CMS Physics Analysis Summary

Contact: cms-pag-conveners-higgs@cern.ch 2019/11/29

Search for a light charged Higgs boson in the $H^{\pm} \to cs$ channel at 13 TeV

The CMS Collaboration

Abstract

A search is conducted for a low-mass charged Higgs boson produced in a top quark decay and subsequently decaying into a charm and an antistrange quark. The data sample was recorded in proton-proton collisions at $\sqrt{s} = 13$ TeV by the CMS experiment at the LHC and corresponds to an integrated luminosity of 35.9 fb^{-1}. The signal search is conducted in the process of top-quark pair production, where one top quark decays to a bottom quark and a charged Higgs boson, and the other to a bottom quark and a W boson. With the W boson decaying to a charged lepton (electron or muon) and a neutrino, the final state comprises an isolated lepton, missing transverse momentum, and at least four jets, of which two are tagged as b jets. To enhance the search sensitivity, one of the jets originating from the charged Higgs boson is required to satisfy a charm tagging requirement. No significant excess beyond standard model predictions is found in the dijet invariant mass distribution. An upper limit in the range 0.20−1.65% is set on the branching fraction of the top quark decay to the charged Higgs boson and bottom quark for a Higgs mass between 80 and 160 GeV.

The discovery of the Higgs boson in 2012 by the CMS [1] and ATLAS [2] experiments at the CERN LHC has ushered in a new beginning in the field of particle physics. The Higgs boson could be the first of many elementary scalars present in nature to be observed in the laboratory. Various extensions to the standard model (SM), such as supersymmetry [3–5] and the two Higgs doublet model (2HDM) [6], predict multiple scalars as the remnants of an additional $SU(2)_L$ complex doublet introduced to address some known limitations of the SM. After spontaneous symmetry breaking, out of the eight degrees of freedom of the two Higgs doublets, three are used to make the W and Z bosons massive, leaving five physical scalar particles. Of these, two (h, H) are neutral Higgs bosons which are CP-even (scalar), one (A) is neutral and CP-odd (pseudoscalar), and the remaining two are charged Higgs bosons $H^{\pm}$.

The 2HDM can be classified into different categories depending on the type of interaction of the two doublets with quarks and charged leptons. For example, in the type I 2HDM, fermions have Yukawa couplings only to the second doublet. The nature of the Yukawa coupling determines the branching fraction of the charged Higgs boson decays into different final states. We are interested in the decay channel $H^+ \to c\bar{s}$ (and its charge conjugate), whose branching fraction can range up to 100%, depending on the type of Yukawa couplings. The latter is expressed in terms of the parameter $\tan\beta = v_2/v_1$, where v_1 and v_2 are the vacuum expectation values of the two Higgs doublets. In the minimal supersymmetric standard model (MSSM), this is the dominant decay channel for low values of $\tan\beta$ [7]. In this analysis, we assume that $\mathcal{B}(H^+ \to c\bar{s}) = 100\%$.

There have been many earlier searches for charged Higgs bosons at LEP, the Tevatron, and the LHC. At LEP, it was expected to be dominantly produced by the process $e^+e^- \to H^+H^-$. The search was conducted by assuming that H^+ decays only to $c\bar{s}$ and $\tau\overline{\nu}_\tau$. Assuming that the sum of the branching fractions $\mathcal{B}(H^+ \to \tau^+\nu_\tau) + \mathcal{B}(H^+ \to c\bar{s}) = 1$, a lower limit of 79.3 GeV was obtained for the charged Higgs mass at 95% confidence level (CL) [8–10]. Under a more general assumption $\mathcal{B}(H^+ \to \tau^+\nu_\tau) + \mathcal{B}(H^+ \to q\overline{q}') = 1$, a slightly less stringent constraint of 76.3 GeV was obtained at 95% confidence level [11].

Limits on charged Higgs production at hadron colliders have been set by the Tevatron and LHC, assuming the production mode $t \to H^+b$. The CDF collaboration [12] set a 95% CL upper limit on the branching fraction $\mathcal{B}(t \to H^+b)$ of 10–30% for a charged Higgs lying in the mass range 60–150 GeV, assuming that H^+ decays dominantly to $c\bar{s}$. Similar limits have been obtained by the D0 experiment [13]. Using 7 TeV data, the ATLAS Collaboration set an upper limit at 95% CL on the product $\mathcal{B}(t \to H^+b)\mathcal{B}(H^+ \to \tau^+\nu)$ of 0.23–1.3% for a charged Higgs mass in the range 80–160 GeV [14]. A search for a charged Higgs boson decaying into $c\bar{s}$ was performed with 8 TeV data by the CMS Collaboration, which set an upper limit at 95% CL on $\mathcal{B}(t \to H^+b)$ in the range 1.2–6.5% [15].

As illustrated in Fig. 1, in the signal process for H^+ production, one of the top quarks decays to H^+b and the other to $W^-\overline{b}$, with H^- production proceeding by the charge conjugate of this process. The principal SM background to this search consists of $t\bar{t}$ pair production where one of the top quarks decays by $t \to W^+b$ and the other top quark decays by $\bar{t} \to W^-\overline{b}$. The W^+/H^+ decays hadronically into "light" jets (not from a b quark), whereas the W^- decays leptonically (in the $t\bar{t}$ case, this is called the "semileptonic" decay channel); we define two channels depending on whether the lepton produced in the W^- decay is an electron or a muon (events with tau leptons are not considered).

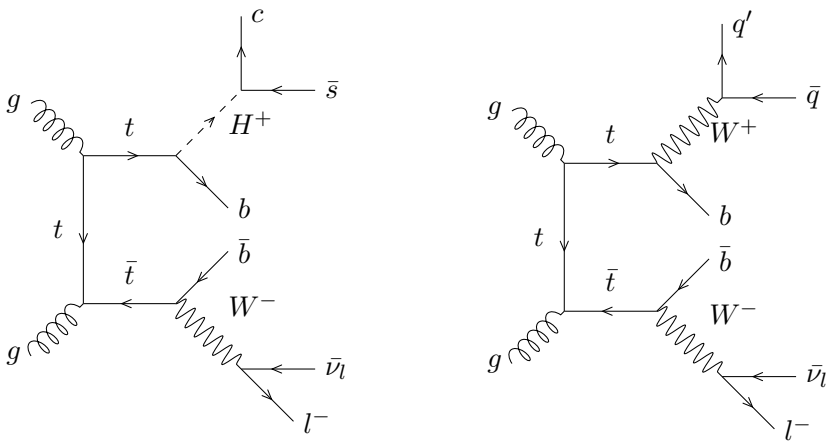

Figure 1: Production of $t\bar{t}$ from gluon-gluon scattering. The left plot shows the signal process in which the $t\bar{t}$ pair decay products include a charged Higgs boson. The right plot shows the SM decay of the $t\bar{t}$ pair in the semileptonic decay channel.

1 The CMS detector

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T. Within the solenoid volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter (HCAL), each composed of a barrel and two endcap sections. Silicon pixel and tracker detector identifies the trajectory of charged particles and accurately measures their transverse momentum up to $|\eta| \leq 2.5$. Forward calorimeters extend the pseudorapidity coverage provided by the barrel and endcap calorimeter. Segmented calorimeters provide sampling of electromagnetic and hadronic showers up to $|\eta| \leq 5$. Muons are detected in gas-ionization chambers embedded in the steel flux-return yoke outside the solenoid in the range of $|\eta| \leq 2.4$. Events of interest are selected using a two-tiered trigger system [16]. The first level (L1), composed of custom hardware processors, uses information from the calorimeters and muon detectors to select events at a rate of around 100 kHz within a time interval of less than 4 μs. The second level, known as the high-level trigger (HLT), consists of a farm of processors running a version of the full event reconstruction software optimized for fast processing, and reduces the event rate to around 1 kHz before data storage. A more detailed description of the CMS detector, together with a definition of the coordinate system used and the relevant kinematic variables can be found [17].

