Combined search for the Standard Model Higgs boson in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector


Published in: Physical Review D. Particles, Fields, Gravitation, and Cosmology

DOI: 10.1103/PhysRevD.86.032003

Link to publication

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).
I. INTRODUCTION

Probing the mechanism for electroweak symmetry breaking (EWSB) is one of the prime objectives of the Large Hadron Collider (LHC). In the Standard Model (SM) [1–3], the electroweak interaction is described by a local gauge field theory with an $SU(2)_L \otimes U(1)_Y$ symmetry, and EWSB is achieved via the Higgs mechanism with a single $SU(2)_L$ doublet of complex scalar fields [4–9]. After EWSB the electroweak sector has massive $W$ and $Z$ bosons, a massless photon, and a massive $CP$-even, scalar boson, referred to as the Higgs boson. Fermion masses are generated from Yukawa interactions with couplings proportional to the masses of fermions. The mass of the Higgs boson, $m_H$, is a free parameter in the SM. However, for a given $m_H$ hypothesis the cross sections of the various Higgs boson production processes and the branching fractions of the decay modes can be predicted, allowing a combined search with data from several search channels.

Combined searches at the CERN LEP $e^+e^-$ collider excluded the production of a SM Higgs boson with mass below 114.4 GeV at 95% confidence level (CL) [10]. The combined searches at the Fermilab Tevatron $p \bar{p}$ collider excluded the production of a SM Higgs boson with a mass between 147 GeV and 179 GeV, and between 100 GeV and 106 GeV at 95% CL [11]. Precision electroweak measurements are sensitive to $m_H$ via radiative corrections and indirectly constrain the SM Higgs boson mass to be less than 158 GeV [12] at 95% CL.

In 2011, the LHC delivered an integrated luminosity of 5.6 fb$^{-1}$ of proton-proton ($pp$) collisions at a center-of-mass energy of 7 TeV to the ATLAS detector [13]. Of the 4.9 fb$^{-1}$ collected, the integrated luminosity used in the individual Higgs search channels is between 4.6 fb$^{-1}$ and 4.6 fb$^{-1}$, depending on the data quality requirements specific to each channel.

This paper presents a combined search for the SM Higgs boson in the decay modes $H \rightarrow \gamma\gamma$, $H \rightarrow ZZ$, $H \rightarrow WW$, $H \rightarrow \tau^+\tau^-$, and $H \rightarrow b\bar{b}$, with subsequent decays of the $W$, $Z$, and $\tau$ leading to different final states. Some searches are designed to exploit the features of the production modes $pp \rightarrow H$ (gluon fusion), $pp \rightarrow q\bar{q}H$ (vector boson fusion), and $pp \rightarrow VH$ with $V = W$ or $Z$ (associated production with a gauge boson). In order to enhance the search sensitivity, the various decay modes are further subdivided into subchannels with different signal and background contributions and different sensitivities to systematic uncertainties. While the selection requirements for individual search channels are disjoint, each selection is, in general, populated by more than one combination of Higgs boson production and decay. For instance, Higgs boson production initiated by vector boson fusion (VBF) can contribute significantly to a search channel optimized for gluon fusion production.

The ATLAS Collaboration has previously published a similar but less extensive combined search for the Higgs boson [14] in data taken at the LHC in 2011. The CMS Collaboration has also performed a combined analysis of Higgs searches with data collected in 2011 and has obtained similar results [15]. In comparison to the analysis of Ref. [14], the $H \rightarrow \tau^+\tau^-$ and $H \rightarrow b\bar{b}$ channels have been added, the $H \rightarrow WW \rightarrow \ell^+\ell^-\nu\bar{\nu}$ analysis has been updated and extended to cover the mass range of 110–600 GeV, and the $H \rightarrow WW \rightarrow \ell\nu qq'$, $H \rightarrow ZZ \rightarrow \ell^+\ell^-\nu\bar{\nu}$, and $H \rightarrow ZZ \rightarrow \ell^+\ell^- q\bar{q}$ analyses have been...
updated to use the full 2011 data set. Both the \( H \rightarrow WW^{(*)} \rightarrow \ell^+\ell^−\nu\bar{\nu} \) and \( H \rightarrow WW \rightarrow \ell\nuq\bar{q}' \) analyses include a specific treatment of the 2-jet final state, which is targeted at the VBF production process.

The different channels entering the combination are summarized in Table I. After describing the general approach to statistical modeling in Sec. II, the individual channels and the specific systematic uncertainties are described in Secs. III and IV, respectively. The statistical procedure is described in Sec. V and the resulting exclusion limits and compatibility with the background-only hypothesis are presented in Secs. VI and VII, respectively.

II. STATISTICAL MODELING

In this combined analysis, a given search channel, indexed by \( c \), is defined by its associated event selection criteria, which may select events from various physical processes. In addition to the number of selected events, \( n \), each channel may make use of an invariant or transverse mass distribution of the Higgs boson candidates. The discriminating variable is denoted \( x \) and its probability density function (pdf) is written as \( f(x|\alpha) \), where \( \alpha \) represents both theoretical parameters such as \( m_H \) and nuisance parameters associated with various systematic effects. These distributions are normalized to unit probability. The predicted number of events satisfying the selection requirements is parametrized as \( \nu(\alpha) \). For a channel with \( n \) selected events, the data consist of the values of the discriminating variables for each event \( D = \{x_1, \ldots, x_n\} \). The probability model for these types of data is referred to as an unbinned extended likelihood or marked Poisson model \( f \), given by

\[
f(D|\alpha) = \text{Pois}(n|\nu(\alpha)) \prod_{e=1}^{n} f(x_e|\alpha), \quad (1)
\]

For each channel several signal and background scattering processes contribute to the total rate \( \nu \) and the overall pdf \( f(x|\alpha) \). Here, the term process is used for any set of scattering processes that can be added incoherently. The total rate is the sum of the individual rates

\[
\nu(\alpha) = \sum_{k \in \text{processes}} \nu_k(\alpha) \quad (2)
\]

and the total pdf is the weighted sum

\[
f(x|\alpha) = \frac{1}{\nu(\alpha)} \sum_{k \in \text{processes}} \nu_k(\alpha) f_k(x|\alpha). \quad (3)
\]

Using \( e \) as the index over the \( n_e \) events in the \( e \)th channel, \( x_{ce} \) is the value of the observable \( x \) for the \( e \)th event in channels 1 to \( e_{\text{max}} \). The total data are a collection of data from individual channels: \( D_{\text{com}} = \{D_1, \ldots, D_{e_{\text{max}}}\} \). The combined model can then be written as follows:

\[
f_{\text{com}}(D_{\text{com}}|\alpha) = \prod_{e=1}^{e_{\text{max}}} \text{Pois}(n_e|\nu_e(\alpha)) \prod_{e=1}^{n_e} f_e(x_{ce}|\alpha). \quad (4)
\]

A. Parametrization of the model

The parameter of interest is the overall signal strength factor \( \mu \), which acts as a scale factor to the total rate of signal events. This global factor is used for all pairings of production cross sections and branching ratios. The signal strength is defined such that \( \mu = 0 \) corresponds to the background-only model and \( \mu = 1 \) corresponds to the SM Higgs boson signal. It is convenient to separate the full list of parameters \( \alpha \) into the parameter of interest \( \mu \), the Higgs boson mass \( m_H \), and the nuisance parameters \( \theta \), i.e. \( \alpha = (\mu, m_H, \theta) \).

032003-2
Each channel in the combined model uses either the reconstructed transverse mass or the invariant mass of the Higgs candidate as a discriminating variable. Two approaches are adopted to model the signal pdfs at intermediate values of \( m_H \) where full simulation has not been performed. The first, used in the \( H \to \gamma \gamma \) channel, is based on a continuous parametrization of the signal as a function of \( m_H \) using an analytical expression for the pdf validated with simulated Monte Carlo (MC) samples. The second, used in channels where pdfs are modeled with histograms, is based on an interpolation procedure using the algorithm of Ref. [24].

### B. Auxiliary measurements

The nuisance parameters represent uncertain aspects of the model, such as the background normalization, reconstruction efficiencies, energy scale and resolution, luminosity, and theoretical predictions. These nuisance parameters are often estimated from auxiliary measurements, such as control regions, sidebands, or dedicated calibration measurements. A detailed account of these measurements is beyond the scope of this paper and is given in the references for the individual channels [16–23].

Each parameter \( \alpha_p \) with a dedicated auxiliary measurement \( f_{\text{aux}}(D_{\text{aux}}|\alpha_p, \alpha_{\text{other}}) \) provides a maximum likelihood estimate for \( \alpha_p, \sigma_p \), and a standard error \( \sigma_p \). Thus, the detailed probability model for an auxiliary measurement is approximated as

\[
f_{\text{aux}}(D_{\text{aux}}|\alpha_p, \alpha_{\text{other}}) \to f_p(a_p|\alpha_p, \sigma_p).
\] (5)

The \( f_p(a_p|\alpha_p, \sigma_p) \) are referred to as constraint terms.

