Princeton/HEP/92-04
September 4, 1992

Maximum-Likelihood Analysis of CP-Violating Asymmetries

K.T. McDonald
Joseph Henry Laboratories, Princeton University, Princeton, NJ 08544

Abstract

We deduce estimates of the statistical precision of analyses of CP-violating asymmetries in the B^0-$\overline{B}^0$ system via the maximum-likelihood method. In the case of B^0 decays to a CP eigenstate f the decay-time distributions have the form

$$N_{\pm}(t) = \frac{N}{2} e^{-t}(1 \pm A \sin xt),$$

where N is the total number of decays to state f, A is the CP-violating parameter which is a simple function of parameters of the C-K-M matrix, $x = \Delta M/\Gamma$ is the mixing parameter, and $+(-)$ labels decays in which the B was born as a $B^0(\overline{B}^0)$. The estimated error on the measurement of A can be written in terms of 'dilution factors' as

$$\sigma_A = \frac{1}{D\sqrt{N}},$$

where

$$D = \frac{x}{1+x^2}$$

for a time-integrated analysis;

$$D = D_t \sqrt{\frac{2x^2}{1+4x^2}}$$

for a time-dependent analysis;

$$D = D_t \frac{x}{1+x^2} \sqrt{\frac{1+2x^4}{1+4x^2}}$$

for an analysis based only on the shape of the decay distribution; and

$$D_t = e^{-x^2\sigma_t^2/2}$$

represents the effect of time resolution σ_t. Results are also presented for simultaneous analysis of the CP-violating parameter A and the mixing parameter x, and for analysis of the mixing parameter via decays to non-CP eigenstates. We end with an analysis of asymmetries appropriate for study of CP violation at an e^+e^- collider.

Contents

1 Introduction

Following discussion at the mini-workshop on B physics at the SSC Laboratory, June 29-30, 1992, Milind Purohit pointed out that an optimum analysis of CP-violating asymmetries would be based on the maximum-likelihood method. This should yield greater statistical precision than the methods presented in refs. [1] and [2]. Here we deduce the size of the error on various asymmetries via the likelihood technique.

The principal example we consider is the case of neutral-B-meson decay to a CP eigenstate f. Here we suppose that we have a sample of N decays of either a B^0 or $\overline{B}^0$ to state f in an experiment where there are equal numbers of B and $\overline{B}$'s produced. Then following eq. (26) of ref. [2] the time distribution of the observed decays can be written

$$N_\pm(t) = \frac{N}{2} e^{-t}(1 \pm A \sin xt), \tag{1}$$

where throughout this note time is measured in units of the B^0 lifetime, and A is a simple function of the parameters of the C-K-M matrix (in the Standard Model). The subscript $+$ means that the decay occurred for a B that was a B^0 at $t = 0$, while subscript $-$ means the B was a $\overline{B}^0$ at $t = 0$. These initial conditions must be determined by observation of the second B in the event. For hadroproduction of B's the effect of tagging the second B factorizes from the analysis of the first and we do not consider the second B in this note (except in sec. 11 on e^+e^- colliders).

When A is nonzero there is CP violation, which manifests itself both in the difference between the shape of distributions $N_+(t)$ and $N_-(t)$, and in the difference between the total number of decays of each type:

$$N_\pm \equiv \int_0^\infty N_\pm(t)dt = \frac{N}{2}\left(1 \pm A\frac{x}{1+x^2}\right). \tag{2}$$

Eventually we will wish to consider the effect of the experimental resolution in time t on the analysis. It is felicitous that this has only a minor effect on the formalism, so we prepare the general case now. We designate σ_t as the r.m.s. time resolution, which means that the observed decays distributions can be obtained by convolution [3]:

$$\begin{aligned} N_\pm(t) &= \frac{N}{2}\int_{-\infty}^{\infty} \frac{e^{-(t-t')^2/2\sigma_t^2}}{\sqrt{2\pi}\sigma_t} dt' e^{-t'}(1 \pm A \sin xt') \\ &= \frac{N}{2} e^{\sigma_t^2/2} e^{-t}\left(1 \pm A e^{-x^2\sigma_t^2/2} \sin x(t-\sigma_t^2)\right) \\ &\approx \frac{N}{2} e^{-t}(1 \pm A e^{-x^2\sigma_t^2/2} \sin xt), \end{aligned} \tag{3}$$

using integral 3.896.4 of ref. [4], and where the approximation holds well when $\sigma_t \ll 1$ (*i.e.*, when the time resolution is much better than a lifetime), as is expected to be the case when a silicon vertex detector is used.

