Search for the Higgs boson in the \( H \to ZZ^{(*)} \to 4\ell \) decay channel with the ATLAS experiment

Detailed PhD Research Project

**PhD Candidate:** Valerio Ippolito  
**Supervisor:** Prof. Carlo Dionisi

Università degli Studi di Roma “La Sapienza”  
Dottorato in Fisica, XXVI Ciclo

1 The Higgs boson

A number of experiments in the last decades have shown how accurate is the description of the sub-atomic world provided by the Standard Model of particle physics (SM)[1]. Precision measurements at LEP and Tevatron have been able to probe the validity of the predictions of this theory, thus showing the reliability of its current description of the electroweak interactions.

Nevertheless, the nature of the spontaneous breaking of the electroweak gauge symmetry, which is necessary to explain the non–zero mass of its gauge bosons, is still an unsolved puzzle from an experimental point of view. In the SM, the introduction of a scalar field with a non–vanishing vacuum expectation value gives rise to an additional particle \( H \), the Higgs boson[2]. The mass of the electroweak gauge bosons arises from their coupling to the Higgs field, while fermion masses are explained through Yukawa coupling between them and the Higgs field — the coupling constants being proportional to their masses.

The Higgs boson mass itself, \( m_H \), is a free parameter of the theory. At different values of \( m_H \) correspond different production cross sections and dominant decay channels: different experimental strategies therefore arise, depending on the \( m_H \) range being probed, to reconstruct an hypothetical Higgs signal. Figure 1 shows the Feynman diagrams for various Higgs production mechanisms: in \( pp \) collisions, gluon–gluon fusion (ggF) gives the dominant contribution, while vector boson fusion (VBF) and associate production of an Higgs boson with a vector boson are less significant. The production cross section and decay branching ratios are shown, as a function of \( m_H \) and for a center–of–mass energy \( \sqrt{s} = 7 \) TeV, in Fig. 2.

No evidence for an Higgs boson has been observed so far. Experimental limits from direct searches exist on its existence for various values of \( m_H \) (Fig. 3), as well as indirect limits from precision measurements of electroweak parameters sensitive to \( m_H \),
Figure 1: Feynman diagrams for Higgs production via a) gluon fusion, b) vector boson fusion, c) associate production with a $W$ boson and d) with a $t\bar{t}$ pair.

Figure 2: Left: Higgs production cross sections in $pp$ collisions at $\sqrt{s} = 7$ TeV. Right: Higgs decay branching ratios for various processes, as a function of its mass.
Figure 3: Observed (full line) and expected (dashed line) 95% CL upper limits on the SM Higgs boson production cross section divided by the Standard Model expectation as a function of $m_H$, obtained by the ATLAS collaboration.

which seem to favour an Higgs boson with mass close to the lower limit set by LEP ($m_H > 114$ GeV). Despite no single experiment has reported an evidence for a signal, tantalizing hints at a level close to $3\sigma$ in the $H \to \gamma\gamma$ channel have been reported in the 125 GeV Higgs mass region by both the ATLAS and CMS collaborations and a slight excess in both $ZZ$ (ATLAS) or $ZZ$ and $WW$ channels (CMS) has been reported recently using data collected at the LHC in 2011[3, 4].

2 $H \to ZZ^{(*)} \to \ell^+\ell^-\ell^+\ell^-$: a golden channel

Various are the decay channels one might exploit in this low mass region ($114 < m_H < 150$ GeV); the dominant cross section times branching ratio channels are $H \to \gamma\gamma$ and $H \to W^+W^-$, experimentally reconstructed respectively asking two high energy photons in the final state and one or two leptons plus missing energy, compatible with a leptonic decay of one or both the $W$ bosons.

The $pp \to H \to ZZ^{(*)} \to 4\ell$ process, on the other hand, has a lower cross section ($2 \div 5$ fb, to be compared with the $\sim 40$ fb of $pp \to H \to \gamma\gamma$), but profits from the clean signature arising from the presence of leptons in the final state, an ingredient which is able to strongly suppress the contribution of the QCD background typical of an hadron collider. This channel features therefore a high signal/background ratio — the main background being the diboson production $pp \to ZZ^{(*)} \to 4\ell$, with minor contribution from $Z + jj$ and $t\bar{t}$ — with the Higgs evidence being a narrow peak in the invariant mass distribution of the four lepton system, rising over the backgrounds.

