CDF note 7159

The "Lifetime" Distribution of $X(3872)$ Mesons Produced in $\bar{p}p$ Collisions at CDF

The CDF II Collaboration
URL http://www-cdf.fnal.gov
(Dated: August 3, 2004)

The "lifetime" distribution of $X(3872) \to J/\psi\pi^+\pi^-$ decays produced at the Tevatron is analyzed to quantify what fraction of this new meson have an apparent lifetime, and thus arise from the decay of b-hadrons, in contrast to those produced promptly. We also study a similar $\psi(2S) \to J/\psi\pi^+\pi^-$ sample for comparison. We find, for the selected samples, that $28.3 \pm 1.0\,(stat) \pm 0.7\,(syst)\,\%$ of the $\psi(2S)$ candidates are from the b-decays, whereas $16.1 \pm 4.9\,(stat) \pm 2.0\,(syst)\,\%$ of the X-mesons arise from such decays. These fractions are dependent upon the transverse momentum of the selected mesons, but for this particular sample with its modest p_T cut the production is dominated by prompt sources rather than b-decays.

Preliminary Results — Summer 2004

I. INTRODUCTION

At the Lepton-Photon Symposum held at Fermilab in August 2003 the Belle Collaboration announced the discovery of a new narrow charmonium-like state decaying into $J/\psi\pi^+\pi^-$ [1]. This meson, later dubbed the $X(3872)$, was observed in exclusive $B^+ \to XK^+$ decays. The obvious interpretation of this new resonance is as the long-sought $^3D_2$ charmonium state. However, the $^3D_2$ is expected to be lighter, and also have a prominent decay into $\chi_1\gamma$, which Belle failed to observe. Furthermore, as Belle emphasized, the $X(3872)$ mass is very close to the $D^0\overline{D}^{*0}$ mass, and they conjectured that the $X(3872)$ could be a D^0-$\overline{D}^{*0}$ "molecule." The true nature of the $X(3872)$ stands as a mystery [2].

The CDF collaboration presented preliminary results the following month confirming the $X(3872) \to J/\psi\pi^+\pi^-$ in about 220 pb^{-1} of $\bar{p}p$ collisions, and measured its mass with a value and precision similar to Belle. These results have been finalized and accepted for publication [3]. A distinctive feature of the CDF observation relative to Belle's is the production of this state. Belle observed the $X(3872)$ in exclusively reconstructed B-decays with a sample of about ~ 35 candidates. On the other hand, the CDF sample is about 700 candidates—albeit on a very large background—and can be expected to be a mix coming from b-hadron decays as well as directly produced X-mesons. The Belle observation demonstrates that B-decays can be a significant source of the $X(3872)$, but otherwise its production rate in particle collisions is not known. In particular, without further study it was not known whether or not the CDF X-sample could in fact arise entirely from b-hadrons.

If the structure of the $X(3872)$ lies beyond the conventional quark model there is currently little theoretical guidance as to its production rate. However, interpreted as a standard $c\bar{c}$-bound state there is a large body of experimental and theoretical work on charmonium production to draw upon. For a recent experimental survey with references to theory see Ref [4]. Differences between the production of the $X(3872)$ as an exotic hadron versus as a conventional charmonium state may offer insights of whether or not this new state has an exotic quark composition, or is simply a new member of the $c\bar{c}$-family—albeit one which poorly fits prior potential model predictions. The $\psi(2S)$ provides a good reference state because it is prominantly produced and it has a decay topologically and kinematically similar to the $X(3872)$, making experimental comparisons more direct. Furthermore, $\psi(2S)$ production in $\bar{p}p$ collisions has been studied in its own right, both experimentally [5] and theoretically [6].

A simple hypothetical consideration highlights the point. The yield of the $X(3872)$ originally reported by CDF is *potentially* quite comparable to the observed $\psi(2S)$ rate. Applying the same reconstruction and selection cuts for both resonances—with the one exception of further demanding that the dipion mass for X-candidates be above 500 MeV/c^2—we reported 730 $X(3872)$ candidates compared to 5790 $\psi(2S)$'s. Application of the dipion cut to the $X(3872)$ selection but not for the $\psi(2S)$ has a small effect on the relative efficiencies because little if any $X(3872)$ signal falls below 500 MeV/c^2. Even though the selections are equivelant there will still be some efficiency difference between the $\psi(2S)$ and the $X(3872)$ because of their differing masses, and significantly, the p_T-spectrum of X-production is unknown. Ignoring the the efficiency correction, and to be concrete, assuming that the unknown $X(3872) \to J/\psi\pi^+\pi^-$ branching ratio to be 10%, the CDF yields would appear to translate into an $X(3872)$ production rate that is $\sim 40\%$ of the $\psi(2S)$ rate. This would easily place the $X(3872)$ rate observed by CDF in the natural realm of conventional quarkonia production. This kind of argument in favor of a charmonium interpretation is at best circumstantial in the absence of reliable production models for the alternate interpretations of the $X(3872)$. But more immediately, such a rate argument can not yet be made, one way or the other, because: the X-branching ratio is unknown, the necessary efficiency corrections depend upon the unknown production spectra of the X, and finally, the observed sample consists of an unknown mixture of X's coming from b-decays and non-b sources.

A well established technique of addressing the last issue is the statistical separation of these two components by measuring the apparent "lifetime" distribution, *i.e.* the decay distance traversed by moving X-mesons. Being a non-weak decay, the physical lifetime of the $X(3872)$ is too short to have a decay vertex measurably displaced from its production point, and thus any displacement observed beyond detector resolution may be attributed to the decay distance of a b-hadron producing that $X(3872)$.

A step in this direction has been taken by the DØ Collaboration [7], which has recently reported observing the $X(3872)$ at the Tevatron as well. They considered the relative signal yields for candidates below "proper decay distances" of 100 μm, to the number above. The ratio of yields for the $X(3872)$ was statistically indistinguishable from the same ratio for the $\psi(2S)$, indicating that in this sample the two states have roughly similar "lifetime" characteristics. The relative contributions of prompt and long-lived sources for this $100\,\mu$m subdivision is, however, not specifically quantified. Furthermore, in this sample the $\psi(2S)$ appears to be in a mass region with a much different detection efficiency than for the $X(3872)$, highlighting the importance of explicit efficiency corrections.