Hence an analysis of distributions of the form

$$N_\pm(t) = \frac{N}{2} e^{-t}(1 \pm a \sin xt) \tag{4}$$

includes the effect of time resolution if we write

$$a = AD_t, \qquad \text{with} \qquad D_t \equiv e^{-x^2\sigma_t^2/2} \tag{5}$$

where D_t is the 'dilution factor' associated with finite time resolution.

We anticipate that an analysis of B^0-$\overline{B}^0$ mixing will be similar to that of CP violation. In the case of mixing, we take $N_+(t)$ to be the distribution of decays in which the B was born as a B^0 and decayed as a B^0 (or was born as a $\overline{B}^0$ and decayed as a $\overline{B}^0$), while $N_-(t)$ is the distribution of decays in which the B was born as a B^0 and decayed as a $\overline{B}^0$ (or was born as a $\overline{B}^0$ and decayed as a B^0. For this we must be able to tell whether the particle was a B^0 or $\overline{B}^0$ at the time of decay, and so we cannot use the CP eigenstates discussed above – unless there is CP violation, for which case the statistical precision will typically be greatly reduced, as discussed later.

The mixing time distributions have the well-known form

$$N_\pm(t) = \frac{N}{2}e^{-t}(1 \pm \cos xt), \tag{6}$$

leading to integrated numbers of events

$$N_\pm = \frac{N}{2}\int_0^\infty e^{-t}(1 \pm \cos xt) = \frac{N}{2}\left(1 \pm \frac{1}{1+x^2}\right). \tag{7}$$

As before, the effect of a time resolution σ_t is readily included via convolution with a gaussian:

$$\begin{aligned} N_\pm(t) &= \frac{N}{2}\int_{-\infty}^{\infty} \frac{e^{-(t-t')^2/2\sigma_t^2}}{\sqrt{2\pi}\sigma_t} dt' e^{-t'}(1 \pm \cos xt') \\ &= \frac{N}{2}e^{\sigma_t^2/2}e^{-t}\left(1 \pm e^{-x^2\sigma_t^2/2}\cos x(t-\sigma_t^2)\right) \\ &\approx \frac{N}{2}e^{-t}(1 \pm e^{-x^2\sigma_t^2/2}\cos xt), \end{aligned} \tag{8}$$

Hence a general mixing analysis will deal with distributions of form

$$N_\pm(t) = \frac{N}{2}e^{-t}(1 \pm a\cos xt), \tag{9}$$

which are closely related to those for CP violation given in eq. (4).

2 The Maximum-Likelihood Method

We recall the technique of data analysis via maximizing the likelihood by the example of N data points x_i sampled from a gaussian distribution of mean a and variance σ:

$$P(x,a) = \frac{e^{-(x-a)^2/2\sigma^2}}{\sqrt{2\pi}\sigma}. \tag{10}$$

The probability (or likelihood) of observing the data set $\{x_i\}$ is then

$$\mathcal{L}(a) = \prod_{i=1}^{N} P(x_i, a). \tag{11}$$

The idea of the maximum-likelihood method is that $\mathcal{L}$ is approximately gaussian in the parameter a (whether or not $P(x, a)$ is a gaussian function of x), and hence the value of a that maximizes $\mathcal{L}(a)$ is the best estimate of a. Further, an excellent estimate of the error on the measurement of a follows from the second derivative of $\ln \mathcal{L}$:

$$\mathcal{L} = \prod_{i=1}^{N} e^{-(x_i-a)^2/2\sigma^2}, \tag{12}$$

$$\ln \mathcal{L} = -\frac{1}{2}\sum_{i=1}^{N} \frac{(x_i - a)^2}{\sigma^2}, \tag{13}$$

$$\frac{d \ln \mathcal{L}}{da} = \frac{1}{2}\sum_i \frac{x_i - a}{\sigma^2}, \tag{14}$$

$$\frac{d^2 \ln \mathcal{L}}{da^2} = -\sum_i \frac{1}{\sigma^2} = -\frac{N}{\sigma^2}. \tag{15}$$

The maximum of $\mathcal{L}$ and for $\ln \mathcal{L}$ occur at the same value of a, namely $a = \sum_i x_i/N$ as expected. We identify

$$-\frac{d^2 \ln \mathcal{L}}{da^2} \equiv \frac{1}{\sigma_a^2} \tag{16}$$

to find that $\sigma_a = \sigma/\sqrt{N}$ as expected.

