CMS Physics Analysis Summary

Search for exotic decays of the Higgs boson to a pair of new light bosons with two muon and two b jets in final states

The CMS Collaboration

Abstract

We report the results of a search for exotic decays of a Higgs boson with $m_h = 125$ GeV to a pair of new light bosons, $a_1$, where one of the light bosons decays to a pair of muons and the other one decays to a pair of b quarks. Such signatures are predicted in a number of well motivated extensions of the standard model, including the next-to-minimal supersymmetry and generic two Higgs doublet models with an additional scalar singlet. A data sample corresponding to an integrated luminosity of 19.7 fb$^{-1}$ recorded with the CMS detector in 2012 is exploited in $\mu^+\mu^-b\bar{b}$ final states where no statistically significant excess is observed with respect to the standard model backgrounds for different $m_{a_1}$ hypotheses above 25 GeV and below $m_h/2$. Upper limits are set on $\sigma_{ggF} \times \text{Br}(h \rightarrow \mu^+\mu^-b\bar{b})$ and on the branching ratio itself with former ranging between 4 to 12 fb, depending on the $m_{a_1}$ values. The limit on the branching ratio is $\text{Br}(h \rightarrow \mu^+\mu^-b\bar{b}) < 9 \times 10^{-4}$ for the entire mass range.
1 Introduction

The discovery of a Higgs-like boson with a mass of about 125 GeV, referred to as “the Higgs” or “h” hereafter, with the ATLAS and CMS experiments [1, 2] at the CERN LHC [3] has opened a new era in the history of particle physics. So far the precise measurements of the Higgs boson cross section and properties have been consistent with the predictions for the standard model (SM) Higgs boson [4, 5]. The precision of these measurements are however not sufficient to substantially exclude the Higgs boson decays into new particles; instead encouraging direct searches in the framework of physics beyond the standard model.

The analysis of the existing data, including constraints from the Higgs boson observation in different SM channels, allows for Higgs boson decays to states beyond the standard model (BSM) with a rate of $O(20\%–50\%)$. There are several experimental and theoretical studies that fit the couplings of the Higgs boson to SM and constrain the Higgs boson branching ratio (Br) into a so far undetected or an invisible final state [6–11]. Assuming the SM production for the Higgs boson, a $Br(h \rightarrow BSM)$ of 20% is allowed at 95% CL. This value increases to 30% if some new physics modifies the loop-induced couplings to gluons and photons. More conservative approaches allow for even larger rates [12, 13]. Added to the experimental results, the ultimate precision of the LHC indirect measurement of the Higgs branching fraction to new particles is expected to be of $O(5\%–10\%)$ [14–16]. Therefore $Br(h \rightarrow BSM) \sim 10\%$ seems to remain a reasonable targets during the physics program of the LHC.

The next-to-minimal supersymmetric standard model (NMSSM) and extensions to two Higgs doublet models (2HDM) are examples of well motivated models predicting the Higgs boson decays into a pairs of new scalar singlets [17]. After the electroweak symmetry breaking, the Two Higgs Doublet models allow for a pair of charged Higgs bosons, a neutral pseudoscalar, and two neutral scalar mass eigenstates. In the decoupling limit, the lightest scalar eigenstate could be identified as the recently discovered boson, h, with $m_h = 125$ GeV. The model can be extended further by adding a complex scalar singlet ($a_1$) with no direct Yukawa couplings. The new singlet is therefore expected to decay to SM fermions by virtue of mixing with the Higgs sector. The mixing needs to be small to preserve the SM-like nature of the h boson.

For the NMSSM the additional singlet is introduced to resolve the so-called $\mu$-problem of the MSSM superpotential while reducing the fine-tuning of model parameters. The Higgs sector of NMSSM contains seven higgs particles: three CP-even, $h_{1,2,3}$, with $m_{h_1} > m_{h_2} > m_{h_3}$, where either of $h_1$ or $h_2$ could be considered as the discovered Higgs boson (h) with a mass close to 125 GeV. There are two CP-odd Higgs particles, $a_1, a_2$, with $m_{a_1} > m_{a_2}$ and a pair of charged Higgs bosons, $H^\pm$. Based on the decay of the CP-odd bosons and their mass, different lines of research are followed looking for NMSSM signatures [18–20].

