

A method to estimate the tau fake rate: from dijet to W +jets

Stephane Tourneur, Aurore Savoy-Navarro

LPNHE - University of Paris 6, France

Abstract

We propose a method to estimate the probability for a jet passing some denominator cuts to be falsely identified as a semihadronically decaying tau. A fake rate is derived progressively, starting from simple unbiased QCD dijet events, where precisions of the order of 6% are achieved, to event domains dominated by QCD events with high jet multiplicity, ending with regions dominated by $W \rightarrow l\nu + \text{jets}$.

Contents

1	Introduction	2
2	Definition of the jet to tau fake rate	2
3	Estimate of the jet to tau fake rate in the dijet events	3
4	Estimate of the jet to tau fake rate in the multijet events	5
4.1	The jet to tau fake rate matrix $F_{kR}(\text{jet } E_T, \text{Sum } E_T)$	5
4.2	Corrections depending on the number of jets and uncertainties on the estimate of the jet to tau fake rate in QCD events	8
5	The jet to tau fake rate related distributions	9
6	The jet to tau fake rate for $W \rightarrow l\nu + \text{jets}$ events	12
7	Conclusion	13

1 Introduction

The goal was to estimate the number of jets that are misidentified as tau leptons (labelled as **jets faking taus**), especially those from the dominant background of the $t\bar{t}$ tau dilepton analysis, i.e. the process: $W \rightarrow l\nu + \geq 3 \text{ jets}$. The Monte Carlo (Pythia and Herwig) overestimates the jet to tau fake rate by factors less than 2 [1,2]. In the absence of a tuned MC tool able to reproduce the behaviour of the jets faking taus in different phase space regions of interest, the attempt was made to estimate it from the data only.

Indeed the inherent difficulty in this study, is that there is not enough data available that describe the jets in the needed signal phase space region. The signature of this signal is characterized by one lepton (electron or muon), a missing transverse energy (\cancel{E}_T) larger than 20 GeV, a high-activity in the detector summarized by ($Ht > 205 \text{ GeV}$) and at least three jets (of which one is identified as a lepton tau decaying hadronically).

Our method consists in estimating the jet to tau fake rate due to the jets produced in the process $W \rightarrow l\nu + \geq 3 \text{ jets}$ from the jet to tau fake rate computed on several sets of QCD dijet events. This is achieved following the three steps:

1. Estimate of a tau fake rate in the *unbiased* QCD dijet events.
2. Extrapolation to the QCD events with higher jet multiplicities and transverse energy, using the Top Multijet (SumEt) sample.
3. Verification that the obtained tau fake rate also applies to the jets produced in the process $W + \text{jets}$ and thus validates the method.

2 Definition of the jet to tau fake rate

The jet to tau fake rate is defined as the average probability for a jet which passes the set of denominator cuts, which are enumerated below, to pass the remaining numerator cuts (also enumerated here below) and thus to be misidentified as a tau lepton.

The definition of the denominator is based on a high-quality selection, namely: The muon and the electron vetos are applied, the jet is required to be in the central region, the tau track must be of good quality and compatible with the interaction point. Furthermore it is based on the TauFinder algorithm in order to be able to use the tau variables at the output of the jet to tau fake rate. Otherwise, we would end up with some jets which have a non zero probability to be a tau, and thus cannot be rejected although they do not have a TauFinder object associated to them.

1. The denominator is defined with the following series of conditions:
 - The TauFinder algorithm requests:
 - A seed tower of $> 6 \text{ GeV}$ transverse energy
 - A seed track with transverse momentum p_T of $> 4.5 \text{ GeV}/c$

- \leq six neighbouring towers with > 1 GeV each
 - A cluster with $|\eta| < 1.1$
 - The fiducial requirements are: $9 \text{ cm} < \text{seed track } |z_{CES}| < 216 \text{ cm}$
 - The transverse momentum of the tracks and neutral pions contained in the tau jet must be less than $15 \text{ GeV}/c$
 - The z-distance to the vertex of the tau lepton, $\tau|z_0|$, must be less than 60 cm .
 - The impact parameter of the tau lepton, $\tau|d_0|$, must be less than 0.2 cm .
 - The electron veto defined by: $\frac{E_{had}}{\sum P} > 0.15$ must be applied.
 - The muon veto, defined by cluster E_T / seed track p_T , must be less than 0.5 .
 - The seed track quality defined as at least 3 stereo and axial superlayers with ≥ 5 hits must be applied.
2. The numerator is defined with the following series of requirements
- The transverse mass of the tracks plus the neutral pions must be less than $1.8 \text{ GeV}/c^2$
 - The relative calorimeter isolation computed in a cone in ΔR of 0.4 must be less than 0.1
 - The track isolation, defined as the number of tracks with a transverse momentum of at least $1 \text{ GeV}/c$ in a conical region between 10 and 30 degrees must be equal to zero.
 - The $\text{Pi}0$ isolation defined as the number of $\text{Pi}0$ in a conical region between 10 and 30 degrees must be equal to zero

3 Estimate of the jet to tau fake rate in the dijet events

Several triggers selecting so-called QCD events, i.e. based on events with jets are at disposal. The samples used for estimating the jet to tau fake rate in the dijet data are:

- The trigger **ST05** selects events with a single calorimeter tower collecting more than 5 GeV .
- The trigger **Jet20** selects events with at least one jet defined in a cone in ΔR of 0.7 , with more than 20 GeV transverse energy and with a single tower of more than 5 GeV transverse energy at level 1.