The method is readily extended to distributions that depend on multiple parameters. We will later consider two parameters, say a and b, for which the likelihood function $\mathcal{L}(a, b)$ formed from products of the probabilities $P(x_i, a, b)$ is expected to be gaussian in a and b:

$$\mathcal{L}(a, b) \propto \exp\left\{-\frac{1}{2}\left(\frac{(a - a_{\rm true})^2}{\sigma_a^2} + \frac{2(a - a_{\rm true})(b - b_{\rm true})}{\sigma_{ab}^2} + \frac{(b - b_{\rm true})^2}{2\sigma_b^2}\right)\right\}. \tag{17}$$

Hence our estimates on the errors of the fitted values of a and b will be

$$\frac{1}{\sigma_a^2} = -\frac{\partial^2 \ln \mathcal{L}}{\partial a^2}, \qquad \frac{1}{\sigma_{ab}^2} = -\frac{\partial^2 \ln \mathcal{L}}{\partial a \partial b}, \qquad \frac{1}{\sigma_b^2} = -\frac{\partial^2 \ln \mathcal{L}}{\partial b^2}. \tag{18}$$

3 Analysis of a Simple Asymmetry

As a preliminary example of the maximum-likelihood method, we consider the case when the data can take on only two values, labelled + and −, with probability

$$P_\pm = \frac{1 \pm a}{2}, \tag{19}$$

where a is the asymmetry parameter. For an experiment in which N_+ and N_- events are observed, we form the likelihood function

$$\mathcal{L} = \left(\frac{1 + a}{2}\right)^{N_+} \left(\frac{1 - a}{2}\right)^{N_-}. \tag{20}$$

The needed derivatives of $\ln\mathcal{L}$ are

$$\begin{aligned}
\ln\mathcal{L} &= N_+\ln(1+a) + N_-\ln(1-a) + \text{constant}, &(21)\\
\frac{d\ln\mathcal{L}}{da} &= \frac{N_+}{1+a} - \frac{N_-}{1-a}, &(22)\\
\frac{d^2\ln\mathcal{L}}{da^2} &= -\frac{N_+}{(1+a)^2} - \frac{N_-}{(1-a)^2}. &(23)
\end{aligned}$$

On setting the first derivative to zero, we find the usual expression for the asymmetry:

$$a = \frac{N_+ - N_-}{N_+ + N_-}. \tag{24}$$

From this we express N_+ and N_- in terms of a and $N = N_+ + N_-$ to evaluate the error on the estimate of a as

$$\sigma_a = \sqrt{\frac{1-a^2}{N}}, \tag{25}$$

using eq. (16). This agrees with the usual analysis based on the binomial distribution.

4 Time-Integrated Analysis of CP Violation

After these lengthy preliminaries, we turn to the analysis of CP-violating asymmetries, beginning with the case where the data in integrated over time to yield the total numbers of events given in eq. (2). In this case we study a simple asymmetry related by

$$a = A\frac{x}{1+x^2} = AD_{t-\text{int}}, \tag{26}$$

where A is the CP-violating factor introduced in eq. (1), and we define

$$D_{t-\text{int}} \equiv \frac{x}{1+x^2} \tag{27}$$

as the dilution factor due to time integration.

From eq. (25) we estimate the error on the measurement of A as

$$\sigma_A = \frac{\sigma_a}{D_{t-\text{int}}} = \frac{1}{D_{t-\text{int}}}\sqrt{\frac{1-A^2D^2_{t-\text{int}}}{N}} \approx \frac{1}{D_{t-\text{int}}\sqrt{N}} = \frac{1+x^2}{x\sqrt{N}}, \tag{28}$$

where the approximation holds for small values of $AD_{t-\text{int}}$.

The error on A is large for both large and small values of the mixing parameter x. The minimum error as a function of x occurs if $x = 1$, for which $\sigma_A = 2/\sqrt{N}$. As $x \approx 1/\sqrt{2}$ for the B^0_d meson, a time-integrated analysis is rather effective in this case.