2 Data and simulation

The data used for the analysis was collected by the CMS detector in 2016 in proton-proton (pp) collisions at $\sqrt{s} = 13$ TeV, with an integrated luminosity of 35.9 fb^{-1}. The events selected online require, at the L1 trigger level, a muon candidate with $p_{\mathrm{T}} > 22$ GeV or an electron or photon candidate with $p_{\mathrm{T}} > 30$ GeV, or 22 GeV if it is isolated; at the HLT level, an isolated muon (electron) with $p_{\mathrm{T}} > 24$ (27) GeV is required.

As shown in Fig. 1, the charged Higgs is assumed to decay into $c\bar{s}$ or $\bar{c}s$ only. As a result, in the final state, there will be four jets (two b jets, one c jet, one s jet), one lepton (e or μ; τ is not considered in this analysis) and missing transverse momentum ($p_{\mathrm{T}}^{\mathrm{miss}}$), which is attributed to the neutrino. The SM processes that give the same final states (four jets + one lepton + missing

transverse momentum) are considered as background channels for this analysis. The simulated signal and background samples are generated using the MADGRAPH5_aMC@NLO [18] and POWHEG v2 [19–21] generators at parton level. In all cases, these parton level events are hadronized using PYTHIA 8 [22] with the CUETP8M1 [23] tune, and then passed to GEANT4 [24] for simulation of the CMS detector response. Finally, the events are reconstructed after complete detector simulation using the same reconstruction process as for data.

The SM $\mathrm{t\bar{t}}$ + jets process is an irreducible background, and represents the largest contribution, about 94% of the total expected background in the signal region. The parton level SM $\mathrm{t\bar{t}}$ + jets events are generated using POWHEG v2 at next-to-leading order (NLO) [25]. The NNPDF30_nlo_as_0118 [26] PDF set is used for this purpose. The events are hadronized and simulated as above. The next-to-NLO cross-section for $\mathrm{t\bar{t}}$ + jets is estimated to be $831.76\,^{+20}_{-29}$ (scale) ± 35 (PDF + α_S) pb [27]. For simplicity, the top quark mass is taken to be 172.5 GeV.

The charged Higgs signal samples are generated using MADGRAPH5_aMC@NLO. Only $\mathrm{H^+}$ samples are generated, and $\mathrm{H^-}$ production is assumed to be the same. The signal sample is generated for several mass points in the range of 80 to 160 GeV (80, 90, 100, 120, 140, 150, 155, and 160 GeV). The cross section for the signal is 831.76×0.21 pb, where 831.76 pb is the inclusive SM $\mathrm{t\bar{t}}$ + jets production cross section and the factor of 0.21 is the branching fraction of $\mathrm{W^-} \to \ell^- \overline{\nu}_\ell$ (where ℓ = e or μ) [28]. Furthermore, the signal events are scaled by a factor of $2 \times 0.065 \times (1 - 0.065) = 0.12$, where 6.5% is the maximum observed upper limit on $\mathcal{B}(\mathrm{t} \to \mathrm{H^+b})$ obtained at 8 TeV [15].

The single top quark production processes, where a top quark is produced with jets in the s-channel, t-channel, or tW-channel, can also mimic the signal topology. The s-channel single top production samples are generated using MADGRAPH5_aMC@NLO [18], while the t-channel and tW-channel samples are generated using POWHEG [29, 30]. The production of W and Z bosons with jets, and vector boson fusion, are also considered as background processes. The inclusive W + jets and Z/γ + jets samples are generated using MADGRAPH5_aMC@NLO, with the MLM technique [31] used to avoid double counting. The vector boson fusion samples (WW/WZ/ZZ, collectively referred to as "VV") are generated using PYTHIA 8.

Furthermore, SM events containing only jets produced through the strong interaction, referred to as quantum chromodynamics (QCD) multijet events, can also produce a final state identical to the signal topology, even though these events contain only quarks at the parton level. QCD multijet events can have reconstructed leptons from the misidentification of bottom and charm hadrons, and $p_{\mathrm{T}}^{\mathrm{miss}}$ due to the poor measurement of hadronic activity inside the CMS detector. The simulation of QCD multijet events is computationally intensive, resulting in a limited number of such events being available. Because of this lesser number of events compared to the cross section, the statistical uncertainty is high. Also, the simulated QCD multijet events are not well modeled for high jet multiplicities. In view of this, a data-driven approach is used to make a more precise estimation of the QCD multijet background.

With the exception of the QCD multijet background, the expected yield for each background process is determined from simulation. For the QCD multijet background, a method known as the "ABCD method" is used. First, a normalization is determined from the (low $p_{\mathrm{T}}^{\mathrm{miss}}$, isolated) and (low $p_{\mathrm{T}}^{\mathrm{miss}}$, anti-isolated) regions; then, the QCD shape is determined from the (high $p_{\mathrm{T}}^{\mathrm{miss}}$, anti-isolated) region. By using the normalization obtained on the shape, the expected QCD multijet contribution is determined in the signal region (high $p_{\mathrm{T}}^{\mathrm{miss}}$, isolated). The low and high $p_{\mathrm{T}}^{\mathrm{miss}}$ regions are defined by $p_{\mathrm{T}}^{\mathrm{miss}} < 20$ GeV and $p_{\mathrm{T}}^{\mathrm{miss}} > 20$ GeV, respectively. In the muon channel, the isolated and anti-isolated regions are defined by $I_{\mathrm{rel}}^{\mu} < 0.15$ and $0.15 < I_{\mathrm{rel}}^{\mu} < 0.4$, where I_{rel} is the relative isolation variable. The relative isolation of a lepton is defined as the

ratio of the sum of p_T for all the other particles to the p_T of the corresponding lepton within a cone of $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.4$ around the lepton. For the electron channel, the isolated region corresponds to $I^e_{\text{rel}} < 0.08$ (0.07) and the non-isolated region to $0.08 < I^e_{\text{rel}} < 0.3$ (0.07 $< I^e_{\text{rel}} < 0.3$) for electrons in the barrel (endcap) regions.

3 Event reconstruction

The physics objects of interest are mainly leptons, jets, missing transverse momentum, vertices of pp collisions, and displaced vertices from the decay of bottom or charm hadrons. The particle flow (PF) algorithm [32] is used to reconstruct these objects by optimally using various subsystems of the CMS detector.

The collision vertices are obtained using reconstructed tracks in the silicon tracker. First, candidate vertices are obtained by clustering tracks using the deterministic annealing algorithm [33]. Subsequently, candidate vertices with at least two tracks are fitted using the adaptive vertex fitter [34]. A primary vertex associated with a hard interaction is expected to be accompanied by a large number of tracks.

The reconstructed vertex with the largest value of summed physics-object p_T^2 is taken to be the primary pp interaction vertex. The physics objects are the jets, clustered using the jet finding algorithm [35, 36] with the tracks assigned to the vertex as inputs, and the associated missing transverse momentum, taken as the negative vector sum of the p_T of those jets. Further, the reconstructed primary vertex is required to be within 24 cm along the beam axis and within 2 cm in the transverse direction from the nominal pp interaction region.

