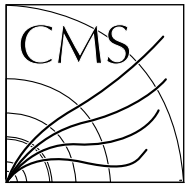
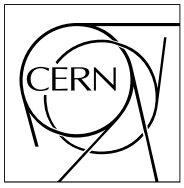
Available on CMS information server **CMS NOTE 2006/101**



The Compact Muon Solenoid Experiment

CMS Note

Mailing address: CMS CERN, CH-1211 GENEVA 23, Switzerland



11 May 2006

Study of MSSM H/A$\rightarrow \tau\tau \rightarrow e\mu + X$ in CMS

S. Lehti

Helsinki Institute of Physics, Helsinki, Finland

Abstract

The electron+muon final states from the H/A$\rightarrow \tau\tau$ decay channel of the neutral MSSM Higgs bosons produced in association with b quarks in $gg \rightarrow b\bar{b}H/A$ are studied with full simulation in CMS. The associated b jets are used to help extract the signal from the background. The discovery potential is evaluated for 30 fb^{-1} at low luminosity. The effect of systematic uncertainties of the background determination on the discovery reach is discussed.

1 Introduction

Since the Higgs mechanism is a cornerstone of the Standard Model and its supersymmetric extensions, one of the most important physics objectives of the LHC is to discover the Higgs boson. In the Minimal Supersymmetric extension of the Standard Model (MSSM) two Higgs doublets are required to preserve supersymmetry leading to five elementary Higgs particles, two CP-even (h,H), one CP-odd (A) and two charged Higgs bosons ($H^{\pm}$). At lowest order all couplings and masses of the MSSM Higgs sector are determined by two independent input parameters, which are generally chosen as the CP-odd Higgs boson mass m_A, and the ratio of the vacuum expectation values of the Higgs fields $v_2/v_1 \equiv \tan\beta$.

In MSSM the Higgs boson couplings to down type fermions, such as b's and τ's, are tanβ enhanced. This enhancement leads to large cross sections in the associated production process $gg \to b\bar{b}H/A$, and to a significant branching ratio in the $H/A \to \tau\tau$ decay channel at large tanβ ($\gtrsim$10). This behaviour makes the $H/A \to \tau\tau$ channel a potential discovery channel for heavy neutral Higgs bosons H and A.

The results from LEP including all data, up to the highest energy 209 GeV, show no significant Higgs signal [1]. The no mixing scenario [2] is almost completely excluded. The search results for the m_h-max scenario [2], which has maximal mixing in the squark (stop) sector, are shown in Figs. 1 and 2 as exclusion regions in the (m_A,tanβ) and (m_h,tanβ) parameter space. The excluded tanβ region shown in Fig. 2 depends on the top quark mass. The most recent determination of the top mass is 172.7$\pm$2.9 GeV/c^2 [3]. The LEP results suggest that the lightest Higgs boson h is close to its expected upper mass limit and mixing in the squark sector is large. Large mixing and light stop quark have, however, only a small effect on the Higgs boson production in association with b quarks. In this work the m_h-max scenario is assumed with the following SUSY parameters: X_t = 2000 GeV/c^2, M_2 = 200 GeV/c^2, μ = 200 GeV/c^2, $M_{\tilde{g}}$ = 800 GeV/c^2 and $M_{\tilde{q},\tilde{l}}$ = 1 TeV/c^2. The top mass is set to 175 GeV/c^2.

The different leptonic and hadronic final states from $H/A \to \tau\tau$ have been investigated earlier in [4, 5, 6, 7, 8] in CMS. The two-lepton final states, presenting a fraction of 12.5% of the $H/A \to \tau\tau$ decays, have been shown to be accessible for $m_A \lesssim 300 GeV/c^2$ and tan$\beta \gtrsim$ 15. The leptonic final state gives a clean signal which is easy to trigger. Two possibilities exist, either to select any-two-lepton final states, which have larger signal rate, or electron+muon final states for which the background is easier to suppress. Here the electron+muon final state with updated cross sections and with full CMS simulation is studied. Cuts are optimized to get visible signal over the background with high statistical significance.

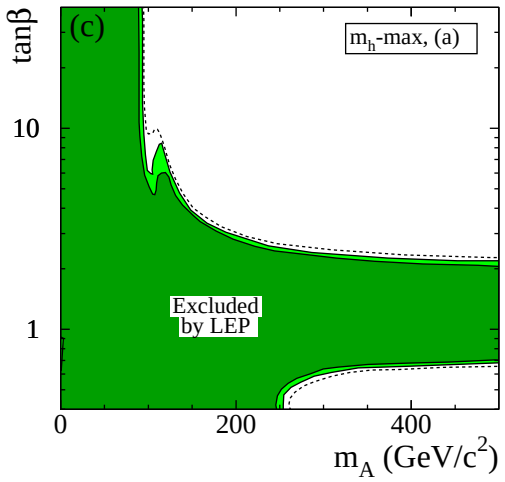


Figure 1: LEP exclusion zone for 95% CL (light color), 99.7% CL (dark color) and for 95% CL on the basis of Monte-Carlo simulations (dashed) [1].

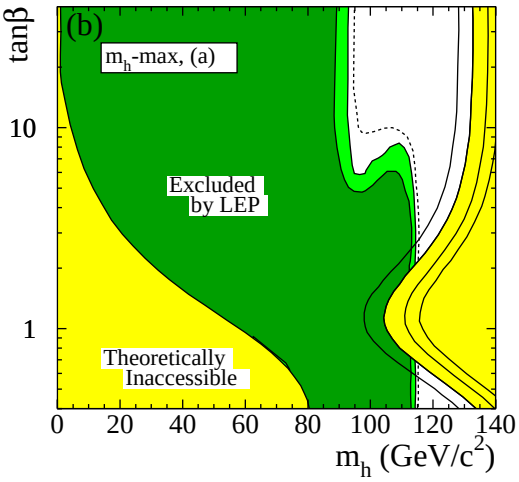


Figure 2: The LEP exclusion regions as in Fig. 1 and for top mass of 169.3, 174.3, 179.3 and 183.0 GeV/c^2 (curves on the right hand side from left to right) [1].

2 Cross sections

2.1 Signal

The dominant production mechanism at large $\tan\beta$ is the Higgs boson production in association with two b quarks, representing about 90% of the total cross section. The second most important production mechanism is Higgs boson production via loop mediated gluon fusion. The LO Feynman graphs for the associated production process are shown in Fig. 3. Other production mechanisms are e.g. associated production with top quarks and vector boson fusion, but their contribution to the total cross section is small. The event selection method used in this study selects effectively the events with associated b quarks only and the other production methods can be safely neglected.

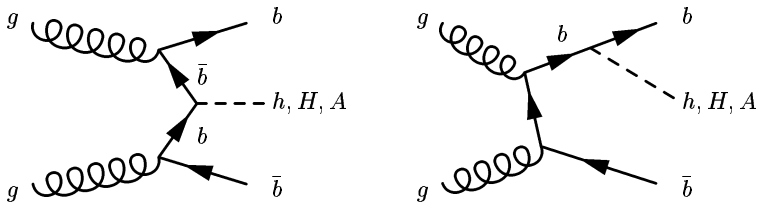


Figure 3: Example LO Feynman graphs contributing to Higgs boson production at LHC.

The cross sections for the two most important Higgs boson production mechanisms are shown in Fig. 4 for m_h-max scenario. The cross sections for the associated production and for the loop mediated gluon fusion are calculated with FeynHiggs [9]. At low m_A the light CP-even Higgs boson h production is important, the cross section is large and the h mass is close to CP-odd Higgs boson mass therefore contributing to signal events. At large m_A the h and A mass are separated, and signal events consist of H and A bosons only. The CP-odd Higgs boson production is shown with a dashed line in the figure.

