# CMS Physics Analysis Summary

Contact: cms-phys-conveners-ftr@cern.ch 2022/03/15

# Prospects for HH measurements in the WW$\gamma\gamma$ and $\tau\tau\gamma\gamma$ final states in proton-proton collisions at $\sqrt{s} = 14$ TeV at the High Luminosity-LHC

The CMS Collaboration


## Abstract

Prospects for nonresonant Higgs boson pair production measurements in the WW$\gamma\gamma$ and $\tau\tau\gamma\gamma$ final states in proton-proton collisions at a centre-of-mass energy of 14 TeV with the upgraded Phase-2 CMS detector at the High-Luminosity LHC (HL-LHC) are presented. The study is performed using generated events that passed DELPHES fast simulation of the CMS detector with an average of 200 proton-proton interactions per bunch crossing and an integrated luminosity of 3000 fb$^{-1}$.

# 1 Introduction

Since the discovery of a Higgs boson on July 4, 2012 by the CMS and ATLAS experiments at the CERN LHC [1–3], physicists have sought to deepen their understanding of the Higgs boson and its properties. Experimentally, measuring the Higgs trilinear coupling ($\lambda_{HHH}$) is a crucial test of the electroweak symmetry breaking mechanism. Among all possible production modes of Higgs boson pair final states at the LHC, gluon fusion ($ggF$) $gg \rightarrow HH$ is the dominant production mode. The interfering Feynman diagrams contributing to the gluon fusion HH production in pp collisions at leading order (LO) are shown in Fig. 1.

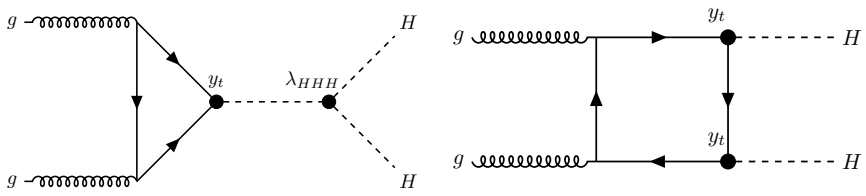



Figure 1: Feynman diagrams for leading order Higgs boson pair production via gluon fusion.

In this analysis, a di-Higgs search in the WW$\gamma\gamma$ and $\tau\tau\gamma\gamma$ channels is performed benefitting from the sensitive $H \rightarrow \gamma\gamma$ process which provides a clean and distinguishable signature. This is the first study providing the expected significance numbers in these channels at a centre-of-mass energy of 14 TeV with integrated luminosity of 3000 fb$^{-1}$ at the HL-LHC using events simulated with the DELPHES [4] package.

In the WW$\gamma\gamma$ process, three final states are possible, as the W boson can decay both leptonically and hadronically: Semi-Leptonic ($WW \rightarrow qql\nu$), Fully-Leptonic ($WW \rightarrow l\nu l\nu$) and Fully-Hadronic ($WW \rightarrow qqqq$) decay modes, where $l = e, \mu$ or $\tau$. For the Fully-Hadronic decay mode, QCD induced processes are a major background and given the limited size of the Monte Carlo samples, this background is not well described. Hence, this final state is not considered in this study. The other two final states are tagged with separate selections and categorization, and their corresponding signal and background estimates are combined with these from the $\tau\tau\gamma\gamma$ channel to improve the overall analysis sensitivity.

# 2 The Phase-2 CMS detector

The CMS detector [5] will be substantially upgraded in order to fully exploit the physics potential offered by the increase in luminosity, and to cope with the demanding operational conditions at the HL-LHC [6–10]. The upgrade of the first level hardware trigger (L1) will allow for an increase of L1 rate and latency to about 750 kHz and 12.5 $\mu$s, respectively, and the high-level software trigger (HLT) is expected to reduce the rate by about a factor of 100 to 7.5 kHz. The entire pixel and strip tracker detectors will be replaced to increase the granularity, reduce the material budget in the tracking volume, improve the radiation hardness, and extend the geometrical coverage and provide efficient tracking up to pseudorapidities of about $|\eta| = 4$. The muon system will be enhanced by upgrading the electronics of the existing cathode strip chambers (CSC), resistive plate chambers (RPC) and drift tubes (DT). New muon detectors based on improved RPC and gas electron multiplier (GEM) technologies will be installed to add redundancy, increase the geometrical coverage up to about $|\eta| = 2.8$, and improve the trigger and reconstruction performance in the forward region. The barrel electromagnetic calorimeter (ECAL) will feature the upgraded front-end electronics that will be able to exploit the information from single crystals at the L1 trigger level, to accommodate trigger latency and bandwidth requirements, and to provide 160 MHz sampling allowing high precision timing

