DØNote 4539-CONF

Observation of the B_c Meson and a Study of its Properties in the $J/\psi\mu X$ Final State

The DØ Collaboration
URL http://www-d0.fnal.gov
(Dated: August 16, 2004)

Using 210 pb^{-1} of data collected by the DØ experiment, we observe and study the properties of the B_c meson in the $B_c^+ \to J/\psi\mu^+ X \to \mu^+\mu^-\mu^+ X$ final state (charge conjugate decay modes are implicit throughout this paper unless otherwise specified). We observe 231 candidates, with the number of signal estimated by a combined lifetime and mass fit to be $95 \pm 12 \pm 11$. The combined fit estimates the mass to be $5.95^{+0.14}_{-0.13} \pm 0.34$ GeV/c^2 and the lifetime to be $0.448^{+0.123}_{-0.096} \pm 0.121$ ps.

Preliminary Results for Summer 2004 Conferences

I. INTRODUCTION

The B_c meson is the last of the ground state B mesons to be definitively observed. The phenomenology of the B_c system is as rich as the B_u and B_d systems, and offers opportunities to further test phenomenological models used to describe heavy flavor meson properties. There are many current phenomenological models that predict B_c lifetime, mass, and branching fractions to a rich array of final states [1] [2] [3]. Most of the models agree the mass should be around 6.4 GeV, and the lifetime should be more similar to that of charmed mesons, rather than B mesons (most models agree lifetime should be between about 0.3 to 0.5 ps). The properties of the past RunI observation of the B_c meson agree, within large uncertainties due to the limited statistics of the data, with these predictions [4].

The B_c meson decays a significant fraction of the time to final states containing a J/ψ mesons. Such mesons are easily triggered by the $\mu^+\mu^-$ final state, even in the busy event environment of a hadron collider such as the Tevatron. The upgraded DØ detector [5] has tracking in a solenoidal magnetic field along with excellent muon fiducial coverage, making our experiment a promising venue for significantly improving the current understanding of the B_c meson.

In this analysis we identify B_c candidates in the $B_c^+ \to J/\psi\mu^+X$ final state. We estimate backgrounds using a $J/\psi+1$ track data control sample selected with the same kinematic cuts used for the $B_c^+ \to J/\psi\mu^+X$ candidate sample. Backgrounds arising from prompt and displaced J/ψ mesons are treated separately in this analysis. We verify the background modelling ansatz using a sample of data $B_c^+ \to \psi(2S)\mu^+X$ candidates (where the $\psi(2S)$ is identified in the $\mu^+\mu^-$ final state); theoretical models predict the $B_c^+ \to J/\psi\mu^+\nu$ branching fraction to be approximately 5 to 100 times larger than that of $B_c^+ \to \psi(2S)\mu^+\nu$ [3], thus the $B_c^+ \to \psi(2S)\mu^+X$ candidates will be almost completely dominated by background.

We then perform a combined unbinned likelihood fit to the invariant mass and observed "pseudo"-proper time distributions (this is not the same as the true proper time of the B_c because of the undetected particles, such as the neutrino, in the final state)[15] to determine the B_c mass and lifetime.

Full details of the data selection and analysis procedure can be found in reference [6].

II. ANALYSIS

A. Data

The data set consists of $\sim$210 pb^{-1} of data that satisfy the DØ dimuon trigger conditions, and are processed with a recent version of the DØ reconstruction software.

B. Event Selection

Events are selected that include at least one pair of oppositely charged muons within $\Delta R < 2$ of each other. There are three layers in the muon system of the DØ detector, and muons are identified in this analysis using the DØ standard muon identification criteria that requires a muon candidate to have an associated hit in at least one of the layers of the muon chambers [7]. Each muon must also have at least one associated hit in each of the the silicon microvertex and central tracking detectors and have $1.5 < p_T < 300$ GeV/c. The dimuon system is required to have invariant mass within 0.25 GeV/c^2 of the J/ψ mass [8].

The dimuon pairs are passed on to a vertexing algorithm that determines whether or not the two muons originate from a common 3D vertex. A J/ψ mass constraint is applied to each vertex candidate; if the mass constrained vertex fit successfully converges, the algorithm then attempts to associate other tracks in the event to the vertex. Only tracks with $p_T > 1.5$ GeV/c are considered, and the best $J/\psi+1$ track vertex is taken to be the vertex that has maximal vertex fit probability $\mathcal{P}_{\rm fit}$, where $\mathcal{P}_{\rm fit}$ is the probability of achieving a χ^2 per degree of freedom larger than that returned by the vertex fit. The J/ψ and $J/\psi+1$ track vertices must have $p_T > 4$ GeV/c and $p_T > 5$ GeV/c, respectively, and the mass constrained J/ψ vertex and the J/ψ+1 track vertex must each have $\mathcal{P}_{\rm fit} > 0.05$.

$B_c^+ \to J/\psi\mu^+X$ candidates are selected from the $J/\psi+1$ track sample by requiring the extra track be a muon and that the $J/\psi\mu$ invariant mass of $J/\psi+1$ track candidate be less than 8 GeV. The extra muon in the data $B_c^+ \to J/\psi\mu^+X$ candidates will be referred to as the **third muon**.