In the present analysis we statistically separate and quantify the prompt and long-lived production components of the $X(3872)$ and the $\psi(2S)$ in the previously published CDF sample. This quantification *is sample specific*. In particular, the relative production fractions of prompt and long-lived $\psi(2S)$ mesons is known to be p_T-dependent at the Tevatron [5]. The close similarity between these two states in mass and decay should result in relatively modest efficiency differences in this sample *except* for the unknown production p_T-spectrum of the $X(3872)$. Until the efficiencies are established, our results nevertheless enable one to make a more informed comparision of the $X(3872)$ and the $\psi(2S)$ production.

II. DETECTOR & DATA SAMPLE

This analysis is a direct extension of the original CDF observation and mass analysis of the $X(3872)$ [3]. The same data sample, reconstruction, and selection serve as the starting point. We briefly review the salient points of that previous analysis and extend the discussion with new features.

Our analysis is based on an integrated luminosity of about 220 pb^{-1} of $\bar{p}p$ collisions at $\sqrt{s} = 1.96\,TeV$ which were collected with the CDF II detector between February 2002 and August 2003. The CDF II detector is described in detail elsewhere [8], and only a few major features are recalled here. The $X(3872) \to J/\psi\pi^+\pi^-$ sample is obtained by triggering on dimuon pairs from $J/\psi \to \mu^+\mu^-$ decays. Muon triggers are the product of a successful spatial match between tracks in triggered muon chambers which surround the main CDF detector, and tracks constructed by a "fast-track processor" (XFT) using hits from the central drift chamber (COT). The latter chamber is sensitive to ionizing particles over radii from 43 to 132 cm from the beamline. The COT is constructed in 8 superlayers of 12 wires each, with the layers alternating between axial and stereo measurements for full three dimensional track reconstruction. Furthermore, it is immersed in a 1.4 T solenoidal magnetic field for the measurement of charged particle momenta transverse to the beamline, p_T. The minimum p_T for an acceptable XFT muon candidate is 1.5 or 2.0 GeV/c, depending upon which region of the muon system is being matched. All subsystem triggers are fed into the final "Level-3" trigger system, which consists of a large network of PCs running offline-like reconstruction code and selection algorithms. In particular, the complete digitized COT information is available for track reconstruction and a more refined matching to muon chamber tracks. Only opposite-sign muon-pairs with masses between 2.7 and 4.0 GeV/c^2 are accepted by the Level 3 low-mass dimuon trigger for permanent recording.

One other detector component which plays a critical role in the current analysis is the SVX tracker. This device consists of five concentric layers of double-sided silicon sensors surrounding the beampipe at radii from 2.5 to 10.6 cm. Information from the SVX was used as available in the X-mass measurement, but in that application its value was largely peripheral. On the other hand, the spatial precision obtained with the SVX is the fundamental ingredient which makes it possible to distinguish the displaced vertices of b-decays. Correspondingly, effective use of the SVX imposes some additional selection requirements beyond that used in the mass analysis.

The dimuon triggers were processed with standard reconstruction software and detector calibrations and alignments. A loose $J/\psi \to \mu^+\mu^-$ sample was created as a standard starting point for all J/ψ-based analyses.

III. $J/\psi\pi\pi$ EVENT SELECTION

The $J/\psi\pi\pi$ candidate selection is based on adding pairs of reconstructed charge tracks, assumed to be pions, to the $J/\psi \to \mu^+\mu^-$ candidates. All the same selection requirement used in the mass analysis [3] are again applied. First, a preselection is made demanding:

- Minimum hit requirements for the SVX and COT on all tracks;
- $\mu^+\mu^-$ mass window from 2.8-3.2 GeV/c^2;

- Track $p_T > 350$ MeV/c^2 for pions;
- Vertex fit of $\mu^+\mu^-\pi\pi$ system (dimuons constrained to world average J/ψ mass) must have $\chi^2 < 40$ (6 DoF).

Because of the high multiplicity of charged tracks produced in Tevatron events our principle backgound is the association of random tracks with real J/ψ's. Multiplicity fluctuations in the vicinity of the J/ψ, particularly in the form of jets, has a long tail whereby a few events can contribute dozens of $J/\psi\pi\pi$ candidates in a single event. Such events inject a large background contribution relative to a small potential signal. We therefore entirely reject events that have more than 12 $J/\psi\pi\pi$ preselection candidates with masses below 4.5 GeV/c^2. The value of this cut was obtained by an optimization study based on our $\psi(2S)$ signal, and its specific value is directly dependent upon how the preselection sample is defined. This number appears rather large at this stage because of the loose selection. After the full selection the average number of $J/\psi\pi^+\pi^-$ candidates with masses between 3.65 and 4.0 GeV/c^2 is less than 1.2 per event, where of course the average is over events with at least one valid candidate.

A tighter selection is imposed to improve the signal-to-noise. We continue to reproduce the mass analysis with the following requirements:

- Vertex fit of $\mu^+\mu^-$ must have $\chi^2 < 15$ (1 DoF);
- Dimuon mass from the above vertex fit must be within 60 MeV/c^2 of world average;
- Transverse momentum of the J/ψ greater than 4 GeV/c;
- Opposite-sign pions;
- Transverse momentum of the pions greater than 400 MeV/c;
- Vertex fit of $J/\psi\pi^+\pi^-$ tightened to $\chi^2 < 25$;
- $\Delta R \equiv \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} \leq 0.7$ for both pions.

In this last requirement $\Delta\phi$ is the difference in azimuthal angle between the pion under consideration and the full $J/\psi\pi^+\pi^-$ candidate, and $\Delta\eta$ is the difference in pseudorapidity. This selection replicates the mass distribution of Ref [3], and the resulting $J/\psi\pi^+\pi^-$ mass distribution is reproduced in the left plot of Fig. 1. Peaks for the $\psi(2S)$ and $X(3872)$ are clearly visible.

As originally noted by Belle, the dipion system of the $X(3872)$ signal strongly favors high values. The right plot of Fig. 1 shows the $J/\psi\pi^+\pi^-$ mass distribution for candidates with $M(\pi^+\pi^-) < 500$ MeV/c^2 by the open circles, and for those above 500 MeV/c^2 by the solid points. One is unable to discern any signal in the low-mass plot. This translates into a significant reduction in background in the high-mass subsample with apparently no loss of signal.