5 Time-Dependent Analysis of CP Violation

We now determine what additional statistical power can be expected if we perform an analysis of the time-dependent CP-violating decay distributions given in eq. (1). The likelihood function is then

$$\mathcal{L} = \prod_i e^{-t_i}(1 + A\sin xt_i) \prod_j e^{-t_j}(1 - A\sin xt_j), \tag{29}$$

where subscript i labels events in which the B was born as a B^0, and j labels events in the the B was born as a $\overline{B}^0$. This form of the likelihood function is normalized to include information both on the shape as well as the integral of the decay distributions.

According to eq. (16) we estimate the error on the measurement of A as

$$\frac{1}{\sigma_A^2} = -\frac{d^2 \ln \mathcal{L}}{dA^2} = \sum_i \frac{\sin^2 xt_i}{(1 + A\sin xt_i)^2} + \sum_j \frac{\sin^2 xt_j}{(1 - A\sin xt_j)^2}. \tag{30}$$

We estimate the sums by integrals according to

$$\sum_{i(j)} f(t) \approx \frac{N}{2} \int_0^\infty dt e^{-t}(1 \pm A\sin xt) f(t), \tag{31}$$

which leads to

$$\frac{1}{\sigma_A^2} = N \int_0^\infty \frac{dt e^{-t}\sin^2 xt}{1 - A^2\sin^2 xt} \approx N \int_0^\infty dt e^{-t}\sin^2 xt = \frac{2x^2 N}{1 + 4x^2}, \tag{32}$$

where we ignore the time-varying term in the denominator for small A, and we have used integral 3.895.1 of ref. [4].

The full integral can be expressed as an infinite series on expanding the denominator in a Taylor series. Keeping the first correction we find that

$$\frac{1}{\sigma_A^2} \approx N \int_0^\infty dt e^{-t}\sin^2 xt(1 + A^2\sin^2 xt) = \frac{2x^2 N}{1 + 4x^2}\left(1 + \frac{12A^2x^2}{1 + 16x^2}\right). \tag{33}$$

Thus even for $A = \frac{1}{3}$ the correction is at most 8% for any value of x.

We summarize the result (32) by writing

$$\sigma_A \approx \frac{1}{D_{t-\rm dep}\sqrt{N}} \qquad \text{with} \qquad D_{t-\rm dep} \equiv \sqrt{\frac{2x^2}{1 + 4x^2}}. \tag{34}$$

The time-dependent dilution factor $D_{t-\rm dep}$ is larger than the time integrated factor (from eq. (27)) for any value of x, and consequently the time-dependent analysis is always more powerful statistically, as is to be expected.

In particular, the time-dependent analysis remains very powerful for large x, where a time-integrated analysis yields no information. Indeed, for the time-dependent analysis,

$$\sigma_A \approx \sqrt{\frac{2}{N}} \qquad \text{for large } x. \tag{35}$$

This result also compares favorably with that reported in refs. [1] and [2], where it was argued that the effective dilution factor at large x is the average of $\sin xt$ over a half-cycle, namely $2/\pi$, leading to $\sigma_A \approx \pi/2\sqrt{N}$.

6 Analysis of the Shape of the Time Distribution

M. Purohit has noted [5] that one could also perform an analysis of CP violation based only on the shape of the decay distributions, ignoring the CP-violating asymmetry in the integrated decay rates. Such an analysis would be the only one possible if the experiment consisted of B's born only as B^0 (or only as $\overline{B}^0$).

The analysis is based on a likelihood function in which the decay distribution is normalized to one (using the notation of eq. (29) and assuming equal numbers of B^0 and $\overline{B}^0$ initially):

$$\mathcal{L} = \prod_i \frac{e^{-t_i}(1 + A\sin xt_i)}{1 + A\frac{x}{1+x^2}} \prod_j \frac{e^{-t_j}(1 - A\sin xt_j)}{1 - A\frac{x}{1+x^2}}. \tag{36}$$

Approximating the sums in the second derivative of $\ln\mathcal{L}$ by the appropriate integrals, and again neglecting a factor in A^2 in the denominator, we have

$$\sigma_A \approx \frac{1}{D_{\text{shape}}\sqrt{N}} \qquad \text{with} \qquad D_{\text{shape}} \equiv \frac{x}{1+x^2}\sqrt{\frac{1+2x^4}{1+4x^2}}. \tag{37}$$

This result is, of course, poorer than the full time-dependent analysis (eq. (34)), but approaches the same accuracy for large x where only the shape matters. The shape analysis is less powerful than the time-integrated analysis (eq. (28)) for $x < \sqrt{2}$, which includes the case of B_d^0 mesons.