capability for photons. The hadronic calorimeter (HCAL), consisting in the barrel region of brass absorber plates and plastic scintillator layers, will be read out by silicon photomultipliers (SiPMs). The endcap electromagnetic and hadron calorimeters will be replaced with a new combined sampling calorimeter (HGCal) that will provide highly-segmented spatial information in both transverse and longitudinal directions, as well as high-precision timing information. Finally, the addition of a new timing detector for minimum ionizing particles (MTD) in both barrel and endcap regions is envisaged to provide the capability for 4-dimensional reconstruction of interaction vertices that will significantly offset the CMS performance degradation due to high PU rates.

A detailed overview of the CMS detector upgrade program is presented in Ref. [6–13], while the expected performance of the reconstruction algorithms and pile-up mitigation with the CMS detector is summarised in Ref. [14].

## 3 Simulated samples

Signal Monte Carlo $gg \rightarrow HH$ samples are generated using POWHEG V2 [15–18] at next-to-leading order (NLO) in QCD including the full top quark mass dependence with the SM parameters. Subsequent decays of the Higgs boson pairs into WW or $\tau\tau$ and a pair of photons are implemented using PYTHIA 8.212 [19]. The signal samples are generated separately for the three possible final states in $WW\gamma\gamma$. For $\tau\tau\gamma\gamma$ signal samples, all possible decays for taus are allowed.

The analysis is affected by backgrounds from single Higgs boson production and by non-resonant backgrounds with continuum $m_{\gamma\gamma}$ spectra. The event generator MADGRAPH5_AMCATNLO [20, 21] with the FxFx merging scheme [22] was used for the generation of the background from SM single Higgs boson production, including gluon-gluon fusion (ggH), vector-boson fusion (VBFH), associated production with a Z or W boson (VH) and associated production with a top quark pair (ttH) , while the top quark associated production with a Higgs boson (tHq) was done using MADGRAPH version-2.7 at LO.

The continuum background contribution comes from various SM processes. Most of the dominant backgrounds across all the final states are due to the $\gamma\gamma$+jets processes that are modeled with the SHERPA v.2.2.1 generator [23]. $\gamma$+jets, QCD-induced processes and WW processes are modeled with the PYTHIA 8 [19] generator. Drell Yan and W production processes in association with photons and jets are modeled using MADGRAPH5 version-2.7 at LO. While $t\bar{t}$ are generated using POWHEG v2, $t\bar{t}W$, $t\bar{t}\gamma$, $t\bar{t}\gamma\gamma$, $Z\gamma$ are modeled using MADGRAPH5_AMCATNLO [20–22].

All signal and background samples are simulated with the Phase-2 upgraded CMS detector geometry using DELPHES fast simulation with average pile-up of 200 interactions and at $\sqrt{s} =$ 14 TeV.

## 4 Object Selections

Photons used in this analysis are required to have a transverse momentum ($p_T$) above 25 GeV with the leading photon above 35 GeV within $|\eta| < 2.5$ excluding the ECAL barrel and endcap transition region ($1.442 < |\eta| < 1.556$). The relative isolation of the photon candidate, defined as sum of the $p_T$ of all the particles within a cone of size 0.3 around the photon and divided the sum by the photon $p_T$, is required to be less than 0.3 and to pass the loose identification criteria corresponding to 90% signal efficiency.

