CDF/ANAL/EXOTIC/CDFR/9239
Version 2.0
March 17, 2008

Study of Fake-Lepton Rates for the High-p_T Like-Sign Dilepton Analysis to Search for the Wh Production

Toru Okusawa, Yoshihiro Seiya, Takayuki Wakisaka, and Kazuhiro Yamamoto

Osaka City University, Japan

Abstract

Fake-lepton rates are investigated using generic jet samples. These rates are used in our like-sign dilepton (ee, $e\mu$, and $\mu\mu$) analysis for the fermiophobic higgs and high-mass Standard-Model higgs searches to estimate fake-lepton backgrounds. The rates are defined by those relative to the rates of isolated tracks with certain energy depositions in the calorimeters which we call the denominator objects. The numerator objects are isolated lepton objects passing our selection criteria in the jet samples as usual. The particular definition of the denominator objects is chosen so that real-lepton contamination is minimized in applications especially to opposite-sign dilepton events, which, otherwise, would lead to overestimates of fake-lepton backgrounds.

1 Introduction

Fake leptons are one of major backgrounds in the like-sign (LS) dilepton events as known in the previous analysis [1]. They were estimated by weighting lepton + isolated track events with the expected fake-lepton yield for a given isolated track. These rates used to estimate fake-lepton backgrounds are called the fake-lepton rates, and are defined with respect to some reference rates, the rates of denominator objects. The analysis described in this note to estimate the fake-lepton rates is basically the same as before but we change the definition of the denominator objects. In the previous analysis, we used naive isolated tracks as the denominator, thus, the fake-lepton rates was also called the fake-lepton probabilities. It is good enough if we analyze only LS dilepton events, but now we want to look at opposite-sign (OS) dilepton events as well in order to understand our LS analysis better. In this situation, it is expected that the simple isolated-tracks in the OS combination are significantly contaminated by real leptons from Drell-Yan processes, which leads to overestimates of fake-lepton backgrounds. To avoid this problem and to establish a consistent scheme which can be applied to both the OS and LS cases, we choose isolated tracks that deposites certain energies in the electromagnetic (EM) and hadron (HA) calorimeters in the way such that they are not likely to be induced by real leptons. One could use LS tracks to estimate fake-lepton backgrounds in the OS dilepton events. We, however, reserve a handle to properly take into account of possible charge correlation between two lepton-objects in fake-lepton background events.

2 Fake-lepton backgrounds

The lepton plus fake-lepton backgrounds arise typically from a single lepton event such as $W \rightarrow \ell\nu$. This type of backgrounds consist of one trigger lepton and one fake lepton. The components of the "fake lepton" are

- Fake leptons
    1. Interactive $\pi^{\pm} \rightarrow$ fake electrons,
    2. Overlap of π^0 and a track $\rightarrow$ fake electrons,
    3. Punch-through hadrons $\rightarrow$ fake muons,
- Non-prompt leptons
    1. Residual photon conversions $\rightarrow$ electrons
    2. Decay-in-flight muons from $\pi^{\pm}$ and $K^{\pm} \rightarrow$ muons,
    3. (Semi-)Leptonic decay of heavy-flavor hadrons $\rightarrow$ leptons.

As noted here, we use "fake leptons" as a generic word to mean both the literal fake leptons and non-prompt leptons. Most of the components are considered to be non-isolated and quite common in generic QCD events, while the residual photon-conversions are not necessarily QCD specific, and they are separately estimated from identified conversions with a similar philosophy as the fake-lepton rates. Details of the estimation of residual photon-conversion backgrounds are discussed in a separate CDF note [2]. Correspondingly, contributions of residual conversions are subtracted from fake-electron rates in this study.

3 Measurements of fake-lepton rates

We define the fake rate R_{fake} as a rate of fake leptons relative to isolated tracks with certain energy depositions especially in the hadron calorimeters, which we call "hadronic" tracks:

$$R_{\text{fake}} = \frac{\text{lepton objects passing nominal selections and considered to be fakes}}{\text{isolated tracks with required calorimeter energy depositions}} \tag{1}$$

It is noted that we define the fake rate here per isolated track, not per jet. The event pre-selection and the isolated hadronic-track selection criteria are listed in Table 1. The kinematical cut is $p_T > 6$ GeV/c and we require the same track-quality cuts as used in the LS dilepton analysis [3]. The hadronic-track selection imposes cuts on the energy depositions in the calorimeter towers that a given track is pointing. The first cut in the table is thought to mainly reject real electrons and the other cuts are meant to reject real muons. The effects of these cuts on the selection of denominator objects in inclusive lepton samples are briefly discussed in §5. As to the numerator, we apply the nominal lepton selection cuts to find fake-lepton objects.

We evaluate R_{fake} using inclusive jet samples. The data we use are the `Jet20, Jet50, Jet70,` and `Jet100` generic jet samples from the `0d` through the `0i` datasets[1]. The same good-run filtering is performed as in the LS analysis [3]. In order to remove real-leptons from the numerator side, we impose a W veto and a Z veto. The W veto rejects events with

- electrons with $E_T > 20$ GeV or muons with $p_T > 20$ GeV/c passing our lepton selection [3],
- $\not\!\!E_T > 30$ GeV or transverse mass > 40 GeV/c^2.

The $\not\!\!E_T$ is corrected for muons passing our muon selection with the $p_T > 6$ GeV/c cut. The Z veto rejects events with

- electrons with $E_T > 20$ GeV or muons with $p_T > 20$ GeV/c passing our lepton selection,
- invariant mass with the 2nd-leg objects is in the mass window of 81–101 GeV/c.

The definition of the 2nd-leg objects for the Z veto is given in Table 2. Note that the EM objects for the 2nd-leg are not necessarily in the central region. Finally, we require the hadronic tracks and fake-lepton objects be both separated from the trigger jet in the η-ϕ space by $R > 1.0$ to remove possible trigger biases.

Figure 1 shows the p_T distributions of isolated hadronic tracks, fake electrons, and fake muons for each jet sample combined over the datasets from the `0d` through the `0i`. The corresponding fake rates are shown in Fig. 2. Comparisons of fake-electron and fake-muon rates between different datasets to check the stability are shown in Figs. 3 and 4, respectively. From the rates shown in Fig. 2, we obtain the final fake rates by combining jet samples with the RMS spread as a systematic uncertainty. The result is shown in Fig. 5 with breakdown of fake-muon rates into those of CMUP and CMX muons. To be complete, we show the fake rates as a function of isolation and η in Figs. 6 and 7, respectively. In the current version of the LS analysis, we use the fake rates as a function of p_T and we do not distinguish CMUP and CMX muons.

[1] In this note, we use "jet samples" to mean different types of jet data collected by different triggers (e.g. `Jet20`), while "datasets" to mean different data according to when the data-taking occurred (e.g. `0d`).

4 Subtraction of residual photon-conversions

Residual photon-conversions in the fake-electron objects found in the jet samples are estimated from the amount of identified conversions multiplied with weights for residual photon-conversions infered from them, which we refer to as residual conversion ratios. The details of the method itself is described in [2]. We show only the results in this section. Figure 8 shows the p_T distributions of estimated residual conversions for each jet sample combined over the datasets. The stability over the datasets is checked in terms of the rates relative to the hadronic-track rates and shown in Fig. 9. The fake-electron rates after subtraction of the estimated residual conversions are shown in Fig. 10 for each jet sample combined over the datasets. Similar to the original fake-electron rates, we further combine the corrected fake-electron rates over the different type of jet samples, which is shown in Fig. 11. In this case, the uncertainties shown in the figure include a systematic uncertainty from the residual conversion estimation, quantified by $\pm 1\sigma$ variations of the residual conversion ratios, as well as the RMS spread over the jet samples.

5 Real-lepton contamination and charge correlation

In this section, we briefly check an implication of hadronic-track selection in applications to inclusive lepton data. Figure 12 shows effective cross sections of μ + track objects of a few types. For this plot, the track object search in a given event is performed only when there is exactly one good muon passing all the muon selection cut with the high p_T cut of $p_T > 20$ GeV/c. Also, the simple isolated tracks in this analysis are checked if they are associated with any good leptons or any EM objects with conversion flags, and removed if so. The $\mu\mu$ events and the μ + isolated-track events are therefore mutually exclusive.

First of all, we can confirm from a strong Z peak that simple isolated tracks of the OS case are significantly contaminated by real muons. Since the fake-muon rates are about at a few % level as obtained in the previous sections, the estimated fake backgrounds from the isolated tracks would be also a few % level of dimuon candidates. It is actually not a serious problem compared to the large contributions from real dimuon events, althought the estimated value of fake backgrounds itself would be orders of magnitude larger than the correct answer. We also see that the isolated track rates get larger than those of the dimuon candidates toward low mass regions, which indicate more backgrounds at low masses as expected. For the case of isolated hadronic tracks, the situation is greatly improved. We are picking up relatively more backgrounds with suppressing real muon contamination. Remembering again the fake-muon rates are about a few %, the fake-lepton backgrounds are negligibly small in the OS dimuon events in rather a wide invariant mass range.