Electrons are reconstructed using the PF algorithm [32] based on the tracks in the tracker and energy deposits in the electromagnetic calorimeter [37]. The reconstructed trajectory in the tracker is mapped to the energy deposit in the ECAL to form an electron candidate. The bending direction of the trajectory in the tracker is used to identify the charge of an electron.

Muons, being minimum ionising particles, can traverse a long distance in the CMS detector. The trajectory of the muon is bent due to the presence of a strong magnetic field inside the solenoid and the return magnetic field in the opposite direction outside the solenoid. Muon candidates are identified in the muon detectors and matched to tracks measured in the silicon tracker, resulting in an excellent p_T resolution between 1 and 10% for p_T values up to 1 TeV.

Due to color confinement [38], the quarks and gluons produced in pp collisions cannot exist in free states; instead, they produce a cluster of colorless hadrons, which subsequently decay to leptons and photons. Jets are clustered from the reconstructed particle-flow candidates using the anti-k_T algorithm [35, 36] with a distance parameter of $\Delta R = 0.4$. Each jet is required to pass dedicated quality criteria to suppress the impact of instrumental noise and misreconstruction. Additional proton-proton interactions within the same or nearby bunch crossings (pileup) can contribute additional tracks and calorimetric energy depositions, increasing the apparent jet momentum. To mitigate this effect, tracks identified to be originating from pileup vertices are discarded and an offset correction is applied to correct for remaining contributions. Jet energy corrections are derived from simulation studies so that the average measured response of jets becomes identical to that of particle level jets. In situ measurements of the momentum balance in dijet, γ + jet, Z + jet, and multijet events are used to determine any residual differences between the jet energy scale in data and in simulation, and appropriate corrections are made [39].

The missing transverse momentum vector $\vec{p}_T^{\,\text{miss}}$ is defined as the projection onto the plane perpendicular to the beam axis of the negative vector sum of the momenta of all reconstructed

particle-flow objects in an event. Its magnitude is referred to as p_T^{miss}. Neutrinos, being weakly interacting particles with a very low cross section, cannot be directly detected by the CMS detector and thus contribute to p_T^{miss}. The p_T^{miss} in the simulation and data can have different resolutions. Therefore, in simulated events the p_T^{miss} value is smeared by propagating the jet energy corrections to p_T^{miss}.

There are two b jets in the final state illustrated in Fig. 1 in both the charged Higgs signal process and the SM $t\bar{t}$+jets background. An accurate identification of b jets substantially reduces the SM backgrounds from other processes such as $Z/\gamma +$ jets, VV, $W +$ jets, etc. The combined secondary vertex (CSVv2) method [40] is used to tag a b jet. The algorithm combines information on track impact parameters and secondary vertices within a jet into an artificial neural network classifier that provides separation between a b jet and jets of other flavors. As the charged Higgs boson decays to a charm and an antistrange quark, the identification of charm jets is expected to increase the signal significance. A charm tagger has been recently developed [41] for 13 TeV data. It is based on the CSVv2 method and works similarly to the b tagging procedure.

4 Event selection

In the event topology of interest, there are four jets (two b jets and two light jets), one charged lepton, and p_T^{miss}. Various selection requirements are applied to ensure the resulting events have this topology. In the offline analysis, events that pass the previously described triggers and contain a muon (electron) with $p_T > 26$ (30) GeV and $|\eta| < 2.4$ (2.5) are selected. The signal event topology has only one lepton, so events having a second lepton with $p_T^{\ell} > 15$ GeV are rejected. To eliminate events where the lepton is found within a jet, a requirement on the relative isolation is used. This requirement is $I_{rel}^{\mu} < 0.15$ for muons, and $I_{rel}^{e} < 0.08(0.07)$ for electrons in the barrel (endcap) region. No charge requirement is applied to the lepton.

Jets are selected by requiring $p_T^{j} > 25$ GeV, $|\eta^{j}| < 2.4$, neutral hadron energy fraction < 0.99, neutral electromagnetic energy fraction < 0.99, number of constituents > 1, charged hadron energy fraction > 0, charged hadron multiplicity > 0, and charged hadron electromagnetic energy fraction < 0.99, where more details about these can be found in Ref. [32]; a total of at least four jets is required. The p_T^{miss} is required to be greater than 20 GeV. The events are required to have at least two b jets satisfying the medium b tagger working point (discriminant value > 0.8484) [40]. The corresponding b tagging efficiency is 63% and the probability of a light jet being misidentified as a b jet is 1%.

5 Dijet invariant mass (m_{jj}) distribution

In this analysis, the charged Higgs boson is assumed to decay to $c\bar{s}$ or $\bar{c}s$. The invariant mass of the system of the two light jets (m_{jj}) is thus used as the final observable. The m_{jj} distribution of the two highest p_T light jets is shown in the top row of Fig. 2 for the two channels. If the two observed light jets come from a semileptonic $t\bar{t}$ decay, then the m_{jj} distribution should have a peak at the W boson mass. However, the observed mean of the m_{jj} distribution is much higher (around 138 GeV), reflecting the fact that the two light jets in each event may not necessarily come from the decay of a W boson. In addition, the m_{jj} distribution has a long tail, which might constrain the search for new resonances in the dijet mode.

To select true semileptonic $t\bar{t}$ events, a kinematic fit (KF) is performed on the reconstructed objects using the top kinematic fitter package [42]. The top kinematic fitter takes physics objects such as leptons, jets, p_T^{miss}, and their resolutions as input, and gives improved four-vectors of