Electrons are required to have $p_T$ above 10 GeV within $|\eta| < 2.5$ excluding the ECAL transition region and must be isolated from the photons with an angular separation in the $\eta - \phi$ plane greater than 0.4 ($\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} > 0.4$). The transverse momenta of the muons are required to be above 10 GeV and within $|\eta| < 2.5$ , and they are required to be isolated from the photons and electrons with an angular separation greater than 0.4. Hadronically decaying taus are required to have $p_T > 20$ GeV within $|\eta| < 2.5$, and are required to be separated from photons, electrons and muons with an angular separation greater than 0.2. The relative isolation of the electrons (muons) is required to be less than 0.3 (0.1). Jets are reconstructed using the anti-$k_T$ clustering method with a distance parameter of 0.4. They are required to have $p_T > 30$ GeV, be within $|\eta| < 5$ and to be well separated from the photon and lepton candidates with an angular separation greater than 0.4. Jets are b-tagged using a deep neural network (DNN) based secondary vertex algorithm, DEEPJET [24, 25].

# 5 Event Selections and Categorization

All events are required to have exactly two photons whose invariant mass verifies $100 < m_{\gamma\gamma} < 180$ GeV, making this the common pre-selection to all final states. The analysis is performed in mutually exclusive final states targeting decays of the vector bosons referred to as 1L and 2L final states for $WW\gamma\gamma$ and 1 $\tau$ or 2 $\tau$s final states for $\tau\tau\gamma\gamma$. Here, lepton (L) refers to either an $e^{\pm}$ or a $\mu^{\pm}$.

## 5.1 One Lepton final state

Events fall into this analysis category if they contain at least one pre-selected diphoton pair, and contain exactly one electron or muon passing the selection criteria described above. This final state is expected to be the most sensitive of the three $WW\gamma\gamma$ channels due to the combination of a relatively large $W \to qq$ branching ratio, and the presence of a high energy lepton from the $W \to l\nu$ decay. Here, two multiclass DNNs are trained to separate the di-Higgs signal from the expected single Higgs boson and continuum background where the di-Higgs (HH $\to 2\gamma 2q l\nu$, HH $\to 2\gamma 2\tau$, HH $\to 2\gamma 2l2\nu$) processes are labelled as $HH$, single Higgs backgrounds as $H$ and all the other background samples as the continuum background. Event weights are reweighted such that the learning weight in all classes is similar. This ensures that the network focuses on categorising all classes with an equal importance. Events are classified in "even" and "odd" sets based on the fifth decimal place of leading photon $\phi$. This separation on data sets allows to apply the training performed with even events on the odd data set, and vice versa, to avoid any training bias.

The feature variables used as inputs for the training include the kinematic variables such as $p_T$, $\eta$, $\phi$ and energy values of photons, jets, electrons and muons. For photons, $p_T$ and energy values reweighted by the diphoton mass are used. Additionally, the jet multiplicity, missing transverse energy and the invariant mass of the leading and subleading jets are also utilised in the training.

The DNN score distribution is shown in Figure 2a. Events are then partitioned in four categories making use of the $HH$ node output score. The boundaries shown in Table 1 are chosen such that the expected significance is maximised.

The categorisation allows to deal with different signal over background ratios. Category 4 presents the best signal purity and significance.

| Categories | Definition |
|---|---|
| Category 1 | $0.1 <$ DNN score $< 0.6$ |
| Category 2 | $0.6 <$ DNN score $< 0.8$ |
| Category 3 | $0.8 <$ DNN score $< 0.92$ |
| Category 4 | DNN score $> 0.92$ |

Table 1: One Lepton final state DNN score categories.