One of interesting thing is a comparison between the OS and LS combinations of μ + hadronic tracks (see Fig. 12). The OS rates are larger than the LS rates even outside Z mass regions, which could be a sign of charge correlation besides the effects due to real dimuon events. In order to get a better measure, we investigate real-muon contamination in the hadronic tracks with helps of `PYTHIA` Monte Carlo (MC) events. We use inclusive $W \to \mu\nu$, $Z \to \mu\mu$, and $Z \to \tau\tau$ MC samples[2] to collect μ + hadronic tracks, and compare with the data as given in Fig. 13. We see that the observed μ + hadronic tracks are qualitatively described by MC events, and also see that contributing mass regions are different between the processes. The gradual rise of the rates toward low mass regions can be understood as coming mainly from contributions of W events. The MC predictions of the absolute value do not completely agree with the observed effective cross section due to some unknown reasons. We, however, try to obtain the fraction of real muons in these MC events by looking at `OBSP` information for which systematics would be cancelled to some extent. The components of real muons in the MC events are shown in Fig. 14. The real-muon contamination is due to $Z \to \mu\mu$ events and the small peak around Z mass regions are completely attributed to them as expected[3]. Regarding this, we comment that Z mass regions will be eventually excluded from the final analysis, which provides us a certain safety factor against the problem of real-muon contamination. Real-muon fractions, f, as a function of invariant mass are estimated as shown in Fig. 15, and a correction for real-muon contamination in the observed OS μ + hadronic tracks can be done by multiplying $1 - f$. The OS/LS ratios after the correction is shown in Fig. 16. The result is not statistically powerful, but there seem to still remain indications of the charge correlation in the data which are as much as a factor of 2 in a mass region around 40–50 GeV/c^2. The MC predictions obtained by explicitly rejecting real-muons using `OBSP` information are also included in the figure. It shows a similar level of charge correlation, and further says that the correlation is originating from W events, not from $Z \to \mu\mu$ events. The $Z \to \tau\tau$ strongly favors the OS combination as expected but the contribution itself is relatively small compared to the W events, thus the net effect is driven by the W events. From these investigation, we think that there may be charge correlations between leptons and additional track objects, thus the estimation of the fake-lepton backgrounds in the OS dilepton events using LS combinations is not necessarily ensured to lead to a correct answer. We can also say that the charge correlation may depend on the selection and on the kinematical space that we look at, and not always important especially when W events are not contributing much. For example, it might be reasonable to use LS combinations in the Z mass window. In general, we must always bear this in our mind when we analyze the OS dilepton events.

6 Fake-lepton rates for silicon tracks

It is our plan to require a cut on the number of silicon hits of tracks in the final analysis to further suppress residual photon-conversion backgrounds. The cut is

[2]These are MCs of the electroweak group. The datasets are `wewk8m` ($W \to \mu\nu$), `zewk9m` ($Z \to \mu\mu$), and `zewkat` ($Z \to \tau\tau$).

[3]Contributions from $W \to \mu\nu$ and $Z \to \tau\tau$ events are too small to be clearly visible in the plot.

$$\text{the number of Si hits} \geq 3,$$

and will be included as one of the track-quality cuts. If this cut is truned on, it will change both the denominator rates and the numerator rates, especially the fake-electron rates. We present the main results when the Si cut is included in the analysis. The good-run filtering is properly performed with requiring only the good silicon runs.

Figure 17 shows the p_T distributions of isolated hadronic tracks, fake electrons, and fake muons for each jet sample combined over the datasets. The rates are changed very slightly[4], about 10% for all the objects. The resulting fake rates are shown in Fig. 18. The p_T spectra of residual photon-conversions are shown in Fig. 19. The effective cross sections are reduced to about 30–40% of what is without the Si hit cut, which is due to reductions of identified conversions passing the selection. The corrected fake-electron rates are shown in Fig. 20. As expected, the effect of the conversion subtraction is now small and the corrected fake-electron rates are consistent with the rates obtained without applying the Si hit cut (see Fig. 11). The Si hit cut is, therefore, mainly affecting photon-conversions in the fake-lepton backgrounds.

7 Conclusions

We obtained fake rates to be used for fake-lepton background estimation in the like-sign dilepton analysis. The fake rates are defined as ratios between the fake-lepton objects passing our lepton selection criteria and the isolated tracks with certain energy depositions in the calorimeters, which were both collected using generic jet samples. The denominator was constructed so as to minimize real-lepton contaminations from Drell-Yan processes which was forseen in the application to opposite-sign dilepton events performed as a check of the like-sign analysis.

From the jet samples combining the `0d` through `0i` datasets, we found the fake-electron rates are about 8% to 2% depending on the track p_T, and fake-muon rates are about 4% to 2%. The stability of the fake rates in different datasets was checked to find no significant problem. We also obtained the corrected fake-electron rates by subtracting residual photon-conversions and they were about 1% (absolute) smaller than the rates before subtraction.

Charge correlations between leptons and track objects were investigated using μ + track events. We confirmed possible correlations especially in the kinematical region where W events were expected to contribute. This would validate our motivation for sticking to OS combination of lepton + track events and for using hadronic tracks.

Finally, in order to prepare for the final analysis with a cut on the number of silicon hits of tracks to reduce residual photon-conversion backgrounds in like-sign dilepton events, we obtained the fake rates with the cut imposed as a track-quality cut.

[4]Note that the integrated luminosity is changed by about 7% but we are showing effective cross sections in the plot.

Table 1: Event pre-selection and the denominator track selection.

<u>Event pre-selection</u>
$\lvert z_{\mathrm{pv}}\rvert < 60$ cm
Cosmic-ray veto
<u>Geometrical and kinematical cuts</u>
CDF default tracks in the central region
$p_T > 6$ GeV/c
<u>Track quality cuts</u>
axial ≥ 3 and stereo ≥ 3 (≥ 7 hits)
$\lvert z_0 - z_{\mathrm{pv}}\rvert < 2$ cm
$\lvert d_0\rvert < 0.2$ cm (silicon hits < 3), 0.02 cm (≥ 3)
<u>Isolation cut</u>
$\mathrm{ISO}_{0.4}^{\mathrm{cal}} < 2$ GeV
<u>Hadronic-track selection</u>
HA/EM > 0.2
EM > 1 GeV
HA > 4 GeV

Table 2: Physics objects used to identify and remove Z bosons.

Track object
Opposite-sign
$p_T > 10$ GeV/c
Track cone-isolation < 4 GeV/c
$\lvert z_0 - z_{\mathrm{pv}} \rvert < 10$ cm
EM object
$E_T > 10$ GeV
HAD/EM < 0.12
Fractional isolation, $\mathrm{ISO}_{0.4}^{\mathrm{cal}}/E_T < 0.15$
Muon object
$p_T > 10$ GeV/c
EM < 5 GeV
HAD < 10 GeV
Fractional isolation, $\mathrm{ISO}_{0.4}^{\mathrm{cal}}/p_T < 0.15$
$\lvert z_0 - z_{\mathrm{pv}} \rvert < 10$ cm
$\lvert d_0 \rvert < 0.5$ cm

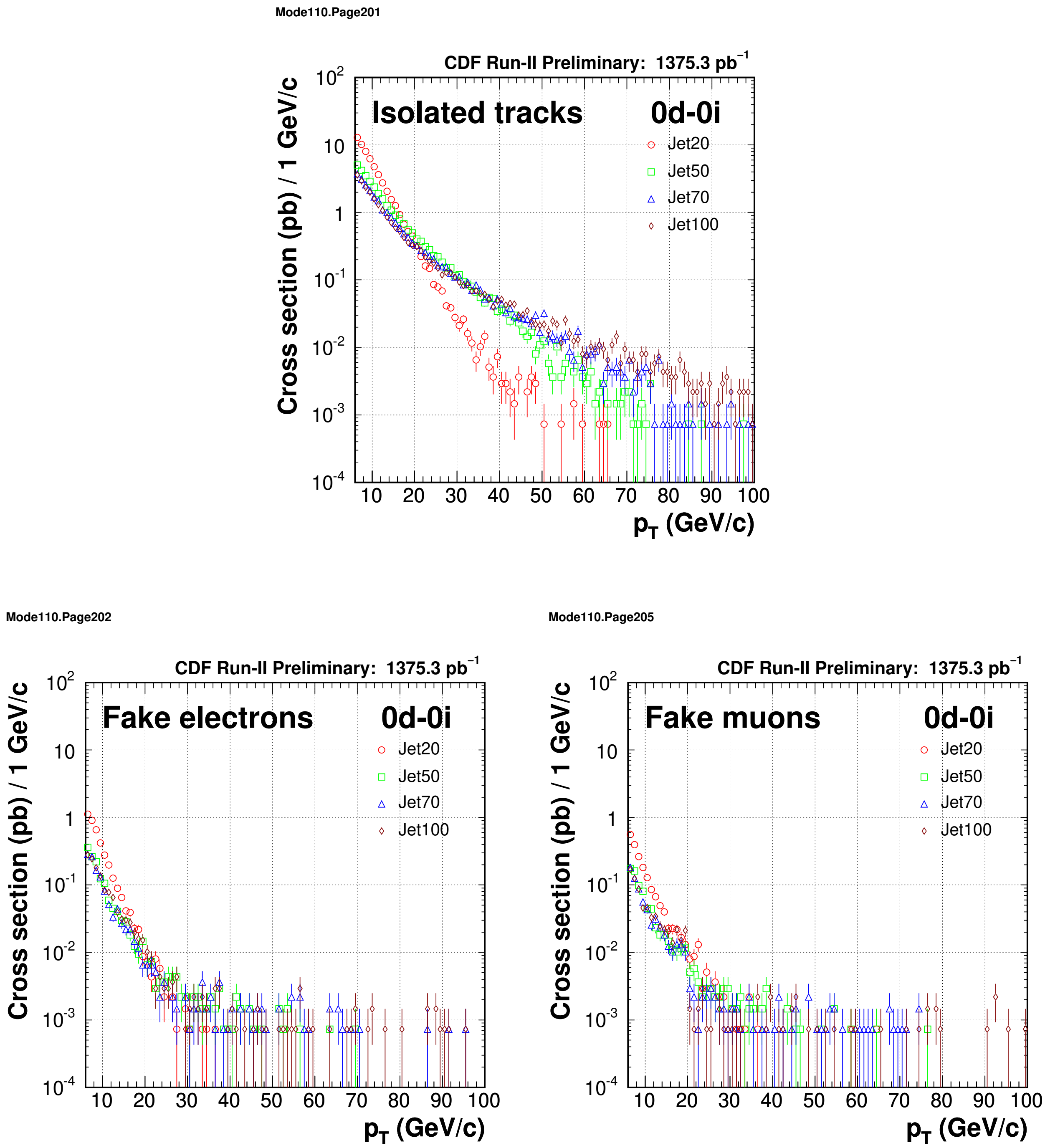
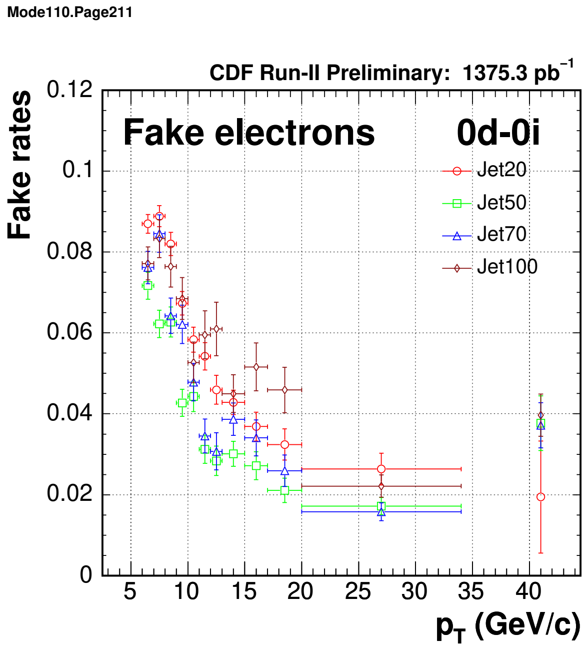