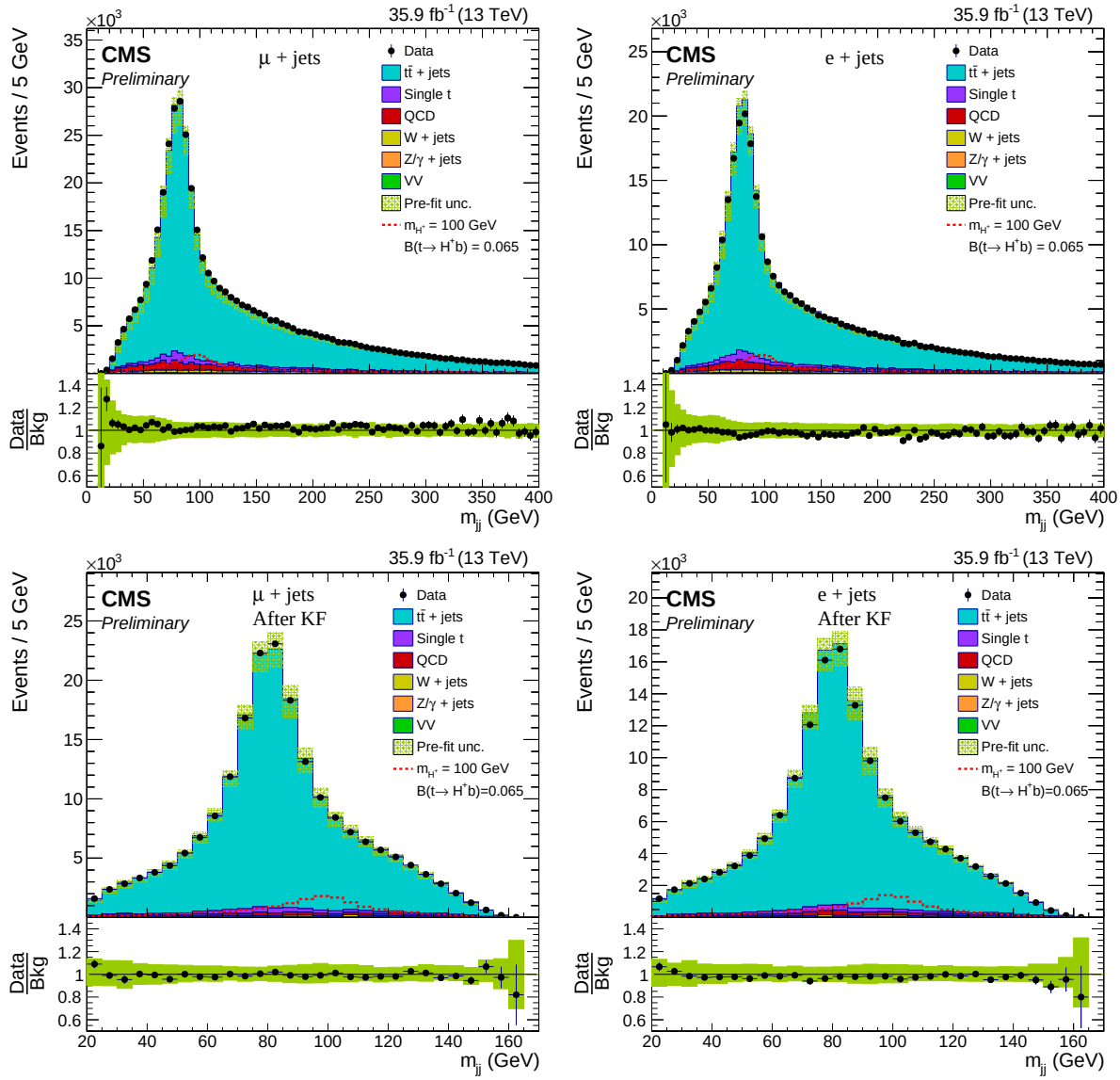

Figure 2: Distributions of m_{jj}, prior to the fit to data, of the two highest p_T light jets for the muon+jets channel (left column) and the electron+jets channel (right column). The two distributions in the top row are obtained using reconstructed jets. On the other hand, the distributions in the bottom row are calculated using kinematic fitted jets after the kinematic fit selection. The mean of the invariant mass distribution from the kinematic fitted jets is closer to the W mass as compared to that of reconstructed jets. The uncertainty band includes statistical as well as systematic uncertainties.

leptons, jets, and a neutrino with associated χ^2 and probability of the fit as output. The x and y components of the neutrino momentum are taken from $p_{\mathrm{T}}^{\mathrm{miss}}$, as the missing transverse momentum is attributed to the neutrino, and the z component of the neutrino momentum, p_z^{ν}, is determined from the fit. The following kinematic constraints are imposed on the semileptonic $\mathrm{t\bar{t}}$ system:

$$m_{\mathrm{inv}}(\mathrm{b_{had}q\bar{q}}) = m_{\mathrm{t}} = 172.5\,\mathrm{GeV} \tag{1a}$$
$$m_{\mathrm{inv}}(\mathrm{b_{lep}}\ell\nu_{\ell}) = m_{\bar{\mathrm{t}}} = 172.5\,\mathrm{GeV} \tag{1b}$$

After the fit, p_z^{ν} is determined from Eq. 1b. For every event, a χ^2 is constructed and minimized by varying the p_{T}, η, and ϕ of each object within their resolution. The values of p_{T}, η, and ϕ are finally selected that minimise the χ^2 and at the same time satisfy Eq. 1. In the output, the top kinematic fitter gives exactly four jets (two b jets, one from each of the leptonic and hadronic t decays, and two light jets from the hadronic t decay), a lepton, and a neutrino. The two light jets coming from the hadronic t decay are further used for charm tagging. Events for which the fit does not converge are discarded. Also, the same requirements as for the reconstructed objects are applied to the kinematically fitted objects. The directions of the kinematically fitted jets and lepton are required to be compatible with those of the reconstructed jets and lepton ($\Delta R < 0.2$), respectively.

Further, events are divided exclusively into loose, medium, and tight categories, based on whether at least one of the light jets passes the loose but neither passes the medium, at least one passes the medium but neither passes the tight, or at least one passes the tight working points of the charm tagging selection requirements [41], respectively. The m_{jj} distributions for the exclusive charm categories are shown in Fig. 3 after a background-only fit to the data. From these figures, it can be seen that the expected signal to background ratio is different in the various charm categories, so partitioning the events into categories results in an improvement in the upper limits on $\mathcal{B}(\mathrm{t} \to \mathrm{H^{+}b})$. The expected event yields with statistical and systematic (as discussed in Section 6) uncertainties for the different charm categories are shown in Table 1.

6 Systematic uncertainties

There are various sources of systematic uncertainty which may arise due to detector calibration effects, uncertainty in the measured reconstructed efficiency, the theoretical modeling of signal events, and other effects.

The uncertainty in the integrated luminosity is 2.5% for 2016 data taking [43]. To account for uncertainty in the total inelastic cross section of 69.2 mb, this is varied by its uncertainty of 4.7% and the simulated events are reweighted to match the resulting pileup distributions. The lepton reconstruction efficiency is different in data and simulated samples; to correct for this, lepton scale factors are applied to the simulated events. The systematic uncertainty in the lepton scale factor is 3.0% [37, 44]. The p_{T} of jets in the simulated samples are corrected using the jet energy scale (JES) and jet energy resolution (JER) scale factors. The jet energy correction is also propagated to correct $p_{\mathrm{T}}^{\mathrm{miss}}$. The systematic uncertainties due to JES and JER on the p_{T} of the jets and $p_{\mathrm{T}}^{\mathrm{miss}}$ are considered by varying these scale factors within their uncertainties. The b and c tagging efficiencies are different in simulation and data, and scale factors are applied to the simulated events. To estimate the corresponding systematic uncertainty, the b and c tag scale factors are varied within their uncertainties, with proper correlations applied. To estimate the systematic uncertainty on the data-driven QCD multijet background estimation, the muon (electron) relative isolation selection value is conservatively changed to 0.17 (0.11)