- The trigger **Jet50** selects events with at least one jet defined in a cone in ΔR of 0.7, with more than 50 GeV transverse energy and with a single tower of more than 5 GeV transverse energy at level 1.
- The trigger **Jet70** selects events with at least one jet defined in a cone in ΔR of 0.7, with more than 70 GeV transverse energy and with a single tower of more than 10 GeV transverse energy at level 1.

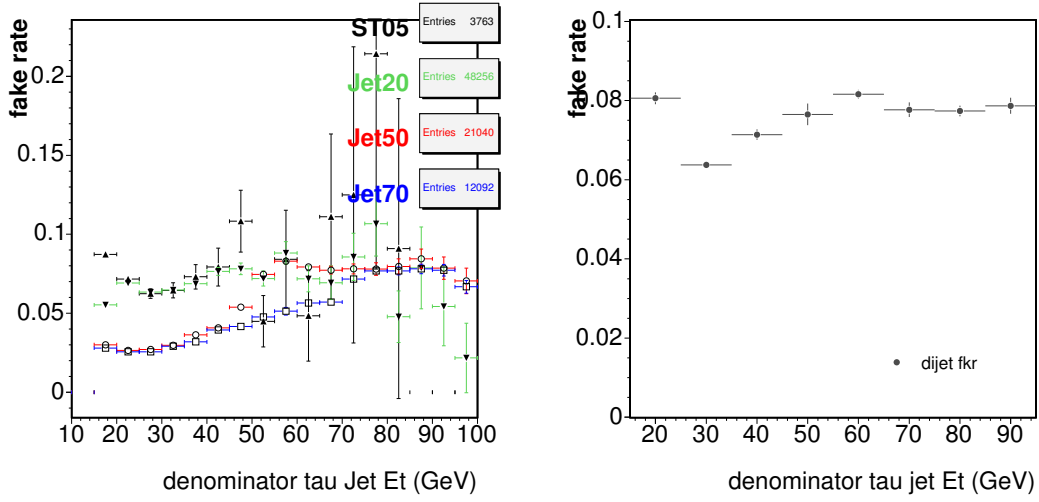
The data samples are dominated by events with two transversely back to back jets with transverse energies close to each other, i.e. well balanced in energy.

It should be noted that the triggers are based on jets with a cone in $\Delta R = 0.7$, whereas the jets used in this analysis are selected with the standard ClusterModule 0.4 cone algorithm: A trigger jet energy is thus always greater than the matching jet energy at the analysis level. Everytime a jet transverse energy (jet E_T) will be mentioned, it refers to a jet with a cone in ΔR of 0.4 and an uncorrected jet E_T , re-clustered at the highest p_T vertex in the event.

Using the ST05 sample, it was checked that the amount of jets with $E_T > 75$ GeV and with no tower with an energy exceeding 10 GeV is very small. Furthermore, the efficiency for a jet passing the tau identification denominator cuts, and with E_T above 25 GeV, to be selected by the **Jet20** trigger, is on the order of 100%. The trigger efficiency plateau is indeed reached around this energy of 25 GeV. Likewise, all tau denominator jets with E_T greater than 55 and 75 GeV will be selected by the triggers **Jet50** and **Jet70** respectively. The denominator selection includes a cut of 6 GeV on the tau seed tower E_T ; it is assumed that the **ST05** trigger is almost 100% efficient in selecting any events containing such a denominator jet. If 100% of the denominator jets are selected by these triggers it means that all these jets are *unbiased* by this trigger selection. The sample of denominator jets, in a given E_T range, extracted from any of the four triggers is thus well representative of the set of denominator jets of the same E_T range produced at Run II, or at least this is a good approximation. This gives us the means to calculate the jet to tau fake rate for unbiased QCD data as a function of the jet transverse energy.

The figure 1(a) shows the jet to tau fake rate as a function of the jet E_T . As expected, the jet to tau fake rates from the four samples agree remarkably well in the unbiased E_T regions, above the trigger thresholds. However, the jet to tau fake rate calculated for jets with E_T below the trigger threshold, shows some discrepancy as expected. Indeed the typical event provided by the jet triggers is a transversely back-to-back dijet event with the two jets of approximately the same E_T and slightly above the trigger threshold. If a jet with E_T below the trigger threshold is selected, chances are high that this jet is coming from the gluon splitting produced by one of the original jets. The splitting leads to two lower energy jets too close to pass the numerator cuts, as these cuts are mainly based on isolation requirements. A lower jet to tau fake rate is thus expected for these *biased* jets and this is observed as shown in fig.1.

