



UNIVERSITÀ DEGLI STUDI DI BARI ALDO MORO

UNIVERSITÀ DEGLI STUDI "ALDO MORO" DI BARI
Dipartimento Interateneo di Fisica "M. Merlin"

DOTTORATO DI RICERCA IN FISICA
Ciclo XXXIV
Settore scientifico disciplinare FIS/01

Dottorando:
Ilirjan Arben Margjeka

Search for a HH production in the $b\bar{b}ZZ(4l)$ final state at
 $\sqrt{s} = 13\text{TeV}$ with full Run II data.

Supervisor: Prof. Nicola De Filippis, Prof. Gabriella Pugliese

Coordinatore: Ch.mo Prof. Giuseppe Gonnella



UniBa
Università degli Studi "Aldo Moro" di Bari
University of Bari "Aldo Moro"

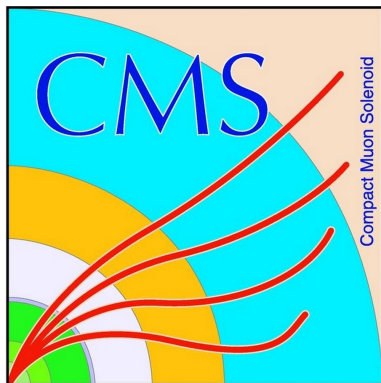
Thesis for the Degree of Philosophiae Doctor
Faculty of Science
Department of Physics

Submitted 26/11/2021;
Approved 10/01/2022;
Date of defense 04/03/2022;

Supervisors
Supervisor: Prof. Nicola De Filippis & Prof. Gabriella Pugliese

Assessment Committee

- 1st opponent
Prof. Dr. 1st opponent,
Name of the institute in the original language
Name of the institute in English
City, Country
- 2nd opponent
Prof. Dr. 2nd opponent,
Name of the institute in the original language
Name of the institute in English
City, Country
- Additional members of the committee and administrator
Prof. committee



Search for a HH production in the $b\bar{b}ZZ(4l)$ final state at $\sqrt{s} = 13\text{TeV}$ with full Run II data.

Introduction

Chapter 1 © The double Higgs (H) production

Chapter 2 © The Large Hadron Collider Experimental Apparatus

Chapter 3 © Graphite resistivity studies on Phase 2 improved RPC detectors

Chapter 4 © Monte Carlo simulation

Chapter 5 © Analysis Workflow

Chapter 6 © Results

Conclusions

Doctoral Theses at the University of Bari, 04/03/2022

Dedicated to each of my beloved family members, who sacrificed everything for my academic and personal formation.

And a special dedication goes to my father and my best friend at the same time, the one who was and is my idol and the best example to be followed.

My dear father, I owe you everything.

"We cannot solve our problems with the same thinking we used
when we created them." - Albert Einstein



"Tell me and I forget. Teach me and I remember. Involve me and I
learn. " - Benjamin Franklin



"Life is really simple, but we insist on making it complicated. " -
Confucius



"I have not failed. I've successfully discovered 10.000 things that
won't work. " - Thomas Edison

Abstract in English . Sommario in Italiano

Abstract in English

A search for resonant and non-resonant double Higgs boson production at the LHC with the CMS experiment is presented. The extension of the HEFT theory for the non-resonant HH production and Kaluza-Klein Bulk Graviton with spin-2 and Radion with spin-0 are the resonances responsible for the Higgs (H) pair production. For both cases, one of the Higgs boson will be studied into a leptonic final state ($H \rightarrow ZZ(4l)$, where, $l = e, \mu$), meanwhile the other one decays into a pair of b -quarks ($b\bar{b}$). The analyzed data corresponds to an integrated luminosity of $137.2fb^{-1}$ of pp -collision at a center-of-mass-energy of $\sqrt{s} = 13$ TeV collected with the CMS detector during the full RunII period. Advanced statistical techniques are employed to extract the results. Upper limits are set on the signal strength modifier μ , defined as the double-Higgs boson rate in the $b\bar{b}4l$ channel to the Standard Model (SM) expectation.