The full time-dependent analysis of the previous section can be considered as the proper combination of the time-integrated and the shape analyses. We readily verify the validity of this by noting that

$$\frac{1}{\sigma^2(\text{time-dependent})} = \frac{1}{\sigma^2(\text{time-integrated})} + \frac{1}{\sigma^2(\text{shape})}, \tag{38}$$

on comparing eqs. (28), (34), and (37).

As a numerical example, we consider the case of $x = 1/\sqrt{2}$, as holds approximately for B_d^0 mesons. We then have

$$\sigma(\text{time-dependent}) = \sqrt{\frac{3}{N}} = \frac{1.73}{\sqrt{N}}, \qquad \sigma(\text{time-integrated}) = \frac{3}{\sqrt{2N}} = \frac{2.12}{\sqrt{N}},$$
$$\sigma(\text{shape}) = \frac{3}{\sqrt{N}}. \tag{39}$$

It is remarkable that the time-dependent analysis is only 20% better than the time-integrated analysis, while the former requires a costly silicon vertex detector.

7 The Effect of Time Resolution

In sec. 1 we noted that the effect on the analysis of a time resolution σ_t is well approximated by a dilution factor $D_t = e^{-x^2\sigma_t^2/2}$ multiplying the CP-violating parameter A (see eqs. 3)-(5)).

Thus the full-time-dependent analysis including time resolution will yield

$$\sigma_A \approx \frac{1}{D_{t-\mathrm{dep}} D_t \sqrt{N}} = \sqrt{\frac{1+4x^2}{2x^2}} \frac{e^{x^2\sigma_t^2/2}}{\sqrt{N}}. \tag{40}$$

The effect of time resolution is only noticeable for $xt \gtrsim 1$, *i.e.*, for large x, in which case

$$\sigma_A \approx e^{x^2\sigma_t^2/2} \sqrt{\frac{2}{N}} \qquad \text{(large } x\text{)}. \tag{41}$$

8 The Effect of a Cut at Short Times

M. Purohit has also pointed out [5] that in a realistic analysis based on decay times reconstructed with a silicon vertex detector there will be a loss of events for times t less than some small time t_0, when the secondary vertex cannot be distinguished from the primary. In this case the full time-dependent analysis proceeds as in sec. 5, except that when estimating sums by integrals we now use

$$\sum_{i(j)} f(t) \approx \frac{N}{2} \int_{t_0}^{\infty} dt e^{-t} (1 \pm A \sin xt) f(t), \tag{42}$$

which leads to

$$\frac{1}{\sigma_A^2} \approx N \int_{t_0}^{\infty} dt e^{-t} \sin^2 xt = \frac{e^{-t_0} N}{2} \left(1 + \frac{2x \sin 2xt_0 - \cos 2xt_0}{1+4x^2}\right), \tag{43}$$

using integral 2.663.1 of ref. [4]. As $c\tau \approx 320$ μm is the B decay length, and the typical resolution of silicon vertex detector is less than 20 μm, the condition $t_0 \ll 1$ lifetime will likely be satisfied. Then for small x we can write

$$\frac{1}{\sigma_A^2} \approx \frac{2x^2 N}{1+4x^2} \left(1 - \frac{t_0^3}{6}(1+4x^2)\right), \qquad (x \ll 1), \tag{44}$$

which implies a very small correction. For large x we have

$$\frac{1}{\sigma_A^2} \approx \frac{N}{2}(1 - t_0), \qquad (x \gg 1), \tag{45}$$

which indicates that the correction for the cut at small times is small but perhaps notable in this case.

9 Simultaneous Analysis of Parameters A and x

In all of the proceeding we have tacitly assumed that the value of the mixing parameter x is known from other studies. This might not be so for B_s^0 mesons.

Here we consider the time-dependent likelihood function (29) to estimate the errors on measurement of both A and x according to the procedure of eq. (18). The effect of time resolution is included as the dilution factor D_t to parameter A. With the same approximation of sums as integrals we find

$$\frac{1}{\sigma_A^2} \approx \frac{2x^2 e^{-x^2\sigma_t^2} N}{1+4x^2}, \qquad \frac{1}{\sigma_{Ax}^2} \approx A e^{-x^2\sigma_t^2} N \int_0^\infty dt e^{-t} t \sin xt \cos xt = \frac{A e^{-x^2\sigma_t^2} N \sin(2\tan^{-1} 2x)}{2(1+4x^2)},$$
$$\frac{1}{\sigma_x^2} \approx A^2 e^{-x^2\sigma_t^2} \int_0^\infty dt e^{-t} t^2 \cos^2 xt = A^2 e^{-x^2\sigma_t^2} N \left(1 + \frac{\cos(3\tan^{-1} 2x)}{(1+4x^2)^{3/2}}\right), \quad (46)$$

using integrals 3.944.5 and 3.944.6 of ref. [4].

