

Observation of $B_s^0 - \bar{B}_s^0$ oscillations

A. Belloni^a On behalf of the CDF Collaboration

^aMassachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139-4307, USA

We report the observation of $B_s^0 - \bar{B}_s^0$ oscillations using 1 fb^{-1} of data from $p\bar{p}$ collisions with the CDF II detector at the Fermilab Tevatron. We measure the probability, as a function of proper decay time, that the B_s^0 decays with the same, or opposite, flavor as its flavor at production, which is determined using opposite-side and same-side flavor identification methods. We find a signal consistent with $B_s^0 - \bar{B}_s^0$ oscillations, with a significance greater than 5σ . We measure $\Delta m_s = 17.77 \pm 0.10(\text{stat}) \pm 0.07(\text{syst}) \text{ ps}^{-1}$.

1. INTRODUCTION

The determination of the $B_s^0 - \bar{B}_s^0$ oscillation frequency Δm_s from a time-dependent measurement of $B_s^0 - \bar{B}_s^0$ oscillations has been a major objective of experimental particle physics [1]. This frequency can be used to extract the magnitude of V_{ts} , one of the nine elements of the Cabibbo-Kobayashi-Maskawa (CKM) matrix [2], and to constrain contributions from new physics. Recently, we reported [3] the strongest evidence to date of the direct observation of $B_s^0 - \bar{B}_s^0$ oscillations. That analysis used 1 fb^{-1} of data collected with the CDF II detector [4] at the Fermilab Tevatron.

The following sections present the improved analysis which achieved the definitive observation of $B_s^0 - \bar{B}_s^0$ oscillations.

2. B MIXING THEORY

The Standard Model describes neutral B meson mixing in terms of a unitary, non-diagonal quark mixing matrix. The eigenstates of the Hamiltonian, with definite mass and lifetime, are different from the eigenstates of the strong interaction. B_s^0 and B^0 mesons thus can mix with \bar{B}_s^0 and \bar{B}^0 , respectively, via box diagrams such as the one depicted in Fig. 1, where the dominant contribution in the loop comes from the top quark.

Neglecting effects from CP violation and the possible lifetime difference between the heavy and

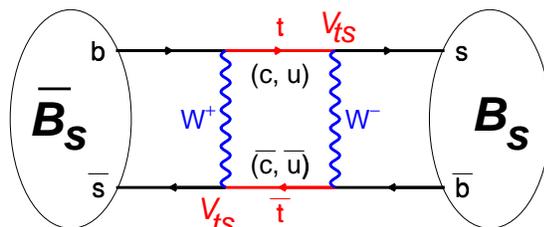


Figure 1. One of the box diagrams which describe B_s^0 mesons mixing. In the case of B^0 mixing, the corresponding diagram is obtained by substituting the $s(\bar{s})$ quarks with a $d(\bar{d})$ quark, respectively.

light mass eigenstates, the probability density function \mathcal{P} for a B_s^0 meson produced at time $t = 0$ to decay as a B_s^0 (\bar{B}_s^0) at time $t > 0$ is:

$$\mathcal{P}_{\pm}(t) = \frac{\Gamma}{2} e^{-\Gamma t} [1 \pm \cos(\Delta m_s t)], \quad (1)$$

where the subscript “+” (“-”) indicates that the meson decays as B_s^0 (\bar{B}_s^0).

The oscillation frequency, Δm_s , is directly related to elements of the Cabibbo-Kobayashi-Maskawa matrix which describes quark mixing. The following relations hold:

$$\Delta m_q \propto m_{B_q} \hat{B}_{B_q} f_{B_q}^2 |V_{tb} V_{tq}^*|^2, \quad (2)$$

where q can be either a d or an s quark. The parameter m_{B_q} is the mass of the B meson of interest. Theoretical uncertainties in the bag factor \hat{B} and the form factor f do not allow for the di-

rect measurement of $|V_{tb}V_{tq}^*|$ with a precision better than 10% using a single Δm_q measurement. However, these uncertainties are strongly reduced in the ratio $\Delta m_d/\Delta m_s$, which results equal to:

$$\frac{\Delta m_s}{\Delta m_d} = \frac{m_{B_s}}{m_{B_d}} \xi^2 \frac{|V_{ts}|^2}{|V_{td}|^2}. \quad (3)$$

The latest lattice QCD calculation gives ξ with a 3 to 4% uncertainty [5]. Since Δm_d has already been measured with a precision of 1% [6], the measure of Δm_s provides a tight constraint on the CKM matrix.

