Statistical and Systematic Errors Introduced into $B-\overline{B}$ Oscillation Measurements by Δz Resolution

R. Cahn

November 9, 1999

Finite resolution of Δz will increase the errors in Δm in the $B-\overline{B}$ mixing measurement and in the measurement of $\sin 2\beta$. A second and perhaps more important error will be induced by the uncertainty in the resolution. If it were known perfectly, its effect could be mostly eliminated, provided the width of the resolution is not too great.

To estimate the errors in quantities like Δm induced by the uncertainty in the Δz resolution we use a maximum likelihood method. Here we compute analytically and numerically without resort to Monte Carlo methods.

1 Maximum Likelihood Method

To analyze data with the unbinned maximum likelihood method we form

$$\mathcal{L} = \sum_j \log f(x_j;\alpha) \tag{1}$$

where f is normalized to unity:

$$\int dx f(x;\alpha) = 1 \tag{2}$$

and where α represents the parameter(s) to be found. If the true value of α is $\overline{\alpha}$, then the events are really distributed as $f(x,;\overline{\alpha})$. The expectation of $\mathcal{L}$ for N data points with the trial values α is

$$\langle\mathcal{L}(\alpha)\rangle = N\int dx f(x;\overline{\alpha}) \log f(x;\alpha) \tag{3}$$

In the event there are separate categories of events (e.g. unmixed and mixed), with distributions $f_i(x;\alpha)$, the proper normalization is

$$\sum_i \int dx f_i(x;\alpha) = 1 \tag{4}$$

and the expectation for the log-likelihood function is

$$\langle \mathcal{L} \rangle = N \int dx f_i(x; \overline{\alpha}) \log f_i(x; \alpha) \tag{5}$$

As described in the *Review of Particle Properties 1998*, pp. 172-177, log-likelihood can be related to χ^2 and thus to traditional measures of uncertainty. The covariance matrix is given by

$$V_{mn} = \left(\langle - \frac{\partial^2 \mathcal{L}}{\partial \alpha_m \partial \alpha_n} \rangle \right)^{-1} \tag{6}$$

The expected uncertainty in measuring α_i is

$$\sigma_i = \sqrt{V_{ii}} \tag{7}$$

which gives the familiar result for a single parameter

$$\sigma^{-2} = N \int dx \frac{1}{f} \left(\frac{\partial f}{\partial \alpha} \right)^2 \tag{8}$$

where all quantities are evaluated at the true value, $\overline{\alpha}$. In this way we can calculate, a priori, anticipated uncertainties in various measured quantities.

2 $B - \overline{B}$ Oscillations

We consider specifically $B - \overline{B}$ mixing with the time behavior for unmixed and mixed flavor combinations. Before smearing in Δz, which is directly related to smearing in t, we have

$$\begin{aligned} f_+(\Gamma, \Delta m, b, t) &= \frac{1}{2} \Gamma e^{-\Gamma t} [1 + b \cos \Delta m t] \\ f_-(\Gamma, \Delta m, b, t) &= \frac{1}{2} \Gamma e^{-\Gamma t} [1 - b \cos \Delta m t] \end{aligned} \tag{9}$$

After smearing, we write

$$\begin{aligned} F_+(\Gamma, \Delta m, b, t, \sigma) &= \int dt' \frac{1}{\sqrt{2\pi}\sigma} e^{-(t-t')^2/2\sigma^2} f_+(\Gamma, \Delta m, b, t') \\ F_-(\Gamma, \Delta m, b, t, \sigma) &= \int dt' \frac{1}{\sqrt{2\pi}\sigma} e^{-(t-t')^2/2\sigma^2} f_-(\Gamma, \Delta m, b, t') \end{aligned} \tag{10}$$

It is possible to express these functions in terms of the complex error function:

$$\mathrm{erfc}(z) = \frac{2}{\sqrt{\pi}} \int_z^\infty e^{-t^2} dt \tag{11}$$

which has the property $\mathrm{erfc}(0) = 1, \mathrm{erfc}(-\infty) = 2$. We find

$$\begin{aligned} F_\pm(\Gamma, \Delta m, b, t, \sigma) &= \frac{\Gamma}{4} e^{-\Gamma t} \mathrm{Re}\left[e^{\sigma^2 t^2/2} \mathrm{erfc}(\frac{-t}{\sqrt{2}\sigma} + \frac{\sigma\Gamma}{\sqrt{2}}) \right. \\ &\left. \pm e^{i\Delta m t} e^{(\sigma\Gamma - i\sigma\Delta m)^2/2} \mathrm{erfc}(\frac{-t}{\sqrt{2}\sigma} + \frac{\sigma\Gamma}{\sqrt{2}} - i\frac{\sigma\Delta m}{\sqrt{2}}) \right] \end{aligned} \tag{12}$$

In fact, we measure only the distance Δz, an unsigned quantity. The original distribution is confined to $t > 0$. We fold the smeared distribution, making a distribution that is a function of $|t|$ alone. We shall not introduce separate notation for this distribution, which is obtained by adding the distributions for t positive and t negative.

We indicate the true values of the parameters with a bar: $\overline{\Gamma}$, $\overline{\Delta m}$, $\overline{b} = 1-2\overline{w}$, $\overline{\sigma}$. We will determine which values of Γ, $b = 1-2w$, and Δm maximize the log-likelihood for values of σ that differ from the true value $\overline{\sigma}$. This will enable us to answer questions like: if the true resolution is a gaussian with width 125 μm and instead we believe the width is 150 μm, what shift in the parameters Γ, Δm, and b will result?