In this analysis, only the gluon fusion production mechanism is considered for the Higgs production with the next-to-leading-order rate of $\sigma_{ggH} \simeq 19.3$ pb [21]. The final results are presented with a branching fraction of 10% for $h \rightarrow a_1a_1$ where, assuming $\tan \beta = 2$, one can obtain $2 \times Br(a_1 \rightarrow b\bar{b}) Br(a_1 \rightarrow \mu^+\mu^-) = 1.7 \times 10^{-3}$ for $m_{a_1} = 30$ GeV in the context of the extended 2HDM [17]. For the set of parameters under discussion and with $25 \leq m_{a_1} \leq 65$ GeV, no strong dependence to $m_{a_1}$ is expected for $Br(a_1 \rightarrow f\bar{f})$, with $f$ being muon or b quark [17].

Figure 1 illustrates the Higgs two-body decay to two $a_1$ bosons with one decaying to a pair of muons and the other to a pair of b quarks. This final state has the advantage of the higher rate and lower background contamination in comparison with the $4\mu$ and $4b$ final states, respectively. We search for an exotic lighter particle, referred to as $a_1$ throughout the text for simplicity, with $25 \leq m_{a_1} \leq 65$ GeV. The lower bound is set in order to avoid the significant
Data and simulated samples

loss of sensitivity of the current search towards $m_{a_1} \approx 20$ GeV and lower. The upper bound is slightly above the kinematic threshold imposed by mass of the Higgs boson. The analysis is performed using the data collected with the CMS detector during 2012, corresponding to an integrated luminosity of 19.7 fb$^{-1}$.

2 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 superconducting solenoid volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass/scintillator hadron calorimeter (HCAL). Muons are measured in gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid. Extensive forward calorimetry complements the coverage provided by the barrel and endcap detectors. Muons measured in the pseudorapidity range $|\eta| < 2.4$ of the muon system are matched to tracks measured in the silicon tracker. This results in transverse momentum resolution for muons with $20 < p_T < 100$ GeV of 1.3–2.0% in the barrel and better than 6% in the endcaps [22].

The calorimetry systems, ECAL and HCAL, with $|\eta| < 3.0$ coverage are used to identify and measure the energy of different particles including hadrons. The CMS detector is nearly hermetic, which permits good measurements of the energy imbalance in the plane transverse to the beam line. 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 in [23].

3 Data and simulated samples

This analysis is performed using the data from the LHC proton-proton collisions at 8 TeV center-of-mass energy. The data sample, corresponding to an integrated luminosity of 19.7 fb$^{-1}$ for the double-muon triggers, was collected with the CMS detector in 2012. The NMSSM benchmark model is used to generate signal samples with PYTHIA 6.4 [24] where the $p_T$ of the Higgs boson is corrected for the next-to-next-to-leading order effects. The Drell-Yan process, $Z/\gamma^*(\to \ell\ell) + \text{jets}$, is modeled with MADGRAPH 5.148 [25] event generator and interfaced with PYTHIA for parton showering. A lower bound of $m_{\ell\ell} > 10$ GeV is applied to avoid divergences at low dilepton invariant masses. Similar generator and showering program is used for t$\bar{t}$ and dibosons (WW, WZ, ZZ) event samples. Single top quark events produced in association with a W boson are generated using POWHEG 1.0 [26–29] interfaced with PYTHIA for parton showering. Despite its small contribution, the Z-boson associated production of the SM Higgs boson is also included in the list of backgrounds. This sample is generated with
POWHEG 1.0 and interfaced with HERWIG++ 2.6 [30] for parton showering. The full CMS detector simulation based on GEANT4 [31] is implemented for all Monte Carlo (MC) generated event samples.