The CDF collaboration already reported a very precise measurement of Δm_s [3], observing $B_s^0 - \bar{B}_s^0$ oscillations with a 3σ significance. In this report we focus on the updated analysis which achieved a significance greater than 5σ , which is conventionally the threshold required to claim definitive observation.

3. DATA SAMPLE

The analysis uses 1 fb^{-1} of data collected by the CDF II detector between March 2002 and February 2006. B_s^0 mesons are reconstructed in fully hadronic $D_s^- \pi^+ (\pi^- \pi^+)$ and semileptonic $\ell^+ D_s^- X$ decays, where the D_s^- candidate is reconstructed in the $\phi^0 \pi^-$, $K^{*0} K^-$ and $\pi^+ \pi^- \pi^-$ final states. In addition, the $D_s^- \pi^+$ and $D_s^- \rho^+$ decays, with $D_s^- \rightarrow \phi^0 \pi^-$, are reconstructed. In the last two modes, the complete decay chains include $D_s^{*-} \rightarrow D_s^- \gamma / \pi^0$ and $\rho^+ \rightarrow \pi^+ \pi^0$, in which the neutral particles are not reconstructed. Although not fully reconstructed, the fraction of momentum carried by the lost particle is very small, in the order of 4%. Throughout the paper, unless otherwise specified, references to a particular process imply the inclusion of the charge conjugate one. An Artificial Neural Network (ANN) performs the selection of B_s^0 candidates for the analysis. Particle identification is directly utilized to suppress B^0 background in both hadronic and semileptonic samples. Particle identification also helps to reduce the combinatorial background, which is dominated by pions, when reconstructing signals that contain kaons in the final state. The signal yield for the hadronic modes, including partially reconstructed decays, is 8,700, while

61,500 candidates are reconstructed in semileptonic decays. The previous analysis utilized 3,600 candidates reconstructed in fully hadronic decays and 37,000 in semileptonic decays. The mass distribution of $B_s^0 \rightarrow D_s^- (\phi^0 \pi^-) \pi^+$ candidates is shown in Fig. 2; the contribution of partially reconstructed B decays is evident in the peaks which populate the lower sideband region. Figure 3 contains the mass distribution of $\ell^+ D_s^- (K^{*0} K^-)$ combinations and, in the inset, of the corresponding $D_s^- \rightarrow K^{*0} K^-$ candidates. The $m(\ell D)$ distribution is very broad, since the not-reconstructed particles of the complete B_s^0 candidate can carry a large amount of momentum.

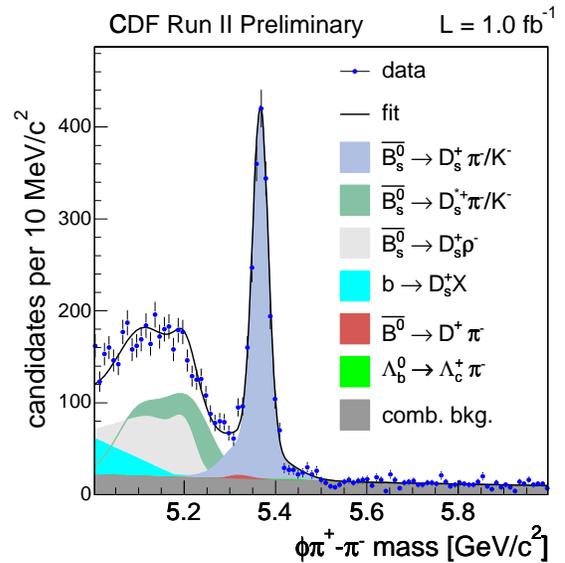


Figure 2. Invariant mass distribution for $B_s^0 \rightarrow D_s^- \pi^+$, $D_s^- \rightarrow \phi^0 \pi^-$, $\phi^0 \rightarrow K^+ K^-$ candidates.

