized by the helicity angle. The B helicity angle is the angle between the beam direction and the flight direction of the B candidate, in the center-of-mass frame.[10] The decay $\Upsilon(4S) \to B\overline{B}$ produces two B mesons in an L=1 state, with the result that for signal events the helicity angle θ_B follows a $\sin^2\theta_B$ $(= 1 - \cos^2\theta_B)$ distribution. In contrast, fake B candidates from $q\overline{q}$ events are not subject to this angular momentum constraint, so the $\cos\theta_B$ distribution for continuum events is more uniform.

  The final cut for the B helicity angle is,

$$|\cos\theta_B| < 0.90. \tag{5.11}$$

  The effect of this cut is shown in Fig. 5.3.

5.5.4 Unused selection variables

In addition to the selection variables listed above, we also consider and decide against using several other selection variables. They are listed below and presented in Fig. 5.4 for the sake of completeness.

- **J/ψ helicity angle** The distributions of the J/ψ helicity angle for signal and background Monte Carlo do not suggest any useful cut for this quantity, so we do not test or use it.

[10] In the reaction, $Y \to X \to a + b$, the helicity angle of particle a is the angle measured in the rest frame of the decaying parent particle, X, between the direction of the decay daughter a and the direction of the grandparent particle Y [67]. In this case, $Y = e^+e^-$, $X = \Upsilon(4S)$, $a = B$, and $b = \overline{B}$.

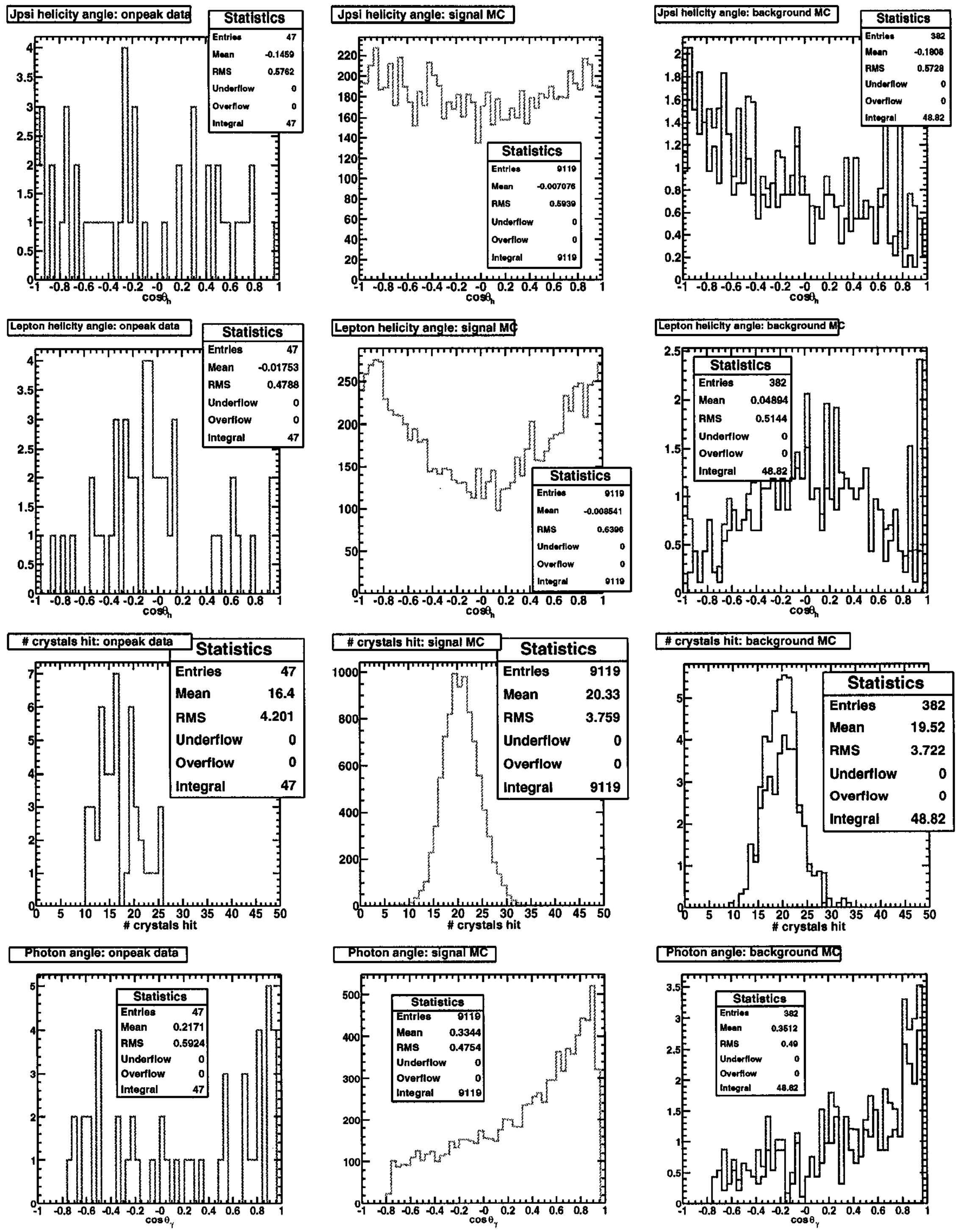

Figure 5.4: Unused selection variables, with all final cuts applied. From top to bottom: J/ψ helicity, lepton helicity, number of crystal hits in a photon cluster, and photon angle. From left to right, the plots are: onpeak data (purple), signal MC (green), and total background MC (red), with the contribution to background from inclusive J/ψ MC events shown in blue.

- **lepton helicity angle** The signal and background distributions for the cosine of the helicity angle are different, suggesting that a cut on this quantity might be useful. However, the optimization shows that the cut is only marginally useful. Also, it is probably safer not to apply a lepton helicity cut, because if there are any New Physics effects that enhance the decay rate enough to make $B^0 \to J/\psi\gamma$ measurable, then they might also make the lepton helicity distribution different from that predicted in Monte Carlo. Therefore we do not use this variable.

- **number of crystal hits** A low cut on the number of EMC hits in the photon cluster targets noise; a high cut targets background from hadronic showers. This cut is useful in some analyses; however, in this study the distributions for signal and background MC are almost indistinguishable (Fig. 5.4), so we do not test or use this variable.

- **geometrical acceptance** [64] The EMC provides coverage between $-0.774 < \cos\theta_{lab} < 0.962$. A cut requiring the photon to come from a tighter angular window could avoid candidates from the edges of the EMC and from beam backgrounds, which are generally not as well reconstructed as other photons. However, as the expected signal yield for $B^0 \to J/\psi\gamma$ is so tiny, higher statistics are more important than photon quality, so we do not test or apply this cut.

5.5.5 Comment on the onpeak plots

We note that the onpeak (left, purple) and total background (right, red) distributions for the selection variables presented are very similar, as expected for a rare decay in which a null or very small signal is expected. This reassures us that the Monte Carlo modeling of the background is correct.

5.6 ΔE, m_{ES}, and the three analysis regions

5.6.1 Kinematic variables

In the analysis of an exclusive B decay, the B meson is reconstructed from its decay products—in this case, a J/ψ and a photon. However, we also know that all B mesons at *BABAR* are produced in the process $e^+e^- \to B\overline{B}$. So to separate signal from background, we demand that the reconstructed B four-momentum be consistent with that expected for a B meson from $e^+e^- \to B\overline{B}$. Because the electron and positron beam four-momenta are known with only very small errors, this imposes a powerful kinematic constraint.

The standard method to impose this constraint is to use the kinematic variables ΔE and m_{ES} [68]. The physical significance of these two quantities is most easily seen in the center-of-mass frame, where they take the following form:

$$\Delta E = E_B^* - E_{beam}^* \tag{5.12}$$

$$m_{\mathrm{ES}} = \sqrt{E_{beam}^{*2} - \mathbf{p_B^*} \cdot \mathbf{p_B^*}} \tag{5.13}$$

(The stars $*$ denote center-of-mass variables.) The quantity ΔE is just the difference between the reconstructed energy of the B candidate and the expected energy from the beam. For a true B meson, ΔE should peak at zero. The beam-energy substituted mass m_{ES} is the invariant mass of the B meson, calculated with ΔE set to zero. A true B meson should have $m_{\mathrm{ES}} = m_B$.

The advantage of using ΔE and m_{ES} is that they are largely uncorrelated and make maximum use of the best-known quantity, the beam four-momentum. They are also Lorentz invariants. Since *BABAR* data are (of course) recorded in the lab frame, in this analysis we calculate ΔE and m_{ES} in the lab frame using the following equations:

$$\Delta E = (2q_\Upsilon \cdot q_B - s)/2\sqrt{s} \tag{5.14}$$

$$m_{\mathrm{ES}} = \sqrt{(0.5s + \mathbf{p_\Upsilon} \cdot \mathbf{p_B})^2 / E_\Upsilon^2 - p_B^2} \tag{5.15}$$

The explicit calculation in terms of the quantities available in the database (and thus, in the ntuples) is given below.

- $q_\Upsilon = (E_\Upsilon, \mathbf{p}_\Upsilon)$ is the four-vector of the $\Upsilon(4S)$—and thus, of the center of mass—in the lab frame. $s \equiv |q_\Upsilon|^2$. (Note: $E_\Upsilon^* = 2E_{beam}^*$.)

- $q_B = q_{J/\psi} + q_\gamma$.

- $q_{J/\psi} = (E_{J/\psi}, \mathbf{p}_{J/\psi})$. $\mathbf{p}_{J/\psi}$ is the reconstructed three-momentum after the two lepton candidates in $J/\psi \to \ell^+\ell^-$ are vertexed, and $E_{J/\psi} = \sqrt{p_{J/\psi}^2 + m_{J/\psi}^2}$, where $m_{J/\psi}$ is the PDG value [10].

- $q_\gamma = (E_\gamma, E_\gamma \hat{\mathbf{n}})$. $\hat{\mathbf{n}}$ is the direction of the photon candidate. It is given by $\hat{\mathbf{n}} = \mathbf{\Delta x}/|\Delta x|$, $\mathbf{\Delta x} \equiv \mathbf{x_{EMC}} - \mathbf{x_{vertex}}$, where $\mathbf{x_{EMC}}$ is the position of the photon bump in the EMC, and $\mathbf{x_{vertex}}$ is the position of the J/ψ vertex. This procedure reevaluates the photon direction assuming the photon was produced at the J/ψ vertex rather than the primary vertex, thus improving the measurement. (The primary vertex is the default assumption during the central online prompt reconstruction, since the reconstruction algorithm makes no assumption about the decay mode in which the photon is produced and therefore its production vertex is not known.)

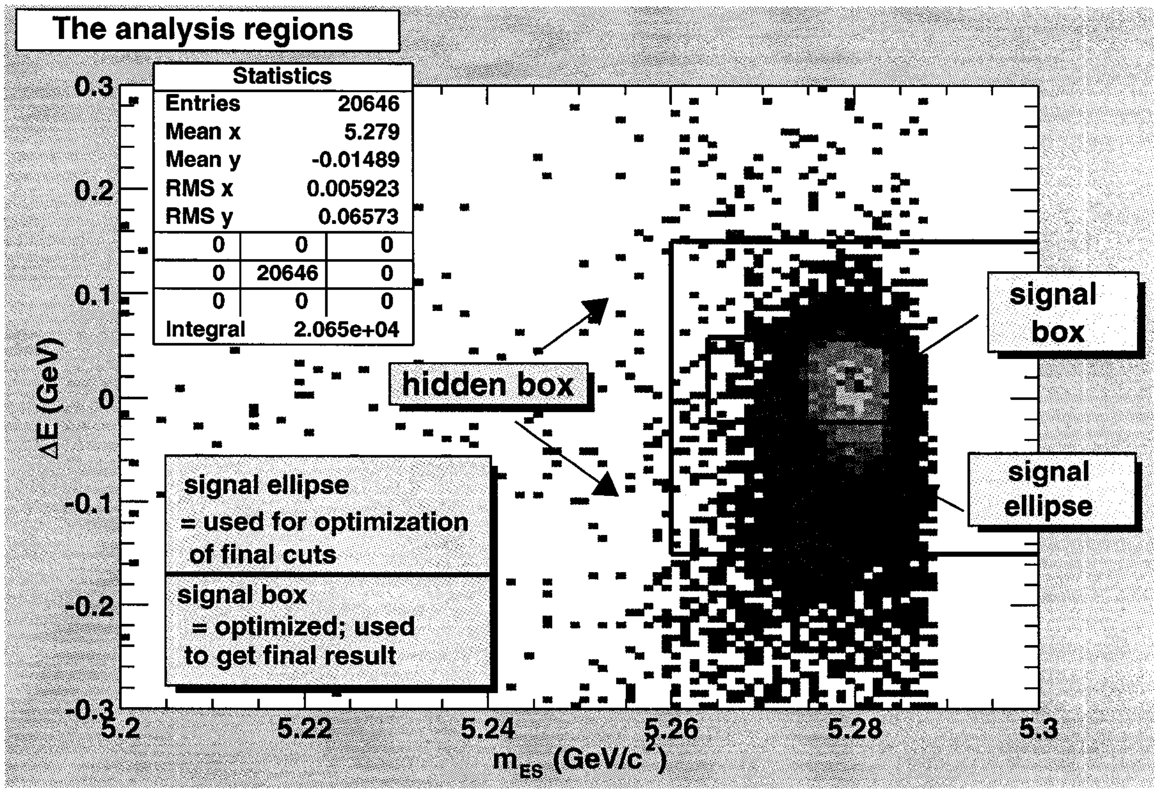

Figure 5.5: The analysis regions. The distribution shown is the signal MC distribution, with no cuts applied (except the preselection). The *analysis window* (AW) is the full ΔE-m_{ES} region shown. For the signal region, we use a *signal ellipse* (SE) to optimize the final cuts; but we use the optimized *signal box* (SB) to obtain the final result. For a blind analysis, we do not look at onpeak data in the *hidden box* until the analysis is complete. The boundaries of the analysis regions are given in Eqs. 5.16-5.19. See the text for details.

5.6.2 The analysis regions

For the blind analysis we define three analysis regions in the ΔE vs. m_{ES} plane, shown in Fig. 5.5. The distribution shown is for signal Monte Carlo; you can see that the signal is indeed concentrated near the expected values of $\Delta E = 0$ and $m_{\mathrm{ES}} = m_B = 5.279\,\mathrm{GeV}/c^2$ for signal events.