A background control sample is selected from the data J/ψ+1 track sample by requiring that the extra track not be associated with a muon.

In order to study the background estimation procedure used in this analysis, we use a sample of $B_c^+ \to \psi(2S)\mu^+X$ candidates; this candidate sample is selected in an exactly analogous fashion to the $B_c^+ \to J/\psi\mu^+X$ candidates, except that a $\psi(2S)$ mass constraint is applied to the dimuon vertex. Similarly, a $\psi(2S)$+1 track background control sample is also selected.

A study of the reconstructed 2D decay length of J/ψ mesons in Monte Carlo generated $B^+ \to J/\psi K^+$ decays that pass similar topological and kinematic cuts to those used in this analysis reveals that the vertexing and selection procedure produces a bias of about -10 μm in the decay length, and that the bias does not appear to be decay length dependent. This small bias is on the order of that occurring from residual misalignments in the silicon detector, and we currently do not correct for either of these slight biases. Instead we assess the sensitivity of the analysis to their presence by performing two separate systematic studies wherein the decay length of the secondary vertices is shifted by $+20$ μm, and then -20 μm.

Table III in Appendix A gives a summary of the number of events in the $B_c^+ \to J/\psi\mu^+ X$ and $B_c^+ \to \psi(2S)\mu^+ X$ candidate samples and the respective background control samples.

C. Monte Carlo

Monte Carlo algorithms are relied upon in this analysis only for a kinematic description of the B_c decay products as a function of B_c mass and p_T. We use EvtGen[9] for this purpose. We use the Isgur-Scora-Grinstein-Wise (ISGW) decay model for the B_c [10]. To estimate the sensitivity of the analysis to the model used, however, we also cross-check the analysis using Monte Carlo B_c decays generated through phase-space, and also with form factors taken from reference [11].

The same kinematic cuts used to identify the data candidates are applied to the Monte Carlo samples. Samples of 50000 B_c decays are generated at mass hypotheses evenly separated by 0.1 GeV/c^2 intervals between 5.5 and 6.7 GeV. Additional samples are generated at finer mass binning close to 6.2 GeV/c^2.

The invariant mass and pseudo-proper time of the Monte Carlo generated decays are Gaussian smeared using the measured invariant mass and pseudo-proper time uncertainties of the data candidates, as determined by the vertex fit.

D. Background Estimation

Background in this analysis arises from three primary sources. The first background consists of prompt J/ψ mesons that vertex with another charged particle in the event that decays-in-flight to a muon or punches through the calorimeter to the muon system. The second consists of J/ψ mesons from B meson decays that have an associated charged particle in the decay chain that decays in flight to a muon or punches through. The third background consists of J/ψ mesons from B meson decays that vertex with another muon from elsewhere in the event (for example, a true muon from the decay of the other b quark in the event, or a decay-in-flight or punch-through charged particle from elsewhere in the event). In this analysis the backgrounds arising from prompt J/ψ's are treated separately from the backgrounds arising from J/ψ's from B meson decays (which we refer to as **heavy flavor background**).

We use the data J/ψ+1 track background control sample to estimate the pseudo-proper time and invariant mass distributions of events in the data $B_c^+ \to J/\psi\mu^+ X$ candidate sample. We break the background sample into a separate prompt and heavy flavor+prompt samples, by making the pseudo-proper time cuts $T<0$ and $T>0$,respectively.

The pseudo-proper time distribution of the heavy flavor component of the background is obtained from the pseudo-proper time distribution of background candidates with $T>0$, minus the pseudo-proper time distribution of candidates with $T<0$ flipped about the $T=0$ axis to obtain the symmetric distribution.

The behavior of any other kinematic quantity in the heavy flavor component of the background is deduced from the distribution of that quantity for the heavy+prompt $T>0$ sample, minus the distribution of that quantity for the prompt $T<0$ sample. Monte Carlo studies show that only a small fraction of heavy flavor background lies below $T=0$, thus to a good approximation this procedure is valid.

E. $\psi(2S)$ Feed-Down and Non-Resonant $B_c^+ \to J/\psi\mu^+\nu\pi^0$ Production

Before we can extract the mass and the lifetime of the B_c from the $B_c^+ \to J/\psi\mu^+ X$ final state, we first determine the fraction of our $B_c^+ \to J/\psi\mu^+ X$ candidates that are actually feed-down from $B_c^+ \to \psi(2S)\mu^+\nu$ decays where the $\psi(2S)$ decays to a final state containing a J/ψ.

To estimate this, we examine the $J/\psi\pi^+\pi^-$ invariant mass of data events that pass a $J/\psi\pi^+\pi^-\mu^+ X$ selection very similar to the $B_c^+ \to J/\psi\mu^+ X$ selection, except that the vertex consists of a J/ψ plus two tracks and a muon, rather than a J/ψ plus a muon with no other specific track requirements (note that the $J/\psi\pi^+\pi^-\mu^+ X$ sample is a subset of the $J/\psi\mu^+ X$ sample). The two tracks are each required to have $p_T > 0.5$ GeV/c. Monte Carlo studies show that $B_c^+ \to \psi(2S)\mu^+\nu$ decays predominantly populate the $J/\psi\pi^+\pi^-\mu^+$ invariant mass region from 4 to 6 GeV/c^2.

