Accuracy of Measurements of a CP-Violating Asymmetry

Kirk T. McDonald
Joseph Henry Laboratories, Princeton University, Princeton, NJ 08544
(March 25, 2005)

1 Problem

This problem concerns the statistical analysis of data to determine an asymmetry parameter that arises in the CP-violating decays of the neutral B mesons to a final state that is a CP eigenstate.

The B mesons are produced, for example, at an e^+e^- collider, or at a $p\bar{p}$ collider (as is most relevant to the concerns of this problem), where there is an equal probability for production of B^0 and $\overline{B}^0$ mesons. Presuming that you know whether the B meson was created as a B^0 and $\overline{B}^0$, the time dependence of number of decays to the final state of interest has the form,

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

where,
- N is the total number of decays,
- $+(-)$ labels decays in which the B meson was born as a $B^0(\overline{B}^0)$,
- A is the unknown CP-violating parameter,
- $x = \Delta M/\Gamma$ is the known mixing parameter that desscribes the B^0-$\overline{B}^0$ oscillations,
- time t is measured in units of the B^0-meson lifetime.

What are the errors $\sigma_A(x, N)$ on the asymmetry parameter $A \ll 1$ that can be obtained from observation of a total of N decays for,

a) An analysis that ignores the decay time t?

b) An analysis that takes the decay time t into account?

What is the ratio of the error $\sigma_A(x, N)$ obtained from analysis (a) to the that from analysis (b) for B_d^0 mesons with $x_d \approx 1/\sqrt{2}$ (measured), and for B_s^0 mesons with $x_s \approx 20$ (estimated), again assuming that $A \ll 1$?

Recall that the smallest statistical errors on parameters deduced from data distributed according to forms such as eq. (1) can be obtained via a so-called maximum likelihood *analysis. In brief, if distribution of observations of a measurable quantity t depends on an unknown parameter a and a known parameter x according to a known functional form $f(t, a, x)$, then the likelihood $\mathcal{L}$ of a set $\{t_i\}$ of N observations is given by,*

$$\mathcal{L}(a) = \prod_{i=1}^{N} f(t_i, a, x). \tag{2}$$

A version of the central limit theorem indicates that for large N the likelihood function (2) has a Gaussian dependence on parameter a,

$$\mathcal{L}(a) \propto e^{-(a-a_0)^2/2\sigma_a^2}, \tag{3}$$

where a_0 is the "true" or "best" value of the parameter a, which maximizes the likelihood function (2). The maximum-likelihood estimate of the error, σ_a, on a_0 follows from eq. (3) as,

$$\frac{1}{\sigma_a^2} = -\frac{\partial^2 \ln \mathcal{L}(a_0, x)}{\partial a^2}\,. \tag{4}$$

2 Solution

This problem is extracted from my earlier note *Maximum-Likelihood Analysis of CP-Violating Asymmetries* (Sept. 4, 1992), http://kirkmcd.princeton.edu/tndc/likelihood.pdf

2.1 Time-Integrated Analysis

In this analysis we do not record the time t of the B-meson decay, and simply integrate over the distributions (1) to obtain,

$$\begin{aligned} N_\pm &= \frac{N}{2}\int_0^\infty e^{-t}(1 \pm A \sin xt)\, dt = \frac{N}{2}\int_0^\infty e^{-t}\left(1 \pm A\frac{e^{ixt} - e^{-ixt}}{2i}\right)\, dt \\ &= \frac{N}{2}\int_0^\infty \left(e^{-t} \pm A\frac{e^{-t(1-ix)} - e^{-t(1+ix)}}{2i}\right)\, dt \\ &= \frac{N}{2}\left[1 \pm \frac{A}{2i}\left(\frac{1}{1-ix} - \frac{1}{1+ix}\right)\right] \\ &= \frac{N}{2}\left(1 \pm A\frac{x}{1+x^2}\right). \end{aligned} \tag{5}$$

We now have only two observables, a decay of a B^0 meson which event we call $+$, and a decay of a $\overline{B}^0$ meson which event we call $-$. The event distributions for these observations have the form,

$$f_\pm = N_\pm = N\frac{1 \pm a}{2}, \tag{6}$$

where the asymmetry a is given by,

$$a = \frac{N_+ - N_-}{N_+ + N_-} = A\frac{x}{1+x^2}\,. \tag{7}$$

This is often written as,

$$a = AD, \qquad \text{where} \qquad D = \frac{x}{1+x^2} \tag{8}$$

where D is called the dilution factor associated with the time-integrated analysis.

