Sensitivity Limits for a Measurement of Dilution in B_s^0-$\bar{B}_s^0$ Mixing

Jamie E. Hegarty

June 29, 2005

Abstract

Using an unbinned likelihood amplitude fitting method with a toy Monte Carlo simulation of B_s^0-$\bar{B}_s^0$ mixing events, we have determined preliminary sensitivity limits at for a measurement of dilution. Our best value using this method is $\Delta m < 13.386 ps^{-1}$.

1 Introduction

At the forefront of High Energy Physics are efforts to verify or modify the Standard Model of Fundamental Particles and Interactions – the theoretical framework which classifies and relates all known elementary particles. The Standard Model describes leptons and quarks as fundamental constituents of matter, which interact via force-carrying bosons, the mediators of fields. There are three generations, or "flavors" of quarks, which may combine via interaction with gluons to form composite hadrons such as baryons and mesons.

This project is concerned with B_s^0 and $\bar{B}_s^0$ mesons and a process known as mixing. A B_s^0 is comprised of an s-quark and a $\bar{b}$-quark, while a $\bar{B}_s^0$ contains a b and an $\bar{s}$. Each of these flavor states exists in a superposition of two mass states, essentially a light B_s^L and a heavy B_s^H. The B_s^L and B_s^H can then be expressed as linear combinations of the flavor states [1]:

$$|B_s^L\rangle = p|B_s^0\rangle + q|\bar{B}_s^0\rangle \quad \text{and} \quad |B_s^H\rangle = p|B_s^0\rangle - q|\bar{B}_s^0\rangle \qquad (1)$$
$$(\text{where} \quad |p|^2 + |q|^2 = 1)$$

The difference in mass between these two is defined by the Standard Model as $\Delta m = M_H - M_L$, where M_H and M_L are the masses of the B_s^H and B_s^L states, respectively. Δm is very small, but always positive. The current experimental lower limit is $\Delta m \geq 14.6 ps^{-1}$ [1]. While our ultimate goal is to actually measure Δm, this project focuses on finding an upper limit for making this measurement.

Mixing describes the process by which a B_s^0 meson may change into a $\bar{B}_s^0$ meson through the exchange of virtual $W^{\pm}$ bosons. Figure 1 illustrates this process. Mixing is a weak, flavor-changing loop interaction dominated by the top quark. Charm or up loops may also occur, although much more rarely. The likelihood of a particular weak flavor-changing decay is proportional to the coupling strength of the quarks to the $W^{\pm}$ field (given by the coupling

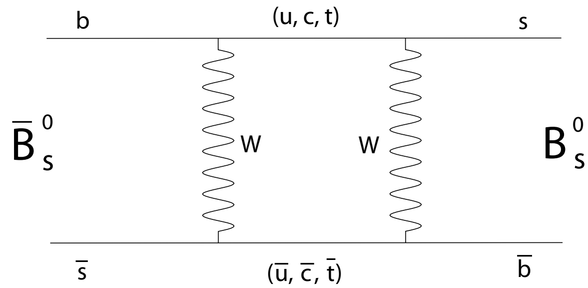

Figure 1: B^0_s-$\bar{B}^0_s$ **mixing via** $W^\pm$ **bosons.**

constant α_W) and to the element of the CKM matrix[1] V which describes the particular interaction [2] [3]. For example, the probability for a t-quark to decay to an s-quark via a virtual W^+ is proportional to $\alpha_W V_{ts}$. Therefore, this element of the CKM matrix is also proportional to the probability that B^0_s mixing will occur, and likewise related to the mass difference Δm. A measurement or limit on Δm then provides useful information about V_{ts}, the CKM matrix, and the Standard Model.

Mixing not only allows a B^0_s to change into a $\bar{B}^0_s$ but for oscillation between the two flavor states. The oscillation frequency of B^0_s-$\bar{B}^0_s$ mixing is directly related to the mass difference Δm between the two mass eigenstates, and so measuring this frequency would necessarily yield measurement of Δm. However, quite a number of factors contribute to the difficulty of making such a measurement, including the high frequency range that Δm is likely to fall into and the inherent smearing of lifetime measurements due to detector resolution. Because of these, fitting data for Δm alone could return several possible frequency values, and one way to narrow down these potential measurements is to fit for and measure the amplitude of the oscillations as well.