- The *analysis window* (AW) is the full region in which the analysis is performed (i.e., the complete histogram in Fig. 5.5). In this study we use the analysis window:

$$5.2 < m_{\mathrm{ES}} < 5.3\,\mathrm{GeV}/c^2 \qquad -0.30 < \Delta E < 0.30\,\mathrm{GeV} \tag{5.16}$$

  Only events in the analysis window are used in the analysis.

- The *signal region* is a small region in which the signal is concentrated. To pass the final selection an event must fall in signal region and pass the final cuts.

We use two signal regions in this analysis: a *signal ellipse* (SE) and a *signal box* (SB). The signal ellipse is a temporary signal region, used for the optimization of the final cuts. The signal ellipse is then discarded, and we optimize the signal region to obtain an optimized signal box. It is this signal box that is used to obtain the final result.

The signal ellipse is a 2.5σ ellipse in ΔE and $m_{\rm ES}$:

$$\left(\frac{m_{\rm ES} - 5.279}{0.0033}\right)^2 + \left(\frac{\Delta E + 0.0027}{0.041}\right)^2 < (2.5)^2 \tag{5.17}$$

(The units are GeV/c^2 and GeV, as usual.) The optimization of the final cuts using the signal ellipse is described in Section 5.8.

The signal box is defined by:

$$5.264 < m_{\rm ES} < 5.284\,\mathrm{GeV}/c^2 \qquad -0.023 < \Delta E < 0.058\,\mathrm{GeV}. \tag{5.18}$$

The optimization of the signal box is described in Section 5.9.

- The *hidden region* (H) is the region kept hidden from the researcher throughout the blind analysis. Prior to unblinding, the experimenter may look at onpeak data *outside* of this region only.[11] Typically the hidden region is a box that surrounds the signal box but is slightly larger, so that events that "leak" out of the tighter signal box remain hidden. For this analysis we use the hidden box:

$$m_{\rm ES} > 5.260\,\mathrm{GeV}/c^2 \qquad -0.15 < \Delta E < 0.15\,\mathrm{GeV} \tag{5.19}$$

This choice is based on Monte Carlo histograms, which showed that most of the signal events are contained in this box (Fig. 5.5).

The hidden region also defines the boundary of the region used in the background estimate. The final background estimate is determined from onpeak data outside of the hidden region.

5.7 The branching fraction

The branching fraction for $B^0 \to J/\psi\gamma$ is given by

$$\mathcal{B} = \frac{n_s}{N_{B\overline{B}}\,\varepsilon_{signal}\,\mathcal{B}(J/\psi \to \ell^+\ell^-)}. \tag{5.20}$$

[11] To be clear: this restriction applies *only* to onpeak data. The researcher may look at Monte Carlo or offpeak data in any region, the hidden region included.

Here, $N_{B\overline{B}}$ is the number of $B\overline{B}$ events in the data samples, obtained from *BABAR*'s `lumi` script, which implements the B-counting method described in *BABAR* Analysis Document 134 [59]. $\mathcal{B}(J/\psi \to \ell^+\ell^-)$ is the branching fraction for $J/\psi \to \ell^+\ell^-$, obtained from the PDG [10] and included to account for the fact that $B^0 \to J/\psi\gamma$ is reconstructed in the $J/\psi \to \ell^+\ell^-$ mode only. n_s is the number of signal events, and ε_{signal} is the efficiency for signal events, defined as the fraction of signal events passing the final selection (final cuts + signal box).

The number of signal events n_s is the observed number of events in the signal region, n_0, minus the expected contribution from background, n_b:[12]

$$n_s = n_0 - n_b. \tag{5.21}$$

During the optimization we minimize a quantity representative of the 90% CL upper limit on the branching fraction, derived assuming that any observed events come from background fluctuations. The method is described in Section 5.8.

For the final result n_0 and n_b are obtained from onpeak data. n_0 is the number of onpeak events in the signal box (kept blind throughout the analysis), and n_b is the number of background events in the signal box, estimated from the number of onpeak events outside of the hidden box. The formula and method are presented in Section 5.15.

5.8 Optimization of the final cuts

In the first stage of the optimization, we hold the signal region constant and optimize the final cuts, using the full set of background and signal MC events. For the optimization, we use as the signal region a 2.5σ signal ellipse ("SE") in ΔE and m_{ES}:

$$\left(\frac{m_{\mathrm{ES}} - m_{\mathrm{ES}0}}{\sigma_{m_{\mathrm{ES}}}}\right)^2 + \left(\frac{\Delta E - \Delta E_0}{\sigma_{\Delta E}}\right)^2 < (2.5)^2 \tag{5.22}$$

We choose a signal ellipse because the signal distribution in ΔE and m_{ES} is elliptical (Fig. 5.6). The values for the means and standard deviations are obtained from Gaussian fits of the ΔE and m_{ES} distributions for signal Monte Carlo samples, shown in Fig. 5.6:

$$m_{\mathrm{ES}0} = 5.279\,\mathrm{GeV}/c^2 \qquad \sigma_{m_{\mathrm{ES}}} = 0.0033\,\mathrm{GeV}/c^2 \tag{5.23}$$

$$\Delta E_0 = -0.0027\,\mathrm{GeV} \qquad \sigma_{\Delta E} = 0.041\,\mathrm{GeV} \tag{5.24}$$

The purpose of the optimization is to find the optimal selection, one that reduces background without losing too many signal events. In analyses with measurable signal, and

[12]To help the reader keep track of all these terms, we have provided a glossary at the end of this chapter.

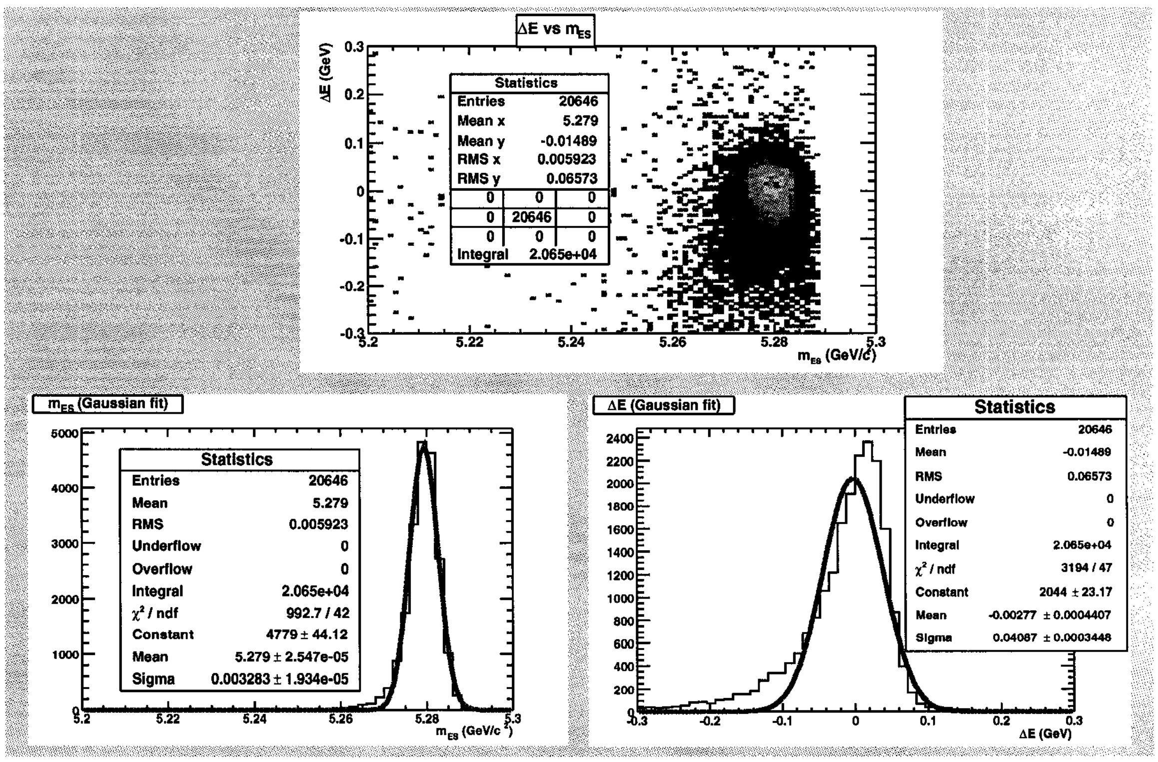

Figure 5.6: ΔE and m_{ES} distributions for the signal Monte Carlo sample. The parameters from the fit are used only to define the signal ellipse in Eqs. 5.23 and 5.24. The final cuts are not applied. Clockwise from the top: ΔE vs. m_{ES} scatter plot, projection on the ΔE axis, and projection on the m_{ES} axis.

thus measurable branching fraction, this is often achieved by maximizing the signal-to-background ratio S/B. However, for rare decay searches the signal is expected to be zero or very close to zero, so this quantity is less appropriate for our analysis. To find a more useful figure of merit, we recall that the aim of the analysis is to measure an upper limit on the branching fraction. Therefore we instead optimize the selection by minimizing a quantity representative of the 90% CL upper limit for the Monte Carlo samples used in the optimization.

To derive the formula we assume that the observed number of events in the signal ellipse is equal to the number of background events expected, $n_0 = n_b$ (see Eq. 5.21). This means that after background subtraction the number of signal events $n_s = 0$, and the uncertainty on n_s is $\sigma(n_s) = \sigma(n_b) = \sqrt{n_b}$, or $\sqrt{B_{SE}}$ in Monte Carlo. Assuming a normal distribution, the 90% CL upper limit is 1.28σ above the peak. The 90% CL upper limit is therefore given by

$$\text{90\% CL upper limit} = \frac{1.28\sqrt{B_{SE}}}{N_{B\bar{B}}\,\varepsilon_{signal}\,\mathcal{B}(J/\psi \to \ell^+\ell^-)}, \tag{5.25}$$

where ε_{signal} is the efficiency for signal events,

$$\varepsilon_{signal} = S_{SE}/S_{total}, \tag{5.26}$$

and S_{total} is the known total number of events in the signal MC sample. S_{total}, $N_{B\bar{B}}$, and $\mathcal{B}(J/\psi \to \ell^+\ell^-)$ are constants, so the quantity that actually varies during the optimization is just $\sqrt{B_{SE}}/S_{SE}$, where S_{SE} and B_{SE} are the number of signal and background events in the signal ellipse, respectively. The number B_{SE} of background events in the signal ellipse is estimated by counting the number of background events in the analysis window and multiplying by the scale factor R_0,

$$B_{SE} = R_0 B_{AW}. \tag{5.27}$$

The advantage of this procedure over simply counting the number of background events in the signal ellipse is that it provides a larger sample of background events. The number of background events in the signal ellipse is relatively low, so it is helpful to increase the size of the background sample to make it more sensitive to the variation of the cuts during the optimization. Also, using background events in the full analysis window reduces bias in that the selection criteria are less likely to be tuned to reject specific background events in the signal ellipse.

The optimization is independent of the value of R_0; we do not need to know R_0 to minimize $\sqrt{B_{AW}}/S_{SE}$. However, R_0 is useful because it allows for an intuitive interpretation of the quantities in the optimization in terms of the estimated number of background events, B_{SE}, and the 90% CL upper limit. To derive the scale factor, we apply a set of reasonable

Table 5.2: Final (optimized) selection cuts.

Variable	Final cut
J/ψ mass	$3.00 < m(J/\psi \rightarrow e^+e^-) < 3.14\,\mathrm{GeV}/c^2$ $3.07 < m(J/\psi \rightarrow \mu^+\mu^-) < 3.13\,\mathrm{GeV}/c^2$
Lepton ID	`PidLHElectron` (likelihood electron selector) `muMicroTight` (cut-based muon selector)
photon LAT	$LAT < 0.40$
π^0 and η vetoes	both applied
Fox Wolfram moment	$R2 < 0.45$
thrust angle	$\cos\theta_t < 0.85$
sphericity angle	$\cos\theta_{sph} < 0.80$
B helicity angle	$\lvert\cos\theta_B\rvert < 0.90$

(but not optimized) selection cuts, and count the number of background events B_{SE} in the signal ellipse and the number B_{AW} in the full analysis window. This gives $R_0 = 0.21$, and we fix the scale factor at this value throughout the optimization.

The next step is to optimize the cuts, one selection variable at a time. The cut for a given selection variable is varied, with the cuts for all of the other variables fixed, and the resulting 90% CL upper limit is calculated. The cut for that selection variable is then fixed at the value that gives the lowest 90% CL upper limit. The procedure is iterative, and is repeated until the cuts stabilize at their optimal values. The final cuts are presented in Section 5.5 and summarized here in Tables 5.2 and 5.3. Also, Figs. 5.7 and 5.8 show plots of ΔE vs. m_{ES} for signal and background Monte Carlo, with and without the final cuts applied, respectively.[13]

5.9 Signal box optimization

In the second stage of the optimization, we fix the final cuts at their optimized values, and optimize the signal region itself. This removes most of the remaining low-ΔE inclusive J/ψ background events missed by the final cuts. The final signal box obtained in the optimization is the one used to determine the final result.

Applying the final cuts eliminates nearly all of the continuum and $B\overline{B}$ background. So the main remaining source of background is inclusive J/ψ events, most of which are concentrated in a region of low ΔE (Fig. 5.8). Checking the Monte Carlo truth information

[13]Note however that the preselection cuts of Section 5.4 were imposed during the production of the ntuples, so the preselection cuts are ALWAYS applied. This is implied throughout this chapter.