We look for evidence of $\psi(2S)$ decays identified in the $J/\psi\pi^+\pi^-$ final state in this sample. Figure 1 shows the $J/\psi\pi^+\pi^-$ invariant mass of events in the data $J/\psi\pi^+\pi^-\mu^+X$ candidate sample. The background is estimated by $J/\psi\pi^+\pi^-\mu^+X$ events with invariant mass not between 4 to 6 GeV, and the background is normalized such that the number of events with $T<0$ is the same in both the candidate and background samples. Subtracting the background reveals that there are at most about 10 to 15 events in a peak close to the $\psi(2S)$ mass at around 3.7 GeV.

The efficiency for selecting $J/\psi\pi^+\pi^-\mu^+X$ decays relative to $J/\psi\mu^+X$ decays of B_c mesons is estimated from a sample of Monte Carlo generated events to be between about 35% to 45%. The branching fraction of $\psi(2S)\to J/\psi\pi^+\pi^-$ is roughly twice that of $\psi(2S)\to J/\psi\pi^0\pi^0$ [12]. Since we have most about 15 $\psi(2S)\to J/\psi\pi^+\pi^-$ candidates in our $J/\psi\mu^+X$ sample, we thus estimate that the $J/\psi\mu^+X$ sample contains at most about 15 $\psi(2S)\to J/\psi\pi^0\pi^0$ events.

Note, however, that B mesons can also decay to final states that contain a $\psi(2S)$. Thus we can expect that some fraction of the $\psi(2S)$ mesons we observe in the $J/\psi\pi^+\pi^-$ invariant mass distribution will be due to B backgrounds. We therefore calculate the central values assuming that $30/2 = 15$ of our $B_c^+\to J/\psi\mu^+X$ events are due to $\psi(2S)$ feed-down decays, and as a systematic study we take a conservative approach and fully vary this number to 0 and then to 30 to assess the sensitivity of the mass and lifetime estimates to the feed-down fraction assumption.

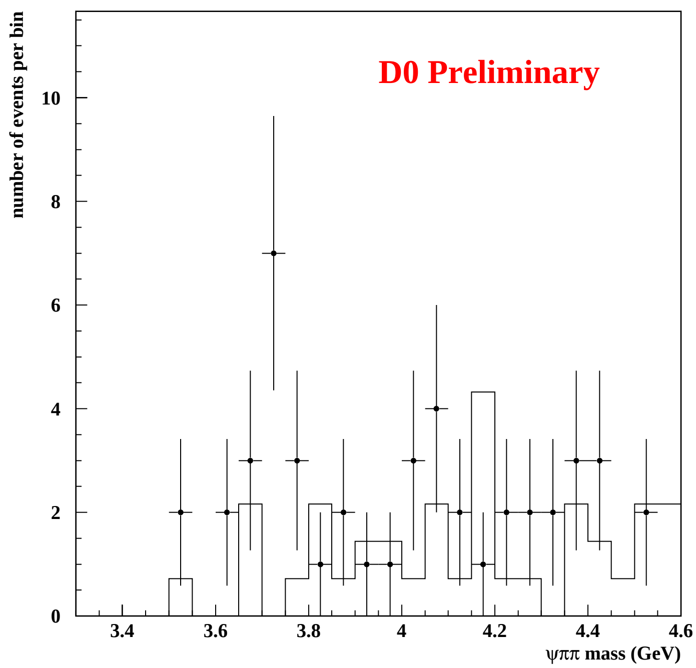

FIG. 1: $J/\psi\pi^+\pi^-$ invariant mass of events in the data $J/\psi\pi^+\pi^-\mu^+X$ candidate sample that have $J/\psi\pi^+\pi^-\mu^+$ invariant mass between 4 to 6 GeV/c^2 (points). The histogram is the background, as estimated by events in the data $J/\psi\pi^+\pi^-\mu^+X$ candidate sample that do not have $J/\psi\pi^+\pi^-\mu^+$ invariant mass between 4 to 6 GeV.

In addition to $\psi(2S)$ feed-down, we also must account for the possible presence of non-resonant $B_c^+\to J/\psi\mu^+\nu\pi^0$ production. In B_u and B_d semileptonic decays, an examination of the inclusive branching fractions versus the branching fractions to exclusive resonant final states shows that about 30% of such decays are non-resonant [12]. An examination of the $B^0\to J/\psi K^+\pi^-$ branching fraction versus the $B^0\to J/\psi K^{*0}$ branching fraction shows that the non-resonant contribution can be as high as 30% in this system ($B^0\to J/\psi K^+\pi^-$ are likely to be a better analogy to $B_c^+\to J/\psi\mu^+\nu$ than semileptonic decays of the B with pseudoscalar D mesons in the final state). We assume that 15% of the signal comes from $B_c^+\to J/\psi\mu^+\nu\pi^0$ decays. We then vary this fraction down to 0 and then to 30% to test the sensitivity of the analysis to assumptions made about the presence of non-resonant decay modes.

F. Combined Mass and Lifetime Analysis

An unbinned likelihood fit is performed to the pseudo-proper time and $J/\psi\mu$ invariant mass of candidates in the $B_c^+ \to J/\psi\mu^+ X$ sample.