The $M(\pi^+\pi^-)$-cut is a powerful tool for improving the sensitivity of the lifetime analysis. A 500 MeV/c^2 cut results in essentially no loss in signal, and this kind of cut should introduce no bias into the lifetime distribution. Even though $\psi(2S)$ decays also favor high dipion masses, a 500 MeV/c^2 cut has a strong effect because of the $\psi(2S)$'s smaller mass. In order to maintain similar efficiencies for both states and yet take advantage of the dipion mass to purify the X-signal we treat the two resonances differently in the lifetime analysis. We analyze the $\psi(2S)$ *without* any dipion cut; and since the 500 MeV/c^2 cut introduces little reduction in X-yield we are free to apply it there with virtually no loss of efficiency.

Up to this point the $J/\psi\pi^+\pi^-$ sample is identical to that of Ref. [3]. However, extending the analysis into the lifetime domain mandates a few additional requirements. The one that introduces the largest change in the sample size is the requirement that the error on the $J/\psi\pi^+\pi^-$ projected displacement "L_{xy}" (see below) from the interaction vertex be less than 125 μm. Smaller sample reductions arise by explicitly demanding SVX operation bits and the availablity of run-averaged beamlines used in determining L_{xy}.

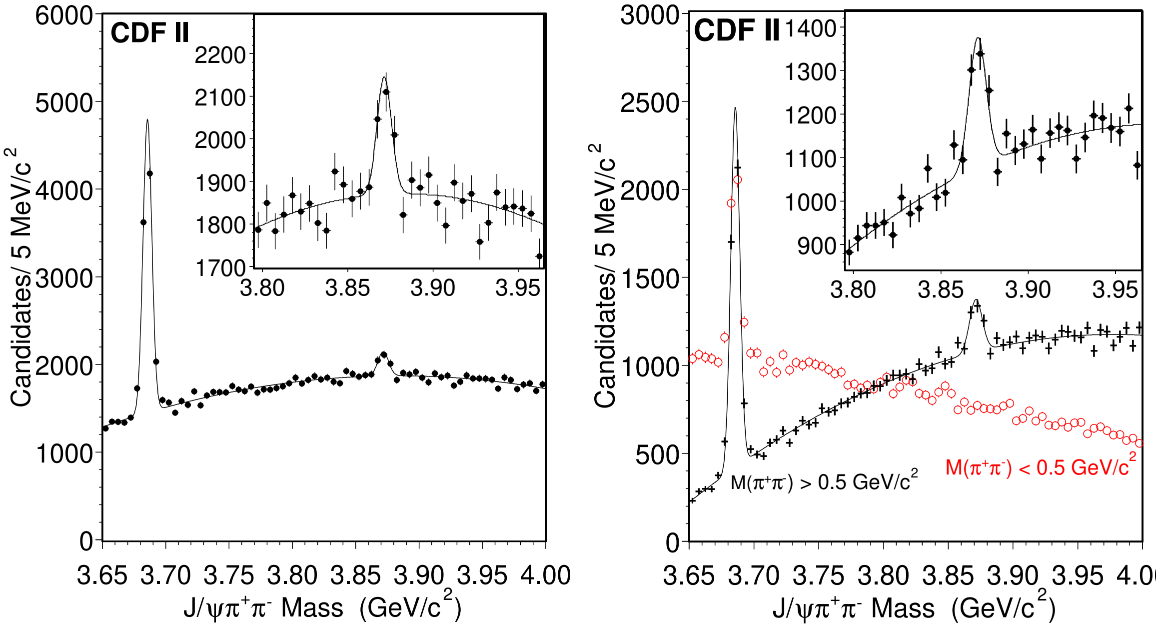

FIG. 1: The $J/\psi\pi^+\pi^-$ mass distributions for the mass analysis of Ref. [3]. Left plot corresponds to the nominal sample and selection (see text). Right plot shows the same sample separated into those candidates with dipion masses less than (circles) or greater than (solid points) 500 MeV/c^2.

IV. MEASURING AND FITTING THE "LIFETIME" DISTRIBUTION

Measuring the "lifetime" distribution of charm and bottom hadrons is a well established technique. Such hadrons moving at reasonable speeds traverse distances which are significant for precision tracking systems. In this analysis the vertex fit of the $J/\psi\pi^+\pi^-$ system determines the transverse (i.e. in "r-ϕ" plane) projection of the decay point, $\vec{x}_{decay}$. The position of the primary interaction vertex, $\vec{x}_{prim}$, is determined from run-by-run beamline averages. The laboratory decay distance is then defined as:

$$L_{xy} \equiv \frac{(\vec{x}_{decay} - \vec{x}_{prim}) \cdot \vec{p}_T}{|\vec{p}_T|}, \tag{1}$$

where $\vec{p}_T$ is the transverse momentum vector of the $J/\psi\pi^+\pi^-$ system. Ideally this quantity can be transformed into the proper decay time:

$$\widehat{ct} \equiv \frac{ML_{xy}}{p_T}, \tag{2}$$

where M is the particle mass. However, the true proper time requires M to be the actual mass and p_T to be the transverse momentum of the decaying particle. In this analysis the decaying particles are an unknown mix over all b-hadrons, and the full p_T is unknown because other decay products are not reconstructed. In lieu of the true proper time, we define a "pseudo-proper time" ct, for which M and p_T are for the observed $J/\psi\pi^+\pi^-$ system. In inclusive analyses where the goal is to measure a physically meaningful particle lifetime the true proper time can be recovered by applying a Monte Carlo-derived correction factor $\langle K\rangle$, so that:

$$\widehat{ct} \equiv \frac{ML_{xy}}{\langle K\rangle p_T}. \tag{3}$$

However, our interest is in the fraction of signal which is long-lived, and not in measuring an average b-hadron lifetime. Therefore we make no such correction, and sometimes refer to ct as the "uncorrected proper time." It is important to stress that the actual X "lifetime" obtained by fitting the ct-distribution is only indirectly related to true b-hadron lifetimes. Furthermore, this relationship is both complex and sample specific.