4. PROPER DECAY TIME

The decay time in the B_s^0 rest frame is obtained as follows:

$$t = L_{xy}(B_s^0) \frac{m_{B_s^0}}{p_T(B_s^0)}, \quad (4)$$

where L_{xy} is the displacement of the B_s^0 decay point with respect to the primary vertex pro-

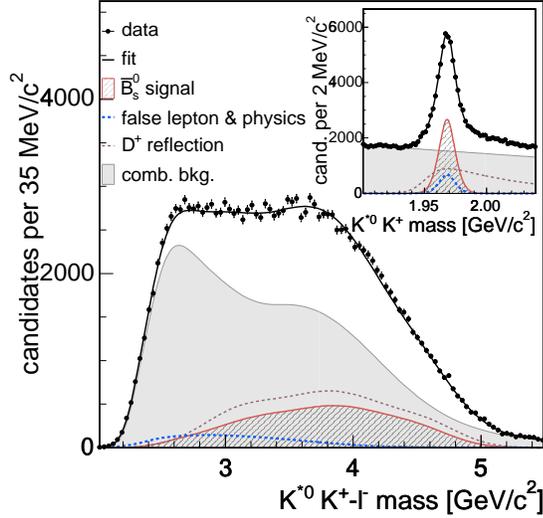


Figure 3. $\ell^+ D_s^- (K^{*0} K^-)$ and D_s^- invariant mass distributions for semileptonic decays.

jected onto the B_s^0 transverse momentum vector p_T . In the case of partially reconstructed and semileptonic B_s^0 decays, where the B_s^0 candidate is not fully reconstructed, a correction factor k has to be included to account for the missing momentum. The expression becomes:

$$t = L_{xy}^{recon} \frac{M_{B_s^0}}{p_T^{recon}} k, \quad k \equiv \frac{L_{xy}(B_s^0)}{L_{xy}^{recon}} \frac{p_T^{recon}}{p_T(B_s^0)}, \quad (5)$$

where L_{xy}^{recon} and p_T^{recon} are the projected displacement and the transverse momentum of the reconstructed decay products.

The k -factor distribution is obtained from Monte Carlo (MC) simulation. The distributions of k -factor for partially reconstructed hadronic B_s^0 decays peak closely to unity and is very narrow, as a consequence of the softness of the lost particle in the decay chain. In the case of semileptonic decays, the distributions vary as a function of the reconstructed ℓD mass, as shown in Fig. 4. In Fig. 5 the proper decay time distribution of $B_s^0 \rightarrow D_s^- (\phi^0 \pi^-) \pi^+$ candidates, both fully and partially reconstructed, is shown.

The determination of the proper decay time resolution is a critical part of the analysis, since it affects dramatically the sensitivity for observing a signal. The most precise determinations

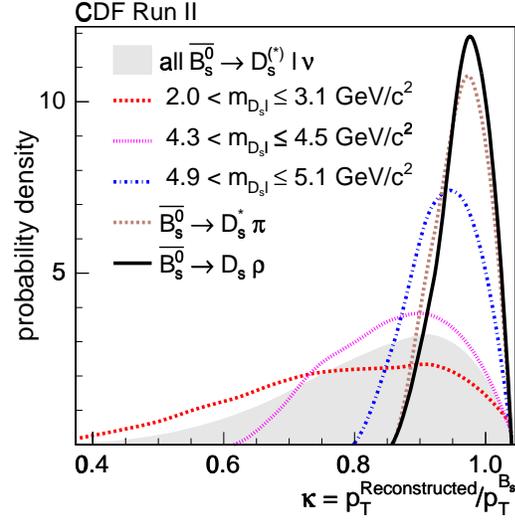


Figure 4. k -factor distribution for several ℓD mass regions ($B_s^0 \rightarrow \ell^+ D_s^- X$, $D_s^- \rightarrow \phi^0 \pi^-$ decays) and for partially reconstructed hadronic decays.

of σ_{ct} come from lifetime measurements in exclusively reconstructed modes with a J/ψ in the final state, where the presence of prompt candidates allows for the direct calibration of the proper decay time resolution. The same direct measurement is not possible in the samples used in this analysis, which are collected with triggers that bias the lifetime distribution. A calibration sample of pseudo- B_s^0 candidates is obtained by associating a track, which is prompt in most of the cases, to a charged D candidate. It is thus possible to measure and calibrate t/σ_t directly on data using the produced set of $D_s^- + track$ candidates, which have the same topology of real signal candidates. Finally, an event-by-event correction depending on the decay topology and the kinematic of the candidate is applied.

5. FLAVOR TAGGING

While the flavor of the B_s^0 candidate at the decay point is unambiguously defined by the charges of its daughter tracks, the flavor at production can be inferred, with a certain degree of uncertainty, by flavor tagging algorithms.

