

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 postprint

Authors: Aad, Georges and others

Date: 2016-09-09T00:00:00+00:00

Abstract

A search for the decay to a pair of new particles of either the 125 GeV Higgs boson (h) or a second CP-even Higgs boson (H) is presented. The dataset corresponds to an integrated luminosity of 20.3 fb^{-1} of pp collisions at $\sqrt{s}=8$ TeV recorded by the ATLAS experiment at the LHC in 2012. The search was done in the context of the next-to-minimal supersymmetric standard model, in which the new particles are the lightest neutral pseudoscalar Higgs bosons (a). One of the two a bosons is required to decay to two muons while the other is required to decay to two τ -leptons. No significant excess is observed above the expected backgrounds in the dimuon invariant mass range from 3.7 GeV to 50 GeV. Upper limits are placed on the production of $h \rightarrow aa$ relative to the Standard Model $gg \rightarrow h$ production, assuming no coupling of the a boson to quarks. The most stringent limit is placed at 3.5% for $m_a=3.75$ GeV. Upper limits are also placed on the production cross section of $H \rightarrow aa$ from 2.33 pb to 0.72 pb, for fixed $m_a = 5$ GeV with m_H ranging from 100 GeV to 500 GeV.

Full Text

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

The ATLAS Collaboration

Abstract

A search for the decay to a pair of new particles of either the 125 GeV Higgs boson (h) or a second CP-even Higgs boson (H) is presented. The dataset corresponds to an integrated luminosity of 20.3 fb^{-1} of pp collisions at $\sqrt{s} =$

8 TeV recorded by the ATLAS experiment at the LHC in 2012. The search was performed in the context of the next-to-minimal supersymmetric standard model, in which the new particles are the lightest neutral pseudoscalar Higgs bosons (a). One of the two a bosons is required to decay to two muons while the other is required to decay to two τ -leptons. No significant excess is observed above the expected backgrounds in the dimuon invariant mass range from 3.7 GeV to 50 GeV. Upper limits are placed on the production of $h \rightarrow aa$ relative to the Standard Model $gg \rightarrow h$ production, assuming no coupling of the a boson to quarks. The most stringent limit is placed at 3.5% for $m_a = 3.75$ GeV. Upper limits are also placed on the production cross section of $H \rightarrow aa$ from 2.33 pb to 0.72 pb, for fixed $m_a = 5$ GeV with m_H ranging from 100 GeV to 500 GeV.

I. INTRODUCTION

Since its introduction, the Standard Model (SM) has successfully predicted several new particles, culminating in the discovery of the Higgs boson (h) in July 2012 \cite{higgs_atlas, higgs_cms}. The h boson discovered by ATLAS and CMS appears to have properties consistent with those predicted by the SM within current experimental uncertainties \cite{atlas_higgs_props, cms_higgs_props}. While the many achievements of the SM are remarkable, several unanswered questions remain, such as the hierarchy problem \cite{weinberg_hierarchy, gildener_hierarchy, weinberg_2, susskind} and the nature of the observed dark-matter abundance in the universe. By introducing an additional symmetry between fermions and bosons, Supersymmetric (SUSY) extensions [?, ?, ?, ?, ?, ?, ?, ?, ?] of the SM are able to address both of these issues.

The minimal supersymmetric SM (MSSM) [?, ?, ?, ?, ?] is the simplest extension to the SM that incorporates supersymmetry. It predicts four additional Higgs bosons, generally assumed to be heavier than the h boson: two neutral states, the H and A bosons, as well as two charged states, the H^\pm charged bosons. However, the recent discovery of only an h boson and several null-result searches for associated enhancements in Higgs boson decay rates \cite{atlas_searches, atlas_searches2, cms_searches} have made the model less attractive. Specifically, the measured mass of the Higgs boson \cite{cms_higgs_mass, atlas_higgs_mass} close to 125 GeV results in the reintroduction of a small hierarchy problem in the MSSM, which is removed in the next-to-MSSM (NMSSM) [?, ?] by allowing for additional contributions to the mass of the Higgs boson from new scalar particles. In particular, the NMSSM contains an additional pseudoscalar Higgs boson (a), generally assumed to have a mass lower than the h boson since its mass is protected by a Peccei-Quinn symmetry [?]. The 95% confidence level (CL) upper limit on the branching ratio (BR) of the h boson decaying to non-SM particles is independently set by the ATLAS and CMS experiments to be 37% and 49% respectively \cite{atlas_br_limit, cms_br_limit}.

The analysis presented in this paper targets the $H \rightarrow aa$ decay for a boson

masses above the τ -lepton pair-production threshold. The current lower limit on m_H for such decays was set by the ALEPH experiment at LEP to be $m_H > 107$ GeV for $\text{BR}(H \rightarrow aa)=1$, under the assumption that the H boson couples to the Z boson with SM-strength Zh coupling [?]. The $D\bar{O}$ experiment at the Tevatron also set limits on $H \rightarrow aa$ in this m_a range in the $\mu\mu\tau\tau$ channel [?]. At the LHC, the CMS experiment looked for direct a boson production with $a \rightarrow \mu\mu$, with m_a between 5.5 GeV and 14 GeV [?]. Another search for the a boson below the τ -lepton pair-production threshold has been performed by the CMS experiment, which looked for the 4μ final state [?].

This analysis uses data corresponding to an integrated luminosity of 20.3 fb^{-1} of pp collisions at $\sqrt{s} = 8$ TeV recorded by the ATLAS detector at the LHC in 2012. A search is performed for events consistent with the $H \rightarrow aa$ process, where one a boson decays to two muons and the other decays to a pair of τ -leptons with at least one τ -lepton decaying to an electron or muon. The search is performed in a range of m_a values from 3.7 GeV to 50 GeV, with m_H set to 125 GeV to be consistent with the measured mass of the h boson, m_h [?], [?]. A range of m_H values from 100 GeV to 500 GeV, with m_a set to 5 GeV, is also considered. For the purposes of this analysis, the assumption is made that there is no coupling of the a boson to quarks, and the event selection has been optimized for sensitivity to $m_a \lesssim 10$ GeV (approximately the $b\bar{b}$ production threshold), for which the decay products of either a boson are likely to overlap.

In this range, restricting the decays of one a boson to a pair of muons (rather than allowing for both to decay to pairs of τ -leptons) reduces the total production rate of the signal by a factor of approximately 100. This decrease in signal efficiency is accepted in exchange for a very high trigger efficiency, a larger signal-to-background ratio, and an expected narrow resonance in the dimuon invariant mass ($m_{\mu\mu}$) spectrum. The latter is used to discriminate between background and signal hypotheses based on templates derived in both data and simulation.

