CMS Physics Analysis Summary

Higgs pair production at the High Luminosity LHC

The CMS Collaboration

Abstract

Studies of the Higgs boson pair production and decays into $bb\gamma\gamma$, $bb\tau\tau$, and $bbWW$ final states are presented. The studies are performed assuming the operational conditions of the High-Luminosity LHC, with an integrated luminosity of 3000 fb$^{-1}$, and the upgraded CMS experiment. Combining the studies of $bb\gamma\gamma$ and $bb\tau\tau$ final states, the expected significance for Higgs boson pair production is 1.9 standard deviation. The resulting expected uncertainty in the signal yield is 54%. The benefits of the CMS Phase-II upgrade, to meet the challenges presented by the high luminosity environment, are emphasized.

Version of November 15, 2015: additional references included.
1 Introduction

The detailed studies of the properties of the observed Higgs boson [1–3] by the ATLAS and CMS Collaborations are central components of the CERN LHC physics program. In the standard model (SM) all properties of the Higgs boson are predicted for a given mass of the Higgs boson. The present measurements of the Higgs bosons couplings to fermions and bosons and of the tensor structure of the Higgs boson interaction with electroweak gauge bosons show no significant deviations [4, 5] with respect to the SM expectations. Physics beyond the SM (BSM) can potentially lead to deviations from the SM predictions.

Studies of Higgs boson pair production at the High-Luminosity LHC (HL-LHC) will provide insight on the Higgs boson trilinear coupling [6–12]. This measurement would directly probe the Higgs field potential since the self-coupling is related to the third derivative of the Higgs potential at its minimum. The process is also sensitive to other BSM effects, as new physics can modify the rate of the production. The dominant Higgs boson pair production mode at the LHC is through gluon fusion. Figure 1 shows the dominant Feynman diagrams. Di-Higgs boson events can be produced via a box diagram and through the Higgs boson self-coupling contribution. The two processes interfere destructively and the cross section is near minimum for the SM. The Higgs boson pair production cross section is about 1000 times smaller than the single Higgs boson production cross section. It should be noted that the cross section increases by a factor of two if the Higgs boson self coupling is zero.

Figure 1: Feynman diagrams contributing to gluon fusion Higgs boson pair production.

The gluon fusion production cross section for the Higgs boson pair production at a center of mass energy of 14 TeV has been calculated to next-to-next-to-leading-order to be 40.7 fb [13, 14]. Other production modes potentially add another 10% to the total production cross section, but have not been included in this study. A large amount of integrated luminosity is required in order for these processes to be observed and measured. The signal events that are expected to be produced per experiment at HL-LHC with 3000 fb$^{-1}$ integrated luminosity are approximately 320, 9000, and 1500 events for $bb\gamma\gamma$, $bb\tau\tau$, and the leptonic decays of $bbWW$ final states, respectively.

The increased instantaneous luminosity required to reach this target is expected to be accompanied by a significant increase in the number of overlapping inelastic collisions (pileup). The HL-LHC is expected to operate with an average of 140 simultaneous pileup events. This presents a serious challenge to the experiments in their ability to deal with this increased level of activity
and energy flow, and to preserve the detector performance under this environment.

As part of a comprehensive strategy to address these issues, CMS has released a technical proposal for the Phase-II upgrade [15] program. The expected performance of this detector at HL-LHC is assumed for these studies and is discussed in the following sections. The impact of some of the individual components of the Phase-II upgrade on the results are highlighted where it is appropriate. In addition, the $bb\gamma\gamma$ results are also shown assuming the detector performance of the so called Phase-I CMS upgrade [16] detector after an assumed integrated luminosity of 1000 $fb^{-1}$, configuration hereafter denoted as “Phase-I aged”.

