CDF/ANAL/TOP/CDFR/6638
August 22, 2003
Version 1.0

Identification of $\tau^- \to \pi^0\pi^-\nu_\tau$ at CDF

Matt Reece and Henry Frisch
Enrico Fermi Institute
University of Chicago

DRAFT

Abstract

We propose identifying tau leptons in top quark decays by focusing on the exclusive decay modes one at a time, rather than bundling them all together with a cone algorithm, in order to exploit the differing characteristic signals and background levels. We have begun by considering the decay $\tau^- \to \pi^0\pi^-\nu_\tau$. This decay accounts for 39.3% of all hadronic tau decays. Using a sample of 45,000 top pair decays created with Pythia and Tauola in which one W is forced to $\tau\nu$ and the other to light quarks, we find that 49% of the $\tau^- \to \pi^0\pi^-\nu_\tau$ decays have the charged track and both photons in a single tower, 20% have the track and one photon in the same tower and the other photon in a second tower, 16% have the two photons in one tower and the track in a different tower, and 15% have the three decay products in three separate towers. Alternatively, 90.1 % of the events are contained in two neighboring towers in eta and 96.4% have a maximum spread of 2 towers in eta. We observe that, with default settings, the TauFinder module makes a CdfTau that matches the HEPG tau 78.4% of the time for $\tau^- \to \pi^0\pi^-\nu_\tau$ decays, identical to the 78.0% we find for all hadronic tau decays. This can be improved to at least 85% with lowered thresholds. To improve efficiency in the case of top dilepton events, in which event topology can reject a large amount of background, we propose an alternative τ identification method based on information in the track and π^0 collections together with calorimetric information from CdfJet objects. We compare a number of variables for the $\tau^- \to \pi^0\pi^-\nu_\tau$ decays and for light quark jets; there are significant differences in the distributions, for example in the mass of the seed track and the nearest track outside the tau cone, and the angle to the nearest track. These plots can be used along with the impact parameter to construct a tau likelihood to select this specific decay mode with (hopefully) better background rejection and efficiency than a generic cone algorithm.

1 Introduction

The decay of the top quark into tau leptons is of particular interest as both the tau and the top are third generation and are heavy, and hence would (must?) couple to whatever mechanisms generate mass. One specific example is the decay of the top into a charged Higgs, H^+ and a b-quark; this channel would enhance tau final states over the electron and muon states, which in the SM have to be at the same rate (to terms of order $(m_{lepton}/M_W)^2$). Within the SM, the addition of the tau channel to the top dilepton analysis in principle can double the available branching ratio from 4/81 to 8/81 (we assume that the tau-tau channel is not included- however, with b-tagging, some acceptance should be gained here as well).

Most previous analyses of top decay [1, 2, 3, 4] have used a cone-algorithm to distinguish the narrow 'jet' from hadronic tau decay from the jets of quarks and gluons (the exception has been the track-based analysis of Michelle Gallinaro, which focused on the one-prong decays exclusively [5]). Tau decay, however, proceeds through a small number of exclusive channels, each containing a specific number (generally small, i.e. 0, 1, or 2) of charged pizeros, and either 1 or 3 tracks. CDF has invested a great deal of effort in identifying pizeros with the CES and separating single photons from pizeros for values of $\mathrm{E_T}$ up to about 30 GeV. The question is can we gain in efficiency and in signal-to-background by treating some of the topologies separately, exploiting the good tracking and the pizero identification? We see this as a subset of tau identification- those channels for which this is too hard, or the gain too small, would then still be lumped into a more general tau identification, such as a tau-likelihood, so that no decay modes with decent signal-to-noise are lost.

The tau lepton decays to $\pi^-\pi^0\nu_\tau$ $(25.41 \pm 0.14)\%$ of the time, generally through the $\rho(770)$ resonance [6]. The branching ratio to all semihadronic decays is 64.7%, so this decay accounts for 39.3% of all hadronic taus. We seek to improve our measurement of this decay. The π^0 decays to two photons 98.8% of the time [6]. The module Pi0ReconstructionModule in Cdfsoft2 searches for strip and wire clusters in the CES associated with a τ. For energetic π^0's, the CES clusters of the photons should overlap, while for very soft photons, they may be well-separated. Even when the photons overlap, the width of the CES cluster can discriminate between a single photon and the pair from a π^0 for cluster $\mathrm{E_T}$ up to around 30 GeV.

Earlier [8] we found that in a $t\bar{t}$ sample, using the Cdfsoft2 release 4.7.1 and the default settings that are in the Production TCL files, only 73.2% of hadronically decaying Monte Carlo τ's with $|\eta| <$ 1.1 and $Et > 15$ GeV (as reported by the Monte Carlo) were reconstructed by TauFinderModule. This is a large loss to take at the beginning of a search for top events involving τ leptons, so we wish to improve the efficiency.

We intend to supplement the tau identification method of the Tau Group [4] by using the CdfTrack and Pi0Candidate collections to identify the $\tau^- \to \pi^0\pi^-\nu_\tau$ decay channel, independently of TauFinderModule. Further calorimetric information may be obtained by searching for a CdfJet associated with a seed track. The plots in Section 5.1 demonstrate the feasibility of this approach.

2 Monte Carlo Sample

Using the HEPG block, we select events containing a τ from the decay of a W^- from top. We require that the τ decay contains a single π^0 and exactly one charged daughter, a π^-. We then locate the two photons from π^0 decay in the decay chain. Using the HEPG four-momentum for the two photons and the π^-, we compute the (η, ϕ) of each of these three particles and the corresponding tower. There are four basic cases to consider:

- Case 1: All three decay products lie in a single tower.
- Case 2: The two photons from the π^0 land in one tower, while the π^- is in another tower.
- Case 3: One photon from the π^0 and the π^- land in one tower, while the other photon is in a different tower.
- Case 4: Each of the three decay products is in a separate tower.

Case 1 represents the narrowest cluster, and so would be expected to have the best signal-to-noise against jets. Case 2 seems convenient for determining π^0 information. Case 4 seems to have the cleanest kinematics for analysis; In our Monte Carlo sample we find that 49% fit case 1, 16% fit case 2, 20% fit case 3, and 15% fit case 4.

Using the information in the HEPG bank, we examined the spread in η of towers involved in the decay, i.e. if t_1, t_2, t_3 are the towers in which the π^0 and the two photons land, then spread $\equiv \max |t_i - t_j|$. The left-hand plot of Fig. 1 shows the spread in η. We find that 52.2% of events have all three decay products in one tower, and 38.7% have a spread of one tower. Only 9.1% of events have more than a one tower spread in eta, and only 3.6% of events have more than a two tower spread.

3 Reconstruction of CdfTau Objects

3.1 TauFinderModule parameters

TauFinderModule [7] has a setting for the minimum E_T required of the "seed tower," as well as of any other tower associated to the tau ("shoulder towers"). It rejects clusters that spread across too many towers. These are controlled by "talk-to" parameters, CaloSeedEt, CaloShoulderEt, and CaloNumberOfTowers. By default, the seed tower is required to have 6 GeV deposited, while shoulder towers have 1 GeV deposited. The number of towers allowed for a tau cluster is set to 6.

