Measurement of $\sigma(p\bar{p} \to Z) \cdot B(Z \to \tau\tau)$ in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV

MEASUREMENT OF $\sigma(p\bar{p}\rightarrow Z)$ . . .

092004-3

PHYSICAL REVIEW D 75, 092004 (2007)
I. INTRODUCTION

Experimental studies of the production mechanism and decay properties of $Z$ and $W$ bosons provide an important insight into the gauge structure of the standard model (SM) of particle physics and have been a major focus of many dedicated studies since their discovery 20 years ago. In addition to providing a precision test of the gauge structure of the electroweak sector of the SM, measurements of the properties of gauge bosons shed light on the properties of the still undiscovered Higgs boson and provide constraints on new physics beyond the SM. Both Higgs and supersymmetry (SUSY) phenomenology predict signatures with multiple tau leptons in the final state [6–8]. In the case of the search for a Higgs boson, $H$, with low to intermediate (below the $WW$ threshold) masses, the final state with tau lepton pairs comprises $10\%$ of all final states, and is smaller only than the $b\bar{b}$ decay mode. Efficient tau lepton reconstruction and identification allow for Higgs searches using the dominant Higgs production mechanism, $gg \rightarrow H$, which is precluded for $b\bar{b}$ final states due to overwhelming QCD $b\bar{b}$ backgrounds. In Ref. [7], it was demonstrated that the $H \rightarrow \tau \tau$ sensitivity is comparable to that of the Higgs searches in the $b\bar{b}$ mode, where Higgs bosons are produced in association with a heavy gauge boson. Another important search is the “golden mode” for the observation of SUSY at hadron colliders [8] in the process of chargino-neutralino production, $p\bar{p} \rightarrow \tilde{\chi}^{\pm}_1 \tilde{\chi}^{-}_1 + X$, with three charged leptons in the final state. Multi-tau final states become dominant for moderately large $\tan\beta$ values, which are preferred by the LEP data if supersymmetric particles exist. Efficient and reliable reconstruction of hadronic tau candidates, as well as robust techniques for background estimation, maintain small systematic uncertainties and therefore maximize the sensitivity of searches for new phenomena. CDF has recently published a paper describing a search for SUSY Higgs decaying to di-taus [9], using Run II data and relying on the techniques described in detail in this paper, and has demonstrated a significantly improved sensitivity compared to an analogous Run I search [10].

This paper describes details of the triggering, reconstruction, and identification for hadronically-decaying
taus developed by CDF in Run II, along with methods used in estimating efficiencies and associated systematics. We also describe a novel data-driven method of estimating backgrounds in the di-tau final state and demonstrate that it significantly outperforms the standard techniques based on measuring rates for generic hadronic jets to be mis-identified as hadronically-decaying tau leptons. In the following sections, we describe the CDF II detector and trigger system and details of the data collection, followed by the description of the reconstruction and identification for electrons and hadronic taus, background estimation, and the procedure to extract the value of the cross section.

II. EXPERIMENTAL APPARATUS

CDF II is a multipurpose detector designed for precision energy, momentum, and position measurements of particles produced in proton-antiproton collisions at a center-of-mass energy of 1.96 TeV at the Fermilab Tevatron. The detector is described elsewhere [11]. Here we briefly describe those elements directly related to this analysis. The CDF coordinate system is defined with respect to the proton beam direction which defines the positive z direction, while the azimuthal angle \( \phi \) is measured around the beam axis. The polar angle \( \theta \) is measured with respect to the positive z direction. The pseudorapidity \( \eta \) is defined as \( \eta = - \ln \tan \frac{\theta}{2} \). Transverse components of particle energy and momentum are conventionally defined as projections onto the plane transverse to the beam line, \( E_T = E \sin \theta \) and \( p_T = |\vec{p}| \sin \theta \). The subdetectors critical to this analysis are the silicon vertex detector (SVX II), the central outer tracker (COT), and the central parts of the calorimeter system, namely, the central electromagnetic (CEM), with the shower maximum detector (CES) embedded inside the CEM, central hadronic (CHA), and wall hadronic (WHA) calorimeters. The SVX II is the component of the CDF II detector closest to the beam line and provides precise determination of the vertex position in the transverse plane via \( r - \phi \) tracking. The SVX II is mounted inside the cylindrical 3.2 m long COT. The COT is an open cell drift chamber covering the radii from 0.40 m to 1.37 m designed for precision measurement of particle trajectories. Particle momenta are determined from the trajectory curvature and the strength of the solenoidal magnetic field (\( B = 1.4 \) T). The chamber contains 96 layers of sense wires grouped into eight superlayers of 12 wires each. The superlayers alternate between purely axial wires and stereo wires tilted by \( 2^\circ \) with respect to the beam line. The COT fully covers the region \(| \eta | < 1\) with a momentum resolution \( \delta p_T / p_T \approx 0.0015 \) (GeV/c)\(^{-1}\). When the SVX II tracking information is available, its hits are added to the track helix, improving the resolution.

The CEM is a lead-scintillator calorimeter, while the CHA and WHA consist of alternating iron and scintillator sheets. The CEM, CHA, and WHA have complete azimu-
missing transverse energy $\hat{E}_T$ defined as

$$\hat{E}_T = - \sum_i E_i \hat{n}_i,$$  

where $E_i$ is the magnitude of the transverse energy contained in trigger tower $i$, and $\hat{n}_i$ is the unit vector from the nominal interaction point to the tower in the plane transverse to the beam direction. Level 3 uses a simplified version of the offline reconstruction code, allowing further refinement of the selection. Events selected by level 3 are written to tape.

III. DATA AND TRIGGER

In Run II, CDF has designed and installed a set of triggers suitable for selecting events with two taus in the final state. The di-tau trigger requires at least two hadronic tau candidates in the final state, each with $p_T > 15$ GeV/c. The tau plus $E_T$ trigger requires a hadronic tau candidate with $p_T > 20$ GeV/c and large missing transverse energy, $E_T > 25$ GeV. For final states with one of the taus decaying leptonically, a set of lepton plus track triggers [12] has proven to be a very powerful and efficient tool. The lepton plus track triggers consist of the electron plus track and two complementary muon plus track triggers covering different ranges in $|\eta|$. These triggers require at least one central electron (muon) candidate with $p_T > 8$ GeV/c and one tau candidate. Tau selection is identical across the lepton plus track triggers and described further in the text. The lepton plus track triggers are highly efficient and have been used in nearly all CDF analyses based on the di-tau signature [9].

The measurement presented here is based on 350 pb$^{-1}$ of Run II data collected using the electron plus track trigger in 2002–2004. The luminosity is measured with an accuracy of 5.8% [13].

At level 1, the electron plus track trigger requires at least one calorimeter trigger tower with EM $E_T > 8$ GeV and a matching XFT track with $p_T > 8$ GeV/c. At level 2, these requirements are reapplied using level 2 calorimeter clusters along with the requirement of an associated CES cluster with a pulse height corresponding to an electron candidate with $E_T > 3$ GeV. There is an additional requirement of a second XFT track with $p_T > 5$ GeV/c separated from the electron candidate by at least 10 degrees in $\phi$. At level 3, the trigger requires a reconstructed electron candidate with $E_T > 8$ GeV and an isolated track (tau seed) with $p_T > 5$ GeV/c. The isolation is defined as a requirement that there be no track with $p_T > 1.5$ GeV/c in the annulus $0.17 < \Delta R = \sqrt{\Delta \phi^2 + \Delta \eta^2} < 0.52$ around the seed track. These selected events provide the data sample for further analysis. Additional details of the electron plus track trigger can be found in [12].

IV. ANALYSIS OUTLINE

The event sample is reduced offline by requiring at least one electron candidate with $E_T > 10$ GeV and a hadronic tau candidate of $p_T > 15$ GeV/c (note that due to escaping neutrinos the full energy of the tau undergoing hadronic decay generally cannot be reconstructed; thus, all energy and momenta measurements discussed in this paper refer to the visible products of tau decays). The candidates must be well separated and satisfy fiducial requirements to ensure high trigger efficiency and robust event reconstruction. Trigger selections effectively require both candidates to be in the central region of the detector. To improve the purity of the signal, a set of electron and tau identification requirements is applied. These selections strongly suppress large background contributions associated with multijet and direct photon production. We then apply “event-level” selections that further suppress jet backgrounds, including those from $W +$ jets events, to improve the sensitivity of the measurement. These “event-level” selections are optimized by comparing the topological and kinematic properties of signal events to the dominant backgrounds. Finally, we define several regions complementary to the primary signal region by loosening some of the analysis selections. These complementary regions are designed to be enriched with one of the dominant backgrounds. We then fit the number of candidates in signal and complementary regions to extract the rate of signal and background and convert the former into a cross section measurement. In the following sections, we describe in detail each of the above steps in the analysis.

V. KINEMATIC AND GEOMETRIC ACCEPTANCE

The acceptance is defined as the ratio of the number of $Z/\gamma^* \to \tau\tau \to e\nu\nu_{\tau\tau}\nu$ events with an electron and hadronic tau candidates satisfying kinematic and geometric requirements to the number of events produced within the mass window $66 < M_{\tau\tau} < 116$ GeV/$c^2$. This definition is consistent with CDF measurements of the $\sigma(p\bar{p} \to Z/\gamma^* \to e^+e^-)$ and $\sigma(p\bar{p} \to Z/\gamma^* \to e^+e^-)$ [4]. In the remainder of this section, we describe the electron and hadronic tau reconstruction, followed by the estimate of the acceptance and its uncertainties.

A. Electron reconstruction

Reconstruction of electron candidates begins with an energy cluster in the EM calorimeter that combines up to two calorimeter towers adjacent in $\eta$. If the hadronic calorimeter energy deposition is small compared to the EM energy and there is a COT track pointing to the cluster, such a cluster becomes an electron candidate. The algorithm then looks for possible CES clusters that could be associated with the electron candidate, and retains this information. The electron candidate energy is calculated as a sum of the tower energies measured in the EM calo-
TABLE I. Kinematic and geometrical selections determining the acceptance requirements for $Z \rightarrow \tau\tau$ events.

| CEM Cluster: $E^\text{EM}_T > 10$ GeV | $|\eta^\text{det}| < 0.9$ |
|---|---|
| Matching Track: $p_T^\text{trk} > 8$ GeV/c |
| $|z^\text{trk}_0| < 60$ cm |
| $|z^\text{trk}_{R=137}\text{cm}| < 150$ cm |
| $9 < |z^\text{trk}_{K\text{ces}}| < 230$ cm |
| $|x^\text{trk}_{K\text{ces}}| < 21.5$ cm |
| Tau cluster: $p_T^\tau > 15$ GeV/c |
| $|\eta^\tau| < 1$ |
| Seed Track: $p_T^\tau > 6$ GeV/c |
| $|\eta^\tau| < 1$ |
| $|z^\tau_{R=137}\text{cm}| < 150$ cm |
| $9 < |z^\tau_{K\text{ces}}| < 230$ cm |
| $|z^\tau_{\text{seed}} - z^\tau_{0\text{-trk}}| \leq 5$ cm |

**B. Hadronic tau reconstruction**

Hadronically-decaying tau candidates are reconstructed by matching calorimeter clusters with tracks. The calorimeter cluster is required to have at least one tower with total energy $E_T > 5$ GeV, and the highest $E_T$ tower in the cluster is called a seed tower. All adjacent towers with transverse energy deposition in excess of 1 GeV are added to the cluster. Only clusters consisting of six or fewer towers are retained for further reconstruction. Tracks are reconstructed in the COT, associated SVX II hits are added to the track. All tracks having reconstructed segments in at least two axial and two stereo COT segments that point to the tau calorimeter cluster are associated with the hadronic tau candidate. The one with the highest $p_T$ is selected as the seed track. In this analysis, we require the seed track $p_T > 6$ GeV/c. If these requirements are satisfied, the cluster-track match becomes a hadronic tau candidate.