II. ATLAS DETECTOR

The ATLAS experiment [?], [?] at the LHC is a multipurpose particle detector with a forward-backward symmetric cylindrical geometry and a near 4π coverage in solid angle.¹ It consists of an inner tracking detector surrounded by a thin superconducting solenoid providing a 2 T axial magnetic field, electromagnetic and hadron calorimeters, and a muon spectrometer. The inner tracking detector covers the pseudorapidity range $|\eta| < 2.5$. It consists

¹ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the center of the detector and the z -axis along the beam pipe. The x -axis points from the IP to the center of the LHC ring, and the y -axis points upwards. Cylindrical coordinates (r, ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the z -axis. The pseudorapidity is defined in terms of the polar angle θ as $\eta = -\ln \tan(\theta/2)$. Angular distance is measured in units of $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

of silicon pixel, silicon microstrip, and transition radiation tracking detectors. Lead/liquid-argon (LAr) sampling calorimeters provide electromagnetic (EM) energy measurements with high granularity. A hadron (iron/scintillator-tile) calorimeter covers the central pseudorapidity range ($|\eta| < 1.7$). The endcap and forward regions are instrumented with LAr calorimeters for EM and hadronic energy measurements up to $|\eta| = 4.9$. The muon spectrometer surrounds the calorimeters and is based on three large air-core toroid superconducting magnets with eight coils each. Its bending power is in the range from $2.0 \text{ T} \cdot \text{m}$ to $7.5 \text{ T} \cdot \text{m}$. It includes a system of precision tracking chambers and fast detectors for triggering. A three-level trigger system is used to select events. The first-level trigger is implemented in hardware and uses a subset of the detector information to reduce the accepted rate to at most 75 kHz. This is followed by two software-based trigger levels that together reduce the accepted event rate to 400 Hz on average depending on the data-taking conditions during 2012.

III. MONTE CARLO SIMULATION

Monte Carlo simulations are used in this analysis to model the $H \rightarrow aa$ signal process and to roughly estimate the composition of the SM background, which is then fitted to the data, with the exception of the low-mass SM resonances (J/ψ , ψ' , $\Upsilon(1S)$, $\Upsilon(2S)$, and $\Upsilon(3S)$). All Monte Carlo (MC) simulated samples are obtained with the full ATLAS detector simulation [\cite{atlas_simulation}](#) in which the particle propagation through the detector is modeled with GEANT4 [\[?\]](#). The effect of multiple proton-proton collisions from the same or nearby beam bunch crossings (in-time and out-of-time pile-up, respectively) is incorporated into the simulation by overlaying additional minimum-bias events generated with the Pythia8 generator [\[?\]](#) onto events with a hard scattering. Simulated events are weighted to match the distribution of the average number of interactions per bunch crossing observed in data, but are otherwise reconstructed in the same manner as data. The event generators, cross sections, underlying-event parameter tunes, and the parton distribution function (PDF) sets used for simulating the SM background processes are summarized in Table I.

In each simulated $H \rightarrow aa$ signal event, a scalar H boson is produced via gluon fusion and is then forced to always decay to two pseudoscalar a bosons. Both the scalar H and pseudoscalar a bosons are expected to have narrow widths, which are negligible when compared to the detector resolution. The H width is set to the value for a SM Higgs boson with the same mass and the a boson width is set to 1 MeV [\cite{cms_direct}](#). This analysis searches for a resonance in the $m_{\mu\mu}$ spectrum and is therefore insensitive to the resolution for m_H . The a bosons are each allowed to decay only to either two muons or two τ -leptons. The decay rate of $a \rightarrow \mu\mu$ is expected to depend on the mass of the a boson as well as the muon and τ -lepton masses [\[?\]](#):

$$\frac{\Gamma(a \rightarrow \mu\mu)}{\Gamma(a \rightarrow \tau\tau)} = \frac{m_\mu^2}{m_\tau^2} \sqrt{\frac{1 - (2m_\mu/m_a)^2}{1 - (2m_\tau/m_a)^2}}$$

The width values calculated from this equation are used by Pythia to generate signal events, under the assumption that $\text{BR}(a \rightarrow \mu\mu) = 1 - \text{BR}(a \rightarrow \tau\tau)$. $\text{BR}(a \rightarrow \mu\mu)$ ranges from 1.25% for $m_a = 3.7$ GeV down to 0.35% for $m_a = 50$ GeV.

IV. EVENT SELECTION

The data used for this analysis were collected using a combination of a single-muon trigger and an asymmetric dimuon trigger. The transverse momentum (p_T) threshold for the single-muon trigger is 36 GeV, while the dimuon trigger thresholds for the highest p_T (leading) and the second-highest p_T (subleading) muons are 18 GeV and 8 GeV, respectively [\cite{atlas_trigger}](#). Muons are reconstructed by combining tracks in the inner detector (ID) with those in the muon spectrometer (MS) [\cite{atlas_muons}](#) and are selected if they satisfy the following criteria: each muon must have $p_T > 16$ GeV, $|\eta| < 2.5$, projected longitudinal impact parameter $|z_0 \sin \theta| < 0.4$ mm, the ratio of the transverse impact parameter d_0 to its estimated uncertainty σ_{d_0} such that $|d_0|/\sigma_{d_0} < 3$, and a track isolation requirement. The transverse and longitudinal impact parameters are defined with respect to the primary vertex, defined as the vertex with the largest $\sum p_T^2$ of the associated tracks. The track isolation of each muon is defined as the scalar sum of the p_T of all tracks with $p_T > 1$ GeV, in a surrounding $\Delta R = 0.3$ cone, excluding those tracks associated with either muon. The track isolation is required to be less than 12% of the muon p_T .

The selected muons are organized into oppositely charged pairs, which are then ordered by the vector sum of the p_T of the constituent muons. The leading pair is identified as the $a \rightarrow \mu\mu$ candidate. Events are required to have an $a \rightarrow \mu\mu$ candidate with $p_T > 40$ GeV.

A. Signal regions

Two signal regions (SR) are designed to select $H \rightarrow aa$ events where one a boson directly decays to a pair of muons, and the other to a pair of τ -leptons. The regions are optimized for the scenario where one τ -lepton decays to either a muon or an electron and the other τ -lepton is identified by selecting one to three additional tracks.² Each signal event must have an $a \rightarrow \mu\mu$ candidate

²The signal region has been designed to select events in which the charged decay product(s) of the second τ -lepton is either an electron, a muon, a single charged hadron or three charged hadrons; each decay product is identified as a selected track. For the signals targeted by this analysis, one or more of the tracks in the last case can have p_T insufficient to be selected. This results in final states with less than three reconstructed tracks.

with a dimuon invariant mass between 2.8 GeV and 70 GeV.³ Additionally, the event must contain a third lepton (muon or electron), with the flavor indicated in the SR name (i.e. SR μ or SR e). In the case where the third lepton is a reconstructed muon, it must have $p_T > 7$ GeV and $|\eta| < 2.5$. If the third lepton is a reconstructed electron, it is required to satisfy identification requirements optimized for good efficiency extending to low p_T \cite{atlas_electrons}, must have $p_T > 7$ GeV and $|\eta| < 2.47$. Electrons in the transition region between the barrel and endcap EM calorimeters ($1.37 < |\eta| < 1.55$) are rejected. The two a bosons are expected to be produced back-to-back in the transverse plane. This topology is enhanced by requiring the azimuthal separation between the tagged $a \rightarrow \mu\mu$ candidate and the third lepton to be greater than 1 rad.