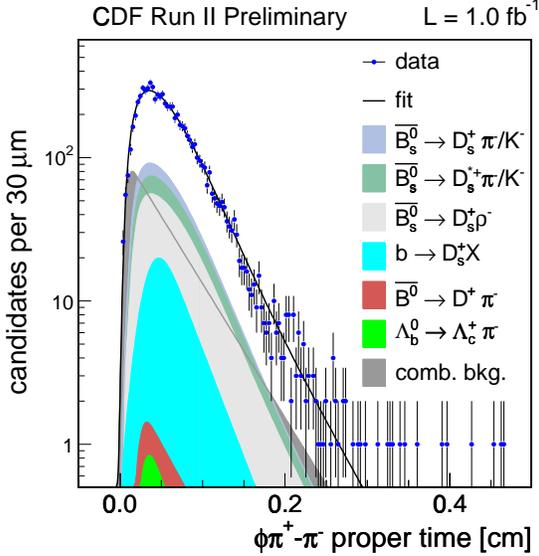


Figure 5. Proper decay time distribution of $B_s^0 \rightarrow D_s^-(\phi^0 \pi^-) \pi^+$ candidates.

Two types of flavor tagging algorithms are utilized at CDF: opposite-side (OS) and same-side (SS) flavor taggers. The performance of a tagger is quantified by its efficiency ϵ , the fraction of candidates to which a tag is assigned, and dilution \mathcal{D} , defined as $1 - 2\mathcal{P}_w$, where \mathcal{P}_w is the probability that the assigned tag is incorrect. The sensitivity for observing a signal is proportional to $\sqrt{\epsilon \mathcal{D}^2}$. The decisions of the SS and OS taggers are combined treating the two taggers as independent.

Table 1

Performance of flavor taggers used at CDF in the hadronic and semileptonic samples. Statistical and systematic errors are added in quadrature.

$\epsilon \mathcal{D}^2$	Semileptonic	Hadronic
OST	1.8 ± 0.1 %	1.8 ± 0.1 %
SST	4.8 ± 1.2 %	3.4 ± 0.9 %

5.1. Opposite-side Flavor Tagging

Opposite-side taggers exploit the fact that at hadron colliders b quarks are mostly produced in

$b\bar{b}$ pairs. Therefore, the flavor of the b quark in the opposite side with respect to the reconstructed candidate is correlated to the flavor at production of the B_s^0 meson of interest.

The soft lepton tagger (SLT) is based on $b \rightarrow \ell^- X$ semileptonic decays; the charge of the lepton, either a muon or an electron, is correlated to the charge of the decaying B meson.

The jet charge tagger (JQT) utilizes the correlation of the charge of a b -jet to the charge of the originating b quark. Different methods of heavy flavor jet tagging algorithms, track displacement, χ^2 probability and maximum transverse momentum, are used and combined in an ANN.

The opposite-side kaon tagger (OSKT) is based on cascade decays $b \rightarrow c \rightarrow s$ and exploits the correlation between the charges of the $K^{+(-)}$ and $B_s^0(\bar{B}_s^0)$ mesons.

An ANN finally combines the pieces of information provided by the three taggers. The performance of opposite-side taggers is independent of the type of B meson produced on the signal side. It is thus possible to exploit high statistic B^+ and B^0 samples to calibrate opposite-side taggers.

The figure-of-merit of the tagger, in the different data samples, is reported in Tab. 1.

5.2. Same-side Flavor Tagging

The same-side (kaon) tagger (SST) is based on the correlation between the b flavor and the charge of the particles produced in association with the B_s^0 candidate during the fragmentation process of the b quark. When a $B_s^0(\bar{B}_s^0)$ meson is formed, a $\bar{s}(s)$ quark is left at the end of the fragmentation chain and may form a $K^+(K^-)$. Thus, if a charged particle is found close to the B_s^0 meson and identified as a kaon, it is likely to be the leading fragmentation track, the charge of which is correlated to the charge of the b quark contained in the reconstructed candidate, at the time of its production. An ANN is utilized to separate kaons from other particles (mainly pions) and select the track which is the most likely to be the leading fragmentation particle. The inputs to the neural network are a combined particle identification likelihood based on information from the Time-of-Flight system and dE/dx , which provides most of the separating power, and

kinematic variables of the event.