At large $\tan\beta$, the Higgs boson decay predominantly into b quarks and τ's. The branching ratios $H \to \tau\tau$ and $A \to \tau\tau$, calculated with program FeynHiggs [9], are shown in Figs. 5 and 6 for H and A, respectively. The effect of the SUSY Higgs mass parameter μ on the branching ratios is shown. At large m_A the Higgs boson decays into neutralinos and charginos are kinematically allowed, which affects the branching ratios to SM particles via total decay width. In the mass region used in this study, $m_A \lesssim 300$ GeV/c^2, the Higgs boson branching ratio to two τ's is almost independent of the Higgs mass. At large $\tan\beta$ the dependence of the $\tau\tau$ branching ratio on Higgs mass becomes even smaller. Due to the chosen $\tau\tau$ final state with one τ decaying to electron, the other to muon, the

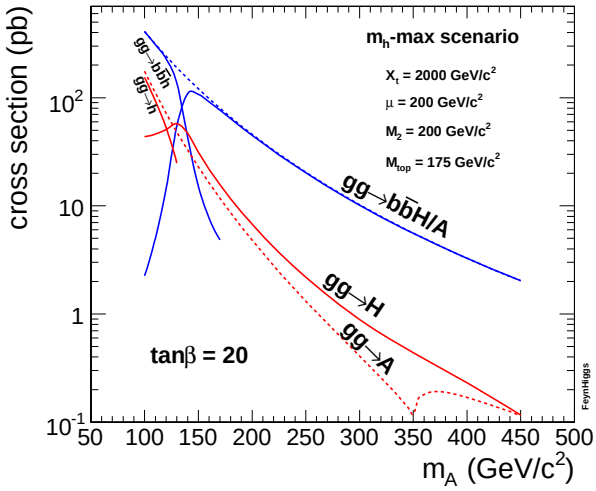


Figure 4: Higgs boson cross sections for associated and loop mediated production in m_h-max scenario for $\tan\beta$ = 20. The dotted line represents the CP-odd Higgs boson production. At low m_A the h production becomes more important than the H production.

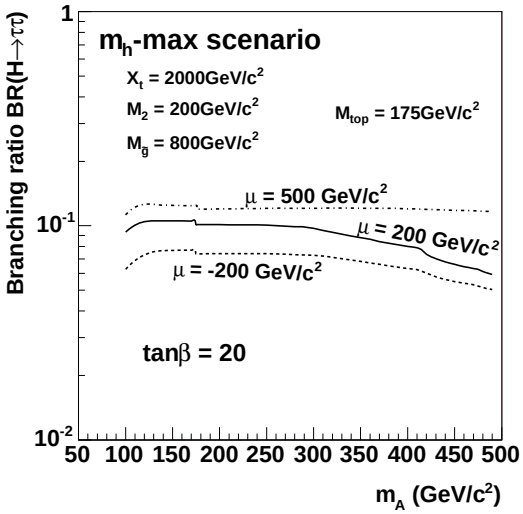


Figure 5: Branching ratio for $H \to \tau\tau$ in m_h-max scenario for tanβ = 20 as a function of m_A for different values of μ.

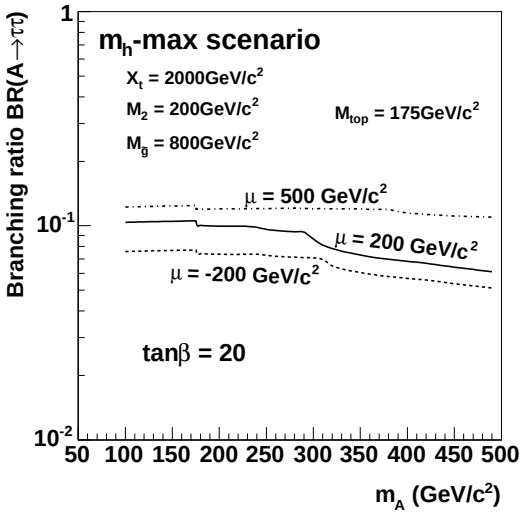


Figure 6: Branching ratio for $A \to \tau\tau$ in m_h-max scenario for tanβ = 20 as a function of m_A for different values of μ.

signal rate is low compared to the hadronic and semileptonic $\tau\tau$ final states.

The signal events are generated with PYTHIA [10] for three Higgs boson masses 140, 200 and 250 GeV/c^2. These mass values are chosen based on fast simulation results [5]: m_A = 140 GeV/c^2 is expected to give best reach in tanβ. The TAUOLA package [11] is used for simulating τ decays. The signal cross section and branching ratio for the m_h-max scenario is calculated with FeynHiggs [9].

2.2 Background

The background comes mainly from two sources, Z/γ^* events decaying into $\tau\tau$, and $t\bar{t}$ events with W decaying to lepton directly or via intermediate τ. The Z/γ^* events with associated genuine b quarks are simulated separately. The Z/γ^* background sample contains also events with two associated b quarks, but to prevent double counting, those events are removed using the available generation level information of the events.

The Z/γ^* background associated with light quark and gluon jets is generated with PYTHIA [10]. An NLO cross section of 233.3 pb [12] calculated with the program MCFM [13] assuming $\tau\tau \to \ell\ell$ final state and $m_{\tau\tau} > 80$ GeV/c^2 is used. The background consisting of Z/γ^* events produced in association with b quarks emitted mostly in the forward direction is generated with CompHEP [14] for $m_{\tau\tau} > 60$ GeV/c^2. For these events cross section

Process		dataset	pb	events
$bbH/A, H/A \to \tau\tau \to \ell\ell + X$	m_A = 140 GeV/c^2	hg03t_h2tau_2l_140	0.39027	38k
$b\bar{b}H/A, H/A \to \tau\tau \to \ell\ell + X$	m_A = 200 GeV/c^2	hg03t_h2tau_2l_200	0.9692	20k
$b\bar{b}H/A, H/A \to \tau\tau \to \ell\ell + X$	m_A = 250 GeV/c^2	hg03t_h2tau_2l_250	0.39027	10k
$Z/\gamma^* \to \tau\tau \to \ell\ell + X$	80< $m_{\tau\tau}$ < 100GeV/c^2	hg03_zg80_100_2tau_2l	223.2	250k
$Z/\gamma^* \to \tau\tau \to \ell\ell + X$	$m_{\tau\tau}$ > 100GeV/c^2	hg03_zg100_2tau_2l	10.1	250k
$b\bar{b}Z/\gamma^*, Z/\gamma^* \to \tau\tau \to \ell\ell + X$	60< $m_{\tau\tau}$ < 100GeV/c^2	hg03_tautaubb_ll_60_100_compHEP	3.29	290k
$b\bar{b}Z/\gamma^*, Z/\gamma^* \to \tau\tau \to \ell\ell + X$	$m_{\tau\tau}$ > 100GeV/c^2	hg03_tautaubb_ll_100_compHEP	0.132	98k
$t\bar{t} \to \ell\ell$+X		eg03_tt_2l_topr+pyth	86.2	392k
$tW \to \ell\ell$+X		eg03_wt_2l_toprex	6.16	200k
$b\bar{b} \to \mu\mu$+X		mu03_bb2mu	18084	1M
$WW \to \ell\ell$+X		hg03_ww_2l	5.21	198k
$WZ \to \ell\ell\ell$+X		hg03b_zw_3l	1.61	100k
$ZZ \to \ell\ell\nu\nu$		hg03b_zz_2l2nu	1.06	100k

Table 1: Cross sections for the signal and background processes and events available in the datasets.

calculated with CompHEP is used. The LO cross section for any-two-lepton final state is 3.4 pb. No p_T or η cuts are applied on b quarks in the $b\bar{b}Z/\gamma^*$ process generation. The b quarks are assumed to have nonzero mass.

The $t\bar{t}$ and single top (tW) events are generated with TopREX [15]. Another $t\bar{t}$ datasample generated with PYTHIA is also used. An inclusive cross section of 840 and 60 pb is used for $t\bar{t}$ and tW events, respectively [16].

The backgrounds from $b\bar{b}$, WW, WZ and ZZ events are also included and are generated with PYTHIA. The contribution from these backgrounds turn out, however, to be small after event selection. The W+jets with jet faking an electron, and Wbb backgrounds are not taken into account as their contribution is estimated to be negligible. The cross sections for signal and main background processes are shown in Table 1.

3 Detector simulation and event reconstruction

The reconstruction is based on official CMS digitised datasets with pile-up included (3.4 minimum bias events superimposed per event crossing for luminosity $2 \times 10^{33} cm^{-2} s^{-1}$). The detector simulation has been done with full GEANT[17] simulation. The CMS detector is simulated with complete ideal detector, no staging and no misalignment of the detector elements is assumed.

The physics objects of the simulation, electrons, muons and jets, are reconstructed using standard methods available in the CMS reconstruction software. Version ORCA_8_7_4 of CMS OO Reconstruction [18] is used. The vertex with highest sum of the transverse momenta of the associated tracks is selected as the primary vertex. Jets are reconstructed in a cone of 0.5 using corrections determined from Monte-Carlo [19]. The electromagnetic+hadron calorimeter tower E and E_T thresholds used in the jet reconstruction are 0.8 GeV and 0.5 GeV, respectively.