The fully frequentist procedure applied for the present analysis includes randomizing the \( a_p \) when constructing the ensemble of possible experiment outcomes. In the hybrid frequentist-Bayesian procedures used at LEP and the Tevatron, the \( a_p \) are held constant and the nuisance parameters \( \alpha_p \) are randomized according to the prior probability density

\[
\pi(a_p|\alpha_p) \propto f(a_p|\alpha_p, \sigma_p)\eta(a_p),
\] (6)

where \( \eta(a_p) \) is an original prior, usually taken to be constant.

The set of nuisance parameters constrained by auxiliary measurements is denoted \( S \) and the set of estimates of those parameters, also referred to as global observables, which augments \( D_{\text{com}} \), is denoted \( G = \{a_p\} \) with \( p \in S \). Including the constraint terms explicitly, the model can be rewritten as

\[
f_{\text{tot}}(D_{\text{com}}, G|\alpha) = \prod_{\alpha=1}^{\text{max}} \left[ \text{Pois}(n_{\alpha}|\nu_\alpha(\alpha)) \right] \prod_{e=1}^{n} f_e(x_{ce}|\alpha) \cdot \prod_{p \in S} f_p(a_p|\alpha_p, \sigma_p).
\] (7)

The use of a Gaussian constraint term \( f_p(a_p|\alpha_p, \sigma_p) = \text{Gauss}(a_p|\alpha_p, \sigma_p) \) is problematic if the parameter is intrinsically non-negative, as is the case for event yields and energy scale uncertainties. This is particularly important when the relative uncertainty is large. An alternative constraint term defined only for positive parameter values is the log-normal distribution, which is given by

\[
f_p(a_p|\alpha_p) = \frac{1}{\sqrt{2\pi} \ln \kappa} \frac{1}{a_p} \exp \left[ -\frac{(\ln(a_p/\alpha_p))^2}{2(\ln \kappa)^2} \right].
\] (8)

The conventional choice \( \kappa = 1 + \sigma_{\text{rel}} \) is made, where \( \sigma_{\text{rel}} \) is the relative uncertainty \( \sigma_p/a_p \) from the observed auxiliary measurement [25].

Using the log-normal distribution for \( a_p \) is equivalent to having a Gaussian constraint for the transformed parameter \( a_p = \ln a_p \) and \( a_p^\prime = \ln \alpha_p \).

For channels that use histograms based on simulated MC samples, the parametric pdf \( f(x|\alpha) \) is formed by interpolating between histogram variations evaluated at \( \alpha_p = a_p \pm \sigma_p \). Since the variations need not be symmetric, the function is treated in a piecewise way using a sixth-order polynomial to interpolate in the range \( \alpha_p \in [a_p - \sigma_p, a_p + \sigma_p] \) with coefficients chosen to match the first and second derivatives [26]. Henceforth, the prime will be suppressed and \( \alpha_p \) will refer to the transformed nuisance parameter.

Not all systematic uncertainties have an associated auxiliary measurement. For example, uncertainties associated with the choice of renormalization and factorization scales and missing higher-order corrections in a theoretical calculation are not statistical in nature. In these cases, the frequentist form of the constraint term is derived assuming, by convention, a log-normal prior probability density on these parameters and inverting Eq. (6).

### III. INDIVIDUAL SEARCH CHANNELS

All the channels combined to search for the SM Higgs boson use the complete 2011 data set passing the relevant quality requirements. The Higgs boson decays considered are \( H \to \gamma \gamma, H \to WW^{(*)}, H \to ZZ^{(*)}, H \to \tau^+ \tau^-, \) and \( H \to b \bar{b} \). In modes with a \( W \) or \( Z \) boson, an electron or muon is required for triggering. In the \( H \to \tau^+ \tau^- \) channel, almost all combinations of subsequent \( \tau \) decays are considered. The results in the \( \gamma \gamma \) and \( \ell^+ \ell^- \ell^+ \ell^- \) modes are the same as in the previously published combination [14], but all other channels have been updated. A summary of the individual channels contributing to this combination is given in Table I.

The invariant and transverse mass distributions for the individual channels are shown in Figs. 1 and 2, with several subchannels merged.

1. \( H \to \gamma \gamma \): This analysis is unchanged with respect to the previous combined search [14,16] and is carried out for \( m_H \) hypotheses between 110 GeV and
FIG. 1 (color online). Invariant or transverse mass distributions for the selected candidate events, the total background and the signal expected in the following channels: (a) $H \rightarrow \gamma\gamma$, (b) $H \rightarrow ZZ^{(*)} \rightarrow \ell^+\ell^-\ell'^+\ell'^-$ in the entire mass range, (c) $H \rightarrow ZZ^{(*)} \rightarrow \ell^+\ell^-\ell'^+\ell'^-$ in the low mass range, (d) $H \rightarrow ZZ \rightarrow \ell^+\ell^-\nu\bar{\nu}$, (e) $b$-tagged selection and (f) untagged selection for $H \rightarrow ZZ \rightarrow \ell^+\ell^-q\bar{q}$, (g) $H \rightarrow WW^{(*)} \rightarrow \ell^+\nu\ell'^-\bar{\nu} + 0$-jet, (h) $H \rightarrow WW^{(*)} \rightarrow \ell^+\nu\ell'^-\bar{\nu} + 1$-jet, (i) $H \rightarrow WW^{(*)} \rightarrow \ell^+\nu\ell'^-\bar{\nu} + 2$-jet, (j) $H \rightarrow WW \rightarrow \ell\nu\ell'\nu + 0$-jet, (k) $H \rightarrow WW \rightarrow \ell\nu\ell'\nu + 1$-jet, and (l) $H \rightarrow WW \rightarrow \ell\nu\ell'\nu + 2$-jet. The $H \rightarrow WW^{(*)} \rightarrow \ell^+\nu\ell'^-\bar{\nu} + 2$-jet distribution is shown before the final selection requirements are applied.

Events are separated into nine independent categories of varying sensitivity. The categorization is based on the pseudorapidity of each photon, whether it was reconstructed as a converted or unconverted photon, and the momentum component of the diphoton system transverse to the diphoton thrust axis ($p_T^\gamma$). The mass resolution is approximately 1.7% for $m_H \sim 120$ GeV.
(2) $H \rightarrow ZZ^{(*)}$: In the $ZZ^{(*)}$ decay mode at least one $Z$ is required to decay to charged leptons, while the other decays to either leptons, neutrinos, or jets.

(i) $H \rightarrow ZZ^{(*)} \rightarrow \ell^+\ell^-\ell^+\ell^-$: This analysis, described in Ref. [17], is performed for $m_H$ hypotheses in the 110 GeV to 600 GeV mass range and is unchanged with respect to the previous combined search [14]. The main irreducible $ZZ^{(*)}$ background is estimated using a combination of Monte Carlo simulation and the observed data. The reducible $Z +$ jets background, which mostly impacts the low four-lepton invariant mass region, is estimated from control regions in the data. The top-quark $(t\bar{t})$ background normalization is validated using a dedicated control sample. Four categories of events are defined by the lepton flavor combinations and the four-lepton invariant mass is used as a discriminating variable. The mass resolution is approximately 1.5% in the four-muon channel and 2% in the four-electron channel for $m_H \sim 120$ GeV.

(ii) $H \rightarrow ZZ \rightarrow \ell^+\ell^-\nu\bar{\nu}$: This analysis [18] is split into two regimes according to the level of pileup, i.e. the average number of $pp$ collisions per bunch crossing. The first 2.3 fb$^{-1}$ of data had an average of about six pileup collisions per event and the subsequent 2.4 fb$^{-1}$ had an average of about 12. The search is performed for $m_H$ hypotheses ranging from 200 GeV to 600 GeV. The analysis is further refined in Ref. [19] where the $m_H$ dependence of the $Z +$ jets background is rescaled by a factor of 5 or 10 for illustration purposes.
categorized by the flavor of the leptons from the $Z$ decay. The selection is optimized separately for Higgs boson masses above and below 280 GeV. The $\ell^+\ell^-$ invariant mass is required to be within 15 GeV of the $Z$ boson mass. The inverted requirement is applied to same-flavor leptons in the $H \rightarrow WW(\gamma) \rightarrow \ell^+ \nu \ell^-\bar{\nu}$ channel to avoid overlap in the selection. The transverse mass ($m_T$), computed from the dilepton transverse momentum and the missing transverse momentum, is used as a discriminating variable.

(iii) $H \rightarrow ZZ \rightarrow \ell^+\ell^- q\bar{q}$: This search is performed for $m_H$ hypotheses ranging from 200 GeV to 600 GeV and is separated into search regions above and below $m_H = 300$ GeV, for which the event selections are independently optimized. The dominant background arises from $Z + \text{jets}$ production, which is estimated using sidebands of the dijet invariant mass distribution in data. To profit from the relatively large rate of $b$-jets from $Z$ boson decays present in the signal compared to the rate of $b$-jets found in the $Z + \text{jets}$ background, the analysis is divided into two categories. The first category contains events in which the two jets are $b$-tagged and the second uses events with less than two $b$-tags. The analysis [19] takes advantage of a highly efficient $b$-tagging algorithm [27] and the sideband to constrain the background yield. Using the $Z$ boson mass constraint improves the mass resolution of the $\ell^+\ell^- q\bar{q}$ system by more than a factor of 2. The invariant mass of the $\ell^+\ell^- q\bar{q}$ system is used as a discriminating variable.