Sommario in Italiano

In questa tesi viene presentata la ricerca sulla produzione di doppio Higgs con il sperimento CMS dell'LHC al CERN.. L'estensione della teoria di campo efficace per la produzione non resonante della coppia del bosone di Higgs così come possibili risonanze massive del gravitone tipo bulk con spin-2 oppure il radione con spin-0 della teoria di Kaluza-Klein e dal modello di Randall-Sundrum, potrebbero essere responsabili per la produzione della coppia dei bosoni di Higgs. Per tutti e due i casi, verrà applicata una analisi sulla coppia di bosoni di Higgs con stato finale di decadimento di due jet b -taggati e quattro leptoni di tipo $l = e, \mu$. I dati sperimentali analizzati ammontano a una luminosità totale di $137.2fb^{-1}$ dalle collisioni tra protoni a energie di centro-di-massa pari a $\sqrt{s} = 13$ TeV. Una analisi statistica avanzata è stata applicata per i risultati finali. Limiti superiori sono stati imposti sulla significanza del segnale $\mu = \sigma_{HH}/\sigma_{HH}^{SM}$, definito come il rapporto del risultato finale ottenuto dall'analisi sul canale $b\bar{b}4l$ sulla base del Modello Standard (SM).

Acknowledgements

I would like to thank a lot my supervisor Prof. Nicola De Filippis and specially my colleague PhD Walaa Elemetenawee for the continues help and support on the Analysis and technical advises, in order to achieve important results on the non-resonant and resonant analysis, where the non-resonant HH analysis was also approved from the CMS Collaboration.

An important thanks goes to Prof. G. Corcella of INFN-LNF (Rome), who gave to our group an important support on the understanding of MC generation and theoretical uncertainties' calculations.

I also would like to add a big thanks to my supervisor Prof. Gabriella Pugliese and the CMS RPC group at B. 904 for the thecnical support on the developement and testing on the new iRPC and read-out electronics, a very important study towards the HL-LHC ugrade of the RPC system.

Of course, It is important to mention also the huge support I've got from the academic personal of the Department of Physics of my home University of Tirana (Albania), where I would like to remind my mentors: Prof. Dr. Artan Boriçi, Prof. Dr. Mimoza Hafizi and Prof. Dr. Sonila Boçi.

Contents

Abstract in English vii

Astratto vii

Astratto vii

Acknowledgement ix

Introduction xxvii

Part I Double Higgs production 1

Chapter 1 The double Higgs (H) production 3

1.1 The standard model (SM) of particle physics. 4

1.2 Structure of the standard model (SM) 5

1.3 The Brout-Englert-Higgs mechanism 9

 The Brout-Englert-Higgs Lagrangian 11

 1.3.1 The experimental observation of the single Higgs (H) boson event at the LHC 13

1.4 The double Higgs production 15

 1.4.1 The SM non-resonant production mechanism of the HH 17

 The EFT Lagrangian for HH processes 17

 1.4.2 Invariant mass $M_{inv}(HH)$ of the non-resonant double Higgs production 19

 1.4.3 Top Quark Mass Uncertainties 20

 1.4.4 Non-resonant HH production mode beyond the SM 22

1.5 Resonant HH production 23

 1.5.1 Higgs Singlet Model 23

 1.5.2 2-Higgs-Doublet Model (2HDM) 24

 1.5.3 Minimal supersymmetric extension of the SM (MSSM) 25

 1.5.4 Kaluza-Klein Bulk Graviton spin-2 and Radion spin-0 26

 1.5.5 Gravity particles 28

Part II The CMS Experiment at the LHC (CERN) 33

Chapter 2 The Large Hadron Collider Experimental Apparatus 35

- 2.1 The Large Hadron Collider (LHC) at CERN 37
- 2.2 LHC experiments 40
- 2.3 The Compact Muon Solenoid (CMS) Experiment 41
- 2.4 The CMS Solenoid Magnet 44
- 2.5 The CMS Tracker 44
- 2.6 The CMS Calorimeter 45
 - 2.6.1 The Electromagnetic Calorimeter (ECAL) 45
 - 2.6.2 The Hadronic Calorimeter (HCAL) 46
- 2.7 The Muon Spectrometer 47
 - 2.7.1 The Drift Tubes 48
 - 2.7.2 The Cathode Strip Chambers 48
 - 2.7.3 Resistive Plate Chambers 49
- 2.8 The Trigger System 49
- 2.9 The CMS upgrade for HL-LHC 50

Part III The Resistive Plate Chambers 55

Chapter 3 Graphite resistivity studies on Phase 2 improved RPC detectors 57

- 3.1 The RPC chamber: Working principle 57
 - 3.1.1 The RPC system upgrade for HL-LHC 59
- 3.2 The Performance Studies of the double coated 1/2 size i-RPC 61
- 3.3 Performance studies of the two resistivity region of the improved RPC module 62
 - 3.3.1 Fast timing pre-amplifiers 65
 - 3.3.2 The custom ASIC discriminators FEBs 67
- 3.4 Conclusions 68