The expectation for the log-likelihood as a function of the unknowns Γ, Δm, and b, with σ fixed (at the wrong value) is

$$\begin{aligned} \langle \mathcal{L} \rangle &= \int_0^\infty dt [F_+(\overline{\Gamma}, \overline{\Delta m}, \overline{b}, t, \overline{\sigma}) \log F_+(\Gamma, \Delta m, b, t, \sigma) \\ &+ F_-(\overline{\Gamma}, \overline{\Delta m}, \overline{b}, t, \overline{\sigma}) \log F_-(\Gamma, \Delta m, b, t, \sigma)] \end{aligned} \tag{13}$$

3 Systematic Errors when the Resolution is Poorly Known

We can explore systematic errors by evaluating the log-likelihood with a parameter fixed at the wrong value. Fixing one parameter will skew the results for others. As we vary σ, the values of Γ, Δm, and b that maximize the log-likelihood vary. The results are shown in three figures, for the input values $\overline{\Gamma} = 0.004\ \mu\mathrm{m}^{-1}$, $\overline{\Delta m} = 0.028\ \mu\mathrm{m}^{-1}$, $\overline{b} = 0.6$, $\overline{\sigma} = 125\ \mu$m. We have chosen to express all quantities as equivalent distances.

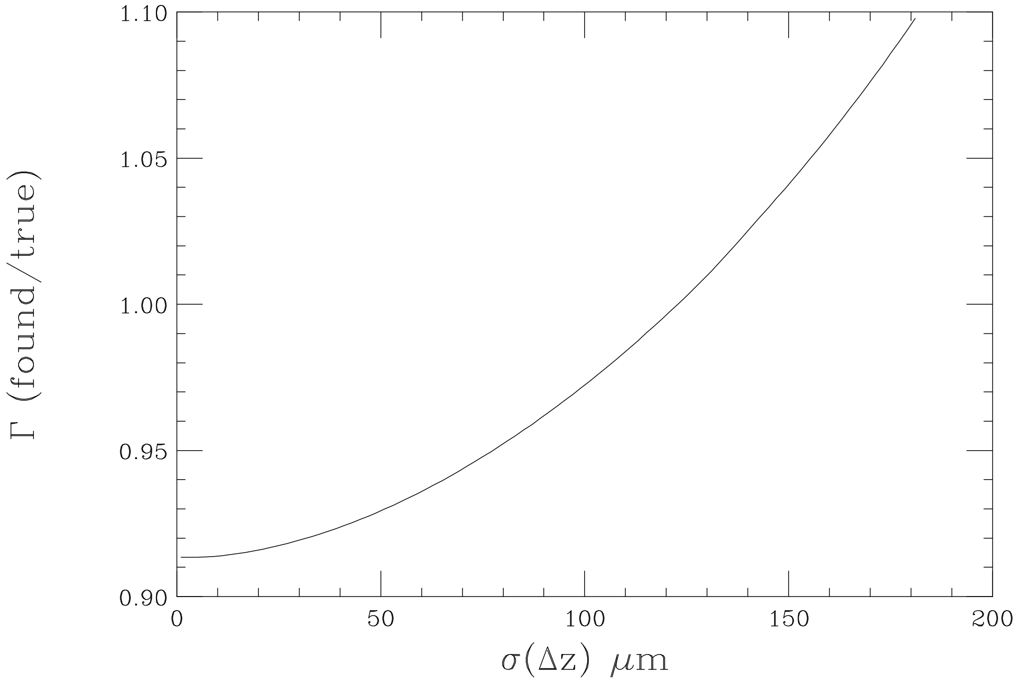

Figure 1: Variation in the inferred value of Γ as a function of the assumed value of σ when the true value is 125 μm.

A comparison of the three can be obtained by noting the percentage shift between their values at $\overline{\sigma} = 125\ \mu$m and at $\sigma = 160\ \mu$m. These shifts are for Γ 5.7%, for Δm 8.2%, and for b 2.1%.

We can make a simple argument to explain the behavior of Δm. Suppose the t distributions are smeared by some general resolution function $G(t-t')$.

$$
\begin{aligned}
F_+(\Gamma, \Delta m, b, t, \sigma) &= \int dt' G(t-t') f_+(\Gamma, \Delta m, b, t') \\
F_-(\Gamma, \Delta m, b, t, \sigma) &= \int dt' G(t-t') f_-(\Gamma, \Delta m, b, t') \qquad (14)
\end{aligned}
$$

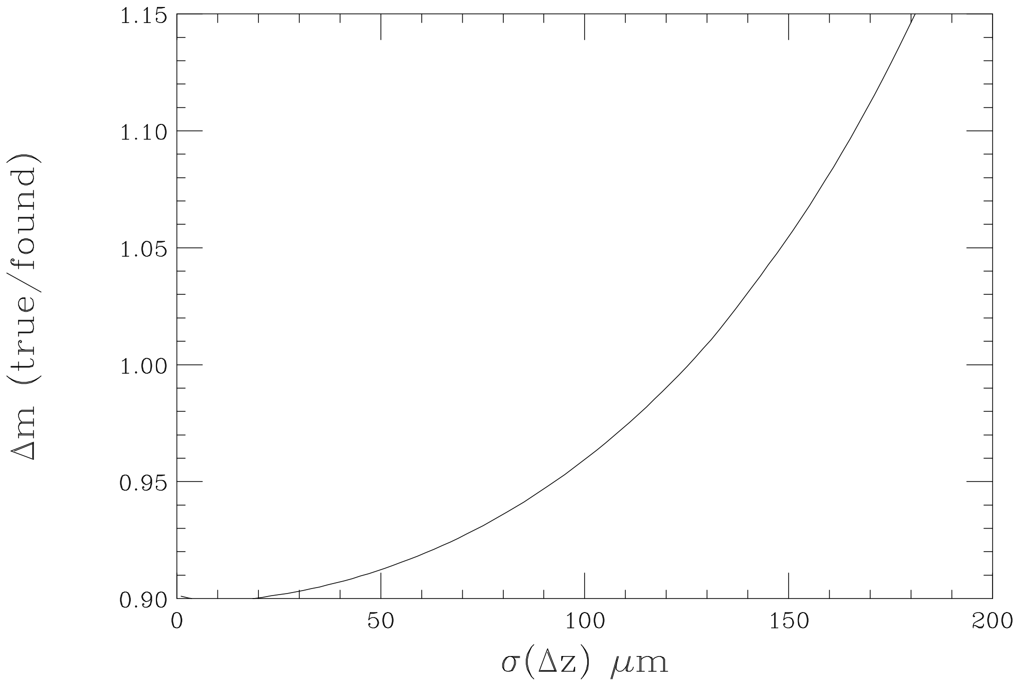

Figure 2: Variation in the inferred value of Δm as a function of the assumed value of σ when the true value is 125 μm.