4 Event selection

4.1 Trigger

The Level-1 trigger involves the calorimetry and muon systems as well as some correlation of information from these systems. The Level-1 decision is based on the presence of local objects such as photons, electrons, muons and jets [20].

The events are triggered with single and double electron and muon trigger. The Level-1 p_T threshold for single muons is 10 GeV/c, for single electrons 20 GeV/c, for double muons 3 GeV/c and for double electrons 10 GeV/c [20]. The trigger is chosen to include all leptonic final states to have the option for studying both any-two-lepton and electron+muon events in parallel. The Level-1 trigger efficiency for the signal is 0.961 at m_A = 200 GeV/c^2.

The selection of electrons in the High-Level Trigger (HLT) proceeds in three steps. The first step uses the calorimeter information alone. The next step demands hits in the pixel detectors consistent with an electron candidate. In the final step the selection of electrons uses full track reconstruction seeded from the pixel hits obtained by the matching step. The muon selection for the HLT proceeds in two steps: firstly, muons are reconstructed in the muon chambers, which confirms the Level-1 decision and refines the p_T measurement using more precise information. In the second step the muon trajectories are extended into the tracker. After each step, isolation is applied to the muon candidates - the calorimeter being used after the first step and the tracker after the second [21]. The HLT p_T threshold for single muons is 19 GeV/c, for single electrons 26 GeV/c, for double muons 7 GeV/c and for double electrons 14.5 GeV/c [21]. The overall trigger efficiency for signal, DY, $t\bar{t}$ and tW backgrounds is found to be 0.822, 0.176, 0.675 and 0.684, respectively. Stronger trigger thresholds and lower efficiency on single and double electrons suppress the electron final states with respect to the muon final states. Therefore there are more muon events than electron events in the signal and background passing the trigger.

In the future also a combined e+mu trigger with symmetric thresholds of 10 GeV/c for the electron and muon will be included. No large gain is expected since events passing e+mu trigger are most probably already triggered by the single muon trigger.

4.2 Offline selection

The basic event selection is a requirement of two isolated oppositely charged leptons (one e and one μ) each with $p_T > 20$ GeV/c in the central detector acceptance region $|\eta| < 2.5$ coming from the reconstructed primary vertex. The electron candidates are required to pass electron identification cuts optimized against W+jet background as described in Ref.[22]. The electron reconstruction efficiency per electron is 0.94 for electrons coming from 200

GeV/c^2 Higgs boson with electron+muon final states, dropping to 0.91 when the electron identification cuts are applied. The leptons are defined isolated when there are no other tracks from the primary vertex with $p_T > 1$ GeV/c within a cone $\Delta R = \sqrt{\Delta\varphi^2 + \Delta\eta^2} \leq 0.4$ around the lepton. The p_T cut and isolation reduce efficiently the backgrounds with soft leptons ($pp \to b\bar{b}$,$c\bar{c}$,..). The p_T spectra for leptons in the signal events with m_A = 200 GeV/c^2 and $b\bar{b}$ background are shown in Fig. 7. Figure 8 shows the electron and muon p_T distributions separately for $b\bar{b}H/A, H/A \to \tau\tau \to e\mu + X$ with m_A = 200 GeV/c^2. The difference between the electron and muon p_T distributions comes from different trigger thresholds, which is higher for electrons.

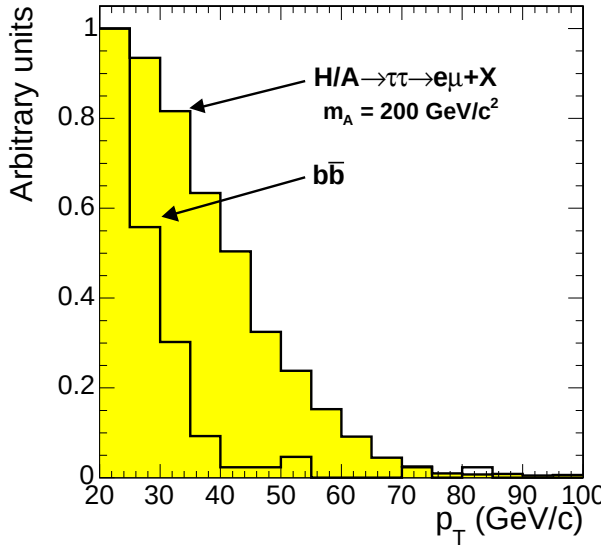


Figure 7: The lepton p_T distribution for 200 GeV/c^2 Higgs boson and for $b\bar{b}$ background.

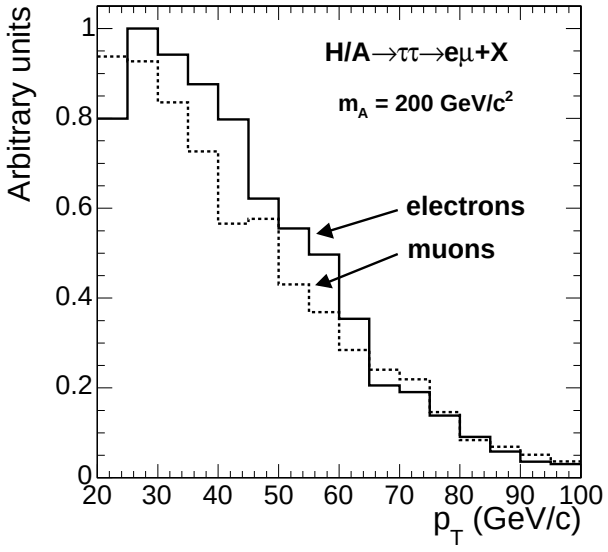


Figure 8: The electron and muon p_T distributions separately for 200 GeV/c^2 Higgs boson.

B jets associated with the Higgs boson provide a powerful tool to separate the $b\bar{b}H/A$ events from the Z/γ^* background. The Z/γ^* events are mostly produced with no significant jet activity, and the associated jets are mostly light quark and gluon jets. Therefore the Z/γ^* background can be suppressed by requiring reconstructed jets present in the event, and even further by requiring that the associated jets are identified as b jets. There are two possibilities, either to require one b tagged jet in the event and veto other jets, or to require two b tagged jets in the event. Here the 1b-tagging option is used in order to have better signal statistics as the Z/γ^* background is sufficiently suppressed. An E_T threshold of 20 GeV is used for the reconstructed jets.

B jets associated with the Higgs bosons are generally very soft, which makes their tagging a challenging task. The p_T and η distributions for b quarks in signal, $b\bar{b}Z/\gamma^*$ and $t\bar{t}$ events are shown in Figs. 9 and 10. For small jet E_T values the track multiplicity and momenta tend to be low, and many jets do not have enough significant tracks to be identified as a b jet. As a consequence the b tagging efficiency is not very high. In this study a b tagging algorithm based on the reconstruction of the secondary decay vertex of the decaying B hadron is chosen [23]. The b-Likelyhood (discriminator) of that algorithm is shown in Fig.11 for b jets in $t\bar{t}$, $b\bar{b}Z/\gamma^*$ and $b\bar{b}H/A$ events and for light quark and gluon jets in Z/γ^* events. A cut of discriminator > 0.25 gives on average 43 % b tagging efficiency per jet ($b\bar{b}H/A$) with ~2% mistagging rate ($Z/\gamma^* + \text{jet}$). Although b tagging works efficiently against backgrounds with light quark and gluon jets, it also suppresses the signal with respect to $t\bar{t}$ events, for which the b jets are more energetic, more central, and easier to reconstruct and b tag.

The $t\bar{t}$ events have more jet activity than $b\bar{b}H/A$, and a veto on additional central jets is used to suppress the $t\bar{t}$ background. Events with jets $E_T > 20$ GeV within the tracker acceptance region in addition to the one b tagged jet are rejected. The number of reconstructed jets $E_T > 20$ GeV within the tracker acceptance region coming from the primary vertex is shown in Fig. 12 for signal and $t\bar{t}$ background. Events with no jets are also shown, and events with one jet only are selected.