These complicated results are perhaps best illustrated by considering the limits of small and large x. For small x:

$$\frac{1}{\sigma_A^2} \approx 2x^2 N, \qquad \frac{1}{\sigma_{Ax}^2} \approx 2AxN, \qquad \frac{1}{\sigma_x^2} \approx 2A^2 N, \qquad (\text{small } x). \quad (47)$$

The result for σ_x suggests that surprisingly good resolution in x can be obtained even when the mixing oscillations are almost indiscernible. However, one must note the correlation of the errors in x and A. More properly we should report the error in x as the the extreme value of the 1-σ error ellipse (from eq. (17)):

$$\frac{A^2}{\sigma_A^2} + \frac{2Ax}{\sigma_{Ax}^2} + \frac{x^2}{\sigma_x^2} = 1. \quad (48)$$

On requiring $dx/dA = 0$ in this we find that the extreme value satisfies $x = A\sigma_{Ax}^2/\sigma_A^2$. Inserting this into eq. (46) we must evaluate to sixth order to find

$$\sigma_x(\text{effective}) \approx \frac{1}{Ax^2\sqrt{112N}}. \quad (49)$$

So indeed for small x it is very difficult to determine x from studies of CP violation.

For large x eq. (46) becomes

$$\sigma_A \approx e^{x^2\sigma_t^2/2}\sqrt{\frac{2}{N}}, \qquad \sigma_{Ax} \to \infty, \qquad \sigma_x \approx \frac{e^{x^2\sigma_t^2/2}}{A\sqrt{N}}, \qquad (\text{large } x). \quad (50)$$

As $x\sigma_t \to 1$, which may well hold for the B_s^0 meson, the resolution in both A and x deteriorate rapidly. It will be advantageous to have determined x in a separate measurement.

10 Analysis of B^0-$\overline{B}^0$ Mixing

As it will be advantageous to deduce the mixing parameter x for the B_s^0 meson from other than CP-violation data, we consider now the statistical power of such an analysis. This is based on eq. (6), or eq. (9) when time resolution is included. We form the likelihood function

$$\mathcal{L} = \prod_i e^{-t_i}(1 + a\cos xt_i) \prod_j e^{-t_j}(1 - a\cos xt_j), \quad (51)$$

where $a = e^{-x^2\sigma_t^2/2}$ is the effect of time resolution, and subscript $i(j)$ refers to events where the B is born as a B^0 and decays as a $B^0(\overline{B}^0)$ (or where the B is born as a $\overline{B}^0$ and decays as a $\overline{B}^0(B^0)$).

Again, approximating sums as integrals in the second derivative of $\ln \mathcal{L}$ we find that

$$\frac{1}{\sigma_x^2} = N \int_0^\infty dt e^{-t} t^2 = 2N, \tag{52}$$

so that

$$\sigma_x = \frac{1}{\sqrt{2N}} \tag{53}$$

if $xt \ll 1$ so that time resolution may be ignored.

When time resolution is significant we find

$$\frac{1}{\sigma_x^2} = N e^{-x^2\sigma_t^2} \int_0^\infty \frac{dt e^{-t} t^2 \sin^2 xt}{1 - e^{-x^2\sigma_t^2} \cos^2 xt}. \tag{54}$$

This integral can be bounded by either considering the denominator to be 1 or $\sin^2 xt$, leading to

$$e^{-x^2\sigma_t^2} N \left(1 - \frac{\cos(3\tan^{-1} 3x)}{(1+4x^2)^{3/2}}\right) \approx e^{-x^2\sigma_t^2} N \leq \frac{1}{\sigma_x^2} \leq 2e^{-x^2\sigma_t^2} N, \tag{55}$$

using integral 3.944.6 of ref. [4], and the approximation holds for large x. This implies

$$\frac{e^{x^2\sigma_t^2/2}}{\sqrt{2N}} \leq \sigma_x \leq \frac{e^{x^2\sigma_t^2/2}}{\sqrt{N}} \tag{56}$$

holds for any x and σ_t. Furthermore, when one is not restricted to the use of decay modes leading to CP eigenstates the total number of events N may be much larger than in (50).