If we treat the width of the resolution function as a small quantity, we can expand to find

$$F_{\pm}(\Gamma, \Delta m, b, t, \sigma) = \frac{1}{2}\Gamma[1 \pm b(\cos \Delta m t + \Gamma \Delta m \langle \sigma^2 \rangle \sin \Delta m t)] \tag{15}$$

where $\langle \sigma^2 \rangle$ is the square of the width of the resolution function.

The asymmetry will go through zero not at $t = \pi/(2\Delta m)$ but at $t = \pi/(2\Delta m) + \Gamma\langle\sigma^2\rangle$. If we were not aware of this effect, we would infer a value of Δm that was too small, with a shift

$$\frac{\delta \Delta m}{\Delta m} = -\frac{2}{\pi}\Delta m \Gamma \langle \sigma^2 \rangle = -\frac{2}{\pi}\frac{\Delta m}{\Gamma}\langle \Gamma^2 \sigma^2 \rangle \tag{16}$$

With $1/\Gamma = 250$ μm, $\langle\sigma^2\rangle^{1/2} = 125$ μm, $x = \Delta m/\Gamma = 0.7$, we find $\delta\Delta m/\Delta m = -0.11$. This is in good agreement with Fig. 3.

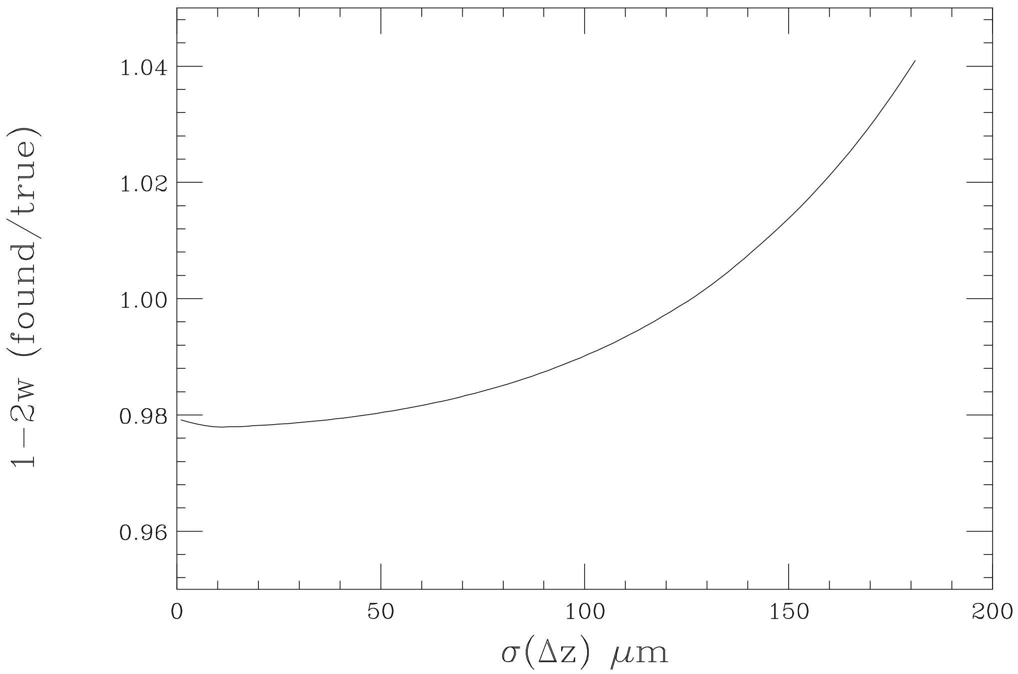

Figure 3: Variation in the inferred value of $b = 1 - 2w$ as a function of the assumed value of σ when the true value is 125 μm.

4 Increase in Statistical Errors Due to Finite Δz Resolution

The uncertainty in a measurement can be predicted from the matrix of second (partial) derivatives of χ^2 evaluated at the true values of the parameters. The diagonal elements of the inverse of this matrix give the squares of the various σs. See Eqs. (6)-(7). It is this inverse matrix that is provided by MINUIT. If we simulate with N events, we get from MINUIT a good estimate of the statistical errors. Our approach is to calculate the expectation of negative log-likelihood (essentially half the χ^2). We then use MINUIT to find the matrix of second derivatives, which gives us the coefficient of $1/\sqrt{N}$ in the statistical errors. The results are shown in the following Figures.