Furthermore, an observation of the Higgs boson in the four lepton channel is cru-
cial to assess the SM nature of this particle. Unlike it happens for example in $H \rightarrow W^+W^- \rightarrow \ell^+\nu\ell^-\nu$, the full final state can be reconstructed, allowing high precision measurements of the Higgs mass from 120 GeV to 1 TeV and of its intrinsic properties, such as spin, CP property and coupling structures.

3 The ATLAS experiment at the LHC

The Large Hadron Collider (LHC[5]) at CERN is a $pp$ and heavy ions collider, with a physics programme ranging from the confirmation of the Standard Model with the observation of the Higgs boson, to new physics searches in the framework of alternative theories. In March 2010 it begun its $pp$ physics program, delivering collisions at $\sqrt{s} = 7$ TeV up to November 2011, with a peak instantaneous luminosity of $3.6 \cdot 10^{33}$ cm$^{-2}$s$^{-1}$.

The ATLAS detector[6] has collected about 4.9 fb$^{-1}$ of integrated luminosity in 2011. ATLAS is a multi-purpose experiment, consisting of a system of high precision detectors in cylindric symmetry around the beam axis. It is composed by an inner tracker (ID) in an axial magnetic field of 2 T, to reconstruct the trajectories of charged particles, an electromagnetic (ECAL) and an hadronic (HCAL) calorimeter for the reconstruction of electromagnetic and hadronic showers, and an outer spectrometer (MS) in a toroidal magnetic field for the reconstruction of penetrating charged particles like muons. Electrons are reconstructed matching an ID track to an energy cluster in ECAL, while muons are reconstructed combining an ID track with a MS track.

4 Search for the Higgs boson in the $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$ channel

The goal of this PhD thesis is to either exclude the SM Higgs boson in the full mass range allowed by the theory, or to observe it in the four lepton channel, measuring its production cross section and properties.

In this section, we will show the results we obtained so far analyzing 2011 collision data[7] and present our plans for 2012 data taking.

4.1 2011 data analysis

4.1.1 Event selection

A total integrated luminosity of about 4.8 fb$^{-1}$, 4.8 fb$^{-1}$ and 4.9 fb$^{-1}$ has been analyzed, respectively, for the $4\mu$, $2e2\mu$ and $4e$ final states. Events are selected using single and dilepton triggers, with a single lepton threshold of transverse momentum $p_T > 18$ GeV for muons and of transverse energy $E_T > 20 \div 22$ GeV for electrons, depending on LHC instantaneous luminosity, and dilepton trigger thresholds of $p_T > 10$ GeV and $E_T > 12$ GeV, respectively.
Leptons are selected if they have \( p_T > 7 \) GeV, pseudorapidity \( |\eta| < 2.7 \) (muons) or \( |\eta| < 2.47 \) (electrons) and if they are isolated (low ECAL energy deposit and \( \sum p_T \) of ID tracks lying within a cone open around the lepton). Higgs candidates are reconstructed searching for two same-flavour opposite-charge lepton pairs, with an additional \( p_T > 20 \) GeV requirement on two of the leptons, a cut on the invariant mass of the lepton pair \( Z_1 \) closest to the \( Z \) mass \( |m_{Z_1} - m_Z| < 15 \) GeV and a \( m_H \)-dependent cut on the (possibly virtual) second lepton pair \( Z_2 \). The significance of the transverse impact parameter \( d_0 \) of the two lowest \( p_T \) leptons must be compatible with their provenance from the primary interaction vertex, to reject \( Z + bb \) backgrounds.

The signal reconstruction and selection efficiencies for \( m_H = 130 \) GeV (\( m_H = 360 \) GeV) are 27% (60%) for the 4\( \mu \) channel, 18% (52%) for the 2\( e2\mu \) channel and 14% (45%) for the 4\( e \) channel. With the integrated luminosity collected in 2011, we expect for the 130 GeV mass hypothesis 1.00 \( \pm \) 0.17 signal events in the 4\( \mu \) channel, 1.22 \( \pm \) 0.21 in the 2\( e2\mu \) channel and 0.43 \( \pm \) 0.08 events in the 4\( e \) channel.