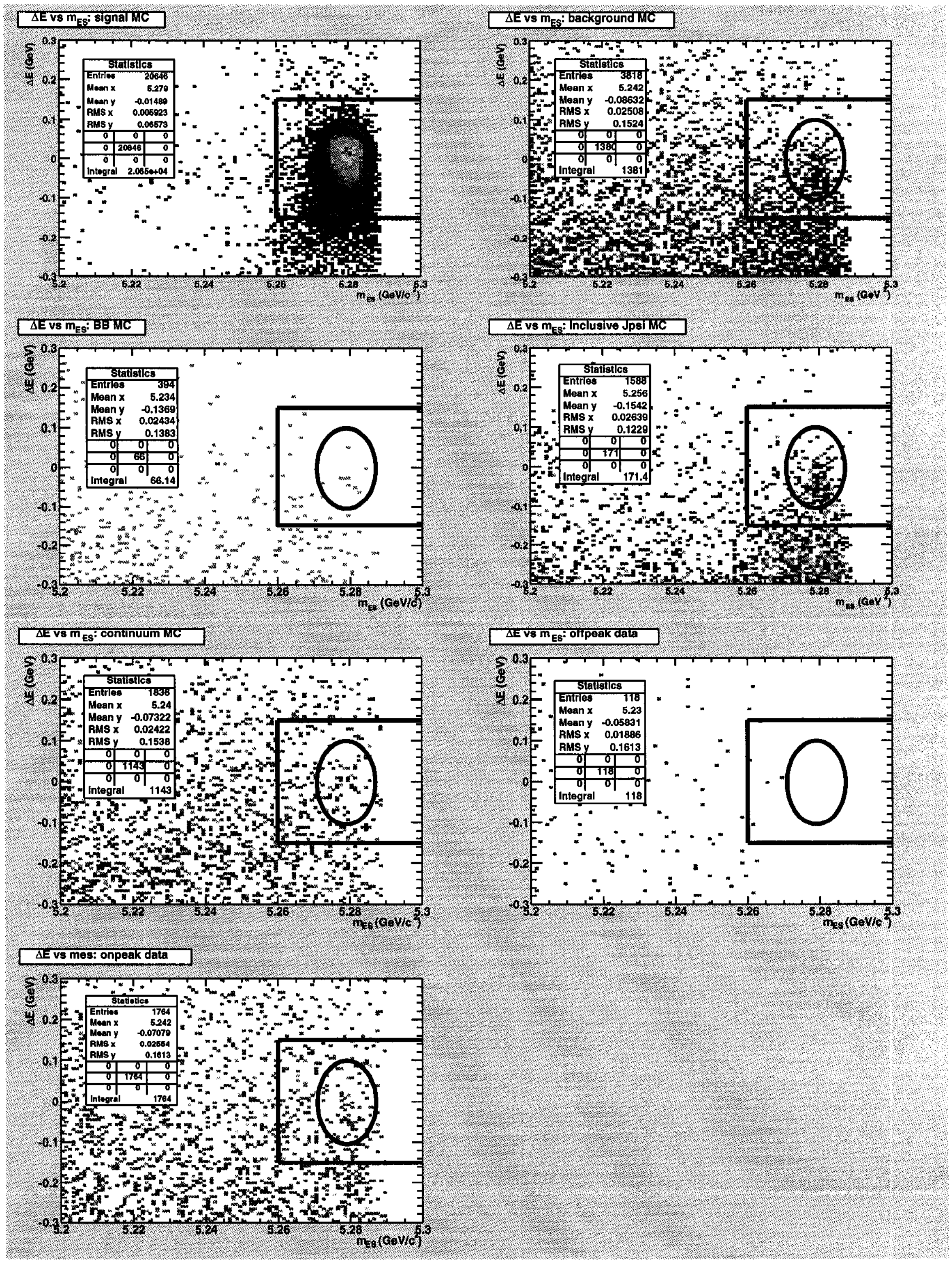

Figure 5.7: ΔE vs. $m_{\rm ES}$ plots, with only the preselection applied (ie, no final cuts). The boundaries shown are those of the hidden box and the signal ellipse (Section 5.6.2), which served as the signal region for the optimization of the final cuts.

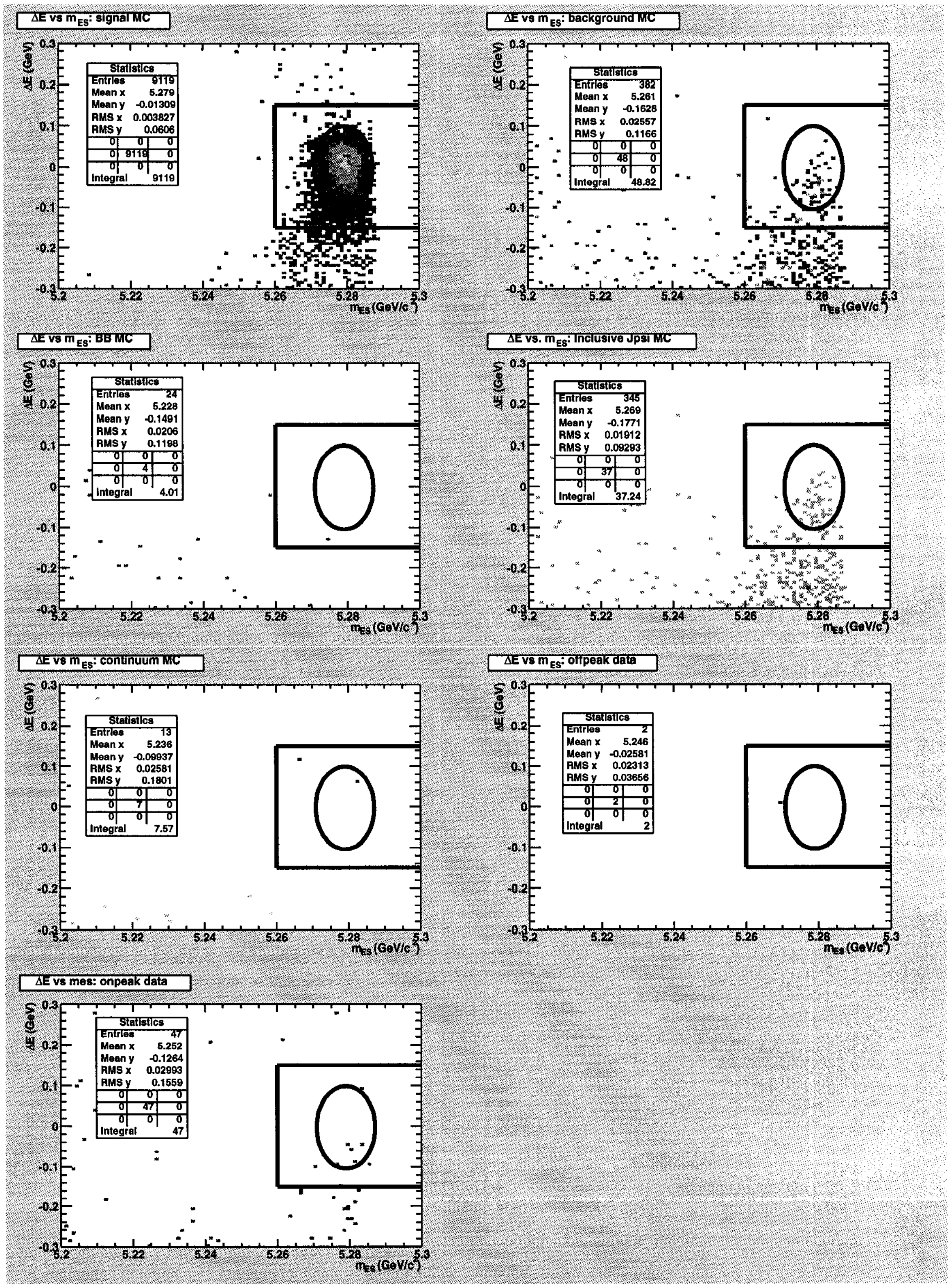

Figure 5.8: ΔE vs. $m_{\rm ES}$ plots with the final cuts (other than the signal region) applied. Nearly all of the continuum and $B\overline{B}$ background is eliminated, so the main remaining source of background comes from inclusive J/ψ events.

Table 5.3: Optimization results for the final cuts. B_{SE}, S_{SE}, ε_{signal}, and the 90% CL upper limit on the branching fraction are described in Section 5.8. For each final cut, we present the effect of turning off that cut only; all the other final cuts are applied. We also present the result with all final cuts applied and with none of the final cuts applied. (Note that the efficiency in the "all final cuts on" column is not the same as the nominal efficiency in Table 5.4 because the results for the present table were obtained for the full MC sample with the signal ellipse as signal region; whereas the final results in Table 5.4 are derived from the half-MC samples "MC-opt" and "MC-final" with the optimized signal box as the signal region.)

Cut turned off	B_{AW}	B_{SE}	S_{SE}	ε_{signal}	90% CL upper limit ($\times 10^{-6}$)
$m(J/\psi \to e^+e^-) > 3.00\,\mathrm{GeV}/c^2$	57.5	12.2	8249	0.212	1.45
$m(J/\psi \to e^+e^-) < 3.14\,\mathrm{GeV}/c^2$	54.5	11.5	8090	0.207	1.44
$m(J/\psi \to \mu^+\mu^-) > 3.07\,\mathrm{GeV}/c^2$	84.8	18.0	8232	0.211	1.77
$m(J/\psi \to \mu^+\mu^-) < 3.13\,\mathrm{GeV}/c^2$	68.8	14.6	8058	0.207	1.63
`PidLHElectron`	73.9	15.7	9128	0.234	1.49
`muMicroTight`	70.9	15.0	9965	0.256	1.33
$LAT < 0.40$	64.4	13.6	8218	0.211	1.54
π^0 veto	54.1	11.5	8067	0.206	1.44
η veto	50.6	10.7	8150	0.208	1.38
$R2 < 0.45$	51.1	10.9	8294	0.213	1.36
$\cos\theta_t < 0.85$	59.8	12.7	8835	0.227	1.38
$\cos\theta_{sph} < 0.80$	88.5	18.7	9319	0.239	1.59
$\lvert\cos\theta_B\rvert < 0.90$	64.3	13.6	8218	0.211	1.54
all final cuts off	1155.2	244.8	17850	0.458	3.01
all final cuts on	31.5	6.7	8188	0.210	1.08

reveals that these are events in which the "X" in $B^0 \to J/\psi X$ is a pion or a state that decays to one or more pions, which nearly always decay via $\pi^0 \to \gamma\gamma$. They are misidentified as $B^0 \to J/\psi\gamma$ events when one of these photons is incorrectly reconstructed as the γ in $B^0 \to J/\psi\gamma$. The two largest single background modes are $B^0 \to J/\psi\ \pi^0$ and $B^0 \to J/\psi\ K^0$, where $K^0 = K^0_S$ or K^0_L decay to pions via $K^0_S \to 2\pi^0$ and $K^0_L \to 3\pi^0$. Presumably the π^0 veto misses these events because the second photon falls outside of the detector acceptance, or is not reconstructed for some other reason.

Because the remaining inclusive J/ψ background is concentrated in a region of low ΔE, we expect that optimizing the signal region will remove much of this background. For our final signal region, we choose a more traditional rectangular box, as the focus in the second stage of the optimization is more on the rejection of the remaining background than on the retention of the elliptical signal.

The procedure for optimizing the signal box is almost the same as that for optimizing the selection cuts. Again, the figure of merit is the 90% CL upper limit given by Eq. 5.25. However, as it is the signal box itself being optimized, to obtain the number of background MC events in the new signal box, B_{SB}, we do NOT scale from the full analysis window via Eq. 5.27. Instead, we count the actual number of background MC events in the signal box. And of course, the signal efficiency is now given by $\varepsilon_{signal} = S_{SB}/S_{total}$, not S_{SE}/S_{total}. We begin with a 2.5σ box in ΔE and $m_{\rm ES}$, vary each boundary with the other three boundaries fixed, and select the boundary that gives the lowest 90% CL upper limit.

With the final, optimized selection cuts applied there are only 59 background events (unweighted `Entries`) remaining in the signal ellipse (Fig. 5.8). With such low statistics there is a risk of "tuning" the final signal box to reject specific background events in the MC sample, thus biasing the result. To avoid bias we therefore use only half of our Monte Carlo sample ("MC-opt") to optimize the signal region, and save the other half ("MC-final") to derive the scale factor B_{SB}/B_{NH} used to obtain the final result (Eq. 5.29).[14]

After three iterations the boundaries stabilize to give the following optimal signal box:

$$5.264 < m_{\rm ES} < 5.284\,{\rm GeV}/c^2 \qquad -0.023 < \Delta E < 0.058\,{\rm GeV}. \tag{5.28}$$

The signal, background, and inclusive J/ψ components of the two Monte Carlo half-samples, with the final cuts applied, are shown in Fig. 5.9. The signal box is superimposed. The results are presented in Table 5.4.[15] For the MC-final (MC-opt) half-sample there are only 6 (4) background events (unweighted `Entries`) in the new signal box, all from inclusive

[14]Of course, this does not prevent the cuts from being tuned to remove specific events in the MC-opt sample. But at least any such tuning will not introduce a bias provided we do not use the MC-opt sample to obtain the final result.

[15]Note that in Fig. 5.9 and Table 5.4 the weights for the MC events — even the signal MC events — are set to twice their usual values to account for the fact that we use only half of the MC sample.

Table 5.4: Results for the optimized signal box. MC-opt is the half-sample of Monte Carlo used to optimize the new signal box; MC-final is the half-sample used to obtain the final result (see Section 5.9). All final cuts are applied. B_{SB}, S_{SB}, ε_{signal}, and the 90% CL upper limit on the branching fraction are described in Section 5.8.

	B_{SB}	S_{SB}	ε_{signal}	90% CL upper limit ($\times 10^{-6}$)
MC-opt	0.863	5696	0.146	0.559
MC-final	1.296	5420	0.139	0.720

J/ψ MC. The low cut on ΔE does indeed eliminate most of the inclusive J/ψ background seen in Fig. 5.8.

Thus, once the signal box is optimized, we have the following things, which we use to obtain the final result:

- A set of final, optimized selection cuts.
- A final, optimized signal box.
- A MC-final sample of Monte Carlo events.

For clarity, we note again that the following are NOT used to obtain the final result:

- The signal ellipse.
- The MC-opt sample.