3.2 Reconstruction

Reconstructed events include a CdfTauCollection created by TauFindModule. We 'match' a CdfTau object to the parent HEPG tau to establish that a HEPG tau was found by TauFinderModule. We consider here only those HEPG taus which are from W and decay to $\pi^-\pi^0\nu_\tau$. We determine the (η, ϕ) of the HEPG tau, then loop over the CdfTauCollection for the event and compute $\Delta R^2 =$

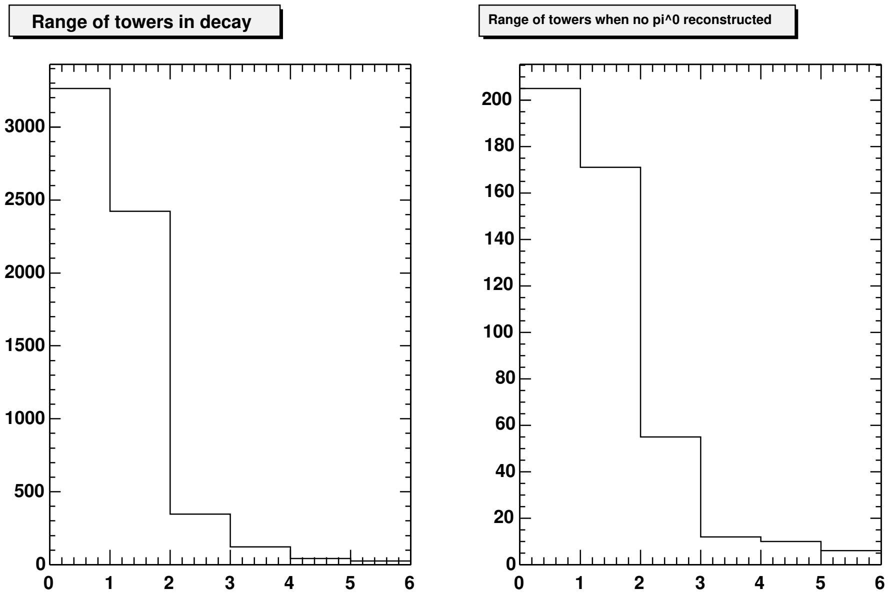

Figure 1: Left: The spread in η, $max|t_i - t_j|$, calculated at HEPG level, between towers in all events where the τ decays as $\tau \to \pi^0\pi^-\nu$. Right: the same spread, but only for those cases in which we do not reconstruct π^0's. The distribution is roughly the same, so it appears that the primary reason we miss the π^0's is *not* that they are too far from the seed track.

Selection	Events	% Passing	% Passing
Total Events	45,000	100%	–
Geometric and Kinematic Acceptance	26,433	59%	–
Tau Decays Hadronically	17,102	38%	100%
TauFinder Created CdfTau (default)	13,346	30%	78%
TauFinder Created CdfTau (low thr)	14,452	32%	85%
Tau Decays Hadronically	17,102	38%	—
Tau decays $\tau^- \to \pi^0\pi^-\nu_\tau$	6,258	14%	100%
TauFinder Created CdfTau (default)	4,907	11%	78%
TauFinder Created CdfTau (low thr)	5,185	12%	83%

Table 1: The efficiency of reconstruction using TauFinder for $t\bar{t}$ events in which the W^- is forced to $\tau\nu$ and the W^+ to $u\bar{d}$.

$(\eta_{HEPG} - \eta_{CdfTau})^2 + (\phi_{HEPG} - \phi_{CdfTau})^2$. The CdfTau with the minimum ΔR^2 from the HEPG tau is considered to 'match' the HEPG tau if $\Delta R < 0.1$. To each CdfTau, there is associated a collection of CdfPi0 objects reconstructed by Pi0ReconstructionModule.

Table 1 summarizes the acceptances at Monte Carlo and the efficiency of TauFinderModule. In a sample of 45,000 events, we find that 26,433 pass geometric and kinematic acceptance cuts. (These are $|\eta| < 1.1$, $E_T > 15$ GeV, and $\mathrm{p_T} > 4.5$ GeV/c for the seed track.) Of these, 6,258 decay

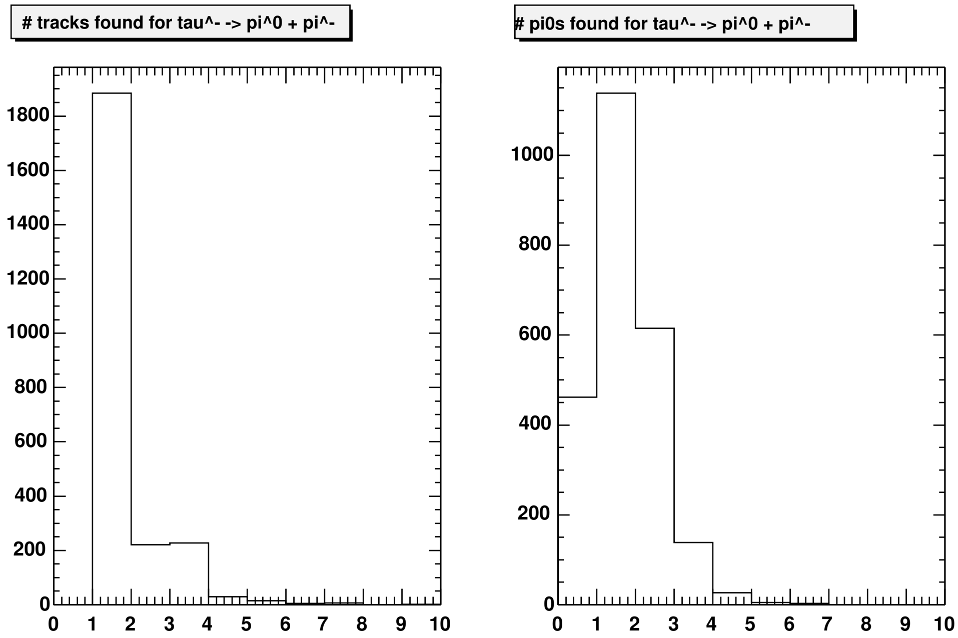

Figure 2: Left: in CdfTaus reconstructing $\tau \to \pi^0\pi^-\nu$, the number of tracks. Exactly one track is found 78.9% of the time. Right: in CdfTaus reconstructing $\tau \to \pi^0\pi^-\nu$, the number of reconstructed π^0's. None are found in 19.3% of the taus, while exactly one is found in 47.7% of the taus.

to $\pi^-\pi^0\nu_\tau$. We find matching CdfTaus for 4,907 of those, i.e., 78.4%. We find a total of 17,102 hadronically decaying taus, of which 13,346 are reconstructed, for an overall reconstruction efficiency of 78.0%. This shows a small improvement over the Cdfsoft2 release 4.7.1 result of 73.2% [8].

Figure 2 shows the distributions in the number of tracks found in reconstructed CdfTau objects for the $\tau^- \to \pi^0\pi^-\nu_\tau$ decay mode on the left; exactly one track is found 78.9% of the time. The right-hand plot shows the number of reconstructed π^0's as given in the CdfTau object; exactly one is found in 47.7% of the taus by Pi0ReconstructionModule. However asking for at least one track and at least one π^0 has a net efficiency of 80.7%.