Part IV $HH \rightarrow b\bar{b}ZZ(4l)$ Analysis 71

Chapter 4 Monte Carlo simulation 73

- 4.1 MC generations studies related to the Monte Carlo generators di-Higgs (H) signal models 73
 - 4.1.1 Non-resonant HH MC generator models 74
 - 4.1.2 Comparisons 75
 - 4.1.3 Theoretical uncertainties 77
 - 4.1.4 The k_λ -factor scan for the λ_{hhh} at 13TeV 80
 - 4.1.5 The resonant HH production 82
- 4.2 Simulations 84
 - 4.2.1 Simulation of the non-resonant double Higgs (H) 2016, 2017 and 2018 Signal Samples 85
 - 4.2.2 Simulation of the resonant KK-Bulk Graviton spin-2 and Radion spin-0 2016, 2017 and 2018 Signal Samples 87
 - 4.2.3 Background samples 90
 - 4.2.4 The Pile-Up reweighting and MC weights 90
- 4.3 Trigger selection and real data samples 93
 - 4.3.1 The 2016, 2017 and 2018 datasets 93

Chapter 5 Analysis Workflow 99

- 5.1 Physics Objects Selection 100
 - 5.1.1 Particle Flow Algorithm 100
 - 5.1.2 Tracking and Clustering 100
- 5.2 ELECTRONS $e^\pm$ 101
- 5.3 MUONS $\mu^\pm$ 104
- 5.4 PHOTONS γ 106
- 5.5 JETS 106
- 5.6 Analysis Strategy 107
 - 5.6.1 Building of the $H \rightarrow ZZ \rightarrow 4l$ System 108
 - 5.6.2 Building the $H \rightarrow b\bar{b}$ system 110
 - 5.6.3 Signal region 112
- 5.7 Background Estimation 113
 - 5.7.1 Irreducible Backgrounds 113
 - 5.7.2 Reducible Backgrounds 114
 - 5.7.3 The Experimental Uncertainties 118

- 5.7.4 The Theoretical Uncertainties 118
- 5.8 Multivariate Analysis 121
 - 5.8.1 The Boosted Decision Trees (BDT) 121

Part V Results 141

Chapter 6 Results 143

- 6.1 Signal yields 143
 - 6.1.1 Non-resonant $HH \rightarrow b\bar{b}ZZ(4l)$ signal region yields 144
 - 6.1.2 Resonant Bulk Graviton spin-2 and Radion spin-0 $HH \rightarrow b\bar{b}ZZ(4l)$ signal region yields 145
- 6.2 Statistical analysis 147
 - 6.2.1 Likelihood function $\mathcal{L}$ and Nuisance parameter θ 148
 - 6.2.2 Quantifying an excess 149
 - 6.2.3 Setting an upper limit 149
 - 6.2.4 Non-resonant SM HH limits for full Run II 150
 - 6.2.5 BSM considerations on the SM HH signal: THE k_λ -SCAN limits for full Run II 150
 - 6.2.6 Resonant HH : limits for full Run II 152
- 6.3 Projected HL-LHC reach for the HH studies 155
 - 6.3.1 Projected non-resonant HH studies for the HL-LHC 155
 - 6.3.2 Thoughts on resonant HH studies for the HL-LHC 158

Part VI Summary and Conclusions 161

Introduction 163

Appendix A Supplementary information for Chapter 1 A-1

- A.1 Effective Field Theory and QCD corrections for ggF-HH production A-1

Appendix B Supplementary information for chapter 5 A-9

- B.1 Tables and Plots related to the Analysis Results A-9
- B.2 Non-resonant HH A-9
- B.3 Resonant HH A-10