Furthermore, the two τ -leptons from the decay of the a boson tend to be highly collimated for low values of m_a . Therefore, the third lepton (muon or electron) is required to have one, two, or three tracks, in a cone of $\Delta R = 0.4$, not including the track of the lepton itself. Tracks are reconstructed in the ID and are required to have $p_T > 1$ GeV, $|\eta| < 2.5$, transverse and longitudinal impact parameters, $|d_0|$ and $|z_0|$, less than 1 mm, at least seven hits in the two silicon tracking detectors, and one hit in the innermost layer of the pixel detector if such a hit is expected. The leading p_T track within the $\Delta R = 0.4$ cone around the third lepton, but not matched to the lepton itself, is denoted the lead-track. It is used to approximate the axis of the second τ -lepton from the $a \rightarrow \tau\tau$ decay. The $\Delta R = 0.2$ cone, centered on this axis, may have no more than three tracks, not including the track of the third lepton. The charge of the lead-track is taken to be the same as its τ -lepton ancestor and is required to be opposite to that of the lepton. Finally, the third lepton is required to pass a cut on its track isolation, defined as the scalar sum of the p_T of the tracks that are within a cone of $\Delta R = 0.4$. The lead-track is excluded from this calculation. The track isolation is required to be less than 12% of the lepton's p_T .

If the event has more than one candidate to be the third lepton, the leading muon is used to define the signal region. If no third muon is identified, the leading electron is used. Events with muon as well as electron candidates are classified as SR μ events in order to maximize the signal-to-background ratio. No veto is applied to events with more than three leptons in order to maintain signal selection efficiency for events where both τ -leptons decay to muons or electrons. In these cases, the additional lepton may be selected as the lead-track. Table II shows the selection efficiency for a signal simulated with $m_a = 5$ GeV and $m_H = m_h = 125$ GeV.

³The $m_{\mu\mu}$ distribution, and ultimately used to fit the observed $m_{\mu\mu}$ spectra, is wider than the range of signal hypotheses (3.7 GeV to 50 GeV) in order to be able to increase the stability of the fit when considering the low-mass SM resonances and the low-end tail of the SM Z resonance.

B. Validation and control regions

Two validation regions (VR) are used to test the methods of the analysis in a signal-free environment. The background in the signal region is expected to be dominated by events with jets that pass the $a \rightarrow \tau\tau$ selection outlined above. Furthermore, the charges of the reconstructed lepton and lead track in such events are not correlated. Events in the validation regions are therefore selected with the same criteria as signal events, with the sole exception that the lepton and lead-track are required to have the same charge. The signal yield in the validation regions is expected to be less than 5% (10%) of that in the muon (electron) signal region, as determined in simulation. As done for the signal region, the two regions are denoted VR_μ and VRe based on the flavor of the third lepton.

Two control regions (CR) are used in this analysis to constrain the SM background, one light-flavor-dominated region CR_j and one heavy-flavor-dominated region CR_b . Events in each control region are required to have an $a \rightarrow \mu\mu$ candidate with a mass between 2.8 GeV (15 GeV) and 70 GeV for CR_j (CR_b).⁴ The two regions are further defined by imposing requirements on jets in the event. The light-flavor-dominated region is used to measure the nonresonant Drell-Yan background as well as the low-mass SM resonances. It is defined by requiring at least one selected jet and exactly zero b -jets. The heavy-flavor-dominated region must have at least two b -jets and is used to measure the $t\bar{t}$ background. All events in the signal regions are explicitly removed from the control regions.

Jets are reconstructed using the anti- k_t algorithm [?] and radius parameter $R = 0.4$ with three-dimensional energy clusters, measured in the calorimeter `\cite{atlas_clusters}`, as input. The calibration of the clusters is performed by applying different weights to the energy deposits arising from the electromagnetic and hadronic components of the showers. The final calibration of the jet energy corrects the response of the calorimeter to match the particle-level jet energy `\cite{atlas_jet_calib}1, atlas{jet}{calib}2`. *Corrections are first determined in simulation and then improved and validated using data. Additional corrections for pile-up from other pp interactions in the same or neighboring bunch crossings are applied as well \cite{atlas_pileup}*. Jets are required to have $p_T > 40$ GeV and $|\eta| < 2.5$. Jets with p_T between 40 GeV and 50 GeV are further required to pass a cut on the jet vertex fraction (JVF), which is defined as the p_T -weighted fraction of tracks inside the jet that are determined to originate from the primary vertex of the event. The JVF is required to be greater than 0.25, but is not applied to jets that have no tracks. Jets can be tagged as containing a b -hadron using a multivariate b -tagging algorithm, and then denoted b -jets. The algorithm is based on the impact parameters of tracks and information from reconstructed secondary vertices `\cite{atlas_btag}`. The operating point chosen for the b -tagging algorithm corresponds to a true b -jet efficiency of 60% and purity of 95%, as determined in $t\bar{t}$ simulation.

⁴The 15 GeV lower $m_{\mu\mu}$ limit for CR_b is chosen such that the region is dominated by the $t\bar{t}$ background.

V. SIGNAL AND BACKGROUND MODELS

An unbinned, log-likelihood fit is performed on the observed dimuon invariant mass spectra in the signal regions (SR μ and SRe) to a combination of background and signal models, the details of which are described in this section.

A. Signal model

A double-sided Crystal Ball (CB) function [\cite{crystal_{ball}1}](#), [crystal_{ball}2](#) is used for the signal model; it was found to be the simplest function that was robust enough to describe the shape of the simulated signal resonances and the observed SM resonances.⁵ The Crystal Ball function is composed of a Gaussian core, with mean μ_{CB} and width σ_{CB} , and power-law distributions of orders 2.5 and 10 for the low-end and high-end tails. The threshold parameter α_{CB} , which is in units of σ_{CB} , determines the point of transition from the core to either tail. The mean of the Crystal Ball function is assumed to be proportional to m_a , with slope a_μ , while its width is assumed to be linearly dependent on both m_a and m_H with slopes a_σ and b_σ respectively.

The Crystal Ball function is used to model the line-shape of the $a \rightarrow \mu\mu$ resonance. A small contribution from $a \rightarrow \tau\tau \rightarrow \mu\mu + 4\nu$ is included as a Gaussian distribution, which, due to the kinematics of the $\tau \rightarrow \mu + 2\nu$ decay, is expected to have a lower mean and worse resolution than the $a \rightarrow \mu\mu$ resonance. The mean, $\mu_{\tau\tau}$, and width, $\sigma_{\tau\tau}$, are set proportional to the corresponding parameters of the Crystal Ball function with the parameters $k_\mu < 1$ and $k_\sigma > 1$. The fraction of $a \rightarrow \tau\tau \rightarrow \mu\mu + 4\nu$ in the total signal is given by $f_{\tau\tau}$. The full signal model is:

$$P_a = \text{CB}(m_{\mu\mu}|\mu_{CB}, \sigma_{CB}) + f_{\tau\tau} \text{GA}(m_{\mu\mu}|\mu_{\tau\tau}, \sigma_{\tau\tau})$$

where:

$$\begin{aligned}\mu_{CB} &= a_\mu \cdot m_a \\ \mu_{\tau\tau} &= k_\mu \cdot \mu_{CB} \\ \sigma_{\tau\tau} &= k_\sigma \cdot a_\sigma \cdot m_a \\ \sigma_{CB} &= a_\sigma \cdot m_a + b_\sigma \cdot (m_H - 100 \text{ GeV})\end{aligned}$$

The values of the parameters b_σ , k_μ , k_σ , and $f_{\tau\tau}$ are determined by fitting the signal simulation, while the parameters α_{CB} , a_μ , and a_σ are measured in data following the procedure described in Sec. V B, and are found to be consistent with simulation. Figure 1 [Figure 1: see original paper] shows the result of a

⁵For a given resonance X , the double-sided Crystal Ball function has been simplified to remove parameters that are common to all resonances. The expression $\text{CB}(m_{\mu\mu}|\mu_X, \sigma_X, \alpha_L = \alpha_{CB}, \alpha_H = \alpha_{CB}, n_L = 2.5, n_H = 10)$ has been shortened to $\text{CB}(m_{\mu\mu}|\mu_X, \sigma_X)$.

simultaneous fit of both signal regions to all simulated signal samples, projected into $\text{SR}\mu$ for one benchmark mass point with $m_a = 5$ GeV and $m_H = m_h = 125$ GeV.

