SLAC–PUB–5644 Rev.
April 1993
(E)

Importance of the $B_s^0 \to \mu^+\mu^-$ Branching Fraction (*)

A. Fridman (**) and A. Snyder
Stanford Linear Accelerator Center, Stanford University, Stanford, California 94309

Summary.—We discuss the $B_s^0 \to \mu^+\mu^-$ decay channel, which is sensitive to deviations from the standard model. We also consider the B_s^0 mixing information that could be obtained from these decays in the framework of the standard model. Finally, we comment on the experimental difficulties of observing the $B_s^0 \to \mu^+\mu^-$ channel.

PACS13.20.–v –Leptonic and semileptonic decays of mesons.
PACS13.20.Jf – Other meson decays.

1. – Introduction

To illustrate the usefulness of large $B\bar{B}$ data samples, we briefly discuss problems related to the measurement of rare B-decays, specifically, the $B_s^0 \to l^+l^-$ channels where $l \equiv \tau, \mu, e$. These decays are expected to have small branching ratios (see below), but to be sensitive to deviations from the standard model. Moreover, these measurements could give additional information about B_s^0 mixing that could complete our knowledge of the mixing observed for the B_d^0 case with the $e^+e^- \to \Upsilon(4S) \to B_d^0\bar{B}_d^0$ experiments [1].

Submitted to Il Nuovo Cimento

(*) Work supported by Department of Energy contract DE–AC03–76SF00515.
(*) LPHHE, Université de Paris VI et VII.

Let us recall that in the standard model the B_s^0 mixing parameter [2,3] x_s is expected to be much larger than x_d, the B_d^0 case ($x_s/x_d \sim 14$). Although the measurement of r_s,

$$(1) \qquad r_s \equiv [N(B_s^0 \to \bar{B}_s^0)]/[N(B_s^0 \to B_s^0)] = [x_s^2/(2+x_s^2)] \ ,$$

appears to be possible [4] (N is the number of events), the extraction of the x_s value becomes difficult if x_s is large [5]. The distribution of the time-dependent rate could, in principle, also allow the x_s value to be obtained [2] if the detector can measure the time-dependence with good precision [4]. Because this measurement will be difficult, an indirect estimate of x_s via branching ratio measurements might be very helpful (see below). In any case, the comparison of x_s measurements using different methods will be a further test of the standard model.

2. – The $B_s^0 \to l^+l^-$ decay

These decays ($l \equiv \tau, \mu, e$) have been investigated theoretically [6-9]. Estimates of the branching ratios [8] are given in Table I for a top quark mass, $m_t = 150$ GeV (here m_k is the mass of the k-type quark or particle). The $B_s^0 \to \tau^+\tau^-$ decay has the largest branching ratio, but will not be easily detected (the momentum of the τ must be measured in order to know if they are coming from the same B_s^0 meson). It may eventually be possible to detect the $B_s^0 \to \mu^+\mu^-$ decay in experiments where very large statistics could be obtained. For this decay the theoretical expression for the branching ratios [8,9] is:

$$(2) \quad \mathrm{BR}(B_s^0 \to \mu^+\mu^-) = \\ = (G_F^2/2\pi) \ (\alpha^2/16\pi^2) \ f_B^2 \, m_B \left(|V_{tb}V_{ts}|^2/\Gamma_s\right) \ m_\mu^2 \, |C(y)|^2 \, .$$

Here G_F is the Fermi coupling constant and α is the fine structure constant. The f_B (usually assumed to be ~ 0.2 GeV) is the B_s^0 decay constant and Γ_s is the B_s^0 decay width. The V_{jk} are the CKM matrix elements and $C(y)$ is a function of

$$y = f\, m_t^2/m_W^2 \quad ,$$

given in [8]. Figure 1 shows the dependence of $\mathrm{BR}(B_s^0 \to l^+l)$ as a function of the top quark mass. We see that the branching ratio increases rapidly with increasing m_t. Notice also from (2) that the $B_d^0 \to \mu^+\mu^-$ branching ratio will be much smaller than that expected in the B_s^0. Their ratio should be of the order of $|V_{td}/V_{ts}|^2 \sim 0.06$.

