CDF/DOC/BOTTOM/CDFR/8791
February 10, 2008 - Version 2.0

Search for Flavor Changing Neutral Current $D^0 \to \mu^+\mu^-$ Decays Using 360 pb^{-1} of CDF Run II Data

E. Berry[1], I.K. Furić[2], R.F. Harr[3], Y.K. Kim[1]

(1) Enrico Fermi Institute, University of Chicago, Chicago, IL 60637
(2) University of Florida, Gainesville, FL 32611
(3) Wayne State University, Detroit, MI 48202

Abstract

We present an updated search for FCNC $D^0 \to \mu^+\mu^-$ decays using 360 pb^{-1} of Run II data. In addition to analyzing more data, the muon coverage is increased by also utilizing muons from the CMX detector. The main source of background in the signal region is identified as coming from $B \to \mu^+\mu^- X$ decays. A likelihood ratio technique is used to optimize the analysis by reducing the dominant background, while keeping the signal region blinded. The upgraded analysis sensitivity for setting a 90% CL limit is 0.65×10^{-6}, which should be compared to the current world best limit of 1.3×10^{-6}.

The blind analysis expects a total of 8.7 ± 1.7 events, of which 4.9 ± 1.3 are expected in the CMU-CMU, 2.7 ± 1.0 in the CMU-CMX, and 1.0 ± 0.5 are expected in the CMX-CMX channels. After unblinding the data, 3 events are found in the CMU-CMU, no events in the CMU-CMX, and one event in the CMX-CMX channel. We use a Bayesian approach to extract the confidence level limits on the branching fraction that allows for the combination of several channels and the inclusion of nuisance parameters. We find $\mathcal{B}(D^0 \to \mu^+\mu^-) < 5.3 \times 10^{-7}$ at the 95% confidence level and $\mathcal{B}(D^0 \to \mu^+\mu^-) < 4.3 \times 10^{-7}$ at the 90% confidence level.

1 Update History

1.1 version 2, Feb. 10, 2008

- Update the limit calculation with results using Bayesian technique.
- Update the affiliations of the authors.
- A few other corrections of typos.

2 Introduction

In the Standard Model, the flavor-changing neutral current decay $D^0 \to \mu^+\mu^-$ is suppressed by the GIM mechanism. Figure 1 shows some of the basic Feynman diagrams contributing to $D^0 \to \mu^+\mu^-$ decays in the Standard Model. The decay rate is dominated by long-range contributions [1], in particular the diagram shown in Figure 2. The total Standard Model branching fraction is estimated [1] to be:

$$\mathcal{B}(D^0 \to \mu^+\mu^-) \geq 4 \times 10^{-13} \tag{1}$$

As opposed to $B_s^0 \to \mu^+\mu^-$, which is enhanced in SUSY models by $\tan\beta$ terms due to couplings with down-type quarks, virtually no enhancement of this decay is expected from the canonical, $\mathcal{R}$-parity conserving SUSY models. However, in $\mathcal{R}$-parity violating SUSY models, enhancements are possible. Reference [1] takes into account other restrictions on $\mathcal{R}$ parity violation and estimates that enhancements of the D^0 branching fraction up to 3.5×10^{-6}. This leaves for seven orders of magnitude of the D^0 branching fraction to be probed for new physics effects. The current best experimental limit on the D^0 branching fraction comes from BaBar, [2], and is equal to:

$$\mathcal{B}(D^0 \to \mu^+\mu^-) \geq 1.3 \times 10^{-6} \tag{2}$$

at 90% C.L. CDF has already performed a search for these rare decays using the first 60 pb^{-1} of Run II data, and set a limit of $\mathcal{B}(D^0 \to \mu^+\mu^-) \geq 2.5 \times 10^{-6}$ [5]. This study builds on the experience of the previous measurement, using more data, extending the muon coverage to also include the CMX detector, and using a muon identification likelihood developed for the $\mathrm{B_s}$-mixing analysis.

3 Measurement Technique

The strategy of the measurement is rather simple, and follows the same strategy of the previous measurement [5]. Hadronic two-body D^0 decays are gathered with the two-track trigger. The combinatorial background is vastly reduced by requiring a $D^{*-} \to D^0\pi^-$ tag. The two-track candidates are reconstructed in the $\mu^+\mu^-$ mass hypothesis, and the D^*- tag mass-difference cut is applied assuming the $\mu^+\mu^-$ mass hypothesis. Both daughter tracks of the D^0 decay are required to be identified as muons in the muon chambers. The invariant mass of the track pair is required to fall inside a search window. Using background subtraction, the number of $D^0 \to \mu\mu$ decays in the search window is determined. The $D^0 \to \pi^+\pi^-$ peak is used as a reference signal to determine the branching fraction:

$$\mathcal{B}(D^0 \to \mu^+\mu^-) = \frac{N(\mu^+\mu^-)}{N(\pi^+\pi^-)} \cdot \frac{\epsilon(\pi^+\pi^-)}{\epsilon(\mu^+\mu^-)} \cdot \mathcal{B}(D^0 \to \pi^+\pi^-), \tag{3}$$

where $N(\mu^+\mu^-)$ is the number of observed dimuon decays, $N(\pi^+\pi^-)$ is the observed number of reference decays, $\epsilon(\pi^+\pi^-)/\epsilon(\mu^+\mu^-)$ is the relative efficiency for triggering and reconstructing $\pi^+\pi^-$ and $\mu^+\mu^-$ decays, respectively, and $\mathcal{B}(D^0 \to \pi^+\pi^-)$ is the reference branching fraction for $D^0 \to \pi^+\pi^-$ decays.

Two detector systems are used for muon identification - the CMU and CMX detectors. This increases the analysis acceptance for two-body D^0 decays by roughly 67%. The muon identification acceptance for the two detectors is not 100% due to detector fiduciality and efficiency turnon for the CMX detector, so the actual increase in expected dimuon acceptance due to adding the CMX detector is about 33%. A study of the single-track and dimuon acceptance is presented in Section 4. The $D^0 \to \pi^+\pi^-$ and $D^0 \to \mu^+\mu^-$ decays have very similar kinematical properties. Since they are gathered through the same trigger path, we expect that the relative trigger efficiency for these decays will be very similar. The relative trigger and reconstruction efficiencies for the two decays are also studied in Section 4.

Different decay processes can mimic $D^0 \to \mu^+\mu^-$ decays. The dominant background before muon identification are $D^0 \to \pi^+\pi^-$ decays. Section 5 explains the measurement of the muon mistag rate and the expected rate from events in which both pions are mistagged as muons. We examine several of the processes which we believe to be the dominant sources of background after muon tagging is applied. Our findings are documented in Section 7. We find that the main background in our search is due to $B \to \mu^+\mu^- X$ decays and estimate the expected sensitivity in Section 11. The measurement is optimized by rejecting the dominant background using a likelihood ratio discriminating prompt charm from B meson decays using the impact parameter and displacement significance of the two-track system. The technique and results are explained in Section 8. We find that roughly 85% of the signal can be retained while rejecting about 60% of the dominant background. After optimization, the backgrounds are re-estimated for the most optimal value of the likelihood ratio cut, and the sensitivity of the optimized analysis is estimated in Section 10.

The entire analysis is performed in an un-biased fashion, not examining the events in the signal search window. Section 12 shows the results found in the open box, and the modified statistical analysis that deals with the low event count found in the signal mass region. Section 13 summarizes the findings of this measurement.

4 Trigger and Muon Identification Efficiency

The total relative efficiency between $\pi^+\pi^-$ and $\mu^+\mu^-$ events at CDF depends on two components. The first component is the relative trigger efficiency. Pions interact hadronically with the detector, giving individual pion tracks a roughly 5% inefficiency. The second component is the muon matching efficiency, or the frequency with which real muon tracks are assigned muon ID. The total relative efficiency is a product of these two values. The calculation of both components is described in the following sections. The product of the trigger and muon identity efficiencies will yield the $\frac{\epsilon(\pi^+\pi^-)}{\epsilon(\mu^+\mu^-)}$ portion of Equation 3.

4.1 Introduction to the muon detector efficiency

The muon subsystems within CDF are not perfect, and there is a chance that a real muon passing through them would not be registered. As with the misidentification fractions, the frequency with which the muon detectors register real muons passing through them, called the "detector efficiency," is dependent on several parameters. We measure this detector efficiency by measuring the frequency with which tracks that have been identified as muons using the COT register as hits within any of the muon subsystems. The process is described in the following subsection.

4.2 Calculation of the muon detector efficiency

In order to calculate the efficiency of the muon subsystems, we must first acquire a clean source of muons. For this, we use $J/\psi \to \mu^+\mu^-$ decays. Usually J/ψ candidates are reconstructed by looking at the invariant mass histogram of pairs of muons. Two confirmed muons would bias our measurement, however, so instead we look at the invariant mass histogram of single 4.0 GeV/c^2 track with muon ID combined with an SVT-confirmed track. This histogram shows a peak at roughly 3.0 GeV/c^2, the accepted rest mass of the J/ψ (see Figure 7). We assume that parent particles under this peak (that are not part of the background) are $J/\psi \to \mu^+\mu^-$ candidates. The tracks with muon ID have 100% detector efficiency by construction, since muon ID is conferred upon those tracks which register within the muon subsystems. The SVT confirmed tracks, however, are unbiased and therefore perfect for our efficiency study. This method was previously used and described in full by Bloom and Dagenhart [6].

In addition to the above triggering requirements, we require each leg of a $J/\psi \to \mu^+\mu^-$ candidate to have at least 10 COT stereo and axial hits and 2 SVX $r - \phi$ hits. Track pairs are formed only between tracks of opposite charge. The muon track must be matched to a muon stub in either CMU, CMP, or CMX and have $p_T > 3.9$ GeV/c. The SVT track is required to have $p_T > 2.0$ GeV/c and impact parameter, d_0 such that $120\mu m < |d_0| < 1mm$. The SVT track is also required to be fiducial at least one muon subdetector, i.e.: $|\eta| < 1.0$.

The final mass spectrum is shown in Figure 7, where it is fit with a gaussian to model the peak, a second, wider gaussian to model the peaks radiative tail, and a line to model the background. As can be seen in the plot, this fitting scenario models the mass histogram rather well. Using this fit, we can see that we have reconstructed 32869 ± 3844 $J/\psi \to \mu^+\mu^-$ candidates.

Once we have reconstructed the J/ψ candidates, our method of calculating the muon detector efficiencies is similar to our method of calculating the misidentification fractions. In order to calculate the efficiency for a given set of parameters, we begin by creating two mass histograms. Both histograms only represent J/ψ candidates whose SVT daughter tracks are fiducial in a given detector. The first is a histogram of those J/ψ candidates within the required parameters whose SVT tracks are fiducial in a given detector, but do not have a muon ID. The second is a histogram of those J/ψ candidates within the required parameters whose SVT tracks are fiducial within the given detector and do have a muon ID.

We then fit these histograms with a gaussian to model the J/ψ peak and a linear fit to model the background. Specifically, we use the following equations:

$$f_{with}(m) = a(m - 3.0) + b + N_{with}\frac{w}{\sqrt{2\pi}\sigma}e^{-\frac{(m-\mu)^2}{2\sigma^2}} \tag{4}$$

$$f_{without}(m) = a(m - 3.0) + b + N_{without}\frac{w}{\sqrt{2\pi}\sigma}e^{-\frac{(m-\mu)^2}{2\sigma^2}} \tag{5}$$

where μ is the mean of the gaussian, b is flat value of the background, a is the slope of the background, N_{with} and $N_{without}$ are the number of tracks with and without muon ID, w is the histogram's average bin width, and σ is the width of the gaussian. Once we have performed these fits, the fitter returns values for N_{with} and $N_{without}$. Knowing these values, we can determine the efficiency for a given set of parameters, ϵ, as:

$$\epsilon = \frac{N_{with}}{N_{with} + N_{without}} \tag{6}$$

The error for this calculation is determined via simple error propagation:

$$\sigma_\epsilon = \sqrt{\frac{N_{without}^2\sigma_{Nwith}^2 + N_{with}^2\sigma_{Nwithout}^2}{(N_{with} + N_{without})^4}} \tag{7}$$

With this method, we can calculate the overall efficiency with which each muon subsystem identifies muons from our $J/\psi \to \mu^+\mu^-$ candidates. The results of this calculation are shown in Table 1.

Muon subdetector	Total efficiency ($\epsilon \pm \sigma_\epsilon$)
CMU	$70.35 \pm 0.64\%$
CMX	$39.35 \pm 1.01\%$

Table 1: Average detector efficiency

We can use the same method to bin the efficiencies of the CMU and CMX detectors as functions of p_T. This is important, because the p_T distribution of muons from $J/\psi \to \mu^+\mu^-$ decays and the p_T distribution of muons from $D^0 \to \mu^+\mu^-$ decays are not identical. Plots of efficiency vs. p_T can be seen in Figure 8. The dependence of efficiency on track p_T is clear for both detectors.

In order to calculate the expected matching efficiency of our three detector configurations on $D^0 \to \mu^+\mu^-$ daughter tracks, we make use of $D^0 \to \mu^+\mu^-$ Monte Carlo and employ a similar method to the one used in section 5.2 to calculate the expected quantity of background from $D^0 \to \pi^+\pi^-$ double misidentifications. Each MC $D^0 \to \mu^+\mu^-$ candidate's daughter tracks are assigned weights. In this case, the track's weight was equal to the matching efficiency for the track's p_T for the of the detector in which the track was fiducial. The MC $D^0 \to \mu^+\mu^-$ candidate's total weight was equal to the product of the weights of its daughter tracks. The sum of all of the candidates' weights is equal to the number of MC $D^0 \to \mu^+\mu^-$ candidates we would expect to have both legs matched to muon ID's. The results of this calculation are shown in Table 2.

Dectectors	N fiducial in MC	Total weight in MC or N($D^0_{matched}$)	Expected $D^0 \to \mu^+\mu^-$ matching efficiency
CMU-CMU	69500 ± 264	30400 ± 174	0.437 ± 0.00301
CMU-CMX	25200 ± 160	6580 ± 80.5	0.257 ± 0.00360
CMX-CMX	17500 ± 133	2810 ± 53.0	0.161 ± 0.00327

Table 2: Expected average dimuon matching efficiency by detector

4.3 Introduction to relative trigger efficiency

After muon matching efficiency, the other important efficiency calculation for our search is the relative trigger efficiency between $\pi^+\pi^-$ and $\mu^+\mu^-$ events. $\pi^+\pi^-$ and $\mu^+\mu^-$ events have similar kinematic properties, but pions have a roughly 5% inefficiency per track, due mainly to hadronic interactions with the detector. In total, therefore, we expect a relative trigger efficiency of roughly 90%. We calculate this relative efficiency by weighing Monte Carlo distributions based on several track parameters as described below.

4.4 Calculation of relative trigger efficiency

To calculate the relative trigger efficiency, we first create normalized histograms of $D \to K\pi$ Monte Carlo and data binned in several parameters. These parameters included the p_T, η, and ϕ of the D^0 and the p_T, η, ϕ, L_{xy}, number of COT axial hits, number of COT stereo hits, and number of $r-\phi$ hits of the kaon and pion tracks. We unit normalized each of these histograms by dividing the number of entries in each bin by the total number of entries in the histogram. These normalized Monte Carlo and data histograms were then overlaid (see Figures 9 through 13).

Each bin was then assigned a weight (ω_b) according to the ratio between the values of the normalized data and normalized Monte Carlo distribution in that bin. Specifically,

$$\omega_b \equiv \left(\frac{N_{MC}}{N_D}\right)\left(\frac{n_{Db}}{n_{MCb}}\right) \tag{8}$$

$$\sigma_{\omega_b} = \omega_b \times \sqrt{\frac{1}{\sqrt{n_{Db}}} + \frac{1}{\sqrt{n_{MCb}}}} \tag{9}$$

where ω_b is the weight assigned to events within a given bin, b. N_D is the total number of data events, N_{MC} is the total number of Monte Carlo events, n_{Db} is the number of data events within the given bin, and n_{MCb} is the number of Monte Carlo events within the given bin. Bins in which Monte Carlo approximates data rather well (i.e. $\frac{n_{Db}}{N_D} \approx \frac{n_{MCb}}{N_{MC}}$) had a weight of one, whereas poorly approximated bins had larger or smaller weights. See Figures

9 through 13 for a graphical representation of these weighting functions in terms of various parameters.

Once we had weighting conventions for several bins of all of the above parameters, we looked at the $D^0 \to \pi^+\pi^-$ and $D^0 \to \mu^+\mu^-$ Monte Carlo. Each bin in the $D^0 \to \pi^+\pi^-$ and $D^0 \to \mu^+\mu^-$ Monte Carlo samples was given the same weight as determined above using the $D^0 \to K\pi$ Monte Carlo. Furthermore, each event within a given bin of these samples was given the same weight (ω_e) as its bin (ω_b). That is, for all events within a given bin, $\omega_e = \omega_b$. The relative trigger efficiency, ξ, may be calculated as the sum of the weights of all events in the $D^0 \to \mu^+\mu^-$ Monte Carlo divided by the sum of the weights of all events in the $D^0 \to \pi^+\pi^-$ Monte Carlo. Defining the total weight for a Monte Carlo sample, $D \to xx$, as Ω_{xx} as follows:

$$\Omega \equiv \sum_e \omega_e = \frac{N_{MC}}{N_D} \sum_e \frac{n_{De}}{n_{MCe}} \tag{10}$$

$$\sigma_\Omega = \sqrt{\sum_e \sigma^2_{\omega_e}} \tag{11}$$

we find:

$$\xi \equiv \frac{\epsilon(\pi^+\pi^-)}{\epsilon(\mu^+\mu^-)} = \frac{\Omega_{\pi^+\pi^-}}{\Omega_{\mu^+\mu^-}} \tag{12}$$

$$\sigma_\xi = \xi \sqrt{\left(\frac{\sigma_{\Omega_{\pi^+\pi^-}}}{\Omega_{\pi^+\pi^-}}\right)^2 + \left(\frac{\sigma_{\Omega_{\mu^+\mu^-}}}{\Omega_{\mu^+\mu^-}}\right)^2} \tag{13}$$

where $\Omega_{\pi^+\pi^-}$ is the total weight for the $D^0 \to \pi^+\pi^-$ Monte Carlo sample, and$\Omega_{\mu^+\mu^-}$ is the total weight for the $D^0 \to \pi^+\pi^-$.

We performed this study using two different samples of Monte Carlo. In the first sample, we allowed hadrons (e.g. the pions) within our simulation to decay. In the second sample, we did not not allow hadrons to decay. The results from both cases were similar. Table 3 shows the results of our trigger efficiency study with hadronic decay turned "ON.". Table 4 shows the results of our trigger efficiency study with hadronic decay turned "OFF."

From the largest discrepancies from the central values found in Tables 3 and 4, we calculate the relative trigger efficiency, $\frac{\Omega_{\pi^+\pi^-}}{\Omega_{\mu^+\mu^-}}$ to be $0.872 \pm 0.001 \pm 0.015$. Note that the chief source systematic uncertainty is due to poor modelling of the number of COT hits by the Monte Carlo. This is an effect of the Monte Carlo generation of only D^0 signal decays, without the underlying event and other detector noise. Figures 9 and 10 provide an excellent illustration for this discrepancy between Monte Carlo and data.

5 Muon Misidentification Rates

In this section, we discuss our methods for muon identification and how we applied those methods to determine the misidentification fractions of kaon and pion tracks. These rates are important, because they allow us to predict the amount of background $\mu^+\mu^-$ events we

Parameter	$\Omega_{\pi^+\pi^-}$	$\Omega_{\mu^+\mu^-}$	$\Omega_{\pi^+\pi^-}/\Omega_{\mu^+\mu^-}$
$p_T(D)$	102800 ± 300	121900 ± 340	0.844 ± 0.00357
$\eta(D)$	101900 ± 300	119100 ± 340	0.866 ± 0.00365
$L_{xy}(D)$	101300 ± 300	117900 ± 340	0.869 ± 0.00368
$p_T(\pi)$	99200 ± 300	115800 ± 340	0.857 ± 0.00371
$\eta(\pi)$	101900 ± 300	119200 ± 340	0.855 ± 0.00365
$\phi(\pi)$	102000 ± 300	119200 ± 340	0.856 ± 0.00365
n COT Ax hits (π)	102900 ± 300	113800 ± 340	0.904 ± 0.00389
n COT St hits (π)	100500 ± 300	115500 ± 340	0.870 ± 0.00375
n Silicon hits (π)	102500 ± 300	119500 ± 340	0.858 ± 0.00365
$p_T(K)$	101600 ± 300	118500 ± 340	0.857 ± 0.00366
$\eta(K)$	101800 ± 300	119000 ± 340	0.856 ± 0.00365
$\phi(K)$	102100 ± 300	119200 ± 340	0.865 ± 0.00367
n COT Ax hits (K)	105900 ± 300	118000 ± 340	0.897 ± 0.00380
n COT St hits (K)	97900 ± 300	114600 ± 340	0.854 ± 0.00372
n Silicon hits (K)	101200 ± 300	118100 ± 340	0.857 ± 0.00367

Table 3: Relative trigger efficiency with hadronic decay "ON"

can expect from non-dimuon decays. As a cross-check of these values, we estimated the total amount of background we expect from double-faked $D^0 \to \pi^+\pi^-$ events using two methods. The first involved multiplying average fake rates by the number of $D^0 \to \pi^+\pi^-$ fiducial in a given detector configuration by the average fake rate of each detector within that configuration. The second method involved a weighting system dependent on the kinematic properties of each $D^0 \to \pi^+\pi^-$ candidate. These methods along with the expected amount of background from double faked $D^0 \to \pi^+\pi^-$ are described in further detail in Section 5.4.

5.1 Muon identification

We identify tracks as muon tracks within the CMU and CMX muon subsystems using the following method. First, a track is required to be fiducial in a given detector. Tracks fiducial in CMU have $0.0 < |\eta| < 0.6$ and tracks fiducial in CMX have $0.0 \geq |\eta| < 1.0$. We also require that the track be matched to a muon stub within the given detector.

5.2 Muon misidentification fractions

The muon misidentification fraction is the fraction of pion or kaon tracks fiducial in a given muon detector that are assigned muon ID's. The average misidentification fraction for pion and kaon tracks is on the order of 1%. It is important to have a very thorough understanding of these misidentification fractions, since $D^0 \to KK$,$K\pi$,$\pi^+\pi^-$ events (among others) with both legs misidentified as muons are a key source of background when we calculate $N(\mu^+\mu^-)$ from Equation 3.

Parameter	$\Omega_{\pi^+\pi^-}$	$\Omega_{\mu^+\mu^-}$	$\Omega_{\pi^+\pi^-}/\Omega_{\mu^+\mu^-}$
$p_T(D)$	104900 ± 300	120100 ± 340	0.873 ± 0.00369
$\eta(D)$	103600 ± 300	119100 ± 340	0.867 ± 0.00369
$L_{xy}(D)$	102500 ± 300	117600 ± 340	0.872 ± 0.00372
$p_T(\pi)$	100700 ± 300	115600 ± 340	0.871 ± 0.00375
$\eta(\pi)$	103100 ± 300	119100 ± 340	0.866 ± 0.00369
$\phi(\pi)$	104900 ± 300	119100 ± 340	0.881 ± 0.00372
n COT Ax hits (π)	102000 ± 300	116600 ± 340	0.875 ± 0.00375
n COT St hits (π)	103400 ± 300	114700 ± 340	0.902 ± 0.00387
n Silicon hits (π)	104100 ± 300	119500 ± 340	0.872 ± 0.00369
$p_T(K)$	103100 ± 300	118400 ± 340	0.871 ± 0.00371
$\eta(K)$	103500 ± 300	118950 ± 340	0.870 ± 0.00370
$\phi(K)$	103700 ± 300	119000 ± 340	0.871 ± 0.00370
n COT Ax hits (K)	107000 ± 300	118800 ± 340	0.901 ± 0.00379
n COT St hits (K)	102000 ± 300	116000 ± 340	0.880 ± 0.00378
n Silicon hits (K)	102800 ± 300	118100 ± 340	0.870 ± 0.00371

Table 4: Relative trigger efficiency with hadronic decay "OFF"

5.3 $D^{*+} \to D^0\pi^+$, $D^0 \to K\pi$ decays

$D^{*+} \to D^0\pi^+$, $D^0 \to K\pi$ decays provide a very clean sample to measure muon misidentification fractions for kaons and pions. We begin reconstructing D^{*+} decays by reconstructing D^0 decays. We do this by looking at track pairs, one track of which we assign the mass of a kaon while we assign the other the mass of a pion (a "$K\pi$ mass hypothesis"). In addition, we require that the daughter tracks of D^0 candidates must pass the following specifications to be considered.

- CMX good run bits must be set in addition to usual good run requirements.
- Each track must have 10 COT axial and stereo hits and 2 at least silicon $r - \phi$ hits.
- Since the decay in question is a two-body decay from a neutral meson, all track pairs must be composed of tracks with opposite charge.
- The invariant mass of the tracks must be between 1.7 GeV/c^2 (the limit at which $D^0 \to KK$ appear) and 4.0 GeV/c^2 (the limit at which background from B^0 meson decays can be measured).
- p_T is requried to be greater than 2.0 GeV/c for all tracks.

The resulting mass peak is shown in Figure 4. We can further clean this distribution by adding a third track to the existing topology in an attempt to reconstruct the D^{*+} mass. To do this, we require that there exists one track for every D^0 candidate that meets the following specifications:

- The new track must have charge opposite to that of the track assigned kaon mass.
- The difference between these three tracks' invariant mass and the invariant mass of the original two tracks must be such that 0.144 GeV/c^2 > m(D^{*+} − D^0) > 0.147 GeV/c^2. The mass difference spectrum is shown in Figure 4.

This last set of criteria, known as D^* tagging, completes the selection process and gives a very clean sample of roughly 0.91×10^6 $D^0 \to K\pi$ candidates. The final D^0 mass spectrum also appears in Figure 4.

5.4 Muon misidentification fraction measurement

Once we have a clean sample of $D^0 \to K\pi$ decays, we may calculate the muon misidentification fraction. The method employed for calculating this fraction is the same as the one described in CDF Note 8042 [4].