B. Background model

The full background model consists of several pieces: six SM resonances (J/ψ , ψ' , $\Upsilon(1S)$, $\Upsilon(2S)$, $\Upsilon(3S)$, Z), a $t\bar{t}$ component, and one piece for the nonresonant continuum background (dominated by low-mass Drell-Yan events). Each SM resonance is modeled by the same double-sided Crystal Ball function used for the signal $a \rightarrow \mu\mu$ resonance (see Sec. V A), with the offset b_σ set to zero. The mean, μ_X , and width, σ_X , of each resonance X are assumed to be linearly dependent on its mass (m_X) with the same slopes as in the signal model. The measured value of the mass of each resonance is found to be consistent with the PDG best-fit value [?] and are therefore constrained to the PDG value and its associated uncertainty. The low-mass resonances are combined into two composite models, ψ and Υ , by adding the resonances of the higher spin states with fractions $f_{\psi'}$, $f_{\Upsilon 2S}$, and $f_{\Upsilon 3S}$, as shown in Eqs. 4 and 5.

The background from $t\bar{t}$ production is modeled with a Rayleigh distribution [?], defined by multiplying $m_{\mu\mu}$ by a Gaussian distribution with mean set to zero and width, $\sigma_{t\bar{t}}$. The mostly Drell-Yan continuum background is modeled by an exponential decay function, with parameter $\alpha_{\gamma^*} < 0$, multiplied by $m_{\mu\mu}$ raised to the power n_{γ^*} . The full expression for the continuum background is:

$$P_{\gamma^*} = m_{\mu\mu}^{n_{\gamma^*}} e^{\alpha_{\gamma^*} m_{\mu\mu}}$$

A contribution to the background from $b\bar{b}$ production followed by two semileptonic decays of b -hadrons to muons was also considered and found to be small in the signal region, with a $m_{\mu\mu}$ shape similar to the Drell-Yan component but with a rate about 1% as large. This is expected since dimuon events from $b\bar{b}$ are highly suppressed after applying the muon isolation requirements, the muon p_T cuts, and the dimuon p_T cut. Events from double semileptonic b -hadron decays ($b \rightarrow c\mu + X \rightarrow \mu\mu + X$) are found to contribute as well, but only for $m_{\mu\mu} < 3.5$ GeV (which is below the signal region) at a rate of about 10% that of Drell-Yan events.

Lastly, the full model in each region (CRj, CRb, $\text{SR}\mu$ or SRe) is defined by adding four background models: one for the ψ and Υ resonances, one for the $t\bar{t}$ component, and one for the Drell-Yan component. The shape parameters of each component, the fractional contribution of higher spin states, and the relative contribution of Z boson to the total Drell-Yan production (f_Z) are constrained by simultaneously fitting the control regions as described in Sec. VI. These parameters are therefore assumed to be the same in the control and signal regions. The relative contributions of each of the four background components (quantified by the independent parameters $f_{t\bar{t}}$, f_Υ , and f_{Res}) are expected to

vary between the regions and are therefore measured in the fit to the signal regions (SR μ and SR e). The complete background model is:

$$P_b = m_{\mu\mu}^{n_{\gamma^*}} e^{\alpha_{\gamma^*} m_{\mu\mu}} + f_Z \text{CB}(m_{\mu\mu} | \mu_Z, \sigma_Z) + f_{t\bar{t}} R(m_{\mu\mu} | \sigma_{t\bar{t}}) + f_{\text{Res}}(P_\psi + f_\Upsilon P_\Upsilon)$$

where:

$$P_\psi = \text{CB}(m_{\mu\mu} | \mu_{J/\psi}, \sigma_{J/\psi}) + f_{\psi'} \text{CB}(m_{\mu\mu} | \mu_{\psi'}, \sigma_{\psi'})$$

$$P_\Upsilon = \text{CB}(m_{\mu\mu} | \mu_{\Upsilon 1S}, \sigma_{\Upsilon 1S}) + f_{\Upsilon 2S} (\text{CB}(m_{\mu\mu} | \mu_{\Upsilon 2S}, \sigma_{\Upsilon 2S}) + f_{\Upsilon 3S} \text{CB}(m_{\mu\mu} | \mu_{\Upsilon 3S}, \sigma_{\Upsilon 3S}))$$

and:

$$\mu_X = a_\mu \cdot m_X; \quad \sigma_X = a_\sigma \cdot m_X; \quad X = \{J/\psi, \psi', \Upsilon 1S, \Upsilon 2S, \Upsilon 3S, Z\}$$

VI. BACKGROUND MEASUREMENT

A simultaneous fit of the full background model is performed on the $m_{\mu\mu}$ data in the two control regions (CRj and CRb), with $m_{\mu\mu}$ in the range from 2.8 GeV (15 GeV) to 70 GeV, for CRj (CRb). The results of the fit are used to constrain the dependent parameters of the four background components as described in Sec. V B. The fitted values and uncertainties of the relevant parameters are reported in Table III and the fitted $m_{\mu\mu}$ spectra in the two regions are shown in Fig. 2 [Figure 2: see original paper]. A large correlation is found between a_μ , a_σ , and α_{CB} while a large anti-correlation is found between $\sigma_{t\bar{t}}$ and f_Z . The figure also includes a comparison with Drell-Yan and $t\bar{t}$ simulation normalized to the data.⁶ In both control regions, the data are well described by the background model, as indicated by the small residuals shown in the figure and by the good agreement with simulation. The large number of events in the control regions provides a tight constraint on the parameters of the model. The small shift of the resonance mean masses to $(99.86 \pm 0.01)\%$ of their PDG values [?], is consistent with the $\mathcal{O}(0.1\%)$ systematic uncertainty on the muon energy scale. The fractional mass resolution, a_σ , is found to be $(1.68 \pm 0.02)\%$, corresponding to an (84 ± 1) MeV width for a 5 GeV dimuon resonance.

VII. SYSTEMATIC UNCERTAINTIES

A. Signal model

In the signal model for the $a \rightarrow \mu\mu$ resonance, the parameters α_{CB} , a_μ , and a_σ are limited by the experimental resolution, and are thus assumed to be 100%

⁶The Drell-Yan simulation has only been produced with $m_{\mu\mu}$ above 10 GeV, so is only valid for comparing to the data in that region.

correlated with the corresponding parameters of the fits to the SM resonances and are measured in the observed CRj data. To account for the extrapolation uncertainty from parameters for the SM resonances to those for signal, an additional uncertainty is assigned to the Crystal Ball function's threshold parameter, α_{CB} . Due to the large correlation between the three parameters, this additional uncertainty assigned to α_{CB} covers the uncertainty of the other two. The uncertainty is determined by fitting the signal model in each simulated signal sample separately and taking the maximum difference between them. Another assumption in the signal model is that the parameter, b_σ , and the Gaussian description of the $a \rightarrow \tau\tau \rightarrow \mu\mu + 4\nu$ tail, parameterized by k_μ , k_σ , and $f_{\tau\tau}$, are properly described in the simulation. The systematic uncertainty on α_{CB} covers the uncertainty on b_σ , since it is strongly correlated with α_{CB} , a_μ , and a_σ . The other three parameters are similarly correlated with one another and, thus, only one additional systematic uncertainty is introduced. The final results of the analysis are found to have little sensitivity to the amount of $a \rightarrow \tau\tau \rightarrow \mu\mu$, so a conservative 50% systematic uncertainty is assigned to $f_{\tau\tau}$.