Equation (2) contains several quantities that also appear in the parameter x_s describing the B_s^0 mixing [10]:

$$(3) \quad x_s \equiv \Delta M/\Gamma_s = \\ = (G_F^2/2\pi) \ B_B \ f_B^2 \, m_B \left(|V_{tb}V_{ts}|^2/\Gamma_s\right) \ m_t^2 \, |F(y)|^2 \, \eta \, .$$

Notice that B_B is the bag parameter (~ 1) corresponding to the B_s^0 meson, while the QCD correction factor is $\eta \sim 0.85$ [11]. In this equation, the mass of the top quark is not neglected with respect to the W mass. Then in the standard model, x_s will depend on m_t via the $m_t \times F(y)$ factor where [10]

$$4 \quad F(y) = (1/4)\left(1 + \left\{[(3-9y)/(y-1)^2] + [6y^2 \, ln(y)]/(y-1)^3 \right\}\right) \, .$$

Figure 1 also presents the m_t dependence of x_s. It is much weaker than the m_t dependence of the branching ratio. Using (2) and (3), we obtain

$$(5)\quad \mathrm{BR}(B_s^0 \to \mu^+\mu^-) = \\ = x_s \left[(3\alpha^2/16\pi) \times (m_\mu^2/B_B\,\eta) \times \{\, |C(y)|^2 \,/\, [m_t^2\, F(y)\,]\,\} \right] .$$

Using this approach, x_s can be obtained by measuring the branching ratio. Moreover, an increase of x_s increases the branching ratio and facilitates its measurement. This contrasts with r_s and the time-dependence rate measurements discussed above. Note that so far only upper limits on $\mathrm{BR}(B^0 \to \mu^+\mu^-)$ have been measured [16] (3.2×10^{-6}).

By using (4) with known values of m_t and B_B, we obtain the value of $r_s = x_s^2/(2 + x_s^2)$, where the error on r_s will be small if x_s is large. For instance, with 100 $B_s^0, \bar{B}_s^0 \to \mu^+\mu^-$ events, we find $\Delta r_s/r_s \leq 4\ 10^{-3}$ for $x_s \geq 10$. Thus the comparison of the r_s values obtained from (1) and (4) could verify the standard model, and eventually may be used to obtain x_s.

3. – Experimental comments

In the case that $\mathrm{BR}(B_s^0 \to \mu^+\mu^-) \sim 10^{-9}$ (Table I), this decay channel could not be studied with e^+e^- colliders [4,12]. For pp interactions with the LHC (c.m. energy of $\sqrt{s} = 16$ TeV) or SSC ($\sqrt{s} = 40$ TeV) colliders, the production of events with B mesons is much larger [12,13]. Table II indicates the number of $b\bar{b}X$ events (X meaning anything) expected in one year ($10^7 s$) of running with a luminosity of $L = 10^{32}$ cm^{-2} s^{-1}. This is the luminosity usually proposed for studying B-physics, in order to decrease multiple interactions per bunch crossing [13]. Since tagging is not needed for the measurement of the discussed branching ratio, we envision a larger luminosity. For instance, with $L = 10^{33}$ cm^{-2} s^{-1}, we obtain at the SSC (LHC) $\sim 1.4 \times 10^3$ ($\sim 6 \times 10^2$) $B_s^0, \bar{B}_s^0 \rightarrow \mu^+\mu^-$ events in one year of running (corresponding, however, to a number of interactions per second of $N_{int}/s \sim 10^8$).

Let us roughly discuss the properties of a μ-detector necessary to obtain a signal-to-background ratio of S/B$\sim$ 1. Due to the large multiplicity in pp interactions, we expect a large background in the data sample used to study the $B_s^0 \rightarrow \mu^+\mu^-$ channel. The background can initially be decreased with kinematical cuts by using only tracks (charged particles) with large momentum ($p > 5$ GeV/c) and transverse momentum ($p_T > 1$ GeV/c). The momentum cut will decrease the contribution of μ's due to the π and K decays. The number of charged particles pairs having an effective mass corresponding to B_s^0 will be reduced by the p_T cut [12]. At $\sqrt{s} = 40$ TeV (and using tracks with $p_T > 1$ GeV/c), we show in Fig. 2 the efficiency (ϵ) for detecting the $B_s^0 \rightarrow \mu^+\mu^-$ channel as a function of the maximum value of θ due to the size of the chosen detector. Here θ is the emission angle of the charged tracks with respect to the beam line. A minimum value of $\theta > 1^\circ$ is used in the Monte Carlo calculation [14] because of the beam-pipe cut. Figure 2 also indicates the influence of the p limits ($p > 0, 5, 10, 20$ GeV/c).