5.10 Estimate of background in the signal box

For the final result, of course, we use the actual onpeak *data* samples, not the Monte Carlo samples. In this case, the number of signal events, n_s, is calculated by (Eq. 5.21)

$$n_s = n_0 - n_b$$

where n_0 is the number of onpeak events that pass the final selection (final cuts + signal box), and n_b is number of background events that pass the final selection. With the final cuts applied, the number of background events n_b in the signal box is estimated from the number of onpeak events in the not-hidden region, n_1, via

$$n_b = \frac{B_{SB}}{B_{NH}} n_1 \tag{5.29}$$

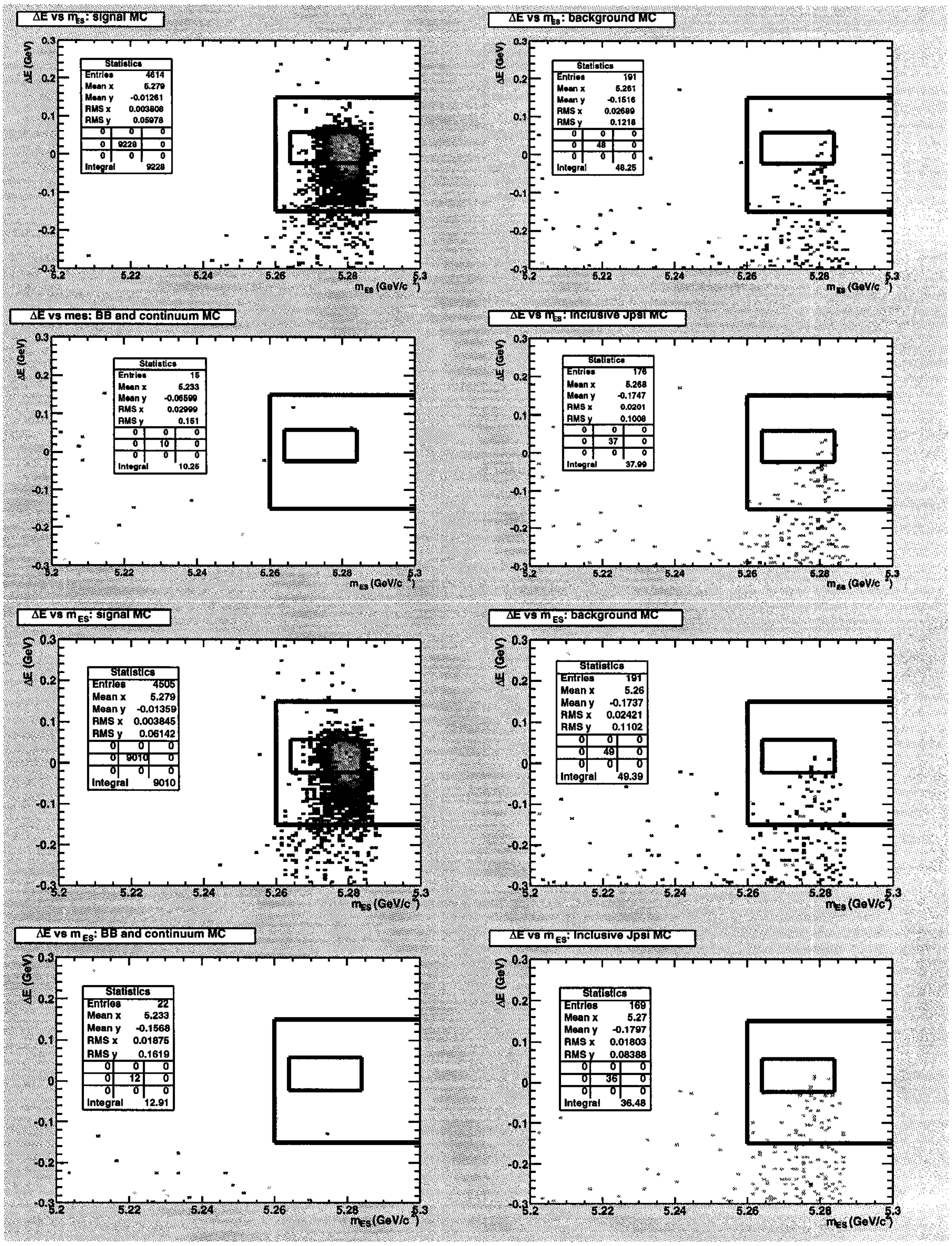

Figure 5.9: The optimized signal box. It is this signal box, not the signal ellipse, that is used to determine the final result (see Section 5.9). Top: signal, background, continuum/$B\overline{B}$ and inclusive J/ψ plots for the "MC-opt" sample used to optimize the new signal box. Bottom: the same plots for the "MC-final" sample used to obtain the final result. (These two MC samples are described in Section 5.9.) All final cuts are applied (except the signal box).

where B_{SB} is the number of background MC-final events in the signal box, and B_{NH} is the number of background MC-final events outside of the hidden box ("NH" stands for "not-hidden").

We have $B_{SB}(MC) = 1.3 \pm 0.5$, $B_{NH}(MC) = 37 \pm 4$, and $n_1 = 38 \pm 6$, giving a background estimate of $n_b = \frac{1.3}{37} \times 38 = 1.3$. The systematic error on this quantity is just the sum (in quadrature) of the statistical errors on each quantity, $\sigma_b = 1.3 \times \sqrt{\left(\frac{0.5}{1.3}\right)^2 + \left(\frac{4}{37}\right)^2 + \left(\frac{6}{38}\right)^2} = 0.6$, or 46%.

5.11 Background cross-checks

The aim of background cross-checks is to confirm that background MC is correctly modeling the true background distribution in onpeak data. While we can never be 100% certain what is signal and what is background in data, for a rare decay like $B^0 \to J/\psi\gamma$ we expect all except possibly one or two data events to be background events. So we can cross-check our background estimate by comparing the distributions of various quantities in background MC and data, ensuring that they are similar and that we obtain consistent background estimates from each. Although the blind analysis forbids us to look at onpeak data in the hidden region, we are free to use onpeak data in the not-hidden region for our cross-checks.[16]

We perform two cross-checks of the Monte Carlo modeling of the background, using the full Monte Carlo and data samples. We consider two quantities:

- N_{cont}, the number of continuum events, can be obtained from two sources:
  - N_{cont} = the number of offpeak data events in the analysis window. Offpeak data is collected at an energy about 40 MeV below the $\Upsilon(4S)$ resonance energy; however, for the cross-check ΔE and m_{ES} are calculated with the beam energy fixed to the onpeak energy. Otherwise the 40 MeV shift in the center of mass energy would make the comparison between offpeak and continuum samples invalid.
  - N_{cont} = the expected number of continuum events from continuum Monte Carlo. This is the number of continuum MC events in the analysis window, scaled to the offpeak luminosity.

The results are presented in Table 5.5. Monte Carlo predicts 0.80 ± 0.22 continuum events, and we observe 2 offpeak data events. The (Poisson) probability of observing

[16] Actually, similar cross-checks can be performed in studies of non-rare decays as well. In this case, the restriction to the not-hidden region has a dual purpose: first, to stay blind; and second, to restrict the estimate to a kinematic region expected to contain only background events. For rare decays the latter is of course less of a concern.

Table 5.5: Cross-check between continuum and offpeak data predictions for the number of continuum background events in the analysis window. We use the full data and MC samples for the cross-check. The error on the number of `Entries` is just the statistical error, $\sqrt{\texttt{Entries}}$. The scale factor is used to weight the events to the offpeak luminosity. 2 offpeak data events are observed, compared to a predicted 0.80 ± 0.22 from continuum MC. The probability of observing 2 or more events when 0.80 are expected is a relatively high 0.2, so the results are consistent.

	`Entries`	Error(`Entries`)	scale	N_{cont}	Error(N_{cont})
c	9	3	0.058	0.52	0.17
uds	4	2	0.070	0.28	0.14
τ	0	0	0.082	0	0
total continuum MC				0.80	0.22
offpeak data	2		1	2	

2 or more events when 0.80 are expected is 0.2, which is relatively high. So the results from continuum Monte Carlo and offpeak data are consistent.

- N_{back}, the number of background events in the not-hidden region, can be estimated from two sources:
  - N_{back} = the number of onpeak data events in the not-hidden region.
  - N_{back} = the number of background Monte Carlo events in the not-hidden region, scaled to the onpeak luminosity.

  The results are presented in Table 5.6. For the not-hidden region, Monte Carlo predicts 36.2 ± 2.6 background events, and we observe 38 background events. So again the predictions from background Monte Carlo and onpeak data are consistent.

We note that the background estimates do not take into account differences between tracking and PID efficiencies in data and Monte Carlo. These could cause the background estimate to vary by about 10%, but this would not change the fact that the results from MC and data are consistent.

5.12 Efficiency corrections

The signal efficiency ε_{signal} in Eq. 5.20 is defined as the fraction of signal events in data that pass the final selection (final cuts + signal box). However, it is determined from Monte Carlo

Table 5.6: Cross-check between Monte Carlo and onpeak data predictions for the number of background events in the not-hidden region. We use the full data and MC samples for the cross-check. The error on the number of `Entries` is just the statistical error, $\sqrt{\texttt{Entries}}$. The scale factor is used to weight the events to the onpeak luminosity. 38 onpeak data events are observed, compared to a predicted 36.2 ± 2.6 from background MC. So the results from background Monte Carlo and onpeak data are consistent.

	`Entries`	Error(`Entries`)	scale	N_{back}	Error(N_{back})
c	8	2.83	0.55	4.38	1.55
uds	2	1.41	0.66	1.32	0.93
τ	0	0	0.77	0	0
$B^0\overline{B}^0$	5	2.24	0.17	0.85	0.38
$B^{\pm}$ data	18	4.24	0.17	2.99	0.70
Inclusive J/ψ	247	15.72	0.11	26.66	1.69
Total background MC				36.21	2.60
onpeak data	38		1	38	

samples. To correct for differences between Monte Carlo and data the standard procedure at *BABAR* is to apply corrections to the Monte Carlo efficiency. Each correction comes with its own systematic error.[17] The required corrections and systematic errors, or the method to obtain these, are determined by the appropriate Analysis Working Group for the part of the analysis in question. For $B^0 \to J/\psi\gamma$ we need to correct for the photon and J/ψ reconstruction efficiencies:

- **Photon reconstruction efficiency** The method for correcting the efficiency for photons and other neutral particles is determined by the *BABAR*'s Neutrals Working Group. For isolated photons no killing but smearing and energy rescaling have to be applied.[18] Studies similar to ours indicate that there should be no efficiency correction for the photon in $B^0 \to J/\psi\gamma$ [64]. But we must still include the associated systematic error of 2.5%.

- **J/ψ reconstruction efficiency** The J/ψ is reconstructed from its decay to a lepton pair, $J/\psi \to \ell^+\ell^-$. For the leptons efficiency corrections are needed both for tracking and particle identification (PID):

[17]Systematic errors are errors not due to statistical fluctuations in the data sample under study [69].

[18]Killing, smearing, and energy rescaling refer to *BABAR* efficiency correction strategies for data/MC differences in PID, resolution, and energy distributions, respectively.

Table 5.7: Signal efficiency and efficiency corrections.

Nominal efficiency	0.139		
Source	Correction	Cumulative efficiency	Systematic error (%)
Photon reconstruction	1.0	0.139	2.5
Lepton tracking	0.984	0.137	4
Lepton PID	0.95 (est.)	0.130	2 (est.)
Corrected efficiency		0.130	5.1

- **tracking efficiency** Tracking efficiency is the responsibility of the *BABAR*'s Tracking Efficiency Task Force. The Task Force's recommendations depend on the properties of the leptons in the given analysis. In our analysis, the lepton selection is imposed via the `JpsiELoose` and `JpsiMuLoose` tag bits. This means that the lepton candidates are derived from `ChargedTracks`. Also, they have well measured momenta $p_t > 180\,\mathrm{GeV}/c$. For this situation the Task Force recommends an efficiency correction of 0.8% per track, with an associated systematic error of 2% per track. As these are correlated errors we add them to obtain an efficiency correction of 1.6% (multiplicative correction 0.984) and systematic error of 4% per J/ψ.
- **PID efficiency** The performance of the particle selectors is different for data and Monte Carlo. We use a reasonable estimate of 0.95 or 5% for the lepton PID efficiency correction, and 2% for the associated systematic error.

The efficiency and efficiency corrections are summarized in Table 5.7.

5.13 Quantities and errors used in the final calculation

We use the following quantities in the final calculation of the 90% CL upper limit on the branching fraction. They are summarized in Table 5.8.

- $N_{B\overline{B}}$ *BABAR* counts the number of produced $B\overline{B}$ pairs, $N_{B\overline{B}} = 123.3 \times 10^6$, using the B-counting method described in *BABAR* Analysis Document 134 [59]. The systematic error is 1.1%, or 1.4×10^6 $B\overline{B}$ pairs [70].

- $\mathcal{B}(J/\psi \to \ell^+\ell^-)$ The branching fraction for $J/\psi \to \ell^+\ell^-$ decays is the sum of the $J/\psi \to e^+e^-$ and $J/\psi \to \mu^+\mu^-$ branching fractions, obtained from the PDG [10].

  $\mathcal{B}(J/\psi \to e^+e^-) = (5.93 \pm 0.10) \times 10^{-2}$

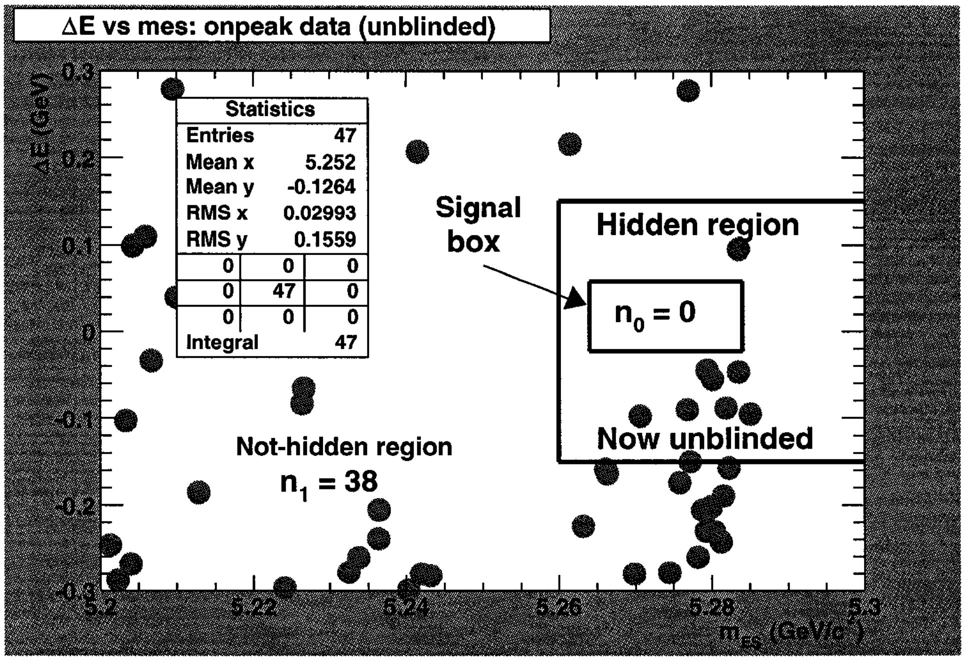

Figure 5.10: Obtaining the final results. All final cuts are applied. n_1 is the number of onpeak events in the not-hidden region, used to estimate n_b (Section 5.10). n_0 is the number of onpeak events in the signal box. n_0 was kept blind throughout the analysis.

$$\mathcal{B}(J/\psi \to \mu^+\mu^-) = (5.88 \pm 0.10) \times 10^{-2}$$

$$\mathcal{B}(J/\psi \to \ell^+\ell^-) = \mathcal{B}(J/\psi \to e^+e^-) + \mathcal{B}(J/\psi \to \mu^+\mu^-) = 0.1181$$

Assuming fully correlated errors, the systematic error is 0.0020, or 1.7%.