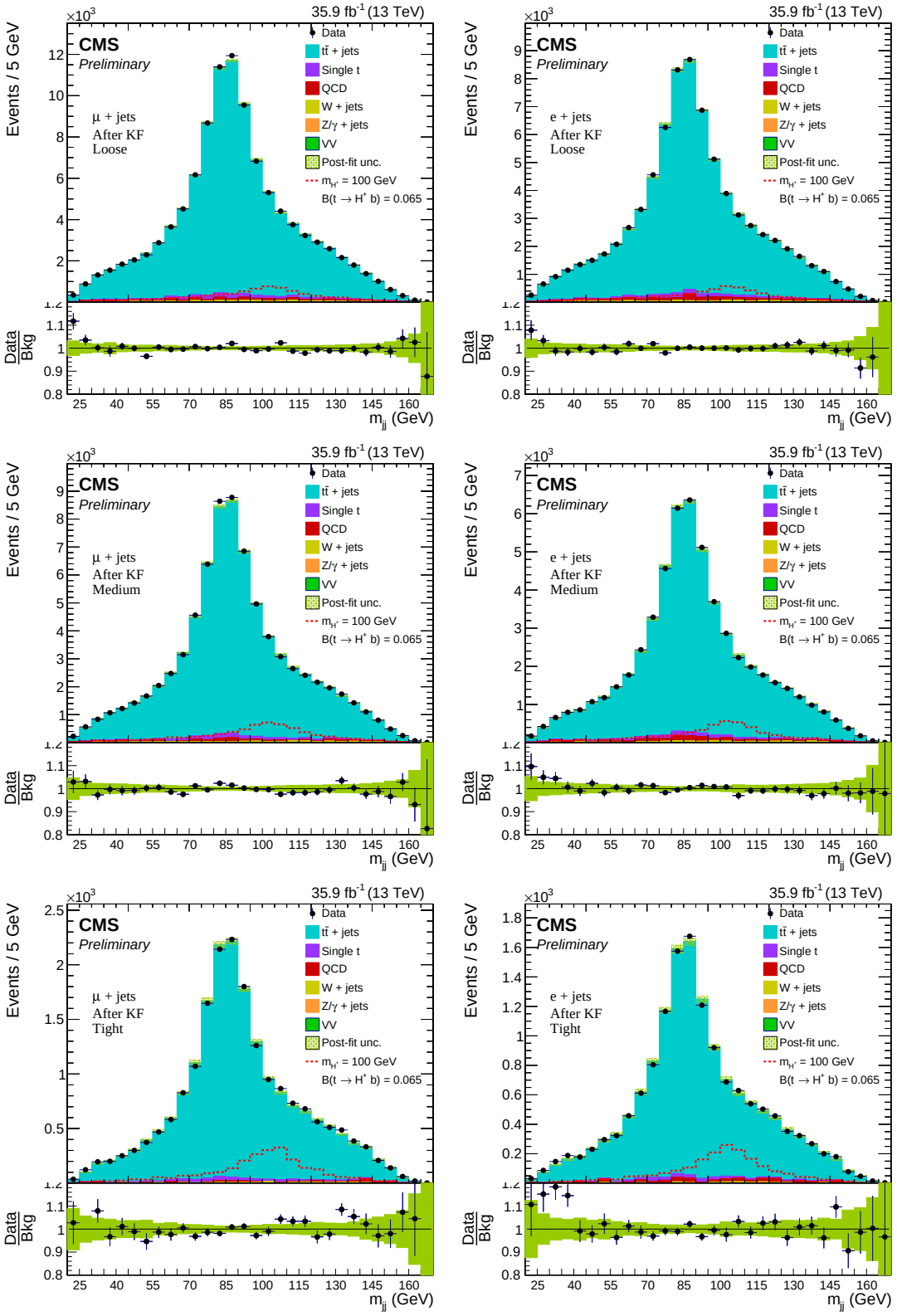

Figure 3: Distributions of m_{jj}, after a background-only fit to the data, in the exclusive charm categories for the muon + jets (left column) and electron + jets (right column) channels. The upper row shows the exclusive loose category, the middle row shows the exclusive medium category, and the lower row shows the exclusive tight category. The expected signal significance (prior to the fit) can be observed to vary across the different categories. The uncertainty band includes statistical as well as systematic uncertainties after the background-only fit.

Table 1: Expected event yields for different signal mass scenarios and backgrounds in each of the channels and event categories. The number of events, along with the uncertainty (including statistical and systematic effects), is shown. The yields for background processes are obtained after a background-only fit to the data. The total uncertainty on the background process is calculated by taking into account all the positive as well as negative correlations among the fit parameters.

Process	Loose		Medium		Tight	
	μ + jets	e + jets	μ + jets	e + jets	μ + jets	e + jets
$m_{H^+} = 80$ GeV	7171 ± 682	5375 ± 503	6137 ± 556	4683 ± 440	2471 ± 282	1826 ± 206
$m_{H^+} = 90$ GeV	7171 ± 722	5556 ± 539	6300 ± 607	4818 ± 472	2442 ± 277	1843 ± 211
$m_{H^+} = 100$ GeV	7382 ± 722	5480 ± 523	6591 ± 642	4911 ± 418	2570 ± 293	1967 ± 221
$m_{H^+} = 120$ GeV	7073 ± 692	5291 ± 528	6378 ± 606	4732 ± 415	2466 ± 265	1923 ± 217
$m_{H^+} = 140$ GeV	5716 ± 628	4338 ± 442	5043 ± 511	3816 ± 405	1869 ± 212	1472 ± 170
$m_{H^+} = 150$ GeV	4218 ± 480	3215 ± 367	3584 ± 407	2774 ± 320	1233 ± 157	1015 ± 138
$m_{H^+} = 155$ GeV	3447 ± 419	2522 ± 336	2758 ± 361	2205 ± 296	935 ± 152	762 ± 101
$m_{H^+} = 160$ GeV	2593 ± 318	2053 ± 252	2210 ± 299	1699 ± 246	682 ± 108	506 ± 68
SM $t\bar{t}$ + jets	99130 ± 458	71874 ± 498	72481 ± 354	52470 ± 329	18624 ± 136	13397 ± 134
Single t	2618 ± 212	1918 ± 165	1865 ± 150	1357 ± 114	421 ± 35	304 ± 27
QCD multijet	2207 ± 301	2153 ± 481	1385 ± 214	1426 ± 255	232 ± 50	364 ± 61
W + jets	1260 ± 146	1008 ± 92	872 ± 114	649 ± 57	112 ± 17	83 ± 14
Z/γ + jets	167 ± 21	220 ± 26	118 ± 15	126 ± 15	51 ± 7	30 ± 3
VV	62 ± 6	43 ± 6	47 ± 11	11 ± 4	11 ± 3	4 ± 1
All background	105444 ± 606	77216 ± 719	76768 ± 455	56039 ± 436	19451 ± 151	14182 ± 151
Data	105485	77244	76811	56051	19451	14179

and the corresponding changes in the QCD yields are determined. To account for the statistical uncertainty, in each bin of m_{jj}, one shape nuisance parameter is considered for all background and charged Higgs samples.

Each distribution for simulated events is normalized to the expected number of events in data, using the factor $L_{data}\sigma_{sim}/N_{sim}$, where L_{data} is the integrated luminosity in the data sample, N_{sim} is the total number of events in the simulated sample, and σ_{sim} is the cross section for the simulated process considered; the uncertainties in σ_{sim} thus contribute to the uncertainty in each background prediction. It is found that the p_T distribution of t quarks in $t\bar{t}$ events in data is softer compared to that from simulated samples [45]. This is corrected by applying the following weight as a function of p_T for SM $t\bar{t}$ + jets and charged Higgs samples:

$$w_{top} = \sqrt{SF(t)SF(\bar{t})}, \text{ where } SF \equiv \exp(0.0615 - 0.0005 p_T). \quad (2)$$

The values in the exponent are given in Ref. [45]. The generator-level p_T of the t and $\bar{t}$ are used to calculate SF. To evaluate the systematic uncertainty in w_{top}, it is varied to 1 and w_{top}^2. The SM $t\bar{t}$ + jets sample was generated with m_t = 172.5 GeV. To evaluate the effect of the chosen m_t on the m_{jj} distribution, alternate $t\bar{t}$ + jets samples with m_t = 171.5 and 173.5 GeV are considered. In the simulated samples, the NLO matrix element parton shower matching is varied by the damping parameter h_{damp}. Additional SM $t\bar{t}$ + jets samples are generated by varying h_{damp} up and down and are used to observe the effect of h_{damp}. Similarly, SM $t\bar{t}$ + jets samples where the renormalization and factorisation scales have been varied up and down are used to evaluate the uncertainties due to these scales.