With the initial tau candidate reconstructed, the algorithm associates additional tracks and reconstructed neutral pions (if found) with the tau candidate. The seed track is used to define signal and isolation cones. The signal cone is defined as a cone with an opening angle $\alpha_{\text{trk}}$ around the seed track:

$$\alpha_{\text{trk}} = \min[0.17, \max(5 \text{ GeV}/E^\text{calo}, 0.05)],$$

where $E^\text{calo}$ is the energy (in GeV) of the calorimeter cluster associated with the tau candidate, and the other parameters are in radians. The energy dependence in the cone definition accounts for collimation of the decay products of highly boosted tau leptons; the lower bound on the cone size is driven by resolution effects.

COT tracks in the signal cone that have (i) at least two axial and two stereo COT segments with at least 6 hits in each segment and (ii) $E_T > 1$ GeV/c in the fiducial volume of the COT, consistent with the trigger requirements. The latter is enforced by the requirement that the track trajectory extrapolation to the radius of the outermost axial superlayer be within the fiducial volume of the tracking chamber, $|z^\text{trk}_{R=137}\text{cm}| < 150$ cm. The track trajectory extrapolation must point to a fully instrumented fiducial region of the CES detector in both the $z$ and local $x$-directions: $9 < |z^\text{trk}_{K\text{ces}}| < 230$ cm and $|x^\text{trk}_{K\text{ces}}| < 21.5$ cm. The track is also required to originate close to the center of the detector by demanding that the track trajectory at the point of closest approach to the beam line must satisfy $|z^\text{trk}| < 60$ cm.
Note that this algorithm makes no distinction between true \( \pi^0 \) mesons and photons, e.g. two photons from a sufficiently energetic \( \pi^0 \) meson \((E_T \approx 10 \text{ GeV})\) are reconstructed as a single cluster, while for a lower momentum \( \pi^0 \) meson the algorithm typically finds two clusters, i.e. “resolved photons.” Distinguishing between these two cases is not necessary for hadronic tau reconstruction since it has little effect on the measured parameters of a reconstructed tau candidate. Therefore, we take a “reconstructed \( \pi^0 \)” to be either a true \( \pi^0 \) candidate or a single photon coming from either a \( \pi^0 \to \gamma \gamma \) decay or from any other source.

As in the case for tracks, for \( \pi^0 \)'s we define a cone around the seed track

\[
\alpha_{\pi^0} = \min[0.17, \max(5 \text{ GeV}/E^{\text{calo}}, 0.10)].
\] (4)

All \( \pi^0 \) candidates found inside the cone with \( E_T > 1 \text{ GeV} \) that have matches in both the \( x \)- and \( z \)-views of CES they are associated with the tau candidate.

Selected hadronic tau candidates are required to satisfy fiducial requirements to ensure efficient reconstruction and triggering (Table I). As in the case of electron tracks (Sec. VA), the tau seed track is required to be fully contained in the fiducial volume of the COT, and its trajectory extrapolation must point to a fully instrumented region of the CES detector in the \( z \)-direction, \( 9 < |z^{\text{CES}}| < 230 \text{ cm} \), to ensure efficient \( \pi^0 \) reconstruction. Figure 1 shows the efficiency of the hadronic tau reconstruction and fiducial selections, described above, as a function of the true visible transverse momentum of a hadronic tau for \( Z/\gamma^* \to \tau \tau \) events as obtained using PYTHIA Monte Carlo [14] and the GEANT-based CDF II detector simulation [15].

The energy of the tau candidate can be measured by using either the calorimeter information or the reconstructed tracks and \( \pi^0 \)'s. Because of the relatively poor energy resolution of the hadronic calorimeter, the latter is preferred. The four-vector momentum of a tau, \( p_T \), is defined as the sum of four-vectors of tau tracks with \( p_T > 1 \text{ GeV}/c \) and \( \pi^0 \)'s with \( E_T > 1 \text{ GeV} \) (both are assumed massless) associated with the tau candidate:

\[
p_T = \sum_{\Delta \Theta < \alpha_{\pi^0}} p^{\text{trk}}_T + \sum_{\Delta \Theta < \alpha_{\pi^0}} p^{\pi^0}_T,
\] (5)

where \( p^{\text{trk}} \) and \( p^{\pi^0} \) denote four-vector momenta of contributing tracks and \( \pi^0 \) candidates, and \( \Delta \Theta \) is the angle between the seed track and a particular track or \( \pi^0 \) candidate. The transverse momentum of a tau candidate is defined as

\[
p_T^x = \sqrt{p_T^x + p_T^z},
\] (6)

where \( p_T^x \) and \( p_T^z \) are the \( x \)- and \( y \)-components of the tau momentum four-vector.

Finally, a correction is applied to the hadronic tau energy, calculated above, that compensates for two effects: (i) inefficiency of \( \pi^0 \) reconstruction; and (ii) false \( \pi^0 \) reconstruction. While neither of these effects are dominant—in this analysis, the correction is invoked for approximately 10% of the candidates—they can result in substantial mismeasurement of a hadronic tau candidate energy. Such mismeasurements are undesirable not only for true hadronically-decaying taus, but also for the background jets or electrons misidentified as taus. In particular, false \( \pi^0 \) reconstruction can promote low energy hadronic jets or electrons misidentified as hadronic taus into the signal region.

The \( \pi^0 \) reconstruction inefficiency results in an underestimate of the tau candidate energy and is caused by either dead wires in the CES or the overlap of a \( \pi^0 \) meson with a track. When additional EM energy, not assigned to any of the \( \pi^0 \) candidates, is detected, the tau candidate energy is corrected by assuming the presence of an additional \( \pi^0 \) with an energy equal to the excess calorimeter energy, and a direction coinciding with the hadronic tau candidate.

For true hadronic taus, false \( \pi^0 \) reconstruction occurs when a charged hadron or an electron (from a photon conversion in \( \tau \to \pi^0 + X \to \gamma \gamma + X \)) deposits a large fraction of its energy in the EM calorimeter. If the charged particle generates a wide shower in the CES, its size may exceed the five-strip/wire limit for 1D CES clusters and thus be split into two clusters. In this case, the cluster

\[\text{FIG. 1 (color online). Efficiency of the hadronic tau reconstruction and fiducial selections as a function of the true visible (neutrino contribution excluded) transverse momentum of a tau for } Z/\gamma^* \to \tau \tau \text{ events as obtained using PYTHIA MC and CDF II detector simulation. The efficiency is calculated for generated hadronic tau decays with } |\eta| < 0.9.\]
closest to the charged particle is not considered a $\pi^0$ candidate, since it is close to a track. The other CES cluster is reconstructed as a $\pi^0$ candidate with an energy nearly equal to the full EM energy in this calorimeter tower leading to an overestimate of the hadronic tau energy. The same effect occurs when a hadronic jet or an electron from electroweak boson decay is misidentified as a hadronic tau promoting backgrounds to higher energies. In the case of electrons misidentified as tau candidates, false $\pi^0$ reconstruction can overestimate the hadronic tau candidate energy by up to a factor of 2, since the candidate energy is computed by adding the charged track $p_T$ and the neutral pion candidate $E_T$ essentially counting electron transverse momentum twice. These cases are identified when most of the tau energy is deposited in the EM calorimeter ($E_{EM}^T/E_{HAD}^T > 0.8$) and when $p_T^\tau$ using tracks and neutral pions disagrees significantly with the calorimeter measurement ($> 3\sigma$ using the hadronic calorimeter energy resolution). For these candidates the $p_T^\tau$ is computed using the calorimeter measurement only.

Figure 2 compares the pure calorimeter-based energy calculation to the one using tracks and $\pi^0$ candidates with appropriate corrections described above by showing the difference between the reconstructed tau candidate energy obtained from the CDF II detector simulation and the true visible energy at the generator level. The figure demonstrates that particle based reconstruction has significantly better energy resolution and is more suitable for hadronic tau energy measurement in our detector.

**FIG. 2** (color online). Comparison of the true hadronic tau transverse energy (as obtained from MC) and the reconstructed tau candidate energy using CDF II detector simulation for $Z \rightarrow \tau\tau$ events. The dashed line corresponds to using only calorimeter energy, while the solid line corresponds to the combination of track and $\pi^0$ energies with appropriate corrections.

**FIG. 3** (color online). Distribution of $M_{\tau\tau}$ for PYTHIA $Z/\gamma^* \rightarrow \tau\tau \rightarrow e\nu\tau_h\ell\nu$ process for all generated events (open histogram) and events passing acceptance requirements (shaded histogram). The latter is scaled by a factor of 10 for illustration purposes. Note that the events were generated with the lower cutoff of $\sqrt{s} = 20$ GeV.

### C. Acceptance definition

Unlike the measurements in $ee$ or $\mu\mu$ channels, in the $\tau\tau$ case the escaping neutrinos do not allow the invariant mass of the $Z/\gamma^*$ decay products to be fully reconstructed, making it impossible to determine whether a particular pair of reconstructed tau candidates has originated from within a particular mass window. While most of the events do originate from within the chosen mass window $66 < M_{\tau\tau} < 116$ GeV/c$^2$, there is a fraction of true $Z/\gamma^* \rightarrow \tau\tau$ events that come from outside the window as illustrated in Fig. 3. To account for this we use a two-step procedure to define the acceptance. We first introduce the “raw” acceptance, $\alpha_0$ as the fraction of $Z/\gamma^* \rightarrow \tau\tau \rightarrow e\nu\tau_h\ell\nu$ events produced within the mass window $66 < M_{\tau\tau} < 116$ GeV/c$^2$ ($M_{\tau\tau}$ is an invariant mass of the two parent taus before decay) that have at least one electron and one tau candidate in the central part of the detector. Electron and tau candidates are required to be separated in $\eta$-$\phi$ space by $\Delta R > 0.9$ and the $z_0$ of electron track and tau seed track must be compatible with coming from the same vertex. Detailed selections are listed in Table I, and the raw acceptance is calculated using the PYTHIA MC and the formula:

$$\alpha_0 = \frac{N_{\text{accepted reqs}}(66:116)}{N_{\text{generated}}(66:116)}.$$  \hfill (7)

where $N_{\text{generated}}(66:116)$ is the total number of $Z/\gamma^* \rightarrow \tau\tau \rightarrow e\nu\tau_h\ell\nu$ events generated in the mass range $66 < M_{\tau\tau} < 116$ GeV/c$^2$ (where $M_{\tau\tau}$ is calculated using MC
generator information) and \(N_{\text{accepted}}^{\text{req}}\) (66:116) is the number of events that pass the acceptance requirements listed in Table I. The acceptance is \(\alpha_0 = 0.0463 \pm 0.0002\), where the uncertainty quoted at this point is purely based on MC statistics.

Then we apply corrections to the raw acceptance, including the correction for tau pairs produced outside the chosen mass window, as discussed in the following section.

**D. Corrections and systematic uncertainties**

We apply corrections to the acceptance and calculate systematic uncertainties to account for both physics and detector simulation effects. This is presented below.

1. **Systematic uncertainties due to parton distribution functions**

In order to estimate the uncertainty due to the parton distribution functions (PDFs), we parametrize acceptance as a function of the Z boson rapidity and use the CTEQ6 methodology [16] to estimate PDF uncertainties. For each of the 20 eigenvectors in the CTEQ6 parametrization, we separately sum up in quadrature positive and negative corrections to the acceptance. If both “positive” and “negative” shifts for any of the eigenvectors result in a positive (or negative) change in the acceptance, we use the one resulting in the larger change as positive (negative) uncertainty and the negative (positive) uncertainty is set to zero for this particular eigenvector. The calculation results in 2.2% variation of the geometrical and kinematic acceptance, which is assigned as a systematic uncertainty.