(3) $H \rightarrow WW(\gamma)$: Two sets of channels are devoted to the decay of the Higgs boson into a pair of $W$ bosons, namely, the $\ell^+\nu\ell^-\bar{\nu}$ and $\ell\nu q\bar{q}'$ channels.

(i) $H \rightarrow WW(\gamma) \rightarrow \ell^+\nu\ell^-\bar{\nu}$: The updated analysis [20] is performed for $m_H$ values from 110 GeV up to 600 GeV. Events with two leptons are classified by the number of associated jets (0, 1 or 2), where the 2-jet category has selection criteria designed to enhance sensitivity to the VBF production process. The events are further divided by the flavors of the charged leptons, $e$, $\mu$, and $\mu$, where the mixed mode ($e\mu$) has a much smaller background from the Drell-Yan process. As in the case of $H \rightarrow ZZ \rightarrow \ell^+\ell^-\nu\bar{\nu}$, the samples are split according to the pileup conditions and analyzed separately. Each subchannel uses the $WW$ transverse mass distribution, except for the 2-jet category, which does not use a discriminating variable.

(ii) $H \rightarrow WW \rightarrow \ell\nu q\bar{q}'$: This analysis is performed for $m_H$ hypotheses ranging from 300 GeV to 600 GeV. A leptonically decaying $W$ boson is tagged with an isolated lepton and missing transverse momentum ($E_T^{\text{miss}}$). Additionally, two jets with an invariant mass compatible with a second $W$ boson [21] are required. The $W$ boson mass constraint allows the reconstruction of the Higgs boson candidate mass on an event-by-event basis by using a quadratic equation to solve for the component of the neutrino momentum along the beam axis. Events where this equation has imaginary solutions are discarded in order to reduce tails in the mass distribution. The analysis searches for a peak in the reconstructed $\ell\nu q\bar{q}'$ mass distribution. The background is modeled with a smooth function. The analysis is further divided by lepton flavor and by the number of additional jets (0, 1 or 2), where the 2-jet channel is optimized for the VBF production process.

(4) $H \rightarrow \tau^+\tau^-$: The analyses [22] are categorized by the decay modes of the two $\tau$ leptons, for $m_H$ hypotheses ranging from 110 GeV to 150 GeV (the leptonically decaying $\tau$ leptons are denoted $\tau_{\text{lep}}$ and the hadronically decaying $\tau$ leptons are denoted $\tau_{\text{had}}$). Most of these subchannels are triggered using leptons, except for the fully hadronic channel $H \rightarrow \tau_{\text{had}}\tau_{\text{had}}$, which is triggered with specific double hadronic $\tau$ decay selections. All the searches using $\tau$ decay modes have a significant background from $Z \rightarrow \tau^+\tau^-$ decays, which are modeled using an embedding technique where $Z \rightarrow \mu^+\mu^-$ candidates selected in the data have the muons replaced by simulated $\tau$ decays [22]. These embedded events are used to describe this background process.

(i) $H \rightarrow \tau_{\text{lep}}\tau_{\text{lep}}$: In this channel events are separately analyzed in four disjoint categories based on the number of reconstructed jets in the event [22]. There are two categories specifically aimed towards the gluon fusion production process, with or without a jet, one for the VBF production process and one for the Higgs boson production in association with a hadronically decaying vector boson. Each jet category requires at least one jet with $p_T$ above 40 GeV. The collinear approximation [28] is used to reconstruct the $\tau\tau$ invariant mass, which is used as the discriminating variable. All three combinations of $e$ and $\mu$ are used, except in the 0-jet category, which uses only the $e\mu$ candidate events where the effective mass is used as a discriminating variable.

(ii) $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$: There are seven separate categories in this subchannel. The selection of VBF-like events requires two jets with oppositely signed pseudorapidities $\eta$, $|\Delta\eta_{jj}| > 3.0$ and a dijet invariant mass larger than 300 GeV, in which events with electrons and muons are combined due to the limited number of candidates. In the other subchannels, electron and muon final states are considered separately. The remaining candidate events are categorized according to the number of jets with transverse momenta in
excess of 25 GeV, the 0-jet category being further subdivided based on whether the $E_T^{\text{miss}}$ exceeds 20 GeV or not. The missing mass calculator (MMC) technique [29] is used to estimate the $\tau\tau$ invariant mass, which is used as a discriminating variable.

(iii) $H + \text{jet} \rightarrow \tau_{\text{had}}\tau_{\text{had}} + \text{jet}$: Events are triggered using a selection of two hadronically decaying $\tau$ leptons with transverse energy thresholds varying according to the running conditions [22]. Two oppositely charged hadronically decaying $\tau$ candidates are required along with one jet with transverse momentum larger than 40 GeV, $E_T^{\text{miss}} > 20$ GeV, and a reconstructed invariant mass of the two $\tau$ leptons and the jet greater than 225 GeV. In addition to the Z background there is a significant multijet background which is estimated using data-driven methods. The $\tau\tau$ invariant mass is estimated via the collinear approximation and is used as a discriminating variable after further selections on the momentum fractions carried away by visible $\tau$ decay products.

(5) $H \rightarrow b\bar{b}$: The $ZH \rightarrow \ell^+\ell^-b\bar{b}$, $ZH \rightarrow \nu\bar{\nu}b\bar{b}$, and $WH \rightarrow \ell v b\bar{b}$ analyses [23] are performed for $m_H$ ranging from 110 GeV to 130 GeV. All three analyses require two $b$-tagged jets (one with $p_T > 45$ GeV and the other with $p_T > 25$ GeV) and the invariant mass of the two $b$-jets, $m_{bb}$, is used as a discriminating variable. The $ZH \rightarrow \ell^+\ell^-b\bar{b}$ analysis requires a dilepton invariant mass in the range $83 \text{GeV} < m_{\ell\ell} < 99 \text{GeV}$ and $E_T^{\text{miss}} < 50$ GeV to suppress the $t\bar{t}$ background. The $WH \rightarrow \ell v b\bar{b}$ analysis requires $E_T^{\text{miss}} > 25$ GeV, the transverse mass of the lepton-$E_T^{\text{miss}}$ system to be in excess of 40 GeV, and no additional leptons with $p_T > 20$ GeV. The $ZH \rightarrow \nu\bar{\nu}b\bar{b}$ analysis requires $E_T^{\text{miss}} > 120$ GeV, as well as $p_T^{\text{miss}} > 30$ GeV, where $p_T^{\text{miss}}$ is the missing transverse momentum determined from the tracks associated with the primary vertex. To increase the sensitivity of the search, the $m_{bb}$ distribution is examined in subchannels with different signal-to-background ratios. In the searches with one or two charged leptons, the division is made according to four bins in transverse momentum $p_T^V$ of the reconstructed vector boson V: $p_T^V < 50 \text{GeV}$, $50 \text{GeV} \leq p_T^V < 100 \text{GeV}$, $100 \text{GeV} \leq p_T^V < 200 \text{GeV}$, and $p_T^V \geq 200 \text{GeV}$. In the $ZH \rightarrow \nu\bar{\nu}b\bar{b}$ search the $E_T^{\text{miss}}$ is used to define three subchannels corresponding to $120 \text{GeV} < E_T^{\text{miss}} < 160 \text{GeV}$, $160 \text{GeV} \leq E_T^{\text{miss}} < 200 \text{GeV}$, and $E_T^{\text{miss}} \geq 200 \text{GeV}$. No categorization is made based on lepton flavor.

IV. SYSTEMATIC UNCERTAINTIES

The sources of systematic uncertainties and their effects on the signal and background rates $\nu_4 (x; \alpha)$ and discriminating variable distributions $f_4 (x; \alpha)$ are described in detail for each channel in Refs. [16–23]. The sources of systematic uncertainty are decomposed into uncorrelated components, such that the constraint terms factorize as in Eq. (7). The main focus of the combination of channels is the correlated effect of given sources of uncertainties across channels. Typically, the correlated effects arise from the ingredients common to several channels, for example the simulation, the lepton and photon identification, and the integrated luminosity. The sources of systematic uncertainty affecting the signal model are frequently different from those affecting the backgrounds, which are often estimated from control regions in the data. The dominant uncertainties giving rise to correlated effects are those associated with theoretical predictions for the signal production cross sections and decay branching fractions, as well as those related to detector response affecting the reconstruction of electrons, photons, muons, jets, $E_T^{\text{miss}}$, and $b$-tagging. The log-normal constraint terms are used for uncertainties in the signal and background normalizations, while Gaussian constraints are used for uncertainties affecting the shapes of the pdfs.