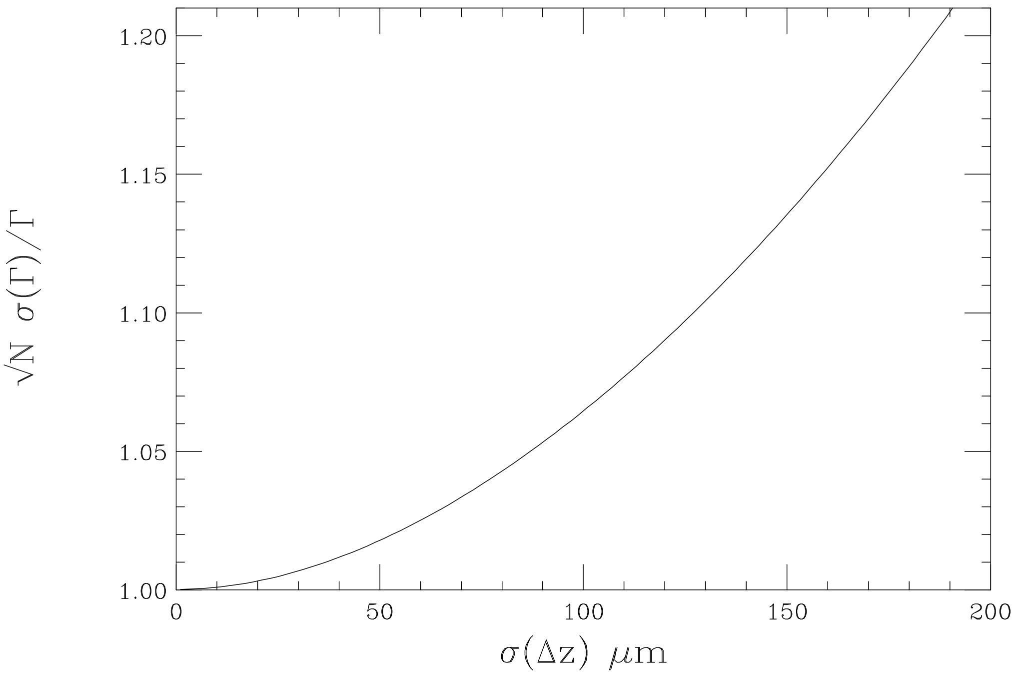

Figure 4: The fractional statistical error in Γ as a function of the width σ of the resolution of Δz, supposing the resolution to be known precisely.

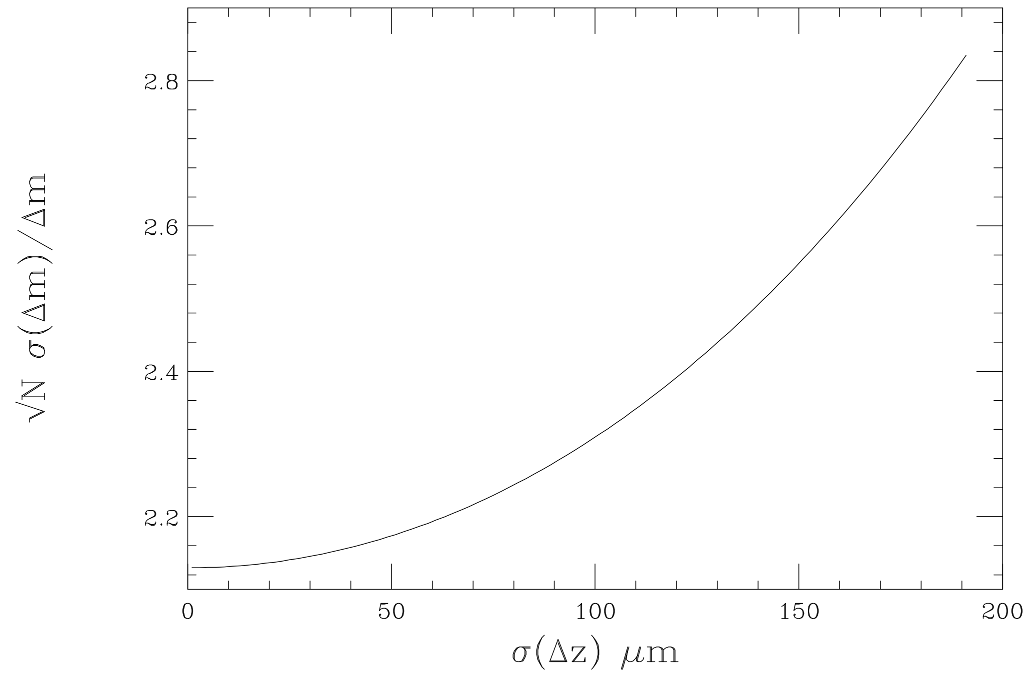

Figure 5: The fractional statistical error in Δm as a function of the width σ of the resolution of Δz, supposing the resolution to be known precisely.

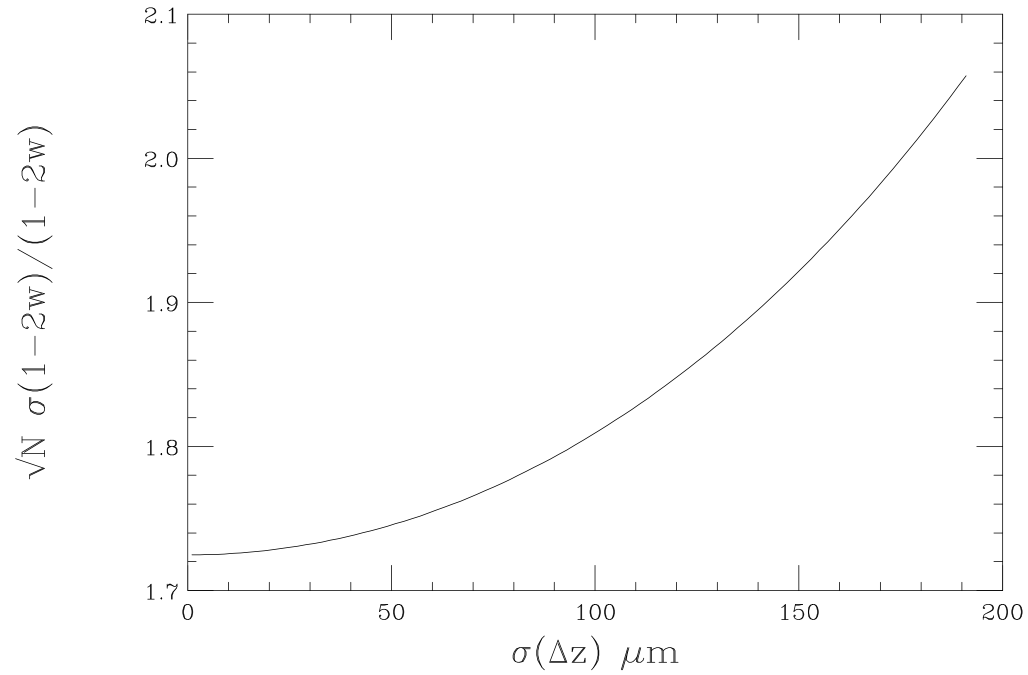

Figure 6: The fractional statistical error in $1-2w$ as a function of the width σ of the resolution of Δz, supposing the resolution to be known precisely.