- n_0 The number of onpeak events passing the final selection (final cuts + signal box). There is no systematic error on this quantity, only statistical error. We observe $n_0 = 0$.

- n_b The estimated number of onpeak background events in the signal box, estimated from onpeak data in the not-hidden region, was calculated in Section 5.10 to be $n_b = 1.3$, with a systematic error of 0.6, or 46%.

- ε_{signal} The nominal efficiency was determined from the "MC-final" half-sample of signal MC with all final cuts applied (Table 5.4). The corrected signal efficiency was calculated from the nominal efficiency in Section 5.12 to be $\varepsilon_{signal} = 0.130$, with a systematic error of 5.1%, or 0.0067.

Table 5.8: Quantities used in the final calculation.

	Value	Error		Value	Error
n_0	0		$N_{B\bar{B}}$ ($\times 10^6$)	123.3	1.4
n_b	1.3	0.6	B ($J/\psi \to \ell^+\ell^-$)	0.1181	0.0020
			ε_{signal}	0.130	0.0067

5.14 Prediction for the number of onpeak events in the signal box, n_0

We can calculate a prediction for n_0 as follows:

Inverting Eq. 5.20 and inserting the results from Table 5.8 and Lu, Wang, and Yang's [1] prediction for the branching fraction (Eq. 1), $\mathcal{B}(B^0 \to J/\psi\gamma) = 7.65 \times 10^{-9}$, gives:

$$\begin{aligned} n_s &= N_{B\bar{B}}\, \varepsilon_{signal}\, \mathcal{B}(J/\psi \to \ell^+\ell^-)\, \mathcal{B}(B^0 \to J/\psi\gamma) \\ &= (123.3 \times 10^6)(0.130)(0.1181)(7.65 \times 10^{-9}) \\ &= 0.0145. \end{aligned}$$

The expected number of onpeak events passing the final selection (final cuts + signal box) is therefore

$$n_0 = n_s + n_b = n_b = 1.3. \tag{5.30}$$

So we expect to observe only about 1 event, from background.

We count $n_0 = 0$ events passing the final selection. The (Poisson) probability of measuring $n_0 = 0$ events when 1.3 are expected is 0.27 or 27%. This is quite large, so the results are consistent with predictions.

5.15 Determination of the 90% CL upper limit on the branching fraction

As emphasized throughout this chapter, in a blind analysis the final result is kept hidden until the analysis is complete. This works because the value of the measurement contains no information about its correctness. So not only is the final result not needed to perform the analysis, but it could be harmful, because knowledge of the final result could bias the experimenters [71]. To avoid bias, therefore, it is important to decide *before* unblinding how the results will be interpreted.

For a hidden signal region analysis, "the final result" is just a single number: the number of onpeak events in the signal region, n_0. So we simply have to decide what is to be done with that number.

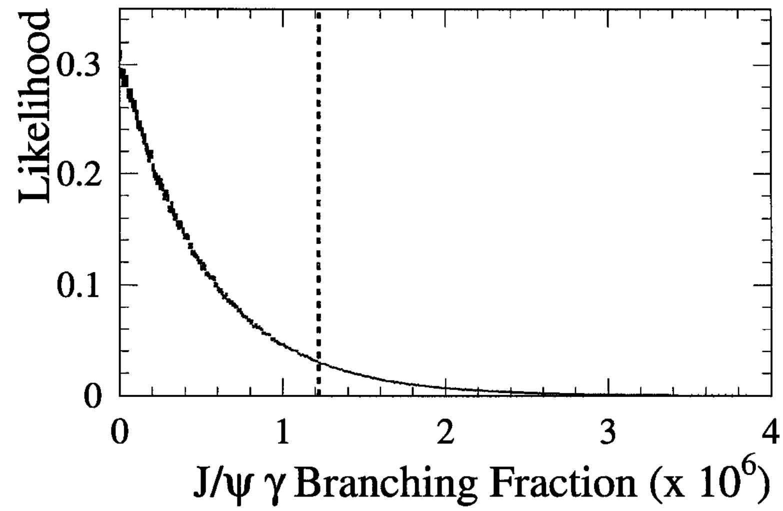

Figure 5.11: Likelihood function of the branching fraction, for $n_0 = 0$. The central value of the branching fraction is the value of $\mathcal{B}$ at which the likelihood function peaks — in this case, zero. The 90% CL upper limit is the value of $\mathcal{B}$ with 90% of the area of the likelihood function to its left, and 10% to its right. This is indicated by the vertical dashed line. We obtain a 90% CL upper limit of 1.2×10^{-6}.

The formula for the branching fraction was given in Eq. 5.20:

$$\mathcal{B} = \frac{n_s}{N_{B\overline{B}}\, \varepsilon_{signal}\, \mathcal{B}(J/\psi \to \ell^+\ell^-)}.$$

So if we had enough events to measure the branching fraction, then we would simply subtract the estimated background n_b from n_0 to get the number of signal events n_s, and then use this result in Eq. 5.20 to obtain the branching fraction.

However, as expected n_0 and n_b are far too small to obtain a measurement of the branching fraction. That is why the aim of the analysis is to set a 90% CL upper limit on the branching fraction, instead. In this case the calculation is not quite so straightforward, because the unlike the branching fraction, the 90% CL upper limit is not a simple analytic function of n_s.

The 90% CL upper limit is defined so that, if we interpret our results to mean "$\mathcal{B}$ < the upper limit," then it is 90% probable that this interpretation is correct. To calculate the 90% CL upper limit we use the likelihood function for the branching fraction, determined using the method used by Chris Hearty in the analysis for $B^+ \to J/\psi p\overline{\Lambda}$ and $B^0 \to J/\psi p\overline{p}$ [72]. An overview of the method is provided in Appendix C.

Fig. 5.11 shows the likelihood function plot for the observed value of $n_0 = 0$. This is a plot of the probability of observing exactly $n_0 = 0$ events — the *likelihood* — as a function of the branching fraction. The likelihood function is used to obtain both a central

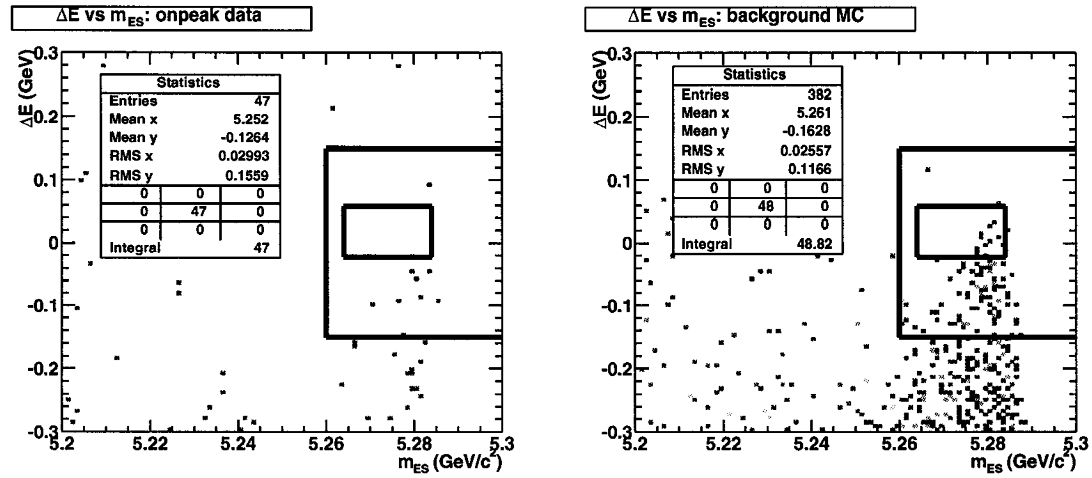

Figure 5.12: ΔE vs. m_{ES} distributions for onpeak data (left) and background MC (right), with all final cuts applied (except the signal box). The onpeak data plot is reproduced from Fig. 5.10. The number of events in the analysis window (the `Integral` value) is 47 for onpeak data, and 48.8 for signal Monte Carlo. As expected, these numbers are consistent.

value and an upper limit on the branching fraction. The central value is the value of $\mathcal{B}$ at which the likelihood function peaks. From Fig. 5.11 you can see that the central value for $\mathcal{B}(B^0 \to J/\psi\gamma)$ is zero. The 90% CL upper limit is the value of $\mathcal{B}$ (indicated by the vertical dashed line) with 90% of the area of the likelihood function to its left, and 10% to its right. We obtain a 90% CL upper limit of 1.2×10^{-6} for $\mathcal{B}(B^0 \to J/\psi\gamma)$.

5.16 Discussion of results

Fig. 5.12 presents the ΔE-m_{ES} distributions for onpeak data (left) and background Monte Carlo (right). (The onpeak plot is the same as the one in Fig. 5.10.) We count $n_0 = 0$ events passing the final selection (final cuts + signal box). This is consistent with the expected $n_0 = 1.3$ events from background; the (Poisson) probability of observing $n_0 = 0$ events when $n_b = 1.3$ are expected is 27%, which is high. As expected, there is no evidence for a signal in this rare decay mode.

We set a 90% CL upper limit of 1.2×10^{-6} on the branching fraction for $B^0 \to J/\psi\gamma$. This is the first measured upper limit for $\mathcal{B}(B^0 \to J/\psi\gamma)$. The result is quite similar to *BABAR*'s recent results for other rare radiative decays to a vector meson [41].[19] For $B^0 \to \rho^0\gamma$ *BABAR* measured a 90% CL upper limit of 1.2×10^{-6}, and for $B^0 \to \omega\gamma$, *BABAR* measured a 90% CL upper limit of 1.0×10^{-6}. The ability to measure or set an upper limit on the branching fraction is a direct reflection of the size of the *BABAR* data set.

[19] Note, however, that the dynamics for $B^0 \to J/\psi\gamma$ are different — see Section 2.4.1

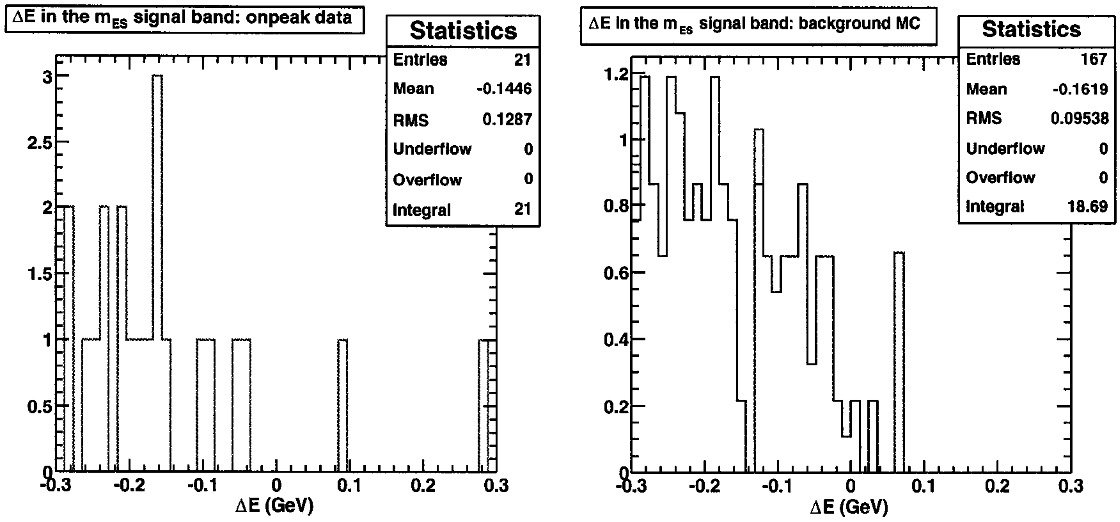

Figure 5.13: ΔE in the m_{ES} signal band, for onpeak data and background Monte Carlo. As in previous plots, the onpeak distribution is shown in purple on the left, and the background distribution is shown in red on the right, with the contribution to the background from inclusive J/ψ shown in blue. As expected, the plots for onpeak data and background MC are similar. We note again that the main source of background is low-ΔE inclusive J/ψ events.

For rare decays like $B^0 \to J/\psi\gamma$, we expect the real onpeak data sample to contain only a few signal events at the most. So if the background Monte Carlo sets accurately model the data, then plots for background MC and onpeak data should be very similar. We already noted in Section 5.5.5 that this is indeed the case in the plots of the selection cuts presented in Section 5.5 (Figs. 5.1 to 5.3).

The ΔE-m_{ES} distributions should also be similar for onpeak data and background MC. We note that the number of events in the full analysis window is 47 for onpeak data, consistent with the 48.82 events observed for background Monte Carlo.[20] This increases our confidence in the Monte Carlo modeling of the background.

The ΔE distributions in the m_{ES} signal band are presented in Fig. 5.13. Again we see that the onpeak and background distributions are very similar. As previously noted, the primary source of background is inclusive J/ψ events in the low-ΔE region. Specifically, the main background modes are $B^0 \to J/\psi\ \pi^0$ and $B^0 \to J/\psi\ K^0$, where $K^0 = K^0_S$ or K^0_L decay to pions via $K^0_S \to 2\pi^0$ and $K^0_L \to 3\pi^0$. All of these events contain neutral pions in the final state, which decay nearly 100% of the time to photons via $\pi^0 \to \gamma\gamma$. They are background

[20] It *looks* like the background MC sample has many more events than the onpeak sample, but that's just because there are more `Entries` in the background MC histogram. It is the `Integral` value that gives the correctly-weighted number of events.

events because one of the photons is misreconstructed as coming from $B^0 \to J/\psi\gamma$ rather than $\pi^0 \to \gamma\gamma$. For some reason, the neutral pion veto misses these events. Perhaps the second photon is not reconstructed, either because it falls outside of the detector acceptance, or for some other reason. In any case, the low cut on ΔE imposed by the optimized signal box eliminates the majority of these remaining background events.

As a side note, it is interesting to calculate the number of events that would have been required for a measurement of the branching fraction. Of course, this number depends on how tightly we define what counts as a "signal." In terms of Poission statistics, a reasonable requirement is to demand that the probability of the observed signal being just a statistical fluctuation be less than 10^{-3} or 10^{-4}. Using a Poisson statistics table and assuming a mean of 1.3 (the expected signal from background) gives a probability of 4×10^{-4} to observe $n_0 = 7$ or more events when 1.3 are expected. So 7 events could be considered a signal.