A. Theoretical uncertainties affecting the signal

The Higgs boson production cross sections are computed up to next-to-next-to-leading order (NNLO) [30–35] in QCD for the gluon fusion ($g g \rightarrow H$) process, including soft-gluon resummation up to next-to-next-to-leading log (NNLL) [36] and next-to-leading-order (NLO) electroweak corrections [37,38]. These predictions are compiled in Refs. [39–41]. The cross section for the VBF process is estimated at NLO [42–44] and approximate NNLO QCD [45]. The cross sections for the associated production processes ($q\bar{q} \rightarrow WH/ZH$) are computed at NLO [46,47], NNLO [48] QCD, and NLO electroweak [47]. The cross sections for the associated production with a $t\bar{t}$ pair ($q\bar{q}/gg \rightarrow t\bar{t}H$) are estimated at NLO [49–53]. The Higgs boson production cross sections and decay branching ratios [54–58], as well as their related uncertainties, are compiled in Ref. [59]. The QCD scale uncertainties for $m_H = 120$ GeV amount to $^{+12}_{-8}$% for the $gg \rightarrow H$ process, $\pm 1\%$ for the $qq' \rightarrow qq'H$ and associated $WH/ZH$ processes, and $^{+3}_{-9}$% for the $q\bar{q}/gg \rightarrow t\bar{t}H$ process. The uncertainties related to the parton distribution functions (PDF) amount to $\pm 8\%$ for the predominantly gluon-initiated processes $gg \rightarrow H$ and $gg \rightarrow t\bar{t}H$, and $\pm 4\%$ for the predominantly quark-initiated processes $qq' \rightarrow qq'H$ and $WH/ZH$ [60]. The theoretical uncertainty associated with the exclusive Higgs boson production process with one additional jet in the $H \rightarrow WW^{(*)} \rightarrow \ell^+\ell^-\nu\bar{\nu}$ channel amounts to $\pm 20\%$ and is treated according to the prescription of Refs. [25,61]. An additional theoretical uncertainty on the signal normalization, to account for effects related to off-shell Higgs boson production and interference with other SM processes, is assigned at high Higgs boson masses ($m_H > 300$ GeV) and estimated as $\pm 150\% \times (m_H/\text{TeV})^3$ [61–64].
B. Theoretical uncertainties affecting the background

In the $H \to \gamma \gamma$ and $H \to WW \to \ell \nu q\bar{q}'$ channels the backgrounds are estimated from a fit to the data. This removes almost all sensitivity to the corresponding theoretical uncertainties. In the case of the $H \to \gamma \gamma$ analysis an additional uncertainty is assigned to take into account possible inadequacies of the analytical background model chosen. Theoretical uncertainties enter in all other channels where theoretical calculations are used for background estimates. In particular, both signal and background processes are sensitive to the parton distribution functions, the underlying event simulation, and the parton shower model.

The $ZZ^{(*)}$ continuum process is the main background for the $H \to ZZ^{(*)} \to \ell^+ \ell^- \ell'^+ \ell'^-$ and $H \to ZZ \to \ell^+ \ell^- \nu \bar{\nu}$ analyses and is also part of the backgrounds in the $H \to ZZ \to \ell^+ \ell^- q\bar{q}$ channel. A NLO prediction [65] is used for the normalization. The QCD scale uncertainty has a $\pm 5\%$ effect on the expected $ZZ^{(*)}$ background, and the effects due to the PDF and $\alpha_S$ uncertainties are $\pm 4\%$ and $\pm 8\%$ for quark-initiated and gluon-initiated processes, respectively. An additional theoretical uncertainty of $\pm 10\%$ on the inclusive $ZZ^{(*)}$ cross section is conservatively included due to the missing higher-order QCD corrections for the gluon-initiated process. This theoretical uncertainty is treated as uncorrelated for the different channels due to the different acceptance in the $H \to ZZ^{(*)} \to \ell^+ \ell^- \ell'^+ \ell'^-$ and $H \to ZZ \to \ell^+ \ell^- \nu \bar{\nu}$ channels and because its contribution to the $H \to ZZ \to \ell^+ \ell^- q\bar{q}$ channel is small.

In most other channels the overall normalization of the main backgrounds is not estimated from theoretical predictions; however, simulations are used to model the pdfs $f(x|\alpha)$ or the scale factors used to extrapolate from the control regions to the signal regions. For example, in the $H \to WW^{(*)} \to \ell^+ \nu \ell'^- \bar{\nu}'$ channel the main backgrounds are continuum $WW^{(*)}$ and $t\bar{t}$ production. Their normalizations are estimated in control regions; however, the factors used to extrapolate to the signal region are estimated with the NLO simulation [66].

C. Experimental uncertainties

The uncertainty on the integrated luminosity is considered as being fully correlated among channels and amounts to $\pm 3.9\%$ [67,68].

The detector-related sources of systematic uncertainty can affect various aspects of the analyses: (a) the overall normalization of the signal or background, (b) the migration of events between categories, and (c) the shape of the discriminating variable distributions $f(x|\alpha)$. Similarly to the theoretical uncertainties, experimental uncertainties on the event yields (a) are treated using a log-normal $f_{\rho}(a_{\rho}|x_{\rho})$ constraint pdf. In cases (b) and (c) a Gaussian constraint is applied.

The experimental sources of systematic uncertainty are modeled using the classification detailed below. Their effect on the signal and background yields in each channel is reported separately in Table II. The various sources of systematic uncertainty have in some cases been grouped for a concise presentation (e.g. the jet energy scale and $b$-tagging efficiencies), while the full statistical model of the data provides a more detailed account of the various systematic effects including the effect on the pdfs $f(x|\alpha)$. The assumptions made in the treatment of systematics are outlined below.

(i) The uncertainties in the trigger and identification efficiencies are treated as fully correlated for electrons and photons. The energy scale and resolution for photons and electrons are treated as uncorrelated sources of uncertainty.

(ii) The uncertainties affecting muons are separated into those related to the inner detector (ID) and the muon spectrometer (MS) in order to provide a better description of the correlated effect among channels using different muon identification criteria and different ranges of muon transverse momenta.

(iii) The jet energy scale (JES) and jet energy resolution (JER) are sensitive to a number of uncertainty quantities, which depend on $p_T$, $\eta$, and flavor of the jet. Measurements of the JES and JER result in complicated correlations among these components. Building a complete model of the response to these correlated sources of uncertainty is intricate. Here, a simplified scheme is used in which independent JES and JER nuisance parameters are associated to channels with significantly different kinematic requirements and scattering processes with different kinematic distributions or flavor composition. This scheme includes a specific treatment for $b$-jets. The sensitivity of the results to various assumptions in the correlation between these sources of uncertainty has been found to be negligible. Furthermore, an additional component to the uncertainty in $E_{\text{miss}}$, which is uncorrelated with the JES uncertainty, is included.

(iv) While the $\tau$ energy scale uncertainty is expected to be partially correlated with the JES, here it is treated as an uncorrelated source of uncertainty. This choice is based on the largely degenerate effect due to the uncertainty associated with the embedding procedure, in which the simulated detector response to hadronic $\tau$ decays is merged with a sample of $Z \to \mu^+ \mu^-$ data events. Furthermore, the uncertainty of this embedding procedure is treated separately for signal and background processes, which is a conservative approach given that the $Z \to \tau^+ \tau^-$ sideband effectively constrains this nuisance parameter.

(v) The $b$-tagging systematic uncertainty is decomposed into five fundamental sources in the $H \to bb$ channels, while a simplified model with a single source is...
TABLE II. A summary of the main correlated experimental systematic uncertainties. The uncorrelated systematic uncertainties are summarized in a single combined number. The uncertainties indicate the ±1σ relative variation in the signal and background yields in (%). The signal corresponds to a Higgs boson mass hypothesis of 125 GeV except for $H \to ZZ \to \ell^+\ell^-q\bar{q}$, $H \to ZZ \to \ell^+\ell^-\nu\bar{\nu}$, and $H \to WW \to \ell^\nu\ell^\nu q\bar{q}$, which are quoted at 350 GeV.

<table>
<thead>
<tr>
<th>$H \to \gamma\gamma$</th>
<th>$H \to ZZ^{(*)}$</th>
<th>$H \to WW^{(*)}$</th>
<th>$H \to \tau^+\tau^-$</th>
<th>$H \to b\bar{b}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Relative uncertainty on signal yields</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Luminosity</td>
<td>±3.9</td>
<td>±3.9</td>
<td>±3.9</td>
<td>±3.9</td>
</tr>
<tr>
<td>$e/\gamma$ efficiency</td>
<td>+13.5</td>
<td>±3.2</td>
<td>±1.3</td>
<td>±1.5</td>
</tr>
<tr>
<td>$e/\gamma$ energy scale</td>
<td></td>
<td></td>
<td>±0.4</td>
<td></td>
</tr>
<tr>
<td>$e/\gamma$ resolution</td>
<td></td>
<td></td>
<td>±0.1</td>
<td></td>
</tr>
<tr>
<td>$\mu$ efficiency</td>
<td></td>
<td></td>
<td>±0.2</td>
<td></td>
</tr>
<tr>
<td>$\mu$ resolution (ID)</td>
<td></td>
<td></td>
<td>±0.1</td>
<td></td>
</tr>
<tr>
<td>$\mu$ resolution (MS)</td>
<td></td>
<td></td>
<td>±0.1</td>
<td></td>
</tr>
<tr>
<td>Jet/$E_T^{miss}$ energy scale</td>
<td></td>
<td>±2.5</td>
<td></td>
<td>+3.5</td>
</tr>
<tr>
<td>Jet energy resolution</td>
<td></td>
<td>±1.1</td>
<td></td>
<td>±4.2</td>
</tr>
<tr>
<td>$\tau$ $b$-tag efficiency</td>
<td></td>
<td></td>
<td>±0.9</td>
<td></td>
</tr>
<tr>
<td>$\tau$ efficiency</td>
<td></td>
<td></td>
<td>±4.2</td>
<td></td>
</tr>
<tr>
<td>Uncorrelated uncertainties</td>
<td>±0.2</td>
<td>±5.0</td>
<td>±12.0</td>
<td>±3.4</td>
</tr>
</tbody>
</table>