After having got rid of the biased jets below the energy thresholds, a much better estimate of the jet to tau fake rate is obtained by adding the contributions of the high



(a) 4 QCD triggers tau fake rates

(b) Estimated tau fake rate in unbiased dijet events

Figure 1: Jet to tau fake rate as a function of the uncorrected jet E_T . A jet is matched to every denominator tau to get the jet E_T .

energy portions of the four QCD triggers (see fig.1(b)). This is the *jettotaufakerate* for generic dijet data. The jet to tau fake rate varies between 6 and 8%, depending on jet E_T .

This fake rate is not directly used in the top in tau analysis but is perfectly suited to inclusive $Z \rightarrow \tau\tau$ or $H \rightarrow \tau\tau$ background estimates. It serves as a basis for estimating the jet to tau fake rates in more and more complicated QCD scenarios.

4 Estimate of the jet to tau fake rate in the multijet events

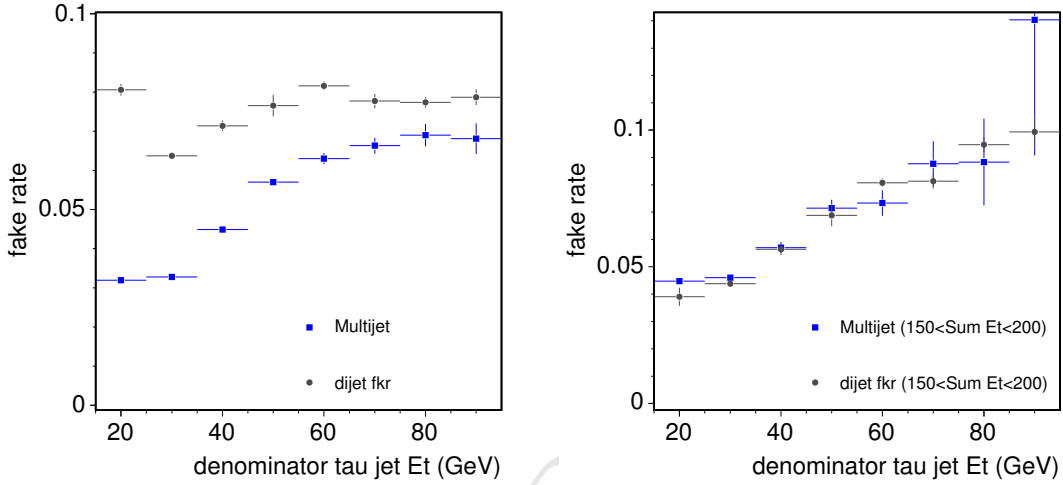
4.1 The jet to tau fake rate matrix $\text{FkR}(\text{jet } E_T, \text{Sum } E_T)$

Estimating the jet to tau fake rate for events with higher E_T activity and higher number of jets is much more difficult than in the case of inclusive dijet events.

The reference trigger here is the **Top Multijet** trigger which was developed and built for selecting the pair of top quarks events where the two produced W bosons decay into quarks. This gives typically six jets events and in any case it ensures a high activity and high jet multiplicity in the events. The trigger selection is made on events with 4 jets with E_T above 15 GeV and a sum of calorimeter transverse energies above 125 GeV. This is also called the **SumEt** trigger.

The data sample provided by the **SumEt** trigger allows to check the jet to tau fake

rate, obtained before for dijet events, in event samples with a higher jet multiplicity and higher activity. Fig. 2(a) shows a comparison between the jet to tau fake rate computed in dijet samples with the jet to tau fake rate computed in the **SumEt** multijet sample. A discrepancy between the jet to tau fake rates in the two samples is observed.



(a) Discrepancy between the jet to tau fake rate from **SumEt** events and the one from the dijet events.

(b) Agreement between the jet to tau fake rate from **SumEt** events and from the dijet events if $150 \text{ GeV} < \text{SumEt} < 200 \text{ GeV}$.

Figure 2: Comparison between the jet to tau fake rates from dijet vs SumEt Multijet events.

Indeed, in order to reproduce the jet to tau fake rate in high jet multiplicity regions, it is required to take into account the higher value of Sum E_T in these events. A new parameter is thus added, namely the sum of the transverse energies of all the calorimeter towers in the event. It is computed at the highest p_T vertex of the event and corrected for the presence of standard central tight muons from the CMUP and CMX muon chambers (as defined in section 2.3.1.2). Fig. 2(b) shows as an example the nice agreement achieved between the jet to tau fake rate from dijet events recalculated from each trigger sample (ST05, Jet20, Jet50, Jet70) when restricting ourselves to events with Sum E_T between 150 and 200 GeV.