5.17 Conclusion

This chapter described the measurement of the 90% CL upper limit on the branching fraction for $B^0 \to J/\psi\gamma$. The result is a 90% CL upper limit of 1.2×10^{-6} for $\mathcal{B}(B^0 \to J/\psi\gamma)$. This is the first upper limit measured for this rare radiative decay. As for the rare radiative decays $B \to \rho\gamma$ and $B \to \omega\gamma$, the measurement for $B^0 \to J/\psi\gamma$ is statistically limited. Therefore our 90% CL upper limit is of the same order of magnitude as *BABAR*'s measured $\mathcal{O}(10^{-6})$ upper limits for $B \to \rho\gamma$ and $B \to \omega\gamma$.

There are a few "loose ends" remaining to be tied up for this analysis. The plan is to undertake a more rigorous study of the efficiency corrections and systematic errors from data/Monte Carlo differences in lepton PID (Section 5.12), using the PID tables of data efficiencies maintained by the Particle Identification Analysis Working Group. However, as in most rare decay searches the systematic error in this analysis is negligible compared to the dominant statistical error, so this should not affect the final result.

With the optimized set of final cuts applied, the main source of background is inclusive J/ψ events in which a photon from $\pi^0 \to \gamma\gamma$ is misreconstructed as coming from a $B^0 \to J/\psi\gamma$ decay. Much of this background is removed by tightening the lower cut on the kinematic variable ΔE. A suggestion for future studies is to find a way to improve the veto of photons from $\pi^0 \to \gamma\gamma$ decays. Another good idea would be to fit the ΔE distribution to obtain a better background estimate with smaller uncertainty. Again, this would not affect the current analysis as it is limited by statistical error. However, it should be important in studies using larger samples of $B\overline{B}$ events.

5.18 Glossary of analysis terms

Monte Carlo counting:

These numbers vary throughout the optimizations.

- B_{AW} = the number of background MC events in the analysis window.
- B_{NH} = the number of background MC events in the not-hidden region.
- B_{SE} = the number of background MC events in the signal ellipse. Estimated from B_{AW}.
- B_{SB} = the number of background MC events in the signal box.
- S_{SE} = the number of signal MC events in the signal ellipse.
- S_{SB} = the number of signal MC events in the signal box.

Onpeak data counting:

These numbers are results from onpeak data; they don't vary.

- n_0 = the number of onpeak events passing the final selection (final cuts + signal box).
- n_1 = the number of onpeak events in the not-hidden region that pass the final cuts.
- n_s = the number of signal events in the signal box. Calculated via $n_s = n_0 - n_b$.
- n_b = the number of background events in the signal box. Estimated from n_1.

Chapter 6

Conclusion

This thesis described the measurement of the 90% CL upper limit on the branching fraction for $B^0 \to J/\psi\gamma$. We obtain a result of

$$\mathcal{B}(B^0 \to J/\psi\gamma) < 1.2 \times 10^{-6} \quad \text{at 90\% CL.}$$

This is the first limit measured for this decay mode.

The next generation of Super$B$-factories will have statistics at least two orders of magnitude higher than those used in this analysis. While the current *BABAR* and Belle experiments cannot measure branching fractions of the order of $\mathcal{O}(10^{-8})$, the upcoming Super*BABAR* and SuperBelle experiments should be able to obtain much more sensitive measurements. If $B^0 \to J/\psi\gamma$ is indeed enhanced by New Physics it may be possible to measure this enhancement at these Super$B$-factories.

Chapter 7

Epilogue

Since the submission of this thesis in February 2004, a number of improvements have been made to the analysis, including:

- An improved veto of photons from $\pi^0 \to \gamma\gamma$ decays.
- The reoptimization of the final selection in light of this change.
- More detailed background studies using MC truth information.
- More detailed background studies using signal MC ntuples for the three main background modes: $B^0 \to J/\psi\ \pi^0$, $B^0 \to J/\psi\ K^0_S$, and $B^0 \to J/\psi\ K^0_L$.
- A study to determine the efficiency corrections and systematic errors required for lepton PID. These replace the estimates in Section 5.12. We also include a new correction for data/MC differences in ΔE.

As a result of these improvements, we have significantly reduced the background and obtained a more accurate calculation of the efficiency. Although the measured 90% CL upper limit on the branching fraction has increased slightly, to 1.6×10^{-6}, we strongly believe that the changes are for the best. In particular, our new estimated result of $n_0 = 0.72$ is even more consistent with the measured $n_0 = 0$; the (Poisson) probability of observing $n_0 = 0$ events when $n_b = 0.72$ are expected is 48.7%, which is quite high.

(I should note also that with the *BABAR* collaboration's help, we were able to make these changes in such a way that our knowledge of the unblinded result did not introduce any bias.)

We expect to publish the results for the $B^0 \to J/\psi\gamma$ analysis sometime this summer (2004).

Appendix A

Measuring $\sin 2\beta$ at *BABAR*

This section gives a very brief explanation of how the the *CP*-violating parameter $\sin 2\beta$ is measured at *BABAR*. This has nothing to do with $B^0 \to J/\psi\gamma$. However, it is useful as background to Chapter 3, given that measurement of $\sin 2\beta$ was the primary motivation for the *BABAR* experiment and therefore determined the experimental objectives of both the PEP-II accelerator complex and the *BABAR* detector, in design and execution.

The operation *CP* is the combined operation of charge conjugation C, which changes a particle into its antiparticle; and the parity operator P, which reverses the sign of spatial coordinates. So, for example, applying *CP* to a left-handed electron gives a right-handed positron. "*CP*-opposite" particles or decays are known as "*CP*-conjugates." *CP* violation refers to the fact that *CP*-conjugate processes may behave differently. For example, they may occur at different rates.

e^+e^- colliders at the $\Upsilon(4S)$ resonance produce $B\overline{B}$ pairs in a coherent, $L = 1$ state, and they remain in that state until one of them decays. To measure *CP* violation, *BABAR* uses events in which one of the B mesons (call it B_{CP}) decays to a *CP* eigenstate f, and the other (call it B_{tag}) decays to a *tagging mode* that reveals its identity (B^0 or $\overline{B}^0$). The *time-dependent CP asymmetry* can be written as:

$$a_f(\Delta t) = \frac{N(\overline{B}_{CP} \to f) - N(B_{CP} \to f)}{N(\overline{B}_{CP} \to f) + N(B_{CP} \to f)} \tag{A.1}$$

The B_{tag} decays to a lepton or kaon, whose charge reveals whether B_{tag} is a B^0 or $\overline{B}^0$ meson. This in turn can be used to determine the probability that its partner B_{CP} is a B^0 or $\overline{B}^0$ meson, as a function of $\Delta t = (t_{CP} - t_{tag})$, the time difference between the B_{CP} and B_{tag} decays. Δt is related to the vertex separation of the two B mesons, $\Delta z = z_{CP} - z_{tag}$, via $\Delta t = \Delta z/\beta\gamma$.

That takes care of the *CP* asymmetry measurement. However, the goal at *BABAR* is not only to measure the time-dependent *CP* asymmetry, but also to relate it to the CKM angle

β.[1] For example, for B decays to charmonium final states ($b \to c\,\bar{c}\,s$) there is a simple relation between $a_f(\Delta t)$ and sin2β:

$$a_f(\Delta t) = -\sin 2\beta \sin(\Delta t \Delta M) \tag{A.2}$$

The mixing frequency ΔM has been known since it was measured by ARGUS back in 1987 [45], and has been determined with greater accuracy since then. You can look it up in the PDG [10]. So to measure sin2β, you need to do three things: measure Δz, and reconstruct the B_{CP} and B_{tag} decays. As described in Chapter 3, both PEP-II and the *BABAR* detector were designed with these three tasks in mind.

[1]A similar method can be used to measure sin2α as well.

Appendix B

BABAR's lepton selectors

The purpose of this appendix is to provide some background information regarding the lepton selectors referred to in Chapter 5.

At *BABAR*, particle identification is standardized for quality and consistency between analyses. Particle identification is implemented via *selectors*, sets of selection variables relevant to the given particle. The electron and muon selectors are developed and maintained by the Electron and Muon AWGs, and coded in the standard `BetaPid` software package.

Muon identification Muon identification is based mostly on information from the IFR. The basic idea is to look for charged particle candidates (that is, particles that leave charged tracks) that do not shower in the calorimeters, and that penetrate far into the IFR. The selector variables for muon identification are [62]:

- The energy released in the EMC.
- The number of IFR hit layers in a cluster.
- Whether or not the cluster has a hit in the inner RPC.
- The first IFR hit layer in the cluster.
- The actual number of interaction lengths traversed by the track in the *BABAR* detector; and the expected number of interaction lengths traversed assuming the particle is a muon.
- The success of the fitting algorithms, as measured by the χ^2 per degree of freedom of the IFR hit strips with respect to the track extrapolation, and with respect to the 3rd order polynomial fit of the cluster.
- The total number of IFR hit strips in a given layer.

- The total number of IFR hit strips in a given cluster.

Electron identification Electron identification is based mainly on information from the EMC and the DCH. The basic idea is to look for charged particle candidates with electromagnetic-like (as opposed to hadronic-like) showers in the EMC. The selector variables for electron identification are [61]:

- E/p, the ratio of energy deposited in the calorimeter to the momentum of the track.
- $\mathrm{d}E/\mathrm{d}x$, the ionization energy loss in the DCH and SVT.
- The number of crystals in the cluster.
- LAT, the lateral energy distribution [63].
- A_{42}, the Zernike moment of order (4,2) [73].
- θ_c, the Cerenkov angle in the DIRC.
- Track-cluster matching in the EMC and DCH.

Cut-based selectors identify the particles via cuts on the selector variables. The user has the choice of five selectors. The first four are `VeryTight`, `Tight`, `Loose`, and `VeryLoose`. The fifth selector is even looser — the `Minimum Ionizing` selector for muons allows even some muons identified only in the EMC to pass the selection; while the `noCal` selector for electrons allows even some electrons with no hits in the EMC to pass the selection. The tighter selectors have higher purity (they reject more background), but the looser selectors have higher efficiency (they accept more signal).

Another option for particle identification is to use a likelihood selector. Instead of using simple cuts the likelihood selector strategy is to calculate the probability that the candidate is an electron, pion, kaon, or proton, and compare the different hypotheses via the *likelihood function.* The electron `Likelihood` selector is more efficient than the cut-based selector.

Appendix C

Monte Carlo method to determine the likelihood function

This appendix is a reference for Section 5.15. It describes the Monte Carlo method used to determine the likelihood function for the branching fraction.

To determine the likelihood function for $\mathcal{B}(B^0 \to J/\psi\gamma)$, we use the same method that Chris Hearty used in the analysis for $B^+ \to J/\psi p\overline{\Lambda}$ and $B^0 \to J/\psi p\overline{p}$ [72]: a Bayesian analysis for the branching fraction $\mathcal{B} > 0$ with a uniform prior, using methods derived from the *BABAR* statistics handbook [69]. We use a Monte Carlo method that calculates the likelihood $\mathcal{L}(\mathcal{B}_i; n_0)$ at each branching fraction point $\mathcal{B}_i$. This method is useful because it provides a way to incorporate the systematic errors in n_b and n_s in the calculation.[1]

The Monte Carlo method evaluates 1000 trials for each branching fraction point. In each trial j, a value is selected from Gaussian distributions for the mean number of expected signal and background events, μ_b and μ_s, and we interpret $\lambda_j = (\mu_b + \mu_s)_j$ as the mean of a Poisson distribution for $\hat{n_0}$. The Poisson probability of observing $\hat{n_0} = n_0$ events is then $P(\hat{n_0} = n_0, \lambda_j) = e^{-\lambda_j}\lambda_j^{n_0}/n_0!$. The average of the Poisson probability over the 1000 cases gives the likelihood $\mathcal{L}(\mathcal{B}_i; n_0)$ for the specified branching fraction point $\mathcal{B}_i$.

The likelihood function is used to obtain both a central value and an upper limit on the branching fraction. The central value is the value of $\mathcal{B}$ at which the likelihood function peaks. The 90% CL upper limit is the value of $\mathcal{B}$ with 90% of the area of the likelihood function to its left, and 10% to its right.

In our analysis we observe $n_0 = 0$ events passing the final selection. The likelihood function for $n_0 = 0$ is shown in Fig. 5.11. The corresponding 90% CL upper limit is $\mathcal{B}(B^0 \to J/\psi\gamma) = 1.2 \times 10^{-6}$.

[1]Incorporating systematic errors in a 90% CL upper limit is a nontrivial problem. See, for example, the discussion by Cousins and Highland [74].

	90% CL upper limit ($\times 10^{-6}$)	
n_0	MC method	Eq. C.1
0	1.21	1.21
1	1.67	1.69
2	2.24	2.22
3	2.87	2.85
4	3.54	3.54
5	4.20	4.23
6	4.86	4.86

Table C.1: Cross-check of the Monte Carlo method to determine the likelihood function. For this cross-check we set the systematic errors to zero in the Monte Carlo method. The results should be consistent with those from Eq. C.1, which also ignores systematic error. The MC method clearly passes the cross-check.

As a cross-check on the Monte Carlo method, we set the systematic errors to zero and compare the results to the upper limit $N(\alpha)$ found using a formula that appeared in the particle data book through 1997, reproduced as Eq. 5.10 in *BABAR* statistics handbook [69]:

$$\alpha = 1 - e^{-(n_b+N(\alpha))} \sum_{n=0}^{n=n_0} \frac{(n_b + N(\alpha))^n}{n!} \Big/ e^{-n_b} \sum_{n=0}^{n=n_0} \frac{n_b^n}{n!}, \tag{C.1}$$

For a 90% CL upper limit, $N(0.90)$, we set $\alpha = 0.90$ and solve for $N(0.90)$.

The results for several different values of n_0 are presented in Table 5.8. The MC method clearly passes the cross-check.

Appendix D

Glossary

analysis window (AW) The full region in ΔE and m_{ES} studied in the analysis. That is, the boundaries of the analysis window are just the boundaries of the ΔE vs. m_{ES} histograms presented in this thesis.

Analysis Working Group (AWG) A group of *BABAR* physicists working together on a common set of physics topics.

AWG *see* Analysis Working Group

***BABAR* Analysis Document (BAD)** Frequently updated internal *BABAR* document available to *BABAR* collaborators. BADs are generally an individual's or group's work in physics analysis, detector performance, analysis tool documentation, upcoming conference talks, or drafts of soon to be published papers.

background Anything that is not a signal (*see* signal).