Figure 1: Track p_T distributions of isolated hadronic tracks (top), fake electrons (lower left), and fake muons (lower right) in the jet samples combined over the datasets.



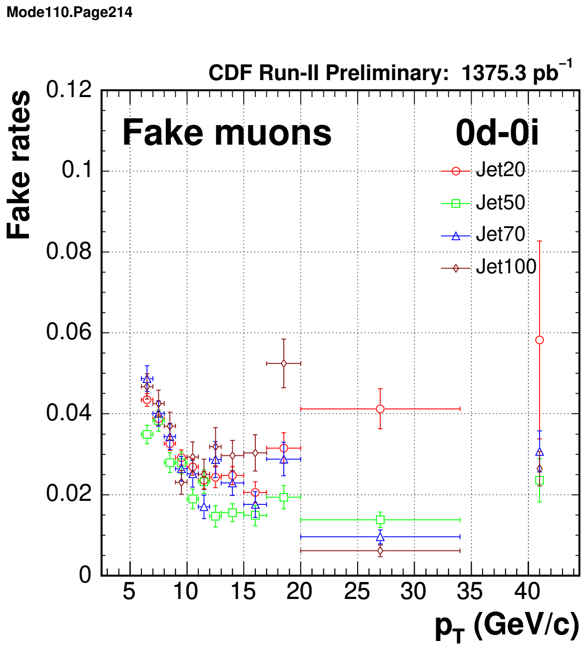


Figure 2: Fake-electron (left) and fake-muon rates as a function of track p_T for each of the jet sample combined over all the datasets. The last data points in the plot correspond to the fake rates including overflow events.

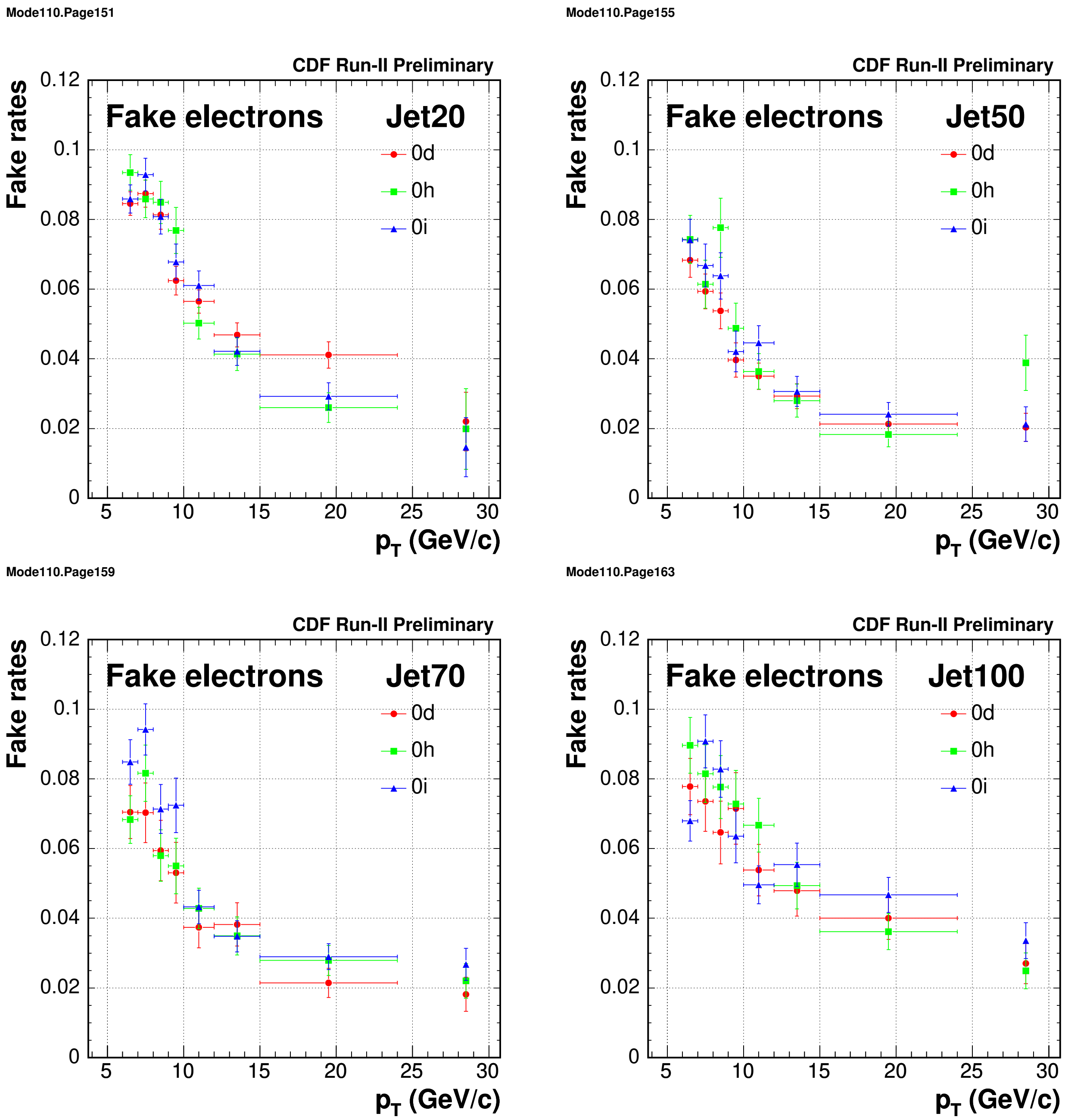


Figure 3: Comparisons of fake-electron rates among the `0d` through the `0i` datasets for each type of jet sample. The last data points in the plot correspond to the fake rates including overflow events.