The τ's from a Higgs boson travel couple of mm before they decay (about 5 mm for 200 GeV/c^2 Higgs boson). Therefore the visible τ decay products, here the measured lepton tracks, are displaced relative to the primary vertex. The measurement of this displacement, the impact parameter, is possible with the innermost tracker layers made of pixel detectors expected to provide a precision of 10 - 30 μm in the transverse plane for tracks with $p_T >$

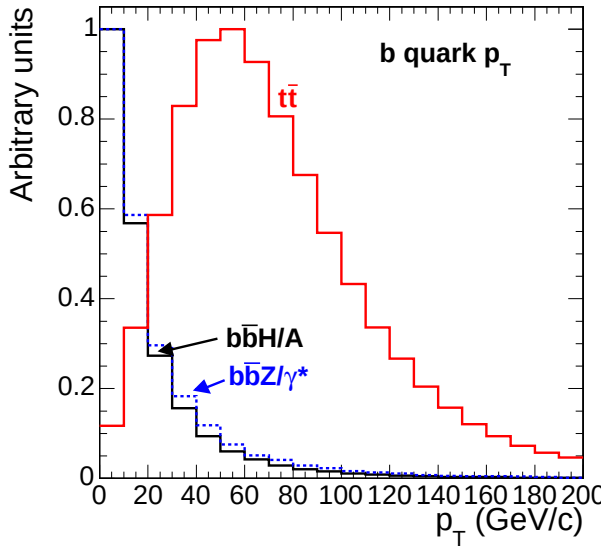


Figure 9: The p_T distributions for b quarks associated with the signal, $b\bar{b}Z/\gamma^*$ and $t\bar{t}$ events.

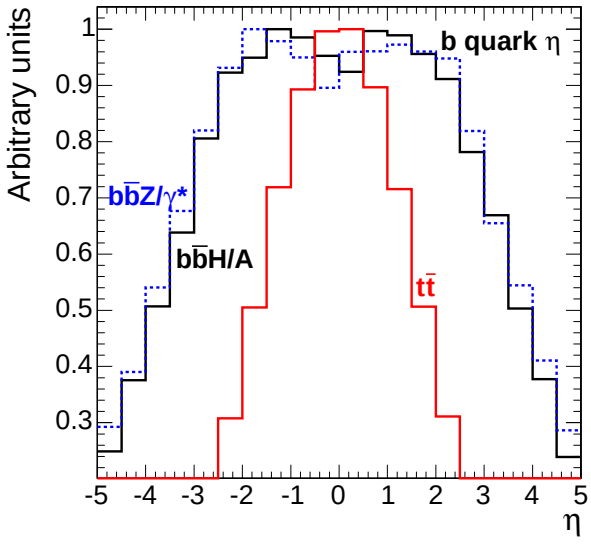


Figure 10: The η distributions for b quarks associated with the signal, $b\bar{b}Z/\gamma^*$ and $t\bar{t}$ events.

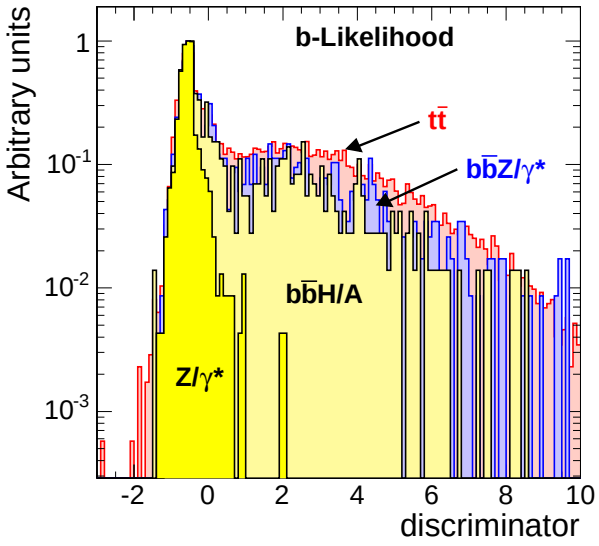


Figure 11: The b likelihood (discriminator) for individual b jets in signal, $b\bar{b}Z/\gamma^*$ and $t\bar{t}$ events, and for light quark and gluon jets in Z/γ^*.

10 GeV/c [24]. The τ lifetime can be exploited by measuring the track impact parameter. Here the transverse impact parameter with respect to detector origin in the transverse plane is used.

The independent track impact parameter measurements for the two τ's are combined quadratically into one variable $\sigma_{\rm ip} = \sigma_{\rm ip}(\tau_1) \oplus \sigma_{\rm ip}(\tau_2)$, where $\sigma_{\rm ip}(\tau_{1,2})$ are the significances of the lepton impact parameter measurements. The $\sigma_{\rm ip}$ distributions for signal and $t\bar{t}$ background are shown in Fig. 13. The combined transverse impact parameter significance is found to suppress efficiently the $t\bar{t}$ events with no genuine τ's, as the leptons originate from W boson decays. In this case impact parameter is expected only due to measurement error. However, the W boson may also decay to lepton via intermediate τ. The fraction of $t\bar{t}$ events with one intermediate τ is 39.9 %, and with two intermediate τ's 3.5 %. The fraction of $t\bar{t}$ events with two intermediate τ's cannot be suppressed by using impact parameter and it remains irreducible.

Since in $t\bar{t}$ events the leptons come from W decays, there are always neutrinos in the final state. For the signal there are neutrinos in the final state but they do not carry as much energy as in $t\bar{t}$ events. The missing E_T (MET)

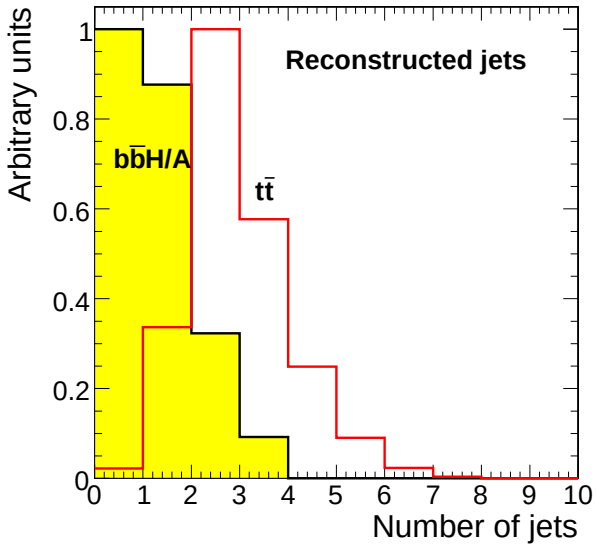


Figure 12: Number of reconstructed jets $\eta < 2.5$ in signal and $\mathrm{t\bar{t}}$ events.

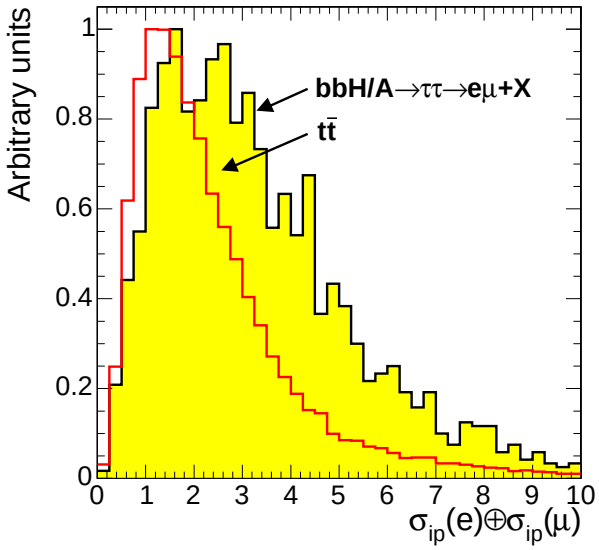


Figure 13: Combined impact parameter significance for the two leptons in signal and $\mathrm{t\bar{t}}$ events.

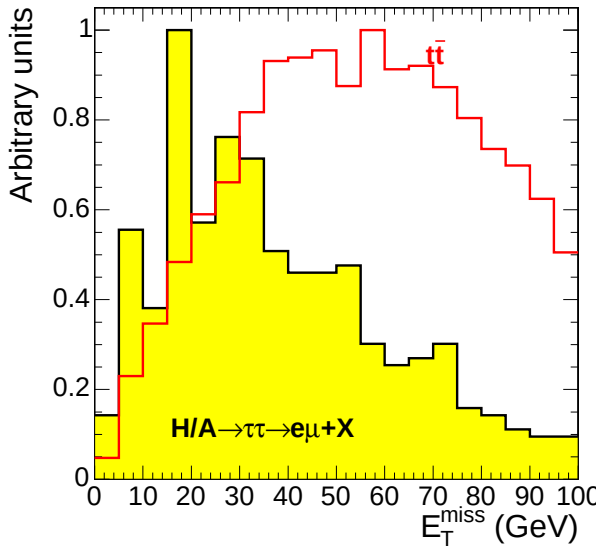


Figure 14: Missing transverse energy distribution for signal and $\mathrm{t\bar{t}}$ background.