By lowering the seed tower threshold to 1 GeV and the shoulder tower threshold to 0.75 GeV, we find that the reconstruction efficiency improves to 85%, recovering roughly half of the previously unreconstructed Monte Carlo taus. This is a large improvement, and we use this as the default reconstruction for the studies described below. Further lowering the seed tower threshold is problematic; at low threshold, a large number of towers will be found above the minimum and the tau candidate will be discarded [9].

The distribution of the energy in HEPG of the π^0 in $\tau^- \to \pi^0\pi^-\nu_\tau$ decays with a matched CdfTau is shown in Figure 3 for those which are reconstructed (in black) and those which are not (in red). Many of the π^0s that have very low energy at the HEPG level are not reconstructed as part of the CdfTau, something that might benefit from further optimization.

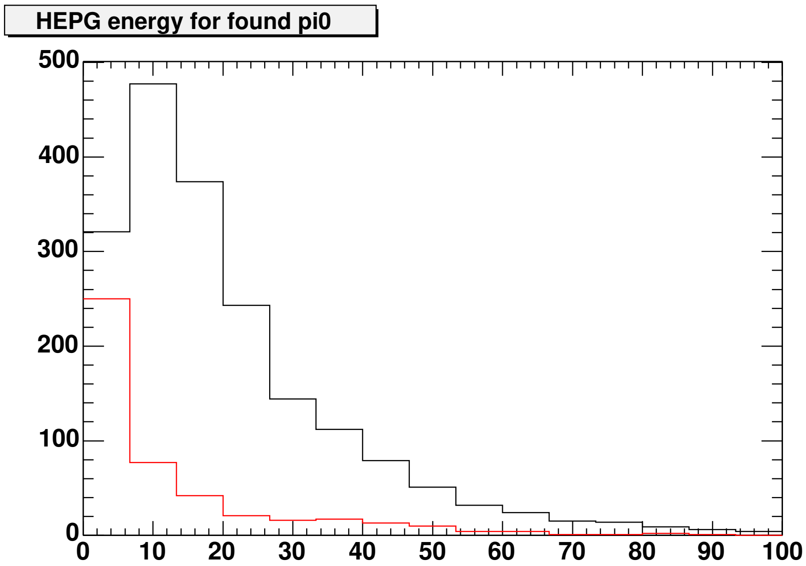

Figure 3: Histogram of the energies of π^0's, as reported by HEPG. The black curve is for the case in which at least one π^0 is reconstructed; the red curve is for the case in which no π^0 is reconstructed. Note that we often fail to reconstruct low-energy π^0's.

4 The Default Tau Group Event Selection

The second step after the reconstruction of CdfTau objects by the TauFinder module is to actually select tau candidates. The tau group's ID cuts are as follows ([4]):

- No tracks or pi0s with $p_T > 1.0$ GeV in isolation annulus ($\alpha_\tau < \alpha < 0.5$), where $\alpha_\tau \equiv \min(10°, 5 \text{ GeV} \cdot \text{rad}/E_\tau)$. [4].
- Calorimeter isolation (in cone of 0.4) $\Sigma E_T < 3$ GeV
- $|z_0|$ of seed track < 60 cm
- $\xi_e = E_{had}/\Sigma(P) > 0.1$ (anti-electron)
- No CMU/P/X stubs in a 15° cone (anti-muon)
- Mass of tracks and π^0's < 1.8 GeV$/c^2$
- Calorimeter mass < 5 GeV$/c^2$

We find that the number of $\tau^- \rightarrow \pi^0\pi^-\nu_\tau$ events passing these cuts is 2,157 out of 6,149 within geometric and kinematic acceptance at the Monte Carlo level, for an efficiency of 35%. On the other hand, 1,385 jets pass these cuts out of 61,053 jets within the geometric and kinematic acceptance (but not requiring a matching CdfTau), giving a fake rate of 2.3%.

5 An Alternative Reconstruction: Using Tracks and π^0's

We attempt to reconstruct τ's from top quark pair production, independently of TauFinderModule, using solely the track and π^0 block. For Monte Carlo τ's decaying as $\tau^- \to \pi^0\pi^-\nu_\tau$, and for the light quarks in Monte Carlo (the u and $\bar{d}$ from W^+ decay), we search for a matching CdfTrack, defined as the highest $\mathrm{p_T}$ track within a cone of 0.1 in $\eta - \phi$ around the Monte Carlo particle's momentum. We require that the $\mathrm{p_T}$ of this matching track be at least 4.5 GeV/c. We find that 97.2% of the Monte Carlo τ's decaying as $\tau^- \to \pi^0\pi^-\nu_\tau$ are matched to CdfTracks in this way. This is better than we have been able to achieve by tuning the parameters of TauFinderModule. Because of this, we would like to try to identify taus in this decay mode solely from the track and π^0 information. A τ could be identified in either this way or using TauFinderModule, so that this should not cost us any τ's. In the case of top dilepton events, having tagged b jets and a good electron or muon should sharply cut down background, so we do not necessarily need extremely strong background rejection. We compare distributions of various quantities to try to determine a method of distinguishing $\tau^- \to \pi^0\pi^-\nu_\tau$ decays from light quark jets in this sample. Many quantities may be computed simply from the track and π^0 collections; for others (e.g. the hadronic content of the jet), we match to a CdfJet and use the calorimetric information it gives us.

5.1 Comparing Distributions from $\tau^- \to \pi^0\pi^-\nu_\tau$ and Light Quark Jets

In order to determine the selection cuts we have plotted a number of quantities from both taus in the $\tau^- \to \pi^0\pi^-\nu_\tau$ decay mode and light quark jets in Figures 4 to 11. These plots are based on the reconstructed CdfTrack and Pi0Candidate collections that are matched to taus or light quark jets in the Monte Carlo (HEPG) bank. We use these to determine a set of tentative ID cuts for the $\tau^- \to \pi^0\pi^-\nu_\tau$ decay mode. There are many features worth noting. The τ's mostly have only one track, while very few have more than three reconstructed tracks. The $\mathrm{p_T}$ distributions for the second and third tracks fall off very sharply for real taus, but not for fakes. The invariant mass of the first and third highest $\mathrm{p_T}$ tracks in a tau is a good discriminant between taus and jets. Taus rarely have more than one π^0. The angle between the seed track and the nearest track not in the tau cone also tends to be much larger for τ's.

These distributions, and the numbers of photons, pizeros, and tracks in the three towers typically associated with a $\tau^- \to \pi^0\pi^-\nu_\tau$ decay, can be the input to a likelihood algorithm much like that for electrons and muons [10].