B. Signal normalization

The dominant source of systematic uncertainty on the signal normalization is found to be the theoretical uncertainty on the rate of SM $gg \rightarrow h/H$ production. In the m_H range relevant for this analysis (from 100 GeV to 500 GeV) the total uncertainty varies from 10% to 11% and is determined from the spread of the cross-section predictions using different PDF sets and their associated uncertainties, as well as from varying the factorization and renormalization scales [?]. A constant 11% is used in this analysis. The next largest systematic uncertainty is on the p_T resolution of the lead-track and is found to be 5%. This uncertainty on the signal normalization is determined by varying the p_T of each track by a conservative $\pm 2\%$, and propagating the effect through the full analysis. Additional sources of systematic uncertainty include those on the trigger efficiency, the lepton reconstruction efficiency, the lepton energy scale and resolution, and the charge of the track. All of these sources were found to contribute a negligible amount to the total uncertainty on the normalization of the signal.

C. Background model

The results of the background measurement reported in Sec. VI are fitted values and associated uncertainties of the dependent parameters of the four background components—Drell-Yan, $t\bar{t}$, ψ , and Υ . In the fit to the signal regions, each parameter is constrained by a Gaussian prior with a mean equal to its fitted value from the background measurement and width equal to the corresponding uncertainty.

Two general assumptions are made in the construction of the background model. First, the chosen functional form accurately describes the background in the signal regions. Second, the dependent parameters are the same in each region. Both of these assumptions introduce a potential bias in the final result, whereby

a nonzero signal may be observed incorrectly. The tolerance of the background model for such a spurious signal is measured for values of m_a and m_H corresponding to each simulated signal point. The measurement is performed using a large sample of signal-free events. Simulated $t\bar{t}$ events with an identified $a \rightarrow \mu\mu$ candidate are used as a large sample of events for the $t\bar{t}$ background component. In lieu of simulation, the observed data in the light-flavor-dominated control region are used for the Drell-Yan component. The simulated $t\bar{t}$ events are then weighted and combined with the data such that the relative contribution of $t\bar{t}$ matches the simulation-based expectation in the signal region with $m_{\mu\mu}$ between 20 GeV and 60 GeV.⁷ Finally, the resulting sample is scaled to the expected normalization of the signal region. The combined signal and background model is fit to the large sample of events. The potential bias is taken to be the measured rate of spurious signal $+1\sigma$ and is found to be between 0.2% and 3.2% of the signal rate normalized to the SM $gg \rightarrow h$ production rate with $\text{BR}(h \rightarrow aa)=100\%$, with a maximum at $m_a = 20$ GeV and $m_H = m_h = 125$ GeV. The measured bias is taken as an additional uncertainty on the signal normalization.

VIII. VALIDATION OF METHODS

To test the methods of the analysis, two validation regions, VR_μ and VR_e (as defined in Sec. IV B), are used in place of SR_μ and SR_e . The validation regions are designed to have properties similar to the signal regions and to also test the robustness of the method against variations in the background composition, since no a priori assumptions are made about the relative contributions from $t\bar{t}$, Drell-Yan, J/ψ or Υ . Furthermore, the validation checks for non-negligible backgrounds with dimuon invariant mass distributions that are unaccounted for in the background model. The results of the simultaneous fit of the full background model (including all relevant systematic uncertainties) to the validation regions are shown in Fig. 3 [Figure 3: see original paper].

In the fit, all region-independent parameters are constrained by the results of the fit to the control regions (reported in Table III). The strong correlations between some of the region-independent parameters, which are reported in Sec. VI, are not explicitly accounted for in the fit to the two validation regions. This simplification is found to have a negligible effect on the results of the fit. The consistency with the background-only model is evaluated by scanning the local p -value as a function of $m_{\mu\mu}$ from 3.7 GeV to 50 GeV, calculated using frequentist hypothesis tests based on the profile-log-likelihood ratio test statistic and approximated with the asymptotic formulae [?]. The p -values are evaluated in 50 MeV intervals below 15 GeV, then 100 MeV intervals up to 30 GeV, and 200 MeV intervals up to $m_{\mu\mu} = 50$ GeV. The minimum local p -value is found for $m_{\mu\mu} = 47.4$ GeV to be 0.0074, corresponding to a local significance of 2.44σ . Correcting for the look-elsewhere effect $\backslash\text{cite}\{\text{gross}_{\text{vitells}}\}$ gives a global p -

⁷For the spurious signal calculation, a narrower $m_{\mu\mu}$ range (20 GeV to 60 GeV instead of 2.8 GeV to 70 GeV) is used to scale the $t\bar{t}$ simulation because it is the range in which the $t\bar{t}$ background is expected to dominate.

value of 0.31, indicating that at least one excess of this magnitude, or larger, is expected from background fluctuations in at least 31% of experiments.

IX. RESULTS

A simultaneous fit of the full background model is performed on the $m_{\mu\mu}$ spectra in the two signal regions (SR μ and SR e), with $m_{\mu\mu}$ in the range from 2.8 GeV to 70 GeV. The observed $m_{\mu\mu}$ distribution and the background-only fit are shown in Fig. 4 [Figure 4: see original paper]; the data are well described by the fit. In the fit, all region-independent parameters are constrained by the results of the fit to the control regions, reported in Table III. The strong correlations between some of the region-independent parameters, which are reported in Sec. VI, are not explicitly accounted for in the fit to the two signal regions. This simplification is found to have a negligible effect on the results of the fit. The fitted values and uncertainties of the remaining parameters, as well as the corresponding values from the fits to the control and validation regions, are shown in Table IV. No significant correlations are found between the parameters listed in Table IV.

The consistency with the background-only model is evaluated by scanning the local p -value as a function of $m_{\mu\mu}$ from 3.7 GeV to 50 GeV, using the same calculation, m_a range, and intervals used in the scan of the validation region. The results of this scan are reported in Fig. 5 [Figure 5: see original paper]. The minimum local p -value is found for $m_{\mu\mu} = 8.65$ GeV to be 0.0223, corresponding to a local significance of 2.01σ . Correcting for the look-elsewhere effect $\backslash\text{cite}\{\text{gross_vitells}\}$ gives a global p -value > 0.5 , indicating that at least one excess of this magnitude, or larger, is expected from background fluctuations in at least 50% of experiments.