| **Relative uncertainty on background yields** |
| Luminosity | | | | | | | | | | |
| $e/\gamma$ efficiency | | | | | | | | | | |
| $e/\gamma$ energy scale | | | | | | | | | | |
| $e/\gamma$ resolution | | | | | | | | | | |
| $\mu$ efficiency | | | | | | | | | | |
| $\mu$ resolution (ID) | | | | | | | | | | |
| $\mu$ resolution (MS) | | | | | | | | | | |
| Jet/$E_T^{miss}$ energy scale | | | | | | | | | | |
| Jet energy resolution | | | | | | | | | | |
| $\tau$ $b$-tag efficiency | | | | | | | | | | |
| $\tau$ efficiency | | | | | | | | | | |
| Uncorrelated uncertainties | ±10.0 | ±4.9 | ±12.0 | ±10.2 | ±6.3 | ±10.3 | ±5.5 | ±2.8 | ±2.9 |

used in the $H \to ZZ \to \ell^+\ell^-q\bar{q}$ channel. The uncertainty in the $b$-veto is considered uncorrelated with the uncertainty in the $b$-tagging efficiency.

The effect of these systematic uncertainties depends on the final state, but is typically small compared to the theoretical uncertainty of the production cross section.

The electron and muon energy scales are directly constrained by $Z \to e^+e^-$ and $Z \to \mu^+\mu^-$ events; the impact of the resulting systematic uncertainty on the four-lepton invariant mass is of the order of ±0.5% for electrons and negligible for muons. The impact of the photon energy scale systematic uncertainty on the diphoton invariant mass is approximately ±0.6%.

D. Background measurement uncertainty

The estimates of background normalizations and model parameters from control regions or sidebands are the main remaining sources of uncertainty. Because of the differences in control regions these uncertainties are not correlated across channels.

In the case of the $H \to b\bar{b}$ channels the background normalizations are constrained both from sideband fits and from auxiliary measurements based on the MC prediction of the main background processes ($Z +jets$, $W +jets$ and $t\bar{t}$).

The uncorrelated sources of systematic uncertainties are summarized as a single combined number in Table II for each channel.

E. Summary of the combined model

To cover the search range efficiently between $m_H = 110$ GeV and $m_H = 600$ GeV, the signal and backgrounds are modeled and the combination performed in $m_H$ steps that reflect the interplay between the invariant mass resolution and the natural width of the Higgs boson. In the low
mass range, where the high mass-resolution $H \to \gamma\gamma$ and $H \to ZZ^{(*)} \to \ell^+\ell^-\ell^+\ell^-$ channels dominate, the signal is modeled in steps from 500 MeV to 2 GeV. For higher masses the combination is performed with step sizes ranging from 2 GeV to 20 GeV. The $m_H$ step sizes are given in Table III.

The combined model and statistical procedure are implemented within the ROOFIT, HISTFACTORY, and ROOSTATS software framework [26,69,70]. As shown in Table I, the number of channels included in the combination depends on the hypothesized value of $m_H$. The details for the number of channels, nuisance parameters, and constraint terms for various $m_H$ ranges are shown in Table IV. For $m_H = 125$ GeV there are 70 channels included in the combined statistical model and the associated data set is comprised of more than 22,000 unbinned events and 8000 bins. For $m_H = 350$ GeV there are 46 channels included in the combined statistical model and the associated data set is comprised only of binned distributions, for which there are more than 4000 bins. Because of the limited size of the MC samples, additional nuisance parameters and constraint terms are included in the model to account for the statistical uncertainty in the MC templates. The difference between the number of nuisance parameters and the number of constraints reported in Table IV corresponds to the number of nuisance parameters without external constraints, which are estimated in data control regions or sidebands.

### Table III

<table>
<thead>
<tr>
<th>$m_H$ (GeV)</th>
<th>Step size</th>
</tr>
</thead>
<tbody>
<tr>
<td>110–120</td>
<td>1 GeV</td>
</tr>
<tr>
<td>120–130</td>
<td>0.5 GeV</td>
</tr>
<tr>
<td>130–150</td>
<td>1 GeV</td>
</tr>
<tr>
<td>150–290</td>
<td>2 GeV</td>
</tr>
<tr>
<td>290–350</td>
<td>5 GeV</td>
</tr>
<tr>
<td>350–400</td>
<td>10 GeV</td>
</tr>
<tr>
<td>400–600</td>
<td>20 GeV</td>
</tr>
</tbody>
</table>

### Table IV

<table>
<thead>
<tr>
<th>$m_H$ (GeV)</th>
<th>Channels</th>
<th>Nuisance</th>
<th>Constraints</th>
<th>MC stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>110–130</td>
<td>70</td>
<td>287</td>
<td>159</td>
<td>180</td>
</tr>
<tr>
<td>131–150</td>
<td>67</td>
<td>210</td>
<td>119</td>
<td>140</td>
</tr>
<tr>
<td>152–198</td>
<td>46</td>
<td>83</td>
<td>80</td>
<td>78</td>
</tr>
<tr>
<td>200–278</td>
<td>52</td>
<td>199</td>
<td>119</td>
<td>120</td>
</tr>
<tr>
<td>280–295</td>
<td>52</td>
<td>208</td>
<td>118</td>
<td>119</td>
</tr>
<tr>
<td>300</td>
<td>58</td>
<td>281</td>
<td>120</td>
<td>121</td>
</tr>
<tr>
<td>305–400</td>
<td>46</td>
<td>269</td>
<td>111</td>
<td>112</td>
</tr>
<tr>
<td>420–480</td>
<td>46</td>
<td>238</td>
<td>111</td>
<td>112</td>
</tr>
<tr>
<td>500–600</td>
<td>46</td>
<td>201</td>
<td>111</td>
<td>112</td>
</tr>
</tbody>
</table>

---

### V. Statistical Procedures

The procedures for computing frequentist $p$-values used for quantifying the agreement of the data with the background-only hypothesis and for determining exclusion limits are based on the profile likelihood ratio test statistic.

For a given data set $D_{\text{com}}$ and values for the global observables $G$ there is an associated likelihood function of $\mu$ and $\theta$ derived from the combined model over all channels including all constraint terms in Eq. (7),

$$L(\mu, \theta; m_H, D_{\text{com}}, G) = f_{\text{tot}}(D_{\text{com}}, G | \mu, m_H, \theta).$$

(9)

The notation $L(\mu, \theta)$ leaves the dependence on the data implicit.

#### A. The test statistics and estimators of $\mu$ and $\theta$

The statistics used to test different values of the strength parameter $\mu$ are defined in terms of a likelihood function $L(\mu, \theta)$. The maximum likelihood estimates $\hat{\mu}$ and $\hat{\theta}$ are the values of the parameters that maximize the likelihood function $L(\mu, \theta)$. The conditional maximum likelihood estimate $\hat{\theta}(\mu)$ is the value of $\theta$ that maximizes the likelihood function with $\mu$ fixed. The tests are based on the profile likelihood ratio $\lambda(\mu)$, which reflects the level of compatibility between the data and $\mu$. It is defined as

$$\lambda(\mu) = \frac{L(\mu, \hat{\theta}(\mu))}{L(\hat{\mu}, \hat{\theta})}.$$  

(10)

Physically, the rate of signal events is non-negative; thus, $\mu \geq 0$. However, it is convenient to define the estimator $\tilde{\mu}$ as the value of $\mu$ that maximizes the likelihood, even if it is negative [as long as the pdf $f_\ell(x | \mu, \theta) > 0$ everywhere]. In particular, $\tilde{\mu} < 0$ indicates a deficit of events with respect to the background in the signal region. Following Ref. [71] a treatment equivalent to requiring $\mu \geq 0$ is to allow $\mu < 0$ and impose the constraint in the test statistic itself, i.e.

$$\lambda(\mu) = \begin{cases} L(\mu, \hat{\theta}(\mu)) & \tilde{\mu} \geq 0, \\ \frac{L(\mu, \hat{\theta}(\mu))}{L(0, \hat{\theta}(0))} & \tilde{\mu} < 0. \end{cases}$$  

(11)

To quantify the significance of an excess, the test statistic $\tilde{q}_0$ is used to test the background-only hypothesis $\mu = 0$ against the alternative hypothesis $\mu > 0$. It is defined as

$$\tilde{q}_0 = \begin{cases} -2 \ln \lambda(0) & \tilde{\mu} > 0, \\ +2 \ln \lambda(0) & \tilde{\mu} \leq 0. \end{cases}$$  

(12)

Instead of defining the test statistic to be identically zero for $\tilde{\mu} \leq 0$ as in Ref. [71], this sign change is introduced in order to probe $p$-values larger than 50%.