2. **Correction for \(Z / \gamma^* \rightarrow \tau \tau\) outside the window**

We use the PYTHIA MC to account for the contribution of true \(\tau\tau\) events reconstructed in this analysis but originating outside the mass window 66 < \(M_{\tau\tau}\) < 116 GeV/c\(^2\). We treat these events as a correction to the acceptance, i.e. as an additional signal contribution. The correction is \(f_{\tau\tau}^\text{corr} = 1.055 \pm 0.004\text{(stat)} \pm 0.002\text{(syst)}\) calculated as the ratio of all events passing acceptance cuts (with any generator di-tau mass, \(M_{\tau\tau}\)), to the number of events passing acceptance cuts and having an invariant mass inside the mass window. Of these missed events, roughly half comes from higher invariant masses than the window we use, while the other half is from below. The systematic uncertainty on \(f_{\tau\tau}^\text{corr}\) is estimated by varying CTEQ6 PDFs as described above. The PDF uncertainty on the number of events outside the window (which constitute about 5% of all events) is roughly 4%. As a result, we obtain an uncertainty of 0.2% on the signal acceptance associated with this correction.

3. **Track reconstruction efficiency**

Another source of systematic uncertainty is the accuracy with which the detector simulation predicts the track reconstruction efficiency. Studies of electron track reconstruction efficiency in \(W \rightarrow e\nu\) events reveal no difference between data and MC, with an uncertainty of 0.4% per track. We assign this as a systematic uncertainty due to knowledge of the tracking reconstruction efficiency for the electron track and for the seed track in one-prong taus. The case of three-prong taus is somewhat different, because the additional tracks create higher occupancies that may have a larger effect on pattern recognition.

To assign a systematic uncertainty for the three-prong taus, we use results obtained in the CDF measurement of \(D^\pm\) meson production [17]. The \(D^\pm\) mesons have multiplicity and momentum distributions similar to those of the three-prong taus, making such comparisons valid. In [17] the track reconstruction efficiencies were measured by embedding simulated \(D^\pm\) mesons in data events and varying detector resolutions. The tracking efficiency was measured for a range of the transverse momenta of \(D^\pm\) mesons, including the range of \(p_T > 20\text{ GeV}/c\) most relevant to this analysis. We assign the uncertainty of 3% from [17] as a systematic error in the knowledge of track reconstruction efficiency for events with a three-prong tau. Note that the three-prong taus constitute about a third of the total used in the cross section measurement, and thus the average systematic uncertainty must be weighted accordingly:

\[
\epsilon \approx f_{1\text{-pr}} \times (\epsilon_e + \epsilon_{1\text{-pr}}) + f_{3\text{-pr}} \times \sqrt{(\epsilon_e^2 + \epsilon_{3\text{-pr}}^2)}, \tag{8}
\]

where \(f_{1\text{-pr}}\) and \(f_{3\text{-pr}}\) are the fractions of the final events with one and three-prong hadronic taus (approximately 2/3 and 1/3 as will be shown in Sec. 4), \(\epsilon_{1\text{-pr}}\) and \(\epsilon_{3\text{-pr}}\) are track reconstruction efficiencies for one and three-prong taus, and \(\epsilon_e\) is electron track reconstruction efficiency. Equation (8) accounts for the 100% correlation of the uncertainties in track reconstruction efficiencies for one-prong taus and electrons. The uncertainty for three-prong taus is measured from a control sample of \(D^\pm\) mesons and is therefore uncorrelated with the electron tracking uncertainty. Using Eq. (8), the average weighted uncertainty is estimated as 1.4% (note that this uncertainty is “per event” and not “per track”).

4. **Other systematic uncertainties**

The MC prediction for the efficiency of the \(|z_{0\text{-tik}}^\text{min}| < 60\text{ cm cut} is checked by comparing its efficiency in minimum bias \(p\bar{p}\) collision data to the MC prediction. No deviation is found and the statistical accuracy of this comparison, 0.3%, is assigned as a systematic uncertainty. Uncertainties associated with the electromagnetic calorimeter energy scale calibrations can affect the efficiency of the electron \(E_T > 10\text{ GeV}\) requirement. We estimate the size of this effect by varying the calibration of the CEM absolute energy scale (±0.3%) and adding additional smearing to MC energy resolution (up to 1.5%) to match
observed differences in the Z mass peak in $Z \rightarrow ee$ data. This has a negligible effect.

The full calorimeter system (electromagnetic plus hadronic) energy calibration affects the tau reconstruction in that it requires the seed tower of the tau candidate to pass a threshold of $E_T > 5$ GeV. We estimate the uncertainty in the acceptance by varying the threshold of the seed tower by 3% (a conservative estimate on the knowledge of the calorimeter energy scale [18]). The resulting change in acceptance is 0.2% making the corresponding systematic uncertainty negligible compared to leading effects.

The knowledge of the material for the tuned simulation is accurate to within about 10% as found from the rate of photon conversions and the rate of tridents in $Z \rightarrow ee$ and $W \rightarrow e\nu$ data (tridents are cases where an electron undergoes strong bremsstrahlung followed by a photon conversion leading to three electrons instead of one). The uncertainty due to the imprecise knowledge of the amount of material in the detector is estimated by comparing the default acceptance to that measured with the case where the amount of material in the detector simulation is shifted. The comparison leads to about a 0.4% effect.

We also study the effect of a cut on the maximum number of towers $N_{\text{tow}}$ in a tau cluster, which is set to six. In fact, there is a difference between data and MC simulation that is related to the deficiencies in the simulation of the lateral profile of a hadronic shower in the calorimeter. Showers are wider in data than in simulation, which results in a larger number of towers in a tau candidate. The effect is found by comparing $N_{\text{tow}}$ distributions in $Z \rightarrow \tau\tau$ MC and in the final selected $Z \rightarrow \tau\tau$ events after subtracting backgrounds and is cross-checked with a clean (≈ 90% pure) sample of $W \rightarrow \tau\nu$ events. To estimate the systematic uncertainty due to this requirement, we compare the efficiency of the $N_{\text{tow}} \leq 6$ cut in the default case with the case where additional smearing is introduced to widen the shower profile. The smearing procedure is tuned using a clean sample of $W \rightarrow \tau\nu$ events in data and MC simulation, and a parametrization is obtained allowing variation in the smeared MC predictions for the size of the shower profile. Because the $N_{\text{tow}} \leq 6$ cut is a very loose requirement, even extreme smearing changes the efficiency by no more than 0.4%, and thus the corresponding systematic uncertainty for this requirement is deemed negligible.

Table II shows the combination of correction factors and associated uncertainties.

### VI. Lepton Identification

Lepton identification selections allow for a significant improvement in the purity of the data by suppressing most of the background associated with jets misreconstructed as lepton candidates. In this section, we describe the electron and hadronic tau identification cuts, their efficiencies and associated systematic uncertainties.

#### A. Electron Identification

We use the standard CDF electron identification procedure described in detail in [4]. We require that the electron candidate has a good quality track associated with it and a good agreement between the energy and track momentum, which reduces background contamination from converting photons and random overlaps of a $\pi^0$ and a charged track. We then require that there is a cluster in the CES chamber that is well matched to the projected electron track position. The quality of the matching is quantified using the distance between the shower position measured in the CES and the extrapolated track position in $z$, $|\Delta Z_{\text{CES}}|$, and $x$ directions, $Q_z, Q_x$ (where $Q_z$ is the charge of the electron or positron, and $Q_x$ accounts for the asymmetry due to the bremsstrahlung effect). The CES cluster should have a shape compatible with the expectation of an electron as determined from testbeam studies, which is achieved by requiring a reasonable value of the $\chi^2$ quantity. If there is more than one calorimeter tower contributing to the cluster, the lateral profile of the cluster should agree with an electron hypothesis. We define a variable, $L_{\text{shr}}$, that embeds this information and use it for our electron selection. We finally require that the electron track has a small impact parameter, $d_0$, to reject electrons from $\gamma$-conversions. We estimate the efficiency of the electron

| Table II. Acceptance corrections and systematic uncertainties. Note that all uncertainties quoted are relative to the acceptance computed in Eq. (7). |
|---------------------------------|-----------------|
| Uncorrected Acceptance         | Multiplicative Correction |
| Track Reconstruction           | $\alpha_0 = 0.0463 \pm 0.0002$(stat) |
| $|z_0| < 60 \text{ cm}$          | $1.000 \pm 0.014$ |
| Calorimeter Scale              | $1.000 \pm 0.003$ |
| Mass Window Cut                | $1.055 \pm 0.003$ |
| Material Uncertainty           | $1.000 \pm 0.004$ |
| PDF                            | $1.000 \pm 0.022$ |
| $N_{\text{tow}} \leq 6$       | 1.0 |
| Corrected Acceptance           | $\alpha = 0.0489 \pm 0.0002$(stat) $\pm 0.0013$(syst) |
identification cuts by using MC events that pass the acceptance cuts. The overall efficiency is corrected using scale factors obtained from the data by comparing electron identification efficiencies in a pure sample of $Z \rightarrow ee$ events [4].

We further require that the electron is well isolated, which helps reject jet induced backgrounds. For the isolation cuts, we use well reconstructed COT tracks and reconstructed $\pi^0$ candidates, as described in the previous section, and define two variables:

$$I^{\Delta R}_{\text{trk}} = \sum_{0<\Delta R<0.4} p_{\text{trk}}^{\Delta R}, \quad I^{\Delta R}_{\pi^0} = \sum_{0.22<\Delta R<0.4} p_{\pi^0}^{\Delta R}, \quad (9)$$

which represent scalar sums of the transverse momenta of tracks and $\pi^0$'s reconstructed in the respective regions in $\eta - \phi$ space around the electron direction. The electron track is excluded from the $I^{\Delta R}_{\text{trk}}$ calculation, and the $I^{\Delta R}_{\pi^0}$ calculation does not include $\pi^0$ candidates reconstructed close to the electron track direction. The latter is to avoid unnecessary decrease in efficiency due to cases when CES clusters produced by a genuine electron are split into two subclusters, with the one further from the electron being incorrectly reconstructed as a $\pi^0$ candidate. Identification selections are summarized in Table III with corresponding efficiencies and corrections.

For true electrons, the two main contributions to the inefficiency of the isolation cuts are bremsstrahlung followed by a conversion and overlap with particles coming from the underlying event. One can disentangle these two effects by studying the density of charged tracks per unit of solid angle, $2\pi dN_{\text{trk}}/d\Omega$, as shown in Fig. 4. The plateau in Fig. 4 corresponds to the underlying event contribution and uncorrelated backgrounds (for example from additional interactions in the same bunch crossing), while the peak above the plateau near the electron direction corresponds to secondary tracks from converted photons emitted by the parent electron. Figure 4 also compares the distribution in our data with MC simulation using PYTHIA. Note that the MC prediction for the plateau has been increased by 10% to agree with the data. This indicates an under-estimation of the soft track multiplicity by the MC.

The enhancement in the track density $2\pi dN_{\text{trk}}/d\Omega$, shown in Fig. 4, in the vicinity of the electron ($\cos\theta \approx 1$)

![FIG. 4 (color online). Density of charged tracks $2\pi dN_{\text{trk}}/d\Omega$ per unit of solid angle plotted as a function of the angle $\theta$ with respect to the direction of the electron in a clean sample of $Z \rightarrow ee$ events. Data (points) is compared to MC simulation using PYTHIA (shown as shaded rectangles); the MC prediction has an added constant (10% of the plateau level) to have the plateau level agree with data.](image-url)
is sensitive to the amount of material in our detector. The enhancement directly measures the rate of the secondary tracks due to Bremsstrahlung that contribute to the inefficiency in the isolation cut. The agreement between data and our MC simulation shown in Fig. 4 indicates that any inaccuracies (including underestimation of the uncorrelated backgrounds) have little effect on the isolation efficiency measurement.