### 4.1.2 Backgrounds

Relevant backgrounds to this search are \( pp \rightarrow ZZ^\ast \rightarrow 4\ell \), \( Z + jj \) and \( tt \). The last two backgrounds are strongly suppressed by the isolation cuts on the four leptons and by the \( m_{Z_1} \) and \( m_{Z_2} \) cuts, while the irreducible background \( pp \rightarrow ZZ^\ast \rightarrow 4\ell \) has the same final state and similar kinematics as the signal, thus it gives the dominant contribution to the overall background.

Background estimation is done both using Monte Carlo (MC) simulation and data-driven techniques (extrapolating from background-enriched control regions). The irreducible background is estimated from MC, with an assigned systematic uncertainty on its yield of 15%. The \( Z + jj \) background is estimated from a data sample obtained selecting events with a \( Z \) boson plus another lepton pair, without applying isolation and impact parameter cuts on these two leptons; a systematic uncertainty of 20 \( \div \) 40% has been assigned to its yield, depending on the flavour of the jets. The \( tt \) background is estimated from MC and its normalization has been verified with a control sample of opposite sign electron–muon pairs with mass consistent with \( m_Z \) and two additional same-flavour leptons.

### 4.1.3 Results

Figure 4 shows the invariant mass distribution of the selected Higgs candidates. We observed 71 events in the full mass range 100 \( \div \) 600 GeV, while expecting 62 \( \pm \) 9 from background.

Limits on \( m_H \) obtained from these distributions are shown in Fig. 5. We observe three excesses of events with respect to the background-only hypothesis for Higgs boson masses of 125 GeV, 244 GeV and 500 GeV, with local significances of 2.1, 2.2 and 2.1 standard deviations. The expected 95% confidence level upper limit on the Higgs boson production cross section in terms of the SM cross section is, for \( m_H = 125 \) GeV,
Figure 4: Invariant mass of the reconstructed Higgs candidates (black dots), with overlaid the background expectation (red histogram) and the signal expectation for various Higgs mass hypotheses, in the low mass range (left) and in the full mass range considered by this analysis (right).

about 1.5$\sigma_{SM}$, comparable to that of the two most sensitive channels in this region (the expected limit from $H \rightarrow \gamma\gamma$ is about 1.5$\sigma_{SM}$ and the one from $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ is about 1$\sigma_{SM}$). Combining this result with all the other searches[3], we obtain the limit shown in Fig. 6: we exclude a SM Higgs boson at the 95% confidence level in the mass ranges from 110.0 GeV to 117.5 GeV, 118.5 GeV to 122.5 GeV, and 129 GeV to 539 GeV.

An excess of events is observed in the combination around $m_H = 126$ GeV, with a local significance of 2.5 standard deviations, the main contributions being those coming from $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$.

4.2 Plans for 2012 data analysis

In 2012, LHC will operate at $\sqrt{s} = 8$ TeV, with a peak instantaneous luminosity of about $6 \cdot 10^{33}$ cm$^{-2}$s$^{-1}$. ATLAS plans to collect 15 ÷ 20 fb$^{-1}$ of integrated luminosity.

The plans of this thesis work for 2012 and 2013 strongly depend on whether this new data will enhance or rule out the significance of the 126 GeV excess. Our strategy will be to maximize the sensitivity of our analysis to a low–mass Higgs signal:

- I will focus on the optimization of the cut–based analysis with MC simulation and 2011 data, assessing its performances with the experimental setting of the 2012 run (pile–up issues, improvements in Higgs mass resolution, acceptance extension, tuning of the rejection of reducible backgrounds);
Figure 5: Observed (full line) and expected (dashed line) 95% CL upper limits on the SM Higgs boson production cross section divided by the Standard Model expectation as a function of $m_H$, obtained from the $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$ analysis.

- I will analyze new data throughout the 2012 run, taking care of candidate reconstruction and data–driven background estimation;

- I will then focus, after the 2012 Summer conferences, on the development of a multi–variate analysis, to maximize our sensitivity using the full signal/background separation power provided by the observables of our candidates, possibly adding a dedicated VBF analysis and spin–parity studies to gain some sensitivity to models other than the SM.

4.2.1 Optimization against event pile–up

LHC will run with an instantaneous luminosity double with respect to the value we had in 2011. This means that, for every $pp$ collision, we expect on average 35 additional $(pile–up)$ interactions on top of the interesting hard interaction.