List of Figures

Chapter 0

Chapter 1

- 1.1 Standard model of elementary particles and their Physics properties as invariant mass, spin and electrical charge: 1) Top left in violet: quarks; 2) Bottom left in green: fermions; 3) The brown column: gauge bosons; 4) top right: The scalar boson, the Higgs particle 4
- 1.2 . Representation of the Higgs potential: The top of the "mexican hat" represents the vacuum state, describing the symmetry of the Higgs potential, while the ground state is not symmetric. In the Higgs mechanism, the rotational degree of freedom along the rim becomes the longitudinal polarization of the $W^\pm$ and Z^0 bosons, which thereby acquire mass. This image is adapted from Ref. [105] 10
- 1.3 Four-lepton invariant mass (left)[146] and photon pair invariant mass (right)[121] distributions in events collected at $\sqrt{s} = 13\text{TeV}$ 13
- 1.4 Left: Total production cross sections for the single Higgs (H) within the SM via gluon fusion, vector-boson fusion, double Higgs-strahlung and double Higgs bremsstrahlung off top quarks. The size of the bands shows the total uncertainties originating from the scale dependence and the $PDF + \alpha_s$ uncertainties [66]. Right: Standard Model Higgs boson decay branching ratios (CERN Yellow Report 4) [36] 14
- 1.5 Normalized Higgs boson coupling constant as a function of the boson or fermion mass: Best fit values as a function of particle mass for the combination of ATLAS and CMS data, with parameters defined as $k_F \hat{m}_F/v$ for the fermions, and as $\sqrt{k_V} \hat{m}_V/v$ for the weak vector bosons, where $v = 246\text{ GeV}$ is the vacuum expectation value of the Higgs field. The dashed (blue) line indicates the predicted dependence on the particle mass in the case of the SM Higgs boson. The solid (red) line indicates the best fit result to the $[M, \epsilon]$ phenomenological model [120] with the corresponding 68% and 95% CL bands. 15
- 1.6 The Feynman diagrams of the HH production via gluon fusion at LO, where the BSM couplings are highlighted in red. 17
- 1.7 Total production cross sections for Higgs pairs within the SM via gluon fusion, vector-boson fusion, double Higgs-strahlung and double Higgs bremsstrahlung off top quarks [112]. The size of the bands shows the total uncertainties originating from the scale dependence and the $PDF + \alpha_s$ uncertainties. 19
- 1.8 Higgs pair invariant mass distribution at leading order for the different contributions to the gluon fusion production mechanism and their interference [36] 21

- 1.9 Higgs pair invariant mass and transverse-momentum distributions for a center-of-mass energy of 14 TeV in various approximations. The full NLO results are shown in red. The red bands show the renormalization and factorization scale dependence obtained from a 7-point scale variation around the central scales $\mu_f = \mu_r = m_{HH}/2$ [18]. 21
- 1.10 The benchmark signal shapes, generating several distributions (1000 events each) of double Higgs boson invariant mass $M_{inv}(HH)$, in LO accuracy using as MC generator MadGraph5_aMC@NLO (MG5_aMC_v2.6.3.2), combining the values of the coupling coefficients from the SM extended Lagrangian in the framework of the EFT 23
- 1.11 Scheme of dimensions on RS theory. The Gravity (Planck) and Weak (TeV) branes are the 4 dimensional boundaries of the extra dimension ϕ compactified on a interval $(0, \pi)$. The figure also illustrate the metric behavior along the extra dimension 27
- 1.12 Representation of the two main benchmarks WED scenarios for the localization of the matter, where left we have Bulk scenario and right we have the RS1 scenario. 29
- 1.13 Production cross sections in pb for the Radion, with $k_l = 35$ and $\Lambda_R = 3\text{TeV}$ [27]. The GF (red) and VBF (Green) production modes for both to bulk (continuous lines) and RS1 (dashed lines) scenario. Right: 13TeV. Left: 100TeV. 31

Chapter 2

- 2.1 . Representation of the CERN accelerator complex [34]. 35
- 2.2 . The mean values of interactions per crossing (pileup) and the minimum bias cross sections for 2011-2012 and 2015-2018, where the distribution of each year is stacked above the previous year [124]. 38
- 2.3 . Total integrated luminosity (top) and peak instantaneous luminosity (bottom) of the LHC as a function of the year, as measured by the CMS experiment [124] 40
- 2.4 . Baseline schedule of LHC and HL-LHC operations [123] 41
- 2.5 . The LHC main experiments: ATLAS, CMS, ALICE and LHCb 42
- 2.6 . Perspective view of the CMS detector layout 43
- 2.7 . The CMS coordinate system 43
- 2.8 . Overview of the CMS tracker system 44
- 2.9 . Longitudinal view of the ECAL layout, representing one quarter of the detector. The barrel and endcaps sections and the preshower detector are shown. 45
- 2.10 . Longitudinal view of the HCAL layout, where are visible the barrel (HB) and endcaps (HE) detectors located inside the solenoid, the outer detector (HO) outside the solenoid, and the forward calorimeter (HF) 46
- 2.11 . Schematic view of the CMS Muon spectrometer. The interaction point is at the lower left corner. The locations of the various muon stations are shown in color (MB = DT = Drift Tubes, ME = CSC = Cathode Strip Chambers, RB and RE = RPC = Resistive Plate Chambers). M denotes Muon, B stands for Barrel and E for Endcap. The magnet yoke is represented by the dark gray areas. 47