B_{AW} In this analysis, the number of background MC events in the analysis window.

beam-energy substituted mass *see* m_{ES}.

`BGFilter` A background filter algorithm that tags physics events as multihadron events, τ events, two-photon events, and so on. An event must have at least one of these tags set to true in order to be written to the central Event Store database.

blind analysis Analysis in which things like event shapes and the best-fit values of the parameters under study are hidden from the experimenter until the analysis is finalized, to avoid bias.

B_{NH} In this analysis, the number of background MC events in the not hidden region.

branching fraction The branching fraction for the decay $X \to Y$ is number of $X \to Y$ events divided by the number of $X \to$ anything events.

B_{SB} In this analysis, the number of background MC events in the signal box.

B_{SE} In this analysis, the number of background MC events in the signal ellipse. Estimated from B_{AW}.

`BtaMicroCandidate` *BABAR*'s C++ class for what is known or assumed about an (alleged, hence candidate) particle. *BABAR*'s main Micro database is a list of `BtaMicroCandidates`.

bump A local energy peak in an EMC cluster, corresponding to a particle candidate. A cluster consists of one or more bumps. *See* cluster.

`CalorNeutral` A list of neutral particle candidates with EMC showers not matched to any charged track.

charged cluster A cluster that is matched to a charged track, and is therefore a charged particle candidate. *See* cluster.

`ChargedTracks` The basic list of charged particle candidates.

cluster (in the EMC) A bunch of hits in the same general region of the EMC that pass a minimum total energy cut. A cluster corresponds to one or more *bumps*, the basic particle candidates in the EMC.

cluster (in the IFR) A bunch of hits in the same general region of the IFR. The basic IFR particle candidate.

collection Subset of *BABAR* data.

composition Composition involves reconstructing the "decay trees" in the event — determining the mother and daughters of each particle.

continuum Background from the nonresonant production of quark-antiquark ($q\bar{q}$) pairs, via $e^+e^- \to q\bar{q}$. In the analysis in this thesis, $\tau^+\tau^-$ events are for convenience also included in this category.

data acquisition (DAQ) The process of getting information from the subdetectors to the online prompt reconstruction (OPR) system, which performs the central *BABAR* reconstruction.

DAQ *see* data acquisition.

DCH *see* Drift Chamber.

DCT *see* Drift Chamber Trigger.

Detector of Internally Reflected Cerenkov radiation (DIRC) One of *BABAR*'s subdetectors, used for dedicated particle identification, particularly K/π separation.

DIRC *see* DIRC.

Drift Chamber (DCH) One of *BABAR*'s subdetectors. The main tracking chamber.

Drift Chamber Trigger (DCT) One of *BABAR*'s two independent triggers. The other comes from the EMC.

ε_{signal} The signal efficiency, defined as the fraction of signal events passing the final selection (final cuts + signal box).

efficiency The efficiency of a selection is the ratio of the number of events passing the selection to the total number of events to which the selection is applied. In general, you want high efficiency for signal events, and low efficiency for background events. By itself, the word "efficiency" usually refers to efficiency for *signal* events.

ΔE A standard kinematic variable for identification of B mesons from $\Upsilon(4S) \rightarrow B\overline{B}$ decays. In the center-of-mass frame, $\Delta E = E_B^* - E_{beam}^*$, where E_B^* is the energy of the reconstructed B meson. For signal events, the expected value of ΔE is zero.

Electromagnetic Calorimeter (EMC) One of *BABAR*'s subdetectors, used mostly to detect electrons and photons.

Electromagnetic Calorimeter Trigger (EMT) One of *BABAR*'s two independent triggers. The other comes from the DCH.

EMC *see* Electromagnetic Calorimeter.

EMT *see* Electromagnetic Calorimeter Trigger.

equivalent luminosity The *equivalent luminosity* is defined as follows: if a Monte Carlo sample contains (for example) N uds events, then the equivalent luminosity of the MC sample is defined as the luminosity of a real data sample that would contain N uds events.

event assembly In DAQ, the assembly of event fragments into a complete event for L3.

event fragment In DAQ, the output of feature extraction.

event shape Event shape variables are used to separate $B\overline{B}$ events, which tend to be isotropic, from the continuum background of jet-like $q\overline{q}$ events.

Event Store Babar's central database.

exclusive Refers to fully reconstructed decays of a specific type. To be contrasted with inclusive, which refers to classes of fully or partially reconstructed decays. For example, $B^0 \to J/\psi\gamma$ is an exclusive decay, but $B^0 \to J/\psi\ X$ is an inclusive decay, because it includes all possible final states X.

feature extraction In DAQ, the transformation of raw data into useful information such as particle hit time and energy. The output of feature extraction is an event fragment.

FEEs *see* front-end electronics.

final selection In this analysis, an event passes the final selection if it passes the final cuts and falls in the signal box. Only onpeak events passing the final selection are counted in n_0.

final cuts The set of selection cuts presented in Section 5.5. Combined with the signal box, makes up the final selection for this analysis. The final cuts are optimized.

Fox-Wolfram moment *see* R2.

front-end electronics (FEEs) The system electronics responsible for amplifying, digitizing, and selecting time slices from the detector signals.

Global Trigger (GLT) Part of the Level 1 Trigger. The GLT receives information from the EMT and DCT, and uses it to decide whether to pass an event.

GLT *see* Global Trigger.

`GoodPhotonsLoose` A list of photon candidates, consisting of candidates from the `CalorNeutral` list that have an electromagnetic shower shape. `CalorNeutral` candidates are bumps in the EMC that are not matched with any charged track. *See also* `CalorNeutral`.

`GoodTracksLoose` A list of charged tracks that pass a "good but loose" selection for well-reconstructed tracks.

helicity angle In the reaction, $Y \to X \to\ a+b$, the helicity angle of particle a is the angle measured in the rest frame of the decaying parent particle, X, between the direction of the decay daughter a and the direction of the grandparent particle Y [67]. In this analysis, the B^0 helicity angle is used as a selection variable to reject continuum background.

hidden region (H) The region of the ΔE-m_{ES} plane kept hidden from the researcher throughout a blind analysis. Prior to unblinding, the experimenter may look at onpeak data *outside* of this region only. The hidden region contains the signal region.

IFR *see* Instrumented Flux Return.

Instrumented Flux Return (IFR) One of *BABAR*'s subdetectors, used to detect muons and neutral hadrons.

interaction point The point of collision of the e^+e^- beams in the PEP-II collider.

integrated luminosity Short for time-integrated luminosity; the number of events per unit area in a data set taken over a given time period.

`isBCMultiHadron` A tag bit used in most B decay studies at *BABAR* to select $B\overline{B}$ events and reject bhabha, two-photon and beam-gas events. Often called the "B-counting" tag bit, since it is used in the determination of the number of $B\overline{B}$ events in a sample.

`JpsiELoose` A tag bit that selects candidate $J/\psi \to e^+e^-$ electron pairs. The pairs are required to pass *BABAR*'s standardized `Loose` cut-based selector for electrons (see Appendix B) and have an invariant mass close to the J/ψ mass. They are also required to pass bremsstrahlung-recovery algorithms to account for energy loss due to radiated photons.

`JpsiMuLoose` A tag bit that selects candidate $J/\psi \to \mu^+\mu^-$ muon pairs. The pairs are required to pass *BABAR*'s standardized `Loose` cut-based selector for muons (see Appendix B) and have an invariant mass close to the J/ψ mass.

kinematic fitting Kinematic fitting uses kinematic constraints to determine the best values for the momenta, mass and energies of a particle.

kinematic plane The ΔE-m_{ES} plane. Because of the kinematic constraints on $e^+e^- \to B\overline{B}$ events, the signal events for exclusive B decays are concentrated about the point $(m_{\text{ES}}, \Delta E) = (m_B, 0)$. (*See also* ΔE, m_{ES}.)

Level 1 (L1) The first level of *BABAR*'s two-level trigger. The L1 trigger performs a fast, loose selection for events of physics interest, mostly on the basis of track and cluster multiplicities. Events passing Level 1 are sent to Level 3.

Level 3 (L3) The second level of *BABAR*'s two-level trigger. (There is no Level 2.) The L3 trigger performs a slower, tighter selection for events of physics interest than Level 1, based on the complete event information.

LAT *see* lateral energy distribution.

lateral energy distribution (*LAT*) *LAT* is a shower-shape variable, used to distinguish electromagnetic from hadronic showers.

luminosity The number of events per unit area, per unit time. However, many people say "luminosity" when what they actually mean is the time-integrated luminosity. You can tell from the units: luminosity is generally measured in $\text{cm}^{-2}\text{s}^{-1}$, while time-integrated luminosity is measured in fb^{-1}. *See also* time-integrated luminosity.

MC *see* Monte Carlo.

MC-opt In this analysis, the Monte Carlo half-sample used to optimize the signal box.

MC-final In this analysis, the Monte Carlo half-sample used to obtain the final result.

Monte Carlo The usual term for simulated data sets at *BABAR* and other particle physics experiments. (The term "monte carlo" actually refers to the computing method used to generate the simulated data sets, in which random numbers are generated and used to simulate physics events and detector responses.)

Micro database The *BABAR* database containing the basic lists of particle candidates, including the basic `BtaMicroCandidate` list as well as sublists like `ChargedTracks` and `GoodPhotonsLoose`.

m_{ES} A standard kinematic variable for identification of B mesons from $\Upsilon(4S) \to B\overline{B}$ decays. In the center-of-mass frame, $m_{\text{ES}} = \sqrt{E^{*2}_{beam} - \mathbf{p}^*_{\mathbf{B}} \cdot \mathbf{p}^*_{\mathbf{B}}}$, where $\mathbf{p}^*_{\mathbf{B}}$ is the three-momentum of the reconstructed B particle. For signal events, the expected value of m_{ES} is m_B.

n_0 In this analysis, the number of onpeak events passing the final selection (final cuts + signal box).

n_1 In this analysis, the number of onpeak events in the not hidden region that pass the final cuts.

Nano database The database containing *BABAR*'s tag bits. *See* tag bits.

n_b In this analysis, the number of background events in the signal box. Estimated from n_1.

$N_{B\overline{B}}$ The number of $B\overline{B}$ events in a sample.

neutral cluster A cluster that it not matched to a charged track. The basic neutral particle candidate.

90% confidence level (CL) upper limit on the branching fraction. The standard way to present the results for a rare decay search. Defined so that, in 90% of similar experiments with the same result for n_0 as the experiment in question, the interpretation, "$\mathcal{B} <$ the 90% CL upper limit" would be correct.

not hidden region (NH) The region of the analysis window outside of the hidden region. The experimenter may look at onpeak data in this region. (*See also* hidden region.)

n_s In this analysis, the number of signal events in the signal box. Calculated via $n_s = n_0 - n_b$.

ntuple A useful, portable data storage format. The ntuples serve as the user's own private copy of the data.

offpeak data Data collected at about 40 MeV below the $\Upsilon(4S)$ resonance energy. Used to study continuum background (as there is no $B\overline{B}$ production at this energy). *See also* onpeak data.

online prompt reconstruction (OPR) The system that performs the central *BABAR* reconstruction. The output is the lists of particle candidates ("`BtaMicroCandidates`") in the Event Store, *BABAR*'s central database. These are the lists used to produce the ntuples.

onpeak data Data collected at the $\Upsilon(4S)$ resonance energy. Almost all of the data collected at *BABAR* is onpeak data. *See also* offpeak data.

OPR *see* Online Prompt Reconstruction.

package A self-contained piece of *BABAR* software intended to perform a well defined task; e.g. find calorimeter clusters, simulate the drift chamber response.

PAW (Physics Analysis Workstation) A software package used for physics analysis.

particle identification (PID) The determination of the most probable identities for each particle. At *BABAR* many aspects of PID are standardized. For example, *BABAR* uses a set of standard selectors to identify leptons (Appendix B).

Particle Data Group (PDG) The physics group responsible for keeping track of the most up-to-date values for basic particle physics parameters and results, like particle masses, other Standard Model parameters, and branching fractions for particle decays. Their Review of Particle Physics, published every two years, is the universally recognized "final word" on particle physics data. Its reference in this thesis is [10].

PDG *see* Particle Data Group.

PID *see* particle identification.

preselection The initial, loose selection applied during ntuple production to select events of physics interest to the user.

primary vertex The reconstructed beam collision point.

primitive A basic data object constructed from the raw hits in the subdetectors, for use by *BABAR*'s Trigger. The idea is that a primitive corresponds to a (possible) particle. The DCT primitives are long and short tracks in the DCH. The EMT primitives are clusters of crystals with energy above a certain threshold.

R_0 In this analysis, a scale factor used during the optimization of the final cuts.

$R2$ The ratio of the second and zeroth Fox-Wolfram moments; a selection variable used to reject continuum background. Continuum events are jetlike and therefore have high values of $R2$, while B events tend to be spherical and therefore have lower values of $R2$.

readout modules (ROMs) Devices that connect to the front-end electronics and read out event data.

release A set of *BABAR* software packages that work together to produce ntuples using the data from the central *BABAR* database. How this is done is described in more detail in Chapter 4. *See also* package, ntuple.

reprocessing Redoing the reconstruction for a data sample, using the latest software release.

Resistive Plate Chambers (RPCs). The active detectors in *BABAR*'s Instrumented Flux Return. They are tracking devices.

rolling calibrations Refers to *BABAR*'s system of using up-to-date calibration information when reconstructing events.

ROMs *see* readout modules.

ROOT A software package for physics analysis.

RPCs *see* Resistive Plate Chambers.

run (of PEP-II) The interval between two beam injections at PEP-II, corresponding to about 300,000 events. This is different from ...

Run (describing data set) Run 1, Run 2, and Run 3 are names of *BABAR* data sets taken over a specific time period. The data for Runs 1-3 used in this analysis is the full data set for 1999-2003.

selector A set of selection criteria designed to select a specific type of event or particle. For example, *BABAR*'s cut-based muon selector consists of a set of cuts designed to accept muons and reject other particles.

S_{SE} In this analysis, the number of signal MC events in the signal ellipse.

S_{SB} In this analysis, the number of signal MC events in the signal box.