All the systematic and statistical uncertainties are shown in Table 2. Among all systematics, the uncertainties on the data-driven QCD multijet background, lepton selection, $t\bar{t}$ + jets cross section, jet energy scale, and b/c tagging have a significant impact on the expected limit. The expected limit changes by 4.7%, 2.7%, 2.7%, 1.3%, and 0.7%, respectively, for the corresponding uncertainties. The effect of each of the remaining systematic uncertainties on the expected limit is estimated to be less than 0.3%.

Table 2: Systematic and statistical uncertainties in %, prior to the fit to data, for the exclusive charm categories in the muon (electron) channel. The "—" indicates that the corresponding uncertainties are not considered for the given process.

Category	Process	Pileup	jet & p_T^{miss}	b & c-jet	Normalization	Statistical	top p_T
Loose	$m_{H^+} = 100$ GeV	0.81 (1.2)	7.4 (6.8)	6.3 (6.4)	6.1 (6.1)	1 (1.2)	0.49 (1)
	$t\bar{t}$ + jets	0.97 (1.2)	6.9 (6.9)	5.9 (5.8)	6.1 (6.1)	0.19 (0.22)	0.73 (1.3)
	Single t	1.1 (0.82)	9.1 (10)	6.8 (6.7)	5 (5)	1 (1.2)	—
	W + jets	1.4 (1.1)	18 (13)	14 (9.7)	5 (5)	6.8 (4.6)	—
	Z/γ + jets	0.4 (2.3)	20 (18)	11 (9.3)	4.5 (4.5)	5.9 (4.4)	—
	VV	4.5 (11)	8.7 (16)	10 (12)	4 (4)	21 (22)	—
	QCD multijet	—	—	—	13 (23)	7.9 (7.7)	—
Medium	$m_{H^+} = 100$ GeV	0.65 (0.36)	6.7 (4.8)	6.9 (6.7)	6.1 (6.1)	1.1 (1.3)	0.65 (1.5)
	$t\bar{t}$ + jets	0.25 (0.45)	6.2 (6.2)	7.4 (7.3)	6.1 (6.1)	0.23 (0.27)	0.62 (1.4)
	Single t	0.73 (0.29)	8.5 (8.6)	7.9 (8.3)	5 (5)	1.3 (1.5)	—
	W + jets	1.5 (2.5)	20 (12)	12 (12)	5 (5)	5 (5.8)	—
	Z/γ + jets	1.5 (3.5)	21 (17)	12 (12)	4.5 (4.5)	6.3 (6)	—
	VV	13 (7.2)	28 (54)	22 (11)	4 (4)	21 (35)	—
	QCD multijet	—	—	—	15 (16)	12 (10)	—
Tight	$m_{H^+} = 100$ GeV	1.3 (1.4)	5.2 (5.6)	9.9 (9.4)	6.1 (6.1)	1.7 (1.9)	0.85 (1.1)
	$t\bar{t}$ + jets	1.3 (0.93)	5.9 (6)	10 (9.5)	6.1 (6.1)	0.44 (0.5)	0.47 (1.3)
	Single t	0.72 (0.38)	7.8 (8.9)	11 (10)	5 (5)	2.6 (3)	—
	W + jets	2.4 (3)	23 (27)	19 (13)	5 (5)	13 (14)	—
	Z/γ + jets	8.8 (4.6)	8.6 (13)	20 (13)	4.5 (4.5)	9.5 (15)	—
	VV	0.66 (8.9)	27 (0)	35 (10)	4 (4)	40 (1e+02)	—
	QCD multijet	—	—	—	11 (13)	28 (18)	—

7 Results

The event yields after all selection requirements have been applied, along with the combined systematic and statistical uncertainty, are shown in Table 1 for the muon + jets and electron + jets channels. From Table 1, it can be seen that the total expected background number of events agrees with the data within the uncertainties. The absence of a charged Higgs signal in the data is characterized by setting exclusion limits on the branching ratio $\mathcal{B}(t \to H^+b)$, assuming that $\mathcal{B}(H^+ \to c\bar{s})$ = 100%. The number of signal events in data, ΔN, is determined by subtracting the predicted number of background events from the total observed number, assuming that the difference is due to signal events only. The difference ΔN is calculated using the following formula:

$$\Delta N = N^{BSM}_{t\bar{t}+jets} - N^{SM}_{t\bar{t}+jets} = 2x(1-x)N^{sig} + [(1-x)^2 - 1]N^{SM}_{t\bar{t}+jets}. \quad (3)$$

where $N^{BSM}_{t\bar{t}+jets}$ is the number of events from beyond the standard model (BSM) decays of $t\bar{t}$, including the production of charged Higgs bosons, N^{sig} is the number of events from the sim-