5 Fitting for σ

An obvious possibility is to try to infer σ from the measured Δz distributions themselves. That is, we can simply regard σ as one more parameter to be fit. In Figs. 7 and 8 we show the distributions for the unmixed and mixed distributions with two values of σ, 125 μm and 150 μm. There is clearly some distribution between the two. The question is how well they can be distinguished statistically.

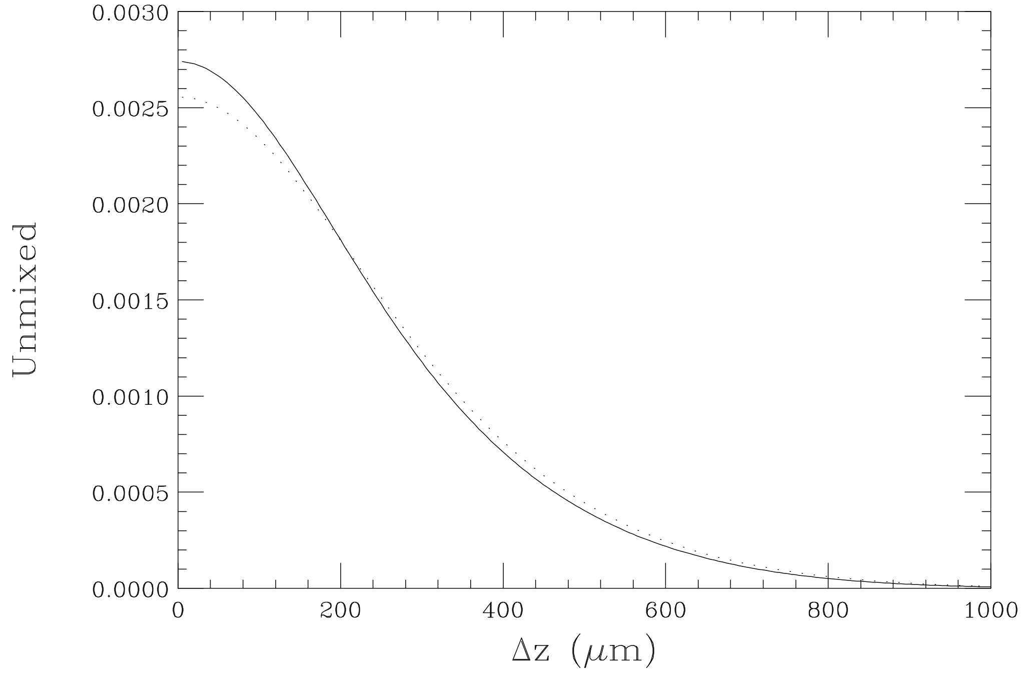

Figure 7: The distribution of unmixed events with no mistagging. The solid curve shows $\sigma =$ 125μm and the dotted curve $\sigma = 150$ μm.

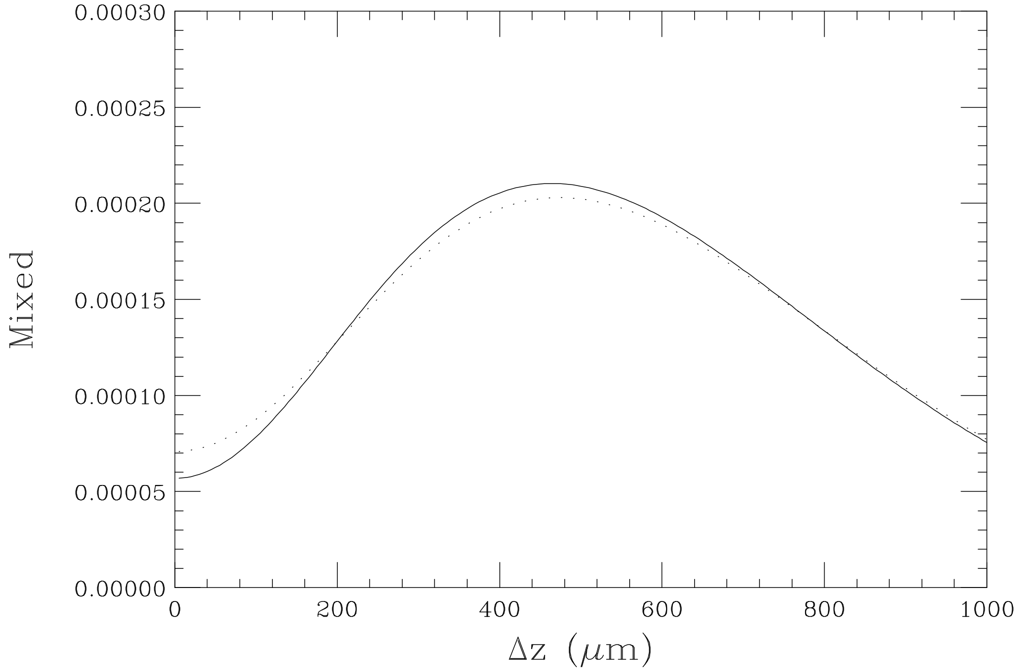

Figure 8: The distribution of mixed events with no mistagging. The solid curve shows $\sigma = 125$ μm and the dotted curve $\sigma = 150$ μm.