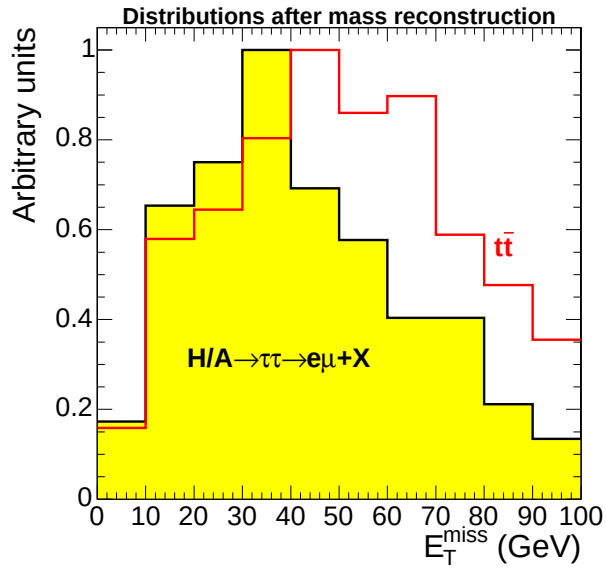


Figure 15: As in Fig. 14, but with correlations taken into account.

is reconstructed from calorimeter towers (HCAL+ECAL) and measured muons. The energy measured in the calorimeter towers is corrected with measured jet energies. Here a correction based on jet calibration made with photon+jet events[19] is used in the MET correction. Jets with uncorrected energy greater than 20 GeV are taken into account. The transverse missing energy distributions are shown in Figs. 14 and 15 for signal and $\mathrm{t\bar{t}}$ background before and after mass reconstruction. The mass reconstruction method is described in Section 4.4. Since mass reconstruction depends strongly on the neutrino collinearity, it effectively selects background events for which the neutrinos are emitted along the leptons. This changes the background MET distribution to look almost like the signal MET distribution, and no cut in missing transverse energy is used. Missing energy is, however, needed in the mass reconstruction.

Figures 16 and 17 show the angle between the two leptons in the transverse plane. This variable has potential separation power as shown in Fig. 16. The correlation with successful mass reconstruction alters these distributions and when these correlations are taken into account the signal separation from the background is difficult. A cut of $\Delta\varphi(e,\mu) < 175^o$ is used only to select events for which the mass reconstruction is possible. A stronger cut gives

better mass resolution, but with diminishing statistics.

The mass reconstruction method gives an estimate for the neutrino momenta, and only events with neutrinos emitted in the lepton direction are selected. In $t\bar{t}$ events the neutrinos are often emitted in the opposite direction. With the neutrino direction cut about 40% of signal events are lost while the suppression factor for $t\bar{t}$ and tW backgrounds is about 6.

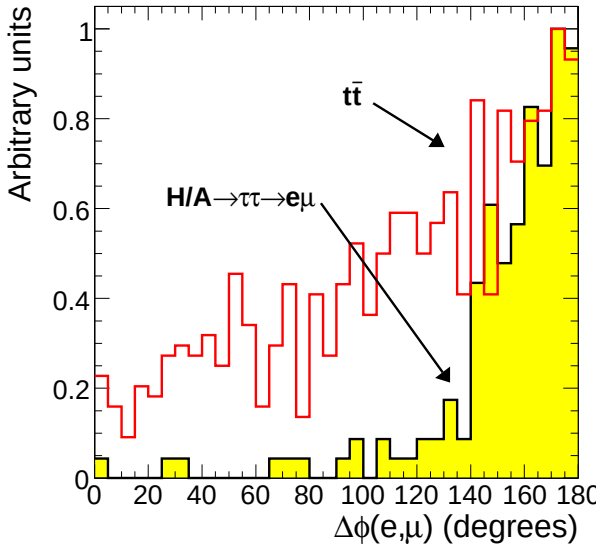


Figure 16: The angle between the two leptons in the transverse plane.

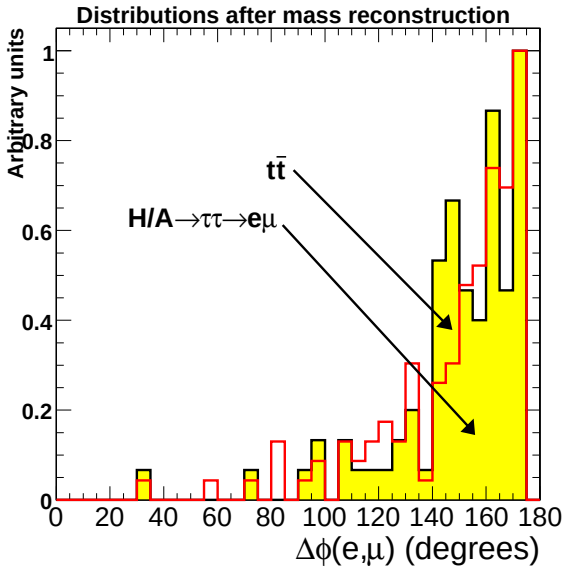


Figure 17: As in Fig. 16, but with correlations taken into account.

4.3 Optimization of cuts

The optimization of cuts has two objectives, to have a visible signal over the background, and to obtain high statistical significance. The selected cuts are chosen to be simple and not too fine tuned.

The lepton p_T cut and isolation are chosen to efficiently suppress the $b\bar{b}$ events. The lepton p_T cut 20 GeV/c is close to the trigger thresholds, in fact it is lower than the trigger threshold for single electrons. The lower than trigger

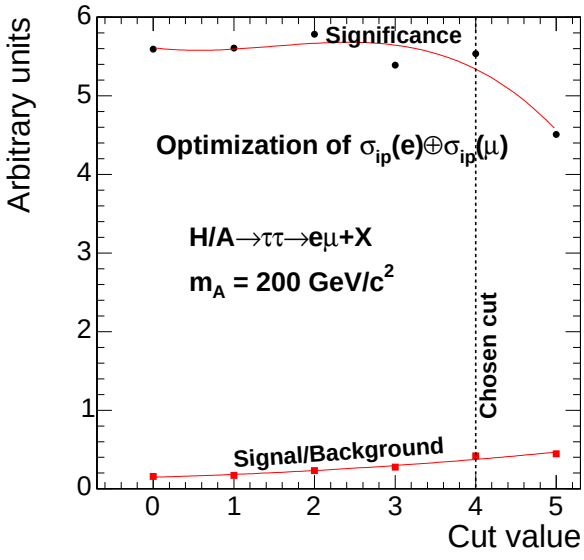


Figure 18: The signal significance and signal-to-background ratio for signal (m_A = 200 GeV/c^2) with respect to $t\bar{t}$ background for the impact parameter cut.

threshold cut is used in order to have more statistics: the lepton p_T distribution decreases exponentially, and there is no reason why one should drop events triggered by single muon trigger only, with successfully reconstructed and identified offline electron. The suppression power of the tracker isolation is optimized by varying the isolation cone and the threshold for track p_T's in the isolation cone. A strong isolation is found to keep the $b\bar{b}$ background best under control.

The τ impact parameter cut is optimized against $t\bar{t}$ background. The signal significance and the signal-to-background ratio as a function of the cut value is shown in Fig. 18. The significance remains high up to $\sigma_{ip} > 4$ while the signal-to-background ratio increases with increasing cut. The value 4 is chosen in order to have simultaneously good significance and good signal-to-background ratio. The result vary between different values of the Higgs mass due to stronger boost for larger mass, but a constant cut as a function of m_A is chosen.

B tagging is optimized by scanning the significance and signal-to-background ratio for signal and all background events. The best significance and signal-to-background ratio is achieved with a very loose discriminator cut. A loose b tagging cut may, however, introduce new potential backgrounds like W+jet with jet faking an electron. Therefore stronger b tagging is used with mistagging rate at about 1% level, as was done in earlier studies [4, 5, 6, 7, 8].

Jet veto is studied with 1b-tagging plus jet veto and 2b-tagging plus jet veto options. Requiring two b jets suppresses the signal too much with respect to the $t\bar{t}$ background. As a result only one b jet per event is selected, despite the two associated b quarks in the signal, and events with more than one jet are vetoed.