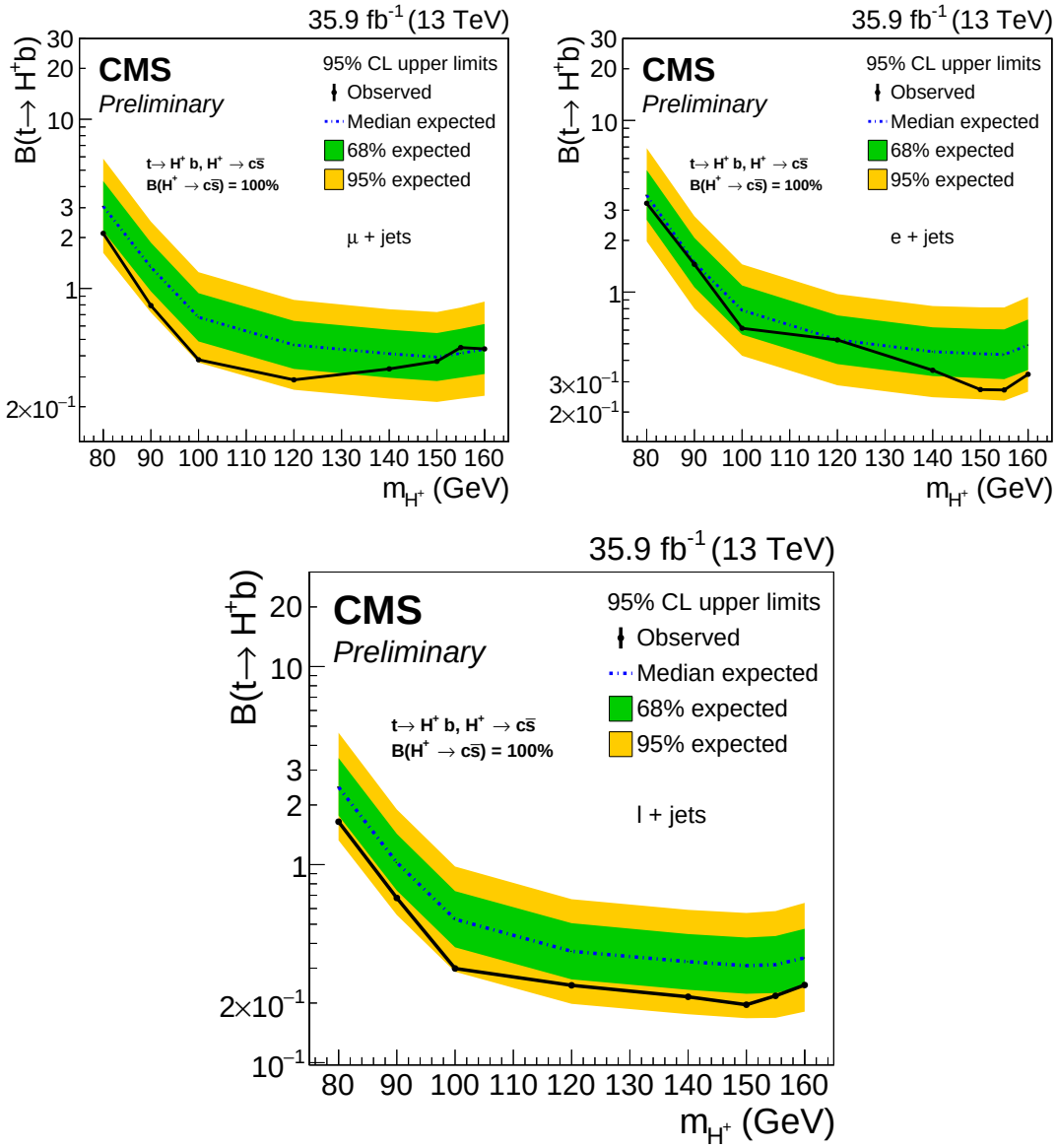

Figure 4: The expected and observed upper limits in % on $\mathcal{B}(t \to H^+b)$ as a function of m_{H^+} using m_{jj} after the individual charm tagging categories have been combined, for the muon + jets (upper left), electron + jets (upper right), and combined (bottom) channels.

ulated signal sample, $N^{SM}_{t\bar{t}+jets}$ is the number of events from SM $t\bar{t}$ + jets process as shown in Fig. 1, and x is the branching ratio of $t \to H^+b$. The values of N^{sig} and $N^{SM}_{t\bar{t}+jets}$ are shown in Table 1 for the muon + jets and electron + jets channels. The factor 2 in Eq. 3 is derived from the assumption that the event yield and $\mathcal{B}(\bar{t} \to H^-\bar{b})$ for H^- are the same as that of H^+ and $\mathcal{B}(t \to H^+b)$, respectively. An asymptotic 95% CL limit on $\mathcal{B}(t \to H^+b)$ is calculated using the CL_s method [46, 47] with likelihood ratios [48]. The exclusion limits as a function of charged Higgs mass using the m_{jj} distribution from combining different exclusive event categories based on charm tagging are shown in Fig. 4 and in Tables 3 and 4. Among the individual categories, the expected limits from the exclusive medium category are best, followed by those from the exclusive loose and tight categories. By construction, the exclusion limits on $\mathcal{B}(\bar{t} \to H^-\bar{b})$ are the same as those on $\mathcal{B}(t \to H^+b)$.

Table 3: Expected and observed 95% CL exclusion limits in % on $\mathcal{B}(t \to H^+b)$ for the muon (electron) channel, after the individual charm tagging categories have been combined.

m_{H^+} (GeV)	Expected					Observed
	-2σ	-1σ	median	$+1\sigma$	$+2\sigma$	
80	1.62 (1.97)	2.18 (2.63)	3.06 (3.66)	4.31 (5.14)	5.85 (6.89)	2.12 (3.29)
90	0.72 (0.80)	0.96 (1.07)	1.34 (1.48)	1.87 (2.07)	2.49 (2.77)	0.79 (1.45)
100	0.36 (0.43)	0.49 (0.56)	0.68 (0.78)	0.94 (1.09)	1.25 (1.45)	0.38 (0.61)
120	0.25 (0.29)	0.33 (0.38)	0.46 (0.53)	0.64 (0.73)	0.86 (0.97)	0.29 (0.53)
140	0.22 (0.24)	0.30 (0.32)	0.41 (0.45)	0.57 (0.62)	0.75 (0.83)	0.33 (0.35)
150	0.21 (0.24)	0.28 (0.31)	0.39 (0.44)	0.54 (0.61)	0.72 (0.81)	0.37 (0.27)
155	0.22 (0.23)	0.30 (0.31)	0.41 (0.43)	0.58 (0.61)	0.77 (0.81)	0.45 (0.27)
160	0.23 (0.26)	0.31 (0.35)	0.44 (0.49)	0.62 (0.69)	0.84 (0.94)	0.44 (0.33)

Table 4: Expected and observed 95% CL exclusion limits in % on $\mathcal{B}(t \to H^+b)$, after the individual charm tagging categories and the electron and muon channels have been combined.

m_{H^+} (GeV)	Expected					Observed
	-2σ	-1σ	median	$+1\sigma$	$+2\sigma$	
80	1.33	1.77	2.46	3.45	4.63	1.65
90	0.56	0.74	1.03	1.43	1.90	0.68
100	0.29	0.38	0.53	0.73	0.98	0.30
120	0.20	0.26	0.36	0.51	0.67	0.25
140	0.18	0.23	0.32	0.45	0.59	0.21
150	0.17	0.22	0.31	0.43	0.57	0.20
155	0.17	0.22	0.31	0.44	0.58	0.22
160	0.18	0.24	0.34	0.47	0.64	0.25

8 Summary

A search for a light charged Higgs boson $H^\pm$ has been performed in the muon + jets and electron + jets channels at $\sqrt{s}$ = 13 TeV, using a data sample with an integrated luminosity of 35.9 fb^{-1}. The observed and predicted number of events are in agreement within the statistical and systematic uncertainties as shown in Table 1. In the absence of observed signal, an exclusion limit at 95% confidence level on the branching ratio $\mathcal{B}(t \to H^+b)$ has been computed by assuming $\mathcal{B}(H^+ \to c\bar{s}) = 100\%$. The observed exclusion limits are in the range, depending on the H^+ mass, 0.29–2.12%, 0.27–3.29%, and 0.20–1.65% for the muon + jets, electron + jets, and combined channels, respectively. The expected exclusion limits from 13 TeV are better by a factor of ≈4, as compared to those obtained from earlier CMS results at 8 TeV [15].

References

[1] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett. B* **716** (2012) 30, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.

[2] ATLAS Collaboration, "Observation of a new particle in the search for the standard model Higgs boson with the ATLAS detector at the LHC", *Phys. Lett. B* **716** (2012) 1, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.

[3] S. P. Martin, "A supersymmetry primer", in *Perspectives on Supersymmetry*, p. 1. 1997. arXiv:hep-ph/9709356. doi:10.1142/9789812839657_0001.

[4] Yu. A. Golfand and E. P. Likhtman, "Extension of the algebra of Poincare group generators and violation of p invariance", *JETP Lett.* **13** (1971) 323.

[5] J. Wess and B. Zumino, "Supergauge transformations in four dimensions", *Nucl. Phys.* **70** (1974) 39, doi:10.1016/0550-3213(74)90355-1.

[6] G. C. Branco et al., "Theory and phenomenology of two-Higgs-doublet models", *Phys. Rept.* **516** (2012) 1, doi:10.1016/j.physrep.2012.02.002, arXiv:1106.0034.

[7] M. Aoki, S. Kanemura, K. Tsumura, and K. Yagyu, "Models of Yukawa interaction in the two Higgs doublet model, and their collider phenomenology", *Phys. Rev. D* **80** (2009) 015017, doi:10.1103/PhysRevD.80.015017, arXiv:0902.4665.