For setting an upper limit on $\mu$, the test statistic $\tilde{q}_\mu$ is used to test the hypothesis of signal events being produced at a rate $\mu$ against the alternative hypothesis of signal events being produced at a lesser rate $\mu' < \mu$:
Again, a sign change is introduced in order to probe \( p \)-values larger than 50%. The test statistic \(-2\ln(\lambda(\mu))\) is used to differentiate signal events being produced at a rate \( \mu \) from the alternative hypothesis of signal events being produced at a different rate \( \mu' \neq \mu \).

Tests of \( \mu \) are carried out with the Higgs mass \( m_H \) fixed to a particular value, and the entire procedure is repeated for values of \( m_H \) spaced in small steps.

**B. The distribution of the test statistic and \( p \)-values**

When calculating upper limits, a range of values of \( \mu \) is explored using the test statistic \( \tilde{q}_\mu \). The value of the test statistic for the observed data is denoted \( \tilde{q}_{\mu, \text{obs}} \). One can construct the distribution of \( \tilde{q}_\mu \) assuming a different value of the signal strength \( \mu' \), which is denoted

\[
 f(\tilde{q}_\mu | \mu', m_H, \theta).
\]

The distribution depends explicitly on \( m_H \) and \( \theta \). The \( p \)-value is given by the tail probability of this distribution, and thus the \( p \)-value will also depend on \( m_H \) and \( \theta \). The reason for choosing the test statistic based on the profile likelihood ratio is that, with sufficiently large numbers of events, the distribution of the profile likelihood ratio with \( \mu = \mu' \) is independent of the values of the nuisance parameters and, thus, also the associated \( p \)-values.

In practice, there is generally some residual dependence of the \( p \)-values on the value of \( \theta \). The values of the nuisance parameters that maximize the \( p \)-value are therefore sought. Following Refs. [25,72–75], the \( p \)-values for testing a particular value of \( \mu \) are based on the distribution constructed at \( \hat{\theta}(\mu, \text{obs}) \), the conditional maximum likelihood estimate with the observed data, as follows:

\[
 p_\mu = \int_{\tilde{q}_{\mu, \text{obs}}}^{\infty} f(\tilde{q}_\mu | \mu, m_H, \hat{\theta}(\mu, \text{obs}))d\tilde{q}_\mu.
\]

The ensemble includes randomizing both \( D \) and \( G \).

Here the distribution of \( \tilde{q}_\mu \) assumes that the data \( D \) as well as the global observables \( G \) are treated as measured quantities; i.e., they fluctuate upon repetition of the experiment according to the model \( f_{\text{(obs)}}(D_{\text{com}}, G | \mu, m_H, \theta) \).

Upper limits for the strength parameter \( \mu \) are calculated using the \( \text{CL}_s \) procedure [76]. To calculate this limit, the quantity \( \text{CL}_s(\mu) \) is defined as the ratio

\[
 \text{CL}_s(\mu) = \frac{p_\mu}{1 - p_b},
\]

where \( p_b \) is the \( p \)-value derived from the same test statistic under the background-only hypothesis,

\[
 p_b = 1 - \int_{\tilde{q}_{\mu, \text{obs}}}^{\infty} f(\tilde{q}_\mu | 0, m_H, \hat{\theta}(\mu = 0, \text{obs}))d\tilde{q}_\mu.
\]

The \( \text{CL}_s \) upper limit on \( \mu \) is denoted \( \mu_{\text{up}} \) and obtained by solving for \( \text{CL}_s(\mu_{\text{up}}) = 5\% \). A value of \( \mu \) is regarded as excluded at the 95% confidence level if \( \mu > \mu_{\text{up}} \).

The significance of an excess is based on the compatibility of the data with the background-only hypothesis. This compatibility is quantified by the following \( p \)-value:

\[
 p_0 = \int_{\tilde{q}_{0, \text{obs}}}^{\infty} f(\tilde{q}_0 | 0, m_H, \hat{\theta}(\mu = 0, \text{obs}))d\tilde{q}_0.
\]

Note that \( p_0 \) and \( p_b \) are both \( p \)-values of the background-only hypothesis, but the test statistic \( \tilde{q}_0 \) in Eq. (18) is optimized for discovery while the test statistic \( \tilde{q}_\mu \) in Eq. (17) is optimized for upper limits.

It is customary to convert the background-only \( p \)-value into an equivalent Gaussian significance \( Z \) (often written \( Z\sigma \)). The conversion is defined as

\[
 Z = \Phi^{-1}(1 - p_0),
\]

where \( \Phi^{-1} \) is the inverse of the cumulative distribution for a standard Gaussian.

**C. Experimental sensitivity and bands**

It is useful to quantify the experimental sensitivity by means of the significance one would expect to find if a given signal hypothesis were true. Similarly, the expected upper limit is the median upper limit one would expect to find if the background-only hypothesis were true. Although these are useful quantities, they are subject to a certain degree of ambiguity because the median values depend on the assumed values of all of the parameters of the model, including the nuisance parameters.

Here, the expected upper limit is defined as the median of the distribution \( f(\mu_{\text{up}} | 0, m_H, \hat{\theta}(\mu = 0, \text{obs})) \) and the expected significance is based on the median of the distribution \( f(p_0 | 1, m_H, \hat{\theta}(\mu = 1, \text{obs})) \). The expected limit and significance thus have a small residual dependence on the observed data through \( \hat{\theta}(\mu, \text{obs}) \).

These distributions are also used to define bands around the median upper limit. The standard presentation of upper limits includes the observed limit, the expected limit, and a \( \pm 1\sigma \) and a \( \pm 2\sigma \) band. More precisely, the edges of these bands, denoted \( \mu_{\text{up} \pm 1} \) and \( \mu_{\text{up} \pm 2} \), are defined by

\[
 \int_0^{\mu_{\text{up} \pm N}} f(\mu_{\text{up}} | 0, m_H, \hat{\theta}(\mu = 0, \text{obs}))d\mu_{\text{up}} = \Phi(\pm N).
\]

**D. Asymptotic formalism**

For large data samples, the asymptotic distributions \( f(\tilde{q}_\mu | \mu', m_H, \theta) \) and \( f(\tilde{q}_0 | 0, m_H, \theta) \) are known and described in Ref. [71]. These formulas require the variance of the maximum likelihood estimate of \( \mu \), given \( \mu' \) is the true value:
\[ \sigma_{\mu'} = \sqrt{\text{var}[\hat{\mu} | \mu']}. \] (21)

One result of Ref. [71] is that \( \sigma_{\mu'} \) can be estimated with an artificial data set referred to as the Asimov data set. This data set is defined as a binned data set, where the number of events in bin \( b \) is exactly the number of events expected in bin \( b \). The value of the test statistic evaluated on the Asimov data is denoted \( \hat{q}_{\mu, A} \), where the subscript \( A \) denotes that this is the Asimov data set associated with \( \alpha \). A convenient way to estimate the variance of \( \hat{\mu} \) is

\[ \sigma_{\mu'} = \frac{\mu - \mu'}{\sqrt{q_{\mu, A}}}. \] (22)

In the asymptotic limit, \( \hat{q}_\mu \) is parabolic; thus \( \sigma_{\mu'} \) is independent of \( \mu \). In Ref. [71], the bands around the expected limit are given by

\[ \mu_{\text{up} + N} = \sigma(\Phi^{-1}(1 - 0.05\Phi(N)) + N). \] (23)

An improved procedure is used here in order to capture the leading deviations of \( \hat{q}_\mu \) from a parabola, corresponding to departures in the distribution of \( \hat{\mu} \) from a Gaussian distribution centered at \( \mu' \) with variance \( \sigma_{\mu'}^2 \). The upper limit \( \mu_{\text{up}} \) for an Asimov data set constructed with \( \alpha_N = (\mu_N, m_H, \hat{\theta}(\mu_N, \text{obs})) \) (using the formulas in Ref. [71]) is found to be more accurate than Eq. (23) in reproducing the bands obtained with ensembles of pseudo-experiments, where \( \mu_N \) is the value of \( \mu \) corresponding to the edges of the \( \mu_{\text{up} + N} \) band. The value of \( \mu_N \) used for the +N\sigma band is the value of \( \mu \) that satisfies \( \sqrt{q_{\mu_N, A}} = N \). The choice of \( \hat{\theta}(\mu_N, \text{obs}) \) is more indicative of \( \hat{\theta} \) for the pseudoexperiments that have \( \mu_{\text{up}} \) near the corresponding band.

