Search for the associated production of a b quark and a neutral supersymmetric Higgs boson which decays to tau pairs
We report results from a search for production of a neutral Higgs boson in association with a $b$ quark. We search for Higgs decays to $\tau\tau$ pairs with one $\tau$ subsequently decaying to a muon and the other to hadrons. The data correspond to 2.7 fb$^{-1}$ of $p\bar{p}$ collisions recorded by the D0 detector at $\sqrt{s} = 1.96$ TeV. The data are found to be consistent with background predictions. The result allows us to exclude a significant region of parameter space of the minimal supersymmetric model.

PACS numbers: 12.10.Dm, 12.60.Fr, 12.60.Jv

The current model of physics at high energies, the standard model (SM), has withstood increasingly precise experimental tests, although the Higgs boson needed to mediate the breaking of electroweak symmetry has not been found. Despite the success of the SM, it has several shortcomings. Theories invoking a new fermion-boson symmetry, called supersymmetry [1] (SUSY), provide an attractive means to address some of these including the hierarchy problem and nonunification of couplings at high energy. In addition to new SUSY-specific partners to SM particles, these theories have an extended Higgs sector. In the minimal supersymmetric standard model (MSSM) there are two Higgs doublet fields which result in five physical Higgs bosons: two neutral scalars...
(h, H), a neutral pseudoscalar (A) and two charged Higgs bosons (H±). The mass spectrum of the Higgs bosons is determined at tree level by two parameters, typically chosen to be tan β, the ratio of the vacuum expectation values of up-type and down-type scalar fields and M\_A, the mass of the physical pseudoscalar. Higher order corrections are dominated by the Higgsino mass parameter μ and the mixing of scalar top quarks.

In this Letter, we present a search for neutral Higgs bosons (collectively denoted \( \phi \)) produced in association with a b quark. The specific Higgs boson decay mode used in this search is \( \phi \rightarrow \tau\tau \) with one of the \( \tau \) leptons subsequently decaying via \( \tau \rightarrow µν_µν_µ \) (denoted \( \tau_µ \)) and the second via \( \tau \rightarrow \mu \) hadrons + \( ν_τ \) (denoted \( ν_τ \)). In the MSSM the Higgs coupling to down-type fermions is enhanced by a factor \( \propto \tan^2 β \) and thus the Higgs production cross section is enhanced by a factor \( \propto \tan β \) relative to the SM, giving potentially detectable rates at the Tevatron. Two of the three neutral Higgs bosons have nearly degenerate masses over much of the parameter space, effectively giving another factor of two in production rate. A previous search in this final state was carried out by the D0 experiment [2]. Searches in the complementary channels \( \phi Z/\phi \phi \rightarrow b\bar{b}τ\tau, ττb \) [3], \( \phi \rightarrow ττ \) [4, 5], and \( \phi b \rightarrow bbb \) [6, 7] have also been carried out by the LEP, D0, and CDF experiments. By searching in complementary channels we reduce overall sensitivity to the particular details of the model. The \( \tau\tau \) final state is less sensitive to SUSY radiative corrections than the \( b\bar{b} \) final state, and has greater sensitivity at low Higgs mass than the \( \phi \rightarrow \tau\tau \) channel, as the \( b \)-jet in the final state reduces the \( Z \rightarrow \tau\tau \) background. Furthermore, an additional complementary channel will contribute to an even stronger exclusion when combining different searches. The result presented in this Letter uses an integrated luminosity of 2.7 \( fb^{-1} \) which is eight times that used for the previous result in this channel. Because of analysis improvements, the gain in sensitivity compared to the prior result is greater than expected from the increased integrated luminosity only. We also extend the Higgs mass search range relative to the previous result in this channel.

The D0 detector [8] is a general purpose detector located at Fermilab’s Tevatron pp collider. The Tevatron operates at a center of mass energy of 1.96 TeV. This analysis relies on all aspects of the detector: tracking, calorimetry, muon detection, the ability to identify detached vertices and the luminosity measurement.