At present, the Phase-II detector simulation includes the upgraded outer tracker, muon systems, and calorimetry. The pixel detector upgrade, however, is still not finalized so the simulation contains the Phase-I pixel detector in the barrel and an extended version of the current pixel detector in the forward detector to provide tracking at higher $\eta$. The primary and secondary vertex reconstruction and identification performance will certainly be better than what is assumed in these studies and should be viewed as a conservative estimate.

# Experimental setup and signal simulation

It is crucial that the Phase-II detector can cope with the challenging environment of HL-LHC, as pileup mitigation, b-tagging, tau-tagging, photon identification efficiencies, and mass resolutions are fundamental to perform measurements on Higgs boson pair production. Triggers are assumed to be 100% efficient in these studies. The DELPHES fast simulation framework [17] is used for $bbWW$ results to model the Phase-II detector. The parameterized performance of the Phase-II detector in Delphes is taken from the corresponding GEANT-based [18] full simulation samples. The $bb\gamma\gamma$ analysis uses Monte Carlo (MC) generator information with smearing functions to model the performance of the detector. A combination of the two approaches mentioned above is used for the $bb\tau\tau$ final state.

Only the dominant gluon fusion inclusive production mode is produced in the signal generation. The samples are generated with MADGRAPH5.2 [19], at leading order (LO), using the results from [20], interfaced with PYTHIA6.4 [21] for parton showering and fragmentation. The generator is also interfaced with TAUOLA [22] for the simulation of the tau lepton decays. The pileup events are simulated in the Delphes samples by randomly placing minimum-bias interactions along the beam axis according to a longitudinal spread taken from the full simulation samples.

# $bb\gamma\gamma$ final state

## Object selection and performance

The signal events of interest contain two high transverse momentum ($p_T$) photons and two high $p_T$ jets originating from b quarks. The photons are rejected if an electron is reconstructed within a distance $\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2}$ (where the $\phi$ is azimuthal angle in radians and $\eta$ is the pseudorapidity) of 0.1 to the photon. A $p_T$ and $\eta$ dependent efficiency is applied to the photons to model the identification and isolation efficiency. The efficiency is about 80% in the barrel, and about 55% in the endcap. The lower efficiency in the endcap is primarily due to the electron veto requirement. The processes involving jets faking photons are among the dominant backgrounds. The rate to misidentify photons, typically referred to as the fake rate, is about $1 \times 10^{-4}$ for gluon jets and about $5 \times 10^{-4}$ for quark jets. The rate to misidentify electrons as photons
3.2 Signal and background estimation

is taken to be as 1% (3%) in barrel (endcap). Di-photon mass width of 1.2 GeV is achieved with the upgraded detector, for the events where both photons are in the barrel, using sophisticated multivariate techniques to calibrate the photon energy. Jets are reconstructed using the anti-$k_T$ clustering algorithm [23] with the resolution parameter equal to 0.5, and corrected for pileup effects using the FastJet technique [24, 25]. Standard CMS jet energy corrections for non-uniformities in $\eta$ and $p_T$ are applied. Jets are tagged as originating from a b quark on the basis of the presence of secondary vertices and large impact parameter tracks, which are exploited for b-tagging in a Combined Secondary Vertex discriminator (CSV) [26–28]. The chosen working point of the CSV discriminator gives, on average, b-tagging efficiencies of about 75% and 65% in the central and forward regions, respectively; with mistagging rates for light and charm jets of about 1% and 20%, respectively. Di-jet mass resolution of about 20 GeV is achieved with the upgraded detector.

Electrons and muons with $p_T > 10$ GeV are selected for the purpose of vetoing events with signatures consistent with a Higgs boson produced in association with a top and anti top quark pair ($t\bar{t}H$). This process contributes significantly to the total background after signal event selection requirements. A non-negligible fraction of such background events contain leptons and can be rejected on this basis. Very loose selection requirements with efficiency in the range between 90% and 95% are placed on the electron and muon candidates in order to suppress this background as much as possible.