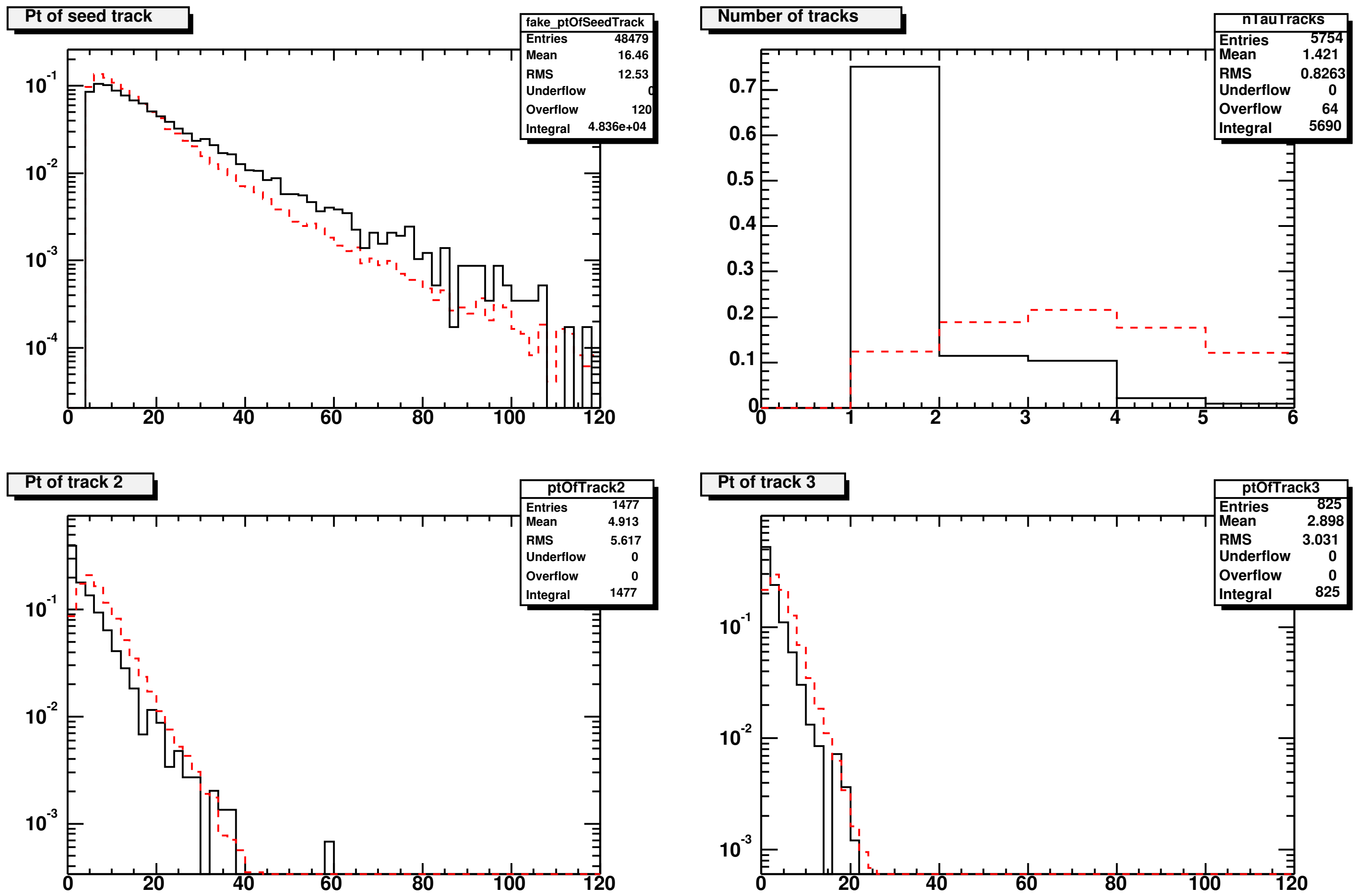

Figure 4: Variables compared between CdfTaus reconstructed from light quark jets (in red, dashed) and taus decaying to $\pi^0\pi^-$ (in black).

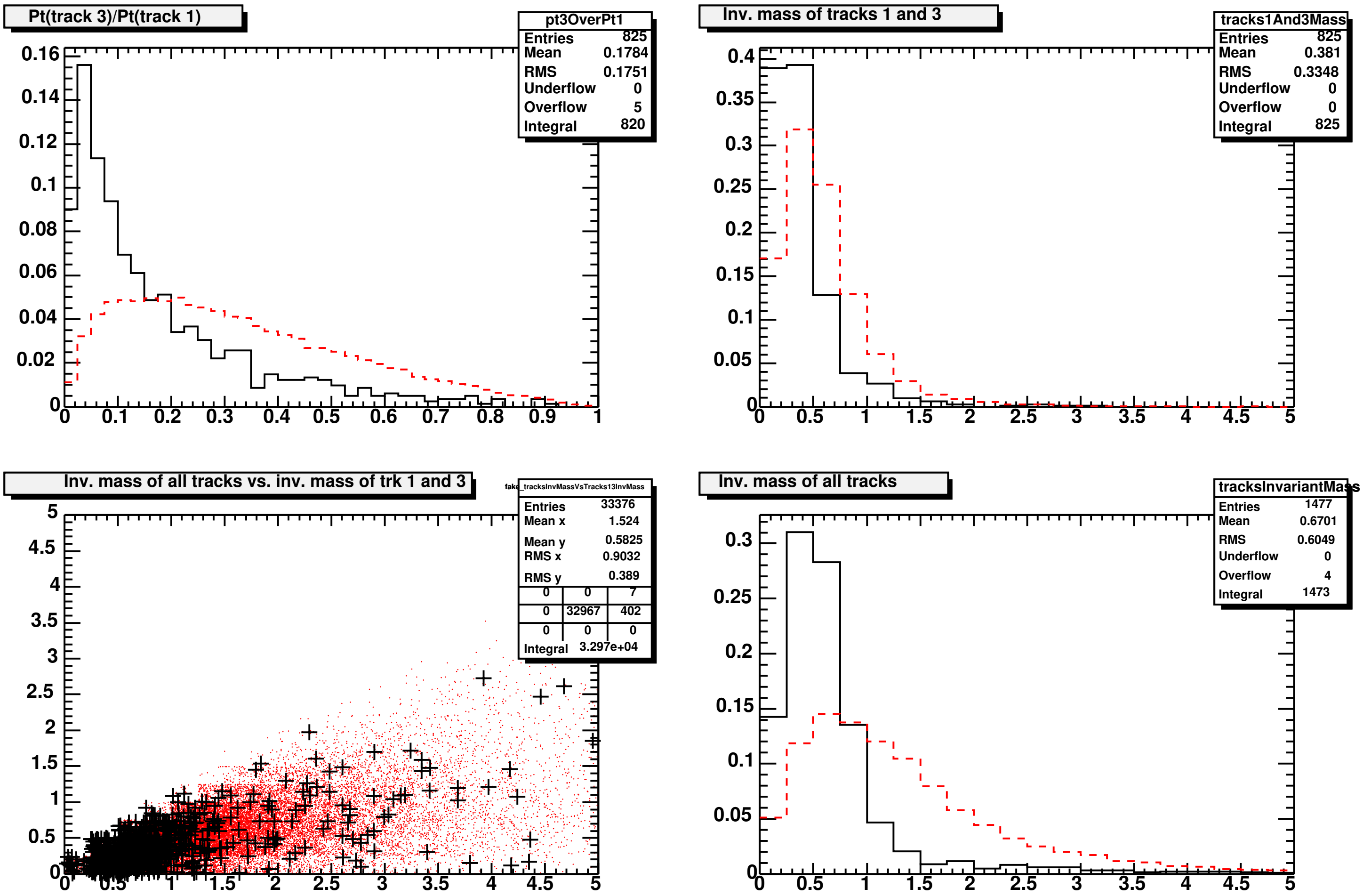

Figure 5: Variables compared between CdfTaus reconstructed from light quark jets (in red, dashed) and taus decaying to $\pi^0\pi^-$ (in black).