The effect of a cut at a small time t_0 is readily considered, as in sec. 8. For small x the correction is fifth order in t_0, while for large x it is third order. That is, the correction is unimportant.

11 Analysis of CP Violation at an e^+e^- Collider

As is now well known, when B's are produced are part of a B^0-$\overline{B}^0$ pair with definite charge conjugation, the analysis of CP violation is more intricate. In particular, if the B's are produced in a C(odd) state, as from $\Upsilon(4S)$ decay, then a time-integrated asymmetry vanishes. However, good statistical power can be recovered by an analysis of time-ordered decay distributions.

Both B's of a produced B-$\overline{B}$ pair must be observed in the CP analysis. We label B_1 as the (neutral) B that decays to the CP eigenstate f, and B_2 as the (charged or neutral) B that decays to a state $g \neq \bar{g}$ that permits us to determine whether B_2 was a particle or antiparticle at the moment of its decay. We can accumulate four time distributions, where one B decays at time t_a and the other at time t_b with $t_a < t_b$:

$$\begin{aligned} I: &\quad \Gamma_{B_1\to f}(t_b)\Gamma_{B_2\to g}(t_a), \\ II: &\quad \Gamma_{B_1\to f}(t_a)\Gamma_{B_2\to g}(t_b), \\ III: &\quad \Gamma_{B_1\to f}(t_b)\Gamma_{\bar{B}_2\to \bar{g}}(t_a), \\ IV: &\quad \Gamma_{B_1\to f}(t_a)\Gamma_{\bar{B}_2\to \bar{g}}(t_b). \end{aligned} \tag{57}$$

The four distributions can be combined to form asymmetries in various ways: most relevant for C(odd) states is

$$A_1(t_a, t_b) \equiv \frac{II + III - I - IV}{I + II + III + IV}, \tag{58}$$

For C(even) states we should consider

$$A_2(t_a, t_b) \equiv \frac{III + IV - I - II}{I + II + III + IV}. \tag{59}$$

The third variation of such asymmetries turns out to vanish and is not considered further:

$$A_3(t_a, t_b) \equiv \frac{I + III - II - IV}{I + II + III + IV}. \tag{60}$$

For the case that mesons 1 and 2 are of the same type the four time distributions take the form

$$\begin{aligned} \Gamma_I(t_a, t_b) &\propto e^{-(t_a+t_b)}[1 \pm A \sin x(t_a \pm t_b)], \\ \Gamma_{II}(t_a, t_b) &\propto e^{-(t_a+t_b)}[1 + A \sin x(t_a \pm t_b)], \\ \Gamma_{III}(t_a, t_b) &\propto e^{-(t_a+t_b)}[1 \mp A \sin x(t_a \pm t_b)], \\ \Gamma_{IV}(t_a, t_b) &\propto e^{-(t_a+t_b)}[1 - A \sin x(t_a \pm t_b)], \end{aligned} \tag{61}$$

where A the CP-violating factor introduced in eq. (1), and the lower sign in the distributions holds for C(odd) states $|B_1\rangle|\overline{B}_2\rangle - |\overline{B}_1\rangle|B_2\rangle$.

Inserting the time distributions into the forms for the asymmetries we have

$$\begin{aligned} A_1 &= \begin{cases} A \sin x(t_a - t_b), & C(\text{odd}), \\ 0, & C(\text{even}), \end{cases} \\ A_2 &= \begin{cases} 0, & C(\text{odd}), \\ A \sin x(t_a + t_b), & C(\text{even}), \end{cases} \\ A_3 &= 0. \end{aligned} \tag{62}$$

Clearly the asymmetry A_1 will be useful at an e^+e^- collider where only C(odd) states are produced.