The Sum E_T parameter is thus introduced as a second parameter to compute the jet to tau fake rate so that it is also applicable to high jet multiplicity QCD events. As a result, a 2-dimensional matrix in (jet E_T , Sum E_T) is obtained (see Fig.3). It includes 16×8 bins, with the jet E_T varying from 15 to 95 GeV and the Sum E_T varying from 0 to 400 GeV. The numbers quoted in this matrix are obtained from the four Jet E_T triggers (always restricted to unbiased energy regions, above trigger thresholds) and from the multijet trigger. It is labelled as the *jet to tau fake rate matrix*. Note that the matrix top left portion is left empty because any event necessarily has a sum E_T greater than the E_T of any jet found in the event. The variations in the rates readable

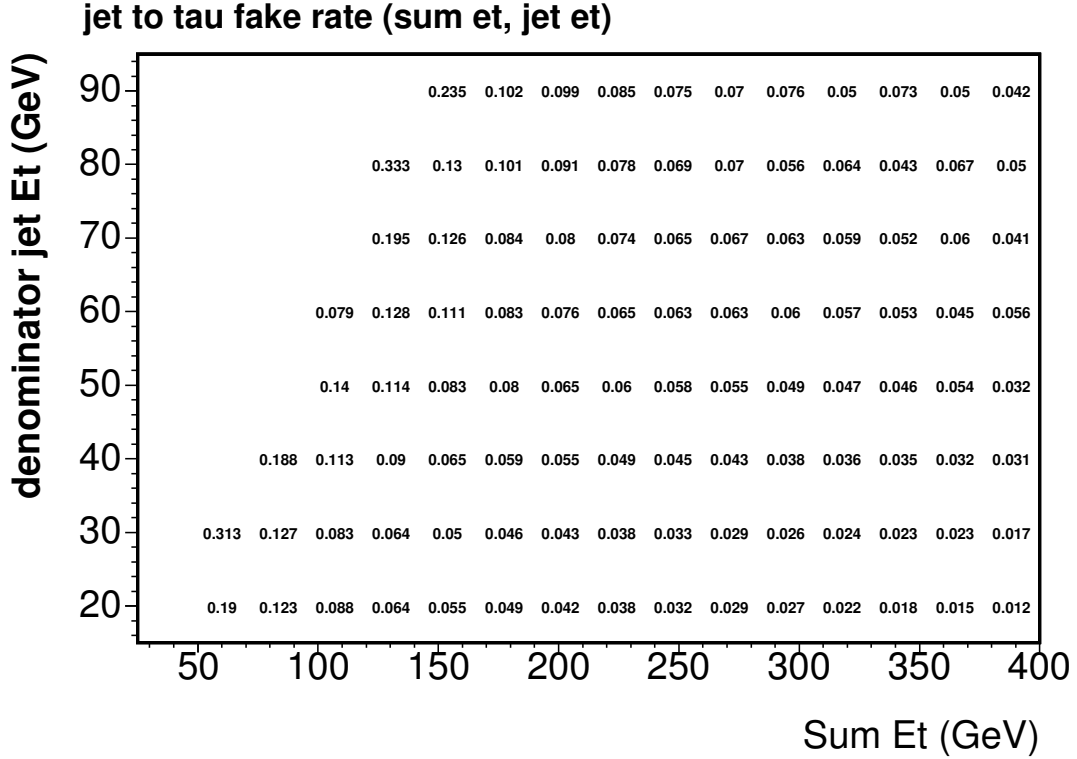


Figure 3: The jet to tau fake rate matrix used in the analysis.

in the matrix top left part are thus simply due to some high statistical fluctuations of small numbers of events.

This matrix makes good predictions for the number of jets faking taus in the high p_T electron trigger (table 1), where no selection except the trigger requirement and the tau lepton identification are applied. Any jet matching the electromagnetic cluster passing the electron trigger requirement is removed from the list of the tau candidates.

	<i>elec_25</i>	<i>elec_35</i>
pred	9436	4675
obs	8588	4310
obs/pred	0.91 ± 0.01	0.92 ± 0.02

Table 1: Predictions and observations of the number of jets faking tau leptons in the high p_T electron trigger. *elec_y* refers to the set of events in the electron trigger where the denominator jet $E_T > y$ GeV.

4.2 Corrections depending on the number of jets and uncertainties on the estimate of the jet to tau fake rate in QCD events

The ability of the jet to tau fake rate matrix to correctly predict the number of jets faking tau leptons in event samples with given jet multiplicities, is now tested. To do this, we take events from the five QCD triggers used to build the matrix, split them according to their event jet multiplicities, and compare the number of identified tau lepton candidates really observed with the number predicted from the jet to tau fake rate matrix. The jet to tau fake rate matrix is then to be considered successful if these two numbers are found equal. Let's remark at this point that this is true only because the number of true tau leptons is negligible compared with the number of jets faking tau leptons in the samples considered for the tests. The reason for this is that the probability for a jet to fake a hadronic tau decay is high (on the order of 1%)¹.