The expected probability density versus mass of signal events is obtained using a kernel method [13][14] from Monte Carlo $B_c^+ \to J/\psi\mu^+\nu$ and $B_c^+ \to \psi(2S)\mu^+\nu$ decays generated at mass hypotheses evenly distributed at 0.10 GeV/c^2 intervals between 5.5 and 6.7 GeV. A mixture of 35000 $B_c^+ \to J/\psi\mu^+\nu$ decays, 7500 non-resonant $B_c \to J/\psi\mu^+\nu\pi^0$ decays, and 7500 $B_c^+ \to \psi(2S)\mu^+\nu$ decays that pass the same kinematic selections as the data $B_c^+ \to J/\psi\mu^+X$ candidates are generated at each mass hypothesis. The $\psi(2S)$ in the Monte Carlo generated $B_c^+ \to \psi(2S)\mu^+\nu$ sample are decayed through the $J/\psi\pi^0\pi^0$ final state. The relative $B_c^+ \to \psi(2S)\mu^+\nu$ feed-down component and the non-resonant $B_c \to J/\psi\mu^+\nu\pi^0$ fraction in the signal are each fixed at 15%. Each Monte Carlo generated decay has the $J/\psi\mu$ invariant mass and pseudo-proper time smeared on a point-by-point basis with the corresponding uncertainties of each event in the data $B_c^+ \to J/\psi\mu^+X$ candidate sample.

The probability distribution of the pseudo-proper time of the signal in the lifetime analysis is obtained from an exponential distribution, convolved on an event-by-event basis with the uncertainty on the pseudo-proper time of each data event, then corrected with an average factor derived from the Monte Carlo generated events to adjust for the fact that the measured pseudo-proper time does not reflect the true B_c proper time because of the undetected neutral particles. The analytic form of the measured proper time probability density for a particular event is:

$$F(T) = \mathcal{A} \sum_{k=1,N} \frac{\exp(-Tc_k/\tau + \sigma^2 c_k^2/2\tau^2)}{\sqrt{\pi}\tau} \mathrm{erfc}\left[\frac{\sigma^2 c_k^2 - \tau T)}{\tau\sigma c_k^2\sqrt{2}}\right], \tag{1}$$

where T is the observed 2D pseudo-proper time of the $\psi\mu$, σ is the measured uncertainty on the 2D pseudo-proper time, and τ is the lifetime of the B_c. N is the number events in the correction factor distribution derived from Monte Carlo generated B_c decays in that bin of ψ and third muon p_T, and c_k is the k^{th} correction factor in that p_T bin. $\mathcal{A}$ is a factor that ensures that $\int_{-\inf}^{+\inf} F(T)dT = 1$. The correction factor c_k is equal to the true proper time of the k^{th} Monte Carlo generated B_c meson, divided by the measured pseudo-proper time of the $\psi\mu$ system in the decay chain of the Monte Carlo generated B_c. The average value of c_k is around 1.1 but depends on the p_T of the J/ψ and the third muon.

The expected probability density versus mass of background events is obtained using a kernel method from the ψ+1 track background control sample; the mass probability density of the prompt background is obtained from events in this sample with pseudo-proper time <0 ps. The mass probability density of the heavy flavor background is obtained from events with pseudo-proper time >2 ps.

The probability density versus measured pseudo-proper time of background events is obtained from the $J/\psi+1$ track background control sample; the probability density of the prompt background is obtained from events in this sample with pseudo-proper time $T<0$. The prompt density distribution is then symmetrized by flipping it about the $T=0$ axis. The pseudo-proper time probability density of the heavy flavor background is obtained from events with pseudo-proper time $T>0$, minus the prompt probability density.

The fit is performed in bins of J/ψ p_T and third muon p_T, and a prompt background fraction and a heavy flavor background fraction are allowed to float in each bin. The lifetime of the signal is also allowed to float in the fit. Because the pseudo-proper time misreconstruction rate appears to be somewhat larger in the $B_c^+ \to J/\psi\mu^+X$ candidates than in the $J/\psi+1$ track background control sample, we introduce a second Gaussian into the fit. The width and normalization of this Gaussian floats in the fit. The fitting procedure is tested and verified using a simulated sample of roughly the same size and composition as the data sample.

The fit is performed under a variety of mass hypotheses. Figure 2 shows the values of $-2\log(\text{likelihood})$ returned by the fit at a variety of mass hypotheses. A full 2D parabolic fit to the values of $-2\log(\text{likelihood})$ returned by the fits yields a mass estimate of $5.95^{+0.14}_{-0.13}$ GeV/c^2 and a lifetime estimate of $0.448^{+0.123}_{-0.096}$ ps. The mass and lifetime are found by the fit to be uncorrelated. The results of the best fit are shown in Figure 3 and Table I.

The number of signal events estimated by the fit is 95 ± 12. The difference in $-2\log(\text{likelihood})$ between the 5.95 GeV/c^2 signal+background fit and a background-only fit is 60 for 5 degrees of freedom.