With no evidence to support the NMSSM hypothesis, a 95% CL limit can be set using the CL_s prescription [?]. Figure 6 [Figure 6: see original paper] shows the observed and expected limits on the rate $(\sigma(gg \rightarrow h) \times \text{BR}(h \rightarrow aa))$ relative to the SM Higgs boson gluon-gluon fusion production cross section (σ_{SM}) , calculated at NLO+NNLL precision [?], as a function of m_a with m_H set to 125 GeV. The limits are evaluated in the same intervals used for the p -value scan. Also shown in the figure is the total rate $(\sigma(gg \rightarrow H) \times \text{BR}(H \rightarrow aa))$ as a function of m_H with m_a set to 5 GeV, evaluated at 50 GeV intervals from $m_H = 100$ GeV to 500 GeV and at $m_H = m_h = 125$ GeV. In both panels of Fig. 6 [Figure 6: see original paper], the observed and expected limits have been scaled by $\text{BR}(a \rightarrow \tau\tau)^2$ to explicitly account for the branching ratios assumed in this analysis and facilitate reinterpretation of the results.

Final states are considered with two muons consistent with the decay of one a boson as well as a third lepton (e or μ) and tracks, consistent with collimated τ -leptons from the other a boson. A scan of the dimuon invariant mass distribution from 3.7 GeV to 50 GeV shows no significant excess of data over SM backgrounds. Limits are set assuming no coupling of the a boson to quarks. The observed 95% CL upper limits on the production rate, $\sigma(gg \rightarrow H) \times \text{BR}(H \rightarrow aa)$, are

consistent with the expected limit and are determined to be from 2.33 pb to 0.72 pb, for m_H between 100 GeV and 500 GeV (and $m_a = 5$ GeV). A 95% CL upper limit for the production of the h boson and its decay rate to two pseudoscalar a bosons is set for m_a from 3.7 GeV to 50 GeV, with the most stringent limit placed at 3.5% for $m_a = 3.75$ GeV.

X. CONCLUSION

A search for the decay of a scalar Higgs boson to two pseudoscalar a Higgs bosons ($H \rightarrow aa$) in the context of the NMSSM is presented with LHC data corresponding to an integrated luminosity of 20.3 fb^{-1} of pp collisions at $\sqrt{s} = 8$ TeV, collected in 2012 by the ATLAS experiment. Final states are considered with two muons consistent with the decay of one a boson as well as a third lepton (e or μ) and tracks, consistent with collimated τ -leptons from the other a boson. A scan of the dimuon invariant mass distribution from 3.7 GeV to 50 GeV shows no significant excess of data over SM backgrounds. Limits are set assuming no coupling of the a boson to quarks. The observed 95% CL upper limits on the production rate, $\sigma(gg \rightarrow H) \times \text{BR}(H \rightarrow aa)$, are consistent with the expected limit and are determined to be from 2.33 pb to 0.72 pb, for m_H between 100 GeV and 500 GeV (and $m_a = 5$ GeV). A 95% CL upper limit for the production of the h boson and its decay rate to two pseudoscalar a bosons is set for m_a from 3.7 GeV to 50 GeV, with the most stringent limit placed at 3.5% for $m_a = 3.75$ GeV.

ACKNOWLEDGMENTS

We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently. We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWFW and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNRF, DNSRC and Lundbeck Foundation, Denmark; EPLANET, ERC and NSRF, European Union; IN2P3-CNRS, CEA-DSM/IRFU, France; GNSF, Georgia; BMBF, DFG, HGF, MPG and AvH Foundation, Germany; GSRT and NSRF, Greece; RGC, Hong Kong SAR, China; ISF, MINERVA, GIF, I-CORE and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; FOM and NWO, Netherlands; BRF and RCN, Norway; MNiSW and NCN, Poland; GRICES and FCT, Portugal; MNE/IFA, Romania; MES of Russia and NRC KI, Russian Federation; JINR; MSTD, Serbia; MSSR, Slovakia; ARRS and MIZŠ, Slovenia; DST/NRF, South Africa; MINECO, Spain; SRC and Wallenberg Foundation, Sweden; SER, SNSF and Cantons of Bern and Geneva, Switzerland; NSC, Taiwan; TAEK, Turkey; STFC, the Royal Society and Leverhulme Trust, United Kingdom; DOE and NSF, United States of America. The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from

CERN and the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (UK) and BNL (USA) and in the Tier-2 facilities worldwide.

REFERENCES

- [1] ATLAS Collaboration, Phys. Lett. B 716, 1 (2012), arXiv:1207.7214 [hep-ex].
- [2] CMS Collaboration, Phys. Lett. B 716, 30 (2012), arXiv:1207.7235 [hep-ex].
- [3] ATLAS Collaboration, ATLAS-CONF-2015-007, <https://cds.cern.ch/record/2002212>.
- [4] CMS Collaboration, (2014), arXiv:1412.8662 [hep-ex]. [5] S. Weinberg, Phys. Rev. D 13, 974 (1976). [6] E. Gildener, Phys. Rev. D 14, 1667 (1976). [7] S. Weinberg, Phys. Rev. D 19, 1277 (1979). [8] L. Susskind, Phys. Rev. D 20, 2619 (1979). [9] H. Miyazawa, Prog. Theor. Phys. 36, 1266 (1966). [10] P. Ramond, Phys. Rev. D 3, 2415 (1971). [11] Y. A. Gol'fand and E. P. Likhtman, JETP Lett. 13, 323 (1971). [12] A. Neveu and J. H. Schwarz, Nucl. Phys. B 31, 86 (1971). [13] A. Neveu and J. H. Schwarz, Phys. Rev. D 4, 1109 (1971). [14] J. L. Gervais and B. Sakita, Nucl. Phys. B 34, 632 (1971). [15] D. V. Volkov and V. P. Akulov, Phys. Lett. B 46, 109 (1973). [16] J. Wess and B. Zumino, Phys. Lett. B 49, 52 (1974). [17] J. Wess and B. Zumino, Nucl. Phys. B 70, 39 (1974). [18] P. Fayet, Phys. Lett. B 64, 159 (1976). [19] P. Fayet, Phys. Lett. B 69, 489 (1977). [20] G. R. Farrar and P. Fayet, Phys. Lett. B 76, 575 (1978). [21] P. Fayet, Phys. Lett. B 84, 416 (1979). [22] S. Dimopoulos and H. Georgi, Nucl. Phys. B 193, 150 (1981). [23] ATLAS Collaboration, J. High Energy Phys. 11, 56 (2014), arXiv:1409.6064 [hep-ex]. [24] ATLAS Collaboration, (2014), arXiv:1412.6663 [hep-ex]. [25] CMS Collaboration, J. High Energy Phys. 10, 160 (2014), arXiv:1408.3316 [hep-ex]. [26] ATLAS Collaboration, Phys. Rev. D 90, 052004 (2014), arXiv:1406.3827 [hep-ex]. [27] U. Ellwanger, J. F. Gunion, and C. Hugonie, (2001), arXiv:hep-ph/0111179 [hep-ph]. [28] D. Curtin et al., Phys. Rev. D 90, 075004 (2014), arXiv:1312.4992 [hep-ph]. [29] M. Lisanti and J. G. Wacker, Phys. Rev. D 79, 115006 (2009), arXiv:0903.1377 [hep-ph]. [30] ATLAS Collaboration (Geneva, 2014) ATLAS-CONF-2014-010, <http://cdsweb.cern.ch/record/1670531>. [31] A. Collaboration, S. Schael, et al., J. High Energy Phys. 05, 49 (2010), arXiv:1003.0705 [hep-ex]. [32] D. Collaboration, V. M. Abazov, et al., Phys. Rev. Lett. 103, 061801 (2009), arXiv:0905.3381 [hep-ex]. [33] CMS Collaboration, Phys. Rev. Lett. 109, 121801 (2012), arXiv:1206.6326 [hep-ex]. [34] CMS Collaboration, Phys. Lett. B 726, 564 (2013), arXiv:1210.7619 [hep-ex]. [35] ATLAS Collaboration, JINST 3, S08003 (2008). [36] ATLAS Collaboration, Eur. Phys. J. C 70, 823 (2010), arXiv:1005.4568 [physics.ins-det]. [37] G. Collaboration, S. Agostinelli, et al., Nucl. Instrum. Meth. A 506, 250 (2003). [38] T. Sjöstrand, S. Mrenna, and P. Z. Skands, Comput. Phys. Commun. 178, 852 (2008), arXiv:0710.3820 [hep-ph]. [39] ATLAS ATL-PHYS-PUB-2012-003, Collaboration, <http://cds.cern.ch/record/1474107>. [40] J. Pumplin et al., J. High Energy Phys. 07, 12 (2002), arXiv:hep-ph/0201195 [hep-ph]. [41] P. Nason, J. High Energy Phys. 11, 40 (2004), arXiv:hep-ph/0409146 [hep-ph].