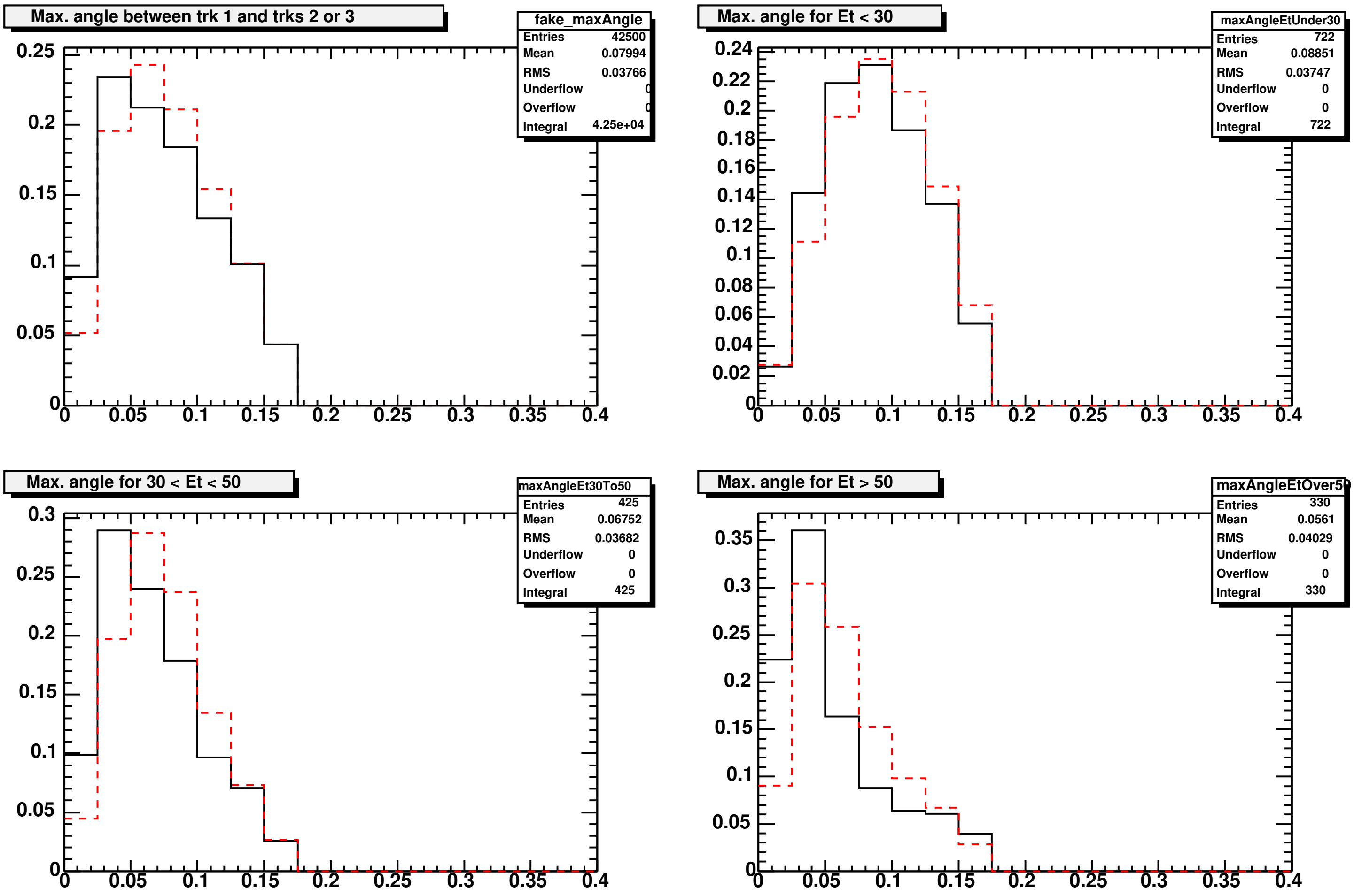

Figure 6: Variables compared between CdfTaus reconstructed from light quark jets (in red, dashed) and taus decaying to $\pi^0\pi^-$ (in black).

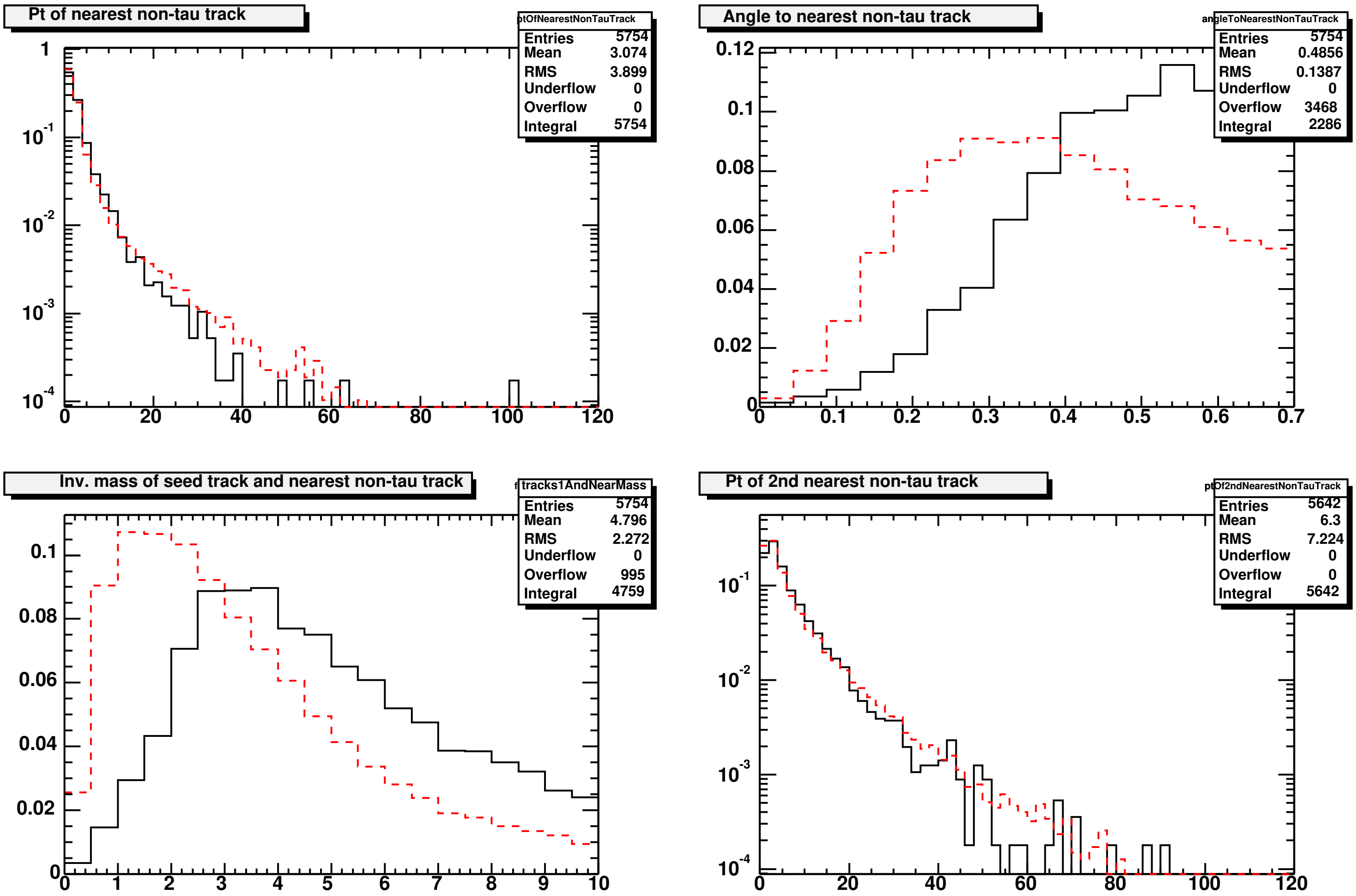

Figure 7: Variables compared between CdfTaus reconstructed from light quark jets (in red, dashed) and taus decaying to $\pi^0\pi^-$ (in black).