4 Event selection and optimization

Events are filtered using a high-level trigger requirement based on the presence of two muons with $p_T > 17$ and 8 GeV. For offline selection, events must contain at least one primary vertex, considered as the vertex of the hard interaction. At least four tracks must be associated to the selected primary vertex. The longitudinal and radial distances of the vertex from the center of the detector must be smaller than 24 cm and 2 cm, respectively. For events with more than one selected primary vertex, the one with the largest $\Sigma p_T^2$ of the associated tracks is chosen for the analysis. As part of the quality requirements, events in which an abnormally high level of noise is detected in the HCAL barrel or endcap detectors have been rejected [32]. Extra selection criteria are applied to leptons and jets, reconstructed using the CMS particle flow (PF) algorithm [33, 34].

In this section we first present a minimal set of requirements, followed by an optimization procedure that maximizes the expected signal significance.

The minimal selection requires two isolated muons originating from the selected primary vertex with $|\eta| < 2.4$. The leading muon $p_T$ should exceed a threshold of 24 GeV whereas the sub-leading one is accepted with $p_T > 8$. GeV.

The isolation variable $I_{\text{rel}}$ is calculated by summing the transverse energy deposited by other particles in a cone of size $\Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2} = 0.4$ around the muon, divided by the muon $p_T$,

$$I_{\text{rel}} = \frac{I_{\text{ch.h}} + \max((I_\gamma + I_{n.h} - I_{\text{PU}}), 0)}{p_T}. \quad (1)$$

Quantities $I_{\text{ch.h}}$, $I_\gamma$, and $I_{n.h}$ are the sum of the transverse energies deposited by stable charged hadrons, photons, and neutral hadrons, respectively. Variable $I_{\text{PU}}$ is defined as $0.5 \times \sum p_T^{\text{PU}}$ which is the sum of transverse momenta of tracks associated to non-leading vertices, used to estimate the contribution of neutral particles from pileup events. The neutral-to-charged particles ratio is approximated to be 0.5 from the isospin invariance. Only muons with the isolation variable satisfying $I_{\text{rel}} < 0.15$ are considered in the analysis. The efficiencies for muon trigger, reconstruction and selection in simulated events are corrected to match those in data.

Jets are reconstructed by clustering the charged and neutral particles using an anti-$k_T$ algorithm [35] with a distance parameter of 0.5. The reconstructed jet energy is corrected for effects from the detector response as a function of the jet $p_T$ and $\eta$. Furthermore, contamination from additional interactions (pileup), underlying events, and electronic noise are subtracted [36]. To achieve a better agreement between data and simulation, an extra $\eta$-dependent smearing is performed on the jet energy of the simulated events [36]. Events are required to have two jets with $|\eta| < 2.4$ and $p_T > 15 \text{ GeV}$, where both jets must be separated from the selected leptons ($\Delta R > 0.5$).

The combined secondary vertex (CSV) algorithm is used to identify jets that are likely to originate from a b quark. The algorithm exploits the track-based lifetime information together with the secondary vertices inside the jet to provide a multivariate discriminator for the b jet identi-
Event selection and optimization

In selected events, the jet with a higher discriminating value is required to meet the loose criteria of the combined secondary vertex algorithm. A set of $p_T$-dependent correction factors are applied to simulated events to account for differences in the $b$ tagging efficiency between data and simulation [37].

Because of the typically small amount of energy carried by neutrinos from possible semileptonic decays in $b$ jets, the imbalance in the transverse momentum of signal events is not expected to be large. The missing transverse energy, $E_T$, is defined as the modulus of $\vec{p}_T$, which is the negative $\vec{p}_T$ sum of all reconstructed PF candidates. The jet energy calibration therefore introduces corrections to the $E_T$ measurement.

The $E_T$ significance, calculated via a likelihood function on an event-by-event basis [38], has shown to provide a better discrimination than $E_T$ against background in the presence of pileup. Events are selected if the value of the $E_T$ significance is less than 6.

The search for a new scalar, and therefore the final limits, are restricted to $m_{a1} \in [25, 65]$. However for the selection, optimization and eventual background modeling, a slightly wider range ($\pm 5$ GeV) is used. This is to ensure a good measurement of the background shape through the entire target search region, including regions near the boundaries. It means events with $m_{\mu\mu}$ out of $[20, 70]$ GeV are discarded.