To begin, two mass histograms are considered for every bin of a parameter under study. The first histogram is that of D^0 candidates whose daughter pions do not have muon ID. The second histogram is that of D^0 candidates whose daughter pions do have muon ID. Since the amount of statistics available in every bin can be significantly smaller than that of the whole sample, the fitting model is simplified with respect to the fit of the entire sample. The background mass spectrum is modelled with a simple linear function, and the signal peak is modelled with a single gaussian distribution. Specifically, the fit functions for the histograms are:

$$f_{with}(m) = a(m - 1.865) + b + N_{with}\frac{w}{\sqrt{2\pi}\sigma}e^{-\frac{(m-\mu)^2}{2\sigma^2}} \tag{14}$$

$$f_{without}(m) = a(m - 1.865) + b + N_{without}\frac{w}{\sqrt{2\pi}\sigma}e^{-\frac{(m-\mu)^2}{2\sigma^2}} \tag{15}$$

where μ is the mean of the gaussian, b is flat value of the background, a is the slope of the background, N_{with} and $N_{without}$ are the number of tracks with and without muon ID, w is the histogram's bin width, and σ is the width of the gaussian. The misidentification fraction, ζ, is calculated as:

$$\zeta = \frac{N_{with}}{N_{with} + N_{without}} \tag{16}$$

The uncertainty on the misidentification fraction is determined by simple error propagation:

$$\sigma_\zeta = \sqrt{\frac{N_{without}^2\sigma_{Nwith}^2 + N_{with}^2\sigma_{Nwithout}^2}{(N_{with} + N_{without})^4}} \tag{17}$$

The process for determining the misidentification fraction of kaons is identical to the above process for pions.

We used the above "fit" method to measure the misidentification fraction of positively charged and negatively charged daughter pion and kaon tracks. We also measured the misidentification fraction of pions and kaons without regard to charge. We calculated misidentification fractions for pion tracks fiducial in CMU and CMX seperately. The results for pions and kaons are shown in Tables 5 and 6, respectively.

	π^-	π^+	Measured avg π	Computed avg π
CMU	$0.501\% \pm 0.0170\%$	$0.531\% \pm 0.0177\%$	$0.518\% \pm 0.0123\%$	$0.518\% \pm 0.0123\%$
CMX	$0.147\% \pm 0.0190\%$	$0.134\% \pm 0.0183\%$	$0.136\% \pm 0.0129\%$	$0.140\% \pm 0.0132\%$

Table 5: Pion misidentification rates

	K^-	K^+	Measured avg K	Computed avg K
CMU	$0.705\% \pm 0.0170\%$	$1.322\% \pm 0.0216\%$	$1.020\% \pm 0.0138\%$	$1.014\% \pm 0.0137\%$
CMX	$0.242\% \pm 0.0153\%$	$0.284\% \pm 0.0158\%$	$0.266\% \pm 0.113\%$	$0.263\% \pm 0.011\%$

Table 6: Kaon misidentification rates

In addition, we measured kaon and pion misidentification fractions as functions of p_T, η, and phi_0. The results of these measurements may be seen in Figures 5 and 6. Negative kaons have a much higher scattering cross-section and are much less likely to punch through to the muon detectors than positive kaons are. This results in a lower fake rate for K^- than for K^+, as seen in these figures.

After calculating the misidentification fractions for both muon subdetectors, we can estimate the number of $D^0 \to \pi^+\pi^-$ candidates that will have both daughter tracks fake muons. To do this, we find the number of $D^0 \to \pi^+\pi^-$ candidates with daughter tracks fiducial in a given pair of subdetectors. We obtained this value by applying a $\pi^+\pi^-$ mass hypothesis to our dataset and making a histogram of the two tracks' invariant mass. This histogram was then fit with an equation similar to Equations 14 and 15 (a gaussian with a linear background).

$$f_{fiducial}(m) = a(m - 1.865) + b + N_{fiducial}\frac{w}{\sqrt{2\pi}\sigma}e^{-\frac{(m-\mu)^2}{2\sigma^2}} \tag{18}$$

The parameters in Equation 18 have similar meaning to those in Equations 14 and 15. a and b, are defined as slope and verticle offset, respectively, of the linear background, and μ and σ defined as the mean and width of the gaussian. w is the bin width of the histogram. $N_{fiducial}$ is the number of $D^0 \to \pi^+\pi^-$ fiducial in the given detector configuration. We used three detector configurations, looking at candidates with both tracks fiducial in CMU ("CMU-CMU" configuration), candidates with one track fiducial in CMU and one track fiducial in CMX ("CMU-CMX" configuration), and candidates with both tracks fiducial in CMX ("CMX-CMX" configuration). As a cross-check, we estimated the total number of doubly-misidentified $D^0 \to \pi^+\pi^-$ candidates using two methods.

The first, more simple method merely involved taking the product of the number $D^0 \to \pi^+\pi^-$ candidates in each detector configuration with the average pion misidentification rate of each detector in that configuration.

In the second method, we binned the pion misidentification rates in pion p_T using the original method with $D^0 \to K\pi$ candidates described above (see Figures 6 and 5). We then ran through $D^0 \to \pi^+\pi^-$ Monte Carlo and assigned each daughter pion track a "weight" equal to its expected misidentification fraction according to its p_T. The total weight of a $D^0 \to \pi^+\pi^-$ MC candidate was taken as the product of the weights of the daughter pions:

$$\omega_i = \zeta_{1i} \cdot \zeta_{2i} \tag{19}$$

$$\sigma_{\omega_i} = \omega_i \cdot \sqrt{\left(\frac{\sigma_{\zeta_{1i}}}{\zeta_{1i}}\right)^2 + \left(\frac{\sigma_{\zeta_{2i}}}{\zeta_{2i}}\right)^2} \tag{20}$$

The sum of all of the weights of the MC candidates in a detector configuration was equal to the expected number of doubly misidentified tracks in Monte Carlo.

$$N_{Doubletag} = \sum_{i=0}^{N(D^0 \to \pi^+\pi^-)} \zeta_{1i} \cdot \zeta_{2i} \tag{21}$$

$$\sigma_{N_{Doubletag}} = \sqrt{\sum \omega_{\sigma_i}^2} \tag{22}$$

We expected the results of these two methods to agree. The results of both procedures are described in Table 7. Note that the results from both methods agree within two sigma.

Detectors	N fiducial in MC	N exp. double misID's in MC (avg)	N exp. double misID's in MC (wt)
CMU-CMU	69500 ± 264	1.860 ± 0.062	1.740 ± 0.001
CMU-CMX	25200 ± 160	0.177 ± 0.017	0.179 ± 0.0004
CMX-CMX	17500 ± 130	0.032 ± 0.004	0.0354 ± 0.0002

Table 7: N expected fake dimuon events from misidentified $D^0 \to \pi^+\pi^-$ in MC

Knowing the number of $D^0 \to \pi^+\pi^-$ candidates we could expect to be misidentified as $D^0 \to \mu^+\mu^-$ in Monte Carlo, we then extrapolated our calculation to data. We did this by simply comparing the number of $D^0 \to \pi^+\pi^-$ candidates we had in each detector configuration in data to that number in Monte Carlo as follows:

$$N_{\text{double misID in Data}} = N_{\text{double misID in MC}} \cdot \frac{N_{\text{fiducial in Data}}}{N_{\text{fiducial in MC}}} \tag{23}$$

and σ_N is determined by simple error propagation. The results of these calculations are listed in Table 8. Again, note that the results of the two methods agree within two sigma.

6 Definition of Signal Search Windows

The $D^0 \to \mu\mu$ decay has a very simple signature. The invariant mass of the two-track system has to be consistent with the world average D^0 mass, and the two tracks have to be matched to muon stubs in the muon chambers.

Detectors	N fiducial in Data	N exp. double misID's in Data (avg)	N exp. double misID's in Data (wt)
CMU-CMU	24400 ± 199	0.654 ± 0.023	0.611 ± 0.005
CMU-CMX	25200 ± 160	0.0651 ± 0.006	0.0658 ± 0.001
CMX-CMX	17500 ± 133	0.0128 ± 0.002	0.0140 ± 0.0003

Table 8: N expected fake dimuon events from misidentified $D^0 \to \pi^+\pi^-$ in data

Detector Configuration	$D^0 \to \pi^+\pi^-$ Yield
CMU-CMU	24400 ± 200
CMU-CMX	9260 ± 130
CMX-CMX	6940 ± 110

Table 9: Yield of untagged $D^0 \to \pi\pi$ decays in different detector configurations

In the analysis, muon identification via likelihood requirements will be required on both tracks in the decay. The signature of the signal decay is an excess of muon pairs peaking at the world average D^0 mass. Due to tracking resolution effects, the mass distribution will be roughly Gaussian. The signal window is chosen to be in the invariant mass of the dimuon pair. It has to be chosen to be wide enough to accept signal with good efficiency, and at the same time, the narrower the search window, the lower the expected background level. One can choose a roughly 95% acceptance window by eye, from 1.845 to 1.890 GeV/c^2.

Figure 14 shows the mass distributions of Monte Carlo simulated $D^0 \to \mu\mu$ decays in the CMU-CMU, CMU-CMX, and CMX-CMX detector configurations. The chosen signal mass window from 1.840 to 1.885 GeV/c^2 provides good acceptance for signal events, as listed in Table 10.

7 Background Sources

There are several different sources of background that need to be considered in this search. The previous analysis [5] has already noticed an unusually large rate of double-tags in the high-mass range. At the time, it was suspected to be a detector effect, due to proximity of the CMU stubs from the two candidates.

With increased sample statistics, there is a chance to study this source of background

Detector Configuration	Signal Window Efficiency
CMU-CMU	[$98.0 \pm 0.1(stat)$] %
CMU-CMX	[$97.4 \pm 0.1(stat)$] %
CMX-CMX	[$96.7 \pm 0.2(stat)$] %

Table 10: Signal window efficiency for different detector configurations

System	Data	Monte Carlo
CMU-CMU	[$95.3 \pm 0.1(stat)$] %	[$95.6 \pm 0.1(stat)$] %
CMU-CMX	[$95.0 \pm 0.2(stat)$] %	[$95.6 \pm 0.1(stat)$] %
CMX-CMX	[$92.7 \pm 0.3(stat)$] %	[$94.8 \pm 0.2(stat)$] %

Table 11: Signal mass window efficiency for $D^0 \rightarrow \pi^+\pi^-$ decays, as calculated from data fits and Monte Carlo simulation.

and try to find if there is underlying physics causing it. Therefore, the purpose of the following sections is to attempt to quantify the expected dimuon rates from different sources of background.

Fundamentally, one can divide the sources of background into two classes: backgrounds due to real muons and backgrounds due to mistags. Obviously, any decay involving a real muon, even if its absolute rate is low, is much more likely to appear in the double-tagged sample because the typical muon efficiency is around 50-70%, while the mis-tag rates for non-muons are below the 1% level. Therefore, we expect an acceptance rate of about 1/10000 for track pairs which are not composed of muons, 1/150 for track pairs in which one track is a real muon, and about 1 in 2 for real dimuons.

The expected background contributions are divided as follows. Backgrounds due to double mistags which are considered are backgrounds due to double-mistagged $D^0 \rightarrow \pi\pi$ decays, double-mistagged $D^0 \rightarrow K\pi$ decays, and double-mistagged combinatorial backgrounds. The term "combinatorial" is slightly more restrictive than usual - a decay is considered "combinatorial" if neither of the two tracks are real muons, but it ends in the double-tagged sample because both tracks are mis-tagged as muons.

Backgrounds due to real muons are broken down into three categories. Backgrounds from semileptonic charm decays are dominated by semileptonic D^0 decays, because other charm decays are suppressed via D^* tagging. Semileptonic B meson decays are another source of real dimuons; within these, decays with one real muon in the final state are treated separately from decays with two real muons in the final state. The following sections discuss the estimates for all the categories listed above.

7.1 Double-Tagged $D^0 \rightarrow \pi\pi$ Decays

The contribution from $D^0 \rightarrow \pi\pi$ decays which have been mis-tagged as dimuon decays is the easiest to calculate. In Section 5, the pion misidentification rates have been calculated. In order to calculate the double-tag misidentification rate, one needs to know the number of $D^0 \rightarrow \pi\pi$ decays that make it into the signal peak, and the misidentification rate for pions.

The efficiency of accepting a $D^0 \rightarrow \pi\pi$ decay into the signal mass window can be calculated from Monte Carlo simulation, or from the data directly. Figures 15 and 16, show the invariant mass of the $D^0 \rightarrow \pi\pi$ decays reconstructed as dimuon decays in data and Monte Carlo simulation, respectively. The corresponding efficiencies of accepting di-pion decays into the signal mass window are listed in Table 11.

The simplest way to calculate the double-tag rate is then to multiply the number of $D^0 \rightarrow \pi\pi$ candidates with the average misidentification rates for positive and negative pions.

System	$D^0 \to \pi^+\pi^-$ Count	Simple Double-Tag Est.	Folding Kinematics
CMU-CMU	24400 ± 199	$0.65 \pm 0.02(stat)$	$0.611 \pm 0.005(stat)$
CMU-CMX	9360 ± 130	$0.065 \pm 0.006(stat)$	$0.066 \pm 0.001(stat)$
CMX-CMX	6940 ± 110	$0.013 \pm 0.002(stat)$	$0.014 \pm 0.003(stat)$

Table 12: Expected rate of double-tags from $D^0 \to \pi\pi$ decays, using the simple calculation, and accounting for kinematic differences between the mistag calibration sample and the target sample.

System	Expected Double-Tag Count
CMU-CMU	$0.61 \pm 0.05\,(tot)$
CMU-CMX	$0.066 \pm 0.001\,(tot)$
CMX-CMX	$0.014 \pm 0.001\,(tot)$

Table 13: Expected rate of double-tags from $D^0 \to \pi\pi$ decays, with total (statistical + systematic) uncertainty.

This makes the assumption that the momentum and pseudorapidity spectra of $D^0 \to K\pi$ and $D^0 \to \pi\pi$ decays are the same, which is a very reasonable approximation given the trigger and selection for both decays is the same. Table 12 lists the results of such a calculation, using only simple error propagation to calculate statistical uncertainties.

A slightly more advanced approach also takes into account the fact that the kinematic properties of the target sample ($D^0 \to \pi\pi$) are somewhat different from those of the calibration sample ($D^0 \to K\pi$). The most dramatic change in mistag rate is observed when studying the mistag rate as a function of track p_T. The expected rate of double-tagged fakes can then be re-calculated by weighting every event in the Monte Carlo sample with the corresponding pion mistag rate from Figure 5. The results of this calculation are also listed in Table 12. The estimate of the mistag rate is taken from the simple calculation performed on data. The relative systematic uncertainty is taken to be the relative difference between the results obtained using the simple and advanced calculations on the Monte Carlo simulation, and is listed in Table 13.

7.2 Double-Tagged $D^0 \to K\pi$ Decays

The contribution of double-tagged $D^0 \to K\pi$ decays can be determined in a similar fashion to $D^0 \to \pi^+\pi^-$ decays. $D^0 \to K\pi$ decays can appear in our search window if the reconstructed $\mu^+\mu^-$ mass of the $K\pi$ pair falls into the search window and both tracks are tagged as muons. In order to estimate the frequency of this happening, we reconstruct $D^0 \to K\pi$ decays in the $\mu^+\mu^-$ mass hypothesis and count how many show up in the signal search window. We then compare the number of candidates reconstructed in the $\mu^+\mu^-$ hypothesis in the signal window, $N_{MC}^{\mu^+\mu^-}$, to the number of candidates reconstructed in the $K\pi$ hypothesis in the signal window, $N_{MC}^{K\pi}$, and the number of $K\pi$ candidates in data, $N_{data}^{K\pi}$. The expected rate

Detectors	MC, $K\pi$ hypo	MC, $\mu^+\mu^-$ hypo	Data, $K\pi$ hypo	exp. 2-tag rate
CMU-CMU	131958	36	558000 ± 4000	0.0070
CMU-CMX	44585	13	208000 ± 1000	0.0011
CMX-CMX	36627	9	174000 ± 1300	0.0001

Table 14: Rates of Monte Carlo simulated events found in the signal mass range in the $K\pi$ and $\mu^+\mu^-$ mass hypotheses, as well as the number of $D^0 \to K\pi$ decays observed in data for the same detector configurations.

of $K\pi$ decays is then:

$$N_{fake}^{K\pi} = \frac{N_{MC}^{\mu^+\mu^-}}{N_{MC}^{K\pi}} \cdot N_{data}^{K\pi} \cdot f_K \cdot f_\pi, \tag{24}$$

where f_K is the average kaon mistag rate and f_π is the average pion mistag rate.

Figure 17 shows the distributions of $D^0 \to K\pi$ decays reconstructed in the $K\pi$ mass hypothesis. Figure 18 shows the same distribution for data. Figure 19 shows the invariant mass of the $D^0 \to K\pi$ Monte Carlo candidates reconstruced in the $\mu^+\mu^-$ mass hypothesis. The signal search window is superimposed on all the distributions. Table 14 lists the rates of Monte Carlo simulated events found in the signal mass range in the $K\pi$ and $\mu^+\mu^-$ mass hypotheses, as well as the number of $D^0 \to K\pi$ decays observed in data for the same detector configurations.

One could be concerned about the fact that the Monte Carlo simulation likely does not simulate the tails of the invariant mass distribution very precisely; still, the expected rate is so minute that the amount of tails should be at least an order of magnitude higher to compare to the $D^0 \to \pi^+\pi^-$ background.

7.3 Double-Tagged Combinatorial Background

This category of backgrounds contains the most generic decays which do not contain real muons, but will end up in the search window because the invariant mass of the track pair falls inside the search window, and tracks were mis-tagged as being muons. When considering that these decays can span the gambit from hadronic B decays, correlated $c - \bar{c}$ production resulting in D mesons decaying with kaons and pions in the final state, all the way to random tracks, it may immediately seem like estimating their contribution precisely will be a daunting task. Luckily, a few things work to the advantage of this analysis. Firstly, the signal that we are reconstructing is not just two tracks which fall into the correct mass window. We require that the decay is D^*-tagged. This puts an extremely narrow window on a three-track topology which is very difficult to mimic with random track combinations. As we have seen in Figure16, the ratio of $D^0 \to \pi^+\pi^-$ to background in the search window is roughly 3:1, so the amount of combinatorial background contribution to double-tags in the signal region is comparable to the double-tag rate from $D^0 \to \pi\pi$ decays to first order.

A more rigorous estimate of the contribution of the combinatorial background to the double-tag rate in the signal search window is provided. In addition to misidentified kaons and pions, as we have already seen, there are real di-muons in the $D^0 \to \mu\mu$ background. The

Detectors	Raw Count	15% kaon	5% kaon	25% kaon	100% kaon
CMU-CMU	4000	0.14	0.12	0.17	0.41
CMU-CMX	2030	0.028	0.024	0.033	0.084
CMX-CMX	1047	0.0025	0.0021	0.0029	0.0074

Table 15: Combinatorial background double-mistag rates assuming different fraction of tracks are kaons (as opposed to pions).

Decay Mode	World Average Br	EvtGen Br
$D^0 \to K^- \mu^+ \nu_\mu$	3.19 ± 0.16 %	3.22%
$D^0 \to K^{*-} \mu^+ \nu_\mu$	1.95 ± 0.25 %	1.98%
$D^0 \to \pi^- \mu^+ \nu_\mu$	$2.4 \pm 0.4 \times 10^{-3}$	0.37%
$D^0 \to K^- \pi^+$	3.80 ± 0.07 %	3.83 %
$D^0 \to \pi^- \pi^+$	$1.364 \pm 0.032 \times 10^{-3}$	0.15 %

Table 16: Branching fractions of relevant semileptonic D^0 decays, reference decays and values used when generating Monte Carlo through EvtGen.

contribution of these decays can be removed by simply anti-selecting events in which both tracks are identified as muons. Since the detector muon identification has high efficiency (90% or better for muons fiducial in the CMU) and low mistag rate (comparable to or less than 1%), the remaining events are an excellent template of combinatorial background. The only question is that of track composition - in Table 15, several scenarios are explored, depending on the kaon content. The central value of the double-tag rate is taken to be the one with 85% pion and 15% kaon tracks. The kaon fraction is varied by ±10% to obtain the systematic uncertainty. As a curiosity, the double-tagged rate is also listed for the case that both tracks in the combinatorial background aere always kaons, which is a vast exaggeration. Figures 30, 31, and 32 show the mass distribution of the combinatorial background templates. The mass range from 1.84 to 1.89 GeV/c^2 is cut out in the plots because this is the signal search window.

7.4 Semileptonic D Decays

In addition to $D^0 \to \pi\pi$ decays ending in the search window, there is also a chance of semileptonic D^0 decays involving a muon appearing in the search window. The kinematics and relevant branching fractions of semileptonic D^0 decays is well known from various measurements [3]. The world average branching fractions for interesting semileptonic D^0 decays, as well as for the reference branching fractions, are listed in Table 16.

The most likely candidates to contribute inside the signal search window are $D^0 \to \pi^- \mu^+ \nu_\mu$ decays because of the largest energy available to the two tracks in the decay. Since the properties of semileptonic D^0 decays are well known, we can use Monte Carlo simulation to predict the invariant mass of semileptonic decays. By generating a mixture of D^0 decays and semileptonic decays in our Monte Carlo sample, we also obtain reference samples of

$D^0 \to K\pi$ and $D^0 \to \pi\pi$ decays for absolute normalization of the rate of semileptonic decays. To be precise, we generate one sample of decays in which the D^0 is forced to decay semileptonically ("enriched sample"), and another sample of decays in which the D^0 is free to decay any which way ("reference sample").

Figures 23, 25 and 27 show the invariant mass distributions of semileptonic D decays from Monte Carlo simulation, for different subdetector configurations, in the enriched sample. The signal search windows are also shown on the figures. Figures 24, 26, and 28 show the mass spectra of the reference sample, with the corresponding $D^0 \to \pi\pi$ peak superimposed.

In this particular case, we are concerned about the situation in which one of the two tracks composing the D^0 candidate is a real muon. Therefore, it is required that one track in the reconstructed candidate be matched to a real muon in the Monte Carlo simulation for all candidates in Figures 23, 25 and 27. The mass distributions are broken down according to the particle type that is matched to the other decay track. The absolute rate of semileptonic D^0 decays in the search window can then be calculated by normalizing the Monte Carlo reconstructed mass spectrum to the calibration sample ($D^0 \to \pi\pi$), and applying the appropriate detector efficiency to every track matched to a muon, and mistag rate to tracks not matched to muons.

Two effects are observed in this study. First, there are no semileptonic decays D^0 decays which end with two real muons. This is very useful because it implies also that there is no source of real dimuons that could contaminate the signal mass window. The second effect observed is that the kinematic cutoff for semileptonic decays in which the other track is either a kaon or a pion is below the signal mass window. We observe no candidates from such decays in the signal mass window in either of the three detector configurations. However, our semileptonic-only sample corresponds to roughly one tenth of the statistics of the data. It is therefore not impossible that with more statistics, we would start to see candidates with one track matched to a real muon and one track matched to a kaon or pion. We would like to quantify what we can learn from not observing any decays in our Monte Carlo in terms of an exclusion limit in the data.

The principle of the calculation follows from the Poisson probability that observing zero events in a sample in which the expected number of 3 events is less than 5%. One can then scale $N^{big}_{\mu X} = 3$ events to the number of events that we would see in data N_{data}:

$$N^{data}_{\mu^+\mu^-} = f^K_{mistag} \cdot \epsilon^\mu \cdot \frac{N^{big}_{\mu X}}{N^{big}_{semil}} \cdot \frac{N^{ref}_{semil}}{N^{ref}_{\pi^+\pi^-}} \cdot N^{data}_{\pi^+\pi^-} \tag{25}$$

where f^K_{mistag} is the average kaon mistag rate, ϵ^μ is the muon identification efficiency, N^{big}_{semil} is the number of semileptonic decays reconstructed in the enriched sample, $N^{ref}_{\pi^+\pi^-}$ and N^{ref}_{semil} are the number of $\pi^+\pi^-$ and semileptonic decays , respectively, reconstructed in the reference sample, and $N^{data}_{\pi^+\pi^-}$ is the number of $\pi^+\pi^-$ decays reconstructed in data. The actual values used in the calculation are listed in Table 17.

In this technique, there is no mechanism to propagate uncertainties due to counting or measurement of all the numbers involved in this scaling, the approach is actually inverted. A frequentist technique is used to set a limit on $N^{data}_{\mu^+\mu^-}$. For a probe value of $N^{data}_{\mu^+\mu^-}$, the corresponding $N^{big}_{\mu X}$ is computed, and the uncertainty on $\sigma^{big}_{\mu X}$ is obtained by error propagation through Equation 25, assuming Gaussian uncertainties. Outcomes of pseudo-experiments

Table 17: Values used in the calculation of the limit on the number of dimuons caused by semileptonic D^0 decays.

Detectors	f^K_{mistag}	N^{big}_{semil}	$N^{ref}_{\pi^+\pi^-}$	N^{ref}_{semil}	$N^{data}_{\pi^+\pi^-}$
CMU-CMU	$(11.8 \pm 3.4) \cdot 10^{-3}$	14193	241	1571	24400 ± 200
CMU-CMX	$(7.7 \pm 2.2) \cdot 10^{-3}$	3414	75	363	9260 ± 130
CMX-CMX	$(3.5 \pm 0.4) \cdot 10^{-3}$	4856	64	525	6940 ± 110

Detectors	Num. Double Tags
CMU-CMU	< 0.48; 95% CL
CMU-CMX	< 0.37; 95% CL
CMX-CMX	< 0.14; 95% CL

Table 18: Upper limits on the semileptonic D^0 decay contribution to the double-tagged rate in the signal mass window.

are thrown following the Poisson distribution. The parameter of the Poisson distribution is Gaussian smeared around $N^{big}_{\mu^+\mu^-}$ with an RMS of $\sigma^{big}_{\mu X}$. The p-value is computed as the fraction of pseudo-experiments in which zero events are observed. The probe $N^{data}_{\mu^+\mu^-}$ for which the p-value is less than 5% is considered the 95% upper limit on the number of dimuons from semileptonic D meson decays. The distributions of p-values as a function of probe $N^{data}_{\mu^+\mu^-}$ for the three subdetector configurations are shown in Figure 29. In Equation 25, the average kaon mistag rate is used because one can roughly assume charge symmetry in the detector at these energies, and the second track in the semileptonic decays is dominantly a kaon track. This is by all means a conservative estimate. The estimated upper limits at the 95% CL level for the contribution of semileptonic D^0 decays are listed in Table 18.