As an example, we present in Table III the ϵ, S/B, and number of $B_s^0 \rightarrow \mu^+\mu^-$ events expected in one year of running, using kinematical cuts. These values were obtained with a Monte Carlo calculation at

$\sqrt{s} = 40$ TeV. We assume that the $B_s^0 \equiv \bar{b}q$ production rate is taken to be 0.15 with respect to all particles containing a $\bar{b}$ quark. We also consider that the $B_s^0 \rightarrow \mu^+\mu^-$ can be detected after a decay length corresponding to the average B_s^0 lifetime. From Table III we see that a background rejection factor of at least $\sim 10^8$ is needed to achieve a workable signal-to-noise (S/B$\sim$ 1). Here, only background due to $B\bar{B}$ events is considered. Thus the detector must have μ identification and B_s^0 vertex reconstruction adequate to obtain this level of background rejection [15]. In any case, detailed investigations have to be done to evaluate the possibility of measuring the $B_s^0 \rightarrow \mu^+\mu^-$. We must also take into account other sources of background, for instance, Drell-Yan μ pair production, as well as $B \rightarrow \mu X, \pi^+\pi^-$ decays and non-$B\overline{B}$ events.

4. – Conclusions

We have discussed the $B_s^0 \rightarrow \mu^+\mu^-$ channel, which is sensitive to deviations from the standard model. In addition, information about B_s^0 mixing was obtained in the framework of the standard model. Using this approach, large values of the x_s mixing parameter were measured. We also discussed the difficulties of measuring the $B_s^0 \rightarrow \mu^+\mu^-$ having theoretical predictions of $\sim 10^{-9}$ for its branching ratio. Clearly, detailed investigations of the various backgrounds have to be made to evaluate the possibility of identifying $B_s^0 \rightarrow \mu^+\mu^-$ events

* * *

It is a pleasure to thank A. Ali, L. Dixon and M. Peskin for useful discussions.

REFERENCES

[1] H. Albrecht et al. (ARGUS): Phys. Lett. **B192**, 245 (1987); M. Artuso et al. (CLEO), Phys. Rev. Lett. **62**, 2233 (1989) .

[2] I.I. Bigi and A.I. Sanda: Nucl. Phys. **B193**, 85 (1981); Nucl. Phys. **B281**, 41 (1987).

[3] The mixing parameter $x_{d,s} = \Delta M_{d,s}/\Gamma_{d,s}$ for the $B^0_{d,s}$ has the inconvenience of having a value between 0 and ∞. Here $\Delta M_{d,s}$ is the mass difference between the heavy and light neutral mesons while $\Gamma_{d,s}$ is their average width; see also the lectures: Mixing and CP Violation in the B-System, A. Fridman: Cern Report CERN–EP/88–123 (1988).

[4] The Physics Program of a High-Luminosity Asymmetric B Factory, SLAC Report 353 (1989).

[5] For instance, the measurement of the small value of r_d ($\sim$ 0.2) allowed one to obtain $x_d = 0.66 \pm 0.11$ [1].

[6] B.A. Campbell and P.J. O'Donnel: Phys. Rev. **25D**, 1989 (1982).

[7] G. Eilam and A. Soni: Phys. Rev. Lett. **215B**, 171 (1988).

[8] A. Ali: DESY Report 91–080 (1991).

[9] M.J. Savage: Report RU–91–22 (1991)

[10] Report CLNS 91–1043 (1991) and references therein.

[11] A.J. Buras, W. Slominski, H. Steger: Nucl. Phys. **B238**, 529 (1984); Nucl. Phys. **B245**, 369 (1984).

[12] A. Fridman and A. Snyder: Proceedings of the Large Hadron Collider Workshop, Aachen, 4-9 October 1990, CERN Report CERN 90–10, Vol. II, 218 (1990).

[13] A. Fridman, *B*-Physics Possibilities with *pp* Colliders: Invited talk to be published in the Proceedings of the 3rd Topical Seminar on Heavy Flavor, San Miniato, 17-21 June 1991, and references therein.

[14] H.-U. Bengtsson and T. Sjöstrand: Comp. Phys. Commun. **46** 43 (1987).

[15] Vertex reconstructions for $B \rightarrow \mu\mu$ are under investigation, see for instance, C.S. Mishra, the E–772 and 789 Collaborations : Dilepton and Dihadron Production in Proton-Nucleus Collisions at 800 GeV, FERMILAB–Conf–90/100–E (1990).