This analysis uses a toy Monte Carlo simulation to generate mixing events for an extremely high frequency Δm, so that the factors affecting the sensitivity of a measurement may be evaluated independently. Only the very basic mixing and detector physics are simulated, and any discrepancies resulting from more complicated effects such as multiple decay modes or background events are ignored. In this way, we have determined preliminary upper limits of sensitivity to changes in resolution for measurements of the B^0_s-$\bar{B}^0_s$ oscillation amplitude, or dilution.

2 Definitions

It is useful to first define some terms:

[1]The Cabibbo-Kobayashi-Maskawa (CKM) matrix V contains nine terms to describe flavor-changing quark interactions and is crucial to the Standard Model [2] [3]:

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

1. **Tagging**: Determining whether a B_s is a B_s^0 or a $\bar{B}_s^0$. This is done by checking the charge of muons on the opposite side of the collision event from the B_s decay.
2. **Unmixed State**: A particle which was tagged as a B_s^0 is reconstructed as a B_s^0, and one tagged as a $\bar{B}_s^0$ is reconstructed as a $\bar{B}_s^0$. Mixing has not occurred.
3. **Mixed State**: A particle which was tagged as a B_s^0 is reconstructed as a $\bar{B}_s^0$, or a $\bar{B}_s^0$ is reconstructed as a B_s^0. Mixing has occurred.
4. **Mistagging**: Incorrectly tagging a B_s^0 as a $\bar{B}_s^0$, or vice versa. The mistagging rate α corresponds to the percent of events for which the tag is assigned incorrectly.
5. **Dilution**: $D = 1 - 2\alpha$, where α is the mistagging rate as described above.
6. **Sensitivity**: The value of Δm for which the error in the fitted value of D (at the 95% confidence limit) is 1, independent of the "actual" value of Δm used in the Monte Carlo.

3 Mixing and Asymmetry Calculations

Before any simulations or fitting can be done, the B_s^0 lifetime distribution must be considered in terms of mixed and unmixed states. The lifetime distributions of the mixed (M) and unmixed (U) states, and the corresponding asymmetry are:

$$U(t) = e^{-\frac{t}{\tau}}[1 + \cos(\Delta m t)], \quad \text{and} \quad M(t) = e^{-\frac{t}{\tau}}[1 - \cos(\Delta m t)]$$

$$\Rightarrow \quad \frac{U(t) - M(t)}{U(t) + M(t)} = \cos(\Delta m t)$$

where t is the lifetime of the B_s^0, τ is the mean lifetime of the B_s^0, and Δm is the mass difference between the B_s^H and the B_s^L. With mistagging rate α included, the distributions of mixed ($N_m(t)$) and unmixed ($N_u(t)$) B_s^0 become:

$$N_m(t) = (1-\alpha)M(t) + \alpha U(t) = (1-\alpha)\{e^{-\frac{t}{\tau}}[1 - \cos(\Delta m t)]\} + \alpha\{e^{-\frac{t}{\tau}}[1 + \cos(\Delta m t)]\}$$
$$N_u(t) = (1-\alpha)U(t) + \alpha M(t) = (1-\alpha)\{e^{-\frac{t}{\tau}}[1 + \cos(\Delta m t)]\} + \alpha\{e^{-\frac{t}{\tau}}[1 - \cos(\Delta m t)]\}$$

The asymmetry may also be calculated with mistagging included:

$$\begin{aligned}
\frac{N_u(t) - N_m(t)}{N_u(t) + N_m(t)} &= \frac{(1-\alpha)U(t) + \alpha M(t) - [(1-\alpha)M(t) + \alpha U(t)]}{(1-\alpha)U(t) + \alpha M(t) + [(1-\alpha)M(t) + \alpha U(t)]} \\
&= (1-2\alpha)\frac{U(t) - M(t)}{U(t) + M(t)} \\
&= (1-2\alpha)\cos(\Delta m t)
\end{aligned}$$

and it becomes clear that the amplitude of oscillation is simply the dilution $D = 1 - 2\alpha$.

However, since $N_m(t)$ and $N_u(t)$ differ by only the sign of the $\cos(\Delta m t)$ terms, an extra parameter "`tag`" may be used, such that `tag` $= 1$ corresponds to the mixed state, and

$\mathtt{tag} = -1$ corresponds to the unmixed state. The separate $N_m(t)$ and $N_u(t)$ can then be generalized to a single distribution function:

$$N(t,\mathtt{tag}) = (1-\alpha)\{e^{-\frac{t}{\tau}}[1-\mathtt{tag}\cdot\cos(\Delta m t)]\} + \alpha\{e^{-\frac{t}{\tau}}[1+\mathtt{tag}\cdot\cos(\Delta m t)]\} \tag{2}$$