3.2 Signal and background estimation

The signal process of interest is the production of two Higgs bosons, one of which decays to a pair of b quarks, and the other decaying to a pair of photons. The main resonant backgrounds are $ZH$, where a Higgs boson is produced in association with a Z boson which subsequently decays to two b-jets, $t\bar{t}H$, where a Higgs boson is produced in association with a top and anti-top quark pair, and $b\bar{b}H$, where a Higgs boson is produced in association with a b and anti-b quark pair. The non-resonant backgrounds include QCD production of $b\bar{b}\gamma\gamma$, QCD production of $jj\gamma\gamma$ with light jets mistagged as b-jets, QCD production of $b\bar{b}\gamma\gamma$ and $b\bar{b}jj$ with one and two jets mis-identified as photons respectively, and QCD production of four jets with two jets mis-identified as photons and two jets mistagged as b-jets, dominated by mistagged charm jets. These background processes have cross sections that are several orders of magnitude larger than the resonant backgrounds, but are suppressed by the low rate for mistags and mis-identified photons. Due to their large cross sections, it is computationally impossible to fully simulate these background events. Instead, we adopt the approach of producing generator particle level Monte Carlo (MC) samples and weighting the events by the corresponding efficiencies and fake rates for selecting the constituent particles. Finally, SM production of $t\bar{t}(\gamma)$ enters as background for our signal event selection for events where both top quarks decay semi-leptonically to electrons and one (or both) of the electrons are mis-identified as photons. A sample of di-electron decays of the $t\bar{t}(\gamma)$ events are weighted by the corresponding electron to photon mis-identification rates to estimate the background contribution.

The cross sections for non-resonant background processes are summarized in Table 1. Recent studies of the non-resonant background production of $b\bar{b}\gamma\gamma$ show that the rate is significantly enhanced when full next-to-leading order (NLO) simulation is performed [29]. The NLO simulation includes both the real and virtual corrections. The k-factor is about 2 and this is included in the table.
Table 1: The production cross section of the non-resonant background processes at a center of mass energy of 14 TeV are shown.

3.3 Event selection

Events containing two photons with $p_T$ greater than 25 GeV and $|\eta| < 2.5$, and two b-agged jets with $p_T$ greater than 30 GeV and $|\eta| < 2.4$ are selected. While the Phase-II upgrades allow b-jet tagging capabilities up to $|\eta|$ ranges of 3.0, the b-tagged jets with $p_T$ greater than 30 GeV for the signal events are predominantly central and therefore only $|\eta| < 2.4$ is required. One of the two photons is required to have $p_T > 40$ GeV. Due to the large amount of background from jets faking photons in the endcap region of the detector the event sample is split up into two categories, one with both photons and in the barrel and one with at least one photon in the endcap. To suppress $t\bar{t}H$ background events, it is required that there are no electrons or muons passing the veto selection and that the number of jets with $p_T > 30$ GeV and $|\eta| < 2.5$ is less than four.

A number of different additional kinematic requirements were investigated in order to improve the signal to background ratio. It is required that the $\Delta R$ between the two photons and the $\Delta R$ between the two b-jets are less than 2.0, and that the minimum of the $\Delta R$ between photons and b-jets is larger than 1.5. With the above kinematic selection requirements, a signal to background ratio of about 1 : 3 is achieved.

The expected event yields for the signal and resonant background processes, and the non-resonant background processes for various stages of the event selection are shown in Table 2. For the event category with both photons in the barrel, the dominant backgrounds are $bb\gamma\gamma$, $jj\gamma\gamma$ primarily consisting of mis-tagged charm jets, and $bbj\gamma$ with one fake photon, while for the event category with at least one photon in the endcap the dominant backgrounds are $bbj\gamma$ and $bbjj$ with one or two fake photons.