From the *Jet20* trigger sample, we make two samples of jets, namely the *Jet20_25* that collects the jets with E_T greater than 25 GeV and that pass the denominator tau cuts, and *Jet20_35* that includes jets with E_T greater than 35 GeV and that pass the denominator tau cuts. Likewise, we define the jet samples *Jet50_55* and *Jet70_75*, from the *Jet50* and *Jet70* samples respectively. Then, *elec_25* and *elec_35* are defined in the same way from the *elec* high p_T electron trigger sample.

The table 2 tests the jet to tau fake rate prediction accuracy inside the six samples ST05, *Jet20_25*, *Jet20_35*, *Jet50_55*, *Jet70_75* and SumEt. It shows three numbers, once for each sample: The observed number of jets identified as tau leptons, the predicted number of jets passing the tau identification selection given by the fake rate matrix, and the ratio between these last two numbers. The results are split into three categories following the event jet multiplicity. The table 3 performs the same tests for the jets in the *elec_25* and *elec_35* samples.

The analysis of these results shows that the number of jets faking tau leptons is still overestimated in events with high jet multiplicity and underestimated in those with low jet multiplicity. A new correction factor depending on the total number of jets in the event, labelled as $f(N \text{ jets})$ is thus applied in order to get more valid predictions for events with more than two jets. The $f(N \text{ jets})$ values must mirror the average of the ratios of observed over predicted numbers of tau leptons measured in the tests. The sample of jets from *Jet20_25* has nice characteristics because it has high statistics in the different jet multiplicities considered and, furthermore, it shows ratios compatible with the average of the ratios calculated in all the other samples. That is why we choose it as our reference sample to decide the values for the $f(N \text{ jets})$ correction factors.

As far as the determination of the systematic error associated with the fake rate goes, we compute it by taking the largest difference between all the ratios $\frac{\text{Number observed taus}}{\text{Number predicted taus}}$ measured in the eight samples and the ratio measured within *Jet20_25*.

One gets:

¹If we were considering other lepton fake rates (like electron), the contamination from true leptons should be taken into account carefully.

- $f(1 \text{ jet}) = 1.06 \pm 0.06$
- $f(2 \text{ jets}) = 0.98 \pm 0.10$
- $f(\geq 3 \text{ jets}) = 0.87 \pm 0.14$

	<i>ST05</i>	<i>Jet20_25</i>	<i>Jet20_35</i>
1 jet (= the tau)	2026/1814 = 1.12 ± 0.02	4967/4674 = 1.06 ± 0.01	613/601 = 1.02 ± 0.04
2 jets	1540/1748 = 0.88 ± 0.03	10035/10230 = 0.98 ± 0.01	2516/2506 = 1.00 ± 0.02
≥ 3 jets	196/239 = 0.82 ± 0.08	1946/2226 = 0.87 ± 0.02	774/874 = 0.89 ± 0.03
	<i>Jet50_55</i>	<i>Jet70_75</i>	<i>SumEt</i>
1 jet (= the tau)	578/537 = 1.08 ± 0.04	214/199 = 1.08 ± 0.08	42/51 = 0.82 ± 0.15
2 jets	4502/4556 = 0.99 ± 0.02	2521/2550 = 0.99 ± 0.02	1002/951 = 1.05 ± 0.03
≥ 3 jets	2832/2822 = 1.00 ± 0.02	2031/2111 = 0.96 ± 0.02	24535/24370 = 1.01 ± 0.01

Table 2: *Predictions and observation of the number of jet to tau fakes within the 5 triggers. Jetx_y means a fake rate applied in the Jetx trigger with only denominator jets with $E_T > y$ GeV*

	<i>elec_25</i>	<i>elec_35</i>
2 jets (=the tau+the trigger object)	6419/6915 = 0.93 ± 0.01	3132/3252 = 0.96 ± 0.02
≥ 3 jets	2169/2521 = 0.86 ± 0.02	1178/1423 = 0.83 ± 0.02