Additionally, Gaussian smearing to account for the proper time resolution of the detector must also be factored in. The time-smeared distributions of the mixed and unmixed B_s^0 are found by convoluting $N(t,\mathtt{tag})$ with the normalized Gaussian $G(t) = \frac{1}{\sqrt{2\pi}\,\sigma}e^{-(\frac{t}{\sqrt{2}\sigma})^2}$:

$$f(t,\mathtt{tag}) = \frac{(1-\alpha)}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{\infty} e^{-(\frac{t-t'}{\sqrt{2}\sigma})^2} e^{-\frac{t'}{\tau}}[1-\mathtt{tag}\cdot\cos(\Delta m t')]\; dt' \;\ldots$$

$$+\frac{\alpha}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{\infty} e^{-(\frac{t-t'}{\sqrt{2}\sigma})^2} e^{-\frac{t'}{\tau}}[1+\mathtt{tag}\cdot\cos(\Delta m t')]\; dt'$$

Integrating and making a few simplifying substitutions yields a function suitable to fit with:

$$f(t,\mathtt{tag}) = (1-\alpha)\frac{e^B}{2\tau}\left[1+\mathrm{erf}(A)-\mathtt{tag}\cdot w_r^- e^{-A^2}\right] + \alpha\frac{e^B}{2\tau}\left[1+\mathrm{erf}(A)+\mathtt{tag}\cdot w_r^- e^{-A^2}\right] \tag{3}$$

$$A = \frac{t}{\sqrt{2}\,\sigma} - \frac{\sigma}{\sqrt{2}\,\tau}\;; \qquad B = \frac{\sigma^2}{2\tau^2} - \frac{t}{\tau}\;; \qquad C = \frac{\sigma\Delta m}{\sqrt{2}}$$

where w_r^- is the real part of the complex error function $w(z)$ [5], evaluated at $z = C - iA$. It is important to note that as A increases, the term $w_r^- e^{-A^2} \to \cos(2AC)$.

4 Monte Carlo Simulation

B_s^0 events are generated using Root [4] to run a toy Monte Carlo (MC) simulation. Proper time resolution σ_t, lifetime-dependent resolution σ_n and mistagging rate α are used as simple parameters to model basic B_s^0 production and detector physics in an idealistic way, while more complex factors such as background events, multiple decay modes, and detector geometry are left out. While a real detector would require far more parameters to be accurately simulated, this MC serves as a preliminary idealistic simulation and is sufficient to serve the purposes of this analysis. The MC runs as follows:

1. For each event, the proper lifetime t_p of the B_s^0 is first selected from a perfect exponential distribution with a mean at $\tau = 1.5\times 10^{-12}s$. Next, a random time "smearing" value, selected from a Gaussian distribution centered at zero, and having width $\sigma = \sigma_t$, is added to the proper time t_p. This smearing is intended to represent the proper time resolution of the detector.

2. The new, smeared time, is then smeared again with a second Gaussian centered at zero, but having a lifetime-dependent[2] width $\sigma = \sigma_n t_p$. This second smearing is intended to account for unmeasured neutrinos in the decay of the B_s^0, which carry away a proportionally larger amount of momentum for long-lifetime B_s^0 than for short-lifetime B_s^0, regardless of the decay mode, thus adding more error to the determination of the decay length of longer-lifetime B_s^0. The final, "measured" time, is then t.

[2]In the rest frame of the B_s^0.

3. Next, it is decided whether or not the B_s^0 in an event mixes by comparing a number selected from a flat random distribution between 0 and 2 with the value of $1+\cos(\Delta m t_p)$ (the mixing comparison). If the first number is larger than the mixing comparison, the B_s^0 is tagged as mixed, and otherwise as unmixed.

4. Finally, the possibility of mistagging is considered, at a percentage rate[3] α. This is the percent of events for which a B_s^0 has been tagged as a $\bar{B_s^0}$, or vice versa. Clearly, in such cases mixing will not be correctly determined. In order to account for this mistagging, a number is selected from a flat random distribution between 0 and 1, and compared to α to determine whether the particle is mistagged, with numbers smaller than α indicating that this is the case. The tag of the particle is then adjusted accordingly.

5 Fitting the Mixing Distributions

Once the MC has been run, we fit the resulting distributions using the function described in Section 3. First, each set of generated events is separated by tag, and then the separated sets are each fit with function f (Equation 3),which has t as the independent variable and the following as parameters: Δm, σ_t, τ, `tag`, and α. RooFit [6] is used to calculate the complex error function.