7.5 Semileptonic B Decays

The calculation of the semileptonic B decay contribution to the total background rate is slightly more complicated than the previous contributions. It is an exercise that builds confidence in the estimate of the real dimuon contribution to the double-tag rate.

In order to estimate the contribution of semileptonic B decays to the double-tag rate, a region of invariant mass inaccessible to D-meson decays, 1.9-4.0 GeV/c^2 is studied. Semileptonic decays are expected to be detected in the sample in which at least one of the decay tracks is tagged as a muon.

Figures 20, 21, and 22 show the same distributions for Monte Carlo simulation. The mass distributions are divided according to what type of particle the other track (which was not required to be a muon) was matched to.

From the mass distribution of combinatorial background, the rate of mis-tagged decays can be estimated. Assuming everything else with at least one muon-tagged track are semileptonic B meson decays, one can predict the total event count in this mass range, and compare it to data. Figures 33, 34 and 35 show the comparison of the predicted and measured mass

Detectors	$1-\mu$ Sample	$2-\mu$ Sample
CMU-CMU	2.9 ± 0.1	1.8 ± 0.1
CMU-CMX	0.81 ± 0.05	0.54 ± 0.06
CMX-CMX	0.58 ± 0.05	0.24 ± 0.04

Table 19: Calculation of the semileptonic B decay contribution, specifically from decays in which one track is a real muon and the other is mistagged as a muon to the double-tagged rate in the signal mass window.

spectra. In general there is reasonably good agreement. This comes as no big surprise, if one recalls that the signature which we are looking for, a D^* tagged decay with one muon matched in the detector, in the mass range accessible to B meson decays, is complex enough that it is difficult to fake with random track combinations.

The discrepancies in the region of low invariant mass indicate that there is likely a systematic disagreement between the data and the Monte Carlo prediction. However, from the Monte Carlo compositions shown in Figures 20, 21, and 22 two predictions can be inferred. The first prediction is the rate of double-tagged semileptonic B decays in which one track is a real muon and the other is a mis-tagged track. These rates are listed in Table 19, and are found to be very small. The other prediction that can be made is the rate of real dimuons that should be seen in the detector, due to B decays, which will be discussed in the next section.

7.6 Real Dimuons from B Decays

Semileptonic decays of B mesons, $B \rightarrow D\mu\nu_\mu X$, followed by semileptonic decays of D mesons, $D \rightarrow \mu\nu_\mu Y$ will result in oppositely charged dimuons in the final state. In addition to these decays, there are also B meson decays involving $J/\psi \rightarrow \mu^+\mu^-$ and $\psi' \rightarrow \mu^+\mu^-$ decays in the final state. All of these are likely to end up in the final double-tagged sample. As mentioned before in Section 7.5, one of the predictions of the single-muon sample composition is also the absolute rate of real dimuons that one should see in the sample. One can obtain that by multiplying the rate of single-muon decays in which the other track was also matched to the muon by the muon matching efficiency. The outcome of this calculation is listed in Table 20.

An alternate calculation is also possible, in which the driving assumption is that the number of double-tags due to single-muons and double-mistags can be predicted from Monte Carlo, and everything else in the high-mass sample are real dimuons from B decays. The agreement of these predictions with the data is shown in Figures 36, 37 and 38. The dimuon rates in the signal region predicted by the two calculations are listed in Table 20. Finally, as a crosscheck, we also fit the mass range between 2.0 and 2.9GeV/c^2 in data and normalized Monte Carlo with a flat distribution and extrapolate into the signal mass range. The outcomes of these crosschecks are also listed in Table 20.

One should not immediately trust that the main source of backgrounds are due to $B \rightarrow \mu\mu X$ decays due to just the mass distribution matching between data and Monte Carlo. There are several additional properties that one can check between the $B \rightarrow \mu\mu X$ Monte

Detector	$1-\mu$ Sample	$2-\mu$ Sample	Data, fit	MC, fit
CMU-CMU	19.5 ± 3.2	17.7 ± 2.9	15.6 ± 0.9	17.7 ± 0.9
CMU-CMX	5.4 ± 1.3	5.3 ± 1.4	5.5 ± 0.5	5.9 ± 0.5
CMX-CMX	4.4 ± 1.3	2.7 ± 0.8	1.8 ± 0.3	2.2 ± 0.3

Table 20: Predicted rates of dimuons due to B decays involving two muons in the final state.

Carlo simulation and the high-mass sideband data to verify this prediction. In Figure 39, we compare the distributions of the two-track transverse momentum (p_T), the displacement of the two-track vertex (L_{xy}), the two-track impact parameter d_0, and the displacement significance ($L_{xy}/\sigma(L_{xy})$) of the two-track system between data and Monte Carlo simulation.

The p_T spectrum does not look like it is a particularly good match between data and Monte Carlo. Since the p_T spectrum is also responsible for the boost , it comes as no surprise that there is also a discrepancy in the L_{xy} distributions. Once could potentially be worried about a contamination from prompt J/ψ's being responsible for a part of the peak in the data. However, the remaining two distributions, the impact parameter and L_{xy} appear to have very good agreement between data and Monte Carlo simulation. This is likely due to the fact that the impact parameter and L_{xy} significance do not directly depend on the p_T spectrum, but more on the decay topologies (opening angles, decay time in the particle's reference frame).

The conclusion from these comparisons is that there is a potential possibility that a fraction of the decays, most likely in the J/ψ peak, is coming from prompt J/ψ production. However, the topological distributions which do not depend so crucially on the transverse momentum spectrum have a good match between data and Monte Carlo. The matching decay topologies and invariant mass spectrum are a compelling argument that the dominant background is coming from $B \to \mu\mu X$ decays. In the following sections, the topological variables will also be used in an attempt to reduce this dominant background.

7.7 Total Expected Background Rate

At this point it is rather obvious that the dominant source of dimuons in the D^* tagged sample are B meson decays involving two real muons in the final state. Therefore the rate prediction from Table 20 dominates the double-tag rate calculation. For completeness, the contribution of every considered source is listed in Table 21. The statistical and systematic uncertainties are added into a total uncertainty. The total dimuon rate is obtained by adding all the contributions. The total uncertainty on the dimuon rate is calculated by adding the individual uncertainties in quadrature.

8 Measurement Optimization

Once it is established that our main background is in fact from real muons and we understand the source, we can attempt to minimize the dominant background in an effort to improve

Source	CMU-CMU	CMU-CMX	CMX-CMX
$D^0 \to \pi^+\pi^-$	0.61 ± 0.05	0.066 ± 0.001	0.014 ± 0.001
$D^0 \to K\pi$	< 0.01	< 0.01	< 0.01
Combinatorial double-mistags	0.14 ± 0.03	0.028 ± 0.005	0.0025 ± 0.0004
$D^0 \to \mu\nu_\mu X$	< 0.48	< 0.37	< 0.14
single-muon B decays	1.8 ± 0.1	0.54 ± 0.06	0.24 ± 0.04
B decays with dimuons	17.7 ± 2.9	5.4 ± 1.4	2.8 ± 0.8
Total	20.3 ± 2.9	6.1 ± 1.4	3.1 ± 0.8

Table 21: Individual contributions to the rate of double-tags in the signal search window.

the sensitivity of the search.

In the following section, we discuss building and testing a likelihood ratio function which we use to select $D^0 \to \mu\mu$ decays from $B \to \mu\mu X$ decays, in order to remove the dominant background, while keeping most of the signal.

The likelihood ratio function takes the form:

$$p(x1, ..., xn) = \frac{\prod_i p^S(x_i)}{\prod_i p^S(x_i) + \prod_i p^B(x_i)} \tag{26}$$

where $x_1, ..., x_n$ are the input variables, $p^S(x_i)$ are the probability density distributions of the x_i variables for $D^0 \to \mu^+\mu^-$ decays, and $p^B(x_i)$ are the probability density distributions of x_i variables for $B \to \mu^+\mu^- X$ decays. This formulation assumes that the different signal and background variables are uncorrelated; correlation between variables induces an automatic penalty of sub-optimal performance.

The variables we consider for separation of signal and background are the momentum of the two-track pair, the L_{xy} of the two-track pair, the impact parameter of the two-track pair momentum with respect to the primary vertex, and the L_{xy} significance of the two-track pair. The momentum is considered because the momentum spectrum of B mesons differs from that of D mesons; while some of the momentum is lost in the particle decay, one would expect to see some signature of the original B decay in the two-track system. The L_{xy} of the two-track pair is chosen because of the different lifetime of B mesons and D^0 mesons. The impact parameter of the two-track pair is considered since most (85%) of the D^0 mesons in the sample are prompt. Hence the momentum of the two-track pair should point straight back at the primary vertex. In $B \to \mu^+\mu^-$ decays, momentum from other decay fragments is not reconstructed, so the two-track momentum does not have to point right back at the primary vertex. Finally, information about the decay topology (track opening angle) is also encoded in the displacement uncertainty, $\sigma(L_{xy})$. The distribution of $L_{xy}/\sigma(L_{xy})$ is equivalent to the distribution of decay time significance, $ct/\sigma(ct)$, so it is independent of the momentum spectrum of the charm and B mesons involved, and simply depends on the lifetimes of the D^0 and B mesons.

Figure 44 show the comparisons of the distributions of these quantities for $D^0 \to \pi^+\pi^-$ and $B \to \mu\mu X$ Monte Carlo simulation. The p_T and L_{xy} distributions show some separation power, but not much. Further, as shown in Section 7.6, the p_T and L_{xy} distributions are poorly matched between the data and Monte Carlo Simulation - one would not like to base a discriminating variable on such inputs. Therefore, we concentrate on the remaining two variables, which have good agreement between data and Monte Carlo, namely the impact parameter and the L_{xy} significance.

The distributions of the two-track system impact parameter and L_{xy} significance are fitted with probability density functions, as shown in Figure 45. The signal and background impact parameter distributions are fit with a double Gaussian, and the signal and background L_{xy} significance distributions are modelled with an exponential convoluted with a Gaussian. All distributions that are being fitted are taken from reallistic Monte Carlo samples.

One can be concerned about the performance of a likelihood derived from Monte Carlo simulation on real data. In particular, the Monte Carlo simulation is generated assuming all D^* are prompt. In reality, we know that about 10-15% of the charm decays we observe in data are in fact coming from decays of B mesons, so the momentum of the D^0 will not necessarily point back directly at the primary vertex. Further, due to the sequential decay, the L_{xy} signficance of the D^0 meson decay is likely to be larger for secondary D^0 decays than for those coming from prompt charm production. In order to evaluate these effects, we compare the Monte Carlo simulated quantities to those extracted from $D^0 \rightarrow K\pi$ decays in data in Figure 46. We also compare the likelihood values obtained from Monte Carlo simulation on prompt charm to those obtained from $D^0 \rightarrow K\pi$ decays in data in Figure 47. A large drop in performance would be found if in the data there were either a strong change in the likelihood distribution shape, or events started suddenly piling up in the zero-bin of the likelihood, where most of the background events are. Instead, only 5% of the data events sit in the zero-bin, and the likelihood distribution looks comparable between data and Monte Carlo simulation of D^0 decays.

The analysis is optimized by selecting events which pass likelihood selection. It is therefore important to determine the efficiency for signal and background events having a certain likelihood or greater. Figure 48 shows the expected signal and background efficiency behaviour as the likelihood cut is being scanned over. For a set of 18 probe values of the likelihood cut, we also estimate the 90% and 95% CL limits using the technique described in Section 11. The results are shown in Figure 49. The non-smooth behavior of the optimization curve is most likely due to statistical fluctuations in the $B \rightarrow \mu^+\mu^- X$ Monte Carlo sample where the background efficiencies are being estimated. For simplicity, we assume that the background level and uncertainty always scale by the background efficiency at a given point. The signal efficiency is estimated from the likelihood output of the $D^0 \rightarrow K\pi$ data. The estimated optimal likelihood cut is a likelihood requirement of $\mathcal{L} \geq 0.35$.

9 In Situ μ Efficiency Cross-Check

The CMU-CMU dimuon sample is sufficiently large for us to attempt an in-situ cross-check of the average muon identification efficiency by comparing the yields of J/ψ's seen in the single-tagged and double-tagged peaks. The fits to the corresponding J/ψ peaks, after applying the bottom/charm separation likelihood cuts, are shown in Figure 59. The measured yields of J/ψ decays in the samples are 368 ± 24 and 196 ± 14, respectively. From this we calculate that the ratio of the two counts is $r = 53.2 \pm 5.2\%$.

The ratio r should not be confused with the single track muon efficiency, ϵ. The prob-

Detectors	$D^0 \to K\pi$ Data	$B \to \mu\mu X$ MC Eff.	$B \to \mu\mu X$ Data Eff.	Comb Bg Eff
CMU-CMU	87%	42%	$(42 \pm 2)\%$	$(25.6 \pm 0.7)\%$
CMU-CMX	86%	42%	$(43 \pm 3)\%$	$(27.4 \pm 1.0)\%$
CMX-CMX	88%	42%	$(38 \pm 5)\%$	$(28.4 \pm 1.3)\%$

Table 22: Signal efficiencies due to the bottom/charm separating likelihood cut for the different detector configurations.

ability that a J/ψ candidate will end up in the single-tag sample, assuming uncorrelated efficiencies for the two J/ψ daughter tracks, is $\mathcal{P}_1 = 1 - (1 - \epsilon)^2$. The probability that a J/ψ candidate will end up in the double-tagged sample is then $calP_2 = \epsilon^2$. The ratio r is then related to the single track efficiency ϵ, as follows:

$$r = \frac{\mathcal{P}_2}{\mathcal{P}_1} = \frac{\epsilon^2}{2\epsilon - \epsilon^2} = \frac{\epsilon}{2 - \epsilon}. \tag{27}$$

Conversely, one can calculate the efficiency ϵ from the rate r using:

$$\epsilon = \frac{2 \cdot r}{1 + r} \tag{28}$$

and by error propagation,

$$\sigma(\epsilon) = \frac{2}{(1 + r)^2} \sigma(r) \tag{29}$$

Using this calculation, we find that the single-track efficiency is $\epsilon = (69.5 \pm 4.4)\%$, which is in good agreement with the value quoted in Table 1, which is $(70.4 \pm 0.6)\%$.

10 Final Background Estimate

Before looking into the signal mass window, we perform a full re-evaluation of the signal efficiencies and expected background estimates, now also requiring that the bottom/charm separating likelihood of all events is greater than $\mathcal{L} \geq 0.35$.

We evaluate the signal efficiency by assuming that the $D^0 \to \mu^+\mu^-$ signal will have a relative contribution from secondary charm ($B \to D^*$) decays which is the same as the secondary fraction of $D^0 \to K\pi$ decays. This is based on the fact that the trigger acceptance efficiency for the two decays is virtually identical, which means that the fraction of secondary D decays will be driven by the ratio of charm and bottom production cross sections, which is the same for both D^0 final states. The signal efficiencies due to the likelihood cut for the different detector configurations are listed in Table 22.

The expected background levels from different sources are estimated separately. The expected contributions from $D^0 \to \pi^+\pi^-$ and $D^0 \to K\pi$ decays are multiplied by the expected signal efficiencies, as their decay topologies match those of the signal, so the likelihood distributions are expected to be similar. The contribution from combinatorial background is re-evaluated following the procedure in Section 7.3, but also requiring that the candidates

Table 23: Calculation of the semileptonic B decay contribution, specifically from decays in which one track is a real muon and the other is mistagged as a muon to the double-tagged rate in the signal mass window, after the Bottom/Charm separation likelihood requirement.

Detectors	$1-\mu$ Sample	$2-\mu$ Sample
CMU-CMU	1.10 ± 0.04	0.54 ± 0.04
CMU-CMX	0.28 ± 0.02	0.13 ± 0.03
CMX-CMX	0.22 ± 0.02	0.07 ± 0.02

Detector	$1-\mu$ Sample	$2-\mu$ Sample	Data, fit	MC, fit
CMU-CMU	7.3 ± 1.2	3.75 ± 1.32	4.9 ± 0.5	5.3 ± 0.5
CMU-CMX	1.9 ± 0.4	2.53 ± 0.96	1.99 ± 0.31	1.96 ± 0.30
CMX-CMX	1.6 ± 0.5	0.96 ± 0.48	0.72 ± 0.19	0.66 ± 0.18

Table 24: Predicted rates of dimuons due to B decays involving two muons in the final state.

pass the bottom/charm separation likelihood requirement. The contributions from semileptonic D decays is scaled by the efficiency of accepting semileptonic Monte Carlo candidates. The contributions from semileptonic B decays and $B \to \mu^+\mu^- X$ decays are evaluated by repeating the procedures from Section 7.5 and 7.6, respectively, requiring that all events in the data and Monte Carlo simulation pass the bottom/charm separation likelihood requirement. The yields of $B \to \mu X$ and $B \to \mu^+\mu^- X$ decays estimated from the single-tagged and double-tagged sample are listed in Tables 23, and 24, respectively.

The list of expected background contributions re-evaluated after applying the likelihood cut is presented in Table 25. After re-evaluating the signal efficiencies and the background contributions, the optimized analysis sensitivity is found to be $0.74 \cdot 10^{-6}$, which is an improvement of roughly 20% over the un-optimized analysis. The distribution of confidence levels as a function of branching fraction, assuming the number of observed events is the same as the number of expected background events, is shown in Figure **??**.

Source	CMU-CMU	CMU-CMX	CMX-CMX
Combinatorial Background	0.040 ± 0.007	0.008 ± 0.001	0.0007 ± 0.0001
$D^0 \to \pi\pi$ Double Tags	0.53 ± 0.005	0.057 ± 0.001	0.012 ± 0.002
$D^0 \to K\pi$ Double Tags	< 0.01	< 0.01	< 0.01
Semileptonic D^0 Decays	< 0.36	< 0.20	< 0.10
Single-Muon B Decays	0.54 ± 0.06	0.13 ± 0.03	0.07 ± 0.02
Di-muon B Decays	3.8 ± 1.32	2.54 ± 0.96	0.96 ± 0.48
Total	4.93 ± 1.33	2.74 ± 0.96	1.04 ± 0.48

Table 25: Background contributions re-evaluated after applying the likelihood cut .

11 Sensitivity and Limits

As already discussed, the interpretation of the measurement is condensed in the following equation:

$$\mathcal{B}(D^0 \to \mu\mu) = \frac{N(\mu\mu)}{N(\pi\pi)} \cdot \frac{\epsilon(\pi\pi)}{\epsilon(\mu\mu)} \cdot \mathcal{B}(D^0 \to \pi\pi) \tag{30}$$

where $N(\mu\mu)$ is the observed number of $D^0 \to \mu\mu$ decays in the search window, $N(\pi\pi)$ is number of reference $D^0 \to \pi\pi$ decays, $\epsilon(\mu\mu)$ and $\epsilon(\pi\pi)$ are the efficiencies for accepting di-muon and di-pion decays, respectively, and $\mathcal{B}(D^0 \to \pi\pi)$ is the branching fraction for the reference $D^0 \to \pi\pi$ decay. The number of observed $D^0 \to \mu\mu$ decays is calculated by subtracting the number of expected decays due to background processes, N_{bkg} from the number of observed decays in the search window, N_{obs}:

$$N(\mu\mu) = N_{\text{obs}} - N_{\text{bkg}}. \tag{31}$$

If a significant excess of observed decays over expected decays is found, the interpretation of the result is trivial. The number of $D^0 \to \mu\mu$ decays is calculated from Eq. 31, and the observed branching fraction is calculated from Eq. 3.

The question is slightly more complicated in the case of setting limits, mostly because a lot of the inputs to Eqs. 3 and 31 have uncertainties. For instance, the efficiencies $\epsilon(\mu^+\mu^-)$, $\epsilon(\pi\pi)$, the number of expected decays N_{bkg} are all measured quantities with associated uncertainties which have to be taken into account when setting limits. Obviously, one also wants to combine the results obtained in the three channels: CMU-CMU, CMU-CMX and CMX-CMX. We use a Bayesian approach to seeing limits on $\mathcal{B}(D^0 \to \mu^+\mu^-)$ that allows for the combination of a number of channels and the inclusion of uncertainties on input quantities.

The Bayesian problem is outlined in CDF note 7587 [7]. We will calculate a limit based on a number of channels where the probability of observing n_k events in the k^{th} channel is

$$\frac{e^{-(se_k+b_k)}(se_k + b_k)^{n_k}}{n_k!} \tag{32}$$

where b_k is the number of background events and se_k is the number of signal events expressed as a product of the signal, s, we wish to extract multiplied by an efficiency factor that yields the number of signal events in the k^{th} channel.

Assuming we knew the branching fraction, we would calculate the expected number of $D^0 \to \mu^+\mu^-$events as

$$N_{\mu\mu} = N_{\pi\pi}\epsilon_{\mu\mu}\frac{\Omega_{\mu\mu}}{\Omega_{\pi\pi}}\frac{\mathcal{B}(D^0 \to \mu^+\mu^-)}{\mathcal{B}(D^0 \to \pi^+\pi^-)}.$$

Letting $s = \mathcal{B}(D^0 \to \mu^+\mu^-)$, we have

$$e = \left[N_{\pi\pi}\epsilon_{\mu\mu}\frac{\Omega_{\mu\mu}}{\Omega_{\pi\pi}}\right]\left(\frac{1}{\mathcal{B}(D^0 \to \pi^+\pi^-)}\right)$$

where the errors on the term in square brackets are uncorrelated between channels, and the error on the term in parenthesis is correlated.

Table 26: The calculation of e for each of the channels used in the analysis. The last row of the table contains the number of events observed after unblinding, as discussed in the next section.

	CMU-CMU	CMU-CMX	CMX-CMX
$N_{\pi\pi}$	24400 ± 200	9260 ± 130	6940 ± 110
$\epsilon_{\mu\mu}$	0.437 ± 0.003	0.257 ± 0.004	0.161 ± 0.003
$\Omega_{\pi\pi}/\Omega_{\mu\mu}$	0.872 ± 0.005		
$N_{\pi\pi}^{eff}$	$9296\pm$	$2074\pm$	$972\pm$
$\mathcal{B}(D^0 \to \pi^+\pi^-)$	$(1.364 \pm 0.032) \times 10^{-3}$		
e $(\times 10^{-4})$	896.5 ± 23.7	200.1 ± 11.7	93.9 ± 3.2
N_{bkg}	4.93 ± 1.33	2.74 ± 0.96	1.04 ± 0.48
N_{obs}	3	0	1

The sensitivity of the measurement can be estimated before opening the box using the estimated efficiencies and backgrounds, and assuming zero signal. We generate 1000 "toy experiments", each with a number of events observed randomly generated according to a poisson distribution with mean equal to the expected background. The distributions of the number of CMU-CMU, CMU-CMX, CMX-CMX, and the sum of the three are shown in Fig. 40.

The limit on the branching ratio is calculated for each experiment at the 68%, 90%, and 95% confidence level. The distributions of limits are displayed in Figs. 41–43. The median of the distribution of limits is quoted as the sensitivity of the measurement. The sensitivities are 0.65×10^{-6} at the 90% confidence level, and 0.78×10^{-6} at the 95% confidence level.

12 Opening the Box

The mechanics of opening the box are rather trivial in this analysis. We count the number of events found in the signal mass window. We find three candidates in the CMU-CMU configuration, none in the CMU-CMX configuration and one candidate in the CMX-CMX configuration, listed in the last row of Table 26. Notice in Fig. 40 the probabilities of observing 3 CMU-CMU events, 0 CMU-CMX events, and 1 CMX-CMX event are reasonable, and the probability of observing 4 events total is about 5%. These observations are consistent with expected background, so we will quote a limit on the branching fraction.

To calculate the limit we lump all the efficiencies and normalizations together into one factor, and solve for the limit on the branching ratio directly. We want s to equal the branching ratio, $\mathcal{B}(D^0 \to \mu^+\mu^-)$, so let

$$e_k = \epsilon_{\mu k} \frac{\Omega_{\mu\mu}}{\Omega_{\pi\pi}} \frac{N_{\pi\pi k}}{\mathcal{B}(D^0 \to \pi^+\pi^-)}$$

such that $se_k = N_{\mu\mu k}$. The effect of the nuisance parameters are included in the limit calculation by allowing for a random, gaussian distributed variations consistent with the gaussian errors on the efficiency and normalization factors. We take a flat prior for s, and for treat the channels separately, except where they share common, correlated errors.

The resulting limit is

$$\mathcal{B}(D^0 \to \mu^+\mu^-) < 0.43 \times 10^{-6} \text{ at the 90\% CL}$$

and

$$\mathcal{B}(D^0 \to \mu^+\mu^-) < 0.53 \times 10^{-6} \text{ at the 95\% CL.}$$

The probability for these limits, assuming there is really no signal, is about 15% and is indicated in Figs. 42 and 43.

13 Conclusions

We have presented an update of the search for FCNC $D^0 \to \mu^+\mu^-$ decays using 360 pb^{-1} of Run II data. The analysis has been expanded to also include muons in the fiducial region of the CMX detector. The increased sample size allows for a more detailed study of the dominant background, resulting in a first identification of $B \to \mu\mu X$ decays as being the most likely physics source of background in the signal mass range. Using a likelihood ratio function, we reject about half of the dominant background while keeping almost 85% of the signal. This improves the sensitivity of the search by roughly 20%. After looking at events in the signal region, we find 3 CMU-CMU, 0 CMU-CMX and 1 CMX-CMX candidates. We set the following limits on the branching ratio of the D^0 decay: $\mathcal{B}(D^0 \to \mu^+\mu^-) < 4.3 \times 10^{-7}$ at the 90% CL, and $\mathcal{B}(D^0 \to \mu^+\mu^-) < 5.3 \times 10^{-7}$ at the 95% CL. This is roughly a factor of three better than the previous world best result [2], and an improvement of roughly a factor of five with respect to the previous CDF analysis [5].