## 5.2 Two Lepton final state

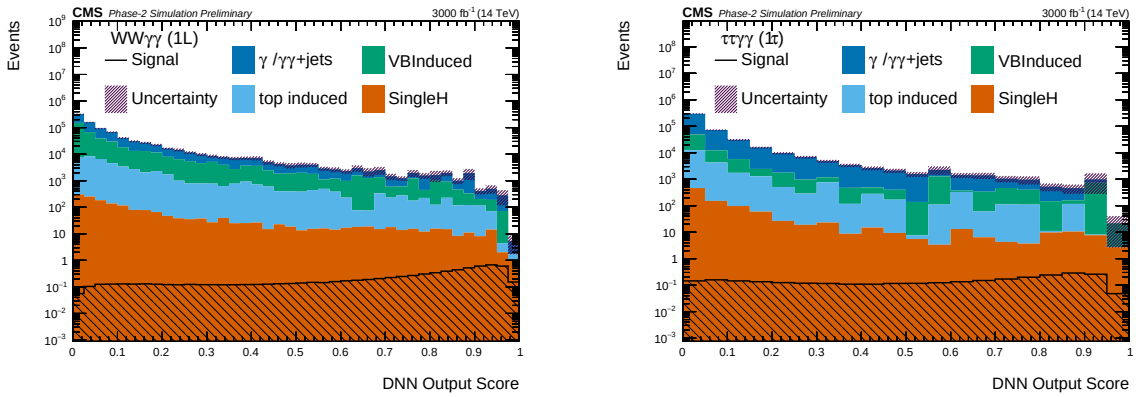
A cut-based analysis is performed here. In this final state, events are required to have at least two oppositely charged leptons ($e^+e^-$, $\mu^+\mu^-$, $e^\pm\mu^\mp$) passing the electron and muon object selections described in Section 4. Furthermore, the events are required to satisfy the selections listed in Table 2, where $\Delta R(l,l)$ is the $\Delta R$ between two leptons, $m_{ll}$ is the mass of dilepton system and $m_{e\gamma}$ is the invariant mass of the leading electron and the leading photon in the events that have at least one electron.

| Variable | Selection |
|---|---|
| $\Delta R(l,l)$ | $> 0.4$ |
| $p_{\mathrm{T}}$ of leading lepton | $> 20\,\mathrm{GeV}$ |
| $p_{\mathrm{T}}$ of subleading lepton | $> 10\,\mathrm{GeV}$ |
| $E_T^{miss}$ | $> 20\,\mathrm{GeV}$ |
| $p_T^{\gamma\gamma}$ | $> 91\,\mathrm{GeV}$ |
| $m_{ll}$ | $< 80\,\mathrm{GeV}$ *or* $> 100\,\mathrm{GeV}$ |
| number of medium-tagged b-jets | $= 0$ |
| $\lvert m_{e\gamma} - m_z \rvert$ | $> 5\,\mathrm{GeV}$ |

Table 2: Selection criteria of the Two Lepton Channel.

## 5.3 One Tau lepton final state

Events fall in this category if they contain two photons within the required invariant mass range and contain exactly one hadronically decaying tau lepton and no electron/muon. Here also, two multiclass deep neural networks (DNNs) are trained following the same strategy as in the one lepton final state. Taus kinematic variables are replacing the ones from the electron or the muon.




(a) One Lepton final state

(b) One tau final state

Figure 2: DNN output scores in the one lepton (left) and one tau (right) final states.

The DNN score distribution is shown in Figure 2b. Here, events are partitioned in two categories making use of the DNN score. Category one corresponds to events where the DNN score lies between 0.1 and 0.65, while events with a DNN score higher than 0.65 are categorized in Category 2.

### 5.4 Two Tau leptons final state

A cut-based analysis is performed in this final state. The events end up in this category if they have at least two taus and no lepton. The two taus are required to be oppositely charged.

## 6 Systematic uncertainties

The contribution of systematic uncertainties have been divided in experimental and theoretical uncertainties. The experimental uncertainties have been added as listed in Table 3. These uncertainties have been applied according to the Yellow Report recommendation described in Ref.[26]. Theoretical uncertainties are added on the ggHH signal and single Higgs boson processes as described in Table 4.

| Uncertainty Source | Input (%) |
|---|---|
| Luminosity | 1 |
| Diphoton trigger | 2 |
| $m_{\gamma\gamma}$ resolution | 5 |
| PhotonID | 0.5/photon |
| electronID | 0.5/electron |
| muonID | 0.5/muon |
| tauID | 2.5/tau |
| Tau energy scale | 3 |
| Jet energy Scale | 3 |
| b-tagging veto | 3 |

Table 3: Experimental uncertainties considered in this study.