The fitting code utilizes the MINUIT Minimization Package [7] to do an *unbinned likelihood fit*, which adjusts parameters of the distribution function f so as to maximize the probability that the sample lifetime distribution matches the function f. This is done by individually comparing each lifetime value t to the distribution function f to determine the probability of selecting t from f, and then multiplying these probabilities together to obtain the overall probability that the lifetime distribution matches the distribution function f. MINUIT then adjusts the fit parameter(s) in small steps, runs the test against f over again until the probability is maximized, and returns the final value of the fit parameter(s) along with the associated error.

By "fixing" all parameters but α during the fit, we have essentially fit for only the amplitude of the oscillation, hence the name *unbinned likelihood amplitude* fitting.

6 Calculating Sensitivity

After the dilution and its associated error have been determined in fitting, the sensitivity can be calculated. This is done by running the MC and fit several times, incrementing the fit value of Δm each time, and plotting the resulting errors in D to determine where it becomes too large for a measurement to be made. This entire process is done for each of a number of conditions, such as fluctuations in the value used for detector resolution, in order to compare the sensitivities. For each set of conditions tested, we calculate the sensitivity of a measurement of dilution ($D = 1 - 2\alpha$) to the conditions as follows:

1. The process of generating events and fitting for the dilution was done 20+ times (a "run"). Each time, the value of Δm used in the fit was incremented by 1 such that it

[3] The dilution $D = 1 - 2\alpha$.

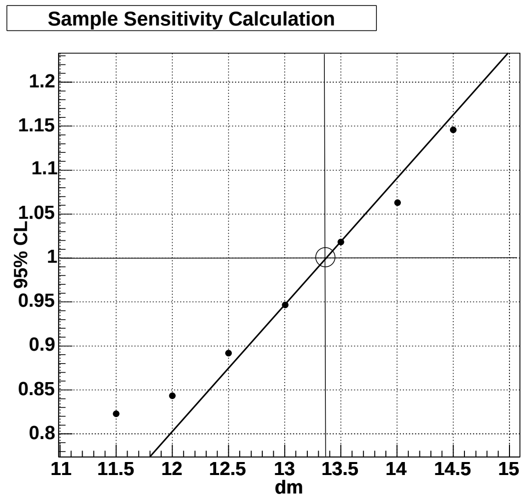

Figure 2: **A sample sensitivity calculation.** The sensitivity is defined to be the value of Δm for which the error in D is 1. *(Dots: Step 1 and 2; Line: Step 3, Circle: Step 4.)*

varied between 0 and 20 over the course of the run[4]. Δm was set to 1000 in the MC so that the sensitivity could be determined regardless of the "actual" value of Δm.

2. The *error* in the dilution D, as reported by MINUIT, was plotted against Δm. *(The dots in Figure 2.)*

3. A straight line crossing through the dilution error value just below 1, and that just above 1, was determined.*(The line in Figure 2.)*

4. The *Sensitivity* was then calculated to be the value of Δm for which this line crossed 1. *(The open circle in Figure 2.)*

7 Results

Using the Monte Carlo, fitting, and sensitivity calculation methods outlined above, we were able to test variations to the detector parameters to determine how they would affect error in a measurement of dilution, regardless of the value of Δm. Sensitivity of a dilution measurement was tested under the following conditions: Variations in proper time resolution σ_t, variations in dilution D, and variations in lifetime-dependent resolution σ_n. Each sensitivity value represents a run of 20 generate-and-fit iterations (4000 events each), with Δm varied between 0 and ~20 over the course of each run, as described in Section 6. Values used are intended to be similar to those found at the Fermilab DØ detector.

Table 1: **Sensitivity of D to variations in σ_t, 4000 events.** σ_t was set to 0.08, 0.10, and 0.12 in the MC, and then for each of these, the fit was done with σ_t fixed at the values listed in the Fit column. The calculated sensitivities are given in the table.

σ_t Fit $\downarrow$ \ MC $\rightarrow$	0.08	0.10	0.12
0.08	22.07	23.56	-
0.10	17.66	17.74	17.88
0.12	-	15.74	15.00

Table 2: **Sensitivity of D to variations in D, 4000 events.** D was set to 0.128, 0.160, and 0.192 in the MC, and then for each of these, a fit was done with D set to the initial values listed in the Fit column. The calculated sensitivities are given in the table.