	σ fixed	σ variable
$\sqrt{N}\sigma(\Gamma)/\Gamma$	1.10	1.56
$\sqrt{N}\sigma(\Delta m)/\Delta m$	2.45	2.88
$\sqrt{N}\sigma(b)/b$	1.86	1.90
$\sqrt{N}\sigma(\sigma)/\sigma$		6.42

Table 1: Statistical uncertainties with the Δz resolution fixed and variable. The values used are $\sigma = 125\ \mu$m and $b = 1-2w = 0.6$. When σ is variable, it is obtained from fitting the distributions for the mixing events. The column with σ fixed corresponds to Figs. 4,5, and 6, for the value $\sigma(\Delta z) = 125\ \mu$m.

	Γ fixed	Γ variable
$\sqrt{N}\sigma(\Gamma)/\Gamma$		1.56
$\sqrt{N}\sigma(\Delta m)/\Delta m$	2.60	2.88
$\sqrt{N}\sigma(b)/b$	1.88	1.90
$\sqrt{N}\sigma(\sigma)/\sigma$	4.58	6.42

Table 2: Statistical uncertainties with Γ fixed and variable. The values used are $\sigma = 125\ \mu$m and $b = 1 - 2w = 0.6$. When Γ is variable, it is obtained from fitting the distributions for the mixing events.

Now fixing the true value of $\overline{\sigma} = 125\ \mu$m, we find the statistical errors indicated in Table 1. While the uncertainty in the Δz resolution is large (6.4% even after 10,000 events), it is good enough, degrading the resolution of Δm by only 15% from what it would be if the Δz resolution were known perfectly.

It might be thought that a substantial improvement would be found if we knew in advance the value of Γ. This seems not to be so, as seen in Table 2. Knowing Γ a priori improves the Δm resolution by 10%.

As a test, Brad Abbott generated events using a"toy" Monte Carlo, providing Δz for both mixed and unmixed events. Initially the events were propertly tagged. With $w = 0$ fixed, we find that the expected fractional errors in Γ, Δm, and σ are $1.34/\sqrt{N}$, $1.40/\sqrt{N}$, and $4.56/\sqrt{N}$. For the example here with $N = 1000$ we expect errors of 4.2%, 4.4%, and 14% respectively. The example in Figs. 9 and 10 shows uncertainties of 4.6%, 4.9%, and 15% in good agreement.

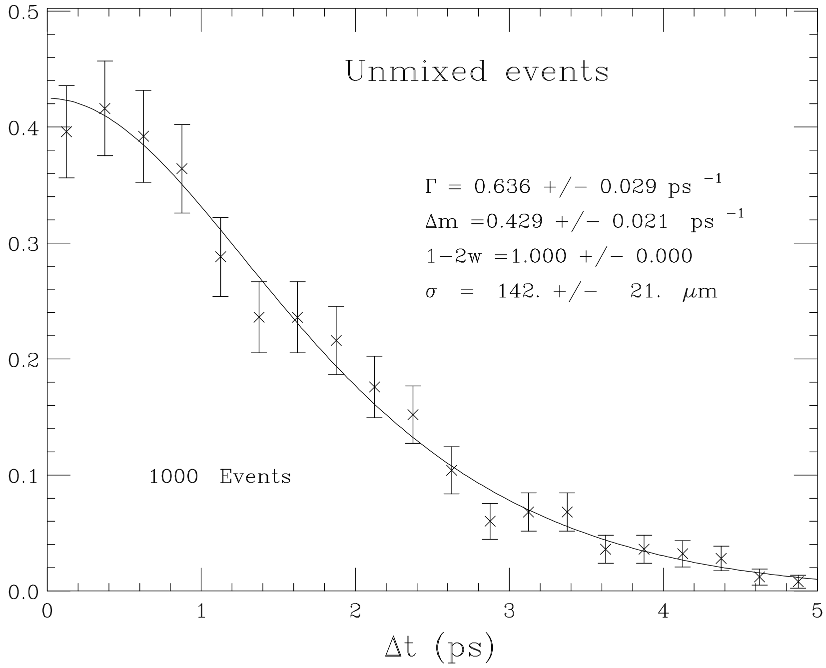

Figure 9: Fit to the unmixed events [out of 1000 total events] with no mistags.

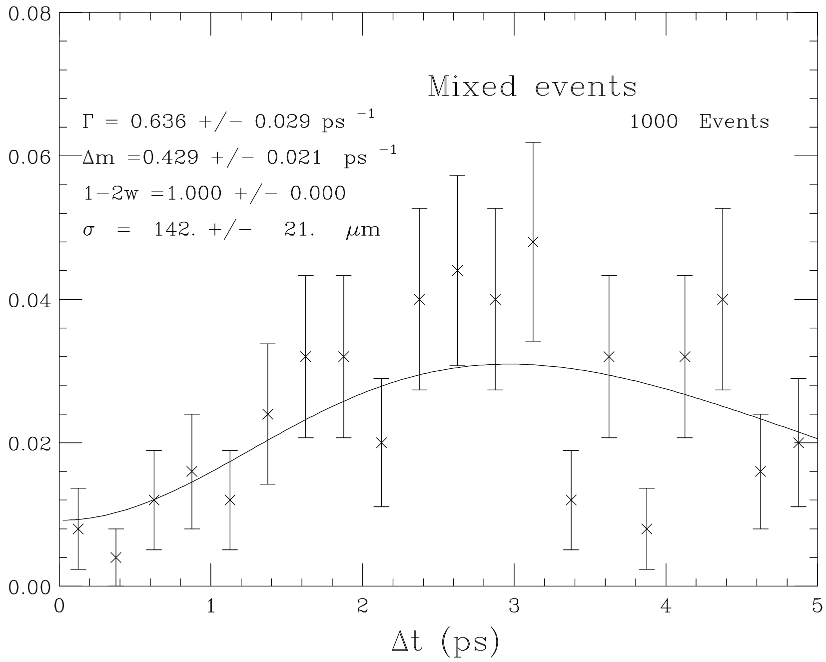

Figure 10: Fit to the mixed events [out of 1000 total events] with no mistags.