1. *Cross-Checks and Systematic Studies*

The estimated systematic uncertainties on the B_c mass and lifetime and on the estimated number of B_c signal in the $B_c^+ \to J/\psi\mu^+X$ sample are show in Table II. The systematic studies are as follows:

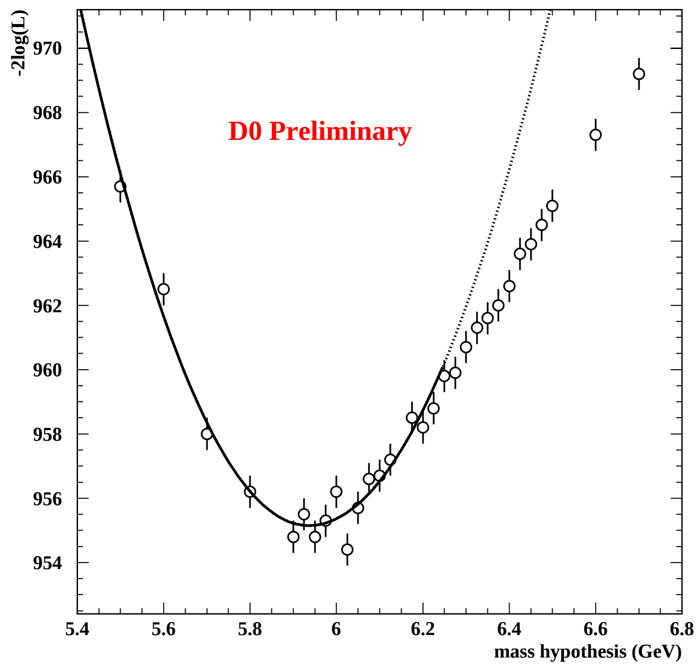

FIG. 2: Distribution of $-2\log(\text{likelihood})$ returned by the combined mass and lifetime likelihood fit at a variety of mass hypotheses.

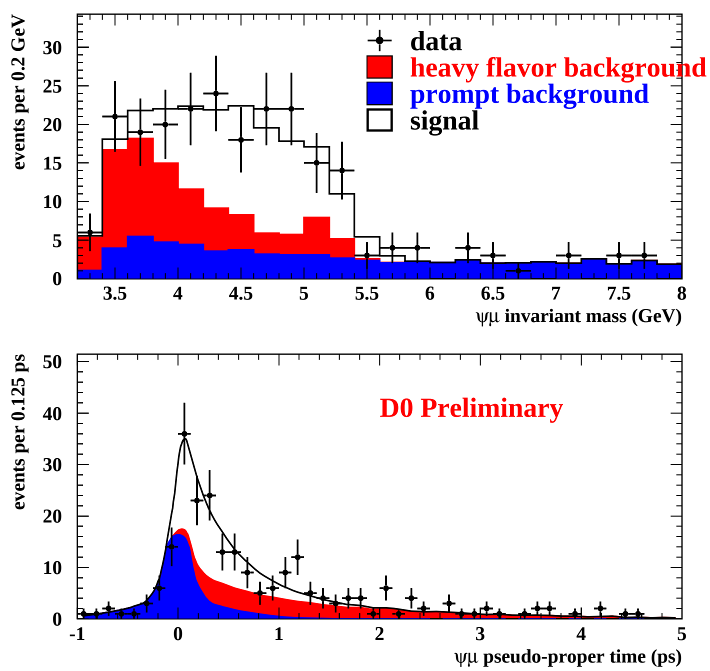

FIG. 3: The $J/\psi\mu$ invariant mass and pseudo-proper time distributions of the data $B_c^+ \to J/\psi\mu^+X$ candidates (points), with the results of the best combined mass and lifetime likelihood fit overlaid. The mass hypothesis used in the Monte Carlo generation of the signal events is 5.95 GeV/c^2.

TABLE I: Number of background and signal estimated by the combined mass and lifetime fit.

Third muon p_T (GeV/c)	J/ψ p_T (GeV/c)	Total #	# signal	# prompt	# heavy flavor
1.5 to 3.0	4 to 8	40	12.2 ± 5.0	18.0	9.8
1.5 to 3.0	above 8	44	5.7 ± 3.1	11.0	27.3
above 3.0	4 to 8	73	51.8 ± 7.8	21.2	0.0
above 3.0	above 8	74	25.8 ± 6.7	15.8	32.4
Total		231	95.5 ± 11.8	66.0	69.5

1. We check the sensitivity of the analysis to the limited statistics of the background sample by breaking the background sample into two halves, and redoing the analysis with each separate subsample. We assess this systematic uncertainty with the spread in the lifetime and mass values returned by the separate analyses divided by $\sqrt{2}$.

2. We assess the sensitivity of the analysis to the relative fraction of the non-resonant $B_c \to J/\psi\mu^+\nu\pi^0$ decays in the B_c candidate sample by redoing the analysis varying the fraction from 0 to 30%.

3. We assess the systematic error associated with the uncertainty in the $\psi(2S)$ feed-down fraction by varying this fraction from 15% down to 0 and up to 30%.

4. Binning the sample in bins of J/ψ p_T and third muon p_T ensures that the analysis should be less sensitive to any suppositions made about the B_c momentum spectrum. To check this, another sample of Monte Carlo B_c events is generated using a momentum spectrum that is scaled by a factor of 0.5 over the original spectrum (i.e., the old probability for generating a B_c with momentum of 2 GeV/c is now the probability for generating a B_c with momentum of 1 GeV/c). Similarly, we also cross-check with another sample of Monte Carlo B_c events generated using a momentum spectrum that is scaled by a factor of 2.0.