To separate the prompt from long-lived components we perform an unbinned likelihood fit in which both mass and lifetime dimensions are simultaneously included. Each candidate is characterized by its mass M, uncorrected proper time, *i.e.*

$$ct \equiv \frac{M(J/\psi\pi^+\pi^-)}{p_T(J/\psi\pi^+\pi^-)} L_{xy}, \tag{4}$$

and the uncertainty on ct, σ_{ct}. The last quantity is derived from the results of the $J/\psi\pi^+\pi^-$ vertex fit. The likelihood model for M, ct, and σ_{ct} may be characterized by the following categories: the signal and background mass distributions, the signal and background proper-time distributions, and the ct-resolution function. Specifically, these components are:

- Mass:
  - Signal: *Gaussian distribution*;
  - Background: *Second order polynomial*;
- Proper Time:
  - Signal: *Exponential convoluted with ct-resolution function*;
  - *Background: Three exponentials convoluted with ct-resolution function*:
    One describing negative-ct with fraction f_- and lifetime λ_-,
    Two describing positive-ct with fractions f_+ and f_{++}, and lifetimes λ_+ and λ_{++};
- ct-Resolution Function: *Two Gaussian distributions.*

The Gaussians for the ct-resolution function are centered at zero, and the widths are determined by σ_{ct} for the candidate multiplied by floating scale factors, a separate one for each Gaussian. The use of scale factors gives the fitter the freedom to adjust the global ct-error scale.

V. THE $\psi(2S)$ "LIFETIME" DISTRIBUTION

Data spanning $J/\psi\pi^+\pi^-$ masses from 3640-3740 MeV/c^2 are used for the $\psi(2S)$ fit. This sample consists of 29,185 candidates. The likelihood fit returns a $\psi(2S)$ signal of 4940 ± 110 candidates. In order to graphically portray the results of the fit we project the likelihood probability density function (PDF) separately onto mass and lifetime histograms of the data. The top plot in Fig. 2 shows the mass projection, and on the bottom is the uncorrected proper-time distribution. Along with the full PDF, the bottom plot also shows the breakdown between prompt and long-lived components for both the signal and the background. The quality of the likelihood fit is monitored by examining the residuals and χ^2's of the projections shown here. All of these show the fits behave well.

Overall, the background strongly dominates the ct-plot in Fig. 2 because of the large amount of background included with a wide mass window. To better expose the likelihood's description of the signal we can project the PDF only for events within ± 2.5 standard deviations of the fitted mass width of the signal. This projection is shown in Fig. 3, and indeed the shaded region now dominates. The projection again describes the data well.

The principal parameter of interest from the likelihood fit is the fraction of $\psi(2S)$ signal which is long-lived. This is $28.3 \pm 1.0\,(stat.)\%$. The "lifetime" of this long-lived exponential is $473 \pm 13\,(stat)\,\mu$m. This value is quite close to the known B^0 and B^+ lifetimes, but as emphasized earlier, the "lifetime" returned from the $\psi(2S)$ fit can not be directly compared to the B^0 and B^+ lifetimes other than to note they are of similar scale. Background parameters for the $\psi(2S)$ fit are listed in the broader discussion of Section VIII.

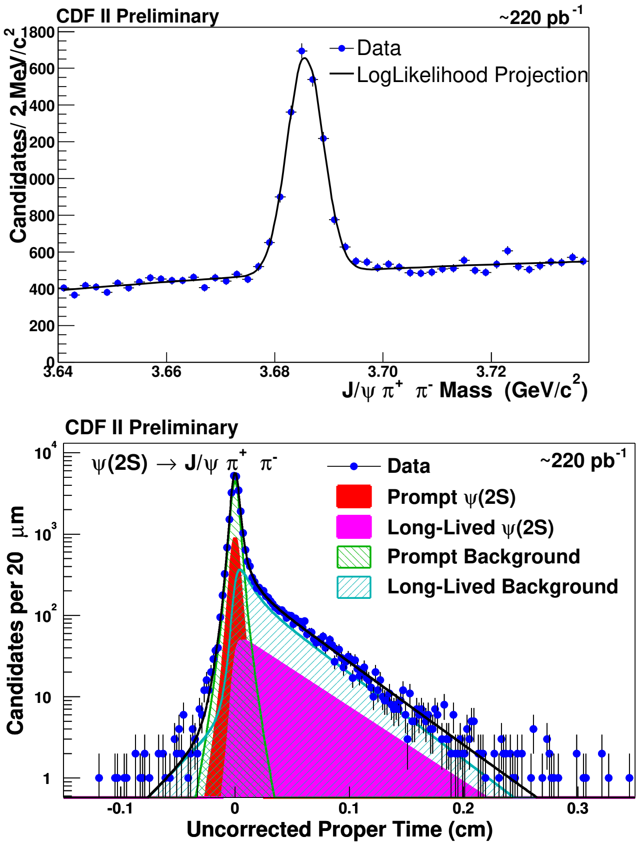

FIG. 2: Projections of the $\psi(2S)$ likelihood fit onto mass (top) and uncorrected proper time (bottom). The different parts of the lifetime probability density function are also shown: the separate prompt and long-lived components of the signal (shaded area) and background (hatched area).

The likelihood fit provides a good description of the data and is a very powerful approach. A much less optimal, but more transparent method, is to perform a simple sideband-subtraction of the uncorrected proper-time distribution. The $\psi(2S)$ signal stands well above the background making this a straightforward operation, the results of which are shown in Fig. 4. The inset shows the full $J/\psi\pi^+\pi^-$ mass window used in the likelihood fit, and the signal and sideband regions used in the subtraction. The uncorrected proper-time distribution from this subtraction is shown as points in the main plot of Fig. 4. The data are fit by a binned-χ^2 with essentially the same model as used in the likelihood fit. Since there is no background after subtraction those parts of the model are removed. Futhermore, as the binned histogram no longer contains the event-by-event σ_{ct} information, the two resolution Gaussians simply have floating widths.

Our fitting model again provides a good description of the data, as is visually apparent in Fig. 4, but now with a χ^2-fit. Examination of fit residuals and χ^2's confirm the visual impression. The χ^2-fit produces a long-lived fraction of $28.1 \pm 1.1\%$, and a "lifetime" of $468 \pm 19\,\mu$m, in excellent agreement with the full likelihood fit. The χ^2-fit errors are marginally larger than for the likelihood. This is sensible in that $\psi(2S)$ parameters are well determined by the signal candidates common to both fits, but the likelihood fit gains a slight benefit from the background candidates helping improve the determination of the ct-resolution.

The fit of the $\psi(2S)$ resonance provides both a good test-case for the the general approach and specific implementation of quantifying the relative prompt and long-lived fractions, as well as a charmonium benchmark for these fractions. We found the lifetime model describes the data well, and both the likelihood and sideband subtraction approaches agree very well. We next apply these same procedures to the $X(3872)$.