| Process | Uncertainty Source | | |
|---|---|---|---|
| | PDF $+\alpha_s$ (%) | QCD Scale (%) | $m_{top}$ (%) |
| ggHH | $\pm$ 3 | +2.1/-4.9 | +4.0/-18 |
| ggH | +4.6/-6.7 | $\pm$ 3.2 | - |
| VBFH | +0.5/-0.3 | $\pm$ 2.1 | - |
| VH | +0.4/-0.7 | $\pm$ 1.8 | - |
| ttH | +6/-9.2 | $\pm$ 3.5 | - |
| tHq | +6.4/-14.7 | $\pm$ 3.6 | - |

Table 4: Theoretical uncertainties considered on the ggHH signal and single Higgs processes.

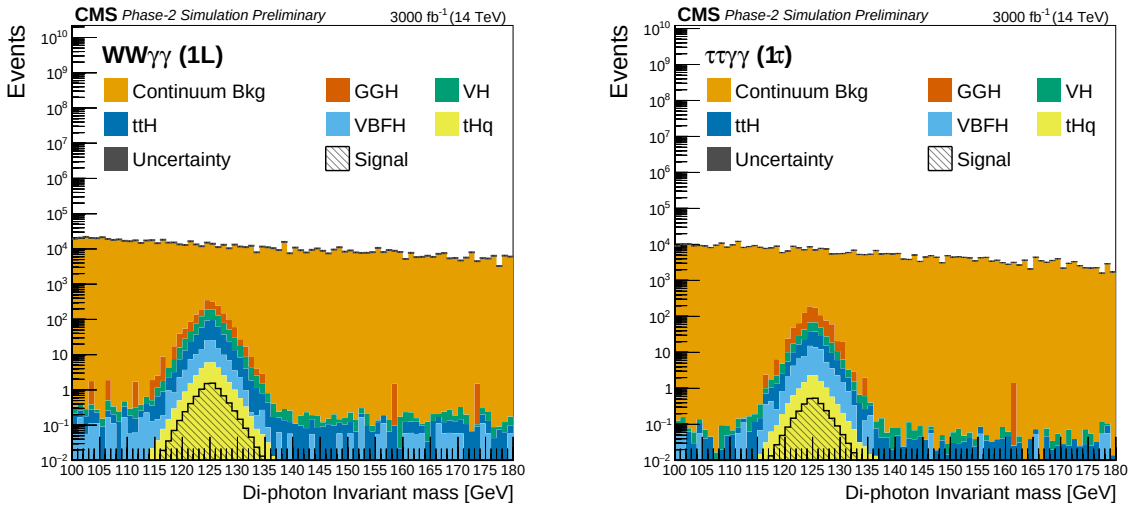
## 7 Results

The two photons invariant mass distributions for the one-leptonic and one tau final states, inclusive of all categories, are shown in Figure 3. The expected significance is extracted by fitting the $m_{\gamma\gamma}$ distributions in all the categories using a binned maximum likelihood approach with all systematic uncertainties treated as nuisance parameters with log-normal distributions using the Higgs Combine tool [27]. Given the presence of high fluctuations in the $m_{\gamma\gamma}$ distribution

for the continuum background across different categories, the shape fitted with an exponential function has been used to describe the continuum background. The correlations among different sources of uncertainties are taken into account while the different final states are considered as independent channels in the fit.

For illustration purposes, the expected Phase-2 $m_{\gamma\gamma}$ distributions are shown in Figure 4 for the 1L final state in its most sensitive category and in the 2 $\tau$ final state, where signal along with the single Higgs and continuum background is described using a Gaussian and an exponential function. The (pseudo-)data are generated according to the fitted signal, single Higgs and continuum background contributions.

The significance values obtained are shown in Table 5 for the WW$\gamma\gamma$ and $\tau\tau\gamma\gamma$ final states along with their combination.