4.4 Mass reconstruction

The Higgs boson mass is reconstructed using the collinear neutrino approximation. Due to neutrinos in the final state a precise mass reconstruction is impossible. In collinear approximation the neutrinos are assumed to be emitted along the leptons, which is a valid assumption for the signal events due to large Lorentz boost of the two τ's. The missing transverse energy is projected along the lepton transverse momentum directions, giving an estimate for the neutrino momentum including the z component of the neutrino momentum. The reconstructed mass is the invariant mass of the summed lepton and neutrino 4-momenta.

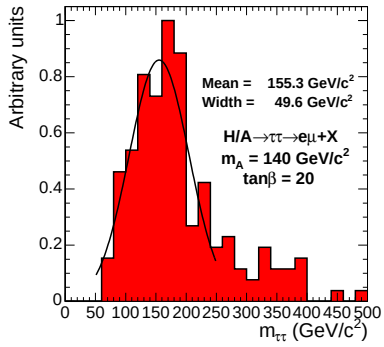


Figure 19: Reconstructed Higgs boson mass distribution for m_A = 140 GeV/c^2.

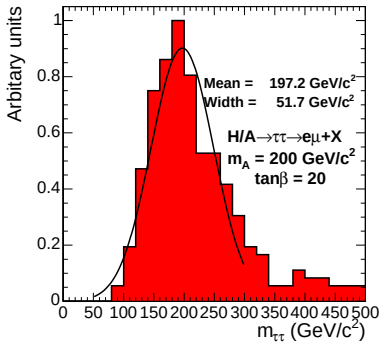


Figure 20: Reconstructed Higgs boson mass distribution for m_A = 200 GeV/c^2.

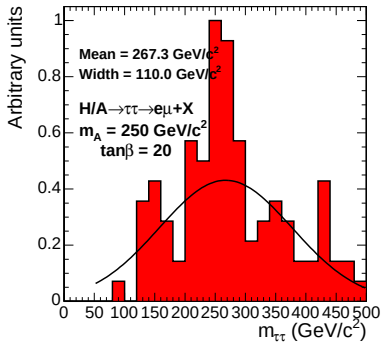


Figure 21: Reconstructed Higgs boson mass distribution for m_A = 250 GeV/c^2.

Mass reconstruction using the collinear approximation is possible when the two leptons are not in a back-to-back configuration in the transverse plane. Events in back-to-back configuration are removed with a cut $\Delta\varphi(\ell_1, \ell_2) < 175^o$, where $\Delta\varphi(\ell_1, \ell_2)$ is the angle between the two leptons in the transverse plane. The reconstructed mass is shown in Figs. 19, 20 and 21 for the signal events with m_A = 140 GeV/c^2, m_A = 200 GeV/c^2 and m_A = 250 GeV/c^2, respectively. The mass window is estimated approximately from the Gaussian fit with mean value as center and with a window of $\pm$ 1 σ width: 100 - 200 GeV/c^2 for m_A = 140 GeV/c^2, 140 - 250 GeV/c^2 for m_A = 200 GeV/c^2 and 160 - 380 GeV/c^2 for m_A = 250 GeV/c^2. The mass resolution can be improved by stronger $\Delta\varphi(\ell_1, \ell_2)$ cut, or by increasing the threshold for b jet transverse energy, but the loss in statistics is significant and they are not used in this study. The reconstructed mass superimposed on the background is shown in Figs. 22, 23 and 24, for m_A = 140 GeV/c^2, m_A = 200 GeV/c^2 and m_A = 250 GeV/c^2, respectively.

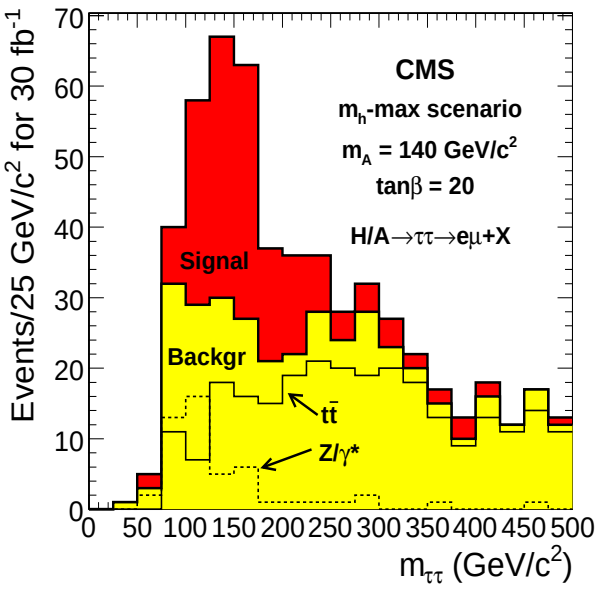


Figure 22: Reconstructed Higgs boson mass over the background for m_A = 140 GeV/c^2, tanβ = 20.

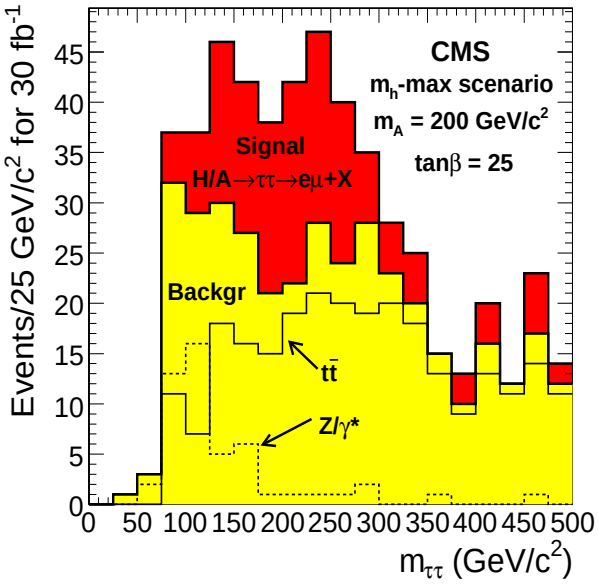


Figure 23: Reconstructed Higgs boson mass over the background for m_A = 200 GeV/c^2, tanβ = 25.

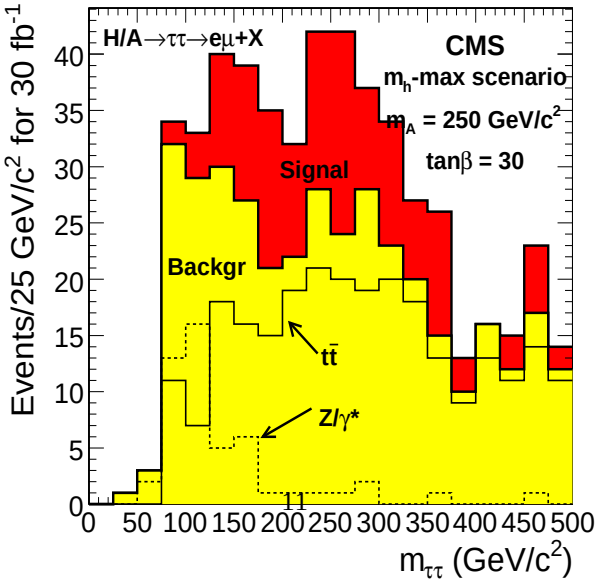


Figure 24: Reconstructed Higgs boson mass over the background for m_A = 250 GeV/c^2, tanβ = 30.

5 Signal-to-background ratios

The cuts used to extract the signal from the background and the number of events after all cuts and in the mass windows for 30fb^{-1} are listed in Tables 2 and 3. The $\mathrm{b\bar{b}}$ background in Table 2 is estimated from $\mathrm{b\bar{b}} \to \mu\mu + \mathrm{X}$ events, as no $\mathrm{b\bar{b}} \to \mathrm{e}\mu + \mathrm{X}$ datasample is yet available. Since the electron trigger efficiency, reconstruction and identification differ from the muons, the number of events in $\mathrm{b\bar{b}} \to \mu\mu + \mathrm{X}$ are not multiplied by two, but the actual factor, less than two, is estimated using Z/γ^* events decaying into electron and muon pairs. The average of double electron and double muon final state efficiency is used for e+μ final state giving factors 1.955, 1.918, 1.933, 1.262 for Level-1, HLT, primary vertex reconstruction and lepton isolation, respectively. The last number 1.262 is used to scale the rest of the cuts which do not depend on lepton reconstruction. Therefore the number of events after all cuts in $\mathrm{b\bar{b}} \to \mu\mu + \mathrm{X}$ channel $\mathrm{N_{b\bar{b}}}$ = 1.6 is scaled to 2.0 in $\mathrm{b\bar{b}} \to \mathrm{e}\mu + \mathrm{X}$ channel. The VV background consists of $\mathrm{WW} \to \ell\ell + \mathrm{X}$, $\mathrm{WZ} \to \ell\ell\ell + \mathrm{X}$ and $\mathrm{ZZ} \to \ell\ell\nu\nu$ events.