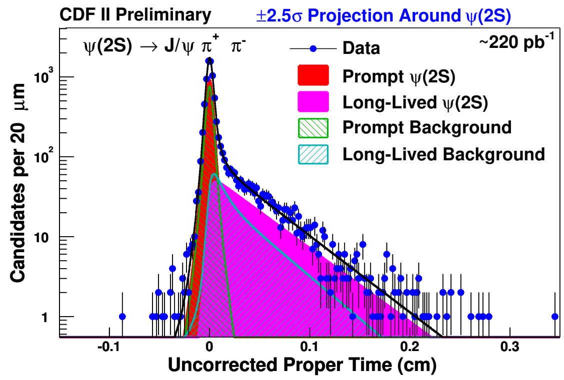

FIG. 3: The uncorrected proper-time projection of the full $\psi(2S)$ likelihood probability density function over the restricted mass range ± 2.5 standard deviations around the $\psi(2S)$ resonance.

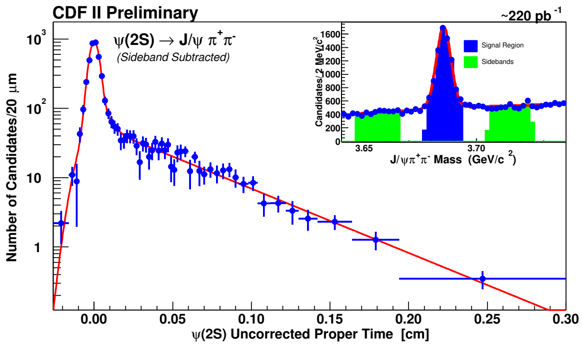

FIG. 4: The uncorrected proper-time distribution for the $\psi(2S)$ obtained from sideband subtraction. The data are shown by the points, and a χ^2 fit by the smooth curve. The inset shows the $J/\psi\pi^+\pi^-$ mass distribution with the signal and sideband regions identified by shading.

VI. THE $X(3872)$ "LIFETIME" DISTRIBUTION

For the $X(3872)$ fit we use a mass window from 3800-3930 MeV/c^2, resulting in a sample of 23,673 candidates. This window is slightly larger than for the $\psi(2S)$. This was done so that the *background* component is close to the same size in both fits, making that part of the fits statistically comperable. Note that the windown had to be *enlarged* rather than reduced because the X-selection has the additional requirement that the *dipion mass be above 500* MeV/c^2.

We use the same likelihood model as before. Even though the X-signal is much weaker than the $\psi(2S)$'s, the fit model does not simplify. The background portion of the sample is still of similar size and warrants the full description. Fitting the 23,673 candidates results in a signal of 605 ± 91 $X(3872)$'s. The likelihood projections onto mass and uncorrected proper-time are shown in Fig. 5. We again obtain good quality fits, but the X-signal is much weaker than the $\psi(2S)$ and the signal components appear deeply buried in the proper-time distribution. However, we see a pattern similar to the $\psi(2S)$: the $X(3872)$ signal is indeed composed of prompt and long-lived components. In particular, the prompt component is quite prominent. Quantitatively, the long-lived fraction is, after a small correction discussed below (Sec. VII), $16.1 \pm 4.9\,(stat)\%$, and has a lifetime of $439 \pm 107\,\mu$m. We see that like the $\psi(2S)$, our X-sample is *dominated by non-b sources*, and nominally to a somewhat greater extent than the $\psi(2S)$. However, the large uncertainty of $\pm 4.9\%$ makes this difference with the $\psi(2S)$ statistically inconclusive.

The X-signal is still background dominated even in a narrow window around the peak. Nevertheless, we follow the $\psi(2S)$ analysis and again make the projection of the likelihood PDF ± 2.5 standard deviations around the fitted X-mass in Fig. 6. The relative size of the signal has increased but, it is still badly overwhelmed by the background as it had to be.

As with the $\psi(2S)$ we can seek confirmation of the likelihood fit in a sideband-subtraction analysis. This is more problematic in that a weak signal is spread across many bins and may be overwhelmed by background fluctuations. Merging low statistics bins together helps compensate, but cannot eliminate the problem. The results of this exercise are shown in Fig. 7. The bin errors are quite large in the proper-time tail, with a few bins having negative values. A χ^2 fit of this distribution returns a long-lived fraction of $20.9 \pm 6.3\%$, and with a lifetime of $431 \pm 109\,\mu$m. The lifetime is almost exactly the same as the likelihood fit, but the long-lived fraction is somewhat larger. Given that the uncertainty on the long-lived fraction is almost 30% larger in the χ^2-fit, the observed difference between the two fractions is not statistically unreasonable. The sideband-subtracted distribution for the X-data certainly has large bin-to-bin fluctuations in the tail, but it must still be taken as good agreement with the likelihood fit.

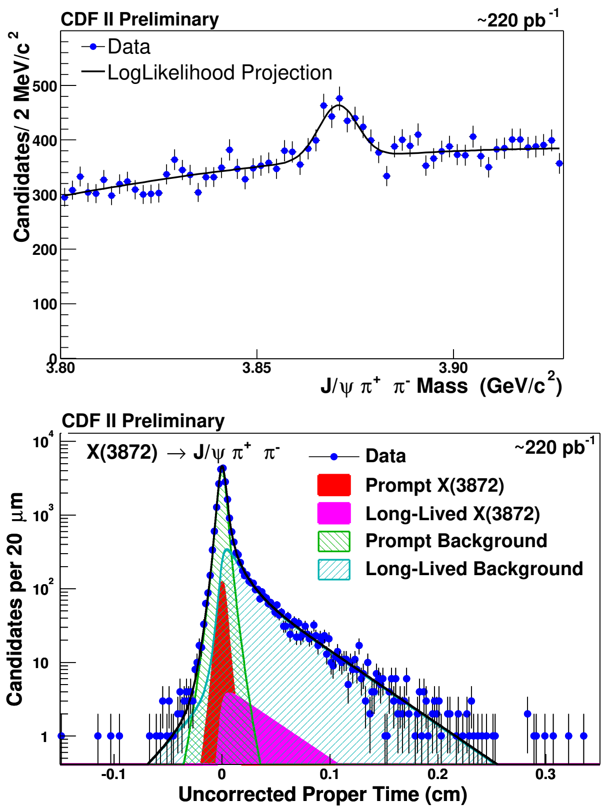

FIG. 5: Projections of the $X(3872)$ likelihood fit onto mass (top) and uncorrected proper time (bottom).