5. We check the sensitivity of the analysis to the momentum binning scheme by doubling the number of J/ψ p_T bins and third muon p_T bins.

6. Monte Carlo studies show that the primary and secondary vertexing algorithms produce a bias in the measured 2D decay length of displace vertices of approximately -10 μm. Additionally, previous DØ analyses of B lifetime measurements have found that alignment effects can produce a bias of up to 5 μm in the decay length. The analysis is repeated, first shifting the decay length of the J/ψ+track vertex by -20 μm, and then by $+20$ μm.

7. We check the sensitivity of the analysis to the selection in $\mathcal{P}_{\rm fit}$ by raising the $\mathcal{P}_{\rm fit}$ cut on both the J/ψ vertex and the J/ψ+muon vertex to $\mathcal{P}_{\rm fit} > 0.1$. Approximately 90% of the signal events in the original selection should now pass the new selection. The lifetime fit with the new selection at a mass hypothesis of 5.95 GeV/c^2 yields an estimate of 79.4 ± 12.8 signal events out of 197 candidates, with an estimated lifetime of 0.431 ± 0.120 ps. The fraction of signal events from the original selection that pass the new selection is thus estimated to be 0.84 ± 0.05, in agreement with the expected fraction of 0.90. The change in the fitted mass using the new selection is $+0.06$ GeV/c^2, and the change in the fitted lifetime is -0.028 ps. To be conservative, we included these changes in our assessment of the systematic error.

8. We assess the sensitivity of the background estimation procedure to the relative fraction of prompt and heavy flavor components in the background control sample by redoing the analysis requiring that the J/ψ p_T be less than 8 GeV/c in the control sample. The relative fraction of the prompt to heavy flavor components of the background is about 2 to 1 with the new selection, whereas with the original selection it is roughly 2 to 3.

9. As a further cross-check, we allow the lifetime to float separately in the fit at a mass hypothesis of 5.95 GeV/c^2 to the $J/\psi + \mu^+$ and $J/\psi + \mu^-$ candidates in the sample. The resulting lifetime from 117 $J/\psi\mu^+X$ candidates is 0.644 ± 0.294 ps, and the lifetime from 114 $J/\psi\mu^-X$ candidates is 0.424 ± 0.116 ps.

10. The mass determined by a fit to the invariant mass only is $6.00^{+0.14}_{-0.16}$ GeV/c^2, and the lifetime determined by a fit to the pseudo-proper time only is $0.323^{+0.138}_{-0.088}$ ps.

Other systematic studies are planned to assess the sensitivity of the analysis to other aspects of the selection and fitting procedure. We believe, however, that the list above includes all dominant systematic effects.

TABLE II: Summary of systematic uncertainties.

	Mass (GeV/c^2)	Lifetime (ps)	# Signal
Statistical	$^{+0.14}_{-0.13}$	$^{+0.118}_{-0.094}$	11.8
Limited statistics of background sample	0.06	0.013	3.0
Fraction non-resonant $B_c^+ \to J/\psi\mu^+\pi^0\nu$	0.14	0.022	6.7
Feed-down fraction from $B_c^+ \to J/\psi(2S)\mu^+\nu$	0.08	0.017	5.4
MC signal modeling: phase space vs. ISGW	0.16	0.023	4.4
MC signal modeling: HQET vs. ISGW	0.06	0.007	1.8
B_c p_T spectrum	0.05	0.004	0.8
Momentum binning	0.14	0.062	0.4
Alignment and primary vertexing algorithm	0.08	0.085	3.1
$\mathcal{P}_{\rm fit}$ selection criteria	0.06	0.028	–
Sensitivity to prompt/heavy relative bkgd fractions	0.15	0.036	–
Total systematic error	0.34	0.121	10.7

G. Other Properties of the B_c Candidates

One of the dominant production mechanisms of B_c mesons requires the production of a $c\bar{c}$ pair off of a b quark. The b and $\bar{c}$ combine to form the B_c leaving a c quark which, upon fragmentation, will likely result in a charmed hadron in relative proximity to the B_c.

In contrast, in B_u and B_d meson production there are usually (but not always) no charmed or b mesons in close angular proximity to the B_u or B_d decay.

One of the hallmarks of the decays of charm and b mesons is that they decay semileptonically a large fraction of the time. Thus more often in B_c production than in B_u and B_d meson production would we naively expect there to be another lepton in relatively close proximity to the B_c candidate.

To study this, we tighten the $J/\psi\mu$ invariant mass selection used in the $B_c^+ \to J/\psi\mu^+X$ candidates selection to be between 3.75 and 6.0/ GeV/c^2 to enhance the signal to background in the sample. B_c candidates that have pseudo-proper time between 0 and 1.5 ps are taken to lie in the signal region, whereas B_c candidates with pseudo-proper time outside this are taken to lie in the background (or "sideband") region. We then count the number of events that have a muon within 90° in ϕ from the B_c candidate. In the J/ψ+track control sample, 1% of events have a muon in close angular association to the J/ψ+track in the signal and sideband regions of the pseudo-proper time distribution. In contrast, there are 108 events in the signal region, 5 of which have a muon in close angular association (1 is expected). There are 26 events and 18 events in the $T>2$ ps and $T<0$ sidebands, respectively, none of which has an associated muon (0.5 are expected). The fraction of muon-associated events in the signal pseudo-proper time region is over 2 standard deviations different from the fraction of muon-associated events in the sideband regions, and also about 2 standard deviations from the number predicted by the J/ψ+track control sample.