14 Figures

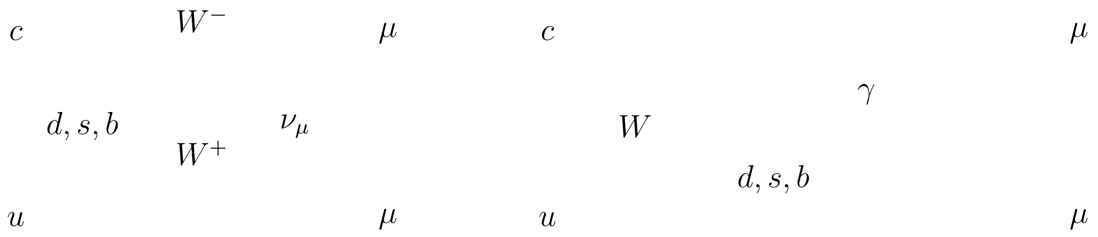

Figure 1: A few of the possible Feynman diagrams for the $D^0 \to \mu^+\mu^-$ decay in the Standard Model

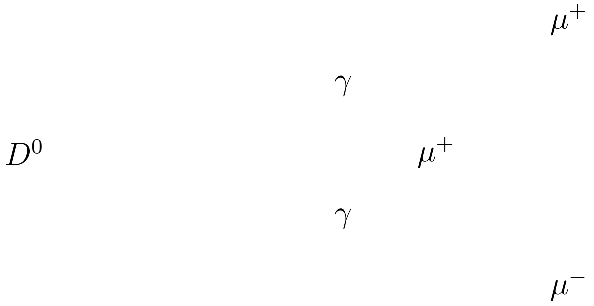

Figure 2: The dominant long-distance Feynman diagram for the $D^0 \to \mu^+\mu^-$ decay in the Standard Model

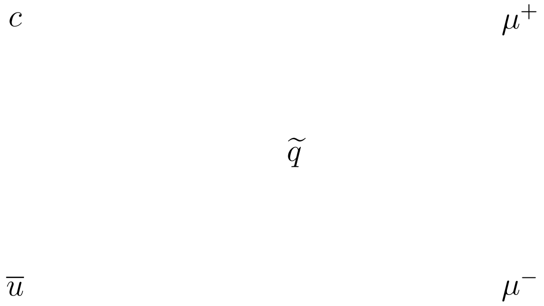

Figure 3: A SUSY contribution to $D^0 \to \mu^+\mu^-$ decays.

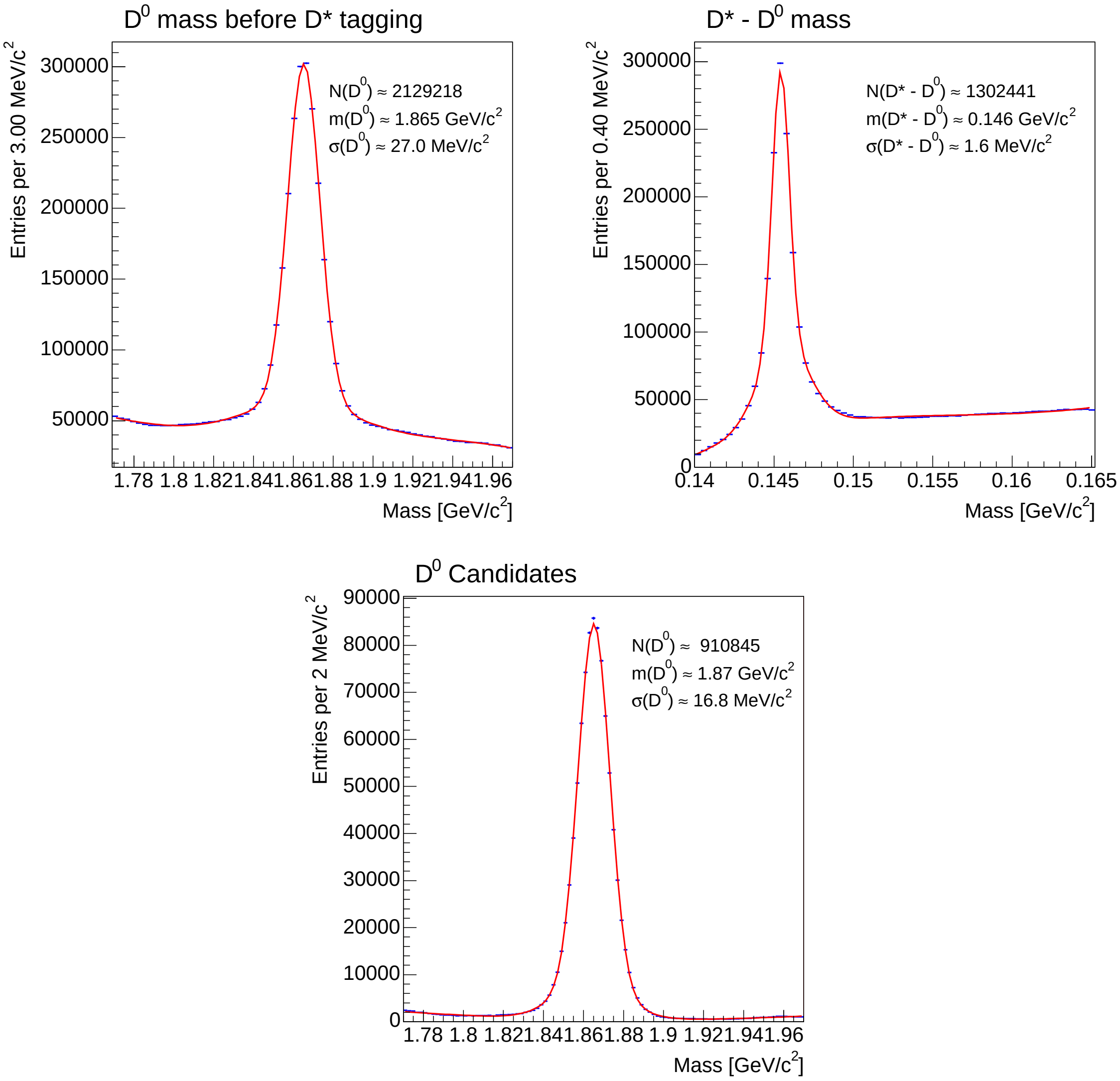

Figure 4: Mass plots. The upper left plot shows the D^0 mass spectrum before D^* tagging. The upper right plot shows the peak in the $m(D^{*+}) - m(D^0)$ mass difference spectrum. We require $0.144\ GeV/c^2 < m(D^{*+}) - m(D^0) < 0.147\ GeV/c^2$. The bottom plot shows the the final collection of D^0 candidates.

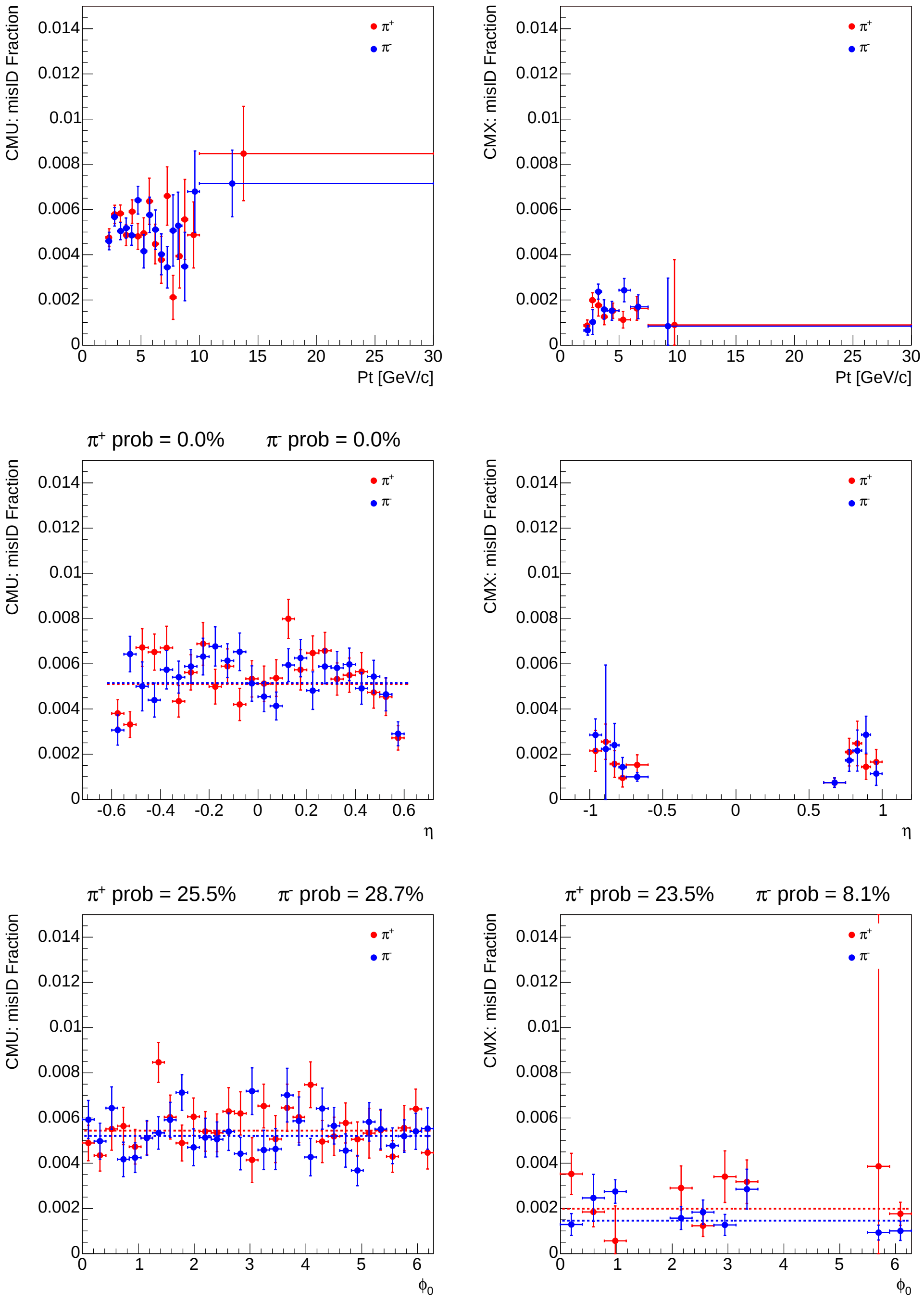

Figure 5: Pion fake rates binned in p_T, η, and ϕ according to charge.

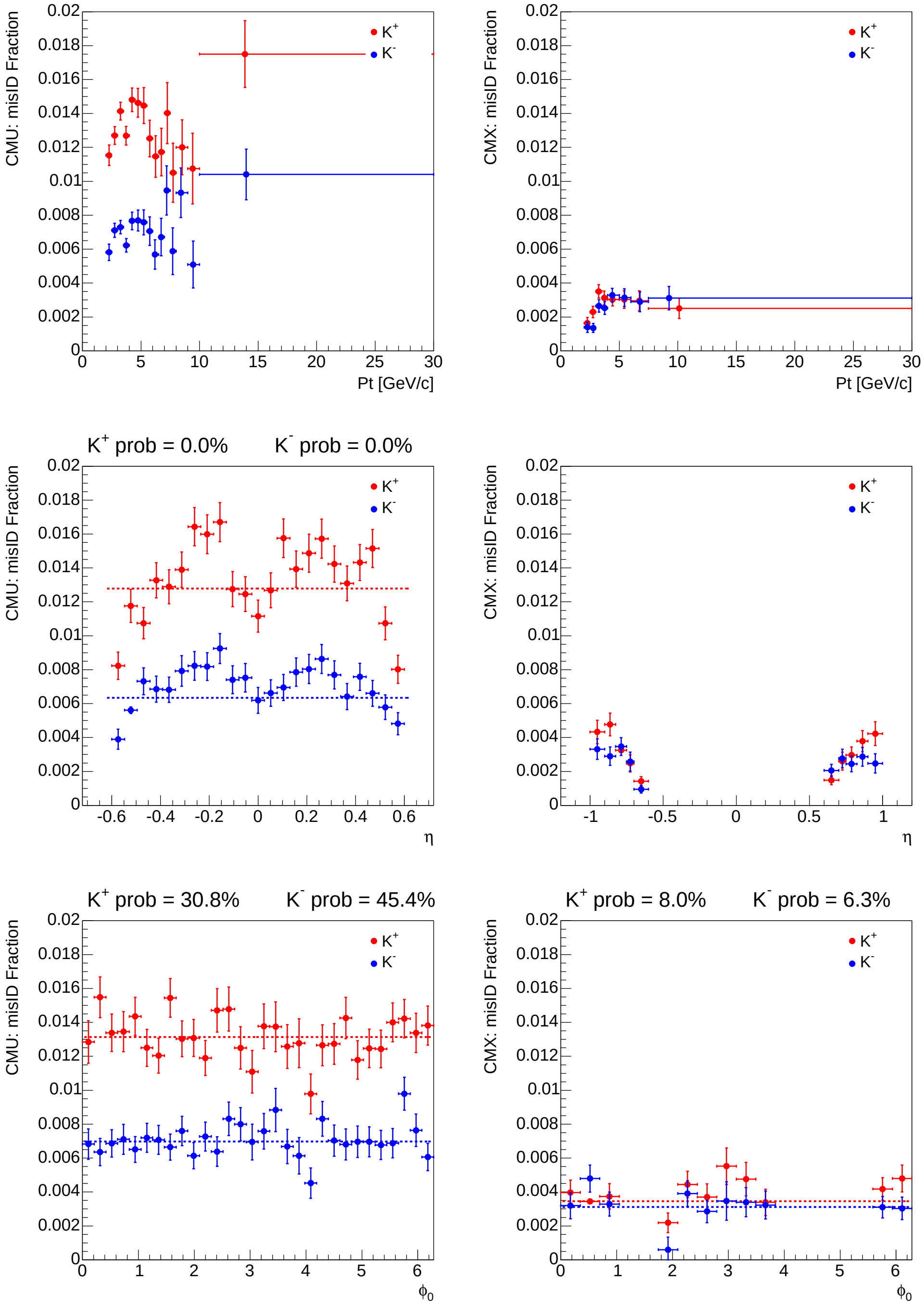

Figure 6: Kaon fake rates binned in p_T, η, and ϕ according to charge. Negative kaons have a much higher scattering cross-section and are much less likely to punch through to the muon detectors than positive kaons are. This results in a lower fake rate for K^- than for K^+, as seen here.

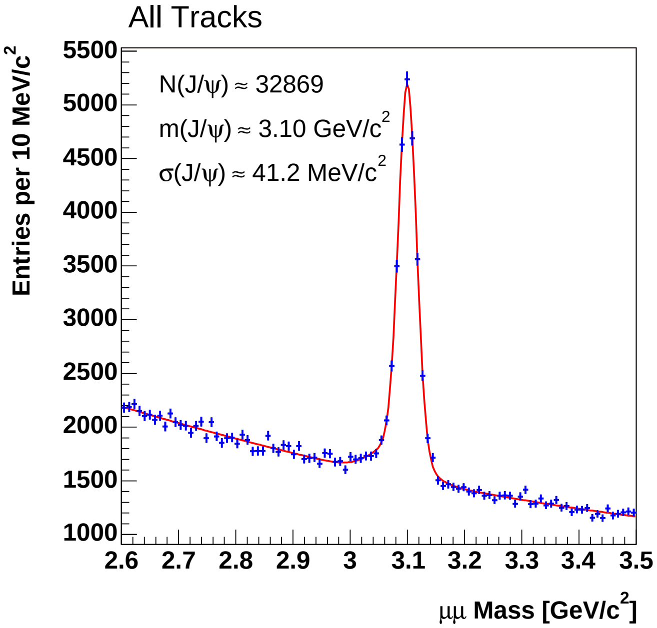

Figure 7: $J/\Psi \rightarrow \mu\mu$ candidates, obtained using mu + LSVT trigger. These candidates are used to calculate muon matching efficiencies.

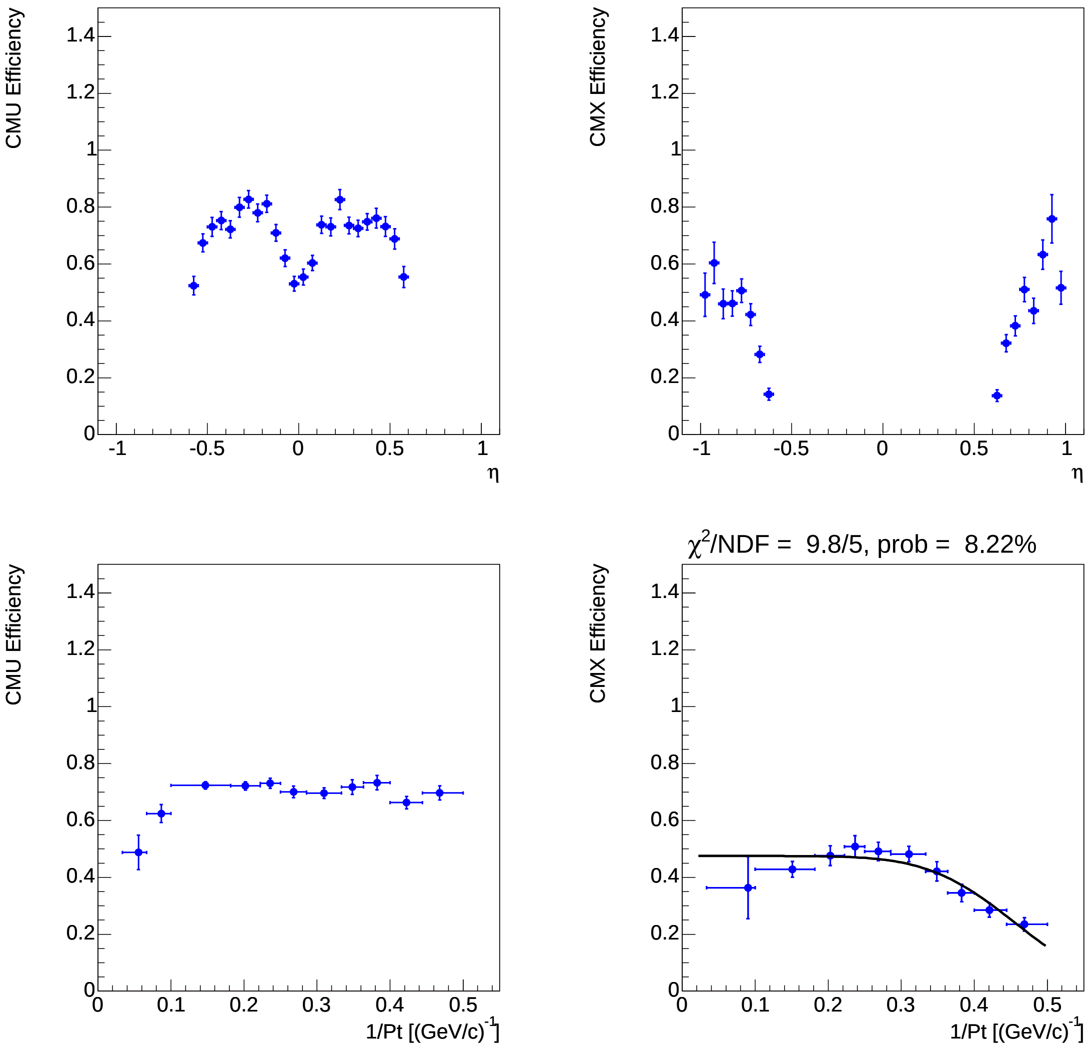

Figure 8: CMU and CMX muon matching efficiencies binned in η and p_T. Note that both CMU and CMP efficiencies are dependent on p_T. This explains the difference in the results of our muon matching efficiency calculation methods.

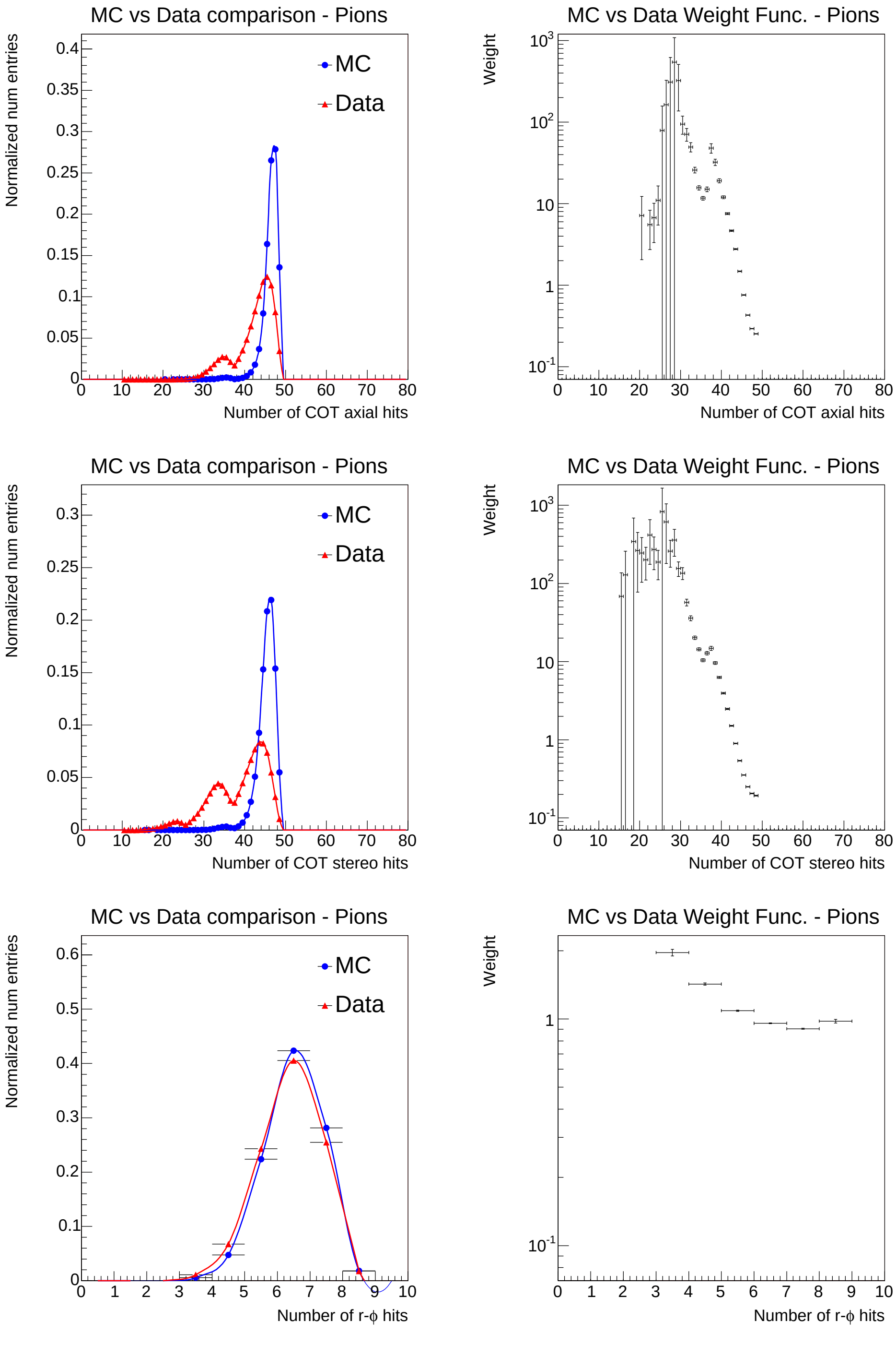

Figure 9: Comparison between Monte Carlo and data measurements of track quality parameters of pion tracks. The plots on the left are unit-normalized histograms. The plots on the right represent the corresponding weighting functions (see equations 8 and 9). These weighting functions can be used to calculate relative trigger efficiency as described in section 6.3.

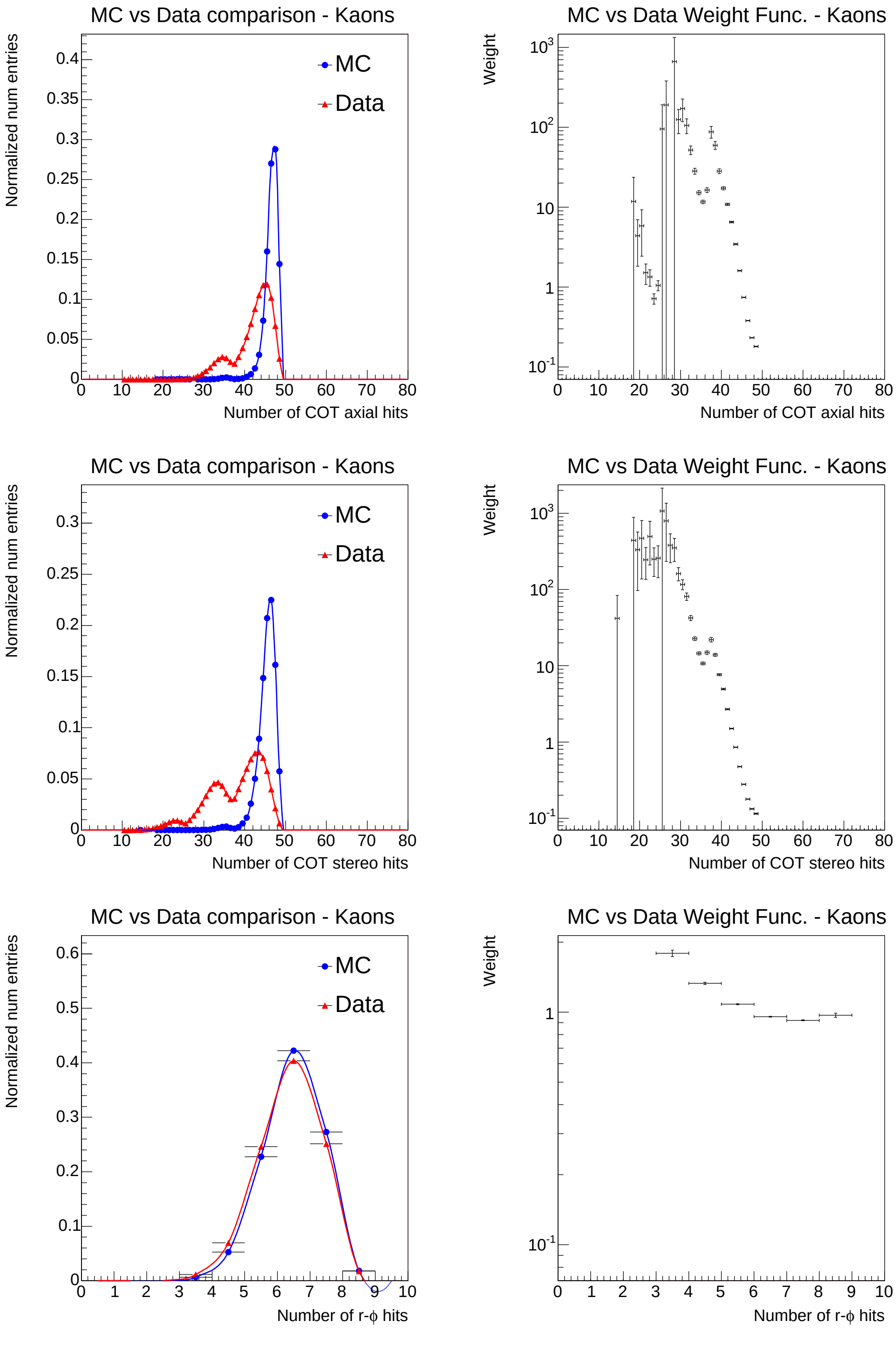

Figure 10: Comparison between Monte Carlo and data measurements of track quality parameters of kaon tracks. The plots on the left are unit-normalized histograms. The plots on the right represent the corresponding weighting functions (see equations 8 and 9). These weighting functions can be used to calculate relative trigger efficiency as described in section 6.3.