With both of these effects under control, we estimate the efficiency of these cuts using our MC simulation. We compare MC predictions of isolation efficiencies measured in data for the \( Z \rightarrow ee \) candidates where we apply the same cut. The ratio of the two is within \( \pm 1.5(1.0)\% \) of unity for track (\( \pi^0 \)) isolation cases. We include these differences as a measure of the associated systematic uncertainty. For isolation selections based on counting neutral pion candidates in the isolation annulus or sum of their transverse energies, we have studied a similar quantity, \( 2\pi dN_{\pi^0}/d\Omega \), around the direction of a tagged electron in \( Z \rightarrow ee \) events. We found it agrees with our MC predictions at a level equivalent to the uncertainty in the isolation cut efficiency of \( \pm 1\% \), indicating that the neutral component of the underlying event is also well simulated in the MC. Note that effects related to Bremsstrahlung affect track isolation only, as \( I^\Delta\xi \) explicitly excludes \( \pi^0 \) candidates close to the electron direction. Therefore, photons emitted by the electron are not counted towards isolation even if they are reconstructed as a separate \( \pi^0 \) candidate as long as the emission angle does not exceed 0.22 rad.

**B. Hadronic tau identification**

Several variables useful for discriminating between tau candidates and backgrounds use track and \( \pi^0 \) candidate information.

The visible mass of a tau candidate, \( M^\tau_{\text{trk}+\pi^0} \), is defined as the invariant mass of the tau momentum four-vector obtained in Eq. (5), before correcting for missed \( \pi^0 \) s. The track mass of a tau candidate, \( M^\tau_{\text{trk}} \), is defined as the invariant mass of the track-only part of the tau momentum four-vector [Eq. (5)]. The charge of a tau candidate is obtained using PYTHIA MC and normalized to a unit area.

Isolation plays an important role in tau identification being the most powerful cut against QCD jet backgrounds. CDF typically uses one of two kinds of track isolation definitions for hadronic tau identification. The first one is the track isolation, \( I^\Delta\theta \), defined as a scalar sum of the transverse momenta of all tracks inside a cone of 30° around the seed track but outside the signal cone, defined by \( \alpha_{\text{trk}} \), in 3D space:

\[
I^\Delta\theta = \sum_{0 < \Delta \theta < 30^\circ} p^\text{trk}_T.
\]  

Note that for a track to be counted in the isolation, we require that the track has \( z_0 \) compatible with that of the seed track: \( |z_0^\text{trk} - z_0^{\text{seed}}| < 5 \text{ cm} \) and at least one stereo and one axial COT segment of at least 6 hits. We also define isolation in \( \eta - \phi \) space:

\[
I^\Delta R = \sum_{0.17 < \Delta R < 0.52} p^\text{trk}_T.
\]  

and use it primarily to confirm the requirements used in the trigger (where 0.17 rad = 10° and 0.52 rad = 30°). A typical cut value is \( I^\Delta\theta/\Delta R < 1 \text{ GeV}/c \). The second type of isolation is based on counting the number of stiff tracks, \( N^\Delta\theta/\Delta R \), in the same isolation cone. To be counted, a track has to have \( p_T > 1 \text{ GeV}/c \) and satisfy the same quality requirements as in the \( \Delta\theta/\Delta R \) cases. This definition has slightly lower background suppression power, but higher efficiency and less dependence on instantaneous luminos-


<table>
<thead>
<tr>
<th>Cut</th>
<th>MC Efficiency [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seed Track Quality $3 \times 2 \times 5$(hits)</td>
<td>99.71 ± 0.03</td>
</tr>
<tr>
<td>$d_{0}^{s\text{-seed}} &lt; 0.2$ cm</td>
<td>97.65 ± 0.09</td>
</tr>
<tr>
<td>$\xi &gt; 0.1$</td>
<td></td>
</tr>
<tr>
<td>$M_{\text{track}}^\text{peak} \leq 1.8$ GeV/c^2 &amp; $M_{\text{track}+\pi^{0}} \leq 2.5$ GeV/c^2</td>
<td></td>
</tr>
<tr>
<td>$N_{\text{trk}}^{\Delta \theta} = 0 &amp; N_{\text{trk}}^{\Delta R} = 0$</td>
<td></td>
</tr>
<tr>
<td>$I_{\text{trk}}^{\Delta \theta} \leq 0.6$ GeV/c</td>
<td></td>
</tr>
<tr>
<td>$N_{\text{cone}}^{\text{trk}} = 1$ or $3$</td>
<td></td>
</tr>
<tr>
<td>Final Efficiency</td>
<td>60.47 ± 0.28(stat) ± 1.8(syst)</td>
</tr>
</tbody>
</table>

TABLE IV. Efficiency of tau ID cuts. The simulation predicted efficiency is used as a default, and the systematic uncertainty is derived from comparisons with data. Only non-negligible systematic uncertainties are quoted.

ity. In this analysis, we choose this second kind of isolation and require $N_{\text{trk}}^{\Delta \theta/\Delta R} = 0$.

In a similar way, we define the tau candidate $\pi^{0}$ isolation as

$$I_{\pi^{0}}^{\Delta \theta} = \sum_{\alpha, \phi < \Delta \theta < \phi'} p_{T}^{\pi^{0}}. \quad (14)$$

Only $\pi^{0}$ candidates that have matches in both wire and strip CES views are used in the isolation.

In choosing tau identification selections, we are careful to avoid cutting on variables that are not well described by the simulation. Problematic variables are related to the modeling of the lateral hadronic shower profiles in the calorimeter, where data suggest that showers are on average wider than predicted by simulation. A full list of tau identification cuts used in this analysis and corresponding efficiencies can be found in Table IV.

Inefficiencies in the tau selection are mainly due to isolation requirements. Because the inner cone of the isolation annulus is large, the inefficiency is dominated by the underlying event tracks and occasional nearby jets. The contribution from the decay products of a hadronically-decaying tau lepton occasionally falling into the isolation annulus is negligible compared to the level of accuracy of this analysis. We verify this by studying the tau isolation efficiencies applied to electron and muon candidates in $Z \rightarrow ee/\mu\mu$ MC samples, which yield the same results as for hadronic tau candidates from $Z \rightarrow \tau\tau$. Figure 6 shows the density $2\pi dN_{\text{trk}}/d\Omega$ of charged tracks with $p_{T} > 1$ GeV/c per unit of solid angle as a function of $\cos \theta$ with respect to the direction of the lepton. We compare three samples: (i) a clean sample of Z→ee data events; (ii) a sample of Z→ee simulated events; and (iii) a sample of Z→ττ simulated events. The good agreement between the first two shows that the efficiency of the isolation selection based on track counting, e.g. $N_{\text{trk}}^{\Delta \theta/\Delta R}$, are well reproduced by the simulation. The comparison of the two simulated distributions demonstrates that the efficiency of such requirements in the Z→ee and Z→ττ samples is the same as long as the region immediately surrounding the lepton is excluded from the calculation. That means that the efficiency of $N_{\text{trk}}^{\Delta \theta/\Delta R} = 0$ requirement is the same in these two samples (for reference, $\cos(10^{0}) = 0.985$ and $\cos(30^{0}) = 0.866$). As a cross-check, we compare the efficiencies of the isolation cuts in the $W \rightarrow e\nu$ and $W \rightarrow \tau\nu$ samples, between MC and data and find a good agreement as well. With this knowledge, the corresponding efficiencies are measured directly from the data using $Z \rightarrow ee$ candidates in a fashion similar to that used for the isolation requirements in the electron
identification. Following this method, we find no disagreement between the data and MC predictions at the level of the precision of these comparisons, which is \( \approx 1.5\% \) (\( \approx 1.0\% \) for \( \pi^0 \)-related isolations). Note that here no scaling is necessary to match the plateau at low \( \cos \theta \) indicating that the lower multiplicity of tracks from the uncorrelated backgrounds observed in Fig. 4 is due to tracks having \( p_T < 1 \text{ GeV}/c \).

Possible systematic biases to the efficiency of the \( \xi \)-cut can be caused by improper simulation of the hadronic calorimeter energy response to charged tracks. To study these effects, we select a sample of isolated charged pions by filtering jet events with exactly one one-prong tau candidate passing all identification criteria. We ensure that these events have no other leptons in order to exclude contamination by electrons from \( Z/\gamma^* \rightarrow ee/\tau\tau \) and also require no additional showers detected in the CES to eliminate photons and neutral pions that may deposit additional energy in the hadronic calorimeter. Using this sample, we compare the efficiencies of the \( \xi > \xi_0 \) cut for several values of \( \xi_0 \) in the data and MC simulation. The comparison shows a good agreement within the statistical precision of \( \approx 2\% \), which is taken as an estimate of the systematic uncertainty for this cut.

Efficiencies related to the seed track quality requirements and the cut on \( |z_{0\text{seed}} - z_{0\text{trk}}| \) are found to be well described in simulation by comparing efficiencies of these cuts applied to \( Z \rightarrow ee/\mu\mu \) data. Inefficiency due to the cut on the impact parameter, \( d_0 \), of the tau seed track is found to be caused by two main effects. The first is the \( d_0 \) resolution in track reconstruction, which is well simulated in MC as known from studies of the \( d_0 \) cut in \( Z \rightarrow ee/\mu\mu \) data and simulation. The second contribution comes from nuclear interactions of charged hadrons in the detector material that is also well simulated. As a cross-check, we repeat the cross section measurement presented in this paper without the \( d_0 \) cut and find that the two are in good agreement indicating that efficiency of this cut is properly simulated in MC simulation.

Efficiencies of the \( M^{\text{track}+\pi^0} \) and \( M^{\text{track}} \) cuts are compared in data and simulation using a clean sample of \( W \rightarrow \tau\nu \) events and are found to agree within the statistical uncertainties which are assigned as a conservative estimate of the systematic error for these cuts.

Finally, the adequacy of the MC in simulating the efficiency of the \( N^{\text{cone}}_{\text{trk}} \) cut depends on the accuracy of the underlying event simulation. To verify that the probability of finding a track with \( p_T > 1 \text{ GeV}/c \) from the underlying event or from another collision event in the same bunch crossing is correct we measure the density of such tracks, \( dN_{\text{trk}}/d\Omega \), in a clean sample of \( Z \rightarrow ee \) events and compare it to the MC predictions (see Fig. 6). While no disagreement is found, scaling of the density in simulation to the one in data changes the efficiency of the cut by 0.3%, which we assign as a systematic uncertainty.

VII. TRIGGER EFFICIENCY

Trigger efficiencies for the lepton plus track trigger are measured for each of the lepton and track triggers (legs) separately using data taken over the same running period as the one used in this measurement.

We use a photon conversion sample to measure the efficiency for the electron leg as a function of the electron \( E_T \) and the associated track \( p_T \). The plateau efficiency for the electron leg is found to be approximately 97% and is primarily determined by the XFT track finding efficiency. This measurement is cross-checked using events tagged as \( Z \rightarrow ee \).

For the track leg, the measurement of the trigger efficiency is made assuming this leg is a hadronic tau. The efficiency is parametrized as a function of several tau variables. In the course of measuring these efficiencies, we verify that the track trigger efficiency is independent of the electron trigger efficiency. We compare the efficiency for events with a single tau candidate and for events with a tau candidate and an additional electron. At level 2 (where an XFT track of \( p_T > 5 \text{ GeV}/c \) is applied), the average efficiency for the track \( p_T \) above 10 GeV/c is approximately 97%, and slowly grows further at higher \( p_T \). To understand trigger efficiency for hadronic taus at level 3, we use a sample of jets passing loose tau identification and carefully study properties of the tau candidates failing the trigger. We find that efficiency for the hadronic taus passing ID selection is generally very high; a small degradation in efficiency (at the level of a \( \approx 0.5\% \)) is primarily due to differences in tracking between offline and level 3, e.g., a track reconstructed offline close but inside the tau signal cone boundary can occasionally be mismeasured by level 3 as falling into the isolation region and vice versa. We defined a variable measuring the distance from the boundary to the track closest to the boundary and use it in the parametrization to account properly for this effect. Another example is that the momentum of a track can occasionally be mismeasured at level 3 to exceed the \( p_T = 1.5 \text{ GeV}/c \) threshold in the isolation cone and veto the event. This second effect is predominantly correlated with the occupancy of soft tracks in the isolation cone, which is used to parametrize the efficiency function. To ensure that there are no unaccounted for correlations in the parameters of the efficiency function, we use a different sample of jet data to compare the performance of the parametrized efficiency with the actual level 3 trigger efficiency obtained by running the level 3 package on the same data. We find excellent agreement indicating that the efficiency parametrization accurately describes trigger behavior. As a cross-check, we repeat the same exercise using a sample of clean \( W \rightarrow \tau\nu \) events and again obtain excellent agreement.