Since the fragmentation process differs among B^+ , B^0 and B_s^0 mesons, it is necessary to rely on MC simulation to measure the performance of the same-side tagger. CDF performed extensive data and MC comparisons, finding reasonably good agreement in all the variables utilized by the neural network. The effectiveness of the tagger in data and Monte Carlo has been compared for different B^+ and B^0 decay modes, showing very good agreement.

Table 1 reports a summary of the performance of the SST in the hadronic and semileptonic data samples, separately.

6. MIXING ANALYSIS

An unbinned maximum likelihood fit framework is utilized to search for B_s^0 oscillations. The likelihood combines mass, proper decay time, proper decay time resolution and flavor tagging information for each candidate. Separate probability density functions are used to describe signal and each type of background. The amplitude scan method [7] was used to search for oscillations. The likelihood term describing the proper decay time of tagged neutral B meson candidates in Eq. 1 is modified by introducing the amplitude \mathcal{A} :

$$\mathcal{L}_{\text{signal}} \propto 1 \pm \mathcal{A} \mathcal{D} \cos(\Delta m t). \quad (6)$$

Then, a scan in Δm is performed by fitting \mathcal{A} for fixed values of Δm . The dilution \mathcal{D} is fixed to the value obtained by the calibration process. In the case of infinite statistic and perfect resolution, it is expected to find $\mathcal{A} = 1$ for the true value of Δm and $\mathcal{A} = 0$ otherwise. In practice, the procedure consists in recording $(\mathcal{A}, \sigma_{\mathcal{A}})$ for each Δm hypothesis. A particular value for Δm is excluded at 95% confidence level if $\mathcal{A} + 1.645\sigma_{\mathcal{A}} < 1$ holds. The sensitivity of a mixing measurement is defined as the Δm value for which $1.645\sigma_{\mathcal{A}} = 1$. The statistical uncertainty in \mathcal{A} , $\sigma_{\mathcal{A}}$, is described to a good approximation by:

$$\frac{1}{\sigma_{\mathcal{A}}} = \sqrt{\frac{S}{S+B}} \cdot e^{-\frac{\Delta m_s^2 \sigma_{\Delta t}^2}{2}} \cdot \sqrt{S \frac{\epsilon D^2}{2}}, \quad (7)$$

where S and B indicates, respectively, the number of signal and background candidates. Eq. 7

clearly shows the importance of understanding the proper decay time resolution, especially for high values of Δm_s , and how the effective statistical power scales with the performance of flavor taggers.

7. RESULTS

The combined amplitude scan is shown in Fig. 6. The sensitivity is 31.3 ps^{-1} . The value of the amplitude is consistent with unity and incompatible with zero around $\Delta m_s = 17.75 \text{ ps}^{-1}$, where $\mathcal{A} = 1.21 \pm 0.20$. For all the other Δm_s values, the amplitude is always consistent with zero. To assess the significance of this deviation, the negative logarithm of the ratio of the likelihood functions for $\mathcal{A} = 1$ (mixing hypothesis) and $\mathcal{A} = 0$ (no-mixing hypothesis) was utilized. Its distribution is shown in Fig. 7. The minimum likelihood ratio Λ is at $\Delta m_s = 17.77 \text{ ps}^{-1}$ and has a value of -17.26 . The significance of the signal in the amplitude is the probability that randomly tagged data would produce a value of Λ lower than -17.26 at any value of Δm_s . The probability is 8×10^{-8} , since only 28 likelihood scans produced $\Lambda < -17.26$ out of the 350 millions generated, well below 5.7×10^{-7} which corresponds to 5σ . The fit for Δm_s , with \mathcal{A} fixed to unity, finds $\Delta m_s = 17.77 \pm 0.10(\text{stat}) \pm 0.07(\text{syst}) \text{ ps}^{-1}$. The dominant contributions to the systematic error come from uncertainties on the absolute scale of the decay-time measurement.

The $B_s^0 - \bar{B}_s^0$ oscillations are depicted in Fig. 8. Candidates in the hadronic sample are collected in five bins of proper decay time modulo $2\pi/\Delta m_s$. In each bin, a fit for \mathcal{A} is performed and the result plotted. The curve corresponds to a cosine wave with amplitude equal to 1.28, which is the fitted value in the hadronic sample. Data are well represented by the curve.