- 2.12 . Structure of a Drift Tube Chamber with three superlayers composed of 4 layers each of drift cells. Right: Layout of a drift cell showing the electric field lines in the gas volume. 48
- 2.13 . 1) Left: Diagram and principle of operation of a CSC endcap muon chamber in CMS; 2) Right: Schematic of the on-chamber and peripheral crate electronics system. 49
- 2.14 . Achitecture of the L1 system [?] 50
- 2.15 . Expected LHC instantaneous and integrated luminosity as a function of time up to 2037 before the pandemic (top) and after the pandemic (bottom). Dates for long-shutdowns (LS) periods are also marked [159] 51

Chapter 3

- 3.1 Scematic structure of the CMS double-gap RPC [127] 57
- 3.2 . Schematic view of the barrel wheels (left) and schematic view of the RPC endcap station (right) 58
- 3.3 Comparison of efficiency curves of CMS RPC double and single-gap, measured with cosmic muons 59
- 3.4 Extention the inner ring of the third (RE3/1) and fourth disk (RE4/1) of the endcap stations from $|\eta| < 1.8$ up to $|\eta| < 2.4$ shown bottom right in the orange box. 60
- 3.5 . Scematic representation of the planned extention of the RE3/1 (left) and RE4/1 (right) endcap. 60
- 3.6 . Schematic structure of the Kodel 1/2 size graphite iRPC at the groud site (left) and high voltage HV site (right), where the two different graphite resistivity $50k\Omega/600k\Omega$ are shown. 61
- 3.7 . The read-out electronics on the strips PCB of the double coated 1/2 size graphite iRPC chamber, where we have in the center eight custom ASIC discriminators FEBs and and in both sides left and right in total sixteen fast charge analog preamplifier. 62
- 3.8 . Full assembled double coated graphite iRPC chamber at the B. 904 at CERN 63
- 3.9 . Cosmic radiation tests with the scintillators (coverd in black tape) as a trigger system on the improved RPC module at B. 904. 63
- 3.10 The patch-panel with the labeled 6 adjustable low voltages (LV) cables for the the powering of the front-end FEBs: V_{TH} , V_{DIS} , V_{AMP} , V_{PU} , V_{LVDS} , V_{PREAMP_Left} , V_{PREAMP_Right} 64
- 3.11 Blue box: Noise window; Red Box: Muon window 65
- 3.12 Efficiency curves on working point $\sim 7kV$ with offline threshold of $10 * \sigma$ of the noise,where similar efficiency for both high and low resistivity regions 66
- 3.13 Cluster size distributions at WP, where we can notice that the average cluster size is slightly lower for high resistivity region and it has a different shape in cluster size, where it is broader for the low ($50k\Omega$) resistivity region 66
- 3.14 Efficiency curves on working point $\sim 7kV$ with offline threshold of $10 * \sigma$ of the noise, where similar efficiency for both high and low resistivity regions with the FEBs 67
- 3.15 Average muon cluster size vs. HV, where the cluster size is lightly lower for high resistivity region with the FEBs 68

- 3.16 . Efficiency curves on three different value of the amplification power-up value of the voltages for the custom ASIC discriminators FEBs in the high resistivity region onf the graphite (600 $k\Omega$): 1) Left: $V_{amp} = 2.5$ V (left); 2) Center: $V_{amp} = 2.25$ V; 3) Right: $V_{amp} = 2.0$ V 68