3.4 Signal extraction

To extract the signal and cross section, the kinematic selection requirements from Section 3.3 are applied, and a two dimensional maximum likelihood fit on the di-photon and di-bjet mass distributions is performed. Probability density functions (PDF’s) are derived for the di-photon mass, $M_{\gamma\gamma}$, and di-bjet mass, $M_{bb}$, distributions for the signal, the resonant background, and the non-resonant background by fitting the distributions from the Monte Carlo simulation samples to particular parametrization of the line-shape for $M_{\gamma\gamma}$ and $M_{bb}$. The distributions of the signal and resonant backgrounds are fitted with a Crystal Ball distribution, while for the non-resonant backgrounds are fitted to a decaying exponential. To model the degraded jet energy resolution under the HL-LHC pileup conditions, the jet energy resolution parameter is then appropriately increased.

The correlations between $M_{\gamma\gamma}$ and $M_{bb}$ are assumed to be negligible. Therefore, the two di-
Table 2: The expected event yields of the signal and background processes for 3000 fb$^{-1}$ of integrated luminosity are shown at various stages of the cut-based selection for the category with both photons in the barrel. The event yields for the category with at least one photon in the endcap are shown inside the brackets. A large fit mass window, 105 GeV to 145 GeV for $M_{\gamma\gamma}$ and 70 GeV to 200 GeV for $M_{bb}$, is used for the likelihood fit analysis described in Section 3.4.

The systematic uncertainties dominate over the statistical uncertainties with 3000 fb$^{-1}$ of integrated luminosity, the main systematic uncertainties are discussed for completeness. The photon selection efficiency systematic uncertainty is dominated by the systematic uncertainty in the efficiency of the electron veto requirement, and is less than 2% [30]. The systematic uncertainty in the b-tagging efficiency is between 2% and 3% depending on the $p_T$ and $\eta$ of the jet [26, 27].

The jet energy resolution has been measured to better than 10% [31], and the photon energy resolution is known to within approximately 15% [32]. These uncertainties are propagated to the signal and background models for $M_{bb}$ and $M_{\gamma\gamma}$ by performing toy experiments. It is found that the average bias induced on the cross section measurement are about 2% and 4% for the jet energy resolution and photon energy resolution, respectively.

The systematic uncertainty in the non-resonant background is evaluated by performing toy
MC experiments where the assumed non-resonant background model used is the product of an exponential and a fourth-degree polynomial fitted to the expected non-resonant background distribution, while an exponential function is used to perform the fit. This results in an average bias of about 12%.

Finally, the systematic uncertainties in the coupling measurement arises from the theoretical uncertainties on the double Higgs boson production cross section. Combining the uncertainties from missing higher order corrections, uncertainties on $\alpha_S$ and the PDF’s, we obtain a potential change in the cross section of about 11% for a center of mass energy of 14 TeV [12].

3.6 Upgrade scenarios

As discussed in Section 3.1, it is critical to achieve a robust reconstruction of the detector objects under the HL-LHC pileup conditions. As the measurement is primarily limited by the number of selected signal events, improving the object selection efficiency is the most important aspect. To explore the effect of the object selection efficiencies and to provide a general goal for the detector upgrade, Figure 3 shows the sensitivity as a function of the relative improvement on the photon selection efficiency and the b-tagging efficiency, respectively, over the current performance estimate under the HL-LHC pileup conditions. It is seen that the measurement can be significantly improved with even a modest improvement in photon selection efficiency or b-tagging efficiency.