VII. SYSTEMATIC UNCERTAINTIES ON THE LONG-LIVED FRACTIONS

Since the determination of the long-lived fractions are at this point entirely sample specific, all systematic uncertainties on the long-lived fractions are related to the sample, the fitting model, or the fitting method. Specifically, we assign systematic uncertainties associated to the mass window chosen, the presence of multiple candiates per event, the fit model, and fitting biases. The same considerations were applied to both the $\psi(2S)$ and $X(3872)$.

The mass window was varied by both shifting the window position (of fixed width), and its width (about a fixed center). The likelihood fit was repeated for several permutations and the variations in the long-lived fraction tracked. Because this particular type of variation is changing the sample composition in the "wings" of the mass distribution, statistical effects may indirectly be introduced via statistical fluctations in the background. The scale of such effects were estimated and found to be reasonably small. We therefore took half of the largest shift in the fraction as the one standard deviation estimate of the systematic effect for the mass window.

Our sample contains a modest contribution of multiple candidates in the same event. Extra candidates principally arise from one J/ψ candidate with different combinations of pion associations. In the mass distribution these different candidates are essentially uncorrelated as changing even one pion has a large effect on the mass. On the other hand, the decay length L_{xy} of a candidate tends to be dominated by the high-p_T muons rather than the low-p_T pions. Therefore different $J/\psi\pi^+\pi^-$ candidates in the same event may be strongly correlated in lifetime. Furthermore, the intrinsic environment of $\bar{p}p$ collisions introduces a ct-dependence in the probability of extra candidates: it is easy to create more candidates near the interaction vertex than it is for a highly displaced J/ψ vertex. This latter fact also undermines what might seem the easy solution of simply rejecting all events that have more than one candidate. Because extra candidates are more likely to

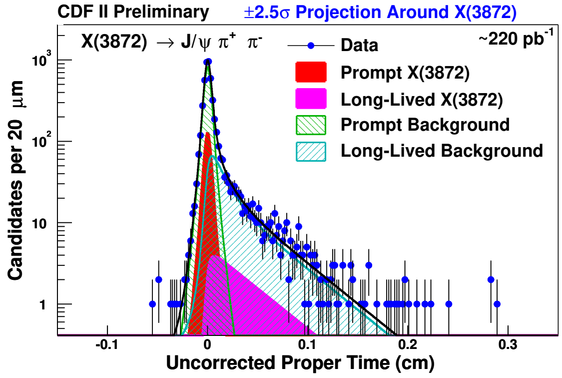

FIG. 6: The uncorrected proper-time projection of the full $X(3872)$ likelihood PDF over the restricted mass range ± 2.5 standard deviations around the $X(3872)$.

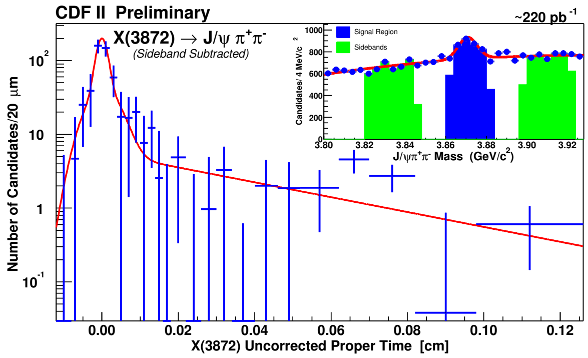

FIG. 7: The uncorrected proper-time distribution for the $X(3872)$ obtained from sideband subtraction. The data are shown by the points, and a χ^2 fit by the smooth curve. The inset shows the $J/\psi\pi^+\pi^-$ mass distribution with the signal and sideband regions identified by shading.

arise near the primary, one is more likely to reject low-ct signal if only one candidate per event is allowed. We therefore assume that the most benign course of action is to leave the sample untouched, and estimate an uncertainty due to multiple candidates by seeing the change that results from various removal schemes. We select only one candidate per event based on: random selection, p_T, dipion mass, vertex error, and vertex fit χ^2. Each selection scheme removes 2872 candidates from the $\psi(2S)$ sample, and 1742 from the X-sample. We again estimate the one standard deviation uncertainty as half of the largest shift in the long-lived fraction.

The particular fit model used has described the data well, and was chosen based on a large body of past experience, but various ingredients are only empirically based. Variations of the model were implemented and fit to the data, these include a double Gaussian description of the signal mass, triple Gaussian ct-resolution, and adding exponetial tails to the resolution function. Again, half the largest change in the fraction is assigned as the systematic.

Finally, we considered whether there were any baises in the fitter. The likelihood fitter code was separately tested and certified to be implemented correctly, and the current studies were directed at fit biases that may arise due to peculiar characteristics of the sample. This issue was explored by generating Monte Carlo "data" fluctuated about true values which are the same as given by the real data fits. Many samples of the same size as real data were generated and fit. The fit results were compared to the generation value in the form of pull distributions, which monitored for deviations from the ideal of zero-centered unit-Gaussians.

The result of the systematic studies for the long-lived fraction is as follows:

Source	$\psi(2S)$	$X(3872)$	
Mass Window	± 0.2	± 1.5	%
Duplicates	± 0.6	± 1.3	%
Fit Model	± 0.3	*negligible*	%
Fit Bias	*negligible*	± 0.4	%
Total	± 0.7	± 2.0	%

where all values are the absolute uncertainties in percent. There are a few features of this table we highlight. The fit model introduces no error for the $X(3872)$ simply because it is such a sufficiently small signal that it does not have the statistical power to discriminate between different descriptions. The $\psi(2S)$ sample on the other hand is quite large and it has some sensitivity to how it is modelled. A more important issue is the results for the fit bias.

The $\psi(2S)$, with its excellent signal to noise, shows no sign of bias, but the weaker $X(3872)$ does. We studied the behavior of fitting various small signals on a large background and found as the signal becomes weaker both the amount of signal and the long-lived fraction becomes increasingly overestimated. For the $X(3872)$ signal fit in the data the long-lived fraction is estimated to be $+0.7\%$ (absolute) too high. Since this is bias rather than a simple uncertainty we *correct the raw likelihood output* by this amount to obtain a long-lived fraction for the X of 16.1%. Furthermore, we estimate the reliablity of this correction to be ±0.4%.