Although the background estimation for this analysis if fully based on data, simulated samples are used to optimize the selection. Figure 2 shows distributions for events passing the minimal selection requirements in data and simulation. The simulated events at this level are used in the optimization procedure. In this figure, data and simulation are compared for the $p_T$ of the di-muon system and the mass and the $p_T$ of the di-jet system. In events with more muons or jets passing the selection criteria, the two with the highest $p_T$ are taken. Using the same selected muon and jet pairs, Fig. 2 also illustrates the distributions of the invariant mass $m_{\mu\mu jj}$ and the transverse momentum $p_{T\mu\mu jj}$ of the four-body system.

The distributions for simulated events follow reasonably those in the data. The yields in data and simulation, presented in Table 1, are also in a reasonable agreement.

Table 1: Event yields for data and simulated processes after the minimal event selection. The expected number of simulated events is normalized to the integrated luminosity of 19.7 fb$^{-1}$.

<table>
<thead>
<tr>
<th>Process</th>
<th>Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Z/\gamma^*+$jets $m_{\ell\ell} &gt; 10$ GeV</td>
<td>87380 ± 687</td>
</tr>
<tr>
<td>Top (tt ($\ell\ell$))</td>
<td>3238 ± 12</td>
</tr>
<tr>
<td>Top (tt ($\ell\ell$) + tW)</td>
<td>500 ± 11</td>
</tr>
<tr>
<td>Diboson</td>
<td>325 ± 4</td>
</tr>
<tr>
<td>Zh</td>
<td>1 ± 0.02</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>91444 ± 686</td>
</tr>
<tr>
<td><strong>Data</strong></td>
<td>92906</td>
</tr>
</tbody>
</table>

After the optimization procedure, the leading muon $p_T$ is required to be greater than 24 GeV where the subleading muon must pass the 9 GeV threshold. The two jets are selected with $p_T > 15$ GeV and are required to pass medium b jet identification criteria. The $E_T$ significance has to be less than 6. Events outside $|m_{\mu\mu bb} - 125| < 25$ GeV window are rejected where here the $m_{\mu\mu bb}$ quantity refers to the invariant mass of the two muon together with the two b-tagged jets. It is expected to be compatible with the mass of the Higgs boson for signal events. For completeness, the expected yields together with the number of data events after the optimized
Figure 2: The distribution of the $p_T$ of the di-muon and di-jet system (a,b), the mass of the di-jet and $\mu\mu jj$ system (c,d) and $p_T$ of the $\mu\mu jj$ system (e) after the minimal selection. Simulated samples are normalized to $19.7 \text{ fb}^{-1}$ with their theoretical cross sections.
selection are provided in Table 2.

Table 2: Expected yields for simulated processes after the optimized selection. The expected number of simulated events is normalized to the integrated luminosity of 19.7 fb⁻¹.

<table>
<thead>
<tr>
<th></th>
<th>Z/γ*+jets (m_{ℓℓ}&gt;10 GeV)</th>
<th>t̅t (ℓℓ)</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Backgrounds</td>
<td>210 ± 35</td>
<td>22 ± 1</td>
<td>3 ± 1</td>
</tr>
<tr>
<td>Total</td>
<td>235 ± 35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data</td>
<td>252</td>
<td></td>
<td></td>
</tr>
<tr>
<td>m_{a1} = 30 GeV</td>
<td>m_{a1} = 40 GeV</td>
<td>m_{a1} = 50 GeV</td>
<td>m_{a1} = 60 GeV</td>
</tr>
<tr>
<td>Signal</td>
<td>1.18</td>
<td>0.972</td>
<td>1.11</td>
</tr>
</tbody>
</table>