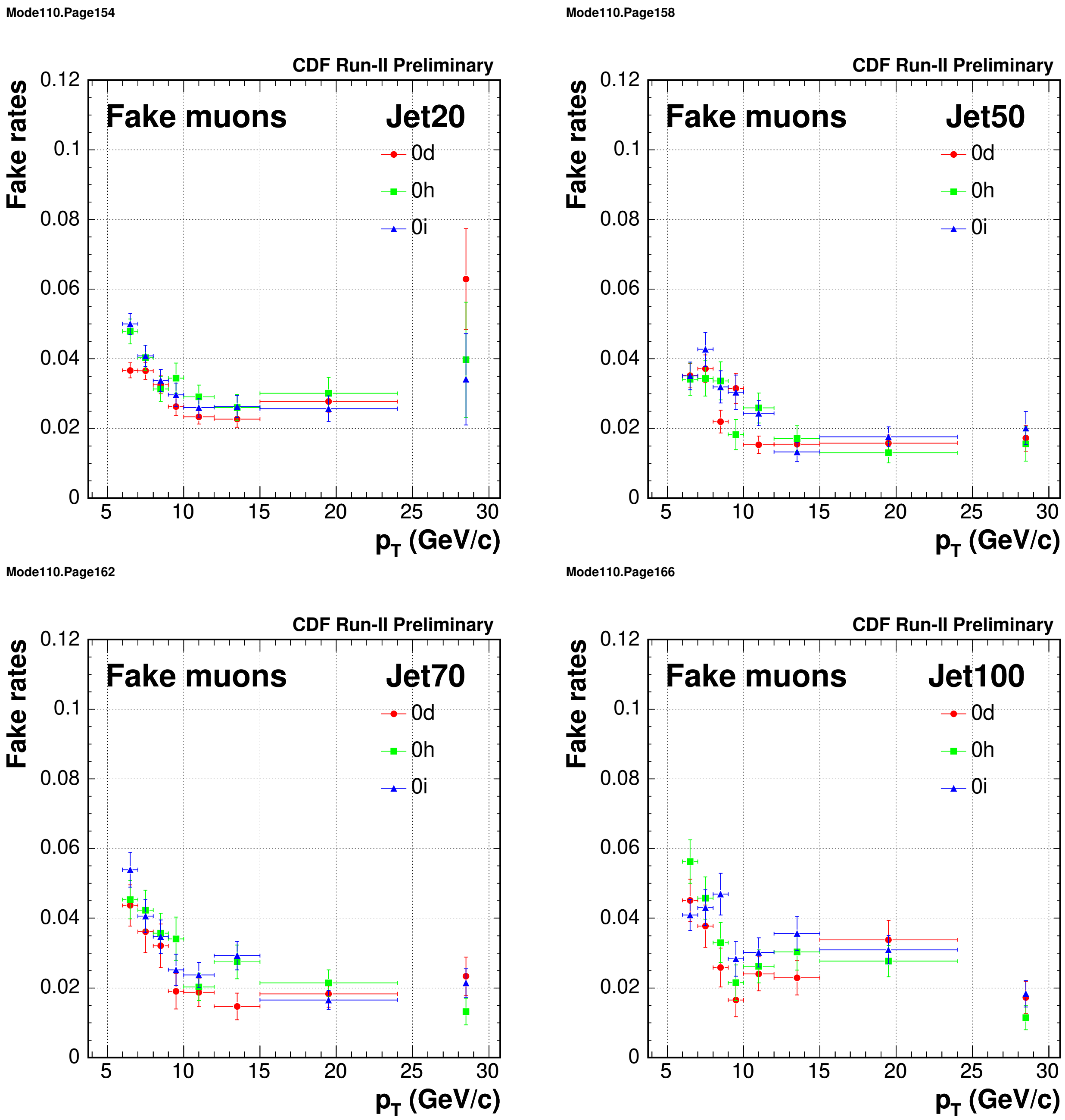


Figure 4: Comparisons of fake-muon rates among the `0d` through the `0j` datasets for each type of jet sample. The last data points in the plot corresponds to the fake rates including overflow events.

Mode110.Page401

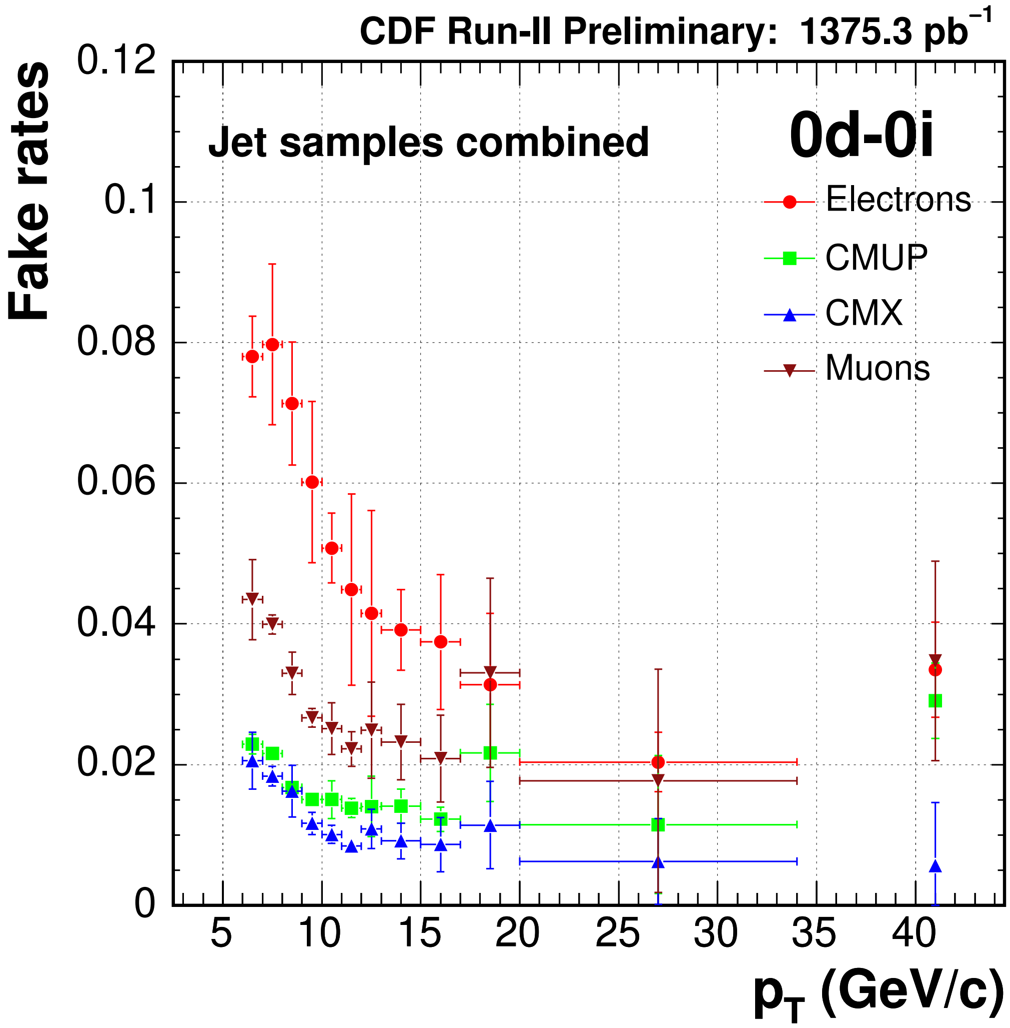


Figure 5: Fake rates as a function of track p_T obtained by combining all the datasets and jet samples. The uncertainties shown in the plots take into account of the RMS spread over jet samples as systematic uncertainties. The last data points in the plot correspond to the fake rates including overflow events.

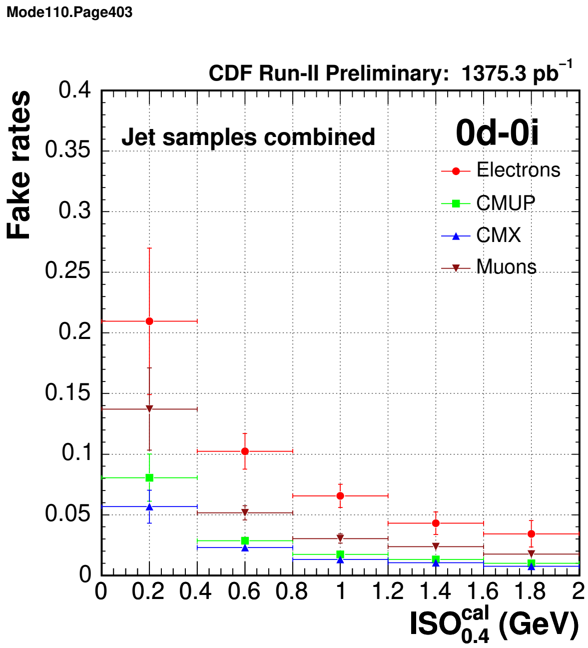


Figure 6: Fake rates as a function of isolation obtained by combining all the datasets and jet samples. The uncertainties shown in the plots take into account of the RMS spread over jet samples as systematic uncertainties.

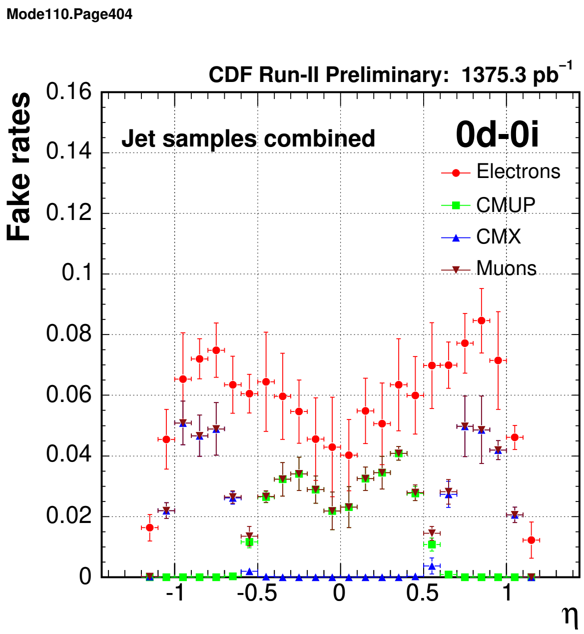


Figure 7: Fake rates as a function of track η obtained by combining all the datasets and jet samples. The uncertainties shown in the plots take into account of the RMS spread over jet samples as systematic uncertainties.