This search requires reconstruction of muons, hadronic τ decays, jets (arising from b quarks) and neutrinos. Muons are identified using track segments in the muon system and are required to have a track reconstructed in the inner tracking system which is close to the muon-system track segment in \( \eta \) and \( \varphi \). Here \( \eta \) is the pseudorapidity and \( \varphi \) is the azimuthal angle in the plane perpendicular to the beam. Jets are reconstructed from calorimeter information using the D0 Run II cone algorithm [9] with a radius of \( R = 0.5 \) in \( (y, \varphi) \) space, where \( R = \sqrt{2(D_\eta^2 + D_\varphi^2)} \) and \( y \) is the rapidity. Jets are additionally identified as being consistent with decay of a b-flavored hadron (b-tagged) if the tracks aligned with the calorimeter jet have high impact parameter or form a vertex separated from the primary interaction point in the plane transverse to the beam as determined by a neural network (NN\_v) algorithm [10]. Hadronic τ decays are identified [11] as clusters of energy in the calorimeter reconstructed [9] using a cone algorithm of radius \( R = 0.3 \) which have associated tracks. The τ candidates are then categorized as being one of three types which correspond roughly to one-prong τ decay with no π⁰s (called Type 1), one-prong decay with π⁰s (Type 2) and multiprong decay (Type 3). A final identification requirement is based on the output value of a neural network (NN\_v) designed to separate τ leptons from jets. The τ is characterized as being consistent with decay of a vector boson or from a \( b \) quark to satisfy \( p_T > 15 \text{ GeV/c} \) and \( |y| < 2.5 \) and using the CTEQ6L1 [13] parton distribution functions (PDF). The TAUOLA [14] program is used to model τ decay and EVTGEN [15] is used to decays hadrons. The dependence of the Higgs boson decay width on tan β is included by reweighting PYTHON samples, and the kinematic properties are reweighted to the prediction of the NLO program MCFM [16]. The generator outputs are passed through a detailed detector simulation based on GEANT [17]. Each GEANT event is combined with collider data events recorded during a random beam crossing to model the effects of detector noise, pileup, and additional pp interactions. The combined output is then passed to the D0 event reconstruction program. Simulated signal samples are generated for different Higgs masses ranging from 90 to 320 GeV/c².

Backgrounds to this search are dominated by \( Z + \text{jets} \), \( τ \ell \), and multijet (MJ) production. In the MJ background the apparent leptons primarily come from semileptonic \( b \) hadron decays, not τ decays. Additional backgrounds include \( W + \text{jets} \) events, SM diboson production and single top quark production. Except for the MJ contribution, all background yields are estimated using simulated events, with the same processing chain used for signal events. The \( Z + \text{jets} \), \( W + \text{jets} \) and \( τ \ell \) samples are generated using ALPGEN [18] with PYTHIA used for fragmentation. The diboson samples are generated using PYTHIA. For simulated samples in which there is only one lepton arising from the decay of a W boson or from \( τ \ell \rightarrow \ell + \text{jets} \), the second lepton is either a jet misidentified as a τ or a muon+jet system from heavy flavor decay in which the muon is misidentified as being isolated from other activ-
Corrections accounting for differences between data and the simulation are applied to the simulated events. The corrections are derived from control data samples and applied to object identification efficiencies, trigger efficiencies, primary \( p\bar{p} \) interaction position (primary vertex) and the transverse momentum spectrum of \( Z \) bosons. After applying all corrections, the yields for signal and each background are calculated as the product of the acceptance times efficiency determined from simulation, luminosity and predicted cross sections.

The initial analysis step is selection of events recorded by at least one trigger from a set of single muon triggers for data taken before the summer of 2006. For data taken after summer 2006 we require at least one trigger from a set of single muon triggers and muon plus hadronic \( \tau \) triggers. The average trigger efficiency for signal events is approximately 65% for both data epochs.

After making the trigger requirements a background-dominated pre-tag sample is selected by requiring a reconstructed primary vertex for the event with at least three tracks, exactly one reconstructed hadronic \( \tau \), exactly one isolated muon, and at least one jet. This analysis requires the \( \tau \) candidates to satisfy \( E_T > 10 \text{ GeV}, p_T^{\tau} > 7(5) \text{ GeV}/c \) and \( NN_{\tau} > 0.9 \) for Type 1(2) taus, \( E_T^{\tau} > 15 \text{ GeV}, p_T^{\tau} > 10 \text{ GeV}/c \) and \( NN_{\tau} > 0.95 \) for Type 3 taus. Here \( E_T^{\tau} \) is the transverse energy of the \( \tau \) measured in the calorimeter, \( p_T^{\tau} \) is the transverse momentum sum of the associated track(s). The muon must satisfy \( p_T^{\mu} > 12 \text{ GeV}/c \) and \( |\eta| < 2.0 \). It is also required to be isolated from activity in the tracker and calorimeter [19]. Selected jets have \( E_T > 15 \text{ GeV}, |\eta| < 2.5 \). The \( \tau \), the muon and jets must all be consistent with arising from the same primary vertex and be separated from each other by \( \Delta R > 0.5 \). In addition, the muon and \( \tau \) are required to have opposite charge, and the \( (\mu, E_T) \) mass variable \( M \equiv \sqrt{2E_T^{\mu}E_T^{\tau}/p_T^{\mu}(1 - \cos(\Delta \phi(\mu, E_T^{\mu}))) \) must satisfy \( M < 80, 80, \) and \( 60 \text{ GeV}/c^2 \) for events with \( \tau \)s of Type 1, 2 and 3 respectively. Here \( E_T^{\mu} \) is the energy of the muon, and \( \Delta \phi \) is the opening angle between the \( E_T \) and muon in the plane transverse to the beam direction.