[42] S. Frixione, P. Nason, and C. Oleari, *J. High Energy Phys.* 11, 70 (2007), arXiv:0709.2092 [hep-ph]. [43] M. Cacciari, M. Czakon, M. L. Mangano, A. Mitov, and P. Nason, *Phys. Lett. B* 710, 612 (2012), arXiv:1111.5869 [hep-ph]. [44] P. Baernreuther, M. Czakon, and A. Mitov, *Phys. Rev. Lett.* 109, 132001 (2012), arXiv:1204.5201 [hep-ph]. [45] M. Czakon and A. Mitov, *J. High Energy Phys.* 12, 054 (2012), arXiv:1207.0236 [hep-ph]. [46] M. Czakon and A. Mitov, *J. High Energy Phys.* 01, 080 (2013), arXiv:1210.6832 [hep-ph]. [47] M. Czakon, P. Fiedler, and A. Mitov, *Phys. Rev. Lett.* 110, 252004 (2013), arXiv:1303.6254 [hep-ph]. [48] M. Czakon and A. Mitov, *Comput. Phys. Commun.* 185, 2930 (2014), arXiv:1112.5675 [hep-ph]. [49] P. Z. Skands, *Phys. Rev. D* 82, 074018 (2010), arXiv:1005.3457 [hep-ph]. [50] H.-L. Lai et al., *Phys. Rev. D* 82, 074024 (2010), arXiv:1007.2241 [hep-ph]. [51] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, *J. High Energy Phys.* 07, 1 (2003), arXiv:hep-ph/0206293 [hep-ph]. [52] T. Sjöstrand, S. Mrenna, and P. Z. Skands, *J. High Energy Phys.* 05, 26 (2006), arXiv:hep-ph/0603175 [hep-ph]. [53] N. Kidonakis, *Phys. Rev. D* 82, 054018 (2010), arXiv:1005.4451 [hep-ph]. [54] G. Corcella et al., *J. High Energy Phys.* 01, 10 (2001), arXiv:hep-ph/0011363 [hep-ph]. [55] ATLAS Collaboration, ATLAS-PHYS-PUB-2011-009, <http://cdsweb.cern.ch/record/1363300>. [56] S. Heinemeyer et al., (2013), arXiv:1307.1347 [hep-ph]. [57] ATLAS Collaboration, *Eur. Phys. J. C* 75, 120 (2014), arXiv:1408.3179 [hep-ex]. [58] ATLAS Collaboration, *Eur. Phys. J. C* 74, 3130 (2014), arXiv:1407.3935 [hep-ex]. [59] ATLAS Collaboration, ATLAS-CONF-2014-032, <http://cds.cern.ch/record/1706245>. [60] M. Cacciari, G. P. Salam, and G. Soyez, *J. High Energy Phys.* 04, 63 (2008), arXiv:0802.1189 [hep-ph]. [61] W. Lampl et al., ATLAS-LARG-PUB-2008-002, <http://cdsweb.cern.ch/record/1099735>. [62] ATLAS Collaboration, *Eur. Phys. J. C* 73, 2304 (2013), arXiv:1112.6426 [hep-ex]. [63] ATLAS Collaboration, *Eur. Phys. J. C* 75, 17 (2015), arXiv:1406.0076 [hep-ex]. [64] ATLAS Collaboration, ATLAS-CONF-2013-083, <http://cds.cern.ch/record/1570994>. [65] ATLAS Collaboration, ATLAS-CONF-2011-102, <http://cdsweb.cern.ch/record/1369219>. [66] M. Oreglia, “A Study of the Reactions $\psi' \rightarrow \gamma\gamma\psi$,” (1980), appendix D. PhD thesis, Univ. Stanford. [67] ATLAS Collaboration, ATLAS-CONF-2014-031, <http://cds.cern.ch/record/1706244>. [68] P. D. Group, K. A. Olive, et al., *Chin. Phys. C* 38, 090001 (2014). [69] A. Papoulis, *Probability, random variables, and stochastic processes*, 4th ed. ed. (McGraw-Hill, Boston, 2002). [70] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Eur. Phys. J. C* 71, 1554 (2011), arXiv:1007.1727 [physics.data-an]. [71] E. Gross and O. Vitells, *Eur. Phys. J. C* 70, 525 (2010), arXiv:1005.1891 [physics.data-an]. [72] A. L. Read, *J. Phys. G* 28, 2693 (2002).

TABLES AND FIGURES

Table I: Simulated signal and background samples. The background simulation is used to understand the background composition and $\mu\mu$ invariant mass distribution. The term “tune” refers to the choice of parameters used for the underlying-event generation. Limits are set on the signal cross section, so the

entry here is listed as ‘...’ .

Process	Generator	Cross section	Tune	PDF set
$H \rightarrow aa$	Pythia8 [?]	...	AU2	CTEQ6L1
$Z/\gamma^* + \text{jets}$	Powheg [?, ?] + Pythia8	NNLO+NNLL [?, ?, ?, ?, ?, ?]	CT10	$\text{\cite{atlas_tune}}$
$h \rightarrow ZZ$	Alpgen [?]+Pythia6 [?]	NNLO+NNLL [?]	Perugia2010	CTEQ6L1
$t\bar{t}$	Pythia8	NNLO+NNLL [?]	Perugia2010	CTEQ6L1
Z/γ^*	Powheg+Pythia8	NNLO+NNLL [?]	Perugia2010	CTEQ6L1
$t\bar{t}$	Powheg+Pythia8	NNLO+NNLL [?]	Perugia2010	CTEQ6L1

Table II: Relative efficiency for each selection step for a signal simulated with $m_a = 5$ GeV and $m_H = m_h = 125$ GeV. The first entry is a generator-level selection relative to the number of generated $h \rightarrow aa \rightarrow \mu\mu\tau\tau$ events. The generator-level selection requires at least two muons with p_T above 15 GeV and $|\eta| < 2.5$. Generator-level events are further required to have a pair of such muons that have opposite charge and have $p_T(\mu\mu) > 30$ GeV. Only 0.5% of generated events have a $\mu\mu\tau\tau$ final state due to the small rate of $a \rightarrow \mu\mu$. As discussed in Section V, a small number of events with four τ -leptons are also accepted, but the rate is too small to impact the efficiency reported in this table. The top section of the table shows the efficiency for selecting $a \rightarrow \mu\mu$ candidates while the bottom part is for the selection of events in the two signal regions.