We calculate the average trigger efficiency by convoluting the pythia MC predictions with parametrized trigger
measures the separation of the two tracks in the \( r \) plane, while the second variable \( S_{xy} \), the distance between the 2D track trajectories at the point of their closest approach to each other, controls their separation in the \( r - \phi \) plane. There is an exclusion to this selection criterion to address “trident” electrons, which are cases when a genuine electron emits a photon that converts into a pair of electrons leading to a system of three close tracks. To avoid cutting out “tridents,” we keep an event if a track associated with the electron candidate is tagged as a conversion partner to more than one other track. Conversion removal cuts have only a small effect on the signal, while removing most of the \( \gamma + \) jet background and a sizable fraction of the multijet background.

Next, we remove events consistent with the topology of \( Z \rightarrow ee \) candidates using the reconstructed mass of the electron and any additional loose electron candidates (calorimeter-based removal) or tracks (track based removal) in the event. In the case of the calorimeter-based removal, we identify all additional electron candidates in the event satisfying “second cluster” requirements in Table VII. If the invariant mass of the leading electron and any additional “second cluster” falls into the chosen mass range, an event is excluded from further analysis. Similarly, for the track based removal, we calculate the invariant masses of two track systems of the leading electron track and any other track in the event satisfying the isolation requirement (see Table VII). If the calculated mass for any such pair falls into the chosen mass range, the whole event is tagged as a \( Z \rightarrow ee \) candidate and not considered in further analysis. The window for the track based removal is wider to accommodate for possible bremsstrahlung effects, and both mass ranges are chosen to preserve high signal efficiency while rejecting most of the \( Z \rightarrow ee \) background events.

It will be shown in the remainder of this section that for the remaining events the best discriminators between signal and the dominant background, QCD jet production, rely on the measurement of missing transverse energy, \( \slashed{E_T} \), and the \( \slashed{E_T} \) direction with respect to other objects (leptons or jets) in the event. It is therefore important to measure \( \slashed{E_T} \) well.

### A. Corrections to missing transverse energy calculation

The raw missing transverse energy is defined as a 2D vector and calculated using the calorimeter tower information:

\[
\slashed{E} = \left( \begin{array}{c} -\sum_i E_i \sin \theta_i \cos \phi_i \\ -\sum_i E_i \sin \theta_i \sin \phi_i \end{array} \right),
\]

where \( E_i \) is the energy of the \( i \)th tower, \( \theta_i, \phi_i \) are the polar and azimuthal angles of the center of the tower with respect to the measured interaction position, and the sum is taken over all calorimeter towers.

Further improvement in the resolution is obtained by applying corrections for jets, electrons and hadronic taus. Unlike the symmetric underlying event contribution, each of these has a large contribution to the total energy sum and a well defined direction. Therefore biases on these quantities have a large effect on the magnitude and direction of

### VIII. EVENT TOPOLOGY CUTS

Additional selections based on the kinematics of the signal events and the dominant background contributions allow strong background suppression with a moderate decrease in the efficiency of signal events selection thus improving precision of this measurement.

First, we require that the two leptons have opposite charges (OS). This cut alone diminishes the background contribution from QCD jet production by a factor of 2 with essentially no effect on the signal yield. Second, we remove events where the leading electron candidate is tagged as a conversion. This is achieved by a requirement that there be at least one good electron candidate where the track associated with the candidate is not a part of a conversion pair, defined as two opposite charged tracks satisfying requirements in Table VI. The \( \Delta \cot \theta \) variable measures the separation of the two tracks in the \( r - z \) plane, while the second variable \( \Delta S_{xy} \), the distance between the 2D track trajectories at the point of their closest approach to each other, controls their separation in the \( r - \phi \) plane. There is an exclusion to this selection criterion to address “trident” electrons, which are cases when a genuine electron emits a photon that converts into a pair of electrons leading to a system of three close tracks. To avoid cutting out “tridents,” we keep an event if a track associated with the electron candidate is tagged as a conversion partner to more than one other track. Conversion removal cuts have only a small effect on the signal, while removing most of the \( \gamma + \) jet background and a sizable fraction of the multijet background.

Next, we remove events consistent with the topology of \( Z \rightarrow ee \) candidates using the reconstructed mass of the electron and any additional loose electron candidates (calorimeter-based removal) or tracks (track based removal) in the event. In the case of the calorimeter-based removal, we identify all additional electron candidates in the event satisfying “second cluster” requirements in Table VII. If the invariant mass of the leading electron and any additional “second cluster” falls into the chosen mass range, an event is excluded from further analysis. Similarly, for the track based removal, we calculate the invariant masses of two track systems of the leading electron track and any other track in the event satisfying the isolation requirement (see Table VII). If the calculated mass for any such pair falls into the chosen mass range, the whole event is tagged as a \( Z \rightarrow ee \) candidate and not considered in further analysis. The window for the track based removal is wider to accommodate for possible bremsstrahlung effects, and both mass ranges are chosen to preserve high signal efficiency while rejecting most of the \( Z \rightarrow ee \) background events.

It will be shown in the remainder of this section that for the remaining events the best discriminators between signal and the dominant background, QCD jet production, rely on the measurement of missing transverse energy, \( \slashed{E_T} \), and the \( \slashed{E_T} \) direction with respect to other objects (leptons or jets) in the event. It is therefore important to measure \( \slashed{E_T} \) well.

### A. Corrections to missing transverse energy calculation

The raw missing transverse energy is defined as a 2D vector and calculated using the calorimeter tower information:

\[
\slashed{E} = \left( \begin{array}{c} -\sum_i E_i \sin \theta_i \cos \phi_i \\ -\sum_i E_i \sin \theta_i \sin \phi_i \end{array} \right),
\]

where \( E_i \) is the energy of the \( i \)th tower, \( \theta_i, \phi_i \) are the polar and azimuthal angles of the center of the tower with respect to the measured interaction position, and the sum is taken over all calorimeter towers.

Further improvement in the resolution is obtained by applying corrections for jets, electrons and hadronic taus. Unlike the symmetric underlying event contribution, each of these has a large contribution to the total energy sum and a well defined direction. Therefore biases on these quantities have a large effect on the magnitude and direction of
the missing transverse energy. Jet energy corrections [18] are obtained from dijet balancing (cross-checked using photon-jet events) and on average compensate for unequal energy responses for neutral and charged hadrons and for geometrical instrumental deficiencies, such as energy losses in the spaces between the towers and nonlinearity in the calorimeter response. In this measurement, jets are defined using an iterative jet clustering algorithm [19] with $\Delta R = 0.4$. Jet corrections are applied to jets with $p_T^{\text{raw}} > 10$ GeV and $|\eta| < 2.5$. If a particular jet has corrected energy $p_T^{\text{cor}} > 15$ GeV, the missing transverse energy is corrected for this jet as follows:

$$\vec{E}_T^\prime = \vec{E}_T - (\vec{p}_T^{\text{cor}} - \vec{p}_T^{\text{raw}}),$$  \hspace{1cm} (16)

where $\vec{E}_T$ is the corrected value of missing transverse energy and $\vec{E}_T$ is the value before correction.

The correction for electrons compensates for the varying response of the calorimeter depending on the position of the electron within the tower, as the response is lower for an electron passing close to the edge of the tower [4]. The correction for hadronic tau candidates is similar to the jet correction in this case, the best measurement of the energy is obtained from the track and $\pi^0$ information with appropriate corrections described in Sec. V B. Correction to the missing transverse energy is applied as follows:

$$\vec{E}_T^\prime = \vec{E}_T - (\vec{p}_T^{\text{cor}} - \vec{p}_T^{\text{cal}})$$  \hspace{1cm} (17)

where $\vec{p}_T^{\text{cal}}$ is calculated as the vector sum of momenta of the towers assigned to the tau cluster and all adjacent towers. This last addition is necessary to compensate for known deficiencies in the simulation of the lateral shower that lead to wider shower profiles in the data than in the simulation and are discussed further in this section. Figure 7 compares the resolution of the missing transverse energy measurement prior to correction and after all corrections for jets, electrons and hadronic taus. Note that we do not correct $E_T$ for possible muon candidates as the target decays or false muons produced as a result of misidentifying jets as muons or misreconstructed tracks. Correcting for such muon candidates would lead to additional systematic effects stemming from the reliance on the simulation in describing these nontrivial effects.

**B. Optimization of the event topology cuts**

To suppress $W +$ jet and QCD backgrounds, we define two additional variables. The first is the transverse mass, $M_T$, of the electron and the missing transverse energy, defined as

$$M_T(e, \vec{E}_T) \equiv \sqrt{2 \times p_T^e \vec{E}_T \times (1 - \cos\Delta\Theta)},$$  \hspace{1cm} (18)

where $\Delta \Theta$ is the 2D angle in the $r - \phi$ plane between the electron track and the missing $E_T$ direction. The second makes use of the transverse momentum of the electron and the missing energy:

$$p_T^e(e, \vec{E}_T) \equiv |\vec{p}_T^e + \vec{E}_T|$$  \hspace{1cm} (19)

To show the effectiveness of these variables Fig. 8(a) shows a distribution of $M_T$ vs $p_T^e$ for signal $Z \to \tau^+ \tau^-$ events, as predicted by the PYTHIA MC, after the acceptance and identification requirements for electron and tau candidates. Figure 8(b) shows a similar distribution for a sample where the electron and tau candidates have the same sign of charge (“like-sign” or LS). It is dominated by backgrounds. The optimal choice of cuts is shown in Fig. 8. The optimization is done by maximizing the significance defined as $S/\sqrt{S + B}$. In the process of optimization, we vary the cut on $M_T$ (the horizontal line in the figures) and the position of the kink in the $M_T$ and $p_T$ dependent cut (the break point in the other line). This results in the optimization of three parameters. The above definition of significance assumes negligible systematics in determination of $B$. For comparison, we used a different definition of the significance, $S/\sqrt{S + 2B}$, which corresponds to a scenario in which the systematic uncertainty in the background estimation is driven by the statistical uncertainty, $\sigma_{\text{syst}} = \sqrt{B}$, and find that the optimal choice of the cuts remains nearly unchanged.

To model backgrounds, we use the LS data combined with the OS component of the $W +$ jets contribution, as predicted by MC that is scaled to fit the CDF data. LS data
are taken after applying all the other cuts used in the analysis and give a good first order prediction for the QCD induced jet backgrounds. The scaling for $W + \text{jets}$ is required to make PYTHIA MC predictions agree with the data ($f^W \approx 0.73$ as shown later in the paper). The scaling factor is not unity because the PYTHIA MC does not correctly predict the absolute probability of a jet to be misidentified as a hadronic tau. This is not surprising as PYTHIA was never tuned to simulate properly the tails of QCD induced jet backgrounds. The area bounded by the solid lines and including the right bottom quadrant on the plots indicates the signal region.

FIG. 8 (color online). Distribution of $M_T$ vs $p_T$ for: (a) signal $Z \rightarrow \tau\tau$ events obtained using the PYTHIA MC simulation; (b) like sign (LS) data after all cuts. The latter sample is dominated by QCD and $\gamma + \text{jets}$ backgrounds. The area bounded by the solid lines and including the right bottom quadrant on the plots indicates the signal region.

Corrections are found to be small, as expected, as our signal candidates rarely have any jets at all, and their contribution to the correction is negligible. Similarly, we turn on and off the correction for electrons, both in the data and in the simulations. There the change in efficiency was also negligible.