Table 2 shows the cross section times branching ratio for the backgrounds for each step of selection. The signal cross sections for $\mathrm{m_A}$ = 140, 200 and 250 GeV/c^2, tanβ = 20 is shown in Table 3, and the number of events after all cuts including the mass window. The dominating background after the selection is the $\mathrm{t\bar{t}}$ background.

	Z/γ^*	bbZ/γ^*	$\mathrm{t\bar{t}}$	tW	bb	VV
$\sigma \times \mathrm{BR}$ (pb)	233.1	3.422	86.2	6.16	36170	7.88
preselections	233	3.40	79.6	6.10	1333	7.84
Level-1 trigger	83.9	1.85	72.2	5.37	811	5.16
HLT	42.6	0.981	53.7	4.17	78.0	4.10
reconstructed primary vertex	40.8	0.952	53.3	4.11	78.1	3.92
isol e+mu,pt	1.10	0.0270	5.65	0.452	0.0378	0.288
$\mathrm{Q_e} + \mathrm{Q}_\mu = 0$	1.09	0.0268	5.62	0.451	0.0374	0.248
$\sigma_{\mathrm{ip}}(e) \oplus \sigma_{\mathrm{ip}}(\mu)$	0.296	0.00745	0.791	0.0550	0.0254	0.0255
njets>0	0.0127	0.00527	0.778	0.0509	0.00654	0.0115
btagging	0.00457	0.00289	0.608	0.0341	0.00312	0.000547
jet veto	0.00344	0.00124	0.0745	0.0166	0.000179	0.000265
$\Delta\varphi(e,\mu)$	0.00295	0.00116	0.0696	0.0159	0.000142	0.000259
E_ν >0	0.00124	0.000486	0.0119	0.00246	0.0000661	0.0000546
Events	37.1	14.6	355.8	73.7	2.0	1.6
100< $\Delta\mathrm{m}_{\tau\tau}$ < 200 GeV/c^2	21	7	56	22	1	0
140< $\Delta\mathrm{m}_{\tau\tau}$ < 250 GeV/c^2	9	1	80	18	1	0
160< $\Delta\mathrm{m}_{\tau\tau}$ < 380 GeV/c^2	8	2	160	33	1	0

Table 2: Background cross section times branching ratio (in pb) for each step of selection, and the number of events after all cuts for 30 fb^{-1}.

$\mathrm{pp} \to \mathrm{b\bar{b}H/A}$	$\mathrm{H/A} \to \tau\tau \to \mathrm{e}\mu + \mathrm{X}$		
$\mathrm{m_A}$ (GeV/c^2)	140	200	250
tanβ	20	20	20
$\sigma \times \mathrm{BR}$ (pb)	3.468	1.123	0.493
preselections	3.468	1.123	0.493
Level-1 trigger	3.238	1.079	0.479
HLT	2.585	0.923	0.419
reconstructed primary vertex	2.434	0.866	0.395
isol e+mu,pt	0.258	0.116	0.0613
$\mathrm{Q_e} + \mathrm{Q}_\mu = 0$	0.256	0.116	0.0612
$\sigma_{\mathrm{ip}}(e) \oplus \sigma_{\mathrm{ip}}(\mu)$	0.0859	0.445	0.0260
njets>0	0.0375	0.0216	0.0130
btagging	0.0177	0.0104	0.00649
jet veto	0.0115	0.00619	0.00390
$\Delta\varphi(e,\mu)$	0.0106	0.00554	0.00351
E_ν >0	0.00601	0.00340	0.00222
Number of signal events in mass window	118	54	40

Table 3: Signal cross section times branching ratio for $\mathrm{m_A} = 140, 200, 250\,\mathrm{GeV}/c^2$ for each step of selection, and the number of events after all cuts for 30 fb^{-1}.

6 Systematic uncertainty on background determination

The uncertainty of the event selection efficiency is related to the uncertainty of the lepton identification, the absolute calorimeter scale and b tagging efficiency. An error in calorimeter scale introduces an error in the jet energy and MET measurement. Here a 3% error on jet energy scale and a 10% error on MET scale is assumed. Varying the jet and MET energy scales by $\pm$3% and $\pm$10% leads to an average 6.0% uncertainty on the number of Z/γ^* events, a 4.3% uncertainty on the number of $\mathrm{b\bar{b}Z}/\gamma^*$ events, a 7.3% uncertainty on the number of $\mathrm{t\bar{t}}$ events and 11.3% uncertainty on the number of tW events and 3.7% uncertainty on the number of $\mathrm{b\bar{b}}$ events passing the event selection cuts. The dominant contribution comes from $\mathrm{t\bar{t}}$, other backgrounds have only a small effect. The contribution to the total background and systematic errors from $\mathrm{b\bar{b}}$ and VV background is so small that it can be safely neglected.

The uncertainty of the b tagging efficiency can be estimated from $\mathrm{t\bar{t}}$ events as in Ref.[25]. A value of 5% is used as a conservative estimate. The mistagging uncertainty is also assumed to be 5% [26]. In Ref.[27] a lepton identification uncertainty of 1% is used for $\mathrm{H} \rightarrow \mathrm{ZZ} \rightarrow 4\mu$ and $\mathrm{H} \rightarrow \mathrm{ZZ} \rightarrow 4\mathrm{e}$. Here a conservative lepton identification uncertainty of 2% is used for both electrons and muons. The uncertainty of the two leptons are added linearly due correlations on errors, giving lepton identification uncertainty of 4% for the two leptons.

The mass resolution for $\mathrm{H/A} \rightarrow \tau\tau \rightarrow \mathrm{e}\mu + \mathrm{X}$ channel is at best about 25%, and the signal and the Z peak cannot be easily separated, as shown in Figs. 22-24. Therefore fitting the signal+background distribution is difficult, and it is not used here to determine the background uncertainty. The background within the mass window consists of mostly $\mathrm{t\bar{t}}$ events. The uncertainty of the $\mathrm{t\bar{t}}$ background is estimated from the theoretical uncertainty of the $\mathrm{t\bar{t}}$ cross section as in Ref.[28]. The theoretical NLO cross section calculation uncertainty for $\mathrm{t\bar{t}}$ events consists of uncertainty due to scale variation and the choise of PDF. The scale uncertainty is taken to be at the level of 5% according to Ref.[29]. The recent results on PDF uncertainty suggests an uncertainty of $\sim$ 2.5 %. Thus a value of 5.6 % is assumed in this study for the uncertainty of $\mathrm{t\bar{t}}$ background. This uncertainty is also assumed for the tW background. The uncertainty of the Z/γ^* cross section is expected to be of the order of 1% [30]. The uncertainty of the $\mathrm{b\bar{b}Z}/\gamma^*$ cross section measurement is studied in Ref.[31]. The uncertainty obtained in this study, 14.2% excluding the luminosity uncertainty, is used for the $\mathrm{b\bar{b}Z}/\gamma^*$.

The systematic uncertainty of the above measurements including the luminosity uncertainty of 3%[19] is shown in Table 4. The variation of the luminosity and it's effect on pile-up is not taken into account.

The total systematic uncertainty of the background is estimated by summing the uncertainties on the number of events from each background channel contributing to the background within the mass windows. Although the number of background events within different mass windows vary, the total uncertainty is almost independent of the chosen mass window as the dominant contribution comes always from $\mathrm{t\bar{t}}$: a systematic uncertainty of 11.2%, 11.9% and 12.0% is obtained for Higgs masses 140 GeV/c^2, 200 GeV/c^2 and 250 GeV/c^2, respectively.

	Uncertainty (%)			
	Z/γ^*	bbZ/γ^*	$\mathrm{t\bar{t}}$	tW
Calorimeter scale	6.0	4.3	7.3	11.3
b tagging	-	5	5	5
mistagging	5	-	-	-
lepton identification	4	4	4	4
cross section	1	14.2	5.6	5.6
luminosity	3	3	3	3
total	9.3	16.4	11.6	14.5

Table 4: Systematic uncertainties for the main background channels.