For an experiment in which N_+ and N_- events are observed, we form the likelihood function,

$$\mathcal{L} = \prod_+ f_+ \prod_- f_- = N_+^{N_+} N_-^{N_-} = N^N \left(\frac{1+a}{2}\right)^{N_+} \left(\frac{1-a}{2}\right)^{N_-}. \tag{9}$$

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

$$\ln \mathcal{L} = N_+ \ln(1+a) + N_- \ln(1-a) + \text{constant}, \tag{10}$$

$$\frac{\partial \ln \mathcal{L}}{\partial a} = \frac{N_+}{1+a} - \frac{N_-}{1-a}, \tag{11}$$

$$\frac{\partial^2 \ln \mathcal{L}}{\partial a^2} = -\frac{N_+}{(1+a)^2} - \frac{N_-}{(1-a)^2}. \tag{12}$$

On setting the first derivative to zero to find the value of a that maximizes the likelihood, we find the usual expression for the asymmetry,

$$a = \frac{N_+ - N_-}{N_+ + N_-}. \tag{13}$$

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,

$$\sigma_a = \sqrt{\frac{1-a^2}{N}}, \tag{14}$$

using eq. (4). This agrees with the usual analysis based on the binomial distribution.

Finally, we obtain the estimate of the error, σ_i, on the CP-violating asymmetry A from the time integrated analysis by recalling eq. (8),

$$\sigma_i = \sigma_A = \frac{\sigma_a}{D} = \frac{1}{D}\sqrt{\frac{1-A^2D^2}{N}} \approx \frac{1}{D\sqrt{N}} = \frac{1+x^2}{x\sqrt{N}}, \tag{15}$$

where the approximation holds for small values of A.

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 $\sigma_i = 2/\sqrt{N}$. As $x \approx 1/\sqrt{2}$ for the B_d^0 meson, a time-integrated analysis is rather effective in this case.

2.2 Time-Dependent Analysis

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,

$$\mathcal{L} = \prod_+ N_+ \prod_- N_- = N^N \prod_+ e^{-t_+}(1 + A \sin xt_+) \prod_- e^{-t_-}(1 - A \sin xt_-), \tag{16}$$

where subscript $+$ labels events in which the B was born as a B^0, and $-$ labels events in which 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 the maximum likelihood method we calculate,

$$\ln \mathcal{L} = -\sum_{+} t_{+} + \sum_{+} \ln(1 + A\sin xt_{+}) - \sum_{-} t_{-} + \sum_{-} \ln(1 - A\sin xt_{-}), \tag{17}$$

$$\frac{\partial \ln \mathcal{L}}{\partial A} = \sum_{+} \frac{\sin xt_{+}}{1 + A\sin xt_{+}} - \sum_{-} \frac{\sin xt_{-}}{1 - A\sin xt_{-}}, \tag{18}$$

$$\frac{1}{\sigma_A^2} = -\frac{\partial^2 \ln \mathcal{L}}{\partial A^2} = \sum_{+} \frac{\sin^2 xt_{+}}{(1 + A\sin xt_{+})^2} + \sum_{-} \frac{\sin^2 xt_{-}}{(1 - A\sin xt_{-})^2}. \tag{19}$$

We estimate the sums by integrals, weighting events according to the distribution (1),

$$\sum_{\pm} f(t_{\pm}) \approx \frac{N}{2} \int_0^\infty dt_{\pm} e^{-t_{\pm}} (1 \pm A_0 \sin xt_{\pm}) f(t_{\pm}), \tag{20}$$

where A_0 is the “true” value of the asymmetry parameter A.

We readily verify that setting $A = A_0$ causes the integral form of eq. (18) to vanish, so that the “best-fit” value of A according to the maximum likelihood method is indeed the “true” value. We do not, however, obtain a simple analytic form of the estimate of A_0 from the time-dependent data.