[16] L. Pondrom, Proceedings of the 25th International Conference on High Energy Physics, Kent Ridge, Singapore, August 2–8, 1990, and references therein.

TABLE I. *Estimates of* $B_s^0 \to l^+l^-$ *branching ratios;* $m_t = 150$ *GeV,* $f_B = 200$ *MeV [8]. The B meson decay constant is* f_B.

$\tau^+\tau^-$	1.8×10^{-7}
$\mu^+\mu^-$	8.3×10^{-10}
e^+e^-	2.0×10^{-14}

TABLE II. *Comparison of the total (σ_T) and $b\bar{b}X$ [$\sigma(b\bar{b})$] cross section of the LHC and SSC colliders at a luminosity L of* 10^{32} cm^{-2} s^{-1} *[12,13]. The number of $b\bar{b}X$ interactions per year ($N(b\bar{b})/10^7$ s) and the interactions per second (N_{int}/s) are also given.*

	LHC (16 TeV)	SSC (40 TeV)
σ_T	$\sim$ 110 mb	$\sim$ 120 mb
$\sigma(b\bar{b})$	$\sim$ 200 μb	$\sim$ 500 μb
$\sigma(b\bar{b})/\sigma_T$	1/550	1/240
$N(b\bar{b})/10^7 s$	2×10^{11}	5×10^{11}
N_{int}/s	1.1×10^7	1.2×10^7

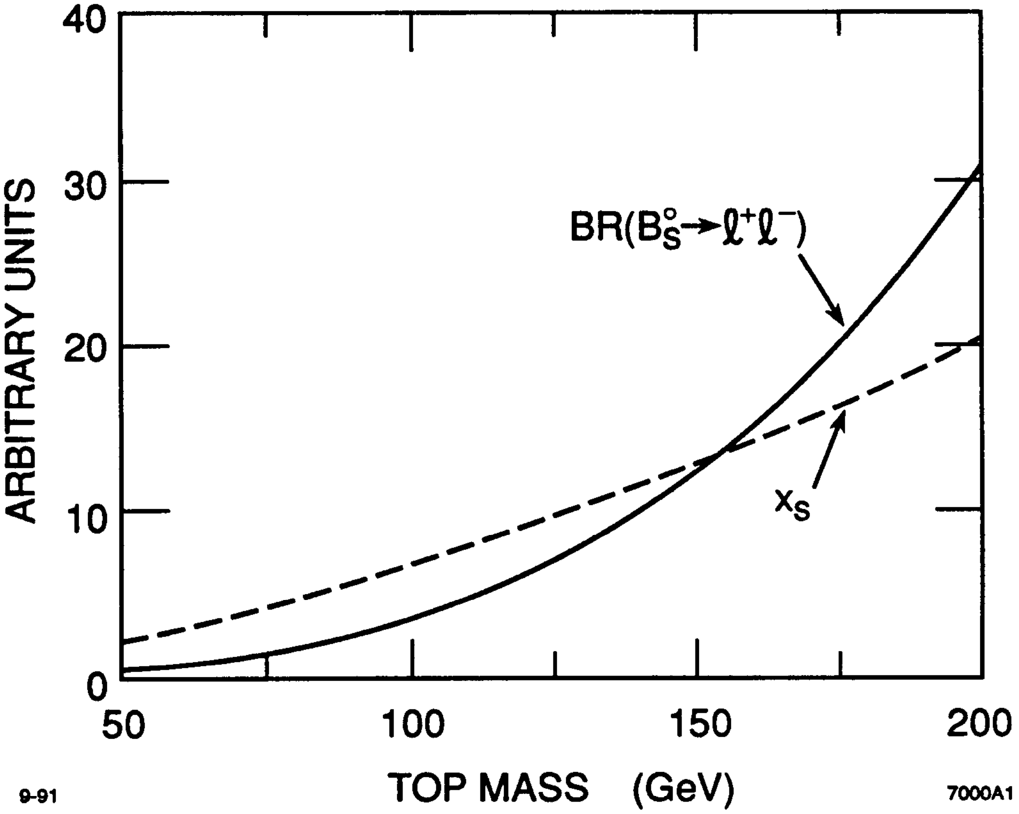
TABLE III. *Kinematical cuts for detecting the* $B_s^0 \to \mu^+\mu^-$ *decay, estimated by using a Monte Carlo calculation [14] at the SSC center-of-mass energy. Given are the fraction of the events passing the kinematical cuts,* ϵ*, and the corresponding signal-to-background ratio, S/B. The cuts are the* μ *momentum (p), its transverse momentum* $p_T \geq 1.5$ *GeV/c, and* $1^\circ < \theta < 20^\circ$*. Here,* $N(\mu\mu)$ *is the number of* $B_s^0 \to \mu^+\mu^-$ *events passing the cuts in one year of running at* $L = 10^{33}$ cm^{-2}s^{-1}.