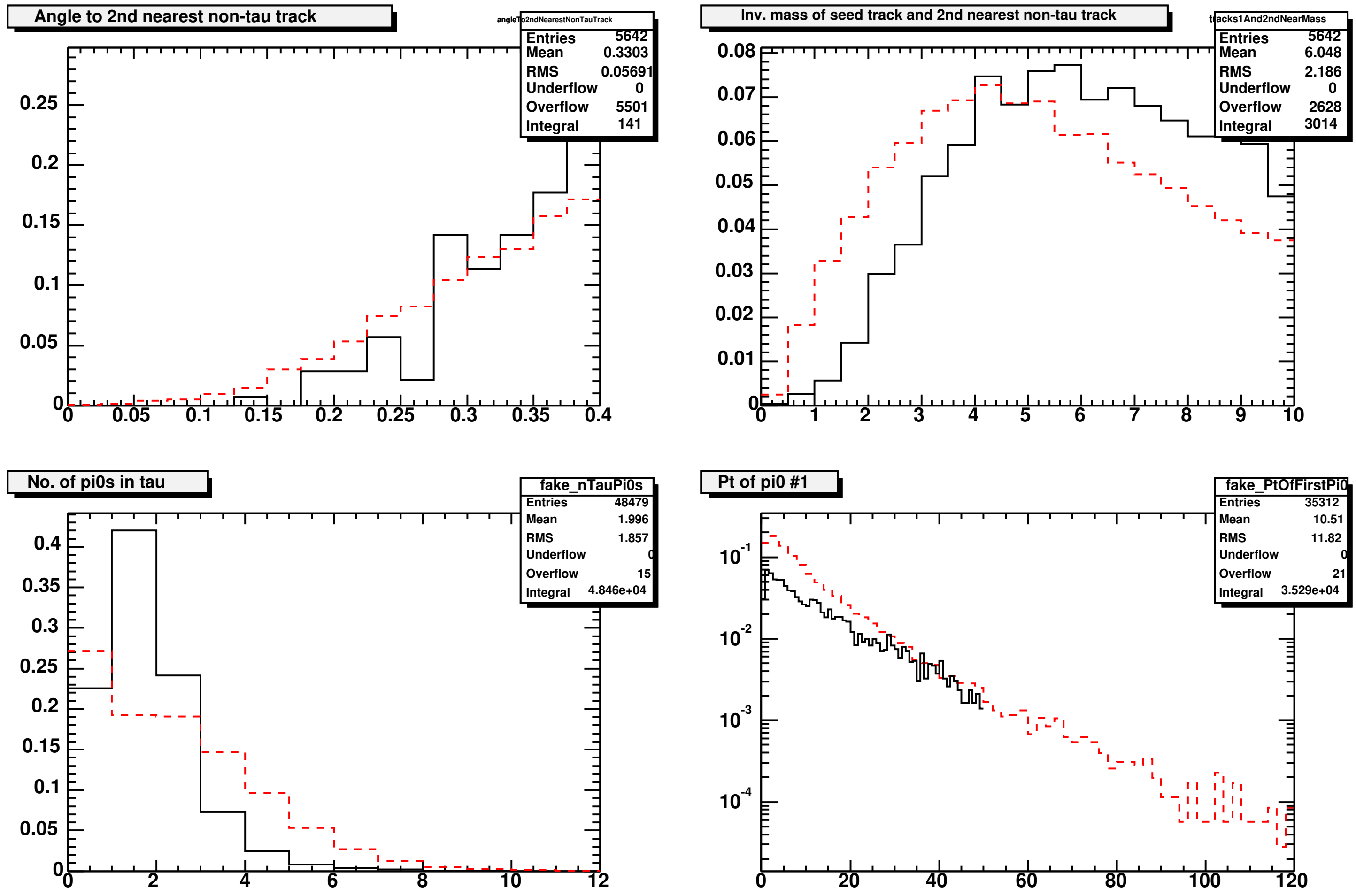

Figure 8: Variables compared between CdfTaus reconstructed from light quark jets (in red, dashed) and taus decaying to $\pi^0\pi^-$ (in black).