Turning to the estimate of the error, σ_A, we combine eqs. (19) and (20), and set A to A_0 to find,

$$\begin{aligned}
\frac{1}{\sigma_A^2} &\approx \frac{N}{2} \int_0^\infty dt_{+} \frac{e^{-t_{+}} \sin^2 xt_{+}}{1 + A_0 \sin xt_{+}} + \frac{N}{2} \int_0^\infty dt_{-} \frac{e^{-t_{-}} \sin^2 xt_{-}}{1 - A_0 \sin xt_{-}} \\
&= N \int_0^\infty \frac{dt e^{-t} \sin^2 xt}{1 - A_0^2 \sin^2 xt} \approx N \int_0^\infty dt e^{-t} \sin^2 xt \\
&= N \int_0^\infty dt e^{-t} \left(\frac{e^{ixt} - e^{-ixt}}{2i} \right)^2 = \frac{N}{4} \int_0^\infty dt \left(2e^{-t} - e^{-t(1-2ix)} - e^{-t(1+2ix)} \right) \\
&= \frac{N}{4} \left(2 - \frac{1}{1-2ix} - \frac{1}{1+2ix} \right) = \frac{2x^2 N}{1+4x^2},
\end{aligned} \tag{21}$$

where we ignore the time-varying term in the denominator for small A_0,

We summarize the result (21) of the time-dependent analysis by writing,

$$\sigma_A = \sigma_t \approx \frac{1}{D_t \sqrt{N}} \qquad \text{with} \qquad D_t \equiv \sqrt{\frac{2x^2}{1+4x^2}}. \tag{22}$$

The time-dependent dilution factor D_t is larger than the time-integrated dilution factor of eq. (8) 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 (as is the case for B_s^0 mesons), where a time-integrated analysis yields no information. Indeed, for the time-dependent analysis,

$$\sigma_A \approx \sqrt{\frac{2}{N}} \qquad \text{for large } x. \tag{23}$$

Variants on the time-dependent likelihood function (16) are possible. For example, we could analyze only the shapes *of the time distributions, ignoring the difference between the numbers of events N_+ and N_-. In this case, the distributions (1) should be renormalized to unity, and the corresponding likelihood function would be,*

$$\mathcal{L} = \prod_i \frac{e^{-t_i}(1 + A\sin xt_i)}{1 + A\frac{x}{1+x^2}} \prod_j \frac{e^{-t_j}(1 - A\sin xt_j)}{1 - A\frac{x}{1+x^2}} . \tag{24}$$

The error σ_s on the asymmetry parameter from the shape analysis is,

$$\sigma_s \approx \frac{1}{D_s\sqrt{N}} \qquad \textit{with} \qquad D_s \equiv \frac{x}{1+x^2}\sqrt{\frac{1+2x^4}{1+4x^2}}. \tag{25}$$

This result is, of course, poorer than the full time-dependent analysis (eq. (22)), 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. (15)) for $x < \sqrt{2}$, which includes the case of B_d^0 mesons.

The full time-dependent result (22) 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,

$$\frac{1}{\sigma_t^2} = \frac{1}{\sigma_i^2} + \frac{1}{\sigma_s^2}, \tag{26}$$

on comparing eqs. (22), (15), and (25).

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

$$\sigma_t = \sqrt{\frac{3}{N}} = \frac{1.73}{\sqrt{N}} , \qquad \sigma_i = \frac{3}{\sqrt{2N}} = \frac{2.12}{\sqrt{N}} , \qquad \sigma_s = \frac{3}{\sqrt{N}} . \tag{27}$$

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.

In contrast, the mixing parameter of the B_s^0 mesons is too large to have been measured thus far, but it is estimated that $x_s \approx 20$. In this case, eqs. (15) and (22) yields,

$$\sigma_i \approx \frac{1+x_s^2}{x_s\sqrt{N}} \approx \frac{x_s}{\sqrt{N}} = \frac{20}{\sqrt{N}} , \qquad \text{and} \qquad \sigma_t \approx \sqrt{\frac{2x^2}{(1+4x^2)N}} \approx \frac{1}{\sqrt{2N}} . \tag{28}$$

Here, the time-dependent analysis is 28 times better than the time-integrated analysis.