Finally, the analysis sensitivity as a function of the total integrated luminosity (left) and the relative contribution of the non-resonant background (right) is shown in Figure 4.
4 \( bb\ell\ell \) final state

4.1 Object selection and performance

The signal events of interest contain two high \( p_T \) taus and two high \( p_T \) jets originating from b quarks. Di-tau final states \( \tau\mu\tau\mu \) and \( \tau_h\tau_h \) where \( h \) denotes hadronic tau decays, and \( \mu \) denotes tau decays to muons, are considered. The jets are reconstructed using the anti-\( k_T \) algorithm with the resolution parameter equal to 0.4. The jets are corrected for pileup effects using the FastJet technique \([24, 25, 27]\). Jet pileup identification is developed in Delphes using the track related and jet shape variables \([24]\). The efficiency of the jet pileup identification is 0.95 with a pileup jet rate of 0.20. The \( E_{\text{miss}} \) resolution is critical for the di-tau mass reconstruction. Without pileup mitigation the resolution is on average 50 GeV with 140 pileup interactions. With pileup identification the \( E_{\text{miss}} \) resolution in Delphes is reduced to about 25 GeV. The efficiency of selecting jets originating from b quarks and tau hadronic decays are parameterized in Delphes. To further reduce background events with light jets mimicking hadronic tau decays it is required that jets originating from hadronic tau decays contain an isolated track. With the upgraded Phase-II detector and this selection, a tau identification efficiency of about 55\% is possible while keeping miss-tag rates for light-jets to be less than 0.5\%. The working point used for b-tagging has an average efficiency of 0.68 with 0.1 and 0.01 mistag rates from c-quarks.
and light quarks respectively. Similarly, muon efficiencies and resolutions are parameterized in Delphes and efficiencies of about 95% are assumed.

4.2 Event selection and background estimation

Events are selected containing two b-tagged jets with $p_T > 30$ GeV and $|\eta| < 2.4$, and two taus with $p_T > 60$ GeV, or $p_T > 90$ GeV for the leading tau and $p_T > 45$ GeV for sub-leading tau, and $|\eta| < 2.1$ for the $\tau_h \tau_h$ di-tau final state, $p_T > 30$ GeV and $|\eta| < 2.1$ for the $\tau_h$ and $p_T > 30$ GeV and $|\eta| < 2.5$ for the $\tau_\mu$ in $\tau_\mu \tau_h$ di-tau final states.

The background samples are simulated using the techniques that were developed for Snowmass 2013 Energy Frontier for the future hadron colliders [33]. The existing samples were reconstructed using Delphes. A similar approach to the $bb\gamma\gamma$ analysis is adopted by producing a generator particle level $t\bar{t}$ sample and weighting the events by their corresponding efficiencies for selecting the constituent particles. The resolution effects of the detector are also taken into account. The main background is $t\bar{t}$ production with fully leptonic decays. Another source of large background is the Drell-Yan production of a $Z$ boson decaying into a pair of tau leptons produced in association with jets, where light jets are mis-tagged as b-jets. The important single Higgs boson backgrounds are $ZH$ and $t\bar{t}H$ processes. The remaining backgrounds considered are single top and $t\bar{t}$ produced in association with a vector boson, and di-boson processes. The QCD multi-jet background is negligible in the signal region, as verified by studying the LHC data available at $\sqrt{s} = 8$ TeV. A likelihood based mass reconstruction technique (SVFIT) is used to reconstruct the di-tau mass [34]. It is a powerful tool to discriminate against background processes containing a $Z$ boson that subsequently decays to tau leptons.

4.3 Signal extraction

Selections are applied on the di-tau mass, $M_{\tau\tau}$, and the di-b-jet mass, $M_{bb}$, distributions to identify Higgs boson decays to tau and b pairs, respectively. The requirement for $m_{bb}$ is $90$ GeV $< m_{bb} < 130$ GeV, and $110$ GeV $< m_{\tau\tau} < 140$ GeV for $M_{\tau\tau}$. The event selection is summarized in Table 3.

| Event selection | $\geq 2$ b-tagged jets with $p_T > 30$ GeV, $|\eta| < 2.4$ |
|----------------|-------------------------------------------------|
| $\tau_h \tau_h$ final state | |
| $\geq 2$ isolated $\tau_h$-s with $p_T > 60$ GeV or $p_T > 90/45$ GeV, $|\eta| < 2.1$ |
| $\tau_\mu \tau_h$ final state | |
| An isolated $\tau_h$ with $p_T > 30$ GeV, $|\eta| < 2.1$ and an isolated $\tau_\mu$ with $p_T > 30$ GeV, $|\eta| < 2.5$ | $90$ GeV $< m_{bb} < 130$ GeV and $110$ GeV $< m_{\tau\tau} < 140$ GeV |

Table 3: Event selection summary for $bb\tau\tau$ final state.