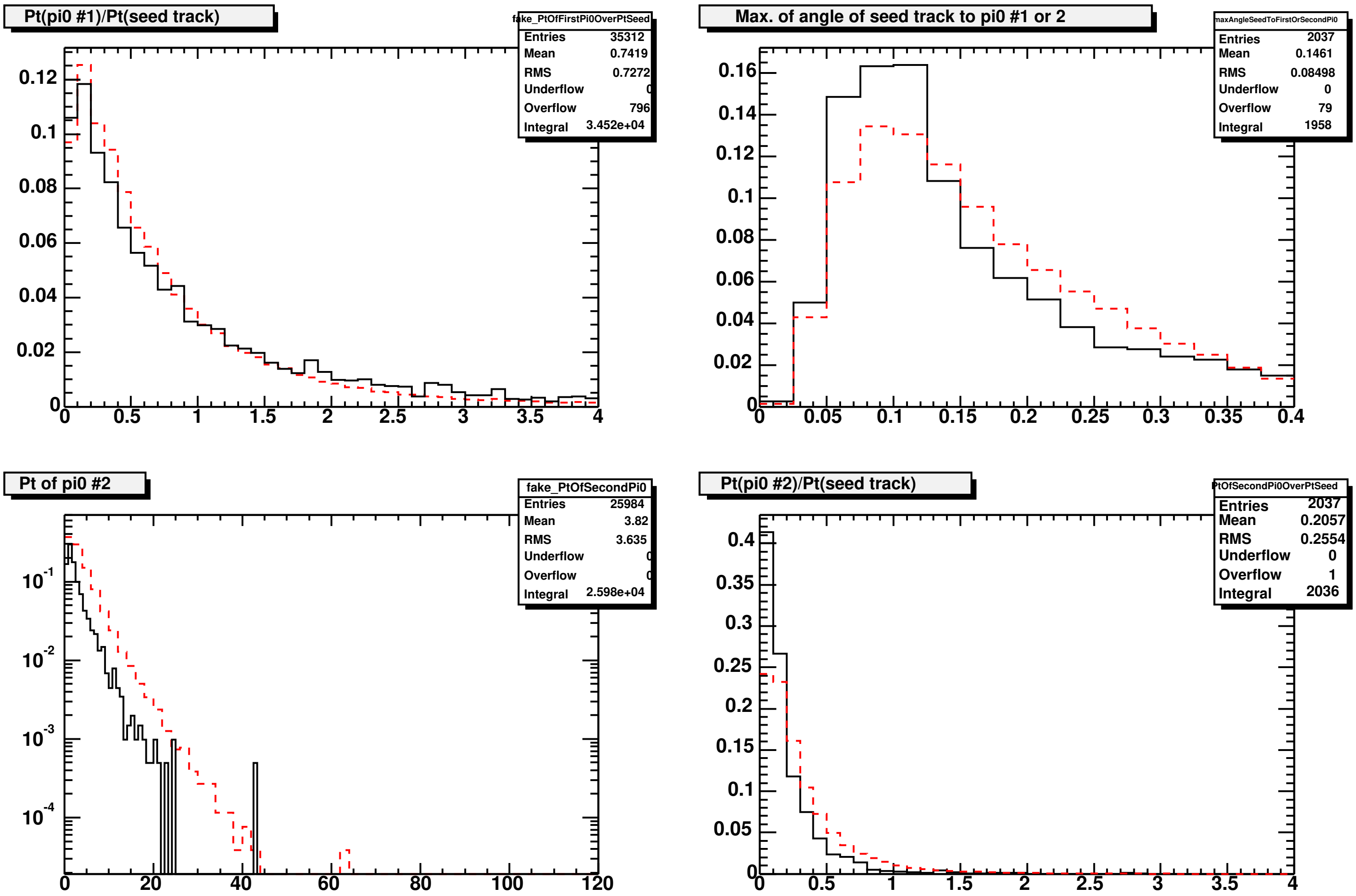

Figure 9: Variables compared between CdfTaus reconstructed from light quark jets (in red, dashed) and taus decaying to $\pi^0\pi^-$ (in black).

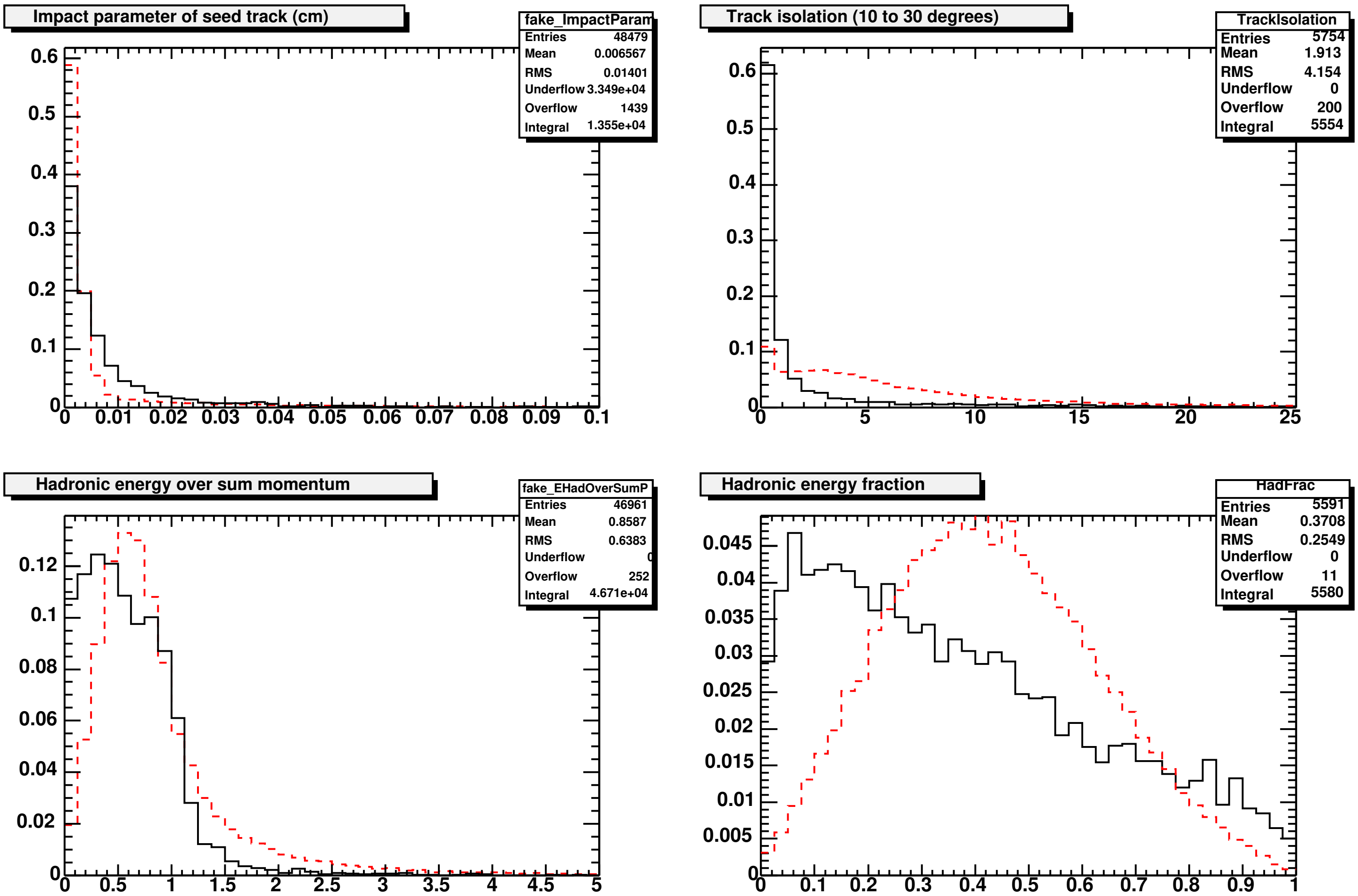

Figure 10: Variables compared between CdfTaus reconstructed from light quark jets (in red, dashed) and taus decaying to $\pi^0\pi^-$ (in black).