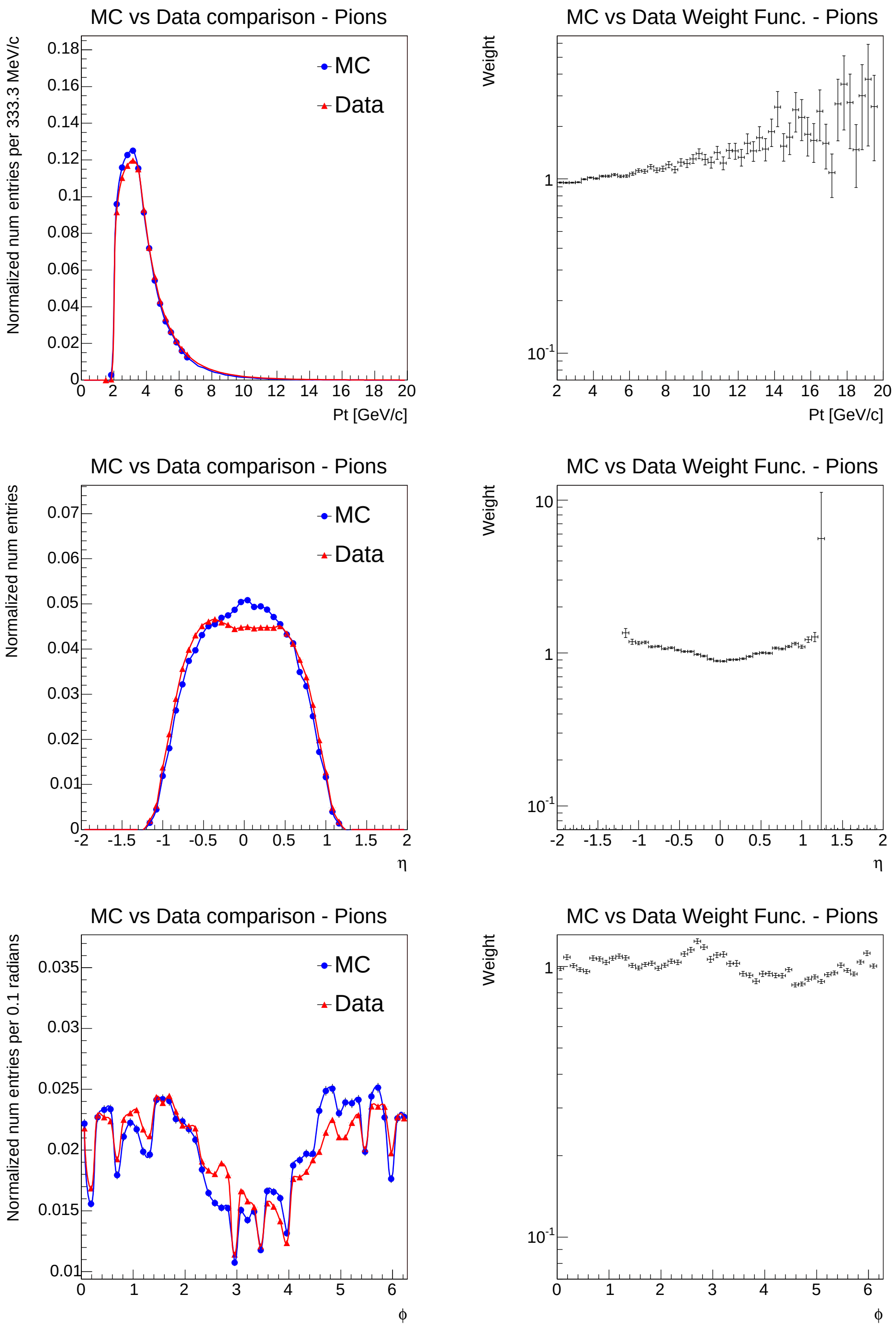

Figure 11: Comparison between Monte Carlo and data for kinematic measurements of pion tracks. The plots on the left are unit-normalized histograms. The plots on the right represent the corresponding weighting functions (see equations 8 and 9). These weighting functions can be used to calculate relative trigger efficiency as described in section 6.3. Agreement between monte carlo and data is good, with the notable exception around $\eta \approx 0$.

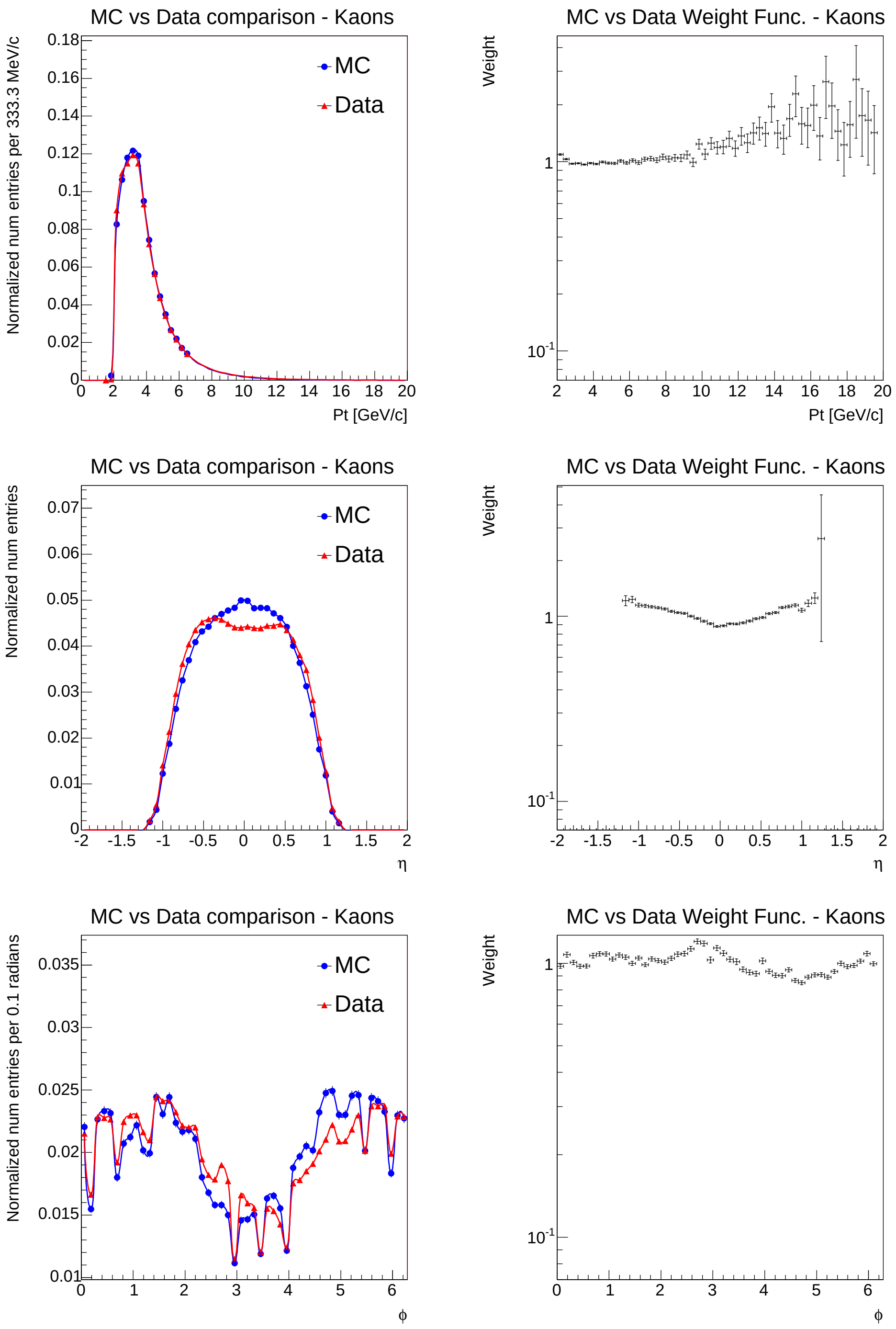

Figure 12: Comparison between Monte Carlo and data for kinematic measurements of kaon tracks. The plots on the left are unit-normalized histograms. The plots on the right represent the corresponding weighting functions (see equations 8 and 9). These weighting functions can be used to calculate relative trigger efficiency as described in section 6.3. Agreement between monte carlo and data is good, with the notable exception around $\eta \approx 0$.

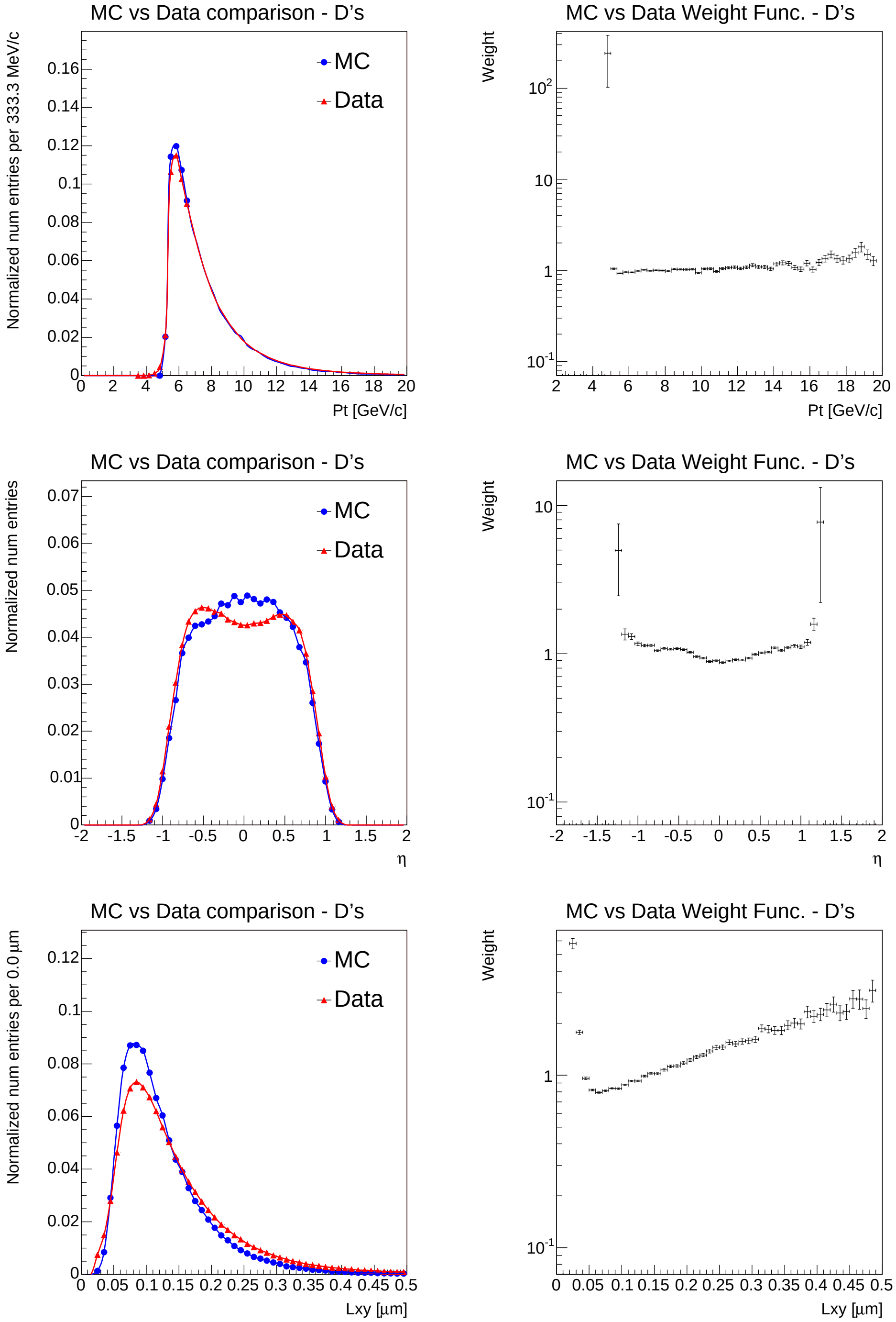

Figure 13: Comparison between Monte Carlo and data PDF's for kinematic measurements of D^0 candidates. Once again, agreement between monte carlo and data is good, with the notable exception around $\eta \approx 0$ The plots on the left are unit-normalized histograms. The plots on the right represent the corresponding weighting functions (see equations 8 and 9). These weighting functions can be used to calculate relative trigger efficiency as described in section 6.3.

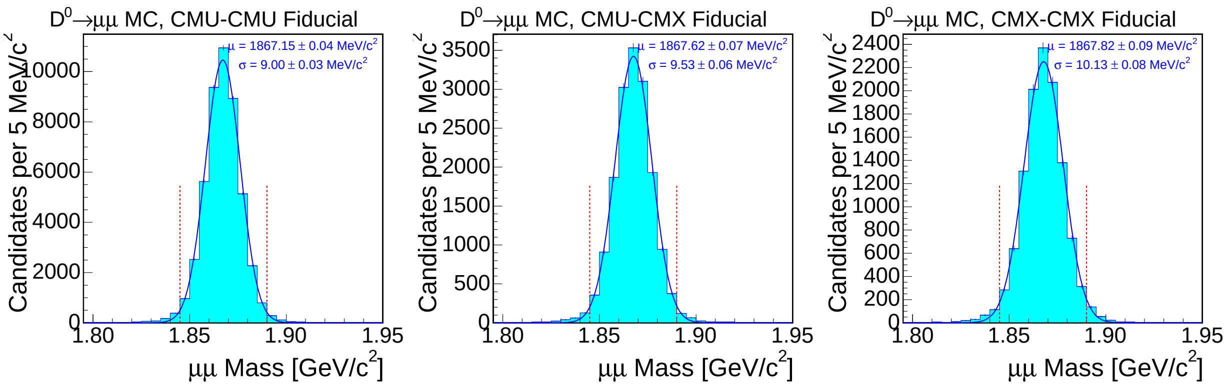

Figure 14: $D^0 \to \mu\mu$ signal mass distribution for the different detector configurations, from realistic Monte Carlo Simulation.

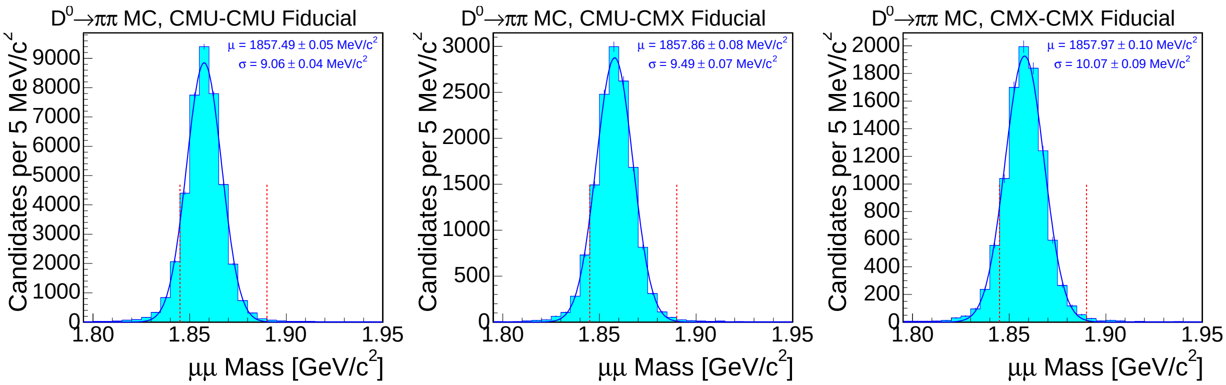

Figure 15: $D^0 \to \pi\pi$ mass distribution for the different detector configurations, from realistic Monte Carlo Simulation.

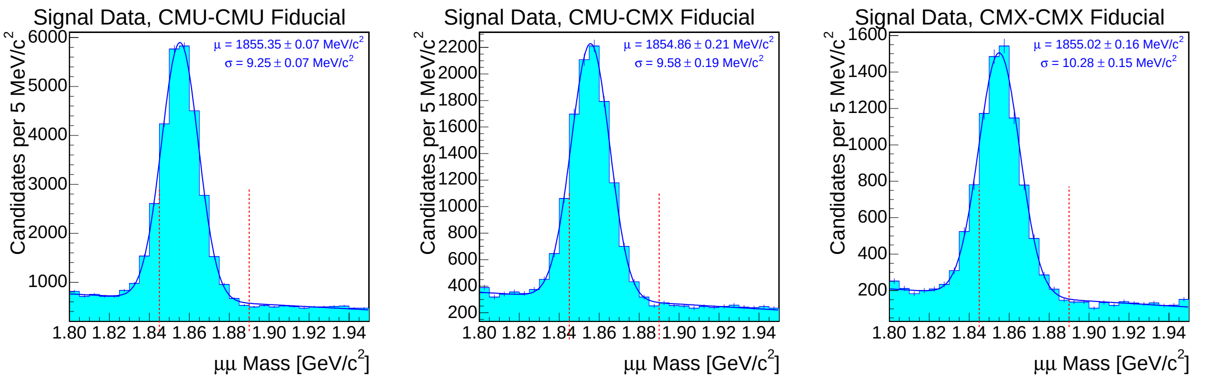

Figure 16: Mass distribution reconstructed in the $\mu\mu$ mass hypothesis, for the different detector configurations, in data.

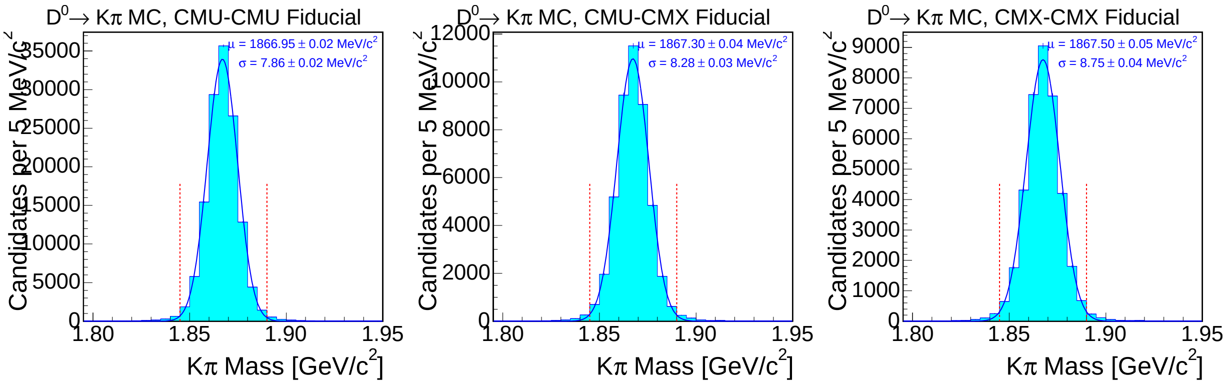

Figure 17: Mass distributions for $D^0 \rightarrow K\pi$ Monte Carlo simulated decays reconstructed in the $K\pi$ mass hypothesis, for different detector configurations.

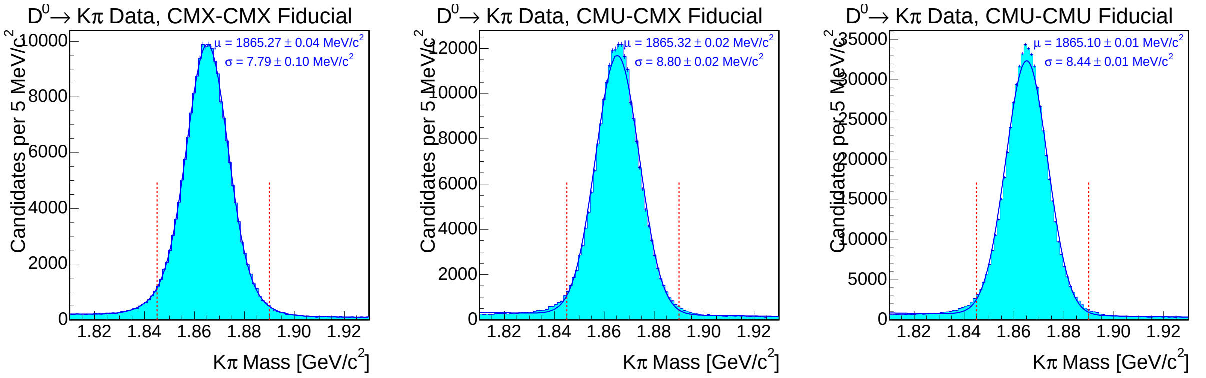

Figure 18: Mass distributions for $D^0 \rightarrow K\pi$ decays in data, reconstructed in the $K\pi$ mass hypothesis, for different detector configurations.

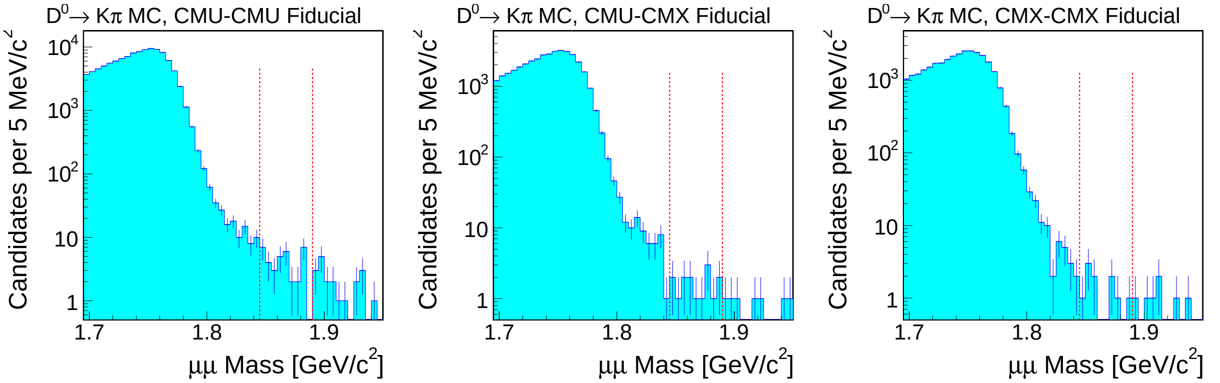

Figure 19: Mass distributions for $D^0 \rightarrow K\pi$ Monte Carlo simulated decays reconstructed in the $\mu^+\mu^-$ mass hypothesis, for different detector configurations.