The background originating from these pile–up interaction consists of energy deposit in the detector. This affects lepton reconstruction performances, since isolation criteria used to select and reconstruct electrons and muons will have lower efficiencies with higher event pile–up. I have shown how this effect is under control in 2011 data in the muon case, and I will work in the optimization of lepton isolation cuts to cope with the more challenging conditions we will have in 2012.

4.2.2 $Z$ mass constraint

The final discriminating variable of our analysis is the reconstructed four lepton invariant mass $m_{4\ell}$; the narrower is the expected signal peak in this distribution, the higher
Figure 6: Observed (full line) and expected (dashed line) 95% CL upper limits on the SM Higgs boson production cross section divided by the Standard Model expectation as a function of $m_H$, obtained combining all ATLAS Higgs searches.
is our signal sensitivity. The width of the signal peak is dominated, for \( m_H < 350 \text{ GeV} \), by the resolution on \( m_{4\ell} \), while at higher masses the dominant contribution comes from the Higgs intrinsic width \( \Gamma_H \). It is therefore in the low mass region that we can profit from enhancing our mass resolution: this can be achieved both dividing the analyzed sample in different subcategories, selected according to the resolution on the single leptons and analyzed separately, or using the fact that one or both the Z bosons are on–shell.

The Z mass constraint method corrects lepton momenta in order for the invariant mass of the lepton pair to be that of the Z boson in that event within intrinsic Z boson natural width and event–by–event resolution; in this way, we are able to shrink the Z peak, leading to a narrower Higgs peak and higher signal sensitivity. Since the Z boson has a non–zero intrinsic width \( \Gamma_Z \), we estimate the true Z mass \( m_{Z}^{\text{true}} \) in each event by maximizing the likelihood

\[
\mathcal{L}(m_{Z}^{\text{true}}) \equiv p(m_{Z}^{\text{true}}|m_{2\ell}) \sim p(m_{2\ell}|m_{Z}^{\text{true}}) \cdot p(m_{Z}^{\text{true}})
\]

against \( m_{Z}^{\text{true}} \), where the first factor on the right–hand term of the equation represents the probability density function (p.d.f.) of the reconstructed dilepton mass \( m_{2\ell} \), the second factor represents the prior p.d.f. for \( m_{Z}^{\text{true}} \) and terms independent from \( m_{Z}^{\text{true}} \) have been neglected.

Using a gaussian resolution p.d.f. — its standard deviation being obtained from the uncertainties on lepton momenta — and a Breit–Wigner distribution with pole \( m_Z \) and width \( \Gamma_Z \) as prior p.d.f. on \( m_{Z}^{\text{true}} \), my work led to an \( O(10\%) \) better mass resolution for every \( m_H \), as estimated from MC simulation. For \( m_H = 125 \text{ GeV} \), this corresponds to a 5% better expected limit.

### 4.2.3 Extending analysis acceptance

Among the various ways to improve the sensitivity of the analysis, the loosening of lepton reconstruction and selection criteria is one of the most important. Extending the geometric acceptance of the detector and lowering the lepton \( p_T \) threshold or the \( m_Z \) cut can lead to significant rises in signal sensitivity. For example, loosening lepton selection criteria and reducing the lower \( m_{Z1} \) cut to take into account the asymmetry of the \( m_Z \) distribution for Z bosons coming from Higgs decays (which in principle does not follow a Breit–Wigner distribution like in the \( pp \rightarrow Z + X \) case), an efficiency rise of O(30%) is obtained in event selection efficiency for \( H \rightarrow ZZ^* \rightarrow 4e \) with \( m_H = 130 \text{ GeV} \). A detailed understanding of the resulting backgrounds, by means of simulation and data–driven techniques, is thus requested, in order to be able to keep background and contamination from hadrons and photons faking leptons under control. I am taking part into these studies, assessing the impact of such an extended analysis in background normalization and composition.
4.2.4 Optimization of $Z + bb/Z + jj$ rejection

Lepton isolations, computed as the sum of the $p_T$s of the ID tracks lying in a certain cone open around the lepton (track isolation) or as the sum of the ECAL energy deposits in the same cone (calorimetric isolation) are of course correlated. Further correlations arise for leptons in the $Z + bb$ background, among these isolation variables and the $d_0$ significance.