To introduce mistags we simply transfer 20% of the unmixed events to the mixed and 20% of the initially mixed back to the unmixed. From the previous tables we find that the expected fractional errors in Γ, Δm, b and σ are $1.56/\sqrt{N}$, $2.88/\sqrt{N}$, $1.90/\sqrt{N}$ and $4.56/\sqrt{N}$. For the example here with $N = 1000$ we expect errors of 4.9%, 9.1%, 6.0, and 20% respectively. The example in the shows uncertainties of 5.2%, 8.9%, 6.0, and 17% in good agreement.

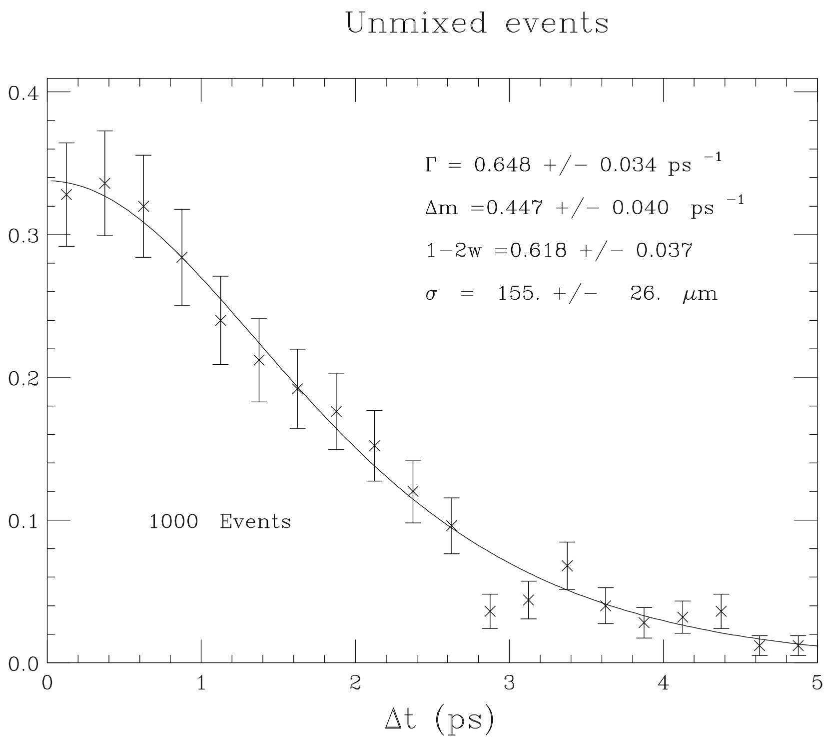

Figure 11: Fit to the unmixed events [out of 1000 total events] with 20% mistags. Here all four parameters are free in the fit.

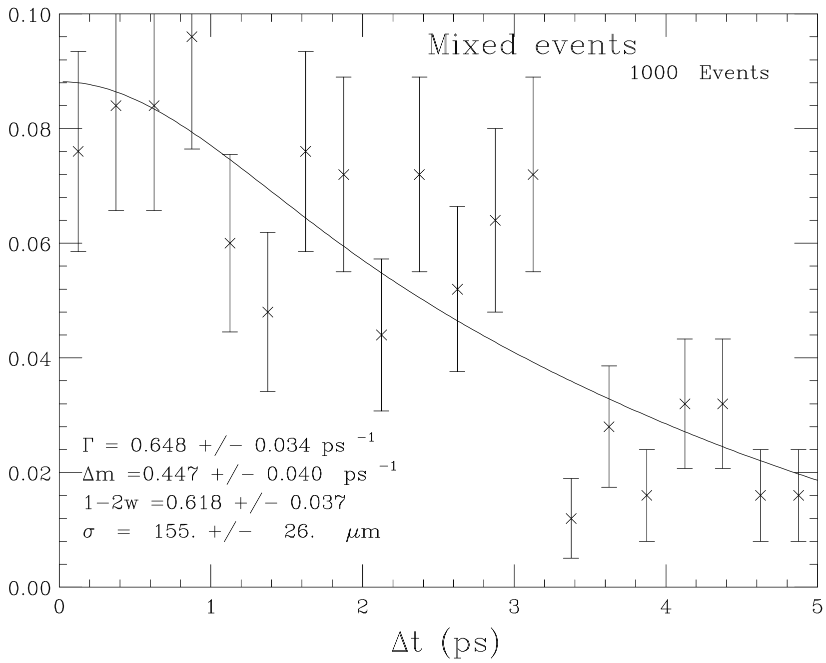

Figure 12: Fit to the mixed events [out of 1000 total events] with 20% mistags. Here all four parameters are free in the fit.

6 Commentary

The neutral B lifetime is known to about 2%. The PDG gives $\tau_{B^0} = 1.54 \pm 0.03$ ps. The mass splitting is known to about 4%. The PDG cites two values, $\Delta m_d = 0.464 \pm 0.018$ ps^{-1} and $\Delta m_d = 0.470 \pm 0.019$ ps^{-1}. These represent a 4% measurement. However, the individual LEP experiments each have measurements only at the 7% level. A 5% measurement by BaBar would be of great interest. Setting aside backgrounds (which we have ignored everywhere), these seems to require around 3300 events, which might be attainable with the event sample through June 2000.