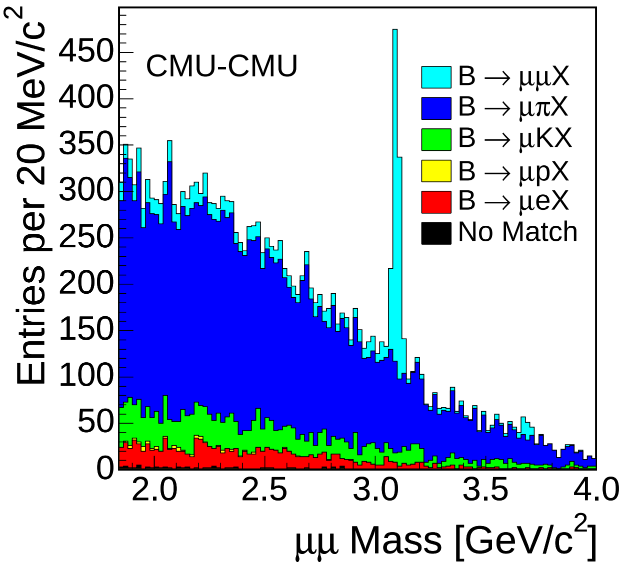

Figure 20: Composition of reconstructed decays in Monte Carlo simulation of B meson decays, in which one track was matched to a HEPG muon, depending on the particle type matched to the other track, for the CMU-CMU detector configuration

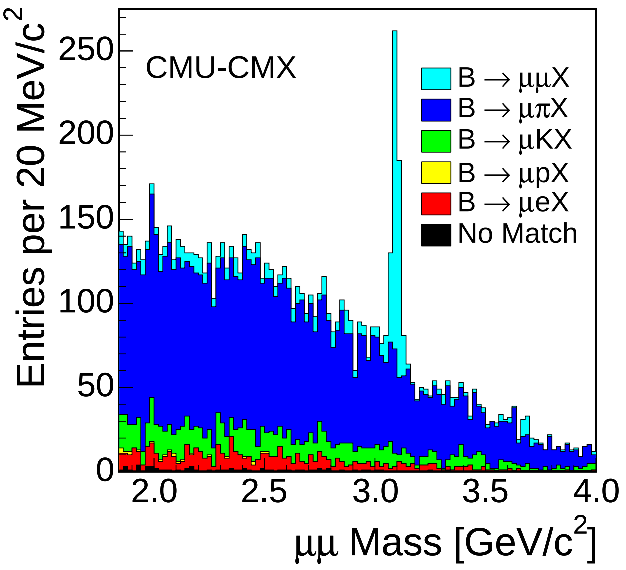

Figure 21: Composition of reconstructed decays in Monte Carlo simulation of B meson decays, in which one track was matched to a HEPG muon, depending on the particle type matched to the other track, for the CMU-CMX detector configuration

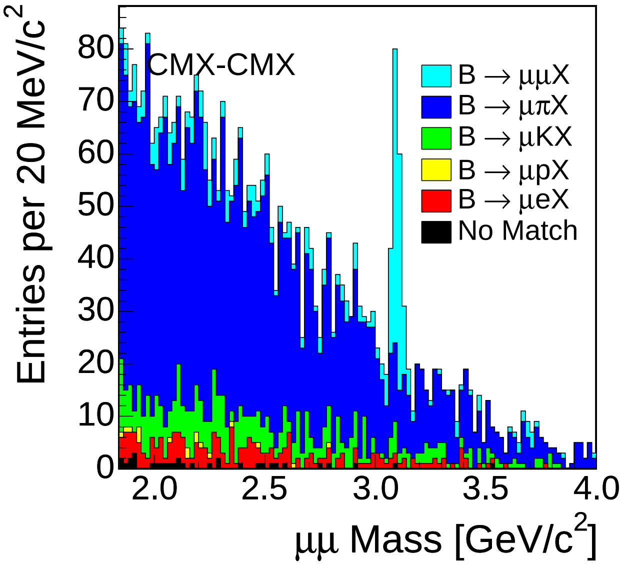

Figure 22: Composition of reconstructed decays in Monte Carlo simulation of B meson decays, in which one track was matched to a HEPG muon, depending on the particle type matched to the other track, for the CMX-CMX detector configuration

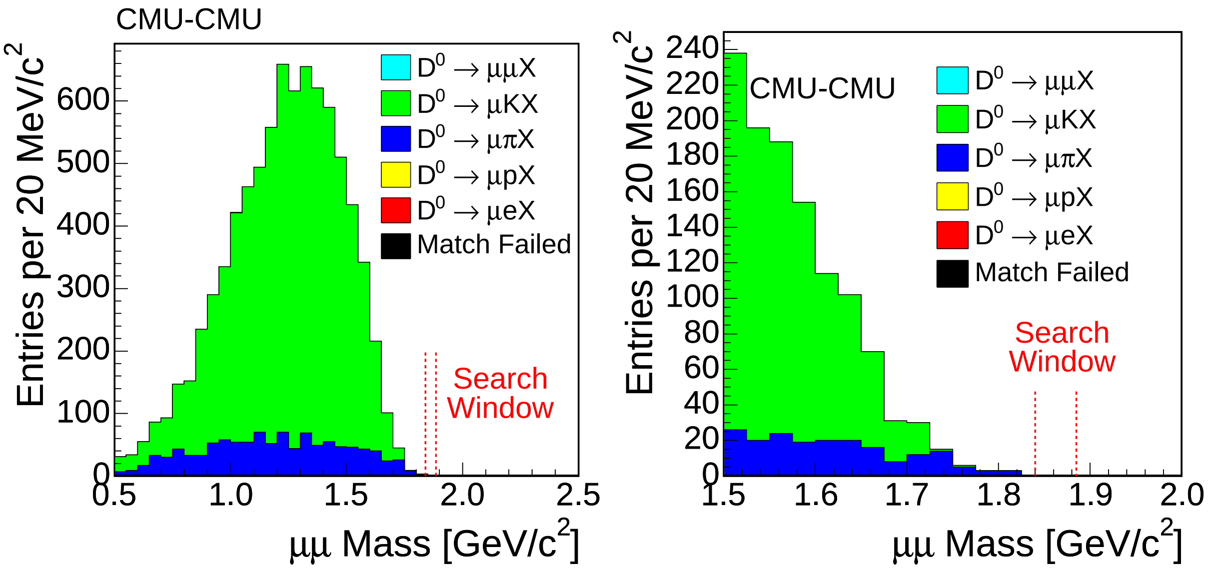

Figure 23: Invariant mass distributions of semileptonic D decays from Monte Carlo simulation, for the CMU-CMU subdetector configuration, in the “enriched” sample. In the left panel, a wide mass range is shown. On the right side, a narrow mass range around the signal search window is shown.

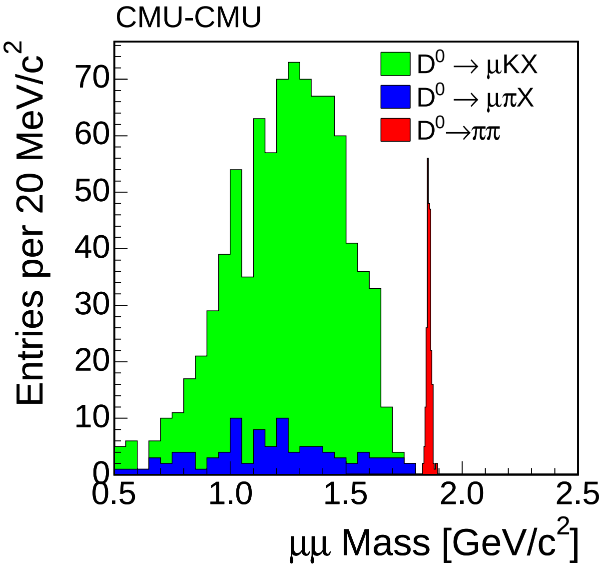

Figure 24: Invariant mass distributions of semileptonic D decays from Monte Carlo simulation, for the CMU-CMU subdetector configuration, in the “reference” sample. The $\pi^+\pi^-$ decay from the same sample is also shown, with finer binning to show the peak width.

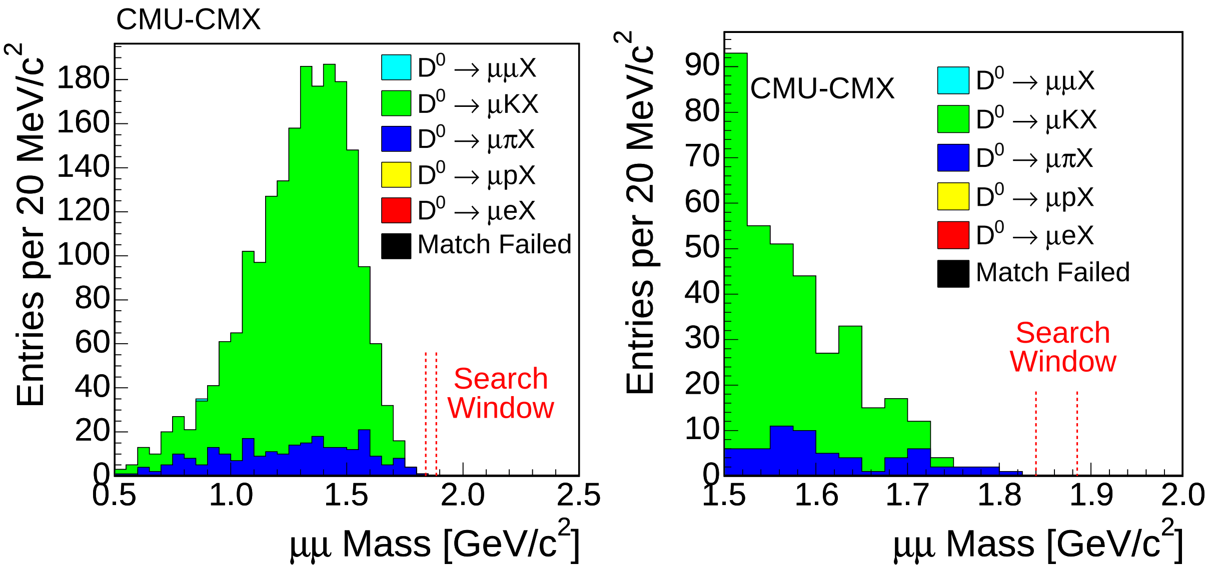

Figure 25: Invariant mass distributions of semileptonic D decays from Monte Carlo simulation, for the CMU-CMX subdetector configuration, in the "enriched" sample. In the left panel, a wide mass range is shown. On the right side, a narrow mass range around the signal search window is shown.

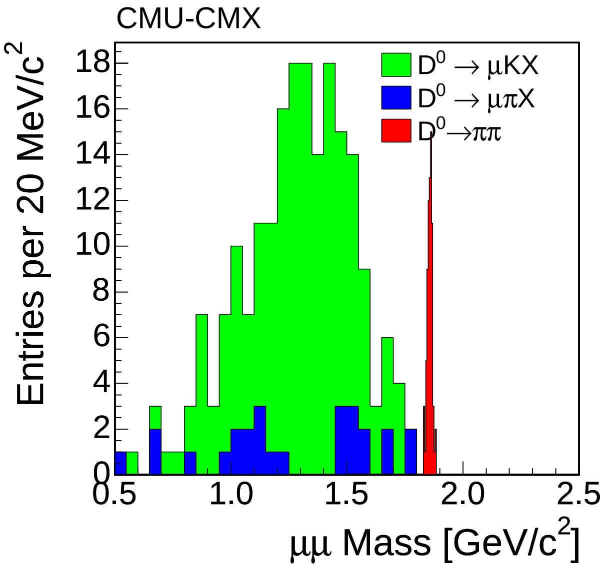

Figure 26: Invariant mass distributions of semileptonic D decays from Monte Carlo simulation, for the CMU-CMX subdetector configuration, in the "reference" sample. The $\pi^+\pi^-$ decay from the same sample is also shown, with finer binning to show the peak width.

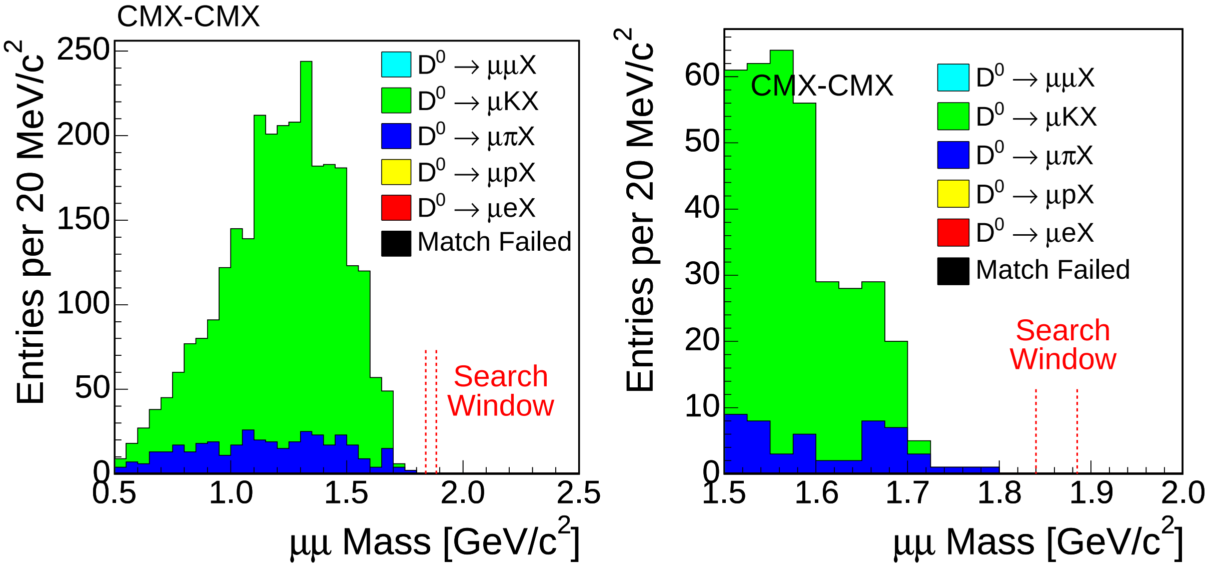

Figure 27: Invariant mass distributions of semileptonic D decays from Monte Carlo simulation, for the CMX-CMX subdetector configuration, in the "enriched" sample. In the left panel, a wide mass range is shown. On the right side, a narrow mass range around the signal search window is shown.

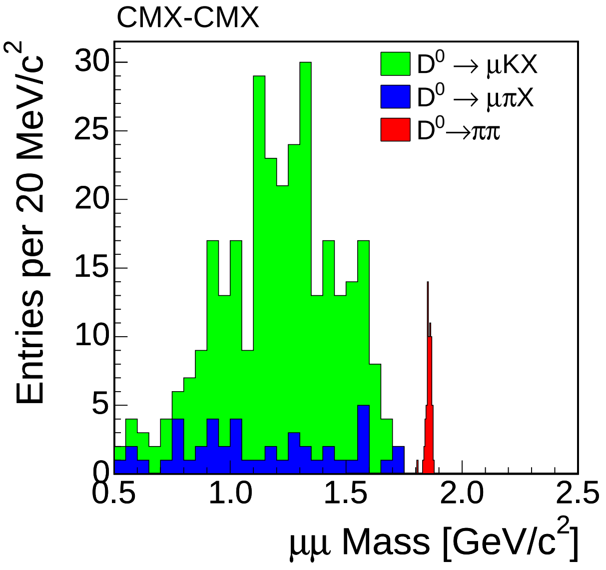

Figure 28: Invariant mass distributions of semileptonic D decays from Monte Carlo simulation, for the CMX-CMX subdetector configuration, in the "reference" sample. The $\pi^+\pi^-$ decay from the same sample is also shown, with finer binning to show the peak width.

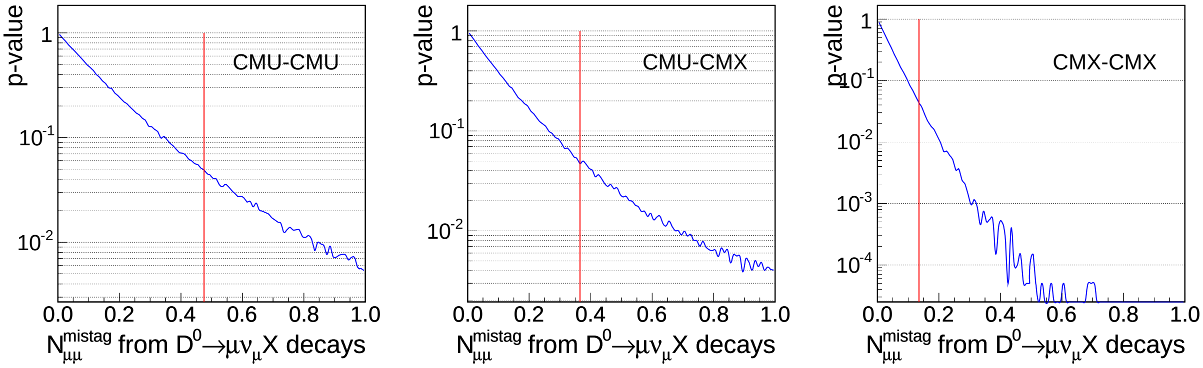

Figure 29: The distributions of p-values as a function of probe $N^{data}_{\mu^+\mu^-}$ for the three subdetector configurations

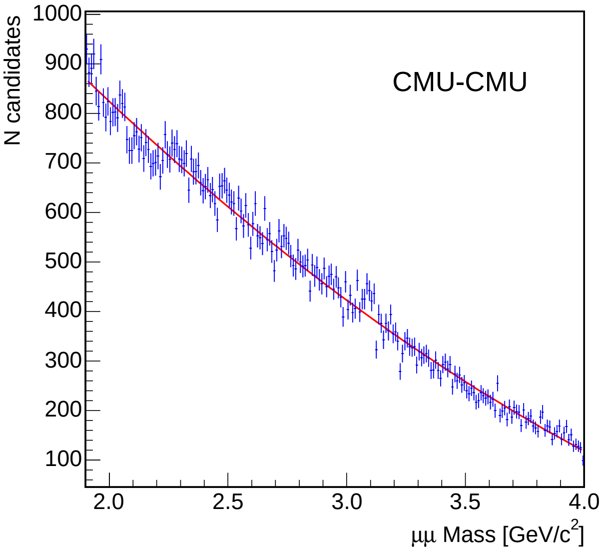

Figure 30: Mass distribution of the combinatorial background templates. The mass range from 1.84 to 1.89 GeV/c^2 is cut out in the plots because this is the signal search window.

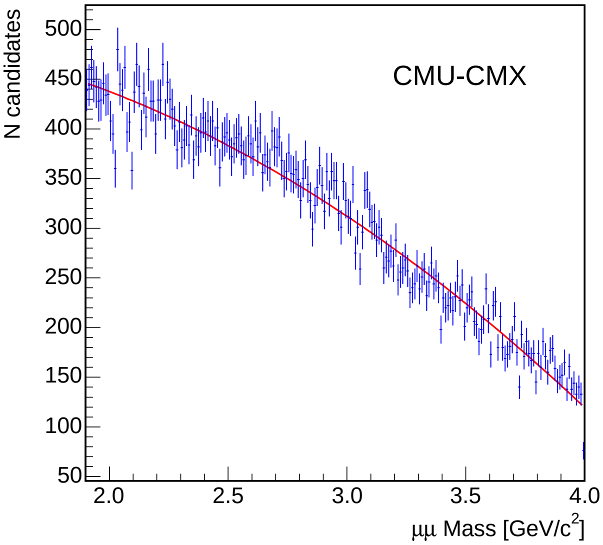

Figure 31: Mass distribution of the combinatorial background templates. The mass range from 1.84 to 1.89 GeV/c^2 is cut out in the plots because this is the signal search window.

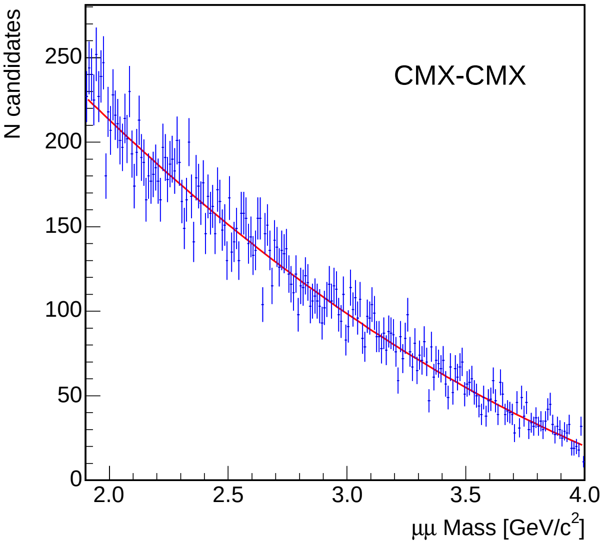

Figure 32: Mass distribution of the combinatorial background templates. The mass range from 1.84 to 1.89 GeV/c^2 is cut out in the plots because this is the signal search window.

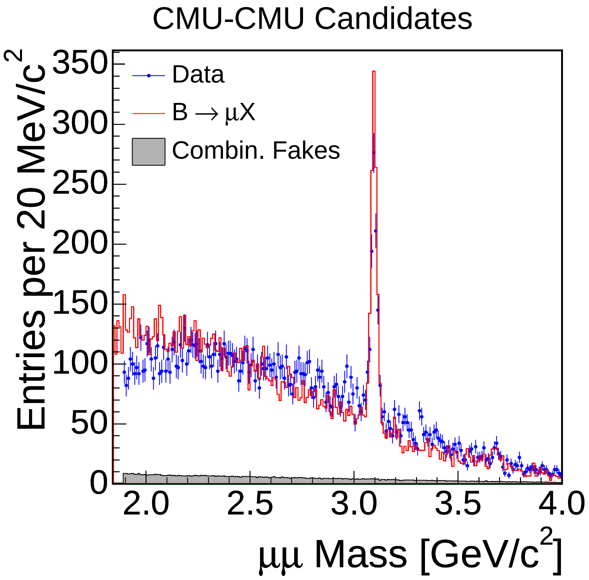

Figure 33: Comparison of the predicted and measured mass spectra for events with at least one leg reconstructed as a muon, in the CMU-CMU detector configuration.

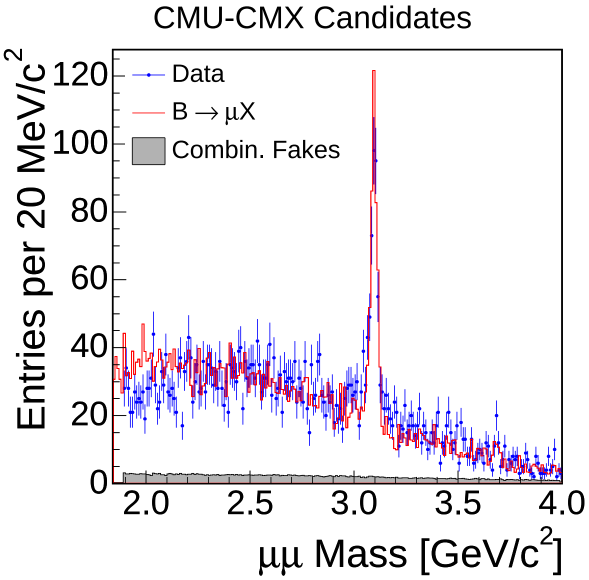

Figure 34: Comparison of the predicted and measured mass spectra for events with at least one leg reconstructed as a muon, in the CMU-CMX detector configuration.

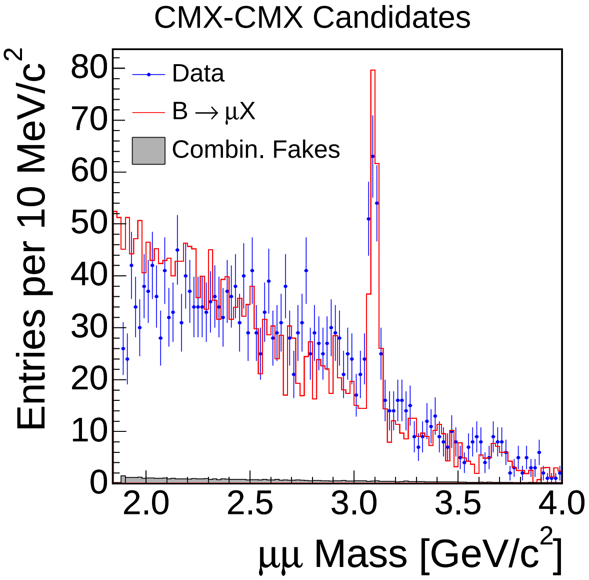

Figure 35: Comparison of the predicted and measured mass spectra for events with at least one leg reconstructed as a muon, in the CMU-CMX detector configuration.

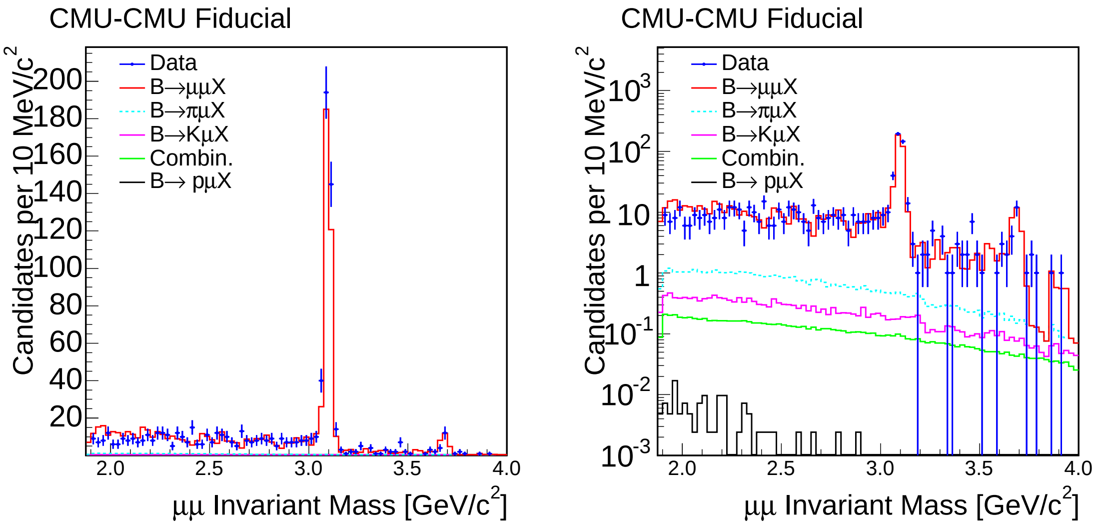

Figure 36: Comparison of the predicted and measured mass spectra for events with at least both legs reconstructed as a muon, in the CMU-CMU detector configuration.

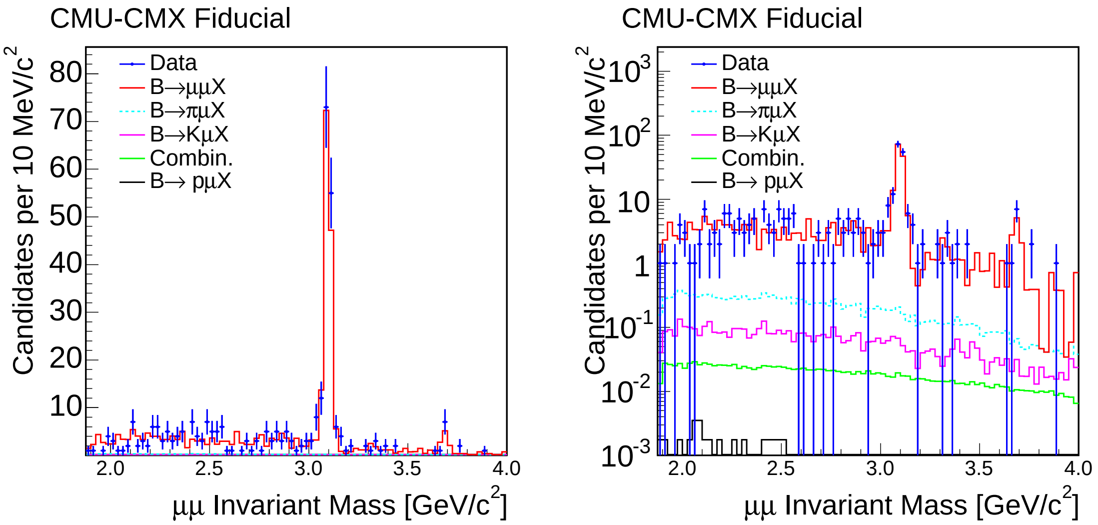

Figure 37: Comparison of the predicted and measured mass spectra for events with at least both legs reconstructed as a muon, in the CMU-CMX detector configuration.