Mode110.Page511

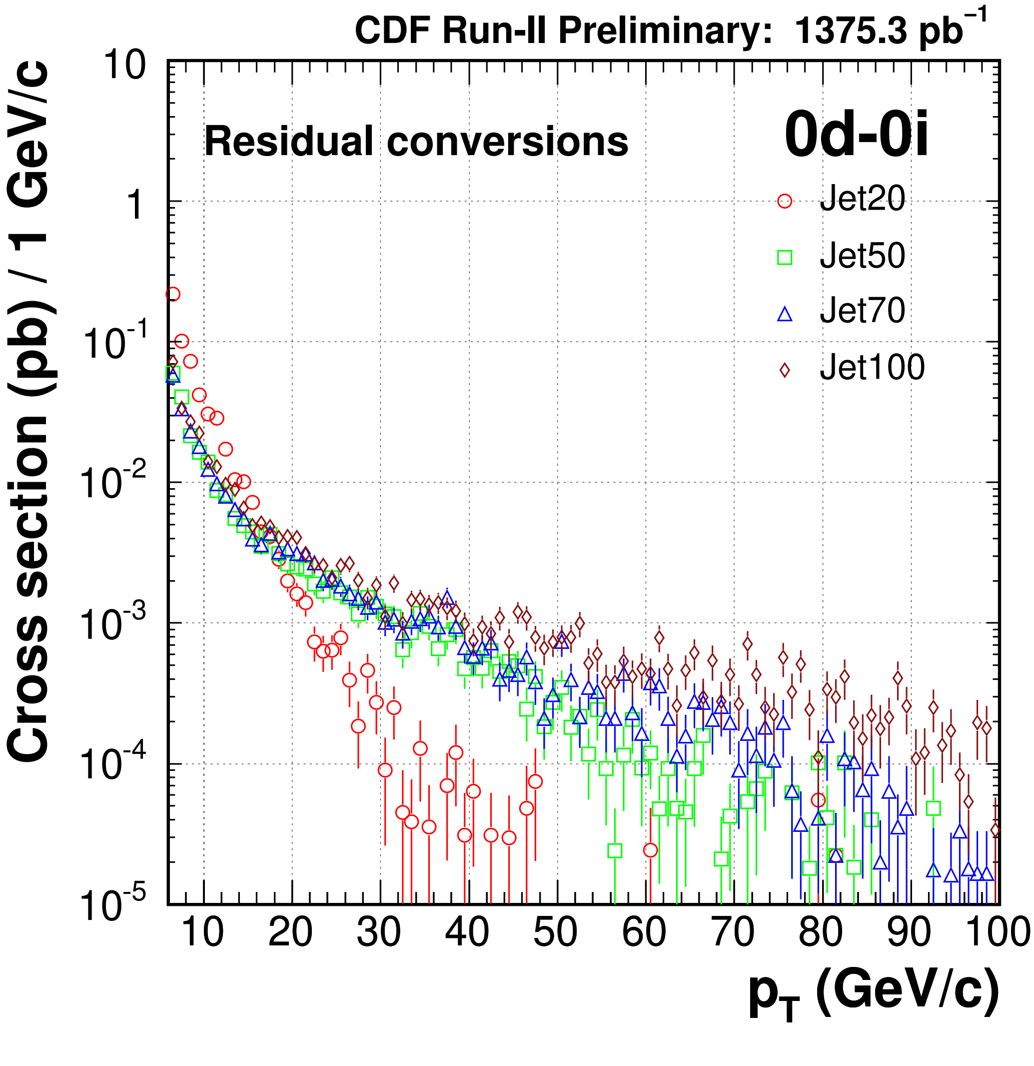


Figure 8: Track p_T distributions of residual photon-conversions in the jet samples combined over the datasets.

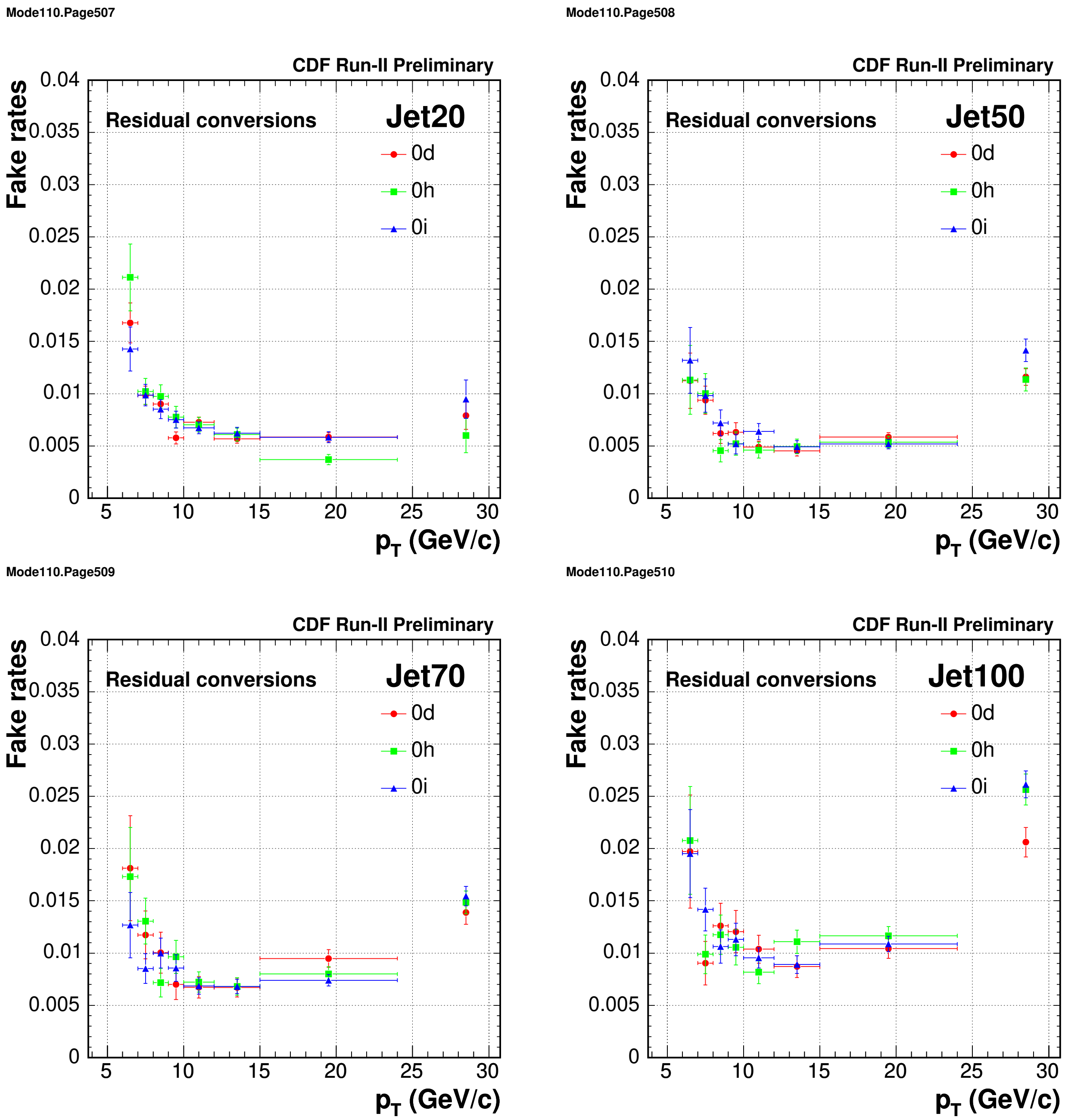


Figure 9: Comparisons of the residual conversion rates relative to the isolated hadronic tracks among the `0d` through the `0i` datasets for each type of jet sample. The last data points in the plot correspond to the fake rates including overflow events.

Mode110.Page612

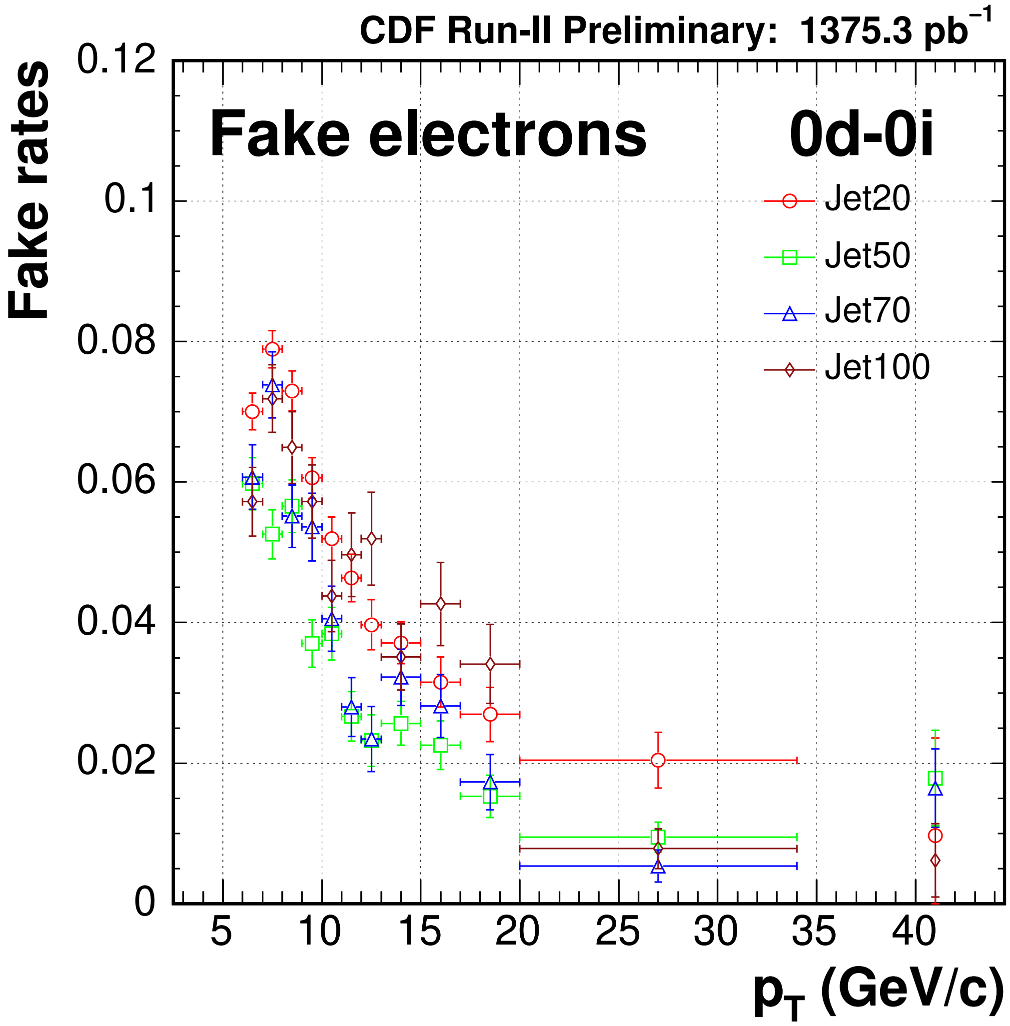


Figure 10: Fake-electron rates as a function of track p_T after subtracting residual photon-conversions for each of the jet sample combined over all the datasets. The last data points in the plot correspond to the fake rates including overflow events.