## 5 Signal and background modeling

The sensitivity of the current search to signal is extracted using a fit to m_{µµ} distribution in data. The signal shape is modeled with a weighted sum of Voigt profile [39] and Crystal ball [40] functions where the mean values of the two are bound to be the same. The Voigt profile function is a convolution of Lorentz and Gaussian profiles,

\[ V(m_{µµ}, σ, γ, m_{a1}) = G(m_{µµ}, σ, m_{a1}) * L(m_{µµ}, γ, m_{a1}) \]

\[ = \left( \frac{e^{- (m_{µµ} - m_{a1})^2 / 2σ^2}}{σ \sqrt{2π}} \right) * \left( \frac{γ}{π((m_{µµ} - m_{a1})^2 + γ^2)} \right), \]

with γ and σ being the width of the respective functions that are both centered at m_{a1}. The Crystal ball function has a Gaussian core and a power-law low-end tail below a certain threshold, α,

\[ CB(m_{µµ}, n, σ_{cb}, α, m_{a1}) = N \left\{ \begin{array}{ll}
    e^{- (m_{µµ} - m_{a1})^2 / 2σ_{cb}^2} & \text{for } \frac{m_{µµ} - m_{a1}}{σ_{cb}} > -α \\
    A \cdot \left( B - \frac{m_{µµ} - m_{a1}}{v_{cb}} \right)^{-n} & \text{for } \frac{m_{µµ} - m_{a1}}{v_{cb}} \leq -α
\end{array} \right. \]

The parameters A, B and N are functions of n, σ_{cb} and α. The initial values for the signal model parameters are extracted from a simultaneous of the model fit to a number of simulated signal samples, namely samples with m_{a1} = \{30, 40, 50, 60\} GeV. All parameters in the signal model are found to be independent of m_{a1}, except σ and σ_{cb}, which change linearly with only their slopes, respectively b_{1} and a_{1}, floating within their uncertainties,

\[ σ_{cb} = \hat{σ}_{cb} + a_{1} \cdot m_{µµ}, \]

\[ σ = \hat{σ} + b_{1} \cdot m_{µµ}. \]

Therefore in the final limit calculation, the signal model with three free parameters, m_{a1}, a_{1} and b_{1}, is interpolated between mass hypotheses of the simulated samples.

The background is evaluated through a fit to the m_{µµ} distribution in data, without any reference to simulation. The uncertainty associated with the choice of the background model is treated in a similar way to other uncertainties for which there are nuisance parameters in the fit. The likelihood function for the signal-plus-background fit has the form of

\[ L(data | s(p, m_{µµ}) + b(m_{µµ})), \]
where \( s(p, m_{\mu\mu}) \) is the parametric signal shape with a set of parameters indicated by \( p \), and \( b(m_{\mu\mu}) \) is the background model. The shape for the background is modeled with a set of analytical functions, using the discrete profiling method [41–43]. In this approach, the background shape is treated as a discrete nuisance parameter in the fit where its phase space contains multiple models, each including their own parameters.

To provide the input background models to the discrete profiling method, the background data are modeled with different parametrization of polynomials together with \( 1/P_n(x) \) functions where \( P_n(x) \equiv x + \sum_{i=2}^{n} a_i x^i \). The degree of polynomials in each category are determined through statistical tests to ensure the sufficiency of number of parameters and to avoid overfitting the data. Every pdf candidate is fitted to data and a \( p \)-value is evaluated for its parameters after the fit, according to the number of degrees of freedom and the relative uncertainty of the parameter. The pdfs whose parameter \( p \)-values are below 5% are discarded.

The input background functions are tried in the minimization of the negative logarithm of the likelihood with a penalty term added to account for the number of free parameters in the background model. The likelihood ratio for the penalized likelihood function can be written as

\[
-2 \ln \frac{\tilde{L}(\text{data}|\mu, \hat{\theta}, \hat{b})}{\tilde{L}(\text{data}|\tilde{\mu}, \tilde{\theta}, \tilde{b})},
\]

where \( \mu \) is the measured quantity. The numerator is the maximum penalized likelihood for a given \( \mu \), at the best-fit values of nuisance parameters, \( \hat{\theta}_\mu \) and of the background function, \( \hat{b}_\mu \). The denominator is the global maximum for \( \tilde{L} \), achieved at \( \mu = \tilde{\mu} \), \( \theta = \tilde{\theta} \) and \( b = \tilde{b} \). A confidence interval on \( \mu \) is obtained with the background function maximizing \( \tilde{L} \) for any value of \( \mu \). This interval is always wider than those evaluated with the fixed functional form from the global best-fit, \( b = \hat{b} \) [41].