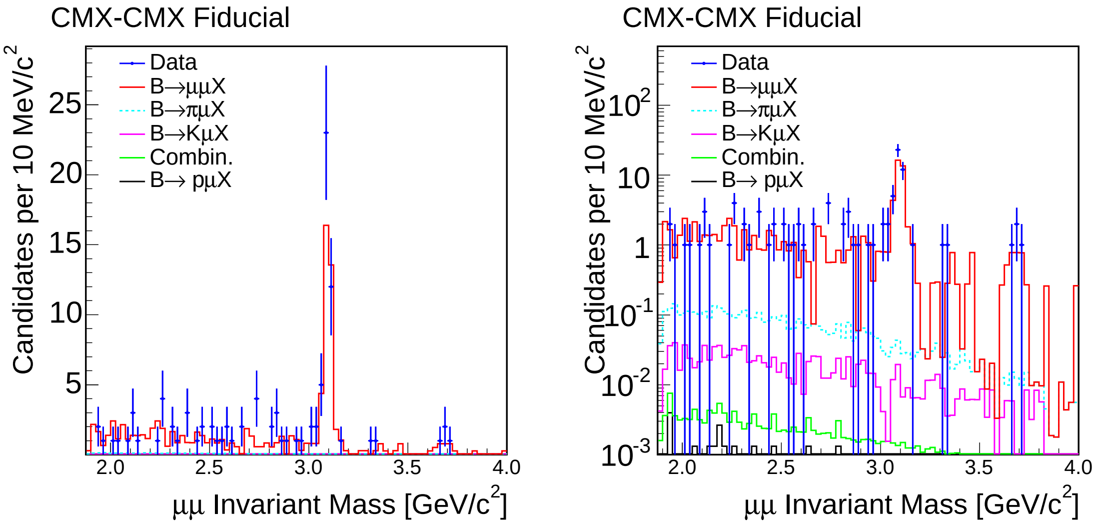

Figure 38: Comparison of the predicted and measured mass spectra for events with at least both legs reconstructed as a muon, in the CMX-CMX detector configuration.

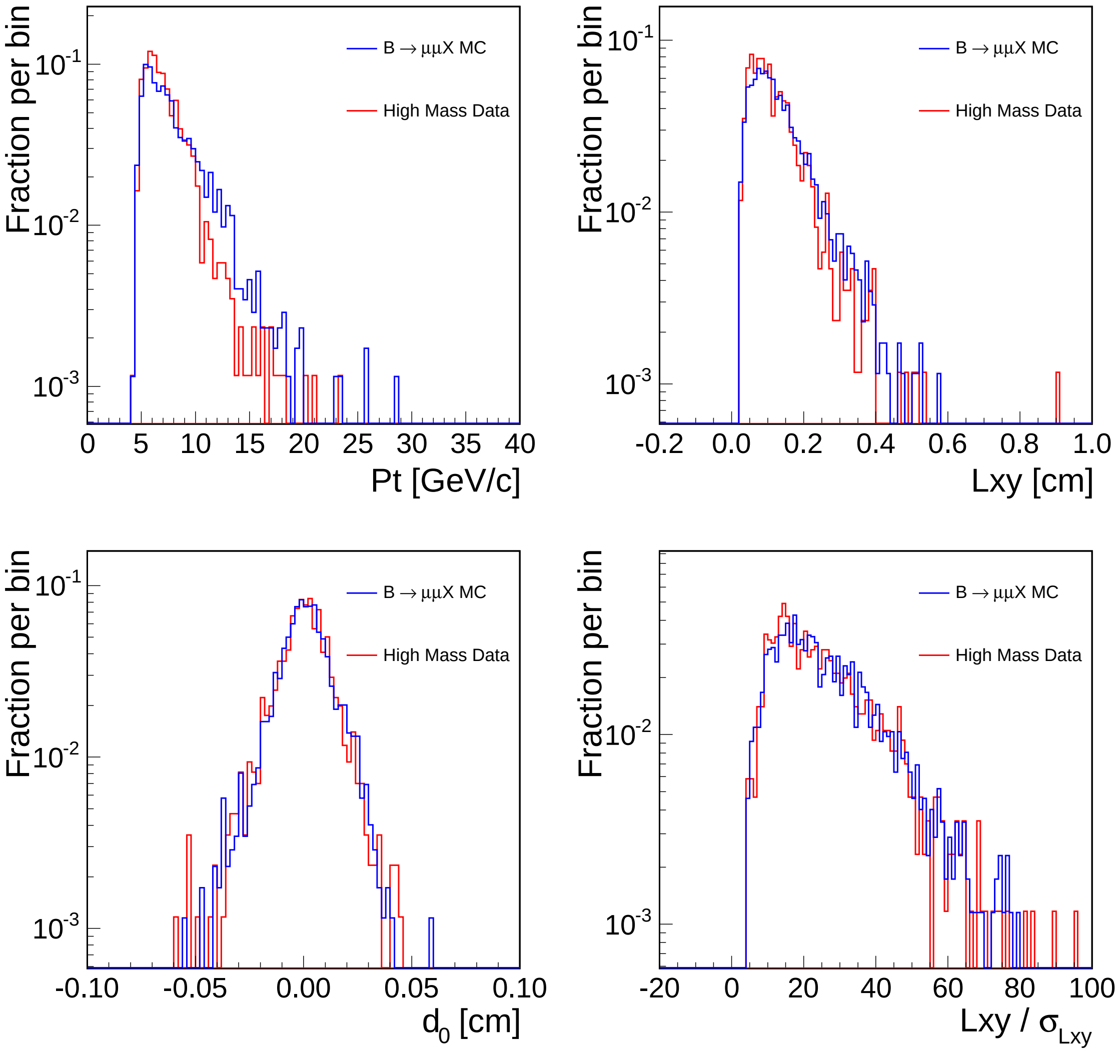

Figure 39: , Comparison of the distributions of the two-track transverse momentum (p_T, upper left), the displacement of the two-track vertex (L_{xy}, upper right panel), the two-track impact parameter d_0 (lower left panel), and the displacement significance ($L_{xy}/\sigma(L_{xy})$, lower right panel) of the two-track system between data and Monte Carlo simulation.

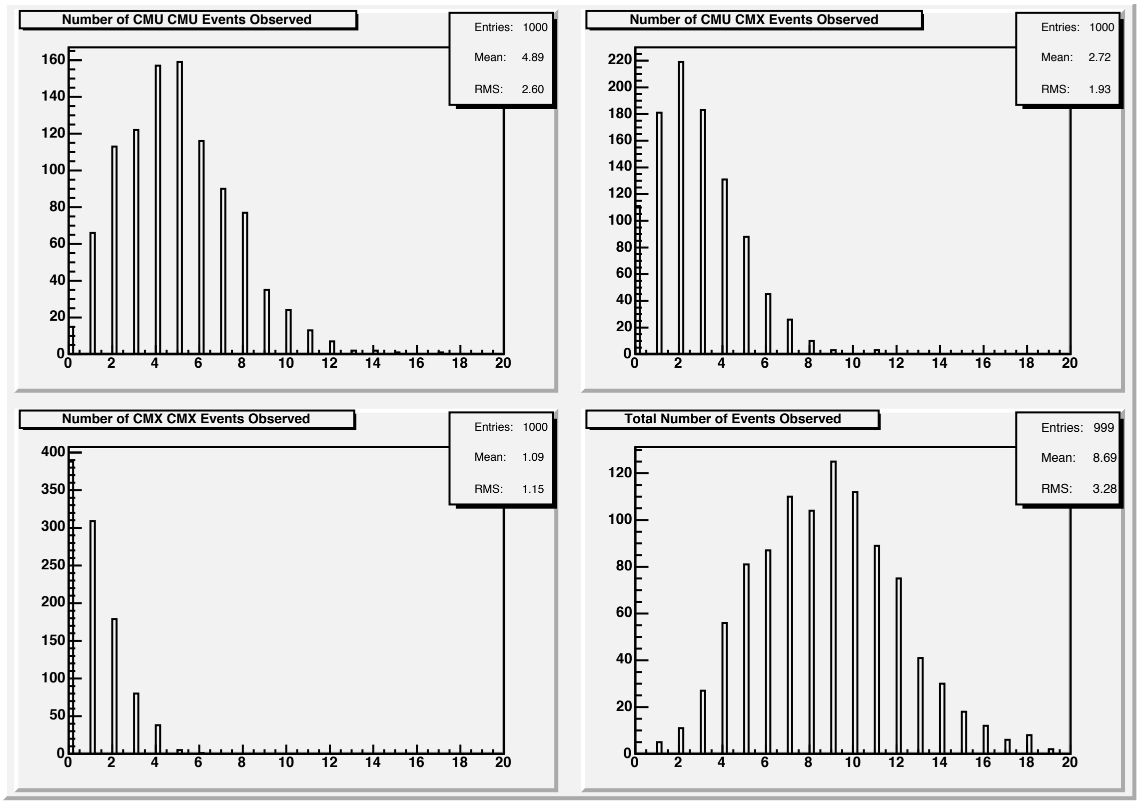

Figure 40: The distributions of numbers of observed events in (clockwise from upper left) CMU-CMU, CMU-CMX, total of all three, and CMX-CMX. The distributions for the three channels are approximately Poisonnian with mean equal to the expected number of background events in that channel, while the total distribution is more complex.

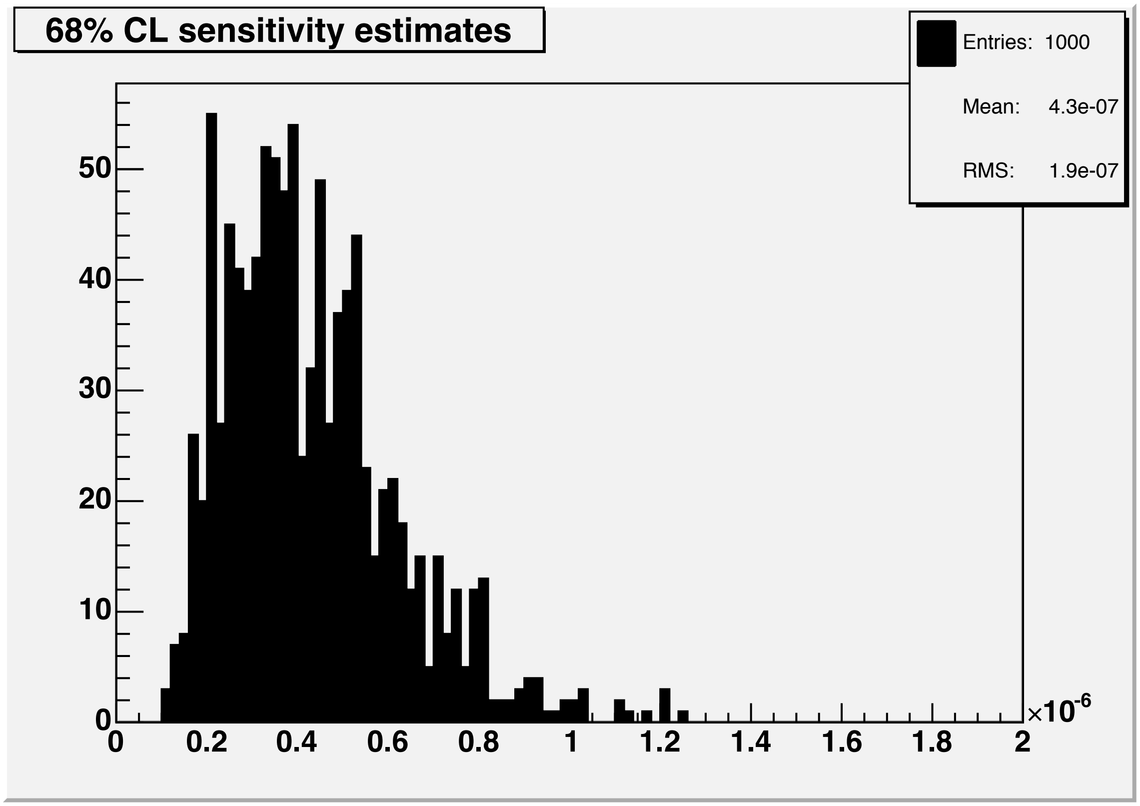

Figure 41: The distribution of limits at the 68% confidence level obtained for 1000 toy experiments.

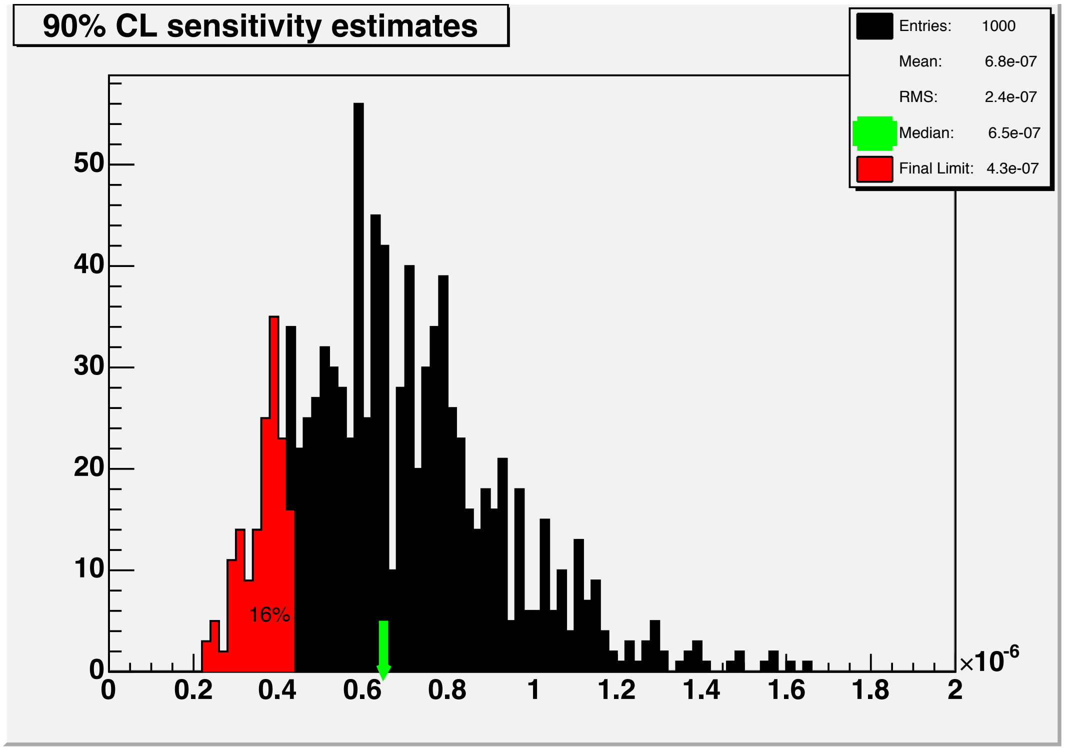

Figure 42: The distribution of limits at the 90% confidence level obtained for 1000 toy experiments. The median of the distribution is indicated by the green arrow. The red shaded area indicates those results equal to or less than the result obtained for our data.

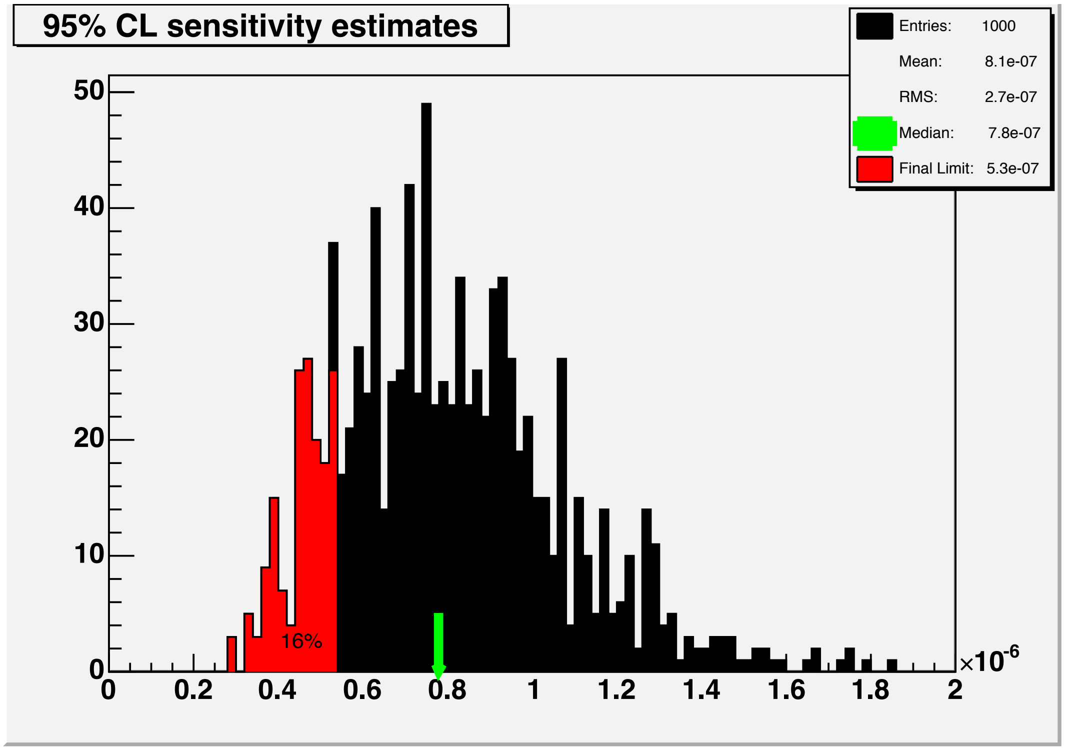

Figure 43: The distribution of limits at the 95% confidence level obtained for 1000 toy experiments. The median of the distribution is indicated by the green arrow. The red shaded area indicates those results equal to or less than the result obtained for our data.

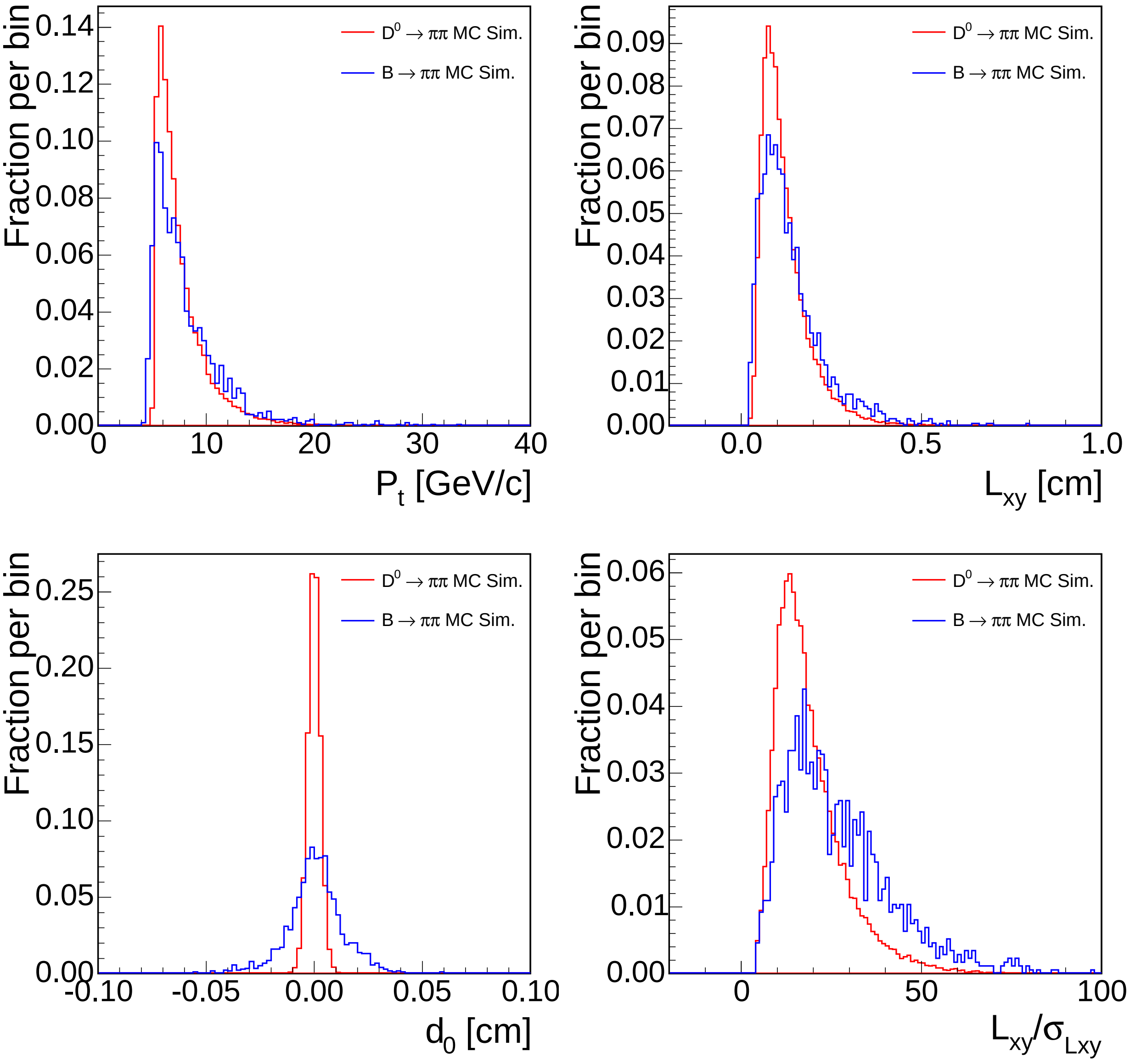

Figure 44: Comparisons of the distributions of various selection variables for $D^0 \to \pi^+\pi^-$ and $B \to \mu\mu X$ Monte Carlo simulation. In the top left panel, the p_T of the two-track pair is shown. In the top right, the L_{xy} of the decay vertex is displayed. On the bottom, the impact parameter distribution is shown in the left and L_{xy} significance in the right panel.

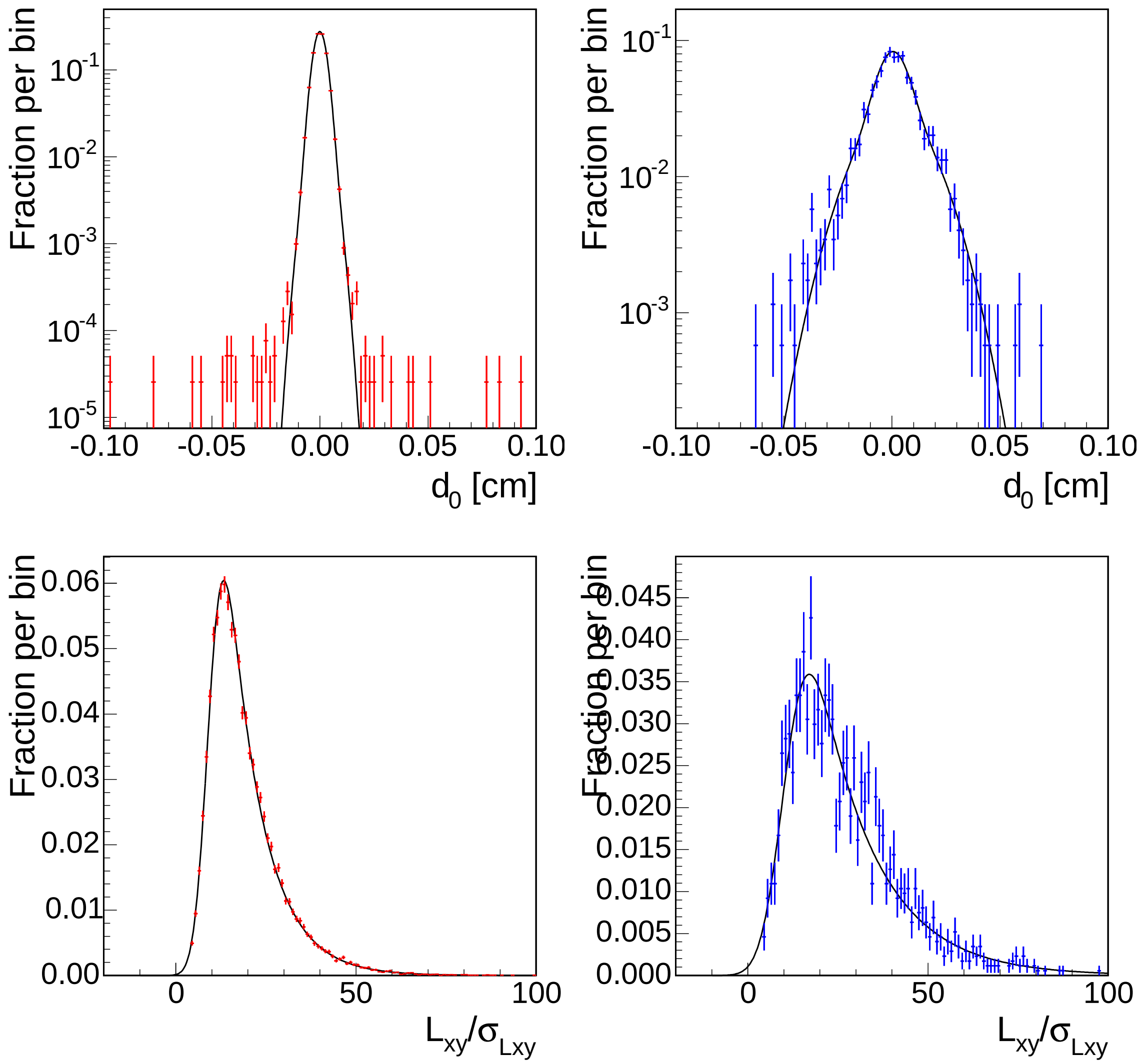

Figure 45: The distributions of the two-track system impact parameter (top) and L_{xy} significance (bottom) for $B \rightarrow \mu^+\mu^-X$ and prompt charm Monte Carlo simulation. The left column fits are fits to prompt charm and the right column fits are fits to $B \rightarrow \mu^+\mu^-X$ Monte Carlo simulation.