(a) One lepton final state

(b) One tau final state

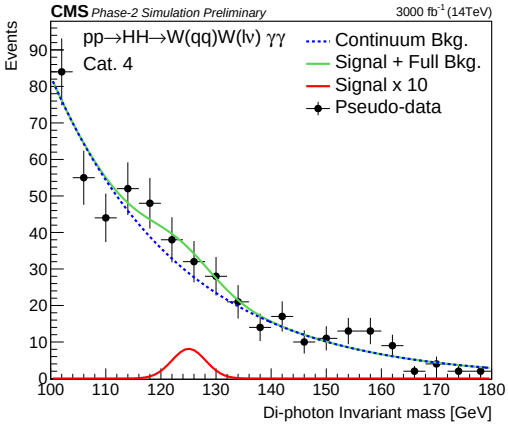
Figure 3: $m_{\gamma\gamma}$ distributions in the $WW\gamma\gamma$, 1L (left) and $\tau\tau\gamma\gamma$, 1$\tau$ (right) final states.

Table 5: Expected HL-LHC significances ($\sigma$) results of the WW$\gamma\gamma$ and $\tau\tau\gamma\gamma$ processes with their combination.

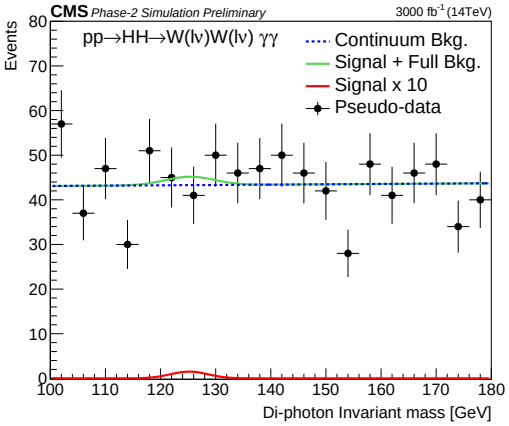
| Final State | Significance (stat+exp+theory) |
|---|---|
| $WW\gamma\gamma$ | 0.21 |
| $\tau\tau\gamma\gamma$ | 0.08 |
| Combination | 0.22 |

## 8 Summary

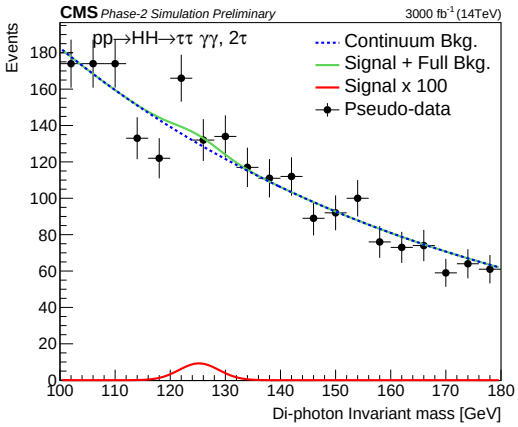
A measurement of non-resonant Higgs boson pair production in the WW$\gamma\gamma$ and $\tau\tau\gamma\gamma$ final states was performed in simulated proton-proton collision events at a centre-of-mass energy of 14 TeV and integrated luminosity of 3000 fb$^{-1}$ at the HL-LHC. Combining all final state categories and including systematic uncertainties, an expected significance 0.22$\sigma$ for signal is measured.




(a) One lepton final state, Category 4




(b) Two leptons final state




(c) Two taus final state

Figure 4: $m_{\gamma\gamma}$ distributions in the $WW\gamma\gamma$, 1L (top left), 2L (top right) and $\tau\tau\gamma\gamma$, 2$\tau$s (bottom) final states.

## Acknowledgments

## References


[1] ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Phys. Lett. B* **716** (2012) 1, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.

[2] CMS Collaboration, "Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC", *Phys. Lett. B* **716** (2012) 30, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.

[3] CMS Collaboration, "Observation of a New Boson with Mass Near 125 GeV in $pp$ Collisions at $\sqrt{s} = 7$ and 8 TeV", *JHEP* **06** (2013) 081, doi:10.1007/JHEP06(2013)081, arXiv:1303.4571.

[4] DELPHES 3 Collaboration, "DELPHES 3, A modular framework for fast simulation of a generic collider experiment", *JHEP* **02** (2014) 057, doi:10.1007/JHEP02(2014)057, arXiv:1307.6346.