A more restrictive \( b \)-tag subsample with improved signal to background ratio is defined by demanding that at least one jet in each event is consistent with a quark production [10]. The \( b \)-jet identification efficiency in signal events is about 35% and the probability to misidentify a light jet as a \( b \) jet is 0.5%.

All backgrounds except MJ are derived from simulated events as described earlier. The MJ background is derived from control data samples. A parent MJ-enriched control sample is created by requiring a muon, \( \tau \), and jet as above, but with the muon isolation requirement removed and with a lower quality \( (0.3 \leq NN_{\tau} \leq 0.9) \) \( \tau \) selected. This is then used to create a \( b \)-tag subsample which requires at least one of the jets to be identified as a \( b \) jet with the same \( b \) jet selection as earlier. The residual contributions from SM backgrounds are subtracted from the MJ control samples using simulated events.

To determine the MJ contribution in the pre-tag analysis sample, a data sample is used that has the same selection as the pre-tag analysis sample except that the muon and \( \tau \) charges have the same sign. This same-sign (SS) sample is dominated by MJ events. After making a subtraction of other SM background processes which contribute to this sample, the number of MJ events in the opposite-sign (OS) signal region is computed by multiplying the SS sample by the OS/SS ratio, \( 1.05 \pm 0.02 \), determined in a control sample selected by requiring a non-isolated muon.

For the \( b \)-tag analysis sample, statistical limitations require a different approach for the MJ background evaluation than for the pre-tag sample. For the \( b \)-tag sample, two methods are used. For the first method, the per jet probability \( P_{\text{tag}} \) that a jet in the SS MJ control subsample would be identified as a \( b \) jet is determined as a function of jet \( p_T \). We apply \( P_{\text{tag}} \) to the jets in the SS pre-tag sample to determine the yield in the \( b \)-tag sample. For the second method, the MJ background is determined by multiplying the \( b \)-tag MJ control sample yield by two factors: (1) the probability that the non-isolated muon would be identified as isolated, and (2) the ratio of events with a \( \tau \) candidate passing the \( NN_{\tau} \) requirements to events with \( \tau \) candidates having \( 0.3 \leq NN_{\tau} < 0.9 \) as determined in a separate control sample. The final MJ contribution in the \( b \)-tag analysis sample is determined using the MJ shape from the first method with the normalization equal to the average of the two methods. We include the normalization difference between the two methods in the systematic uncertainty on the MJ contribution.

The signal to background ratio is further improved using multivariate techniques. Two separate methods are used, one to address the \( t\bar{t} \) background and one to reduce the MJ background. For the \( t\bar{t} \) background, a neural network (\( NN_{\text{top}} \)) is constructed using \( H_T \equiv \Sigma_{\text{jets}} E_T, E_{\text{tot}} \equiv \Sigma_{\text{jets}} E + E_\tau + E_\mu \), the number of jets and \( \Delta \phi(\mu, \tau) \) as inputs. For the MJ background, a simple joint likelihood discriminant \( (LL_{MJ}) \) is constructed using \( p_T^{\mu}, p_T^{\tau}, \Delta R(\mu, \tau), M_\mu, \) and \( M_{\mu\tau} \). Here \( M_{\mu\tau} \) denotes the invariant mass of the muon and tau, and \( M_{\mu\tau} \) is the invariant mass computed from the muon, \( \tau \), and \( E_T \) momentum vectors. The final analysis sample is defined by selecting rectangular regions in the \( NN_{\text{top}} \) versus \( LL_{MJ} \) plane. The regions have been identified for each \( \tau \) type and each Higgs boson mass point separately by optimizing the search sensitivity using simulated events. The signal to background ratio improves by up to a factor of two when applying these requirements.