III. SUMMARY

Using 210 pb^{-1} collected by the DØ experiment, we have observed and study the properties of the B_c meson in the $B_c^+ \to J/\psi\mu^+X \to \mu^+\mu^-\mu^+X$ final state. We observe 231 candidates, with the number of signal estimated by a combined lifetime and mass fit to be $95 \pm 12 \pm 11$. The combined fit estimates the mass to be $5.95^{+0.14}_{-0.13} \pm 0.34$ GeV/c^2 and the lifetime to be $0.448^{+0.123}_{-0.096} \pm 0.121$ ps. This is the first B_c search analysis to achieve an estimated number of signal with significance in excess of 5 standard deviations.

APPENDIX A: SIMPLE COUNTING ANALYSIS

Most theoretical models predict the lifetime of the B_c meson to be significantly shorter than that of B mesons. Under this hypothesis, we can perform a simple counting analysis by normalizing the heavy flavor component of the background to the number of events in the data $B_c^+ \to J/\psi\mu^+ X$ candidate sample that have $T > 2$ ps. Note that if the analysis of the lifetime later shows this short lifetime hypothesis to be incorrect, we would simply redo the counting analysis with a longer pseudo-proper time cut.

As a cross-check to demonstrate that the heavy flavor component of the background control sample correctly describes the heavy flavor background in the data, we examine the pseudo-proper time distributions of events in the data $B_c^+ \to \psi(2S)\mu^+ X$ candidate sample (where the $\psi(2S)$ is identified in the $\mu^+\mu^-$ final state). Theoretical models predict the $B_c^+ \to J/\psi\mu^+\nu$ branching fraction to be approximately 5 to 100 times larger than that of $B_c^+ \to \psi(2S)\mu^+\nu$ [3]. This, taken together with the fact that the $J/\psi \to \mu^+\mu^-$ branching fraction is almost 10 times larger than that of $\psi(2S) \to \mu^+\mu^-$, implies that the $B_c^+ \to \psi(2S)\mu^+ X$ candidates should be almost completely dominated by background.

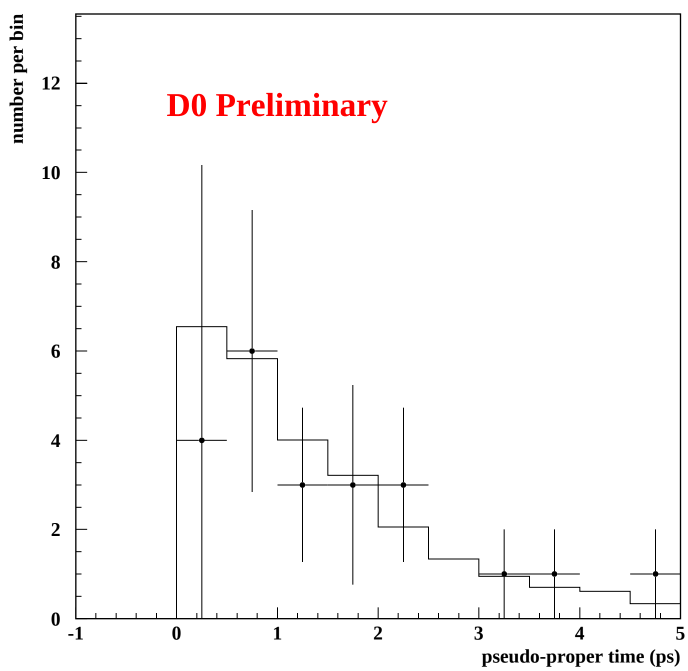

FIG. 4: The $\psi(2S)\mu$ pseudo-proper time distribution of the heavy flavor component of the data $B_c^+ \to \psi(2S)\mu^+ X$ candidates. The background normalization is obtained by assuming that the $B_c^+ \to \psi(2S)\mu^+ X$ candidates with pseudo-proper time greater than 2 ps are all background.

Figure 4 shows the pseudo-proper time distribution of the heavy flavor component of the data $B_c^+ \to \psi(2S)\mu^+ X$ candidates. The corresponding distribution of the heavy flavor component of the data $\psi(2S)$+track background control sample is overlaid. Under the assumption that the $B_c^+ \to \psi(2S)\mu^+ X$ candidates with pseudo-proper time greater than 2 ps are all background, the background normalization ansatz yields an estimate of -1.3 ± 12.4 signal events in the 20 candidates in the heavy flavor component of the data $B_c^+ \to \psi(2S)\mu^+ X$ sample. The heavy flavor component of the data background control sample yields a reasonable description of the corresponding distribution in the data $B_c^+ \to \psi(2S)\mu^+ X$ sample.