Mode110.Page801

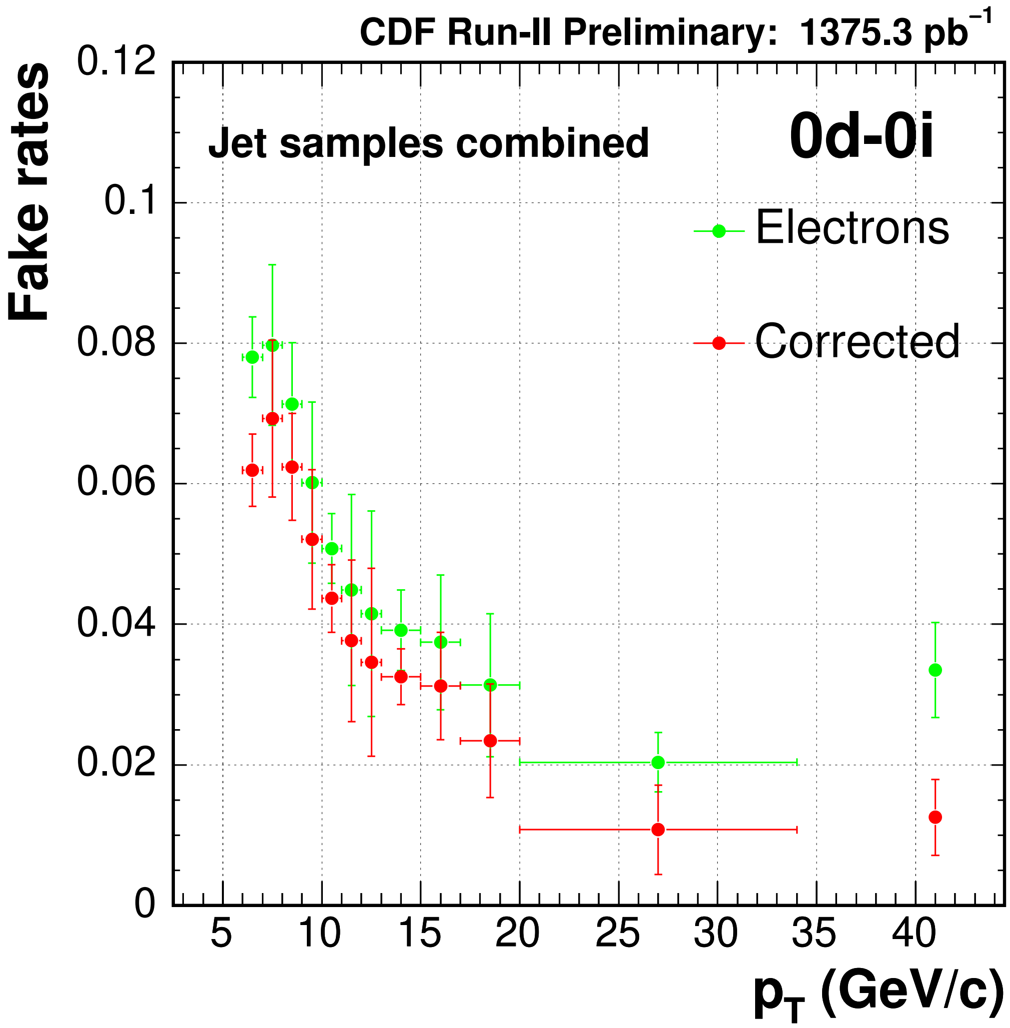


Figure 11: Fake-electron rates as a function of track p_T before and after subtracting residual photon-conversions, with combining all the datasets and jet samples. The uncertainties shown in the plots take into account of the RMS spread over jet samples as systematic uncertainties. The rates after correction also include a systematic uncertainty from the estimation of the residual photon-conversions. The last data points in the plot correspond to the fake rates including overflow events.

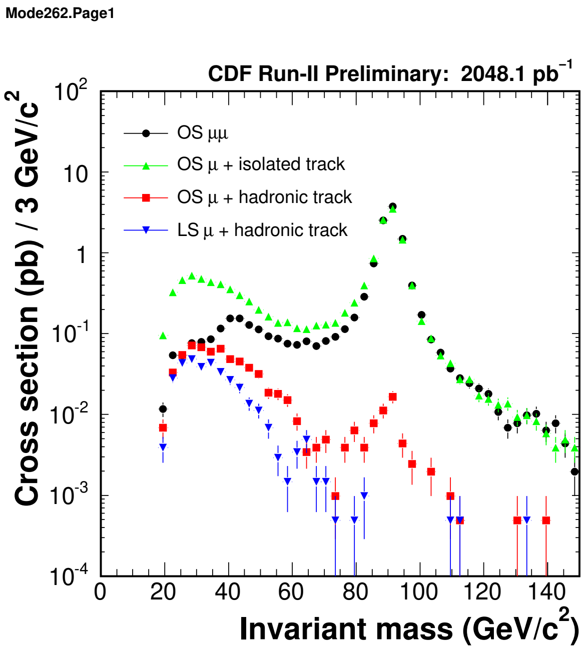


Figure 12: Comparison of the rates of μ + various objects from inclusive muon datasets.

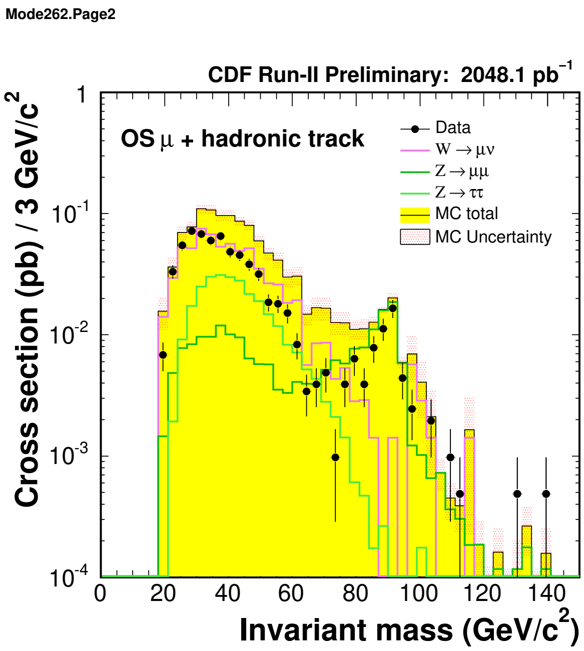


Figure 13: Comparison of the rates of μ + hadronic tracks between the data and the Monte Carlo predictions, where the hatch of the light red color represents statistical uncertainties of the Monte Carlo data.

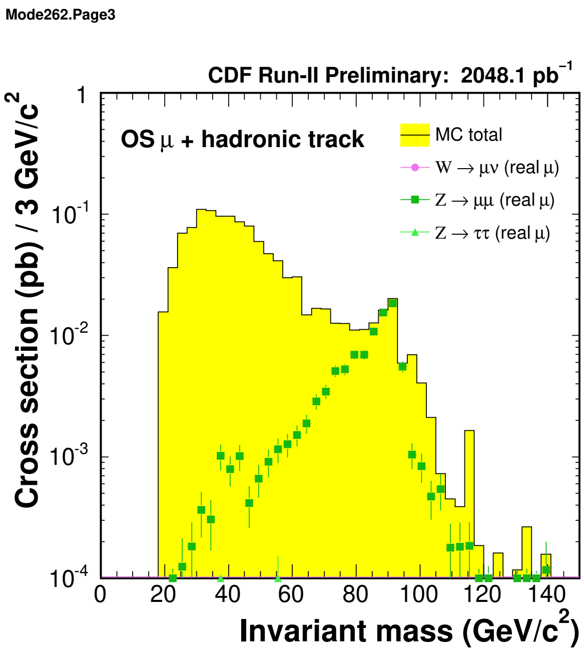


Figure 14: Real-muon components of hadronic tracks in the Monte Carlo events extracted by looking at `OBSP` information.

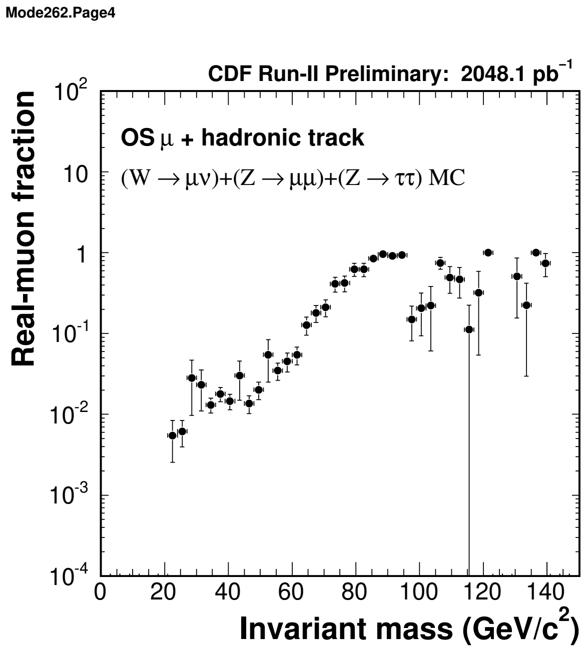


Figure 15: Fraction of real-muons of hadronic tracks for the Monte Carlo events.