S_{total} In this analysis, the number of signal MC events in the full signal MC sample.

signal box (SB) One of the two signal regions used in this analysis. The signal box is the signal region that is optimized, and it is the one used to obtain the final result.

signal ellipse (SE) One of the two signal regions used in this analysis. The signal ellipse is a temporary signal region used to optimize the final cuts. We then discard the signal ellipse and optimize the signal region to obtain an optimized signal box, which is the signal region used to obtain the final result.

signal region A region in one or more sensitive selection variables in which signal events are concentrated. For exclusive B decays the signal region is typically defined in terms of the two standard kinematic variables ΔE and $m_{\rm ES}$. The signal region is then a small region centered at about ($m_{\rm ES}$, ΔE) = $(m_B, 0)$. In this analysis we use two signal regions: a signal box (*see* signal box) and a signal ellipse (*see* signal ellipse).

signal The physics process of interest to the researcher. In this analysis a signal event is a $B^0 \to J/\psi\,\gamma$ event.

Silicon Vertex Tracker (SVT) One of *BABAR*'s subdetectors, used to measure the distance Δz between the two B decays of a $B\overline{B}$ pair. Also provides tracking measurements for particles that don't reach the DCH.

skim A sample of events selected on the basis of tag bits. (*See also* tag bits.) Similar to a stream, but smaller.

SP4, SP5 Names of *BABAR* Monte Carlo sets. Different-numbered SP sets are produced with different releases (versions) of *BABAR* software.

sphericity angle An event-shape variable. The sphericity angle is the angle between the sphericity axis of the B meson and the sphericity axis of the rest of the event. The distribution of its cosine peaks at ± 1 for continuum events, but tends to be uniform for B decays.

stream A sample of events selected on the basis of tag bits. (*See also* tag bits.) Similar to a skim, but larger.

"Summer 2002 dataset" Data collected between *BABAR*'s startup in 1999 and the summer shutdown in 2002.

SVT *see* Silicon Vertex Tracker.

tag bit Loose selectors that flag events of physics interest. Individual tag bits are defined by different *BABAR* Analysis Working Groups. Examples include the `JpsiELoose`, `JpsiMuLoose`, and `isBCMultihadron` tag bits used in the analysis described in this thesis.

thrust angle An event-shape variable. The thrust angle is the angle between the thrust axis of the B meson and the thrust axis of the rest of the event. The distribution of its cosine peaks at ± 1 for continuum events, but tends to be uniform for B decays.

track The basic charged particle candidate. Tracks are reconstructed by "connecting the dots" between the hits in the many layers of the DCH and the SVT.

time-integrated luminosity *see* luminosity.

vertexing Vertexing takes reconstructed decay trees and determines the most likely position of the vertex, and the most likely momenta of the particles at the vertex.

XTC (eXtended Tagged Container) The output of the L3 trigger is stored in XTC files. These serve as the input to the Online Prompt Reconstruction system.

Bibliography

[1] G. r. Lu, R. m. Wang and Y. d. Yang, arXiv:hep-ph/0308256.

[2] T. E. Browder, eConf **C020805**, L07 (2002).

[3] S. L. Glashow, J. Illiopoulos, and L. Maiani, Phys. Rev. **D 2** 1285 (1970).

[4] S. J. Brodsky and S. Gardner, Phys. Rev. D **65**, 054016 (2002).

[5] D. Griffiths, *Introduction to Elementary Particles*, John Wiley and Sons, Inc. (1987).

M E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*, Addison-Wesley (1995). Reading, USA.

C. P. Burgess. *A primer on the standard model of elementary particle physics,* Course notes, McGill University (2003).

B. R. Martin and G. Shaw, "Particle Physics," Second Edition, John Wiley and Sons, England (2002).

[6] E. Noether, Nachr. Gess. Wiss. Göttingen (1918), 171.

[7] P. W. Higgs, Phys. Lett. **12**, 132 (1964).

P. W. Higgs, Phys. Rev. Lett. **13**, 508 (1964).

[8] A. Salam, Reprinted in *Gauge Theory Of Weak and Electromagnetic Interactions*, C. H. Lai (ed.) 188-198. (Singapore: world Scientific, 1981.)

S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967).

S. L. Glashow, Nucl. Phys. **22**, 579 (1961).

[9] G. Arnison *et al.* [UA1 Collaboration], Phys. Lett. B **122**, 103 (1983).

[10] K. Hagiwara *et al.* [Particle Data Group Collaboration], Phys. Rev. D **66**, 010001 (2002).

[11] P. Ball *et al.*, arXiv:hep-ph/0003238.

[12] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).

[13] J. I. Friedman and H. W. Kendall, Ann. Rev. Nucl. Part. Sci. **22**, 203 (1972).

P. N. Kirk *et al.*, Phys. Rev. D **8**, 63 (1973).

[14] J. D. Bjorken, Phys. Rev. D **1**, 1376 (1970).

J. D. Bjorken and E. A. Paschos, Phys. Rev. **185**, 1975 (1969).

[15] G. Hanson *et al.*, Phys. Rev. Lett. **35**, 1609 (1975).

[16] H. J. Behrend *et al.*, Physics Letters **B183**, 400 (1987).

M. Banner *et al.* [UA2 Collaboration], Phys. Lett. B **122**, 476 (1983).

[17] A. D. Sakharov, Pisma Zh. Eksp. Teor. Fiz. **5**, 32 (1967) [JETP Lett. **5**, 24 (1967 SOPUA,34,392-393.1991 UFNAA,161,61-64.1991)].

[18] Q. R. Ahmad *et al.* [SNO Collaboration], Phys. Rev. Lett. **89**, 011301 (2002) [arXiv:nucl-ex/0204008].

[19] Y. Nir, Nucl. Phys. Proc. Suppl. **117**, 111 (2003) [arXiv:hep-ph/0208080].

J. L. Rosner, AIP Conf. Proc. **689**, 150 (2003) [arXiv:hep-ph/0306284].

J. N. Butler [BTeV collaboration], Nucl. Phys. Proc. Suppl. **120**, 325 (2003).

M. Neubert, arXiv:hep-ph/0212360.

H. Landsman and Y. Rozen, AIP Conf. Proc. **549**, 609 (2002).

V. Canale, *Prepared for 7th International Symposium on Particles, Strings and Cosmology (PASCOS 99), Granlibakken, Tahoe City, California, 10-16 Dec 1999.*

M. Beneke, arXiv:hep-ph/0109243.

M. Feindt Nucl. Phys. Proc. Suppl. **116** (2003) 257.

[20] M. Hazumi, Nucl. Phys. Proc. Suppl. **59**, 61 (1997).

[21] B. Aubert *et al.* [BABAR Collaboration], Phys. Rev. Lett. **87**, 091801 (2001) [arXiv:hep-ex/0107013].

[22] K. Abe *et al.* [Belle Collaboration], Phys. Rev. Lett. **87**, 091802 (2001).

[23] Y. Nir, Nucl. Phys. Proc. Suppl. **117**, 111 (2003) [arXiv:hep-ph/0208080].

[24] CLEO collaboration, R. Ammar *et al*, Phys. Rev. Lett. **71**, 674, 3193.

[25] D. Payne [BABAR Collaboration], SLAC-PUB-10238.

[26] A. M. Eisner, arXiv:hep-ex/0308014.

[27] J. M. Cline, "Physics 198-730B: Quantum Field Theory," Course notes, McGill University (2003).

[28] G. 't Hooft, Nucl. Phys. B **33**, 173 (1971).

G. 't Hooft, Nucl. Phys. B **35**, 167 (1971).

[29] C. P. Burgess, *A Primer on the Standard Model of Elementary Particle Physics,* unpublished course notes, McGill University.

[30] A. J. Buras, Lect. Notes Phys. **558**, 65 (2000) [arXiv:hep-ph/9901409].

[31] G. Buchalla, A. J. Buras and M. E. Lautenbacher, Rev. Mod. Phys. **68**, 1125 (1996).

[32] G. Buchalla, arXiv:hep-ph/0202092. Lectures given at 55th Scottish Universities Summer School in Physics: Heavy Flavor Physics, St. Andrews, Scotland, 7-23 Aug 2001. Published in *St. Andrews 2001, Heavy flavour physics* 57-104.

[33] M. Di Pierro, arXiv:hep-lat/0009001.

T. Onogi, Nucl. Phys. Proc. Suppl. **111**, 75 (2002).

[34] M. A. Luty and J. March-Russell, Nucl. Phys. B **426**, 71 (1994) [arXiv:hep-ph/9310369].

[35] A. Manohar and H. Georgi, Nucl. Phys. B **234**, 189 (1984).

[36] G. Buchalla, Nucl. Instrum. Meth. A **462**, 34 (2001).

S. W. Bosch and G. Buchalla, Nucl. Phys. B **621**, 459 (2002).

M. Neubert, AIP Conf. Proc. **602**, 168 (2001) [AIP Conf. Proc. **618**, 217 (2002)] [arXiv:hep-ph/0110093].

M. Beneke, G. Buchalla, M. Neubert and C. T. Sachrajda, Nucl. Phys. B **591**, 313 (2000) [arXiv:hep-ph/0006124].

[37] P. F. . Harrison and H. R. . Quinn [BABAR Collaboration], SLAC-R-0504

[38] J. D. Bjorken, Nucl. Phys. (Proc. Suppl.) **B11** (1989) 325.

[39] S. W. Bosch, arXiv:hep-ph/0208203.

[40] S. W. Bosch and G. Buchalla, Nucl. Phys. B **621**, 459 (2002) [arXiv:hep-ph/0106081].

[41] B. Aubert *et al.* [BABAR Collaboration], arXiv:hep-ex/0306038.

[42] K. Anikeev *et al.*, arXiv:hep-ph/0201071.

[43] E. von Toerne [CLEO Collaboration], eConf **C020805**, TW03 (2002) [arXiv:hep-ex/0301016].

A. Warburton [CLEO Collaboration], arXiv:hep-ex/9906030.

[44] H. D. Schulz, *Prepared for Second International Symposium on The 4th Family of Quarks and Leptons, Santa Monica, California, 23-25 Feb 1989.*

[45] H. Albrecht *et al.* [ARGUS COLLABORATION Collaboration], Phys. Lett. B **192**, 245 (1987).

[46] S. W. Herb *et al.*, Phys. Rev. Lett. **39**, 252 (1977).

[47] R. Poling, "The CLEO experiment: Recent results and future prospects," Presentation at April 2002 meeting of the American Physical Society, Albuquerque, New Mexico (2002).

[48] P. Oddone, "A Conceptual Design for a High Luminosity Linear Collider B Anti-B Factory," Proceedings, UCLA Workshop page 423 (1987).

[49] P. Eerola, Nucl. Instrum. Meth. A **446**, 384 (2000) [Nucl. Instrum. Meth. A **446**, 407 (2000)] [arXiv:hep-ex/9910067].

[50] P. F. . Harrison and H. R. . Quinn [BABAR Collaboration], SLAC-R-0504.

B. Aubert *et al.* [BABAR Collaboration], arXiv:hep-ex/0012042.

[51] D. Boutigny *et al.* [BABAR Collaboration], SLAC-R-0457.

D. Boutigny *et al.* [BABAR Collaboration], SLAC-0443.

[52] J. Va'vra, "Some comments about possible problems of well-made RPCs and remediation of the old RPCs," Presentation for *BABAR* Muon Detection Workshop, November 14, 2002.

[53] R. Claus, P. Grosso, R. T. Hamilton, M. E. Huffer, C. P. O'Grady, J. J. Russell and I. Scott, SLAC-PUB-8134 *Presented at 11th IEEE NPSS Real Time Conference (Santa Fe 99), Santa Fe, NM, 14-18 Jun 1999.*

[54] A. Ryd *et al.* [BABAR Computing Group Collaboration], eConf **C0303241**, MODT012 (2003) [arXiv:cs.dc/0306069].

T. Glanzman, J. Bartelt, R. F. Cowan, S. Dasu, G. Grosdidier, F. di Lodovico and F. Safai Tehrani (BABAR Computing Group Collaboration), SLAC-PUB-8388.

[55] A. Ryd *et al.* "Trigger and filter documentation for the Run1 data." *BABAR* Analysis Document 194, Apr. 2004.

[56] B. Jacobsen [BABAR Collaboration], Nucl. Instrum. Meth. A **389**, 1 (1997).

[57] D. H. Wright *et al.* [BABAR Computing Group Collaboration], eConf **C0303241**, TUMT0006 (2003) [arXiv:hep-ph/0305240].

[58] G. Hamel de Monchenault *et al.*, "Recommendations from the Blind Analysis Group (BAG)", *BABAR* Analysis Document 91, Nov. 2000.

[59] C. Hearty, "Measurement of the Number of Upsilon(4S) Mesons Produced in Run 1 (B Counting)", *BABAR* Analysis Document 134, Jan. 2001.

[60] G. Fox and S. Wolfram, Nucl. Phys. B149, 413 (1979).

[61] U. Langenegger *et al.*, "Cut-based Electron Identification", *BABAR* Analysis Document 90, May 2001.

[62] L. Lista *et al.*, "Muon Identification in the BaBar Experiment", *BABAR* Analysis Document 60, Jun. 2000.

[63] A. Drescher *et al.*, Nucl. Instr. Methods A237, 464 (1985).

[64] M. Convery *et al.* "Improved measurements of Branching Fractions, Direct CP-Asymmetry and Isospin Violation in $B \to K^* \gamma$", BAD 665, December 2003.

[65] S. Brandt *et al.*, Phys. Lett. 12, 57 (1964).
E. Fahri, Phys. Rev. Lett. 39, 1587 (1977).

[66] J. Bjorken and S. Brodsky, Phys. Rev. D 1, 1416 (1970).

[67] H. Hadavand *et al.*, "Measurement of Charged to Neutral Production Ratio from Upsilon(4s) using Charmonium Final States", *BABAR* Analysis Document 529, Nov. 2003.

[68] W. T. Ford, "Choice of Kinematic Variables in B Meson Reconstruction—Take 3", *BABAR* Analysis Document 53, Jun. 2000.

[69] R. Barlow *et al.*, "Recommended Statistical Procedures for *BABAR*", *BABAR* Analysis Document 318, May 2002.

[70] C. Hearty and M. Brodeur. Personal communication, Jul. 2003.

[71] P. F. Harrison, "Blind analysis", J. Phys. G **28**, 2679-2692 (2002).

[72] C. Hearty, "Supporting document for the search for the decays $B^+ \to J/\psi p\overline{\Lambda}$ and $B^0 \to J/\psi p\overline{p}$", *BABAR* Analysis Document 556, March 2003.

[73] R. Sinkus and T. Voss, Nucl. Instrum. Meth. A **391**, 360 (1997).

[74] R. D. Cousins and V. L. Highland, Nucl. Instrum. Meth. A **320**, 331 (1992).