[5] CMS Collaboration, "The CMS Experiment at the CERN LHC", *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[6] CMS Collaboration, "Technical Proposal for the Phase-II Upgrade of the CMS Detector", CMS Technical Proposal CERN-LHCC-2015-010. LHCC-P-008. CMS-TDR-15-02, 2015.

[7] CMS Collaboration, "The Phase-2 Upgrade of the CMS Tracker", CMS Technical Design Report CERN-LHCC-2017-009. CMS-TDR-014, 2017.

[8] CMS Collaboration, "The Phase-2 Upgrade of the CMS Barrel Calorimeters Technical Design Report", CMS Technical Design Report CERN-LHCC-2017-011. CMS-TDR-015, 2017.

[9] CMS Collaboration, "The Phase-2 Upgrade of the CMS Endcap Calorimeter", CMS Technical Design Report CERN-LHCC-2017-023. CMS-TDR-019, 2017.

[10] CMS Collaboration, "The Phase-2 Upgrade of the CMS Muon Detectors", CMS Technical Design Report CERN-LHCC-2017-012. CMS-TDR-016, 2017.

[11] CMS Collaboration, "A MIP Timing Detector for the CMS Phase-2 Upgrade", technical report, CERN, Geneva, Mar, 2019.

[12] CMS Collaboration, "The Phase-2 Upgrade of the CMS Level-1 Trigger", technical report, CERN, Geneva, Apr, 2020. Final version.

[13] CMS Collaboration, "The Phase-2 Upgrade of the CMS Data Acquisition and High Level Trigger", technical report, CERN, Geneva, Mar, 2021. This is a temporary submission, just to get the LHCC reference number, to be used in this and other CMS documents.

[14] CMS Collaboration, "Expected performance of the physics objects with the upgraded CMS detector at the HL-LHC", Technical Report CMS-NOTE-2018-006. CERN-CMS-NOTE-2018-006, 2018.

[15] P. Nason, "A New method for combining NLO QCD with shower Monte Carlo algorithms", *JHEP* **11** (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.

[16] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method", *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.

[17] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP* **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.

[18] G. Heinrich et al., "Probing the trilinear Higgs boson coupling in di-Higgs production at NLO QCD including parton shower effects", *JHEP* **06** (2019) 066, doi:10.1007/JHEP06(2019)066, arXiv:1903.08137.

[19] T. Sjostrand et al., "An Introduction to PYTHIA 8.2", *Comput. Phys. Commun.* **191** (2015) 159, doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012.

[20] J. Alwall et al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", *JHEP* **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.

[21] P. Artoisenet, R. Frederix, O. Mattelaer, and R. Rietkerk, "Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations", *JHEP* **03** (2013) 015, doi:10.1007/JHEP03(2013)015, arXiv:1212.3460.

[22] R. Frederix and S. Frixione, "Merging meets matching in MC@NLO", *JHEP* **12** (2012) 061, doi:10.1007/JHEP12(2012)061, arXiv:1209.6215.

[23] Sherpa Collaboration, "Event Generation with Sherpa 2.2", *SciPost Phys.* **7** (2019), no. 3, 034, doi:10.21468/SciPostPhys.7.3.034, arXiv:1905.09127.

[24] CMS Collaboration, "Performance of the DeepJet b tagging algorithm using 41.9/fb of data from proton-proton collisions at 13TeV with Phase 1 CMS detector", Technical Report CMS-DP-2018-058, Nov, 2018.

[25] E. Bols et al., "Jet Flavour Classification Using DeepJet", *JINST* **15** (2020), no. 12, P12012, doi:10.1088/1748-0221/15/12/P12012, arXiv:2008.10519.

[26] A. Dainese et al., "Report on the Physics at the HL-LHC, and Perspectives for the HE-LHC", technical report, Geneva, Switzerland, 2019. doi:10.23731/CYRM-2019-007.

[27] The ATLAS Collaboration, The CMS Collaboration, The LHC Higgs Combination Group Collaboration, "Procedure for the LHC Higgs boson search combination in Summer 2011", technical report, CERN, Geneva, Aug, 2011.