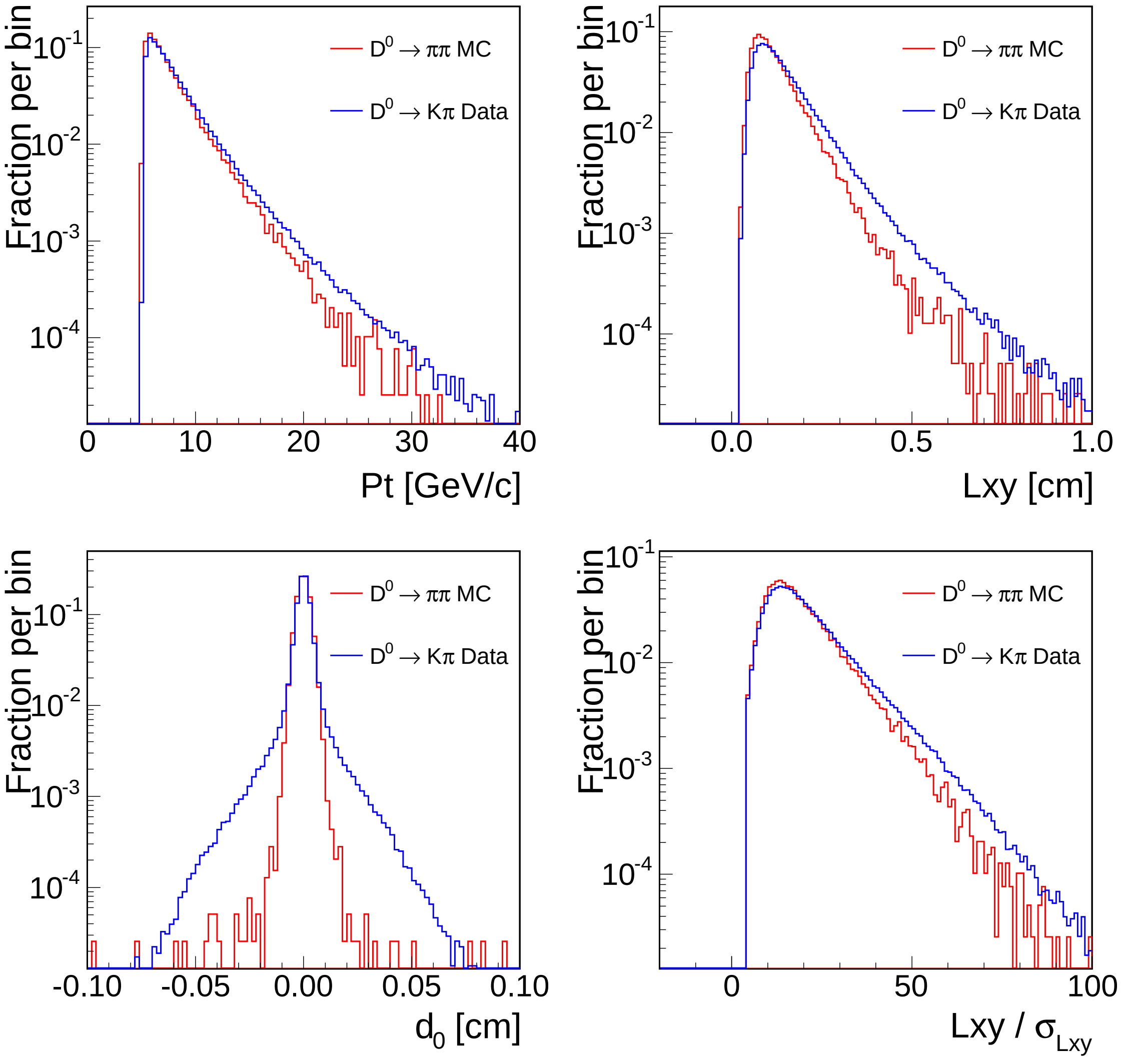

Figure 46: Comparison of Monte Carlo simulated quantities to those extracted from $D^0 \to K\pi$ decays in data. In the top left panel, the p_T distributions are compared. In the top right, the L_{xy} displacement is shown. The two-track impact parameter is compared in the bottom left panel, and the displacement significance is shown in the bottom right panel.

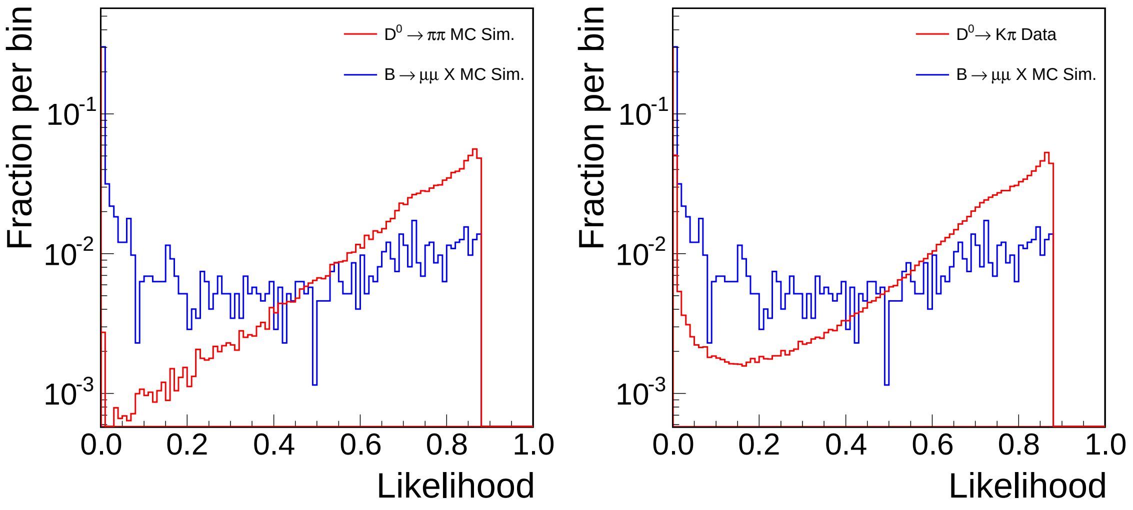

Figure 47: Comparison of likelihood shapes for D^0 events in data and Monte Carlo simulated D^0 decays. In the left panel, the charm likelihood is evaluated on prompt charm Monte Carlo. In the right panel, the charm likelihood is evaluated on $D^0 \to K\pi$ data. In both panels, the background likelihood is evaluated on $B \to \mu\mu X$ Monte Carlo.

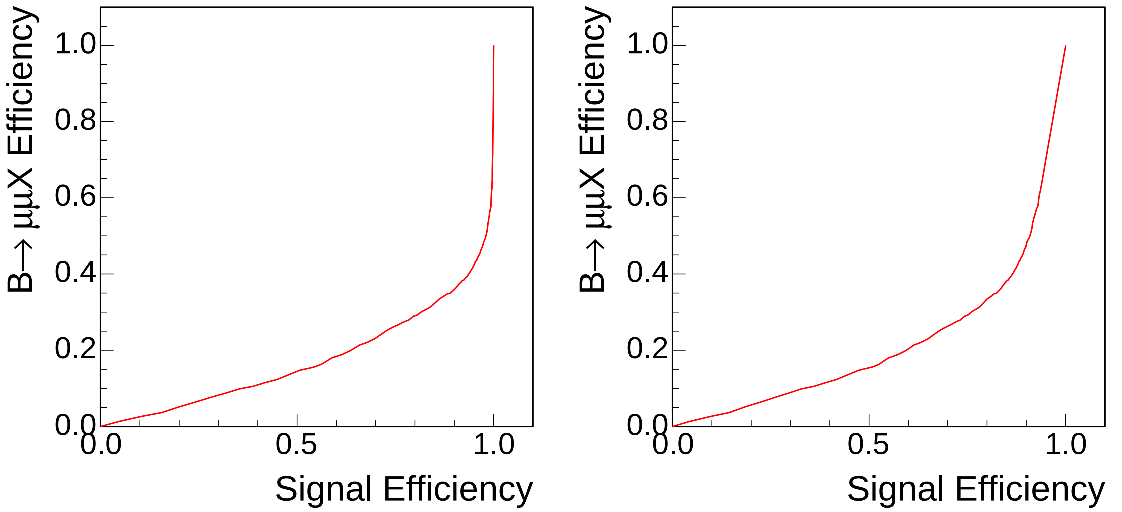

Figure 48: Expected signal and background efficiency behaviour, scanning over likelihood cut values. In the left panel, the signal efficiency is estimated from prompt charm Monte Carlo simulation, and in the right panel, the signal efficiency is estimated from $D^0 \to K\pi$ decays in data.

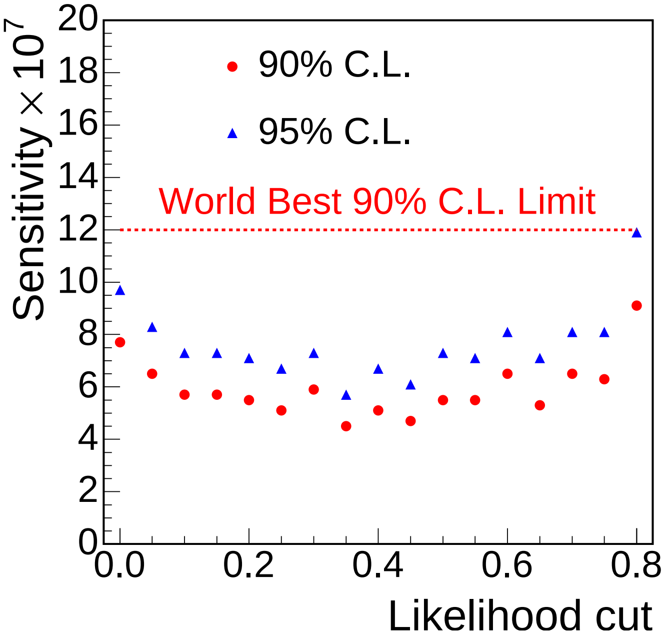

Figure 49: Scan of expected 90% and 95% CL limits for different bottom-charm separating likelihood cuts

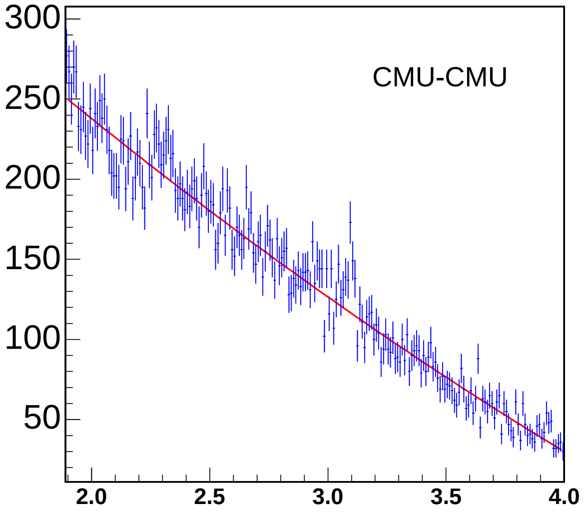

Figure 50: Mass distribution of the combinatorial background templates. The mass range from 1.84 to 1.89 GeV/c^2 is cut out in the plots because this is the signal search window, after applying the Bottom/Charm separation likelihood requirement.

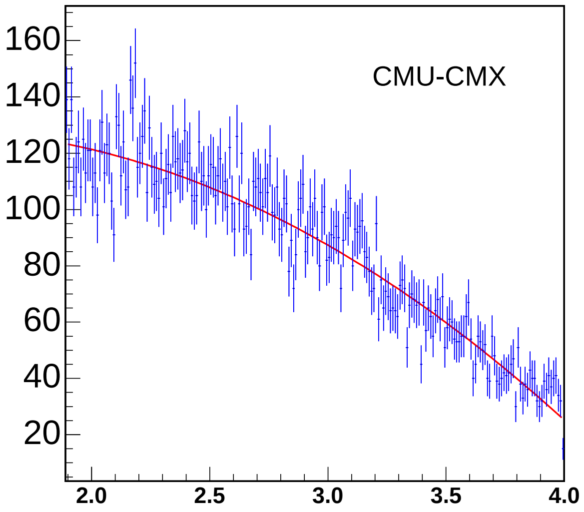

Figure 51: Mass distribution of the combinatorial background templates. The mass range from 1.84 to 1.89 GeV/c^2 is cut out in the plots because this is the signal search window, after applying the Bottom/Charm separation likelihood requirement.

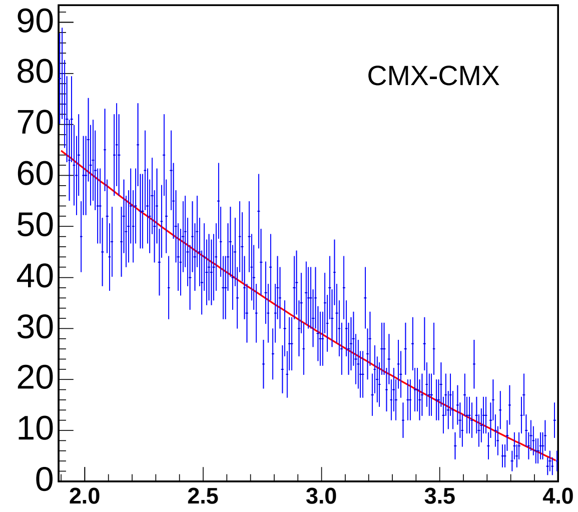

Figure 52: Mass distribution of the combinatorial background templates. The mass range from 1.84 to 1.89 GeV/c^2 is cut out in the plots because this is the signal search window, after applying the Bottom/Charm separation likelihood requirement.

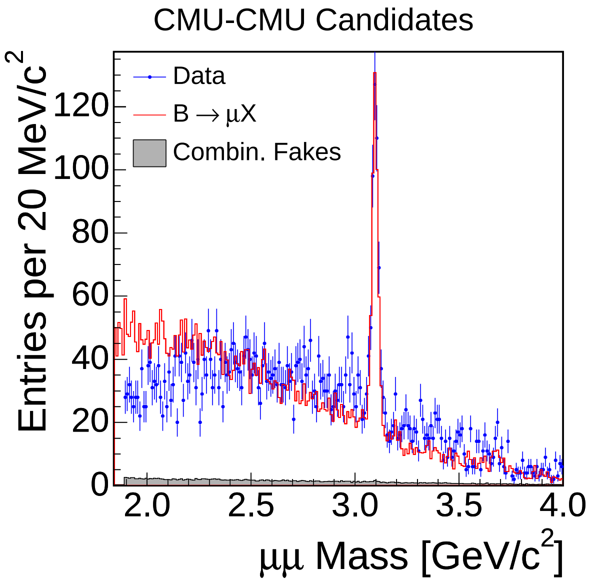

Figure 53: Comparison of the predicted and measured mass spectra for events with at least one leg reconstructed as a muon, in the CMU-CMU detector configuration, after applying the Bottom/Charm separation likelihood requirement.

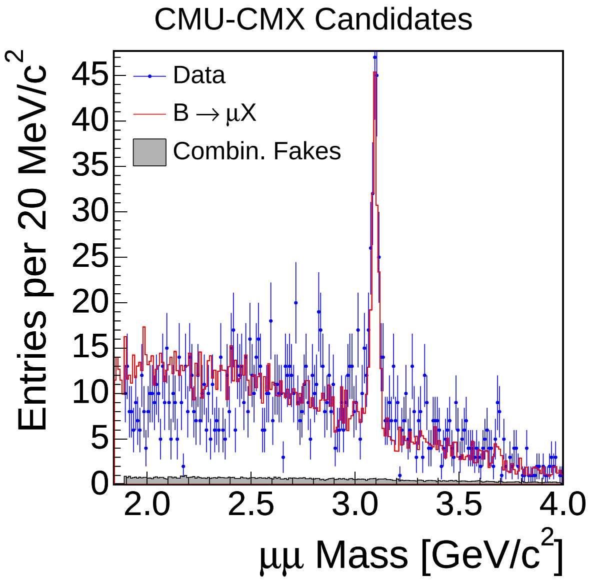

Figure 54: Comparison of the predicted and measured mass spectra for events with at least one leg reconstructed as a muon, in the CMU-CMX detector configuration, after applying the Bottom/Charm separation likelihood requirement.

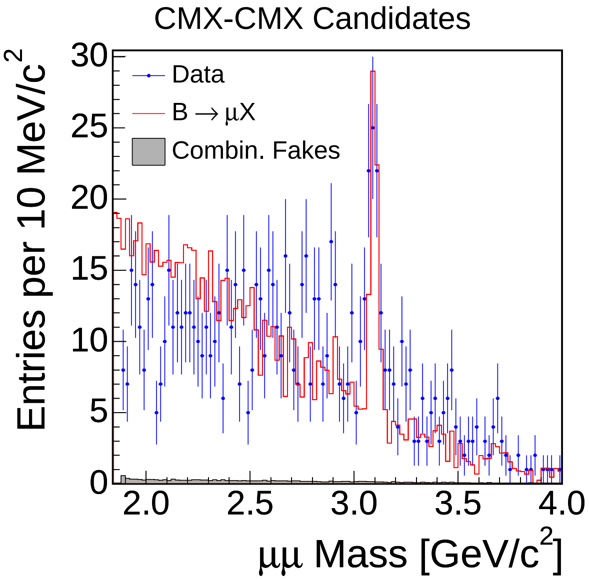

Figure 55: Comparison of the predicted and measured mass spectra for events with at least one leg reconstructed as a muon, in the CMU-CMX detector configuration, after applying the Bottom/Charm separation likelihood requirement.

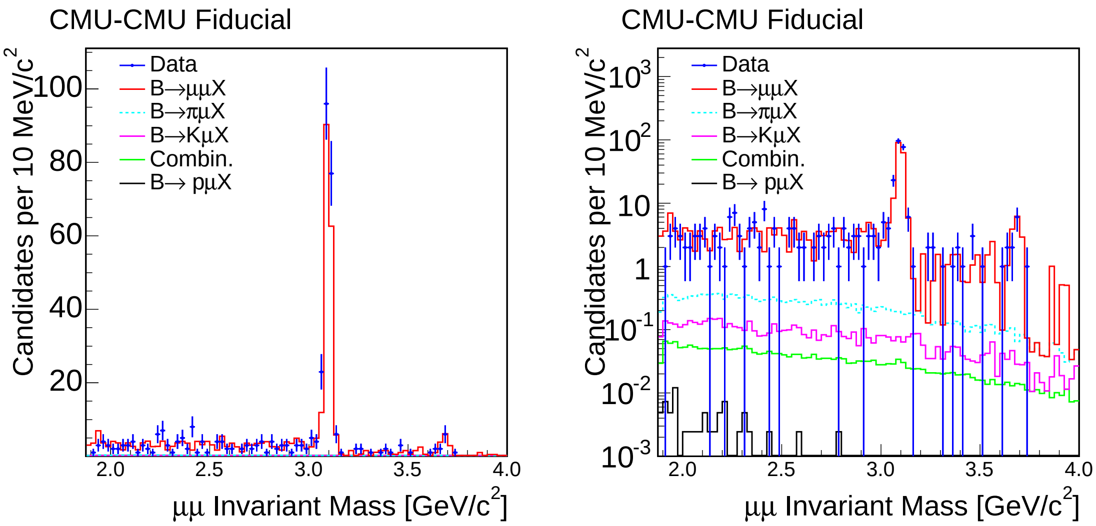

Figure 56: Comparison of the predicted and measured mass spectra for events with at least both legs reconstructed as a muon, in the CMU-CMU detector configuration, after applying Bottom/Charm Likelihood separation requirement.

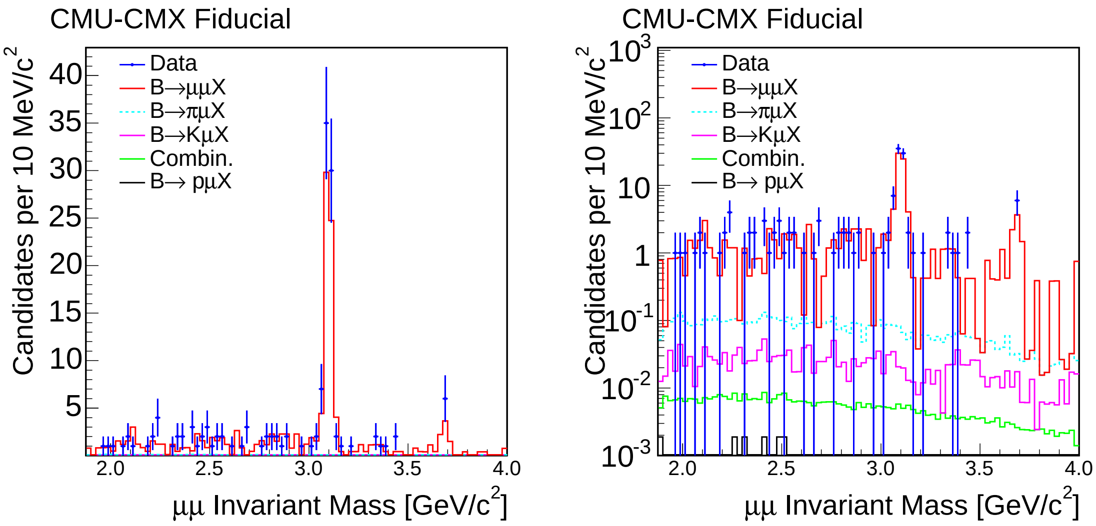

Figure 57: Comparison of the predicted and measured mass spectra for events with at least both legs reconstructed as a muon, in the CMU-CMX detector configuration, after applying Bottom/Charm Likelihood separation requirement.

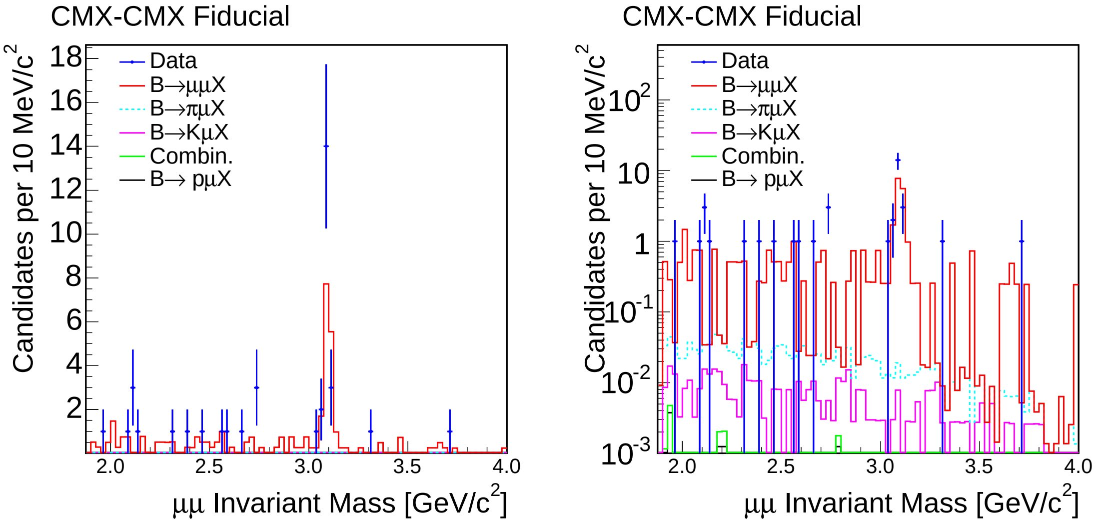

Figure 58: Comparison of the predicted and measured mass spectra for events with at least both legs reconstructed as a muon, in the CMX-CMX detector configuration, after applying Bottom/Charm Likelihood separation requirement.

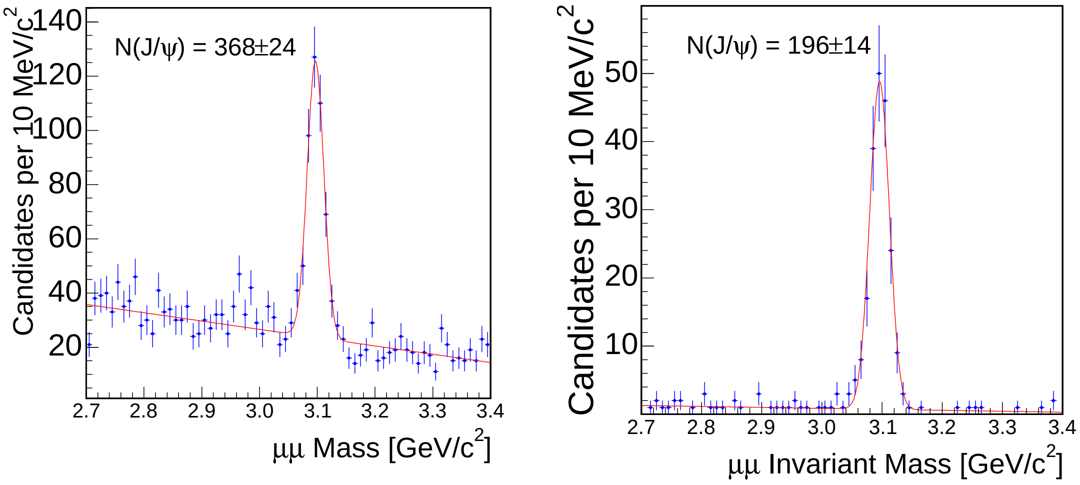

Figure 59: Fits of the J/ψ yield in single-tagged (left panel) and double-tagged (right panel) samples

References

[1] G. Burdman, E. Golowich, J. Hewett, S. Paskvasa, “Rare Charm Decays in the Standard Model and Beyond”, `arXiv:hep-ph/0112235` v2 (March 2002)

[2] B. Aubert *et al.* (BaBar Collaboration) Phys. Rev. Lett. 93 191801 (2004)

[3] W.-M. Yao *et al.*, Journal of Physics **G** 33, 1 (2006)

[4] E. Berry, I.K. Furic., ”Study of Muon Matching Efficiencies and Mistag Rates”, CDF Note 8042 (2006)

[5] W.J. Ashmanskas and R. Harr., “Search for the FCNC decay $D^0 \to \mu\mu$ using the hadronic B trigger”, CDF Note 6723 (2003).

[6] K. Bloom and D. Dagenhart., “Muon-Reconstruction Efficiency”, CDF Note 6029 (2002).

[7] J. Heinrich, “Bayesian limit software; multi-channel with correlated backgrounds and efficiencies”, CDF Note 7587 (2005).