Mode262.Page5

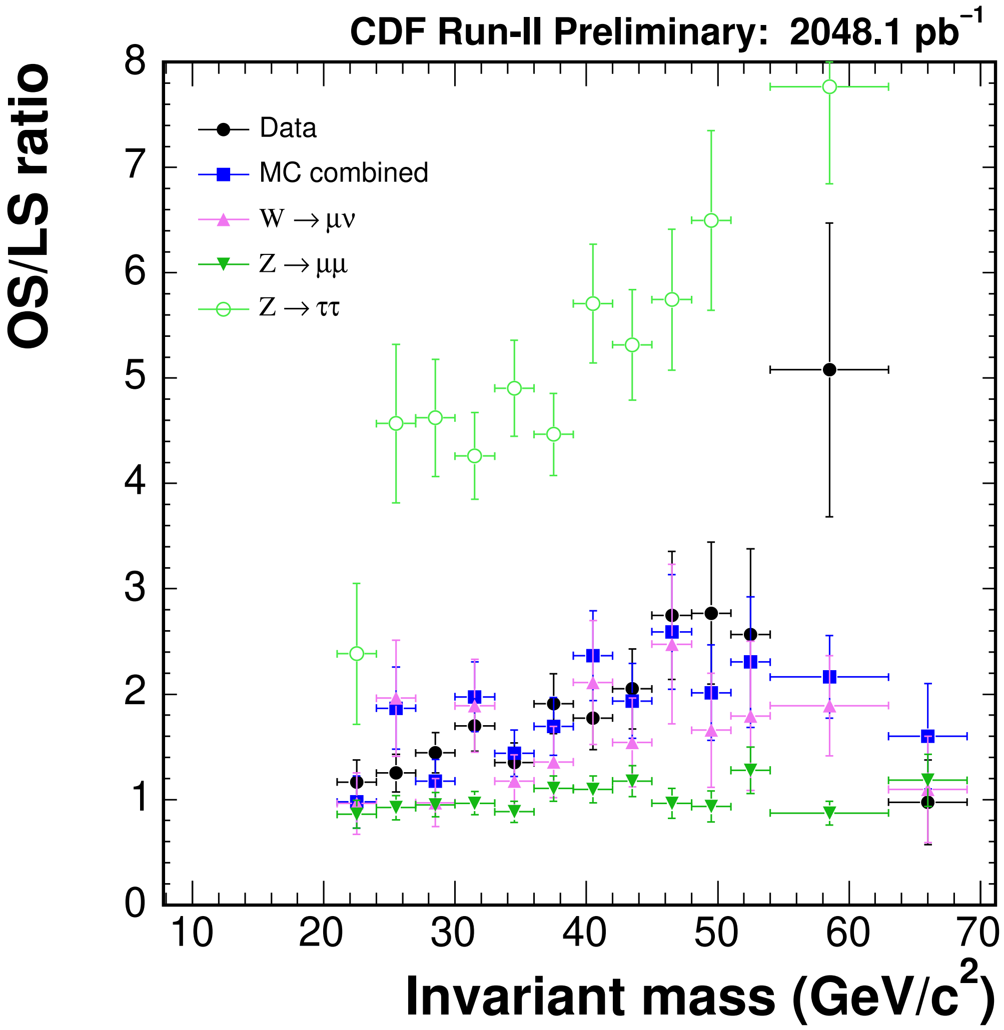


Figure 16: Opposite-sign and like-sign ratios of μ + hadronic track events for the data after subtracting real-muon fractions predicted by Monte Carlo events. The predictions of Monte Carlo events are also shown, which are obtained by explicitly removing real muons using the `OBSP` information.

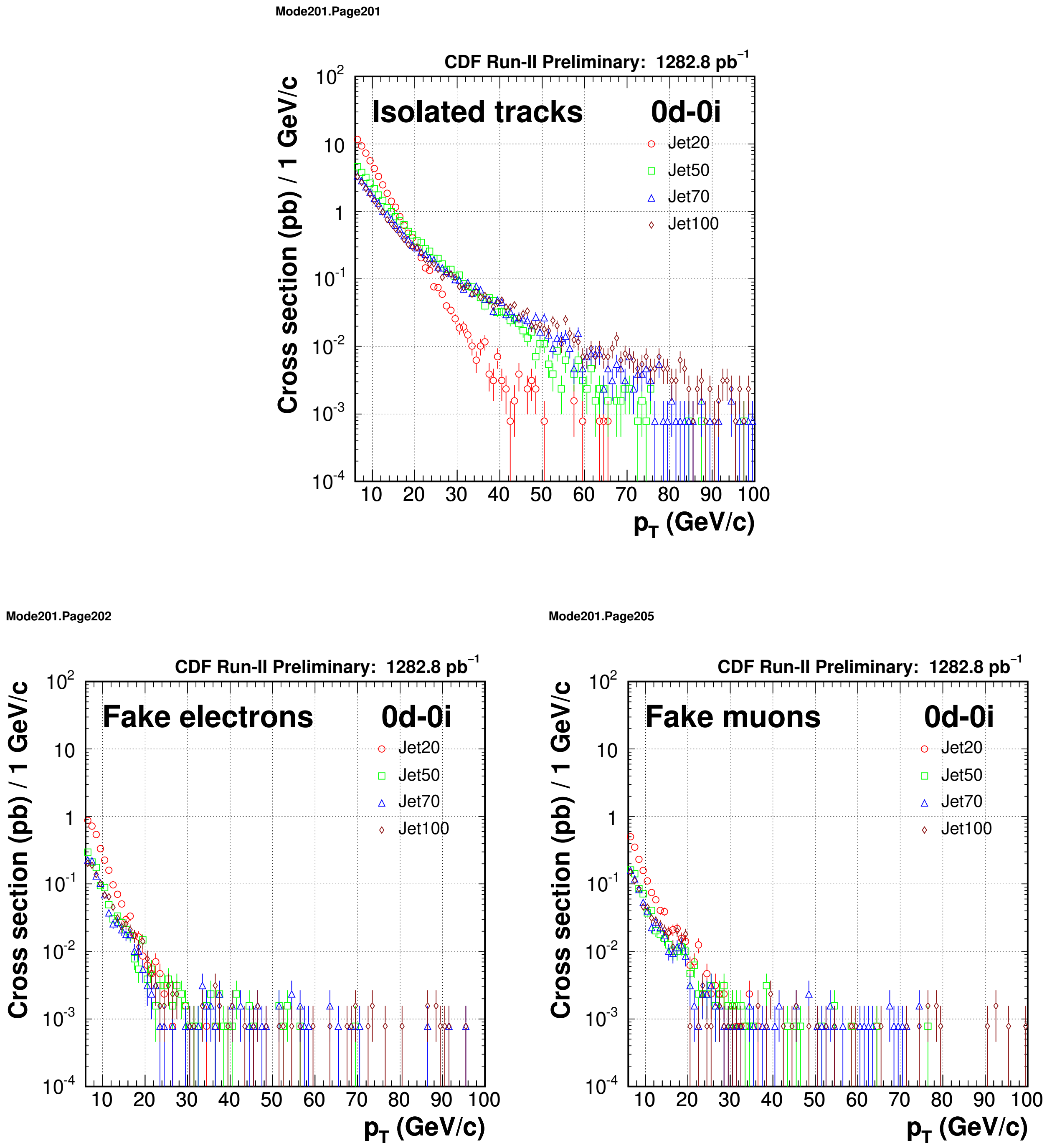


Figure 17: Track p_T distributions of isolated hadronic tracks (top), fake electrons (lower left), and fake muons (lower right) in the jet samples combined over the datasets. The Si hit cut is imposed as a track-quality cut.