A kinematic bounding variable, $m_{T2}$, is introduced to further discriminate the dominant $t\bar{t}$ background from the di-Higgs signal [35]. By construction, $m_{T2}$ is bounded above by the top quark mass for $t\bar{t}$ background events while it is unbounded for di-Higgs boson signal events. For the $\tau_\mu \tau_h$ di-tau final states a boosted decision tree (BDT) discriminant was trained to further exploit the boosted kinematics of di-Higgs boson production. The input variables are the masses, transverse momenta, and $\Delta R$ distances of the di-tau, di-b-jet, and di-Higgs systems. The $m_{T2}$ variable is also included in the training.

Figure 5 shows the predicted distributions of the $m_{T2}$ variable (left) for the $\tau_h \tau_h$ final state and BDT discriminant variable (right) for the $\tau_h \tau_\mu$ final state after the mass window requirements.
Both variables provide good discrimination between signal and background. As expected the $t\bar{t}$ background is bounded above by the top quark mass taking into account the detector resolution effects. Tables 4 and 5 show the expected event yields for 3000 fb$^{-1}$ integrated luminosity at various stages of the event selection for the $\tau_{h}\tau_{h}$ and $\tau_{\mu}\tau_{h}$ channels, respectively.

![Figure 5: $m_{T2}$ (left) and BDT score (right) distributions in $\tau_{h}\tau_{h}$ and $\tau_{\mu}\tau_{h}$ channels, respectively. The yields are the expected SM contributions.](image)

The expected significance for the di-Higgs boson production is 0.5 and 0.7 standard deviations, for $\tau_{\mu}\tau_{h}$ and $\tau_{h}\tau_{h}$ di-tau final states, respectively. For the combination, 0.9 standard deviations are expected. The resulting expected uncertainty in the signal strength is approximately 105%. Theoretical uncertainties in the Higgs boson production are included in this result. Renormalization and factorization scale uncertainties in the di-Higgs boson signal production are 20% for NNLO calculation. The PDF uncertainty is 9%. The systematic uncertainty in the luminosity is taken to be 2.6%. Energy scale uncertainties on jets, tau leptons, and missing energy are also included. A scale uncertainty of 2% is assumed which is comparable to the corresponding uncertainty for Run1 conditions [25] in the barrel and endcap regions. The effect of the jet energy scale uncertainty in the sensitivity is about 5%.

<table>
<thead>
<tr>
<th>Selection</th>
<th>$HH$</th>
<th>$ZH$</th>
<th>$t\bar{t}H$</th>
<th>$Z \rightarrow \tau\tau$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline selection</td>
<td>$23.6 \pm 0.5$</td>
<td>$104.7 \pm 3.5$</td>
<td>$204.6 \pm 5.8$</td>
<td>$479.3 \pm 7.7$</td>
</tr>
<tr>
<td>Mass windows</td>
<td>$8.3 \pm 0.3$</td>
<td>$10.1 \pm 1.0$</td>
<td>$9.8 \pm 1.3$</td>
<td>$60.3 \pm 3.3$</td>
</tr>
<tr>
<td>Signal</td>
<td>$4.9 \pm 0.2$</td>
<td>$6.2 \pm 0.8$</td>
<td>$3.8 \pm 0.8$</td>
<td>$14.7 \pm 1.6$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Selection</th>
<th>$t\bar{t}$</th>
<th>$tW$</th>
<th>$t\bar{t}V$</th>
<th>$VV(V)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline selection</td>
<td>$7662 \pm 69$</td>
<td>$734.4 \pm 19.4$</td>
<td>$189.4 \pm 10.0$</td>
<td>$128.9 \pm 16.7$</td>
</tr>
<tr>
<td>Mass windows</td>
<td>$88.4 \pm 7.5$</td>
<td>$4.8 \pm 1.6$</td>
<td>$2.7 \pm 0.9$</td>
<td>$2.1 \pm 0.7$</td>
</tr>
<tr>
<td>Signal</td>
<td>$3.2 \pm 1.4$</td>
<td>$0.1 \pm 0.1$</td>
<td>$1.3 \pm 0.7$</td>
<td>$1.0 \pm 0.5$</td>
</tr>
</tbody>
</table>