### E. Bayesian methods

A posterior distribution for the signal strength parameter \( \mu \) can be obtained, via Bayes’s theorem, with \( f_{\text{obs}}(D, G | \alpha) \) and a prior \( \eta(\alpha) \). The information about the nuisance parameters from auxiliary measurements is incorporated into \( f_{\text{obs}}(D, G | \alpha) \) through the constraint terms \( f_p(\alpha_p | \alpha_p) \). As in Eq. (6), the prior on \( \eta(\alpha_p) \) is taken as a uniform distribution. The upper limits are calculated for a given value of \( m_H \), so no prior on \( m_H \) is needed. The prior on the signal strength parameter \( \mu \) is taken to be uniform, as this choice leads to one-sided Bayesian credible intervals that correspond numerically to the CLs upper limits in the frequentist formalism for the simple Gaussian and Poisson cases, and have been observed to coincide in more complex situations. The one-sided, 95% Bayesian credible region is defined as

\[ \int_0^{\mu_{\text{up}}} f_{\text{obs}}(D_{\text{sim}}, G | \mu, m_H, \theta) \eta(\mu) \eta(\theta) d\mu d\theta = 0.95. \] (24)

The integration over nuisance parameters is carried out with Markov chain Monte Carlo using the Metropolis-Hastings algorithm in the RooStats package [26].

### F. Correction for the look-elsewhere effect

By scanning over \( m_H \) and repeatedly testing the background-only hypothesis, the procedure is subject to effects of multiple testing referred to as the look-elsewhere effect. In principle, the confidence intervals can be constructed in the \( \mu \sim m_H \) plane using the profile likelihood ratio \( \lambda(\mu, m_H) \). For \( \mu = 0 \) there is no signal present; thus, the model is independent of \( m_H \). This leads to a background-only distribution for \( -2 \ln \lambda(0, m_H) \) that departs from a chi-square distribution, thus complicating the calculation of a global \( p_0 \) when \( m_H \) is not specified. The global test statistic is the supremum of \( q_0(m_H) \) with respect to \( m_H \),

\[ q_0(\hat{m}_H) = \sup_{m_H} q_0(m_H). \] (25)

In the asymptotic regime and for very small \( p \)-values, a procedure [77], based on the result of Ref. [78], exists to estimate the tail probability for \( q_0(\hat{m}_H) \). The procedure requires an estimate for the average number of up-crossings of \( q_0(m_H) \) above some threshold. Due to the \( m_H \)-dependence in the model, changes in cuts, and the list of channels included, it is technically difficult to estimate this quantity with ensembles of pseudoexperiments. Instead, a simple alternative method is used in which the average number of up-crossings is estimated by counting the number of up-crossings with the observed data. When the trials factor is large, the number of up-crossings at low thresholds is also large and thus a satisfactory estimate of the average [25]. This procedure has been checked using a large number of pseudoexperiments in a simplified test case, where it provided a good estimate of the trials factor.

### VI. EXCLUSION LIMITS

The model discussed in Sec. II is used for all channels described in Sec. III and the systematic uncertainties summarized in Sec. IV. The statistical methods described in Sec. V are used to set limits on the signal strength as a function of \( m_H \).

The expected and observed limits from the individual channels entering this combination are shown in Fig. 3. The combined 95% CL exclusion limits on \( m_H \) are shown in Fig. 4 as a function of \( m_H \). These results are based on the asymptotic approximation. The \( \pm 1\sigma \) and \( \pm 2\sigma \) variation bands around the median background expectation are calculated using the improved procedure described in Sec. VD, which yields slightly larger bands compared to those in Ref. [14]. Typically the increase in the bands is of the order of \( \sim 5\% \) for the \( \pm 1\sigma \) band and 10%–15% for the \( \pm 2\sigma \) band. This procedure has been validated using ensemble tests and the Bayesian calculation of the exclusion.
The expected 95% CL exclusion region for the SM (μ = 1) hypothesis covers the m_H range from 120 GeV to 560 GeV. The addition of the H → τ⁺τ⁻ and H → bb channels, as well as the update of the H → WW( → ℓ⁺νℓ⁻→ν̄), brings a significant gain in sensitivity in the low mass region with respect to the previous combined search. For Higgs boson mass hypotheses below approximately 122 GeV the dominant channel is H → γγ. For mass hypotheses larger than 122 GeV but smaller than 200 GeV the H → WW( → ℓ⁺νℓ⁻→ν̄) channel is the most sensitive. In the mass range between ~200 GeV and ~300 GeV the H → ZZ( → ℓ⁺ℓ⁻→ν̄) channel dominates. For higher mass hypotheses, the H → ZZ( → ℓ⁺ℓ⁻→ν̄) channel leads the search sensitivity. The updates of the H → WW( → ℓ⁺νℓ⁻→ν̄), H → ZZ( → ℓ⁺ℓ⁻→ν̄), and H → ZZ( → ℓ⁺ℓ⁻→q̄q̄) channels improve the sensitivity in the high mass region.

The observed exclusion regions range from 111.4 GeV to 116.6 GeV, from 119.4 GeV to 122.1 GeV, and from 129.2 GeV to 541 GeV at 95% CL under the SM (μ = 1) hypothesis. The mass range 122.1 GeV to 129.2 GeV is not excluded due to the observation of an excess of events above the expected background. This excess and its significance are discussed in detail in Sec. VII.

Two mass regions where the observed exclusion is stronger than expected can be seen in Fig. 4. In the low mass range, Higgs mass hypotheses in the 111.4 GeV to 116.6 GeV range are excluded due mainly to a local deficit of events in the diphoton channel with respect to the expected background. A similar deficit is observed in the
of the distribution of $m_H$ from deficits in the high mass region in the range 360 GeV to 420 GeV, resulting from deficits in the $H \rightarrow ZZ^{(*)} \rightarrow \ell^+ \ell^- \ell^+ \ell^-$ and $H \rightarrow ZZ \rightarrow \ell^+ \ell^- \nu \bar{\nu}$ channels. Both fluctuations correspond to approximately 2 standard deviations in the distribution of upper limits expected from background only.

A small mass region near $m_H \sim 245$ GeV was not excluded at the 95% CL in the combined search of Ref. [14], mainly due to a slight excess in the $H \rightarrow ZZ^{(*)} \rightarrow \ell^+ \ell^- \ell^+ \ell^-$ channel. This mass region is now excluded. The CL$_s$ values for $\mu = 1$ as a function of the Higgs boson are shown in Fig. 5, where it can also be seen that the region between 130.7 GeV and 506 GeV is excluded at the 99% CL. The observed exclusion covers a large part of the expected exclusion range.

VII. SIGNIFICANCE OF THE EXCESS

The observed local $p$-values, calculated using the asymptotic approximation, as a function of $m_H$ and the expected value in the presence of a SM Higgs boson signal at that mass are shown in Fig. 6 in the entire search mass range 110–600 GeV. The observed local $p$-values, calculated using the asymptotic approximation, as a function of $m_H$ and the expected value in the presence of a SM Higgs boson signal at that mass. The combined observed local $p_0$ for a background-only experiment to be more signal-like than the observation, for individual channels and the combination, (a) in the full mass range of 110–600 GeV and (b) in the low mass range of 110–150 GeV. The solid curves give the observed individual and combined $p_0$. The dashed curves show the median expected value under the hypothesis of a SM Higgs boson signal at that mass. The combined observed local $p_0$ estimated using ensemble tests and taking into account energy scale systematic (ESS) uncertainties is illustrated in (c); the observed and expected combined results using asymptotic formulas are also shown. The horizontal dashed lines indicate the $p_0$ corresponding to significances of $1\sigma$, $2\sigma$, and $3\sigma$ for (a), (b), and (c).
The largest significance for the combination is observed for $m_H = 126$ GeV, where it reaches $3.0\sigma$ with an expected value in the presence of a signal at that mass of $2.9\sigma$. The observed (expected) local significances for $m_H = 126$ GeV are $2.8\sigma$ ($1.4\sigma$) in the $H \rightarrow \gamma \gamma$ channel and $2.1\sigma$ ($1.4\sigma$) in the $H \rightarrow ZZ^{(*)} \rightarrow \ell^+\ell^-\ell^+\ell^-$ channel. In the $H \rightarrow WW^{(*)} \rightarrow \ell^+\nu\ell^-\bar{\nu}$ channel, which has been updated and includes additional data, the observed (expected) local significance for $m_H = 126$ GeV is $0.8\sigma$ ($1.9\sigma$); the observed significance was previously $1.4\sigma$ [14].

The significance of the excess is not very sensitive to energy scale and resolution systematic uncertainties for photons and electrons; however, the presence of these uncertainties leads to a small deviation from the asymptotic approximation. The observed $p_0$ including these effects is therefore estimated using ensemble tests. The results are displayed in Fig. 6 as a function of $m_H$. The effect of the energy scale systematic uncertainties is an increase of approximately 30% of the corresponding local $p_0$. The maximum local significance decreases slightly to $2.9\sigma$. The muon momentum scale systematic uncertainties are smaller and therefore neglected.

The global $p$-values for the largest excess depend on the range of $m_H$ and the channels considered. The global $p_0$ associated with a $2.8\sigma$ excess anywhere in the $H \rightarrow \gamma \gamma$ search domain of 110–150 GeV is approximately 7%. A $2.1\sigma$ excess anywhere in the $H \rightarrow ZZ^{(*)} \rightarrow \ell^+\ell^-\ell^+\ell^-$ search range of 110–600 GeV corresponds to a global $p_0$ of approximately 30%.