Mode201.Page401

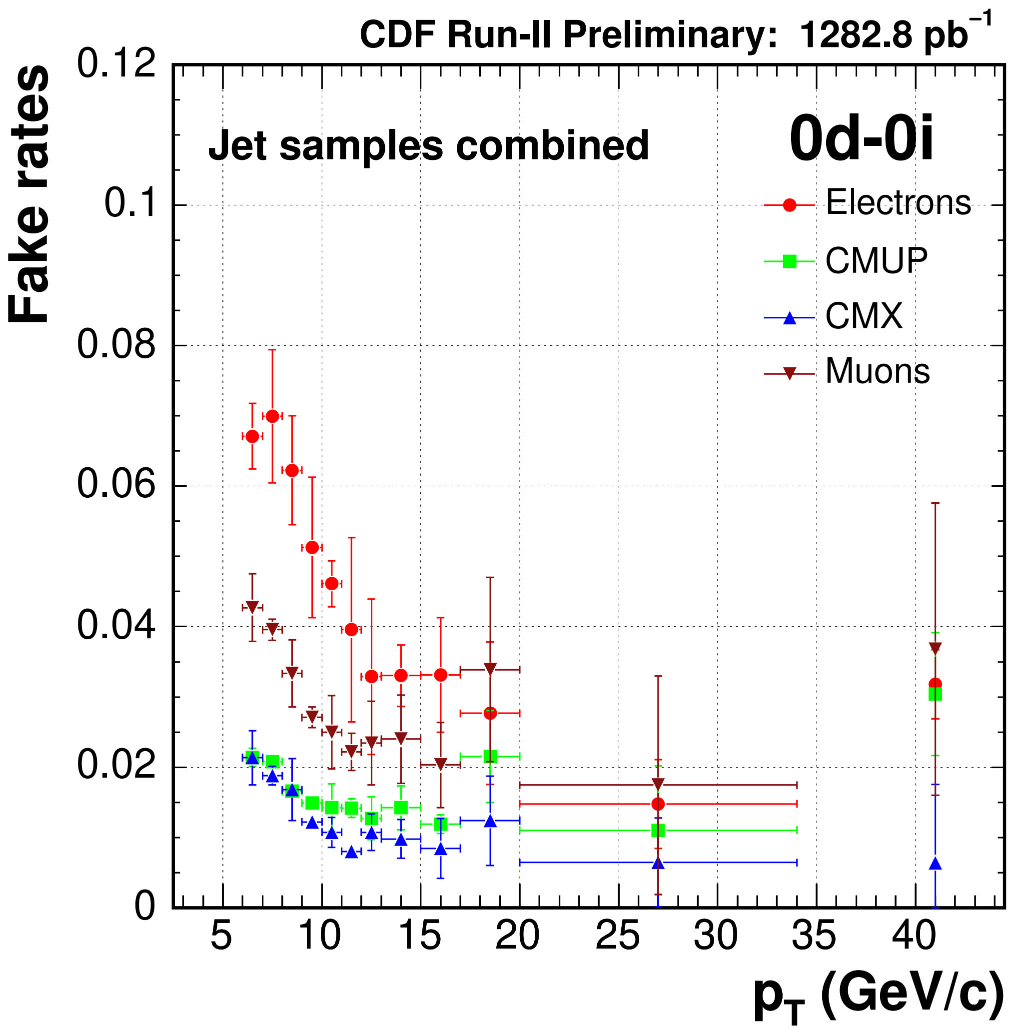


Figure 18: Fake rates as a function of track p_T obtained by combining all the datasets and jet samples. The Si hit cut is imposed as a track-quality cut. The uncertainties shown in the plots take into account of the RMS spread over jet samples as systematic uncertainties. The last data points in the plot correspond to the fake rates including overflow events.

Mode201.Page511

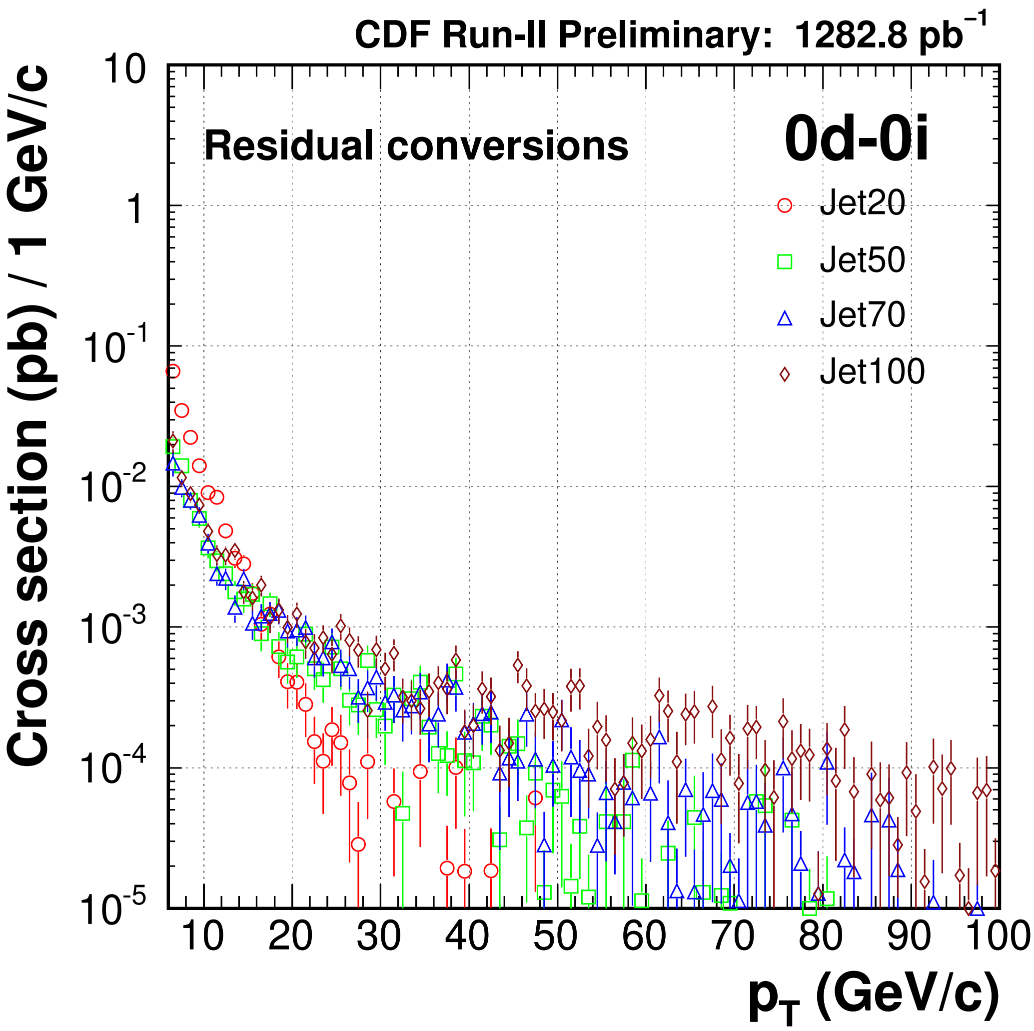


Figure 19: Track p_T distributions of residual photon-conversions in the jet samples combined over the datasets. The Si hit cut is imposed as a track-quality cut.

Mode201.Page801

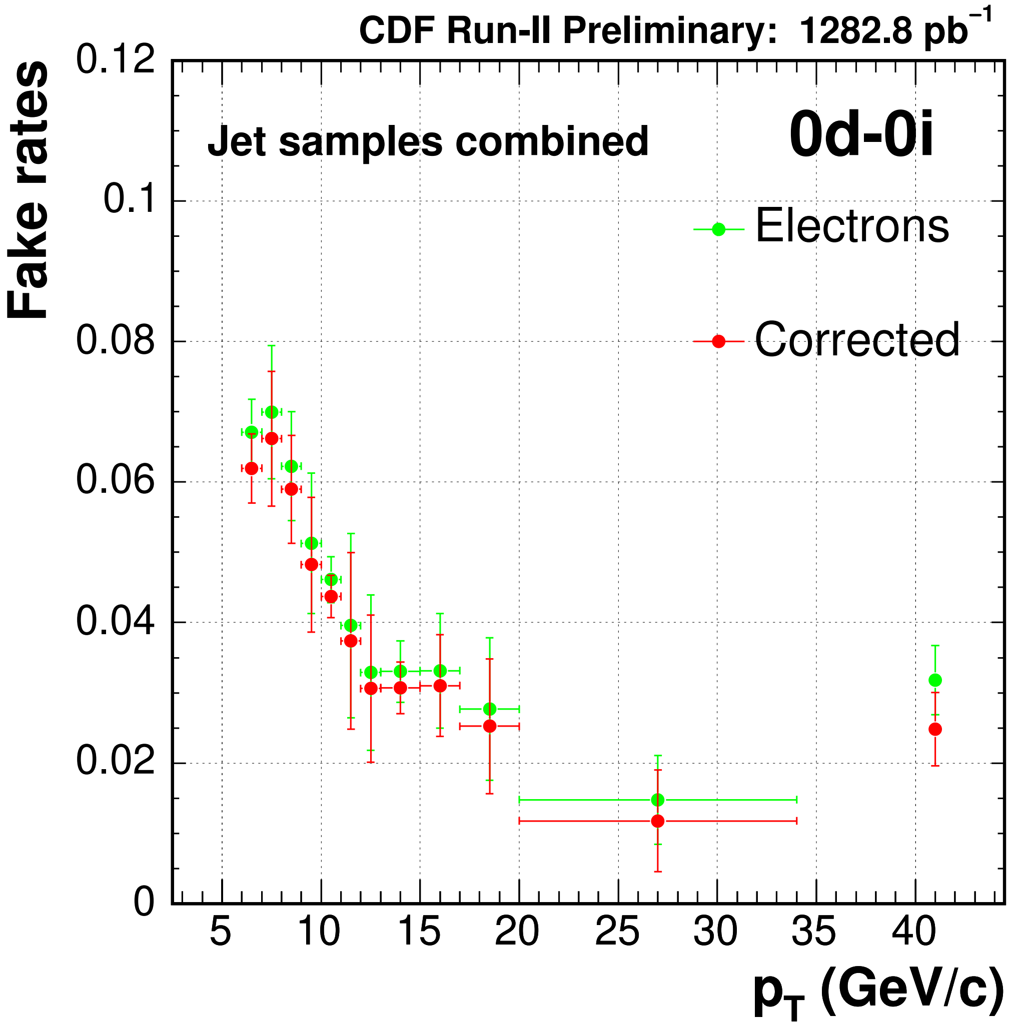


Figure 20: Fake-electron rates as a function of track p_T before and after subtracting residual photon-conversions, with combining all the datasets and jet samples. The Si hit cut is imposed as a track-quality cut. The uncertainties shown in the plots take into account of the RMS spread over jet samples as systematic uncertainties. The rates after correction also include a systematic uncertainty from the estimation of the residual photon-conversions. The last data points in the plot correspond to the fake rates including overflow events.

References

[1] Hirokazu Kobayashi, Kazuhiro Yamamoto, and Yoshihiro Seiya, "Background Estimation for High-p_T Like-Sign Dilepton Analysis", CDF7261 (Version 1.0).

[2] Toru Okusawa, Yoshihiro Seiya, Takayuki Wakisaka, and Kazuhiro Yamamoto, "Study of Conversion Detection Efficiency for the High-p_T Like-Sign Dilepton Analysis to Search for the Wh Production", CDF8078 (Version 2.0).

[3] Toru Okusawa, Yoshihiro Seiya, Takayuki Wakisaka, and Kazuhiro Yamamoto, "Baseline Selection of High-p_T Dilepton Events to Search for the Wh Production", CDF7260 (Version 2.0).