We first present a time-integrated analysis of these asymmetries, as discussed in ref. [1]. Because of the time ordering in the definition of the distributions I-IV, the form of the integrals is

$$\int_0^\infty dt_a \int_{t_a}^\infty dt_b \Gamma_I(t_a, t_b) = \frac{1}{2}\left(1 \pm A \frac{x}{1+x^2}\right), \qquad C(\text{odd}),$$
$$= \frac{1}{2}\left(1 \pm A \frac{2x}{(1+x^2)^2}\right), \qquad C(\text{even}). \tag{63}$$

Thus we can write

$$A_1 = D_{1,t-\text{int}} A \qquad \text{with} \qquad D_{1,t-\text{int}} = \frac{x}{1+x^2} \qquad C(\text{odd}), \tag{64}$$

$$A_2 = D_{2,t-\text{int}} A \qquad \text{with} \qquad D_{2,t-\text{int}} = \frac{2x}{(1+x^2)^2} \qquad C(\text{even}). \tag{65}$$

The above results can be improved upon with a maximum-likelihood analysis. We label events in distributions I, II, III, and IV by indices i, j, k, and l, respectively, to form the likelihood function

$$\mathcal{L} = \prod_i \Gamma_I(t_{ai}, t_{bi}) \prod_j \Gamma_{II}(t_{aj}, t_{bj}) \prod_k \Gamma_{III}(t_{ak}, t_{bk}) \prod_l \Gamma_{IV}(t_{al}, t_{bl}). \tag{66}$$

We again approximate the sums in the second derivative of $\ln \mathcal{L}$ via sums as

$$\sum_i f(t_{ai}, t_{bi}) \to \frac{N}{2} \int_0^\infty dt_a \int_{t_a}^\infty dt_b \Gamma_I(t_a, t_b) f(t_a, t_b), \qquad etc., \tag{67}$$

for a total sample of N events, noting eq. (63). Ignoring the term in the denominator in A^2 the integrals are similar to those encountered previously:

$$\frac{1}{\sigma_A^2} \approx 2N \int_0^\infty dt_a e^{-t_a} \int_{t_a}^\infty dt_b e^{-t_b} \sin^2 x(t_a \pm t_b) = 2N \int_0^\infty dt_a e^{-2t_a} \int_0^\infty ds e^{-s} \sin^2 x(s + t_a \pm t_a), \tag{68}$$

where $s = t_a - t_b$. We characterize the results of the time-dependent analysis via the dilution factors

$$D_{1,t-\text{dep}} = \sqrt{\frac{2x^2}{1+4x^2}}, \quad C(\text{odd}), \qquad \text{and} \qquad D_{2,t-\text{dep}} = \frac{\sqrt{8x^4+6x^2}}{1+4x^2}, \quad C(\text{even}). \tag{69}$$

The dilution factors from the time-dependent maximum-likelihood analysis are larger than those for the time-integrated analysis, and are the best possible. For large x the time-dependent analysis is particularly advantageous.

As was mentioned in sec. 5, the dilution factors for the case of large asymmetry A can be expressed as infinite series, the first terms of which are given in eq. (69). These series have been given in notes by Frank Porter [6].

The effect of time resolution σ_t on the analysis can be calculated as in eq. (3), and can be characterized (for small A) by the dilution factor

$$D_t = e^{-x^2\sigma_t^2} \qquad \text{in the relation} \qquad \sigma_{A_{1,2}} = \frac{1}{D_{1,2} D_t \sqrt{N}}. \tag{70}$$

As both B's must be time-resolved in this analysis the dilution factor D_t is the square of that encountered in the single-B analysis. Viewed another way, since two times are measured for each event at an e^+e^- collider the error on the sum or difference is $\sqrt{2}\sigma_t$. Using this in eq. (5) we also arrive at eq. (70).

12 References

[1] *The Physics Program of a High-Luminosity Asymmetric B Factory at SLAC*, SLAC-353 (Oct. 1989).

[2] BCD Collaboration, *Expression of Interest for a Bottom Collider Detector at the SSC* (EOI0008), submitted to the SSC Laboratory (May 25, 1990).

[3] C. Biino and S. Palestini, *Sensitivity to Mixing and CP Violation in the* B^0-$\overline{B}^0$ *Systems by Decay Distribution Study*, INFN/AE-90-03 (17 April, 1990).

[4] I.S. Gradshteyn and I.M. Ryzhik, *Table of Integrals, Series, and Products*, Academic Press (San Diego, 1980).

[5] M.V. Purohit, *Measuring CP Asymmetries in B Decays*, Princeton U. preprint (July, 1992).

[6] F. Porter, *Dilution Factors in CP Violation Measurement*, BaBar Notes #24 (Jan. 25, 1990) and #26B (Feb. 9, 1990). The author wishes to thank Frank for a critical reading of the present note.