In defining the correction to $E_T$, we are careful to avoid the effect related to the problem of the MC predicting a more narrow lateral profile for taus than indicated by the checks in the data. This problem is traced to the simulation of showers in the hadron calorimeter. A naive correction due to taus could be applied by correcting $E_T$ for the difference between the tau $p_T$ measurement based on tracks and $\pi^0$'s (which should be accurate), and the calorimeter measured tau cluster energy (towers with $E_T > 1$ GeV are clustered). However, wider shower profiles generated by tau candidates in the hadronic calorimeter in the data compared to the simulation lead to a larger amount of unaccounted unclustered energy in the data than in the simulation. This leads to an average effect of “pulling” the $E_T$ into the direction opposite the tau candidate (thus, most of time in the direction of the electron) more often in the data than in the simulation, which in turn leads

C. Efficiency and systematic uncertainties

The efficiency of the conversion removal and the OS requirement for the signal is estimated using MC events. Related systematic effects stem from the knowledge of the amount of the material in the detector and are well understood. The main mechanism for charge misidentification is strong bremsstrahlung off an electron followed by a conversion producing a stiff secondary track misidentified as the original electron track. The systematic uncertainty in the conversion removal is related to the rate of conversions in the data and in the MC simulation. We confirm the MC prediction by comparing the efficiencies of these cuts in the $Z \rightarrow e\bar{e}$ data and in the simulation, and assign the statistical uncertainty of the comparisons as a conservative measure of systematic uncertainty. The results are shown in Table VIII. The efficiency of the $Z \rightarrow e\bar{e}$ removal for the signal is also estimated using MC events and is shown in Table VIII. The systematic uncertainty of 0.2% is estimated using MC simulation by varying the definition of the $Z$ mass window veto to account for track momentum, calorimeter energy resolution and the relative energy deposition in hadronic and electromagnetic calorimeters.

The efficiency of the event topology cuts related to $E_T$ is estimated using MC events. We vary the effect of the $E_T$ corrections for jets, taus and electrons. Effects related to jet

TABLE VIII. Efficiency of event selection criteria. The MC prediction is chosen as the default. The systematic uncertainty for the $M_T$ cut is included in the calculation of the systematic uncertainty for the 2-dimensional cut in the $p_T$ versus $M_T$ plane.

<table>
<thead>
<tr>
<th>Cut</th>
<th>Efficiency [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Q_e \times Q_T = -1$</td>
<td>99.27 ± 0.06 ± 0.2</td>
</tr>
<tr>
<td>Conversion Removal</td>
<td>99.10 ± 0.07 ± 0.2</td>
</tr>
<tr>
<td>Z Removal</td>
<td>97.33 ± 0.12 ± 0.2</td>
</tr>
<tr>
<td>$p_T &gt; 24\text{ GeV}/c$ or</td>
<td></td>
</tr>
<tr>
<td>$M_T &gt; 50 - 1.25 \times p_T \text{ (GeV}/c)$</td>
<td>81.9 ± 0.3 ± 0.1</td>
</tr>
<tr>
<td>$M_T(e, E_T) &lt; 50\text{ GeV}/c^2$</td>
<td>96.9 ± 0.2*</td>
</tr>
<tr>
<td>Final Efficiency</td>
<td>76.0 ± 0.3(stat) ± 0.3(syst)</td>
</tr>
</tbody>
</table>
to higher efficiency for passing event topology cuts in the data events than in the MC simulation.

To remove the effect of this imperfection on the measurement, we calculate the raw calorimeter tau energy using not only the towers assigned to the tau cluster, but also all towers immediately adjacent to the tau cluster. In this case, all the leakage not predicted by MC simulation is absorbed into the tau calorimeter energy, and the redefined tau energy calculation becomes well predicted by simulation.

Possible systematic effects in the estimation of the event topology selection efficiency come from the prediction of the absolute value and direction of the missing transverse energy in the simulation. In \( Z \rightarrow \tau \tau \) events the missing energy is a vector sum of the neutrino momenta and contributions from mismeasurements in the detector (“instrumental” effects). Systematic effects from the former are negligible, while instrumental effects can be appreciable.

To understand the instrumental component, we study \( \gamma + \text{jet(s)} \) events in data with a tagged conversion photon and a jet passing all tau identification requirements. This sample is topologically similar to the \( Z \rightarrow \tau \tau \) sample in terms of the mutual orientation of the electron (photon) and tau (jet reconstructed as a tau) and the instrumental corrections to the \( E_T \). Note that in \( \gamma + \text{jet(s)} \) events the \( E_T \) has only instrumental contributions that can be studied directly. Further, in both samples the instrumental \( E_T \) has two distinct components: one parallel to the direction of electron (or converting photon), \( E_T^\parallel \), and one perpendicular to it, \( E_T^\perp \). \( E_T^\parallel \) is mostly sensitive to the \( E_T \) corrections for electron and tau candidates, while \( E_T^\perp \) primarily depends on the underlying event and additional jets in the event. \( E_T^\perp \) has an average of zero by construction, while \( E_T^\parallel \) is expected to have an average offset due to the subtraction of the underlying event contribution to the measured energy of electron and tau candidates. We compare the offsets and resolution for the \( E_T^\parallel \) and \( E_T^\perp \) in data and simulation and find them to be very similar. To estimate the uncertainty we adjust the simulation so that it matches the data. We then use the adjustment factors to recalculate the \( E_T \) in the \( Z \rightarrow \tau \tau \) MC events as a 2D vector on an event-by-event basis. To do this we combine the adjusted instrumental contribution with the contribution from the neutrinos. The corresponding efficiency of the event topology cuts related to the \( E_T \) selection is recalculated and compared with the efficiency computed using the uncorrected simulation. The difference is less than 0.1%.

To gauge the importance of properly correcting \( E_T \) for taus, we compare the effect on the final cross section of applying the nonoptimized definition of the correction, where the raw tau energy is calculated using towers assigned to the tau cluster, both in data and in simulation. We obtain a 2.4% difference in the measured cross section. Although the statistical uncertainty of each measurement is 8%, there is a large correlation between the two. We conclude this effect could have been a dominant systematic in the event topology systematics if we had not corrected for it.

### IX. ESTIMATION OF BACKGROUND CONTRIBUTIONS

In addition to the \( Z \rightarrow \tau \tau \) signal yield, the selected sample has several background contributions that are not completely removed by the topological cuts. The dominant background contributions at this stage are expected to be due to (i) QCD jet production, (ii) \( \gamma + \text{jet} \) events, and (iii) \( Z \rightarrow ee \) and \( W + \text{jet} \) events. A small contribution is expected from \( t\bar{t} \) and diboson production. In the following, we describe these backgrounds and the methods used to estimate their contribution.

#### A. \( t\bar{t}, \text{diboson and} \ Z \rightarrow ee \) backgrounds

The \( t\bar{t} \) and diboson events constitute an irreducible physics background in the sense that they can satisfy our selection criteria via decay modes with real electrons and hadronic taus. These are expected to be well described by the MC and are found to be small, at the level of a few events compared to over 500 events in the signal sample.

\( Z/\gamma^* \rightarrow ee \) events enter the sample in two ways: (i) one of the electrons passes the hadronic tau requirements by leaving a substantial deposition in the hadron calorimeter, and the reconstructed mass of the pair is outside of the \( Z \) mass window veto cut, or (ii) the recoil jet in \( Z/\gamma^* \rightarrow ee \) is misidentified as a hadronic tau while the event passes the \( Z \rightarrow ee \) candidate removal either because one of the electrons is not reconstructed (e.g., it falls outside the detector coverage) or the invariant mass of the two electron candidates falls outside the \( Z \)-veto mass window.

The main mechanism responsible for electron misidentification as a tau is strong bremsstrahlung. If an electron emits a hard photon, the track associated with the electron has lower momentum, so that the tau candidate associated with the electron passes the \( \xi = \nu^{\text{track}}/\nu^{\text{hadronic}} > 0.1 \) cut. If that occurs, even \( Z/\gamma^* \rightarrow ee \) events originating near the \( Z \)-peak may evade the \( Z \)-veto if the photon emitted by an electron and the electron track deposit energy in adjacent calorimeter towers in \( \phi \). Since the clustering algorithm does not combine adjacent towers in \( \phi \), the affected electron energy is underestimated lowering the reconstructed pair mass. A second mechanism is due to the rare case where an electron deposits a relatively large portion of its energy in the hadron calorimeter. The overall probability of an electron to be misidentified as a tau is relatively well simulated in the MC, as known by measuring the average probability for an electron from \( Z \rightarrow ee \) events to pass the \( \xi \) cut in data and in simulation, which we found to agree to within \( \approx 20\% \). We assign a systematic uncertainty of 20% to these fractions of the \( Z \rightarrow ee \) background, which leads to about 2% uncertainty on our measured cross section. As a cross-check, we modify the \( Z \rightarrow ee \) removal veto re-
requirements, and verify that the change in the event count is consistent with the MC predictions.

A third mechanism, where a jet is misidentified as a hadronic tau candidate, is less important compared to the dominant ones and in the procedure we use is completely accounted for in the background estimation in the $\gamma +\text{jet(s)}$ category (see Sec. IX.D).

**B. $W +\text{jets}$**

The $W +\text{jets}$ background finds its way into our sample when a jet is misidentified as a tau candidate. This background is strongly suppressed by the event topology cuts that require the remaining $W +\text{jets}$ events to have a large transverse momentum $W$ boson ($\approx p_T(e, p_T^\ell)$) and low $M_T$, thus diminishing this contribution. In general, one does not expect a very good agreement between the data and the MC prediction for the $W +\text{jets}$ background in this analysis. The agreement is determined by how well the MC predicts the probability for a jet to be identified as a tau candidate, which relies on the details of the jet fragmentation in a relatively small fraction of the phase space. To estimate the contribution of this background, we introduce a scale factor, $f^W$, to compensate for inaccuracies in the MC predictions for the probability of a jet to be identified as a tau. We then extract the numerical value of $f^W$ by comparing the data and the MC based predictions in the high $M_T$ region which is dominated by $W +\text{jets}$ events. As will be shown later in Sec. X, we find $f^W \simeq 0.73 \pm 0.07$.

**C. Backgrounds due to the QCD jet production**

Light-quark QCD backgrounds get into the sample when one jet is misidentified as an electron (e.g., conversions), while the other jet (or a part of it) passes as a tau candidate. Typically, the track multiplicity for false tau candidates from jets peaks in the two-prong bin for the range of jet $E_T$ and $p_T^\ell$ characteristic of this analysis. Heavy flavor QCD backgrounds have two important features that distinguish the way they populate the $Z \rightarrow \tau\tau$ data set. First, they have “real” electrons from the semileptonic decays of heavy flavor quarks, and, second, false “tau” from heavy flavor jets have a different multiplicity distribution than in the case of lighter quarks. This difference is likely due to the decay modes of $B$ mesons. Some of the “heavy” mesons are ready candidates for false taus (the $D^\ast$ meson has mass of 1.9 GeV/c$^2$ compared to $m_\tau = 1.8$ GeV/c$^2$ and decays into three pions). In this study, we find an important feature common to the QCD backgrounds. If one plots track isolation, $I_{\text{trk}}^{\Delta R<0.4}$ defined in Sec. VIA, to the first order the distribution is uniform for false electrons inside jets. We can estimate the number of events in the signal region ($I_{\text{trk}}^{\Delta R<0.4} < 1$ GeV) by using events in the “sideband” region $2 < I_{\text{trk}}^{\Delta R<0.4} < 8$ GeV:

$$N_{\text{QCD}}^{(0;1)} = r \times N_{(2;8)}, \quad r = 1/6. \quad (20)$$

This technique works well not only for event counting, but for a variety of reasonable kinematic distributions. One can predict the distribution of a variable from the QCD backgrounds in the signal region by measuring the shape of the distribution for events with nonisolated electrons and normalizing according to Eq. (20). As an illustration, Fig. 9(a) shows the distribution of $I_{\text{trk}}^{\Delta R}$, the electron track isolation, for events with one and three-prong tau candidates passing all other selections. The shaded histogram shows the distribution for the LS events dominated by QCD backgrounds and, although there is a significant contamination of $\gamma +\text{jets}$ background in this sample (centrateced at low isolation values), one can still see the nearly flat shape of the isolation distribution for LS events for $I_{\text{trk}}^{\Delta R<0.4} > 1$ GeV/c. To disentangle the QCD backgrounds from the $\gamma +\text{jets}$ contribution, in Fig. 9(b) we

![FIG. 9](color online). (a) Distribution of the electron track isolation, $I_{\text{trk}}^{\Delta R<0.4}$ for OS (points) and LS (histogram) $Z \rightarrow \tau\tau$ candidates after all analysis cuts except the track isolation cut, which is dropped. (b) The same distribution, but with additional cuts, $M_T < 10$ GeV/c$^2$, $10 < E_T < 20$ GeV that suppress $W +\text{jets}$ and $\gamma +\text{jets}$ backgrounds. LS events in this case are a nearly pure sample of QCD jet events.
show the same distribution, but for events passing the additional cuts aimed at suppressing $\gamma +$ jets background ($M_T < 10$ GeV/$c^2$, $10 < E^*_T < 20$ GeV). The histogram for LS events in this sample is now nearly pure QCD backgrounds and one can observe a flat shape of the distribution in the full range of $I_{\text{trk}}^{A,R=0.4}$ proving our initial assumption of uniformity of the QCD background shape. The latter distribution yields $r = 0.17 \pm 0.03$ consistent with $1/6$.