[8] L3 Collaboration, "Search for charged Higgs bosons at LEP", *Phys. Lett. B* **575** (2003) 208, doi:10.1016/j.physletb.2003.09.057, arXiv:hep-ex/0309056.

[9] ALEPH Collaboration, "Search for charged Higgs bosons in e^+e^- collisions at energies up to $\sqrt{s}$ = 209 GeV", *Phys. Lett. B* **543** (2002) 1, doi:10.1016/S0370-2693(02)02380-8, arXiv:hep-ex/0207054.

[10] DELPHI Collaboration, "Search for charged higgs bosons at LEP in general two Higgs doublet models", *Eur. Phys. J. C* **34** (2004) 399, doi:10.1140/epjc/s2004-01732-6, arXiv:hep-ex/0404012.

[11] OPAL Collaboration, "Search for charged Higgs bosons in e^+e^- collisions at $\sqrt{s} = 189 - 209$ GeV", *Eur. Phys. J. C* **72** (2012) 2076, doi:10.1140/epjc/s10052-012-2076-0, arXiv:0812.0267.

[12] CDF Collaboration, "Search for charged Higgs bosons in decays of top quarks in $p\overline{p}$ collisions at $\sqrt{s}$ = 1.96 TeV", *Phys. Rev. Lett.* **103** (2009) 101803, doi:10.1103/PhysRevLett.103.101803, arXiv:0907.1269.

[13] D0 Collaboration, "Search for charged Higgs bosons in top quark decays", *Phys. Lett. B* **682** (2009) 278, doi:10.1016/j.physletb.2009.11.016, arXiv:0908.1811.

[14] ATLAS Collaboration, "Search for a light charged Higgs boson in the decay channel $H^+ \to c\bar{s}$ in $t\bar{t}$ events using pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector", *Eur. Phys. J. C* **73** (2013) 2465, doi:10.1140/epjc/s10052-013-2465-z, arXiv:1302.3694.

[15] CMS Collaboration, "Search for a light charged Higgs boson decaying to $c\bar{s}$ in pp collisions at $\sqrt{s} = 8$ TeV", *JHEP* **12** (2015) 178, doi:10.1007/JHEP12(2015)178, arXiv:1510.04252.

[16] CMS Collaboration, "The CMS trigger system", *JINST* **12** (2017) P01020, doi:10.1088/1748-0221/12/01/P01020, arXiv:1609.02366.

[17] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[18] J. Alwall et al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", *JHEP* **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.

[19] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with parton shower simulations: the POWHEG method", *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.

[20] P. Nason, "A new method for combining NLO QCD with shower Monte Carlo algorithms", *JHEP* **11** (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.

[21] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP* **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.

[22] T. Sjostrand, S. Mrenna, and P. Z. Skands, "A brief introduction to PYTHIA 8.1", *Comput. Phys. Commun.* **178** (2008) 852, doi:10.1016/j.cpc.2008.01.036, arXiv:0710.3820.

[23] CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", *Eur. Phys. J. C* **76** (2016) 155, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815.

[24] GEANT4 Collaboration, "GEANT4—a simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.

[25] J. M. Campbell, R. K. Ellis, P. Nason, and E. Re, "Top-pair production and decay at NLO matched with parton showers", *JHEP* **04** (2015) 114, doi:10.1007/JHEP04(2015)114, arXiv:1412.1828.

[26] NNPDF Collaboration, "Parton distributions from high-precision collider data", *Eur. Phys. J. C* **77** (2017) 663, doi:10.1140/epjc/s10052-017-5199-5, arXiv:1706.00428.

[27] M. Beneke, P. Falgari, S. Klein, and C. Schwinn, "Hadronic top-quark pair production with NNLL threshold resummation", *Nucl. Phys. B* **855** (2012) 695, doi:10.1016/j.nuclphysb.2011.10.021, arXiv:1109.1536.

[28] Particle Data Group Collaboration, "Review of particle physics", *Phys. Rev. D* **98** (2018) 030001, doi:10.1103/PhysRevD.98.030001.

[29] S. Alioli, P. Nason, C. Oleari, and E. Re, "NLO single-top production matched with shower in POWHEG: s- and t-channel contributions", *JHEP* **09** (2009) 111, doi:10.1088/1126-6708/2009/09/111, arXiv:0907.4076. [Erratum: doi:10.1007/JHEP02(2010)011].

[30] E. Re, "Single-top Wt-channel production matched with parton showers using the POWHEG method", *Eur. Phys. J. C* **71** (2011) 1547, doi:10.1140/epjc/s10052-011-1547-z, arXiv:1009.2450.

[31] J. Alwall et al., "Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions", *Eur. Phys. J. C* **53** (2008) 473, doi:10.1140/epjc/s10052-007-0490-5, arXiv:0706.2569.

[32] CMS Collaboration, "Particle-flow reconstruction and global event description with the CMS detector", *JINST* **12** (2017) P10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.

[33] K. Rose, "Deterministic annealing for clustering, compression, classification, regression, and related optimization problems", *IEEE Proc.* **86** (1998) 2210, doi:10.1109/5.726788.

[34] R. Fruhwirth, W. Waltenberger, and P. Vanlaer, "Adaptive vertex fitting", CMS Note 2007/008, 2007.

[35] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm", *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

[36] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet user manual", *Eur. Phys. J. C* **72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.

[37] CMS Collaboration, "Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV", *JINST* **10** (2015) P06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.

[38] A. M. Polyakov, "Quark confinement and topology of gauge groups", *Nucl. Phys. B* **120** (1977) 429, doi:10.1016/0550-3213(77)90086-4.

[39] CMS Collaboration, "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV", *JINST* **12** (2017) P02014, doi:10.1088/1748-0221/12/02/P02014, arXiv:1607.03663.

[40] CMS Collaboration, "Identification of b-quark jets with the CMS experiment", *JINST* **8** (2013) P04013, doi:10.1088/1748-0221/8/04/P04013, arXiv:1211.4462.

[41] CMS Collaboration, "Identification of c-quark jets at the CMS experiment", CMS Physics Analysis Summary CMS-PAS-BTV-16-001, 2016.

[42] J. D'Hondt et al., "Fitting of event topologies with external kinematic constraints in CMS", CMS Note 2006/023, 2006.

[43] CMS Collaboration, "CMS luminosity measurements for the 2016 data taking period", CMS Physics Analysis Summary CMS-PAS-LUM-17-001, 2017.

[44] CMS Collaboration, "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV", *JINST* **13** (2018) P06015, doi:10.1088/1748-0221/13/06/P06015, arXiv:1804.04528.

[45] CMS Collaboration, "Measurement of differential cross sections for top quark pair production using the lepton+jets final state in proton-proton collisions at 13 TeV", *Phys. Rev. D* **95** (2017) 092001, doi:10.1103/PhysRevD.95.092001, arXiv:1610.04191.

[46] A. L. Read, "Presentation of search results: The CL_s technique", *J. Phys. G* **28** (2002) 2693, doi:10.1088/0954-3899/28/10/313.

[47] T. Junk, "Confidence level computation for combining searches with small statistics", *Nucl. Instrum. Meth. A* **434** (1999) 435, doi:10.1016/S0168-9002(99)00498-2, arXiv:hep-ex/9902006.

[48] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", *Eur. Phys. J. C* **71** (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0, arXiv:1007.1727. [Erratum: doi:10.1140/epjc/s10052-013-2501-z].