Table I shows the predicted background and observed data yields in the analysis samples. Between 5% and 10%
of $\phi \to \tau_\mu \tau_\nu$ decays are selected depending on $M_\phi$.

Systematic uncertainties arise from a variety of sources. Most are evaluated using comparisons between data control samples and predictions from simulation. The uncertainties are divided into two categories: (1) those which affect only normalization, and (2) those which also affect the shape of distributions. The sources in the first category include the luminosity (6.1%), muon identification efficiency (4.5%), $\tau$ identification (5%, 4%, 8%), $\tau$ energy calibration (3%), the $t\bar{t}$ and single top cross sections (11% and 12%), diboson cross sections (6%), $Z+(u,d,s,c)$ rate (42%, 5%) and the $W+b$ and $Z+b$ cross sections (30%); those in the second include jet energy calibration (2%-4%), $b$-tagging (3%-5%), trigger (3%-5%), and MJ background (33%, 12%, 11%). For sources with three values, the values correspond to $\tau$ Types 1, 2 and 3 respectively.

After making the final selection, the discriminant $D$ is formed from the product of the $NN_{top}$ and $LL_{MJ}$ variables, $D = LL_{MJ} \times NN_{top}$. The resulting distributions for the predicted background, signal and data are shown in Fig. 1(a). This distribution is used as input to a significance calculation using a modified frequentist approach with a Poisson log-likelihood ratio test statistic [20]. In the absence of a significant signal we set 95% confidence level limits on the presence of neutral Higgs bosons in our data sample. The cross section limits are shown in Fig. 1(b) as a function of Higgs boson mass.

Table 1 shows the predicted background yield, observed data yield and predicted signal yield and their statistical uncertainties at three stages of the analysis. The signal yields are calculated assuming $\tan \beta = 40$ and a Higgs mass of 120 GeV/c$^2$ for the $m_h^{max}$ and $\mu = -200$ GeV/c$^2$ scenario.

Table I: Predicted background yield, observed data yield and predicted signal yield and their statistical uncertainties for the parameter space.

<table>
<thead>
<tr>
<th></th>
<th>Pre-tag</th>
<th>$b$-tagged</th>
<th>Final</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t\bar{t}$</td>
<td>66.0 ± 1.3</td>
<td>39.6 ± 0.8</td>
<td>20.3 ± 0.6</td>
</tr>
<tr>
<td>Multijet</td>
<td>549 ± 26</td>
<td>38.5 ± 2.3</td>
<td>28.1 ± 1.9</td>
</tr>
<tr>
<td>$Z(\rightarrow \tau \tau) +$ jets</td>
<td>1241 ± 8</td>
<td>18.8 ± 0.3</td>
<td>16.3 ± 0.3</td>
</tr>
<tr>
<td>Other Bkg</td>
<td>267 ± 6</td>
<td>5.1 ± 0.1</td>
<td>4.1 ± 0.1</td>
</tr>
<tr>
<td>Total Bkg</td>
<td>2123 ± 28</td>
<td>102 ± 2.4</td>
<td>68.8 ± 2.0</td>
</tr>
<tr>
<td>Data</td>
<td>2077</td>
<td>112</td>
<td>79</td>
</tr>
<tr>
<td>Signal</td>
<td>14.4 ± 0.3</td>
<td>4.8 ± 0.1</td>
<td>4.6 ± 0.1</td>
</tr>
</tbody>
</table>

We thank the staffs at Fermilab and collaborating institutions, and acknowledge support from the DOE and NSF (USA); CEA and CNRS/IN2P3 (France); FASI, Rosatom and RFBR (Russia); CNPq, FAPERJ, FAPESP and FUNDUNESP (Brazil); DAE and DST (India); Colciencias (Colombia); CONACyT (Mexico); KRF and KOSEF (Korea); CONICET and UBACyT (Argentina); FOM (The Netherlands); STFC and the Royal Society (United Kingdom); MSMT and GACR (Czech Republic); CRC Program, CFI, NSERC and WestGrid Project (Canada); BMBF and DFG (Germany); SFI (Ireland); The Swedish Research Council (Sweden); and CAS and CNSF (China).
FIG. 1: (a) The distribution of the final discriminant variable, \( D = N_{\text{lep}} \times LLMJ \). The figure includes all \( \tau \) Types. (b) The cross-section limit as a function of Higgs boson mass. (c) The region in the \( \tan \beta \) versus \( M_A \) plane excluded by this analysis, LEP neutral MSSM Higgs searches, and the previous D0 result in this channel.