**D. $\gamma +$ jet**

We find a sizable background contribution from $p\bar{p} \rightarrow \gamma +$ jet(s) production. The main mechanism is via a photon conversion leading to an electron, while the jet is misidentified as a hadronic tau. Taking into account the explicit removal of recognized conversion electrons and that most of the events pass the high $p_T(e, E_T) > 24$ GeV/$c$ cut, these events typically have conversion electrons with very asymmetric transverse momenta and a relatively low measured $E_T$ due to a good internal momentum balance in these events. In addition, the $\gamma +$ jet(s) events, unlike QCD jet events, typically have an isolated electron. The reason for this is that the conversion removal fails to find events with no visible partner track and, unlike the QCD di-jets, the photon is naturally isolated having few surrounding tracks to boost the value of $I_{\text{trk}}^e$. The $\gamma +$ jet process is the dominant contribution to the backgrounds with an isolated lepton. Note that this background is specific to the $Z/\gamma^* \rightarrow \tau\tau \rightarrow e\nu\nu\tau_h\nu$ mode and is not present in $Z/\gamma^* \rightarrow \nu\nu\tau_h\nu$.

An important feature of the $\gamma +$ jet process is that it is **charge blind**. The number of events in OS data is, on average, the same as in the LS data. To verify this, we use events with an electron tagged as a conversion, but otherwise passing all selection requirements, and measure the charge asymmetry using the leading leg in the conversion pair and the tau candidate charges. The charge **blindness** allows us to estimate the number of $\gamma +$ jets background events in the OS signal by measuring the excess of events with an isolated electron above the flat (QCD) background in the LS data, after subtracting the contamination from $W +$ jets, $Z \rightarrow ee$ and signal.

The shape of the $\gamma +$ jets background is estimated from the data using events with a converted $\gamma +$ tau candidate. We start by requiring that an event has an electron that is tagged as one leg of a conversion pair. We then reconstruct the photon momentum as the sum of the four-vectors of the two conversion electrons. From this point on, we treat the photon as if it were an electron and apply all analysis cuts. The total normalization, $f^\gamma$, can then be obtained from the excess of LS events. As in the $Z \rightarrow ee$ and $W +$ jets case, the overall measurement is not sensitive to the predictions for $\gamma +$ jet(s) events in the nonisolated region, because any discrepancy is absorbed in the QCD estimation by the signal extraction procedure.

**X. SIGNAL EXTRACTION METHOD**

To determine the number of signal and background events, we remove the $M_T(e, E_T)$ and track isolation cuts and drop the OS charge requirement on electron and tau candidates. Then, for OS and LS events separately, we define four regions in the $I_{\text{trk}}^e$ vs $M_T(e, E_T)$ plane, denoted as A-I/II and B-I/II, as defined in Table IX.

With these definitions, the signal candidates are expected to occupy region A-I while the $W +$ jet backgrounds dominate in region B-I. Note that this separation is not exact. Each region has non-negligible contributions from more than one process. To take this into account, for each of the regions, we can write the expected number of events as:

$$N_X = N_X^{Z-\tau\tau} + N_X^{QCD} + N_X^{Wj} + N_X^{Wj + Wj-ee} + N_X^{\text{diboson}}$$

where $X$ denotes regions A-I(II) and B-I(II) for OS or LS data. We then use Eq. (20) to relate the QCD jet production backgrounds in primary and sideband regions.

Expected rates for $N_{Wj}^{QCD}$ are known from MC up to an overall scale factor $f_{OS/LS}^W$ to compensate for MC deficiency in predicting the probability of a jet faking a tau (Sec. IX B); the $N_{\tau\tau}$ distribution among regions is measured using converted photons (Sec. IX D) with an unknown overall scale $f^\gamma$; and the signal $N_{Z-\tau\tau}$ distribution among regions is known from MC simulation, and the normalization is directly related to the cross section being measured in this paper.

The system of eight equations [expanded from Eq. (21) for each of the eight regions] can be solved for eight parameters:

$$(N_A^{Z-\tau\tau}, f_{OS/LS}^W, f^\gamma, (N_A^{QCDOS}, N_A^{QCDLS}), (N_B^{QCDOS}, N_B^{QCDLS}), (N_B^{Wj}, N_B^{Wj})) = (22)$$

introduced above.

Adopting these notations, Tables X and XI illustrate this system of equations. To help readers make their own back-

---

**TABLE IX.** Definition of the regions used in estimating the number of signal events. Region A-I is the signal region. These regions are defined separately for OS and LS data. $(x,y)$ notation indicates the applicable range of values in a particular region for variables indicated in the left column.

<table>
<thead>
<tr>
<th>$M_T$ (GeV/$c^2$)</th>
<th>$I_{\text{trk}}^e$ (GeV/$c$)</th>
<th>Dominant Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0.50$</td>
<td>$0.1$</td>
<td>OS Data</td>
</tr>
<tr>
<td>$50:100$</td>
<td>$(2:8)$</td>
<td>$Z \rightarrow \tau\tau$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>QCD</td>
</tr>
<tr>
<td>$0.1$</td>
<td>$2.8$</td>
<td>LS Data</td>
</tr>
<tr>
<td></td>
<td></td>
<td>QCD</td>
</tr>
</tbody>
</table>

---

**TABLE X.** Definition of regions used in estimating the number of signal events. Region A-I is the signal region. These regions are defined separately for OS and LS data. $(x,y)$ notation indicates the applicable range of values in a particular region for variables indicated in the left column.

<table>
<thead>
<tr>
<th>$M_T$ (GeV/$c^2$)</th>
<th>$I_{\text{trk}}^e$ (GeV/$c$)</th>
<th>Dominant Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0.50$</td>
<td>$0.1$</td>
<td>OS Data</td>
</tr>
<tr>
<td>$50:100$</td>
<td>$(2:8)$</td>
<td>$Z \rightarrow \tau\tau$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>QCD</td>
</tr>
<tr>
<td>$0.1$</td>
<td>$2.8$</td>
<td>LS Data</td>
</tr>
<tr>
<td></td>
<td></td>
<td>QCD</td>
</tr>
</tbody>
</table>

---

**TABLE XI.** Definition of regions used in estimating the number of signal events. Region A-I is the signal region. These regions are defined separately for OS and LS data. $(x,y)$ notation indicates the applicable range of values in a particular region for variables indicated in the left column.

<table>
<thead>
<tr>
<th>$M_T$ (GeV/$c^2$)</th>
<th>$I_{\text{trk}}^e$ (GeV/$c$)</th>
<th>Dominant Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0.50$</td>
<td>$0.1$</td>
<td>OS Data</td>
</tr>
<tr>
<td>$50:100$</td>
<td>$(2:8)$</td>
<td>$Z \rightarrow \tau\tau$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>QCD</td>
</tr>
<tr>
<td>$0.1$</td>
<td>$2.8$</td>
<td>LS Data</td>
</tr>
<tr>
<td></td>
<td></td>
<td>QCD</td>
</tr>
</tbody>
</table>

---

**TABLE XII.** Definition of regions used in estimating the number of signal events. Region A-I is the signal region. These regions are defined separately for OS and LS data. $(x,y)$ notation indicates the applicable range of values in a particular region for variables indicated in the left column.

<table>
<thead>
<tr>
<th>$M_T$ (GeV/$c^2$)</th>
<th>$I_{\text{trk}}^e$ (GeV/$c$)</th>
<th>Dominant Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0.50$</td>
<td>$0.1$</td>
<td>OS Data</td>
</tr>
<tr>
<td>$50:100$</td>
<td>$(2:8)$</td>
<td>$Z \rightarrow \tau\tau$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>QCD</td>
</tr>
<tr>
<td>$0.1$</td>
<td>$2.8$</td>
<td>LS Data</td>
</tr>
<tr>
<td></td>
<td></td>
<td>QCD</td>
</tr>
</tbody>
</table>
of-the-envelope calculations, we scale the numerical values for the contributions with floating normalizations $f^i$ (denoting $Z \rightarrow \tau \tau$, $W +$ jets and $\gamma +$ jets) by corresponding $f^i_0$, the solutions of the system of equations, so that using $f^i/f^i_0 \approx 1$ one achieves an approximate solution of the system. Note that the values for these contributions in region A-I do not have statistical uncertainty, as it will be absorbed into the normalization definition. Uncertainties for other regions reflect the statistical uncertainty in the knowledge of how these contributions are distributed among regions with respect to region A-I.

While the information in Tables X and XI is sufficient to extract the number of expected signal events “by hand,” we choose to fit the expected number of events to the observations in all regions simultaneously and extract all the normalizations.

We build a likelihood function of the form:

$$L = \int \prod_x dN_x P(N_x, D_x) \prod_b d\nu_x \exp \left[ -\frac{(\nu_x - \nu_x')^2}{2\sigma_x^2} \right]$$

(23)

where the first product runs over eight regions, $D_x$ is the number of events in data in each region $X$, and $P(N_x, D_x)$ is the Poisson statistics probability for an expected rate $N_x$ and the number of observed events $D_x$. The second product runs over background type contributions ($Z \rightarrow ee$, $W +$ jets, $\gamma +$ jets, $t\bar{t}$ and diboson) and accounts for statistical uncertainties in the predictions for each of the regions and accuracy in the knowledge of the background distribution over the regions. The latter is applicable to signal, $W +$ jets and $\gamma +$ jets backgrounds as their overall normalizations are included in the fit parameters.

We then perform a transformation of integration variables from $N_x$ to physically important ones [as listed in Eq. (22)]. In the next step, we integrate out all parameters except the signal rate, $N^Z_{A-I}$, considering those to be nuisance parameters. For the integration in the default case, we apply an additional constraint $f^W_{OS}/f^W_{LS} = f^W_{LS}$, as this parameter is related to the accuracy in MC prediction for the probability of a jet to be misidentified as a tau and should not depend on the charge relationship. The fit yields $N^Z_{A-I} = 315 \pm 27$, where the uncertainty corresponds to a 68% C.L. interval around the central value. Figure 10 shows the logarithmic representation of the likelihood log$L/dN^Z_{A-I}$.