Figure 5 shows the J/ψ+muon pseudo-proper time distribution of the heavy flavor component of the data $B_c^+ \to J/\psi\mu^+ X$ candidates. The background is normalized using the same ansatz as before, yielding an estimate of 65 ± 26 signal events in the 183 candidates in the heavy flavor component of the data $B_c^+ \to J/\psi\mu^+ X$ sample. Note that this background estimate effectively includes no shape information from any kinematic distribution (except in the loose sense that a pseudo-proper time cut is used to define a pure background sample), yet this simple counting method yields an estimate of the number of signal that has a significance of about 2.5 standard deviations.[16].

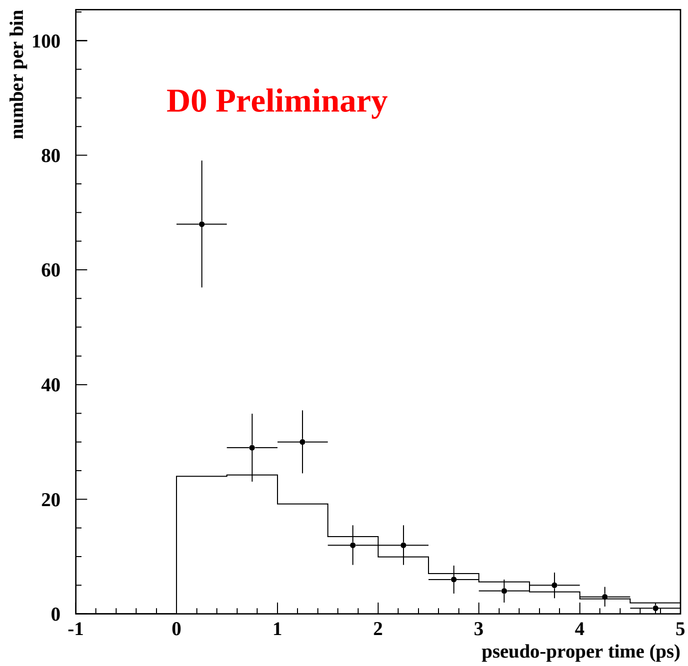

FIG. 5: The $J/\psi\mu$ pseudo-proper time distribution of the heavy flavor component of the data $B_c^+ \to J/\psi\mu^+ X$ candidates. The background normalization is obtained by assuming that the $B_c^+ \to J/\psi\mu^+ X$ candidates with pseudo-proper time greater than 2 ps are all background.

Table III gives a summary of the number of events in the $B_c^+ \to \psi(2S)\mu^+ X$ and $B_c^+ \to J/\psi\mu^+ X$ candidate samples and the number of signal estimated by the simple counting analysis in each.

TABLE III: Summary of the number of events in the data samples.

Sample	Total #	# in Heavy Flavor Component	Estimated # Signal (from Counting Analysis)
$B_c \to J/\psi\mu^+ X$ candidates	231	183	65 ± 26
J/ψ+1 track background	42000	25100	–
$B_c \to \psi(2S)\mu^+ X$ candidates	66	20	-1 ± 12
$\psi(2S)$+1 track background	6300	2500	–

[1] C.Chang and Y.Chen, Phys.Rev **D49** (1994) 3399.
[2] D. Ebert, R.N. Faustov, V.O. Galkin. (2004) hep-ph/0401237
[3] D.Ebert *et al*, Phys.Rev **D68** (2003) 094020, and references therein.
[4] The CDF Collaboration, Phys.Rev.Lett **81** (1998) 2432-2437.
[5] DØ Collaboration, F. Last, Nucl. Instrum. Methods A **123**, 456 (2004).
[6] S.Towers, internal DØNote 4484 (2004).
[7] DØ certified muons for P14
`http://www-d0.fnal.gov/phys_id/muon_id/d0_private/certif/p14/index.html`
[8] G.Borissov *et al*, DØNote 4481 (2004).
[9] D.J.Lange, Nucl.Instrum.Meth **A462** (2001) 152.
[10] D.Scora and N.Isgur, Phys.Rev **D52** (1995) 2783-2812.
[11] J.E.Duboseq *et al* (CLEO), Phys.Rev.Lett. **76** (1996) 3898.
[12] *Review of Particle Physics*, S. Eidelman et al., Physics Letters B**592**, 1 (2004).

[13] D.Scott, *Multivariate Density Estimation*, John Wiley and Sons (1992).

[14] L.Holmstrom *et al*, Comp.Phys.Commun **88** (1995) 195.

[15] Throughout this paper pseudo-proper time is defined as $T = (L_{\psi\mu} \cdot P_{\psi\mu})|L_{\psi}|M^{\psi\mu}/(0.02998|L_{\psi\mu}||P_{\psi\mu}|^2)$, where L_{ψ} and $L_{\psi\mu}$ are the 2D decay length vectors of the J/ψ and J/ψ+muon vertices, respectively, and P_{ψ} and $P_{\psi\mu}$ are their respective 2D momentum vectors. $M^{\psi\mu}$ is the mass of the J/ψ+muon vertex.

[16] Note that because of the truncation at $T = 0$, the estimate of the number of signal in the sample from the simple counting analysis will actually be around 20% too low because around 10% of the pseudo-proper time distribution of signal events spills below $T = 0$ because of detector resolution.