$p \geq$	5 GeV/c	10 GeV/c	20 GeV/c
ϵ	0.27	0.24	0.15
S/B	$\sim 5 \times 10^{-9}$	$\sim 5 \times 10^{-9}$	$\sim 5 \times 10^{-9}$
$N(\mu\mu)$	150	133	84

FIGURE CAPTIONS

Fig. 1. – The $B_s^0 \rightarrow l^+l^-$ branching ratio and the x_s mixing parameter as a function of m_t, the top quark mass.

Fig. 2. – The efficiency ϵ of detecting the $B_s^0 \rightarrow \mu^+\mu^-$ events as a function of the maximum θ value at $\sqrt{s} = 40$ TeV. Here θ is the μ emission angle with respect to the beam line. The curves are given with μ having $p_T > 1$ GeV/c and various p limits.



40
30
20
10
0
ARBITRARY UNITS
$BR(B_s^\circ \to \ell^+\ell^-)$
x_s
50
100
150
200
TOP MASS (GeV)
9-91
7000A1

Fig. 1

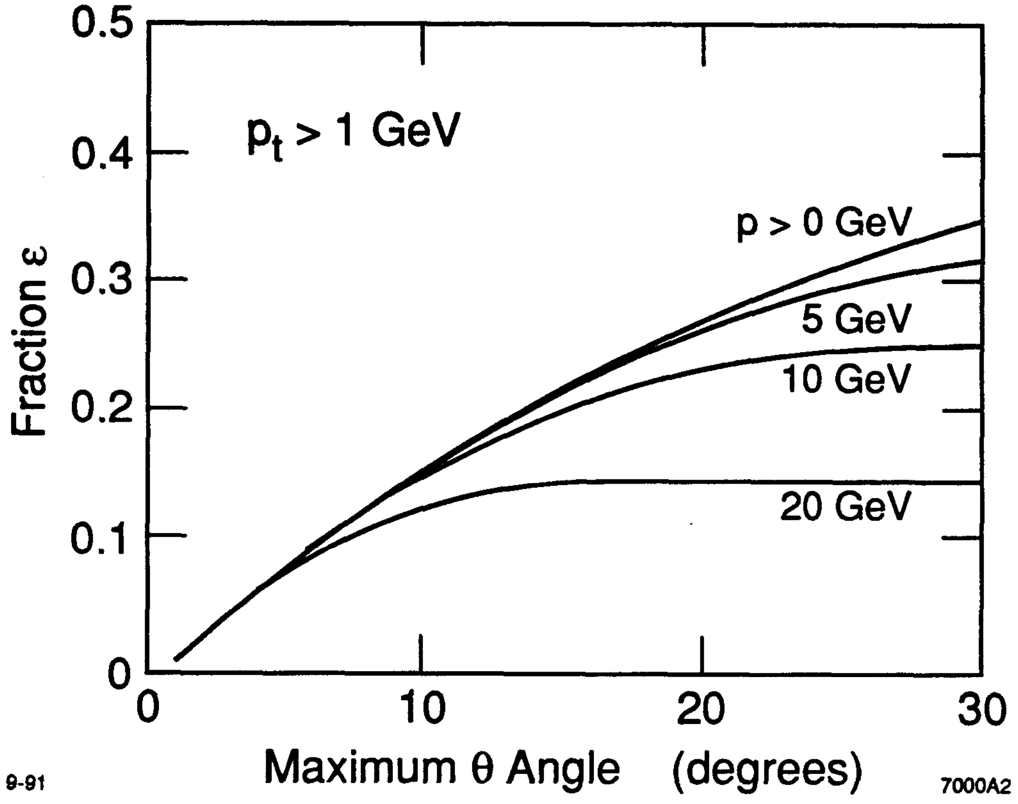

0.5
0.4
0.3
0.2
0.1
0
Fraction ε
p_t > 1 GeV
p > 0 GeV
5 GeV
10 GeV
20 GeV
0
10
20
30
Maximum θ Angle (degrees)
9-91
7000A2

Fig. 2