Table 4: The expected signal yields for 3000 fb$^{-1}$ of integrated luminosity are shown at various stages of the cut-based selection in the $\tau_{h}\tau_{h}$ channel. A loose mass window cut is applied on the $\tilde{t}\bar{t}$ sample at the generation level. For the signal selection stage we require $m_{T2}$ to be greater than 180 GeV. The full $m_{T2}$ distribution is used for the signal extraction.

### 4.4 Trigger performance

The performance of the trigger system is crucial to achieve the result described above, in particular the capability to trigger on charged particles at Level-1. For the $\tau_{h}\tau_{h}$ final state, the di-tau
trigger has an offline threshold of 56 GeV on both tau legs, and the single tau trigger threshold is 88 GeV for the Level-1 sample menu described in this document. These thresholds are significantly higher without the track trigger, 95 GeV on both tau legs for the di-tau trigger and 138 GeV for the single tau trigger. Considering these less performant thresholds, the signal and background yields are reduced by about a factor of two. For the $\tau\mu\tau_h$ final state the situation is similar. The single-muon trigger threshold is 18 GeV with the track trigger and 50 GeV without the track trigger. The thresholds for the muon-tau trigger legs are significantly higher as well. Again, the signal and background yields are reduced by a factor of two by requiring 50 GeV cut on the $p_T$ of the muon and hadronic tau. Thus, in both final states the effect of the trigger performance on the sensitivity of this analysis is significant. The overall sensitivity is reduced by 40%, the equivalent of using only half of 3000 fb$^{-1}$.

### 4.5 Combination of $bb\gamma\gamma$ and $bb\tau\tau$ results

A combination of the $bb\gamma\gamma$ and $bb\tau\tau$ results is performed by performing a simultaneous fit to the generated pseudo-data. A test statistic based on the profile likelihood ratio is used. Systematic uncertainties are incorporated through nuisance parameters and are treated according to frequentist paradigm. A detailed description of the methodology is given in Refs. [36, 37]. The expected uncertainty in the signal yield is approximately 54%. The expected significance for di-Higgs boson production is 1.9 standard deviations.

### 5 $bbWW$ final state

About 1500 $bbWW$ events, where both $W$ decay leptonic, are expected at the HL-LHC, where the leptons are either muons or electrons. The dominant background process is the $t\bar{t}$ production with fully leptonic decay. Other backgrounds have negligible contribution in comparison to $t\bar{t}$ and only the dominant $t\bar{t}$ background is considered. Selected events are required to have two b-tagged jets with $p_T > 30$ GeV and two opposite-sign leptons with muon $p_T > 20$ GeV, electron $p_T > 25$ GeV, and all objects with $|\eta| < 2.5$. Additional requirements that reduce the background include a requirement on the di-lepton mass, $M_{\ell\ell} < 85$ GeV, the di-b-jet mass, $60$ GeV $< M_{bb} < 160$ GeV, the $\Delta R$ between the two leptons, $\Delta R_{\ell\ell} < 2$, the $\Delta R$ between the two b-tagged jets, $\Delta R_{bb} < 3.1$, and $\Delta \phi$ between the di-b-jet and the di-lepton systems, $\Delta \phi_{bb,\ell\ell} > 1.7$.