7 Discovery potential

Table 5 shows the number of signal+background events (measurable variable) and the number of background events (variable to be estimated with Monte-Carlo) for 30 fb^{-1} for three different Higgs boson masses. The mass window $\Delta\mathrm{m}_{\tau\tau}$ is chosen to match the expected Gaussian width of the signal distributions. The statistical significance is calculated with Poisson statistics using program ScPf [32] which allows one to include the systematic uncertainties in the significance determination.

The discovery potential of the $\mathrm{H/A} \rightarrow \tau\tau \rightarrow \mathrm{e}\mu + \mathrm{X}$ channel is shown in Fig. 25 for 30 fb^{-1}. The old fast

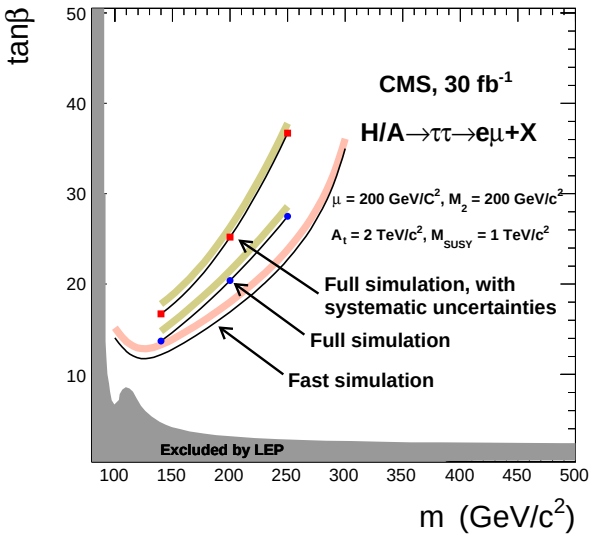


Figure 25: The 5σ-discovery reach for heavy neutral Higgs bosons H and A decaying via $\tau\tau$ to e+μ final state. The dots give the 5σ limit for the studied values of Higgs mass. The fast simulation result is also shown.

simulation results [5] are shown for comparison. The main difference comes from updated signal and background cross sections and worse than expected MET and mass resolution with more background events in the mass window. The cuts are, however, more optimized. The uppermost curve shows the discovery region with the effect of the systematic background uncertainty taken into account. The systematic uncertainties decrease the explorable region.

	$\Delta m_{\tau\tau}$	N_S+N_B (meas.)	N_B (MC)	Significance	
				ΔN_B^{syst} = 0%	$\Delta N_B^{syst} \sim 12\%$
m_A = 140 GeV/c^2, tanβ = 20	100 - 200 GeV/c^2	225	107	9.9	7.3
m_A = 200 GeV/c^2, tanβ = 20	140 - 250 GeV/c^2	163	109	4.8	3.1
m_A = 250 GeV/c^2, tanβ = 20	160 - 380 GeV/c^2	244	204	2.7	1.4

Table 5: Number of measured events in given mass windows and estimated number of background events for 30 fb^{-1} and statistical significance calculated with Poisson statistics.

8 Conclusions

The Higgs boson discovery potential in $H/A \to \tau\tau \to e\mu + X$ channel with Higgs boson produced in association with two b quarks is studied with full simulation. The tagging of one associated b jet is used to suppress the backgrounds. The Higgs boson mass is reconstructed exploiting the collinear neutrino approximation. The 5σ coverage is updated and the effect of systematic errors included. The systematic uncertainties decrease the discovery reach. The largest component in the systematic errors is the uncertainty of the calorimeter energy scale.

The best strategy to improve these results is to improve the MET measurement. The mass resolution depends strongly on the quality of the MET measurement. Smaller width of the mass peak would decrease the number of background events within the mass window. It would also allow separating the Higgs mass peak from the Z peak, and make fitting the signal and background a realistic possibility. The mass width can, however, be improved by selecting events with harder b jets as it results in better defined MET. The signal, for which the associated jets are soft, would be suppressed though, and optimization gives better results for low jet E_T cut. For the same reason no strong $\Delta\varphi(e,\mu)$ cut is used, although improving the mass resolution that way is possible.

The optimization of b tagging gave best results with very loose discriminator cuts. Despite the potential new backgrounds, this is hinting that a Higgs search strategy with no b tagging may be possible. If events with no associated jets are selected, the $t\bar{t}$ background is efficiently suppressed while the Z/γ^* is enhanced. The signal would consist of Higgs bosons produced via loop mediated gluon fusion in addition to associated production $gg \to b\bar{b}H/A$ with no successfully reconstructed jets. This approach remains yet to be studied.

Acknowledgements

The author would like to thank the ARDA/ASAP team for a magnificent grid tool they have created, and M. Chertok, R. Kinnunen, F. Moortgat and I. Tomalin for helpful discussions and reading the manuscript.

References

[1] ALEPH, DELPHI, L3 and OPAL Collaborations, hep-ex/0602042.

[2] M. Carena, S. Heynemeyer, C. Wagner, and G. Weiglein, hep-ph/9912223.

[3] CDF Collaboration, D0 Collaboration, and Tevatron Electroweak Working Group, hep-ex/0507091.

[4] S. Abdullin, S. Banerjee, L. Bellucci, C. Charlot, D. Denegri, M. Dittmar, V. Drollinger, M. Dubinin, M. Dzelalija, D. Green, I. Iashvili, V. Ilyin, R. Kinnunen, S. Kunori, K. Lassila-Perini, S. Lehti, K. Mazumdar, F. Moortgat, T. Muller, A. Nikitenko, I. Puljak, P. Salmi, C. Seez, S. Slabospitsky, S. Stepanov, R. Vidal, W. Wu, H. Yildez, and M. Zeyrek, *CMS Note 2003/006*.

[5] S. Lehti, *CMS Note 2002/035*.

[6] A. Nikitenko and R. Kinnunen, *CMS Note 2003/006*.

[7] D. Denegri and R. Kinnunen, *CMS Note 1999/037*.

[8] R. Kinnunen and A. Nikitenko, *CMS Note 1997/106*.

[9] S. Heinemeyer, G. Weiglein, and W. Hollik, *Comput.Phys.Commun.124:76-89,2000,hep-ph/9812320; Eur.Phys.J.C9:343-366,1999,hep-ph/9812472;Site located at http://www.feynhiggs.de*.

[10] T. Sjostrand, L. Lonnblad, S. Mrenna, and P. Skands, *LU TP* **03-38** (2003) [hep-ph/0308153].

[11] S. Jadach, Z. Was, R. Decker, and J. Kuhn, *Comp. Phys. Commun.* **76** (1993) 361.

[12] J. Campbell, *Private communication*.

[13] *MCFM - Monte Carlo for FeMtobarn processes. Site located at http://mcfm.fnal.gov*.

[14] A. Pukhov, E. Boos, M. Dubinin, V. Edneral, V. Ilyin, D. Kovalenko, A. Kryukov, V. Savrin, S. Shichanin, and A. Semenov, *INP-MSU 98-41/542* [hep-ph/9908288].

[15] S. Slabospitsky and L. Sonnenschein, *Comput. Phys. Commun.* **148** (2002) 87, [hep-ph/0201292].

[16] *CERN Yellow Report* (2000) [hep-ph/0003033].

[17] *CERN Program Library Long Writeup* **W5013** (1994).

[18] *ORCA: CMS Reconstruction Package. Site located at http://cmsdoc.cern.ch/orca*.

[19] *CERN/LHCC 2006-001* **CMS TDR 8.1** (2006).

[20] *CERN/LHCC 2000-038* **CMS TDR 6.1** (2000).

[21] *CERN/LHCC 2002-26* **CMS TDR 6.2** (2002).

[22] G. Davattz, M. Dittmar, and A.-S. Giolo-Nicollerat, *CMS Note 2006/047*.

[23] C. Weiser, *CMS Note 2006/014*.

[24] *CERN/LHCC 98-6* **CMS TDR 5** (1998).

[25] S. Lowette, J. D'Hondt, and J. Heyninck, *CMS Note 2006/013*.

[26] D. Bloch, *Private communication*.

[27] *CERN/LHCC 2006-002* **CMS TDR 8.2** (2006).

[28] M. Baarmand, M. Hashemi, and A. Nikitenko, *CMS Note 2006/056*.

[29] M. Duhssen, hep-ph/0303092.

[30] S. Haywood *et. al.*, hep-ph/0003275.

[31] S. Lehti, *CMS Note 2006/099*.

[32] S. Bityakov, S. Erofeeva, N. Krasnikov, and A. Nikitenko, *Proceedings of PhyStat 2005, September 2005, Oxford, UK*.