### 6 Systematic uncertainties

Final statistical interpretation of the analysis takes into account several sources of systematic uncertainties related to the accuracy in the signal modeling, uncertainties in the signal acceptance and the imprecise knowledge of the shape of the background contributions where the latter and the way it is handled in the fit is discussed earlier.

**Theory uncertainties:** to evaluate the upper limit on \( B(h \to a_1 a_1 \to \mu^+ \mu^- b\bar{b}) \), the Higgs boson production cross section is set to the SM prediction where an uncertainty of 13% is considered for \( \sigma_{gg \to h} \), accounting for parton distribution function, \( \alpha_s \) and scale variations [21].

**Uncertainties in signal shape and acceptance modeling:** an uncertainty of 2.6% is considered for the integrated luminosity of the CMS 8 TeV data [44]. The uncertainty in the amount of pileup interactions per event is estimated by varying total inelastic pp cross section [45] by \( \pm 5\% \). The data-to-simulation correction factors for the trigger efficiency, muon reconstruction and selection efficiencies are estimated using a “tag-and-probe” method [46] in Drell-Yan data and MC samples. These uncertainties cover the pileup dependence of the correction factors. For the jet energy scale, the energy of the jet is varied within a set of uncertainties depending on the jet \( p_T \) and \( \eta \). The jet smearing corrections are altered within their uncertainties [36] to account for the uncertainty arising from the jet energy resolution. The uncertainties associated with the data-to-simulation correction factor for the b tagging efficiencies and misidentification rates are also propagated as systematic uncertainties to the final results [37]. Finally, uncertainties in the knowledge of the parton distribution functions [47] are taken into account.
The uncertainties are found to have negligible effect on the signal model parameters where their corresponding effects on the acceptance are taken into account by introducing nuisance parameters to the fit.

7 Results

The analysis of data yields no significant excess over the SM background prediction. Figure 3 shows the $m_{\mu\mu}$ distribution in data together with the best fit output for a signal-plus-background model at $m_{a_1} = 35$ GeV.

Figure 3: The best fit output to the data for a signal-plus-background model at $m_{a_1} = 35$ GeV, including the uncertainties.

The upper limit on $\sigma_{gg}\to h \times B(h \to a_1 a_1 \to \mu^+ \mu^- b\bar{b})$ is obtained at 95% confidence level (CL) using an asymptotic CLs method [48–51] with systematic uncertainties treated as nuisance parameters. Assuming the SM cross section for the $gg \to h$ process within its theory uncertainty, an upper limit is placed on $B(h \to a_1 a_1 \to \mu^+ \mu^- b\bar{b})$ using the same procedure. Limits are evaluated as a function of $m_{a_1}$.

The observed and expected medians together with the corresponding uncertainty bands are illustrated in Fig. 4 for both scenarii.

8 Summary

A search for $h \to a_1 a_1 \to \mu\mu b\bar{b}$ signal, motivated in NMSSM and extensions to Two Higgs Doublet models, is carried out using 19.7 fb$^{-1}$ of proton collisions at $\sqrt{s} = 8$ TeV. No statistically significant excess is found in data with respect to the SM background prediction. The results of the analysis are presented in the form of upper limits, at 95% CL, on the Higgs cross section times $B(h \to a_1 a_1 \to \mu^+ \mu^- b\bar{b})$ as well as on the Higgs branching ratio to $\mu\mu b\bar{b}$ final state.

References

Figure 4: Observed and expected upper limits at 95% CL on the Higgs boson production times $B(h \rightarrow a_1a_1 \rightarrow \mu^+\mu^-bb)$ (a) and the branching ratio (b) for all $m_{a_1}$ values.


[18] ATLAS Collaboration, “Search for Higgs bosons decaying to $aa$ in the $\mu\mu\tau\tau$ final state in $pp$ collisions at $\sqrt{s} = 8$ TeV with the ATLAS experiment”, arXiv:1505.01609.