D Fit $\downarrow$ \ MC $\rightarrow$	0.128	0.160	0.192
0.128	16.35	16.35	-
0.160	17.75	17.74	17.75
0.192	-	18.90	18.91

7.1 Variations in Proper Time Resolution, σ_t

The effect of $\pm$ 20% variations in σ_t on the sensitivity of D are enumerated in Table 1. Here, the dilution was set to 0.16 both in the MC and as an initial value in the fit, and time-dependent smearing was not used. σ_t was set to values of 0.08, 0.10, and 0.12 in the MC, and for each of these, fitting was done with σ_t values of 0.08, 0.10, and 0.12. The combinations for which the listed sensitivity is blank were not tested.

7.2 Variations in Dilution, D

The effect of $\pm$ 20% variations in the value of D in the MC vs. the initial value of D used in fitting, on the sensitivity of a measurement of D, are listed in Table 2. Here, σ_t was set to 0.10 both in the MC and in the fit, and time-dependent smearing was not used. D was set to values of 0.128, 0.160, and 0.192 in the MC, and for each of these, fitting was done with D at an initial values of 0.128, 0.160, and 0.192. The combinations for which the listed sensitivity is blank were not tested.

7.3 Variations in Lifetime-dependent Resolution, σ_n

Table 3 lists sensitivity of a measurement of D to $\pm$ 20% variations in σ_n vs. variations in accounting for σ_n in fitting[5]. Although the MC makes it easy to include the time-dependent smearing σ_n, this extra Gaussian is not so easily included into the fit function. Therefore,

[4]Values other than 0-20 were used when needed for reasonable precision.

[5]$\sigma_{\text{eff}} = \sqrt{\sigma_t^2 + (\sigma_n t)^2} = 0.14$ for $\sigma_t = 0.10$ and $t = 0.6$, as experimentally determined at DØ.

Table 3: **Sensitivity of D to variations in σ_n, 4000 events.** σ_n was set to 0.130, 0.163, and 0.196 in the MC, and for each of these, fits were done using various methods to account for $sigma_n$ since it couldn't be included directly in the fit. These methods are listed in the Fit column, and the resulting sensitivities are given in the table.

σ_n Fit $\downarrow$ \ MC $\rightarrow$	0.130	0.163	0.196
σ_n ignored	17.799	17.783	17.808
Avg.(σ_n, σ_t)	6.473	5.840	5.162
Eff. σ at t	-	13.386	-

we chose three methods of accounting for this extra factor in the fitting.

In the first row of Table 3, the time-dependent smearing is ignored altogether and $\sigma_{\text{eff}} = \sigma_t$. In the second row, the lifetime values from the MC were histogrammed, and then a bin-weighted average was used to determine an average effective value of σ to use in the fit. In the last row, the actual effective smearing at each lifetime t was computed during fitting, using $\sigma_n = 0.163$. Variations were not tested. This is the most realistic scenario for which we have calculated sensitivity. In all cases, σ_t was set to 0.10 in the MC and fit, and D was set to 0.160 both in the MC and for the initial fit value.

8 Conclusion

The most realistic value we have determined for the sensitivity of a measurement of dilution using an unbinned likelihood amplitude fitting method is $\Delta m < 13.386 ps^{-1}$. This means that with the parameters and method we have used, even in such ideal situations as the toy MC simulates, a measurement of dilution would not be possible for $\Delta m \geq 13.386 ps^{-1}$, because the error would be too high. Since current experiments have set the lower limit of Δm above this value, at $\Delta m \geq 14.6 ps^{-1}$, this measurement may not be possible using this method. However, using other fitting methods or combining multiple methods, we may be able to reduce this error and still make a measurement of dilution, and eventually, of Δm.

References

[1] K. Anikeev *et al.*, "B-Physics at the Tevatron: Run II and Beyond" (2001) hep-ph/0201071.

[2] Griffiths, David. *Introduction to Elementary Particles,* John Wiley & Sons, 1987.

[3] Particle Data Group. *Particle Physics Booklet,* 2004.

[4] The Root Data Analysis Framework, http://root.cern.ch/.

[5] CERNLib C335: Complex Error Function, http://wwwasdoc.web.cern.ch/wwwasdoc/shortwrupsdir/c335/top.html.

[6] The RooFit Toolkit for Data Modeling *(Contains a direct C++ translation of CWERF for use with Root)*, http://roofit.sourceforge.net/.

[7] CERNLib (PackLib) Long Writeup D506: MINUIT Minimization Package, http://wwwasdoc.web.cern.ch/wwwasdoc/WWW/minuit/minmain/minmain.html.