This results in the final values for the long-lived fractions of this sample to be:

$$\boxed{\psi(2S) : 28.3 \pm 1.0\,(stat) \pm 0.7\,(syst)\,\%}$$

$$\boxed{X(3872) : 16.1 \pm 4.9\,(stat) \pm 2.0\,(syst)\,\%}$$

VIII. $\psi(2S)$-$X(3872)$ COMPARISONS & DISCUSSION

We have measured the long-lived fractions of the $\psi(2S)$ and $X(3872)$ in our sample, and find both to be dominated by prompt production. The likelihood results are also substantiated by a simple sideband-subtraction exercise. One other feature of the fits that may interest the reader are the results for the background description. These are summarized in terms of the fraction of the particular exponential and its lifetime. The background fractions are relative to the total background component, so that the fraction of prompt background is simply $1 - f_- - f_+ - f_{++}$. The two sets of background fit results are:

		$\psi(2S)$		$X(3872)$	
Parameter		Value	Error	Value	Error
Fraction Negative-ct Bckgrnd	f_-	0.004	±0.001	0.003	±0.001
Fraction Short Positive-ct Bckgrnd	f_+	0.066	±0.008	0.103	±0.008
Fraction Long Positive-ct Bck	f_{++}	0.188	±0.008	0.130	±0.009
Negative Lifetime (μm)	λ_-	301	±48	274	±53
Short Positive Lifetime (μm)	λ_+	87	±17	116	±12
Long Positive Lifetime (μm)	λ_{++}	409	±12	436	±17

The background parameters for the $\psi(2S)$ and $X(3872)$ are fairly consistent with each other with the exception of the amount of the long-positive exponential. The long-positive exponential also has a lifetime similar to the X-signal, but the error on the X-lifetime is very large and the closeness of the numerical agrement can not be interpreted more than both the $\psi(2S)$ and $X(3872)$ "lifetimes" are roughly similar to the background. This broad similarity is a natural consequence of the fact that the long-lived $\psi(2S)$, $X(3872)$, *and* their respective backgrounds are all coming from b-decays. Finally, we remind the reader that the similarity in the fit errors for the background parameters is due to the fact that the sample sizes were constructed to have approximately the same number of background events.

Finally, some considerations about the significance of the long-lived fraction in the X-sample. Belle's reconstruction of the $X(3872)$ in exclusive decays leaves no doubt that the $X(3872)$ should be present in inclusive B-decays. And therefore the principal interest of our result is the demonstration and quantification of the dominance of promptly produced X's in our sample. However, the mere knowledge of presence is not sufficient to ensure detection. Given the weakness of the long-lived signal we consider how well it is established.

A quick but naive estimation of the significance of the long-lived component can be obtained by dividing the fraction by its error, which gives a value over 3σ. The likelihood fit clearly prefers having a long-lived X-component as the likelihood value without it is worse, at similar level of significance. However, in order to make a more transparent portrayal of the likelihood's description of the long-lived component, we project the PDF for standard likelihood fit onto the data above a specified ct-threshold. The top plot in Fig. 8 shows the results for the $\psi(2S)$ for three different ct-thresholds. The likelihood again describes the data quite well. Since the sample is a mix of prompt and long-lived components one does not *a priori* know how much prompt contamination is present above a given ct. We use the information from the likelihood to indicate by the shaded area the amount of prompt signal present. For $ct > 40\,\mu$m the majority of the $\psi(2S)$ sample is already long-lived, but there is a substantial prompt component. Above $100\,\mu$m our $\psi(2S)$ sample is almost purely long-lived.

The same type of projection for the X, although with different thresholds, is shown in the bottom plot in Fig. 8. Similar features are seen, but with a much weaker signal. The likelihood fit discribes the data well with a long-lived signal present. From this plot we can unambiguously see that the fit is properly including a long-lived X-signal, even if it is not of large statistical significance.

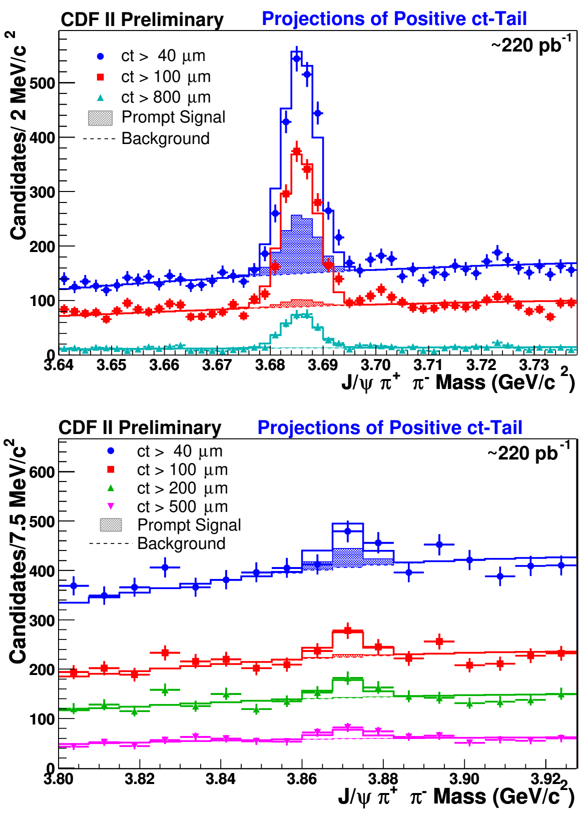

FIG. 8: Projection of the likelihood PDFs in mass for events above a specified (uncorrected) proper-time threshold. Top plot is for the $\psi(2S)$, and on the bottom is for the $X(3872)$. Points are the data, and the histogram is the projected likelihood. The shaded areas indicate the portion of the signal that the fit attributes to the prompt component of the signal.

As a final quantification of the significance of the long-lived X-fraction we quote the result of a Monte Carlo simulation of the fit. Monte Carlo "data" sets were generated with the same overall signal size of the $X(3872)$ in the data fit, but with no long-lived component. The fraction of fits which returned long-lived fractions as large or larger than observed in the X-data is equivalent to a 3-standard deviation Gaussian exclusion of the null hypothesis.