Apart from the signal rate, another interesting parameter obtained in the fit is $f^W$, which determines the normalization of the MC prediction for the $W +$ jets contribution. As discussed in the previous section, it relates to the accuracy

<table>
<thead>
<tr>
<th>Process</th>
<th>A-I</th>
<th>A-II</th>
<th>B-I</th>
<th>B-II</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Z \rightarrow ee$</td>
<td>3.9 ± 0.5</td>
<td>0.32 ± 0.13</td>
<td>0.43 ± 0.15</td>
<td>0.05 ± 0.05</td>
</tr>
<tr>
<td>t\bar{t}, diboson</td>
<td>0.18 ± 0.12</td>
<td>0.01 ± 0.01</td>
<td>0.5 ± 0.2</td>
<td>0.11 ± 0.09</td>
</tr>
<tr>
<td>$W +$ jets</td>
<td>$f^\tau$ (15 ± 2)</td>
<td>$f^\tau$ (1.2 ± 0.7)</td>
<td>$f^\tau$ (28 ± 3)</td>
<td>$f^\tau$ (2 ± 1)</td>
</tr>
<tr>
<td>$\gamma +$ jets</td>
<td>$f^\gamma$ (47.7)</td>
<td>$f^\gamma$ (29.9 ± 1.7)</td>
<td>$f^\gamma$ (1.4 ± 0.4)</td>
<td>$f^\gamma$ (10.0 ± 0.10)</td>
</tr>
<tr>
<td>QCD di-jets</td>
<td>$r \times N^\text{LSQCD}_{A-II}$</td>
<td>$N^\text{LSQCD}_{A-II}$</td>
<td>$r \times N^\text{LSQCD}_{B-II}$</td>
<td>$N^\text{LSQCD}_{B-II}$</td>
</tr>
<tr>
<td>$Z \rightarrow \tau \tau$</td>
<td>$f^\tau$ (3.9 ± 0.9)</td>
<td>$f^\tau$ (0.6 ± 0.3)</td>
<td>$f^\tau$ (0.8 ± 0.2)</td>
<td></td>
</tr>
<tr>
<td>Data</td>
<td>130</td>
<td>386</td>
<td>33</td>
<td>2</td>
</tr>
</tbody>
</table>

Table XI. Number of LS events in each of the regions. Solutions for the QCD contributions are $N^\text{QCDLS}_{A-I} = 60$, $N^\text{QCDLS}_{B-II}$ is negligible.
of the MC in describing the probability of a jet in these events to be misidentified as a hadronic tau. While it is unfair to expect MC to predict such a peculiar quantity accurately, it is interesting that at the position of the minimum $f_0^W = 0.73 \pm 0.07$, which is fairly close to unity showing a decent predicting power of PYTHIA MC for this rate.

To determine the systematic uncertainty on the background estimate, we first calculate the effect on the extracted signal yield by varying $r$ from 1/6 to 0.14 and 0.20, using the measurement of $r$ obtained in Sec. IX C. Note that in this case the procedure effectively reassigns events between the QCD and $\gamma +$ jet categories while the physically important result for $N_{A-I}^{TZ}$ changes by less than 3.0 events, which is assigned as the systematic uncertainty.

Second, we remove the constraint $f_{\text{OS}}^W = f_{\text{LS}}^W$. The resulting change in the fitted $N_{A-I}^{TZ}$ is approximately 1.5 events, which is assigned as a systematic uncertainty.

Third, we vary the prediction for the backgrounds, in which an electron is misidentified as a tau, by $\pm 20\%$ to account for the accuracy in MC predictions for the rate of such misidentifications. This leads to a change in the fitted $N_{A-I}^{TZ}$ by approximately 6 events, and is assigned as a systematic uncertainty.

We then sum up the three uncertainties in quadrature to obtain the systematic uncertainty associated with the assumptions used in the signal extraction method. The result is $N_{A-I}^{TZ} = 315 \pm 27 \pm 7$.

For reference, in Table XII we calculate the expected background contributions for the primary signal region A-I at the point of minimum of the eight parameter likelihood function in Eq. (23) and varying one parameter responsible

<table>
<thead>
<tr>
<th>Process</th>
<th>Yield (in number of events)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Z \rightarrow ee$</td>
<td>$34.5 \pm 1.4$</td>
</tr>
<tr>
<td>$t\bar{t} +$ diboson</td>
<td>$2.0 \pm 0.4$</td>
</tr>
<tr>
<td>$W +$ jets</td>
<td>$37 \pm 4$</td>
</tr>
<tr>
<td>$\gamma +$ jets</td>
<td>$48 \pm 13$</td>
</tr>
<tr>
<td>QCD di-jets</td>
<td>$68 \pm 4$</td>
</tr>
<tr>
<td>Total (sum):</td>
<td>$189.5$</td>
</tr>
<tr>
<td>Data</td>
<td>$504$</td>
</tr>
<tr>
<td>$Z \rightarrow \tau\tau$ (Fit)</td>
<td>$315 \pm 27 \pm 7$</td>
</tr>
</tbody>
</table>

for this particular background with all other fixed at values corresponding to the position of the minimum.

**Consistency checks**

We have analyzed many kinematic distributions to cross-check the background estimates used in this study. As an illustration, Figs. 11–13 show the kinematic distributions for electron and hadronic tau transverse momenta and the invariant mass of the $(e, \tau, E_T)$ system for the selected events in data and the expected background contributions. In all cases, the distributions are normalized to the number of signal and background events as extracted above.

We also compare results obtained from the one- and three-prong tau subsamples and find them to be in statistical agreement with the expected rate for $Z \rightarrow \tau\tau$ events. The measured cross sections for the two subsamples agree within 2(4)% with the average for subsamples with one(three)-prong taus.

We then remove the OS and $N_{\text{trk}} = 1, 3$ requirements and plot the distribution for the number of tau prongs in the data, comparing it with the background predictions, normalized to the summed signal yield in bins 1 and 3 with the OS and $N_{\text{trk}} = 1, 3$ requirements applied. Again we find good agreement as illustrated in Fig. 14. Note that the two-peaked structure characteristic of real taus is clearly present.

To illustrate that MC provides a reasonable description of the $W +$ jets background (apart from an overall scale factor $f^W \approx 0.73 \pm 0.07$), we remove the explicit cut on the transverse mass of the electron and missing transverse energy $M_T(e, E_T) < 50$ GeV/c and plot the distribution in Fig. 15. The region $M_T(e, E_T) < 50$ GeV/c corresponds
FIG. 11 (color online). Distribution of the hadronic tau candidate transverse momentum, $p_T$, for events passing all selections in data (points) compared to the sum of background and signal contributions. Note that this is only visible part of the momentum, as the neutrino escapes undetected. The $Z\rightarrow\tau\tau$ and background predictions are normalized to the number of events extracted from the fit.

FIG. 12 (color online). Distribution of the electron candidate transverse energy, $E_T$, for events passing all selections in data (points) compared to the sum of background and signal contributions. The $Z\rightarrow\tau\tau$ and background predictions are normalized to the number of events extracted from the fit.

FIG. 13 (color online). Distribution of the invariant mass of $e^-\tau^-E_T$ system for opposite sign events in data (points) compared to the sum of background and signal contributions. The $Z\rightarrow\tau\tau$ and background predictions are normalized to the number of events extracted from the fit.

FIG. 14 (color online). Distribution of the prong multiplicity of hadronic tau candidates with the OS and $N_{\text{trk}} = 1, 3$ requirements dropped. Data (points) are compared to the sum of background and signal contributions. The $Z\rightarrow\tau\tau$ and background predictions are normalized to the number of events extracted from the fit (with all cuts applied).
to the region B-I and is enriched with $W + \text{jets}$ backgrounds.

To further validate the procedure, we consider two alternative definitions of the QCD sideband regions: $2 < I < 5$ and $5 < I < 8 \text{ GeV}/c$. In each case, we recalculate the signal yield and find no indication of bias.

Another check we perform is a measurement of the $Z \to ee$ cross section using the same selections as in the main analysis with two exceptions: (i) the electron-rejecting cut and (ii) the $Z \to ee$ removal in the event topology selections is dropped. This measured cross section is in agreement with the published CDF measurement [4]. This comparison confirms that all isolation related selections are well understood. We also verify that variations in the instantaneous luminosity do not bring biases as, e.g., the isolation efficiencies are in fact dependent on the instantaneous luminosity. We study the dependence of the measured $Z \to ee$ cross section for several run ranges in the data and find no evidence of biases.

**XI. CROSS SECTION RESULTS**

With the definitions adopted earlier, we use the following formula to calculate the cross section:

$$
\sigma(Z/\gamma^* \to \tau\tau) = \frac{1}{L} \frac{1}{2B_{\tau e}B_{\tau \tau\text{had}}} \frac{N_{\text{signal}}}{\alpha e_{\text{ID}} e_{\text{trig}} e_{\text{hig}} e_{\text{event}}} (24)
$$

where $N_{\text{signal}}$ is the measured signal yield. The PDG [20] values for the branching ratios are $B_{\tau e} = 17.84 \pm 0.06\%$ and $B_{\tau \tau\text{had}} = 64.79 \pm 0.08\%$.

We summarize all systematic uncertainties presented earlier in Table XIII. Note that the PDF and energy uncertainties are absorbed into the acceptance uncertainty. Other uncertainties are negligible. For example, there is a small effect due to jets that can sometimes be mistaken for hadronic taus in true $Z \to \tau\tau$ events, while the true hadronic tau is not reconstructed. The size of this effect is estimated by dropping matching requirements between reconstructed and generated taus and is found to be small (a fraction of 1%). In addition, this effect is accounted for in the background subtraction scheme, where these events will be treated as charge symmetric backgrounds with an isolated electron.

The cross section is measured to be $\sigma(p\bar{p} \to Z/\gamma^* \to \tau\tau) = 263 \pm 23(\text{stat}) \pm 14(\text{syst}) \pm 15(\text{lumi}) \text{ pb}$. This result is in agreement with a CDF measurement [4] using the $ee$ and $\mu\mu$ channels which yielded $\sigma(p\bar{p} \to Z/\gamma^* \to ll) = 254.9 \pm 3.3(\text{stat}) \pm 4.6(\text{syst}) \pm 15.2(\text{lumi}) \text{ pb}$. A theoretical calculation in the framework of NNLO [21] includes only the $Z$ boson contribution and predicts $\sigma(p\bar{p} \to Z)B(Z \to ll) = 251.3 \pm 5.0 \text{ pb}$. To compare the cross section measurement described in this paper to the theoretical expectation, we convert our measurement of $\sigma(p\bar{p} \to Z/\gamma^* \to \tau\tau)$ into the cross section for pure $Z$ by multiplying it by a correction factor of $F = 1.004 \pm 0.001$ [22]. We obtain $\sigma(p\bar{p} \to Z)B(Z \to \tau\tau) = 264 \pm 23(\text{stat}) \pm 14(\text{syst}) \pm 15(\text{lumi}) \text{ pb}$ as our final result.

**XII. SUMMARY**

We have presented the first CDF measurement of the cross section for the process $p\bar{p} \to Z \to \tau\tau$ in final states where one tau has decayed into an electron and the other into hadrons. The measured value of $\sigma(p\bar{p} \to Z)B(Z \to \tau\tau) = 264 \pm 23(\text{stat}) \pm 14(\text{syst}) \pm 15(\text{lumi}) \text{ pb}$ is in good agreement with the NNLO expectation [21]. The measurement is still statistically limited, but the systematic uncertainties are comparable to the measurements in $ee$ and $\mu\mu$ channels. While this analysis is not designed to target small signals for new physics, the agreement of this measure-
ment with the SM precludes the presence of large signals that could be associated with new phenomena in the final state with an electron and a hadronically-decaying tau.

The efficiency of hadronically-decaying tau reconstruction and identification at CDF, described in this paper, is understood at the 3% level. This is a significant improvement over previous results. The data-driven technique developed in this study determines jet induced backgrounds to approximately 6% relative to the size of these background contributions. This uncertainty is limited by statistics and will improve further with more data. This technique significantly outperforms traditional methods based on measuring the misidentification rates in generic jet data, where the typical uncertainty is well above 20%. These advances improve the CDF experiment’s efficiency for the new physics in channels involving hadronically-decaying tau leptons.

With the significant increase in the data set size in Run II, CDF is working on several important analyses targeting di-tau final states that are expected in Higgs and SUSY models, as well as searches for third generation lepton-quarks and doubly charged Higgs bosons. Identification procedures and background estimation methods in these analyses rely on the techniques described in this paper and improve the sensitivity for possible signals of new phenomena.