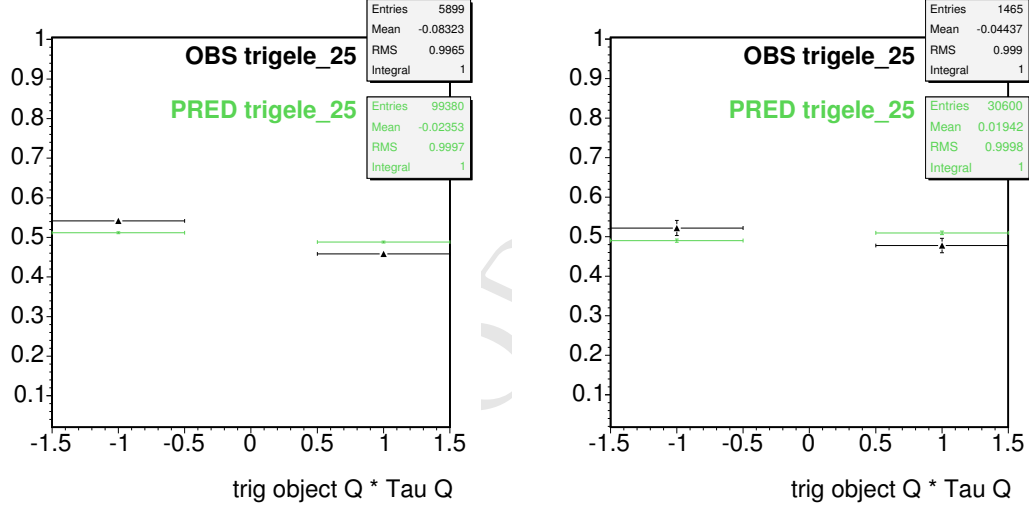
Table 3: *Observations/Predictions of number of jet to tau fakes in the high p_T electron trigger. elec_y refers to the set of events in the electron trigger in which the denominator jet $E_T > y$ GeV*

5 The jet to tau fake rate related distributions

Any estimated jet to lepton fake rate should correctly predict the number of misidentified leptons in any sample. It is however not supposed to be predictive with respect to the distributions of any variable correlated to the variables used for the lepton identification. The case of the tau track isolation variable can be used as an example: Each jet passing the tau identification verifies the track isolation cut and thus has no track with transverse momentum higher than 1 GeV/c, within the tau isolation cone; however, the jet to tau fake rate applied to the denominator jets will sometimes include tracks in the isolation cone and it cannot therefore reproduce the distribution of the number of tracks in the isolation cone. The only way to overcome this issue is to correct the predicted distributions. In the case of the 350 pb⁻¹ top dilepton analysis, the variables used for the signal event selection in addition to the number of jets are: The product of the tau lepton charge with the charge of the other lepton, the activity Ht in the event, defined as the sum of the transverse momentum of the tau, of the other lepton and of the other jets in the event plus the total transverse missing energy (\cancel{E}_T), and the leading jet transverse energy.

The fig.4 shows the predicted and observed distributions for the product of the leptons charges in the case of the sample selected with the high p_T electron trigger. Fig.5 shows the comparison of the measured and predicted H_t and leading jet E_T variables with the same sample of data.

The agreement is rather good. The small discrepancy observed in the lepton charge product must be however corrected. Indeed each time the jet to tau fake rate is applied to the prediction of the lepton charge product, the predicted number of events with opposite charge must be multiplied by a factor 1.07, taking care to keep the total number of predicted events unchanged as this must be a correction to a predicted distribution, not to the total number of events.



(a) Case with 1 jet in the event (=the tau jet)

(b) Case with 2 jets in the event

Figure 4: The predicted lepton charge product distributions obtained from the jet to tau fake rate compared with the observed ones in the elec.25 sample.

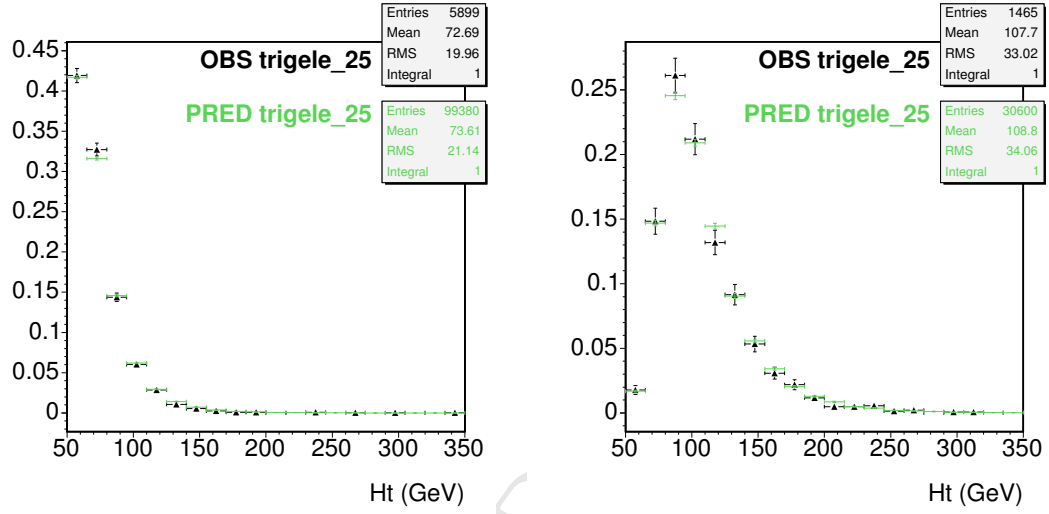
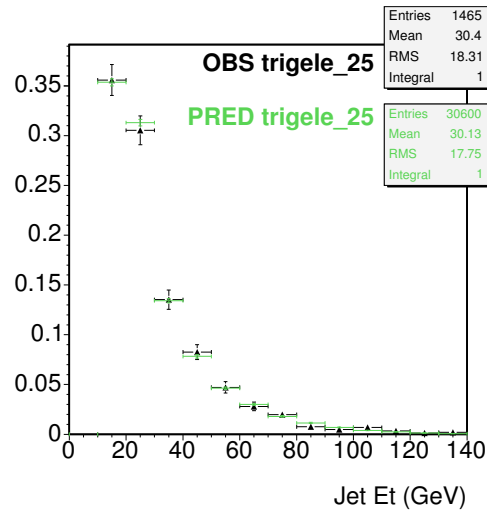
(a) $Ht\ 0j$ (b) $Ht\ 1j$ (c) leading jet E_T