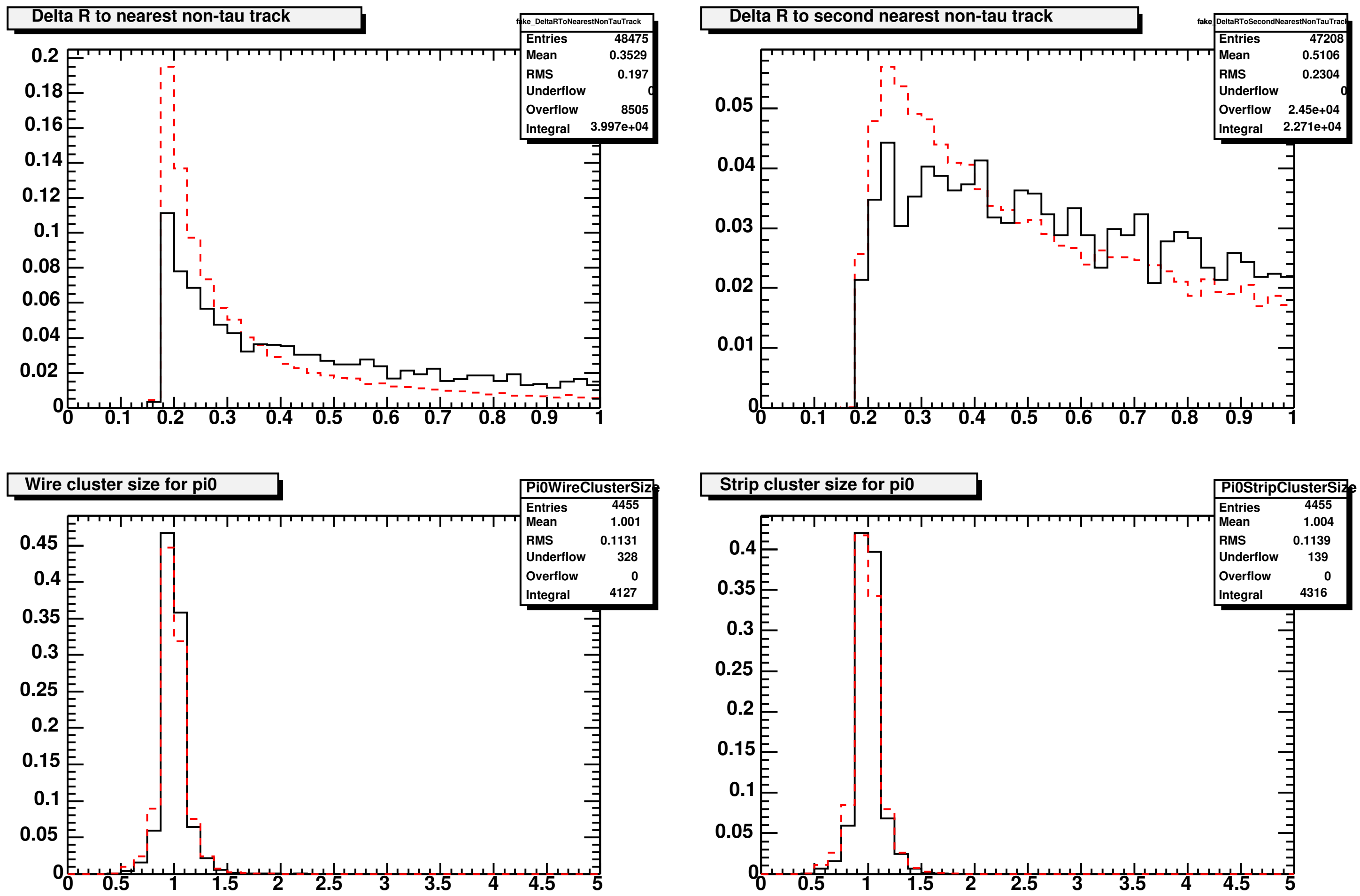

Figure 11: Variables compared between CdfTaus reconstructed from light quark jets (in red, dashed) and taus decaying to $\pi^0\pi^-$ (in black).

We define a tau cone size by $\alpha_\tau = \min(10^\circ, 5\ \mathrm{GeV}\cdot\mathrm{rad}/E_\tau)$. [4] Here E_τ is determined by matching the seed track to a CdfJet, if possible. Based on these distributions, we consider a set of cuts as follows:

- # of tracks ≤ 3
- Track isolation (cone $\alpha_\tau\ -\ 30^\circ$) < 1.0 GeV
- $\mathrm{p_T}$ of second track ≤ 40 GeV/c
- $\mathrm{p_T}$ of third track ≤ 20 GeV/c
- # of π^0's ≤ 2
- Invariant mass of tracks ≤ 2.0 GeV/c^2
- $\frac{\mathrm{p_T(third\ track)}}{\mathrm{p_T(first\ track)}} \leq 0.5$
- Invariant mass of first and third track ≤ 1.0 GeV/c^2
- For the highest $\mathrm{p_T}$ π^0 in the tau: wire width between 0.7 and 1.3, strip width between 0.7 and 1.4
- Hadronic fraction < 0.75
- Calorimeter mass < 4 GeV/c^2
- $\xi_e = E_{had}/\Sigma(P) > 0.1$ (anti-electron)

We find that the number of $\tau^- \to \pi^0\pi^-\nu_\tau$ events passing these cuts is 2,219, of 6,149 (within geometric and kinematic acceptance limits at the HEPG level). This gives an efficiency of 36%. On the other hand, 904 jets pass these cuts, of 61,053 within geometric and kinematic acceptance limits, giving a fake rate of 1.5%. Only 302 hadronic taus in other decay modes pass these cuts, of 10,993 within geometric and kinematic acceptance limits, so that these cuts identify only 2.7% of other decay modes.

6 Comparison: Tau Group Methods and Track/Pi Methods

Taking the OR of our cuts and the tau group's cuts, we find that 3,017 $\tau^- \to \pi^0\pi^-\nu_\tau$ events pass either cut, while 1,868 jets pass either cut. This gives us an overall efficiency of 49% for identifying reconstructed $\tau^- \to \pi^0\pi^-\nu_\tau$ decays, while 3% of the light quark jets are identified as taus. Also, 4,590, or 42%, of the other hadronically-decaying taus in the sample pass either cut (most of these are identified by the tau group's cuts and not by our cuts).

7 Conclusions

Using the CdfTrack and Pi0Candidate collections gives an alternative method to identify τ leptons in the $\tau^- \rightarrow \pi^0\pi^-\nu_\tau$ decay mode, particularly for those which do not deposit large amounts of energy in the calorimeter and so do not form a CdfTau object. The number of identified taus in a Monte Carlo sample using our proposed method is roughly twice the number identified by the standard cuts on CdfTau objects. We plan to incorporate this as part of a general tau likelihood, allowing the amount of signal and background to be tuned, based on the distributions (Figures 4-11) from taus and from light quark jets from top decays.

8 Acknowledgments

We would like to thank Fedor Ratnikov, Anton Anastassov, and Steve Levy for much help.

References

[1] A. Roodman, Ph.D thesis, Univ. of Chicago, *A Measurement of the ppbar $\rightarrow$ W $\rightarrow$ Tau Nu Cross Section Times Branching Ratio as a Test of Lepton Universality*; June 1991.

[2] M. Hohlmann, Ph.D thesis, Univ. of Chicago *Observation of Top Quark Pairs in the Dilepton Decay Channel Using Electrons Muons, and, Taus*; August 1997.

[3] *The $\mu\tau$ and $e\tau$ Decays of Top Quark Pairs Produced in $\bar{p}p$ Collisions at $\sqrt{s}$ = 1.8 TeV*, The CDF Collaboration (M. Hohlmann,M. Gallinaro, et al.) Phys. Rev. Lett. **79**, 3585 (1997).

[4] Tau ID cuts. http://www-cdf.fnal.gov/internal/physics/exotic/tau/tau_id.html, and private communication (the web page is not up-to-date as of this moment).

[5] M. Gallinaro, Ph.D thesis, University of Rome, *Observation of the Top Quark at CDF in the Dilepton Channel with Electrons Muons and Taus*; February 1996.

[6] Particle Data Group; K. Hagiwara et al., Phys. Rev. D66,010001 (2002).

[7] F. Ratnikov, *Tau Reconstruction for CDF Run II*, CDF Note Number: CDF/DOC/CDF/PUBLIC/6252 and private communication.

[8] Using Track Impact Parameter to Distinguish Hadronically-Decaying τ's in Top Quark Decays from Light Quark Jets. M. Reece and H. Frisch. CDF/ANAL/TOP/CDFR/6203 (2002).

[9] We have not done a systematic study, so 85% is a lower bound on the improvement.

[10] See B. Knuteson, CDF5934, CDF 5902, and CDF 6083 for a description of lepton ID using likelihoods for central electrons, plug electrons, and central muons, respectively.