IX. SUMMARY

We have extended the original CDF observation and mass measurement of the $X(3872)$ [3] to the "lifetime" properties of our X-sample. Since the $X(3872)$ does not itself decay weakly, an apparent lifetime for this particle is indicative of feeddown through the weak decay of a parent particle. The presence of two "lifetime" components in this sample, one with experimentally significant displacements of the X-decay vertex relative to the interaction point, and another class compatible with decays at the interaction vertex, is the hallmark signature of two different production modes at work. The displaced candidates are produced via the decay of b-hadrons, and those decaying at the interaction vertex arise from "prompt" sources. Prompt sources may be either direct production from the $\bar{p}p$ interaction, or feedown through some higher, but as yet unknown, state decaying to the $X(3872)$ with a negligible lifetime. The experimental separation into the prompt and long-lived classifications is contingent upon the particle in question having a sufficient boost to travel an experimentally significant distance. A B-meson produced at rest would appear as a prompt candidate. Since the $X(3872) \to J/\psi\pi^+\pi^-$ sample selected required the p_T of the J/ψ candidate to exceed 4.0 GeV/c^2, b-hadrons feeding our sample necessarily have a sufficient boost that they cannot mimic prompt production.

We apply the technique of simultaneously fitting the candidate mass and (uncorrected) proper-time distributions with an unbinned likelihood to statistically separate the prompt and long-lived components in our sample. We first demonstrate the procedure with the similar, and prominent, $\psi(2S) \to J/\psi\pi^+\pi^-$ decays. We obtain good fits and clearly measure separate prompt and long-lived components. These results are confirmed by a simple sideband-subtraction analysis. The likelihood fit finds the following fraction of signal that is long lived is:

$$\boxed{28.3 \pm 1.0\,(stat) \pm 0.7\,(syst)\,\% \quad [\psi(2S)]}$$

The analysis is repeated for $X(3872) \to J/\psi\pi^+\pi^-$ where we establish the presence of both components. Sideband-subtraction provides a similar picture. Quantitatively, we measure the fraction of the X-signal that is long-lived to be:

$$\boxed{16.1 \pm 4.9\,(stat) \pm 2.0\,(syst)\,\% \quad [X(3872)]}$$

It must be emphasized that these fractions must be considered sample-specific at this point because the relative prompt to long-lived mixture depends upon the kinematic regime being sampled. Nevertheless, the main observation is that the $X(3872)$ has a large prompt component at the Tevatron, which goes beyond Belle's original observation of b-handrons as a source. The prompt fraction dominates our sample, by about 5-to-1. The measured long-lived X-fraction is nominally smaller than for the $\psi(2S)$, but given the uncertainties involved no strong statement can be made. In addition, kinematic differences between the $\psi(2S)$ and $X(3872)$ can be responsible for shifts in the relative fraction. The long-lived fraction of X-production is relatively small, and is determined with a substantial uncertainty. We exclude the possibility that there is no long-lived X-component at the 3σ level.

On superficial grounds X-production looks quite similar to charmonium, but further studies are necessary to properly understand the effect of efficiencies. The potential of such production studies would be greatly enhanced by the formulation of specific theoretical predictions of X-production for non-charmonium interpretations of the $X(3872)$.

Acknowledgments

We would like to thank C. Quigg for helpful discussions on this unusual particle.

We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contributions. This work was supported by the U.S. Department of Energy and National Science Foundation; the Italian Istituto Nazionale di Fisica Nucleare; the Ministry of Education, Culture, Sports, Science and Technology of Japan; the Natural Sciences and Engineering Research Council of Canada; the National Science Council of the Republic of China; the Swiss National Science Foundation; the A.P. Sloan Foundation; the Bundesministerium fuer Bildung und Forschung, Germany; the Korean Science and Engineering Foundation and the Korean Research Foundation; the Particle Physics and Astronomy Research Council and the Royal Society, UK; the Russian Foundation for Basic Research; the Comision Interministerial de Ciencia y Tecnologia, Spain; and in part by the European Community's Human Potential Programme under contract HPRN-CT-20002, Probe for New Physics.

[1] K.Abe *et al.* (Belle Collaboration), contributed paper to Lepton-Photon, hep-ex/0308029;
S.-L. Choi *et al.* (Belle Collaboration), Phys. Rev. Lett. **91**, 262001 (2003).

[2] For an entry point into the literature on X-interpretations see: C. Quigg, *Les Recontres de Physique de la Vallée d'Aoste*, 2004 (hep-ph/0403187).

[3] D. Acosta *et al.* (CDF II Collaboration), accepted for publication in Phys. Rev. Lett., hep-ex/0312021.

[4] T. Skwarnicki, *21st International Symposium On Lepton And Photon Interactions At High Energies*, 11-16 August 2003, Fermilab, Int. J. Mod. Phys. **A19**, 1030 (2004) [hep-ph/0311243].

[5] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **79**, 572 (1997); T. Daniels, Ph.D. dissertation, Massachussetts Institute of Technology, 1997.

[6] G.Bodwin, E Braaten, and G. Lepage, Phys. Rev D **51**, 1125 (1995); **55**,5853(E) (1997); E. Braaten and S. Fleming, Phys. Rev. Lett. **74**, 3327 (1995); M. Cacciari *et al.*, Phys. Lett. B **356**, 553 (1995); E. Braaten, and Y. Chen, Phys. Rev D **54**, 3216 (1996); P. Cho and A.K. Leibovich, Phys. Rev D **53**, 150 (1996); *ibid* **53**, 6203 (1996); P. Cho and M. Wise, Phys. Lett. B **346**, 129 (1995); M. Beneke and M. Kramer, Rev D **55**, 5269 (1997); E. Braaten, B. Kniehl, and J. Lee, Rev D **62**:094005 ,2000.

[7] V.M. Abazov *et al.* (DØ Collaboration), submitted to Phys. Rev. Lett., hep-ex/0405004.

[8] F. Abe, et al., Nucl. Instrum. Methods Phys. Res. A **271**, 387 (1988); D. Amidei, et al., Nucl. Instum. Methods Phys. Res. A **350**, 73 (1994); F. Abe, et al., Phys. Rev. D **52**, 4784 (1995); P. Azzi, et al., Nucl. Instrum. Methods Phys. Res. A **360**, 137 (1995); The CDFII Detector Technical Design Report, Fermilab-Pub-96/390-E