Figure 5: distributions from fake rate

6 The jet to tau fake rate for $W \rightarrow l\nu + \text{jets}$ events

The jet to tau fake rate obtained is defined as $fkr(\text{jet } E_T, \text{sum } E_T) \times f(N \text{ jets})$.

As previously discussed it applies to QCD backgrounds with both low or high activity and a number of jets up to 3. Can one apply it to the case of jets from $W \rightarrow l\nu + \text{jets}$ events? The presence of the W decaying into a lepton should not modify the shape of the jets. The fact that these events have a rather high activity and a large number of jets is taken into account by the Sum E_T parameter and the N jets factor.

However, is the proportion of quark to gluon jets the same in W+jets than in QCD events? This is an important question because, in QCD, gluons have higher coupling strengths than quarks to emit extra gluons and, therefore, gluon jets tend to be broader than the quark jets [7] [8]. Because of this, one could expect a lower value for the jet to tau fake rate in the case of gluon jets than in the case of quark jets.

How to verify this possible feature of W+jets events in data? To do so, it is needed to select regions dominated by W+jets events. This is achieved by requiring that the events fulfill the following criteria: $\cancel{E}_T > 20 \text{ GeV}$, one central electron with $E_T > 20 \text{ GeV}$, and extra jets. But the additional requirement of one well-identified tau lepton results in a dramatic drop in the statistics and therefore it is not anymore possible to perform this test on the remaining sample (see first line of table 4). The requirements on \cancel{E}_T and on the electron can hardly be loosened if one wants to stick in regions dominated by W+jets events.

We made an attempt to compute the ratios of jet to tau fake rates obtained in Pythia dijet events and Pythia inclusive W events [2] and to use these ratios to correct the jet to tau fake rate computed on QCD data. But this did not work properly.

Using data, the only way is to loosen the tau identification. To achieve this, seven selections based on the same tau denominators are defined. Firstly, in order to keep the tests independant from the original jet to tau fake rate numbers and to stay far from the top signal region in the 3 jets bins, they all require the denominator tau not to pass successfully the whole tau identification. The simple criteria applied are gathered in the first column of Table 4.

For each of these seven selections, the exact same method as the one previously described is applied to define a new jet to tau fake rate matrix and the new f(N jets factors) from the ST05, Jet20, Jet50, Jet70 and SumEt samples.

The results of the seven tests are shown in Table 4. The numbers quoted in the Table show the comparison between the observed number of identified tau leptons with the predicted number of jets faking taus obtained from our jet to tau fake rate. Backgrounds such as $Z \rightarrow \tau\tau$, WW, electrons faking taus and $t\bar{t}$ have been subtracted from the number of observed events because these events contain true taus or sources of fake taus other than jets. Here below are listed the conclusions:

- In all jet bins, the results in the ratio of predicted to observed taus are compatible with 1; it means that the jet to tau fake rate defined for the QCD samples can be applied to the regions dominated by W+jets events.

$\tau^i ID$	$e+\cancel{E}_T+1 \text{ jet}$	$e+\cancel{E}_T+2 \text{ jets}$	$e+\cancel{E}_T+\geq 3 \text{ jets}$
1 complete τ ID	$61/72 = 0.85 \pm 0.12$	$28/24 = 1.17 \pm 0.20$	-
1 denom. τ w $M < 1.8 \text{ GeV}$	$389/392 = 0.99 \pm 0.06$	$137/144 = 0.95 \pm 0.08$	$57/48 = 1.19 \pm 0.14$
1 denom. τ w $\text{cal.iso} < 0.1$	$431/420 = 1.03 \pm 0.05$	$189/195 = 0.97 \pm 0.07$	$69/66 = 1.05 \pm 0.12$
1 denom. τ w $N \pi^0 \text{ iso} = 0$	$673/642 = 1.05 \pm 0.04$	$266/269 = 0.99 \pm 0.06$	$95/94 = 1.01 \pm 1.03$
1 denom. τ w $N \text{ trk iso} = 0$	$191/217 = 0.88 \pm 0.07$	$80/81 = 0.99 \pm 0.11$	$31/26 = 1.19 \pm 0.20$
1 τ ID w $\text{cal.iso} > 0.1$	$68/64 = 1.06 \pm 0.12$	$17/19 = 0.89 \pm 0.23$	$8.1/5.1 = 1.59 \pm 0.44$
1 τ ID w $M > 1.8 \text{ GeV}$	$143/127 = 1.13 \pm 0.09$	$55/53 = 1.04 \pm 0.14$	$18/17 = 1.06 \pm 0.24$
1 τ ID w $N \text{ trk iso} > 0$	$44/51 = 0.86 \pm 0.14$	$27/24 = 1.12 \pm 0.20$	$11.1/7.1 = 1.56 \pm 0.38$