Using the measured value of Δm_s , together with the world average values of $\Delta m_d = 0.507 \pm 0.005$ [6], $m_{B^0}/m_{B_s^0} = 0.98390$ [8], with negligible uncertainty, and $\xi = 1.21^{+0.047}_{-0.035}$ [5], we extract $|V_{td}/V_{ts}| = 0.2060 \pm 0.0007(\text{exp})^{+0.0081}_{-0.0060}(\text{theor})$, where the first error refers only to the contribution of the Δm_s uncertainty, while among all other external errors the theoretical one largely

dominates.

In conclusion, we report the first observation of $B_s^0 - \bar{B}_s^0$ oscillations from a decay-time-dependent measurement of Δm_s . The signal significance exceeds 5σ , thus qualifying as a definitive observation. The uncertainty in lattice calculations makes further improvement in the precision of the Δm_s measurement unnecessary.

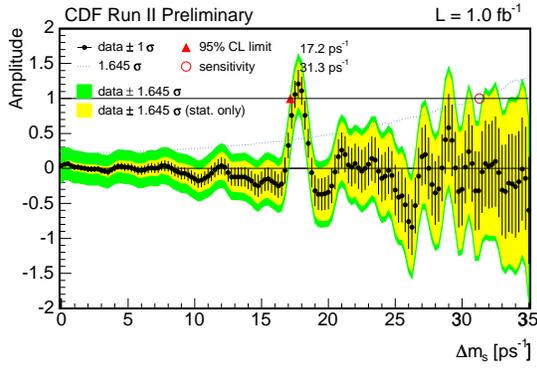


Figure 6. Δm_s amplitude scan. The dotted line represents $1.645\sigma_A$ and indicates a sensitivity of 31.3 ps^{-1} .

REFERENCES

1. C. Gay, *Annu. Rev. Nucl. Part. Sci.* **50** (2000), 577-641.
2. N. Cabibbo, *Phys. Rev. Lett.* **10** (1963), 531; M. Kobayashi and T. Maskawa, *Prog. Theor. Phys.* **49** (1973), 652.
3. A. Abulencia *et al.*, *Phys. Rev. Lett.* **97** (2006), 062003.
4. D. Acosta *et al.* (CDF Collaboration), *Phys. Rev.* **D71** (2005), 032001.
5. M. Okamoto, *Proc. Sci. LAT2005* (2005), 013.
6. W.-M. Yao *et al.*, *J. Phys.* **G33** (2006), 1.
7. H. G. Moser, A. Roussarie, *Nucl. Instr. and Meth.* **A384** (1997), 491-505.
8. D. Acosta *et al.*, *Phys. Rev. Lett.* **96** (2006), 202001.

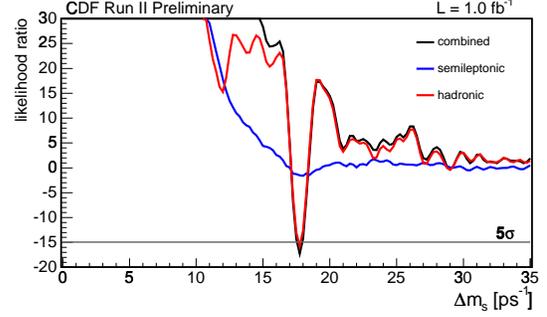


Figure 7. The logarithm of the ratio of likelihoods for amplitude equal to one and amplitude equal to zero. The horizontal line indicates $\Lambda = -15$, which corresponds to a probability of 5σ in the case of randomly tagged data.

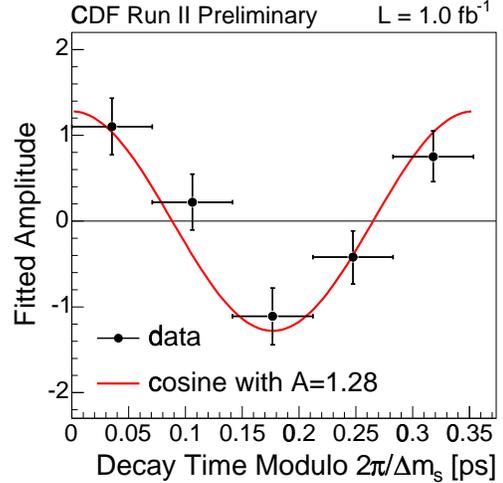


Figure 8. The $B_s^0 - \bar{B}_s^0$ oscillations signal, in the hadronic sample, measured in bins of proper decay time modulo the measured oscillation period $2\pi/\Delta m_s$. The figure is described in the text.