A neural network (NN) discriminator based on the kinematic properties of the event is trained to further reduce the background. The NN takes into account the correlations among the input

<table>
<thead>
<tr>
<th>Selection</th>
<th>$HH$</th>
<th>$ZH$</th>
<th>$t\bar{t}H$</th>
<th>$Z \rightarrow \tau\tau$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline selection</td>
<td>$39.2 \pm 0.6$</td>
<td>$213.7 \pm 4.5$</td>
<td>$1175.8 \pm 12.9$</td>
<td>$3711.7 \pm 34.1$</td>
</tr>
<tr>
<td>Mass window</td>
<td>$13.1 \pm 0.4$</td>
<td>$24.4 \pm 1.5$</td>
<td>$37.7 \pm 2.3$</td>
<td>$234.5 \pm 8.6$</td>
</tr>
<tr>
<td>Signal</td>
<td>$6.1 \pm 0.3$</td>
<td>$8.6 \pm 0.9$</td>
<td>$4.5 \pm 0.8$</td>
<td>$9.7 \pm 1.7$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Selection</th>
<th>$t\bar{t}$</th>
<th>$tW$</th>
<th>$t\bar{t}V$</th>
<th>$VV(V)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline selection</td>
<td>$(3.84 \pm 0.00) \times 10^3$</td>
<td>$(3.72 \pm 0.00) \times 10^4$</td>
<td>$4154 \pm 39$</td>
<td>$1418 \pm 90$</td>
</tr>
<tr>
<td>Mass window</td>
<td>$(2.3 \pm 0.0) \times 10^4$</td>
<td>$1232 \pm 33$</td>
<td>$119.4 \pm 7.4$</td>
<td>$49.6 \pm 2.3$</td>
</tr>
<tr>
<td>Signal</td>
<td>$63.5 \pm 7.3$</td>
<td>$29.2 \pm 4.8$</td>
<td>$3.9 \pm 1.3$</td>
<td>$2.7 \pm 0.7$</td>
</tr>
</tbody>
</table>

Table 5: The expected signal yields for 3000 fb$^{-1}$ of integrated luminosity are shown at various stages of the cut-based selection in the $\tau\mu\tau_h$ channel. A loose mass window cut is applied on the $t\bar{t}$ sample at the generation level. For the signal selection stage we require the BDT variable to be greater than 0.05. The full BDT distribution is used for the signal extraction.
variables used for the training. The signal selection is obtained by applying a threshold on the NN discriminator leading to 3875 background events and 37.1 signal events. Figure 6 shows the expected uncertainty in the di-Higgs boson signal yield as a function of the background uncertainty from 0% to 5%. The results suggest a promising contribution of this final state when combined with the other final states at the HL-LHC. The results are especially interesting for BSM scenarios.

6 Summary

Studies of the Higgs boson pair production and decays into $bb\gamma\gamma$, $bb\tau\tau$, and $bbWW$, where the $W$ bosons decay leptonically, final states are presented. The studies are performed assuming the operational conditions of the HL-LHC, with an integrated luminosity of 3000 fb$^{-1}$, and the upgraded CMS experiment. Combining the studies of $bb\gamma\gamma$ and $bb\tau\tau$ final states, the expected significance for Higgs boson pair production is 1.9 standard deviation. The resulting expected uncertainty in the signal yield is 54%. We emphasized the benefits of the CMS Phase-II upgrades to meet the challenges presented by high luminosity environment.

This summary describes the first detailed studies of Higgs boson pair production with the CMS Phase-II detector at the HL-LHC. Further improvements of the sensitivity are possible by using more sophisticated reconstruction and analysis techniques. Additional di-Higgs boson production and decay modes remain unexplored. Among these the $bbbb$ final state promises the largest potential for improvement.

References