Table 4: Predictions of number of jet to tau fakes with the standard Tau ID (upper line) and 7 selections of denominator taus failing to pass the tau selection

- Half the difference observed between the largest and the smallest ratio is taken as a systematic error for each jet multiplicity:
 - 1 jet: $(1.13-0.85)/2 = 0.14$
 - 2 jets: $(1.17-0.89)/2 = 0.16$
 - ≥ 3 jets: $(1.59-1.01) = 0.29$

These systematic errors are then added quadratically with the systematic errors of the jet to tau fake rate applicable to the non W+jets QCD backgrounds.

Finally, the jet to tau fake rate obtained for the W+jets events is found to be the same as the one for the other QCD events, apart from a larger systematic error. This jet to tau fake rate for W+jets events, expressed as a function $\text{FkR}(\text{jet } E_T, \text{sum } E_T) \times f(N \text{ jets})$, with FkR defined by the matrix of Fig.3, and with updated values for the function $f(N \text{ jets})$, namely:

- $f(1 \text{ jet}) = 1.06 \pm 0.15$
- $f(2 \text{ jets}) = 0.98 \pm 0.19$
- $f(\geq 3 \text{ jets}) = 0.87 \pm 0.28$

7 Conclusion

A method has been established to estimate the jet to tau fake rate in the dijet events, the events with a high jet multiplicity, and the W+jets events. A different jet to tau fake rate has been defined and can be applied for each of the three following analysis cases:

- Jets faking taus as backgrounds for $Z \rightarrow \tau\tau$ or $H \rightarrow \tau\tau$ signals: The jet to tau fake rate is given by the function of jet E_T plotted in Fig.1.
- Jets faking taus as backgrounds for $Z \rightarrow \tau\tau + \text{jets}$ signal: The jet to tau fake rate is given by $\text{FkR}(\text{jet } E_T, \text{sum } E_T) \times f(N \text{ jets})$, where FkR is the matrix reproduced in Fig.3 and $f(N \text{ jets})$ has the following values:

- $f(1 \text{ jet})=1.06\pm0.06$
- $f(2 \text{ jets})=0.98\pm0.10$
- $f(\geq 3 \text{ jets})=0.87\pm0.14$
- $W + \text{jets}$ process where a jet fakes a tau lepton: The jet to tau fake rate is given by $FkR(\text{jet } E_T, \text{sum } E_T) \times f(N \text{ jets})$, where $f(N \text{ jets})$ has the following values:
 - $f(1 \text{ jet})=1.06\pm0.15$
 - $f(2 \text{ jets})=0.98\pm0.19$
 - $f(\geq 3 \text{ jets})=0.87\pm0.28$

The later jet to tau fake rate is the one to be applied to predict the number of expected background events due to jets faking taus in the $t\bar{t} \rightarrow l\tau\nu\nu qq$ analysis. This is given by

$$FkR(\text{jet}E_T, \text{sum}E_T) \times 0.87 \pm 0.28,$$

where FkR is the matrix of Fig.3. This gives a conservative uncertainty of 30%.

This fake rate has been applied with success to make predictions of the number of expected events and distributions in N jets control regions for the $t\bar{t} \rightarrow l + \tau$ analysis, described in the note 7908, 8627 and 9027.

References

- [1] S.Tourneur et al., Tau fake rate, Tau meeting, 12/01/2005
- [2] S.Tourneur et al., Tau fake rate, Tau meeting, 12/08/2005
- [3] Sarah Demers et al., Event Selection and Acceptance for Top to Tau Dilepton Analysis, internal CDF note 6921, March 2004
- [4] Tony Vaiciulis et al., Determining the Jet Fake Rate to Hadronic taus from the Data, CDF note 6784, May 2003
- [5] <http://www-cdf.fnal.gov/internal/people/links/StephaneTourneur/tauIDplots/list.html>
- [6] S.Tourneur, [Top Properties Workshop, 11-10-2005](#)
- [7] Y. Dokshitzer et al., Basics of perturbative QCD, ed. J.Tran Thanh Van, Editions Fontières, Gif-sur-Yvette (1991).
- [8] Ya.I. Azimov et al., Z.Phys. C **27** (1985) 65; Z.Phys. C **31** (1986) 213.