7 Comparison with Time-Independent Measurement

A measurement of $x = \Delta m/\Gamma$ is possible without any determination of Δz. From Eq. (9) we see that the expected number of unmixed and mixed events, U and M are

$$\begin{aligned} U &= \frac{N}{2}\left(1 + \frac{b}{1+x^2}\right) \\ M &= \frac{N}{2}\left(1 - \frac{b}{1+x^2}\right) \end{aligned} \tag{17}$$

so that the asymmetry is given by

$$A = \frac{U-M}{U+M} = \frac{b}{1+x^2} \tag{18}$$

If b is perfectly known, the statistical uncertainty in A is given by

$$\sigma(A)^2 = \frac{1}{N}\left(1 - \frac{b^2}{(1+x^2)^2}\right) \tag{19}$$

from which we find

$$\sigma(x)_{TI} = \frac{1+x^2}{2xb}\sqrt{(1+x^2)^2 - b^2}\frac{1}{\sqrt{N}} \tag{20}$$

The corresponding time-dependent measurement of x with Γ, Δm, and b known gives an uncertainty in x determined by

$$\sigma(x)_{TD}^{-2} = N \int_0^\infty dy y^2 e^{-y} \frac{b^2 \sin^2 xy}{1 - b^2 \cos^2 xy} \tag{21}$$

Numerically, we find that for $x = 0.7$, the time-dependent measurement gives an error only 0.4 times that of the time-independent measurement, for the same number of events. This result is relatively independent of the value of b, which is assumed known.

In any event, we should be able to do whatever CLEO did and since even a time-independent approach can give x, the uncertainty in x should not depend much on the Δz resolution.

Accordingly, we calculate the systematic uncertainty in x coming from an uncertainty in the Δz resolution by taking Γ, x, and b as our independent parameters. The result for x is shown in Fig. 13.

As expected, x is far less sensitive to the Δz resolution than Δm is.

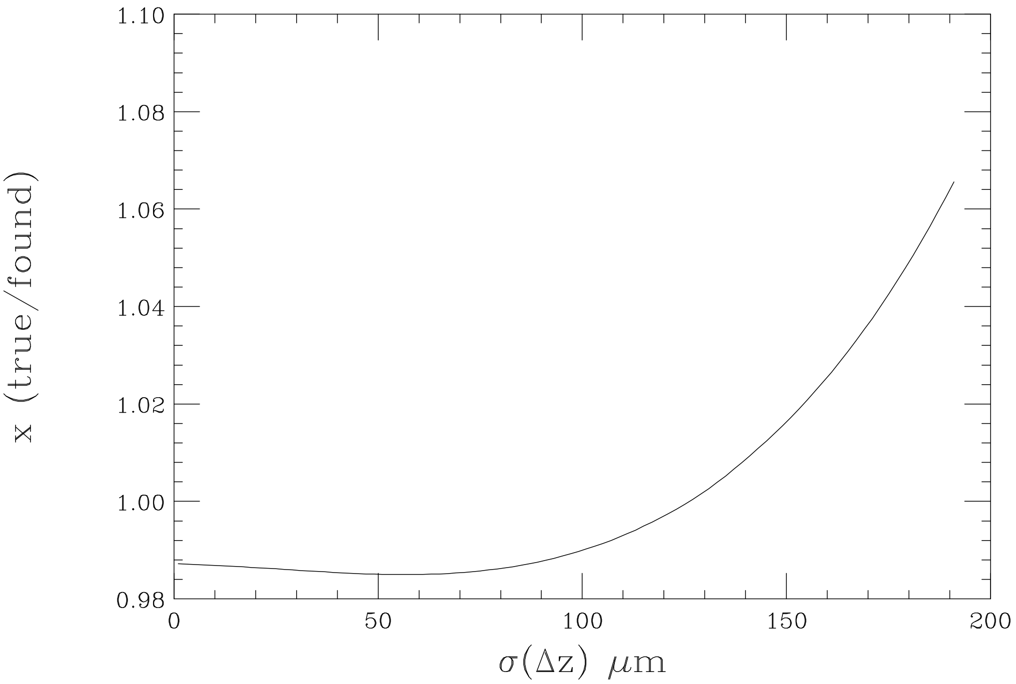

Figure 13: Variation in the inferred value of x as a function of the assumed value of σ when the true value is 125 μm.

	True	Found	Fractional uncertainty $\times\sqrt{N}$
Γ (μm)	4.00×10^{-2}	4.03×10^{-2}	2.80
Δm (μm)	2.80×10^{-2}	2.81×10^{-2}	4.77
b	0.60	0.60	2.30
σ_1 (μm)	80.	82.7	45.
σ_2 (μm)	200.	201.	67.
narrow fraction	0.75	0.74	57.

Table 3: Attempt to infer all the parameters of a double Gaussian fit from the observed $B - \overline{B}$ distributions.

8 Systematic Errors from using single Gaussian

It is the practice in BaBar to use a double Gaussian, one narrow and one wide, to represent the resolution function for Δz. We take here the narrow Gaussian to have a width of 80 μm and the wide one a width of 200 μm, with 80% of the distribution in the narrow Gaussian.

It seems improbable that these three numbers, two widths and a fraction, can be inferred from the oscillation data. However, there is nothing to stop us from trying. Results are shown in Table 8.

Indeed, the uncertainties in the widths and fraction in the narrow Gaussian are very large. The fractional uncertainties in Δm and b are larger than when we used just a single Gaussian.

Suppose we analyze the double Gaussian distribution using just a single Gaussian. In Table 8 we see that the inferred resolution is near the rms value. The errors in Δm and b are similar to what we found when the true distribution came from a single Gaussian. However, there is a shift in the deduced value of Δm, a reduction of -2.85%.

	True	Found	Fractional uncertainty $\times\sqrt{N}$
Γ (μm)	4.00×10^{-2}	3.94×10^{-2}	1.47
Δm (μm)	2.80×10^{-2}	2.72×10^{-2}	2.64
b	0.60	0.59	1.85
σ_1 (μm)	80.	104.	10.9
σ_2 (μm)	200.		
narrow fraction	0.75	1.0	

Table 4: Analysis of a double Gaussian using just a single Gaussian. The statistical errors are small, but there is a sytematic error in Δm.