Selection	Relative efficiency (%)	SR μ (%)	SRe (%)
Generator-level	82.7 ± 0.0	-	-
Pass trigger	67.6 ± 0.3	95.5 ± 0.4	93.7 ± 0.7
Two selected muons	77.8 ± 0.3	91.4 ± 0.5	82.8 ± 1.1
Opposite charge (μ, μ)	100.0 ± 0.0	91.2 ± 0.9	88.1 ± 1.1
$p_T(\mu\mu) > 40$ GeV	98.1 ± 0.1	75.5 ± 0.9	84.6 ± 1.3
$2.8 \text{ GeV} < m_{\mu\mu} < 70 \text{ GeV}$	100.0 ± 0.0	-	-
Third lepton	18.2 ± 0.3	-	-
$\Delta\phi(\mu\mu, \ell) > 1$	95.5 ± 0.4	-	-

Selection	Relative efficiency (%)	SR μ (%)	SRe (%)
1,2 or 3 nearby tracks	91.4 ± 0.5	-	-
Opposite charge (ℓ , lead-track)	91.2 ± 0.9	-	-
Lepton isolation	75.5 ± 0.9	-	-

Table III: Measured values and uncertainties for the SM background model' s parameters, which are constrained by a simultaneous fit to the control regions. The errors on the fit parameters are statistical only.

Parameter	Value
a_μ	$(99.86 \pm 0.01)\%$
a_σ	$(1.68 \pm 0.02)\%$
α_{CB}	1.49 ± 0.03
$\sigma_{t\bar{t}}$	(60.7 ± 3.8) GeV
f_Z	$(23.4 \pm 0.5)\%$
$f_{\psi'}$	$(6.3 \pm 0.3)\%$
$f_{\Upsilon 2S}$	$(46.8 \pm 1.4)\%$
$f_{\Upsilon 3S}$	$(49.9 \pm 0.6)\%$
n_{γ^*}	(-31 ± 3) TeV $^{-1}$
α_{γ^*}	-0.75 ± 0.02

Table IV: Measured values and uncertainties of region-dependent parameters. The $m_{\mu\mu}$ distribution is fit between 2.8 GeV and 70 GeV for all regions, except for CRb, which has a lower bound at 15 GeV. There is no contribution to the total background from the ψ or Υ resonances.

Parameter	CRj	CRb	VR μ	VRe	SR μ	SRe
f_{Res}	14.7 ± 0.1	18.8 ± 2.3	12.2 ± 2.3	15.2 ± 1.6	11.8 ± 1.6	-
$f_{t\bar{t}}$	6.1 ± 0.9	87.2 ± 5.1	28.2 ± 3.2	34.2 ± 3.6	20.4 ± 4.1	23.5 ± 5.0
f_Υ	32.6 ± 0.3	35.8 ± 6.0	36.3 ± 9.2	25.8 ± 4.9	24.5 ± 6.6	-

Figure 1: Simulated dimuon invariant mass ($m_{\mu\mu}$) distribution and the result of the simultaneous fit projected into SR μ for one benchmark mass point with $m_a = 5$ GeV and $m_H = m_h = 125$ GeV. The best fit of the $\mu\mu$ resonance model

to the $h \rightarrow aa$ signal simulation in SR μ is shown in blue with its uncertainty as a yellow band. Also shown are the forced $\pm 1\sigma$ systematic uncertainty variations, defined in Sec. VII, in α_{CB} and $f_{\tau\tau}$ (dashed magenta). The top plot shows the simulation and fits on a linear scale, the middle on a logarithmic scale and the % residual of each fit is shown at the bottom. The errors on the signal simulation are statistical only.

Figure 2: Observed $m_{\mu\mu}$ distribution in CRj (top) and CRb (bottom) and the SM background model after a simultaneous fit. The Z/γ^* component of the fit is the combination of the Z boson resonance and the γ^* continuum models. The % residual of the fit is shown below each plot. Simulated SM backgrounds are shown in the stack, with the Z/γ^* sample only valid above $m_{\mu\mu} > 10$ GeV. The two insets show magnified versions of the J/ψ and Υ resonances.

Figure 3: Observed $m_{\mu\mu}$ distribution in VR μ (top) and VRe (bottom) and the background-only fit. The Z/γ^* component of the fit is the combination of the Z boson resonance and the γ^* continuum models. The % residuals are shown below each plot. Bins below 4 GeV are 200 MeV wide, between 4 GeV and 15 GeV they are 500 MeV wide, and above 15 GeV they are 2 GeV wide. Simulated SM backgrounds are shown in the stack, with the Z/γ^* sample only valid above $m_{\mu\mu} > 10$ GeV.

Figure 4: Observed $m_{\mu\mu}$ distribution in SR μ (top) and SRe (bottom) and the background-only fit. The Z/γ^* component of the fit is the combination of the Z boson resonance and the γ^* continuum models. The % residuals are shown below each plot. Bins below 4 GeV are 200 MeV wide, between 4 GeV and 15 GeV they are 500 MeV wide, and above 15 GeV they are 2 GeV wide. The expected distribution from a signal with $\text{BR}(h \rightarrow aa)=10\%$ is shown for three different m_a hypotheses (5 GeV, 10 GeV, and 20 GeV). Simulated SM backgrounds are shown in the stack, with the Z/γ^* sample only valid above $m_{\mu\mu} > 10$ GeV.

Figure 5: Observed p -value as a function of $m_{\mu\mu}$, with downward fluctuations of the data represented by a p -value of 0.5. The p -values are evaluated in 50 MeV intervals below 15 GeV, then 100 MeV intervals up to 30 GeV, and 200 MeV intervals up to $m_{\mu\mu} = 50$ GeV. The p -values shown have not been corrected for the look-elsewhere effect.

Figure 6: Observed (solid red) and expected (dashed black) limits with the expected $\pm 1\sigma$ and $\pm 2\sigma$ bands shown in green and yellow respectively. The top figure shows the expected and observed limits on the rate $(\sigma(gg \rightarrow h) \times \text{BR}(h \rightarrow aa))$ relative to the SM Higgs boson gluon-gluon fusion production cross section (σ_{SM}) as a function of m_a with m_H set to 125 GeV. The limits are evaluated in 50 MeV intervals below 15 GeV, then 100 MeV intervals up to 30 GeV, and 200 MeV intervals up to $m_a = 50$ GeV. Shown in the bottom figure is the total rate $(\sigma(gg \rightarrow H) \times \text{BR}(H \rightarrow aa))$ as a function of m_H with m_a set to 5 GeV, evaluated at 50 GeV intervals from $m_H = 100$ GeV to 500 GeV and at $m_H = m_h = 125$ GeV. The width of the black band in the bottom figure indicates the theoretical uncertainty on the SM $gg \rightarrow H$ cross section [?]. In

both figures, the observed and expected limits have been scaled by an $\mathcal{O}(1)$ parameter, $\text{BR}(a \rightarrow \tau\tau)^2$, to account for the branching ratios assumed in this analysis and facilitate reinterpretation of the results.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv – Machine translation. Verify with original.