Multivariate techniques (MVA) can be used to take into account the different correlations for signal and background events, in order to discriminate signal from background and among backgrounds. I am evaluating the impact of using MVA discriminants like Fisher and BDT in $Z + bb$ and $Z + jj$ rejection (dominant reducible backgrounds for the muon and electron cases respectively) in both background rejection and discovery sensitivity.

4.2.5 Multivariate analysis

Significant improvements in lepton and Higgs candidate selection can be obtained extending MVA to the full selection, using those kinematic variables for which distributions are significantly different between signal and background. In this way, we can exploit all the features of our signal signature against those of both the reducible and irreducible backgrounds, using the full information we have on their kinematics (for the dominant $pp \rightarrow ZZ^{(*)} \rightarrow 4\ell$ background, there is no constraint for the reconstructed Higgs candidate to have spin zero like the SM Higgs signal does: this leads to different angular distributions of the decay products).

For example, taking into account the decay angles of leptons and $Z$ bosons in the Higgs center of mass system and the invariant mass of the two reconstructed $Z$s, a $10 \div 20\%$ improvement in the discovery significance at $m_H > 200$ GeV can be achieved, while for $m_H = 130$ GeV preliminary studies show an improvement of the order of $5 \div 7\%$. Multivariate techniques are the natural and most powerful means to obtain the best signal/background discrimination, provided a reliable understanding of the input distributions in both data and simulation has been obtained: I will work on this topic, taking care of the choice of the best set of input variables and of controlling the reliability of MC simulation used to train the discriminators.

4.2.6 Dedicated VBF analysis

The vector boson fusion process $qq \rightarrow H qq$, where two gauge bosons are emitted by the two initial state quarks and create an Higgs boson, has still a relevant cross section, ranging from one order of magnitude below $gg \rightarrow H$ at low masses to the same order of magnitude at high masses. A $H \rightarrow 4\ell$ analysis optimized for VBF must exploit the characteristic signature with two jets emitted at small angle, with significant enhancement of the signal/background ratio.

If we measure the ratio between the yield of events with such two jets and the yield of events obtained from the inclusive analysis, the correlated uncertainties (e.g.
luminosity) cancel out, and we can probe the ratio between the coupling constants of the Higgs boson to vector bosons and fermions. In presence of an Higgs boson, 2012 data would not suffice for a precision measurement of this ratio, but we could start to be sensitive to exotic scenarios predicting VBF/ggF ratios other than the SM one.

4.3 Perspectives

Data taken in 2012 will allow to either rule out or confirm the existence of an Higgs boson in the low mass region. Two different scenarios therefore arise.

4.3.1 First scenario: signal at \( m_H \sim 125 \text{ GeV} \)

In case a signal is found and confirmed, my work will focus on the possible studies to assess its SM nature. I will try to exploit the available statistics, which is not expected to suffice for a precision measurement of the intrinsic properties (mass, spin, CP, VBF/ggF ratio) of the signal, but can be enough to put limits on and possibly exclude alternative models other than the SM.

4.3.2 Second scenario: no Higgs boson is found

In case the low mass excess is not confirmed by new data, further studies will be needed to push the SM Higgs exclusion limit up to the high mass region (600 GeV ÷ 1000 GeV), where \( H \rightarrow ZZ \rightarrow 4\ell \) — together with \( H \rightarrow ZZ \rightarrow q\ell q\ell \) and \( H \rightarrow ZZ \rightarrow \ell\ell\nu\nu \) — is one of the most sensitive channels, and to optimize a multi-variate analysis, using the most advanced knowledge of detector performances and calibration, to be able to exclude a non-SM Higgs production cross section as low as possible.

In the absence of an Higgs boson, the Standard Model predicts a vector boson scattering cross-section which violates unitarity at the TeV scale. The further step of my work, in this case, will be to set up a vector boson scattering analysis, in the perspective of the high statistics LHC run starting in 2014, which will be able to provide the statistics (\( O(100 \text{ fb}^{-1}) \)) needed for these studies.

References


[3] The ATLAS Collaboration, “An update to the combined search for the Standard Model Higgs boson with the ATLAS detector at the LHC using up to 4.9 fb\(^{-1}\) of \(pp\) collision data at \(\sqrt{s} = 7\) TeV”, ATLAS-CONF-2012-019