The global probability for a $2.9\sigma$ excess in the combined search to occur anywhere in the mass range 110–600 GeV is estimated to be approximately 15%, decreasing to 5%–7% in the range 110–146 GeV, which is not excluded at the 99% CL by the LHC combined SM Higgs boson search [62]. The data are observed to be consistent with the background-only hypothesis except for the region around

range and in the low mass range. The asymptotic approximation has been verified using ensemble tests which yield numerically consistent results.
FIG. 9 (color online). The best-fit signal strength $\mu$ as a function of the Higgs boson mass hypothesis for the channels (a) $H \to \gamma\gamma$, (b) $H \to ZZ^{(*)} \to \ell^+\ell^-\ell^+\ell^-$ in the low mass region, (c) $H \to ZZ^{(*)} \to \ell^+\ell^-\ell^+\ell^-$ across the full search range, (d) $H \to ZZ \to \ell^+\ell^-\ell^+\ell^-\nu\bar{\nu}$, (e) $H \to ZZ \to \ell^+\ell^-q\bar{q}\ell^+\ell^-q\bar{q}$, (f) $H \to ZZ^{(*)}$ for all subchannels across the full search range, (g) $H \to WW^{(*)} \to \ell^+\ell^-\nu\bar{\nu}$ in the low mass region, (h) $H \to WW^{(*)} \to \ell^+\ell^-\nu\bar{\nu}$ across the full search range, (i) $H \to WW \to \ell\nu q\bar{q}$, (j) $H \to WW^{(*)}$ for all subchannels in the full mass range, (k) $H \to \tau\tau$, and (l) $H \to bb$. The band shows the interval around $\bar{\mu}^\text{best}$ corresponding to a variation of $-2\ln(\lambda(\mu)) < 1$. 

032003-16
$m_H = 126$ GeV. The observed and expected ratio $-2 \ln(\lambda(1)/\lambda(0))$ is shown in Fig. 7, which indicates a departure from the background-only hypothesis similar to the signal-plus-background expectation.

The best-fit value of $\mu$, denoted $\tilde{\mu}$, is displayed for the combination of all channels in Fig. 8 and for individual channels in Fig. 9 as a function of the $m_H$ hypothesis. A summary of $-2 \ln(\mu(\mu)) < 1$ intervals at three specific Higgs boson mass hypotheses ($m_H = 119$ GeV, 126 GeV, and 130 GeV) for each Higgs decay mode and the combination is given in Fig. 10. The bands around $\tilde{\mu}$ illustrate the $\mu$ interval corresponding to $-2 \ln(\mu(\mu)) < 1$ and represent an approximate $\pm 1\sigma$ variation. While the estimator $\tilde{\mu}$ is allowed to be negative in Figs. 8 and 9 in order to illustrate the presence and extent of downward fluctuations, the $\mu$ parameter is bounded to ensure non-negative values of the probability density functions in the individual channels. Hence, for negative $\tilde{\mu}$ values close to the boundary, the $-2 \ln(\mu(\mu)) < 1$ region does not reflect a calibrated 68% confidence interval. It should be noted that the $\tilde{\mu}$ does not directly provide information on the relative strength of the production modes. The excess observed for $m_H = 126$ GeV corresponds to a $\tilde{\mu}$ of 1.1 $\pm 0.4$, which is compatible with the signal strength expected from a SM Higgs boson at that mass ($\mu = 1$).

**VIII. CONCLUSION**

A combined search for the Standard Model Higgs boson has been performed with the ATLAS detector in the $\sqrt{s} = 7$ TeV $pp$ collision data collected in 2011, corresponding to an integrated luminosity of 4.6–4.9 fb$^{-1}$. The channels used in this combination are $H \rightarrow \gamma\gamma$, $H \rightarrow bb$, $H \rightarrow \tau^+\tau^-$, $H \rightarrow ZZ^{(*)} \rightarrow \ell^+\ell^-\ell^+\ell^-$. $H \rightarrow ZZ \rightarrow \ell^+\ell^-q\bar{q}$, $H \rightarrow ZZ \rightarrow \ell^+\ell^-\nu\bar{\nu}$, $H \rightarrow WW^{(*)} \rightarrow \ell^+\nu\ell^-\bar{\nu}$, and $H \rightarrow WW \rightarrow \ell_1\ell_2q\bar{q}$. The observed exclusion ranges at the 95% CL are 111.4 GeV to 116.6 GeV, 119.4 GeV to 122.1 GeV, and 129.2 GeV to 541 GeV, while Higgs boson masses between 120 GeV and 560 GeV are expected to be excluded at the 95% CL. The mass region between 130.7 GeV and 506 GeV is excluded at the 99% CL.

The local significance of the observed excess when all channels are combined is 2.9$\sigma$, in good agreement with the expected significance in the presence of a SM Higgs boson with $m_H = 126$ GeV of 2.9$\sigma$. An estimate of the global probability for such an excess to occur anywhere in the full explored Higgs boson mass region (from 110 GeV to 600 GeV) is approximately 15%. The global $p_0$ in the range not excluded at the 99% CL by the LHC combined Higgs boson search results [62] (from 110 GeV to 146 GeV) is approximately 5%–7%.

**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; BMWF, 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 and ERC, European Union; IN2P3-CNRS, CEA-DSM/IRFU, France; GNAS, Georgia; BMBF, DFG, HGF, MPG and AvH Foundation, Germany; GSRT, Greece; ISF, MINERVA, GIF, DIP and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; FOM and NWO, Netherlands; RCN, Norway; MNI Sw, Poland; GRICES and FCT, Portugal; MERSYS (MECTS), Romania; MES of Russia and ROSATOM, Russian Federation; JINR; MSTD, Serbia; MSSR, Slovakia; ARRS and MVZT, Slovenia; DST/NRF, South Africa; MICINN, 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.
COMBINED SEARCH FOR THE STANDARD MODEL HIGGS


COMBINED SEARCH FOR THE STANDARD MODEL HIGGS ... PHYSICAL REVIEW D 86, 032003 (2012)


(ATLAS Collaboration)

1Physics Department, SUNY Albany, Albany, New York, USA
2Department of Physics, University of Alberta, Edmonton, Alberta, Canada
3Department of Physics, Ankara University, Ankara, Turkey
4Division of Physics, TOBB University of Economics and Technology, Ankara, Turkey
5LAPP, CNRS/IN2P3 and Université de Savoie, Annecy-le-Vieux, France
6High Energy Physics Division, Argonne National Laboratory, Argonne, Illinois, USA
7Department of Physics, University of Arizona, Tucson, Arizona, USA
8Department of Physics, The University of Texas at Arlington, Arlington, Texas, USA
9Physics Department, National Technical University of Athens, Zografou, Greece
10Institut de Física d’Altes Energies and Departament de Física de la Universitat Autònoma de Barcelona and ICREA, Barcelona, Spain
11Institut de Física de Partícules UPF, Barcelona, Spain
12aInstitute of Physics, University of Belgrade, Belgrade, Serbia
12bVinca Institute of Nuclear Sciences, University of Belgrade, Belgrade, Serbia
13Department for Physics and Technology, University of Bergen, Bergen, Norway
14Physics Division, Lawrence Berkeley National Laboratory and University of California, Berkeley, California, USA
15Department of Physics, Humboldt University, Berlin, Germany
16Albert Einstein Center for Fundamental Physics and Laboratory for High Energy Physics, University of Bern, Bern, Switzerland
17School of Physics and Astronomy, University of Birmingham, Birmingham, United Kingdom
18Physics Department, Bogazici University, Istanbul, Turkey
19Division of Physics, Dogus University, Istanbul, Turkey
18cDepartment of Physics, Gaziantep University, Gaziantep, Turkey
18dDepartment of Physics, Istanbul Technical University, Istanbul, Turkey
19aINFN Sezione di Bologna, Italy
19bDipartimento di Fisica, Università di Bologna, Bologna, Italy
18bDepartment of Physics, Bogazici University, Istanbul, Turkey
20Physikalisches Institut, University of Bonn, Bonn, Germany
21Department of Physics, Boston University, Boston, Massachusetts, USA
22Department of Physics, Brandeis University, Waltham, Massachusetts, USA
23aUniversidade Federal do Rio de Janeiro COPPE/EE/IF, Rio de Janeiro, Brazil
23bFederal University of Júlio de Santa Maria (UFSJ), Santa Maria, Brazil
23cInstituto de Fisica, Universidade de Sao Paulo, Sao Paulo, Brazil
24Physics Department, Brookhaven National Laboratory, Upton, New York, USA
25aNational Institute of Physics and Nuclear Engineering, Bucharest, Romania
25bUniversity Politehnica Bucharest, Bucharest, Romania
25cWest University in Timisoara, Timisoara, Romania
26Departamento de Física, Universidad de Buenos Aires, Buenos Aires, Argentina
27Cavendish Laboratory, University of Cambridge, Cambridge, United Kingdom
28Department of Physics, Carleton University, Ottawa, Ontario, Canada
29CERN, Geneva, Switzerland