Chapter 4

- 4.1 a) Left: Feynman diagramm at LO of the HH-box and triangle production used in MadGraph5_aMC@NLO (`loop_sm_hh` with MG5_aMC_v2_5_5 and `BSM_gg_hh` ($k_\lambda = k_t = 1$ with MG5_aMC_v2_6_1)[82]; b) Right: Feynman diagramm at NLO of the HH production used in MadGraph5_aMC@NLO and POWHEG-BOX-V2 (`SMEFT_UF0-no_b_mass` with MG5_aMC_v2_6_0 and `ggHH_EWChL` with POWHEG-BOX-V2) [82][81] 74
- 4.2 Comparison between the at NLO-MC model `ggHH_EWChL` within the POWHEG-BOX-V2 (in red) and the LO-MC model `loop_sm_hh` (green) and `BSM_gg_hh` (in grey) within the MadGraph5_aMC@NLO. a) Top left: Distribution of the M_{inv}^{HH} peaking around 400GeV; b) Top right: Distribution of the p_T^{HH} peaking around 20GeV; c) Center left: Distribution of the $p_T^{H_1}$ of the leading H -boson peaking around 150GeV; d) Center right: Distribution of the $p_T^{H_2}$ of the second H -boson peaking around 150GeV; e) Bottom: Distribution of the angular separation ΔR_{HH} 76
- 4.3 Comparison between the at NLO-MC model `ggHH_EWChL` within the POWHEG-BOX-V2 (in red) and the NLO-MC model `SMEFT_UF0-no_b_mass` in blue within the MadGraph5_aMC@NLO; a) Top left: Distribution of the M_{inv}^{HH} peaking around 400GeV; b) Top right: Distribution of the p_T^{HH} peaking around 20GeV; c) Center left: Distribution of the $p_T^{H_1}$ of the leading H -boson peaking around 150GeV; d) Center right: Distribution of the $p_T^{H_2}$ of the second H -boson peaking around 150GeV; e) Bottom: Distribution of the angular separation ΔR_{HH} 78
- 4.4 Comparision between two different PDF-sets usage; a) Top left: Distribution of the M_{inv}^{HH} peaking around 400GeV; b) Top right: Distribution of the p_T^{HH} peaking around 20GeV; c) Bottom left: Distribution of the $p_T^{H_1}$ of the leading H -boson peaking around 150GeV; d) Bottom right: Distribution of the $p_T^{H_2}$ of the second H -boson peaking around 150GeV 80
- 4.5 The k_λ -scan $[-10; 10; 1]$ of the trilinear and quartic Higgs self coupling constant $\lambda_{3H,4H}$ within the PDF, $\alpha(Q_s^2)$, scale and top quark mass uncertainties 81
- 4.6 Mass distributions of the k_λ -scan within the restricted range $[-2; 4]$ of the trilinear and quartic Higgs self coupling constant $\lambda_{3H,4H}$, where we can notice the maximun interference evolution between the triangle and box Feynman diagrams from $k_\lambda = 2.2$ and $k_\lambda = 2.6$ 82
- 4.7 The Feynman diagrams of the resonant HH production via gluon fusion at LO, where a massive X resonance decays into a pair of Higgs bosons. 82
- 4.8 Plots of the computed cross sections in linear (right) and log (left) mode of the event generation $X(BK - Gr/Radion) \rightarrow HH$. 84
- 4.9 Schematic representation of the data tiers: Reconstructed (RECO) Data and Analysis Object Data (AOD). The RECO data contains objects from all stages of reconstruction coming from the hits of Physics objects (muons, electrons, etc.) of the particle shower during the pp -collisions, which were converted into detector response digital information from the CMS trigger system (RAW), meanwhile the AODs are derived from the RECO information to provide data for physics analyses in a convenient, compact format. 85

- 4.10 a) Top-left: Invariant mass of the di-Higgs M_{inv}^{HH} ; b) Top-center: Transverse momentum of the di-Higgs p_T^{HH} ; c) Top-right: Isolation of the two Higgs bosons within the cone with radius ΔR_{HH} ; d) Bottom-left: transverse momentum of the first Higgs boson $p_T^{H_1}$; e) Bottom-right: transverse momentum of the second Higgs boson $p_T^{H_2}$ 86
- 4.11 f) Left: Mass distribution of the Higgs boson reconstructed at GEN-SIM level from the two ZZ -bosons decaying into leptons ($l = e, \mu$); f) Right: Mass distribution of the Higgs boson reconstructed at GEN-SIM level from two b-tagged jets 86
- 4.12 Kinematical distributions of the Radion spin-0 at 250GeV, 500 GeV, 1000GeV at the construction step within the signal region of $|m_{4l} - 125.09| - 10\text{GeV} \leq m_{4l} \leq 125.09 + 10\text{GeV}$ and $\geq 2\text{jets}$: a) Left: invariant mass distribution M_{inv}^{HH} Radion spin0 at 250GeV, 500 GeV, 1000GeV; b) Centre: invariant mass distribution M_{inv}^H from the $H \rightarrow ZZ(4l)$ Radion spin0 at 250GeV, 500 GeV, 1000GeV; c) Right: invariant mass distribution M_{inv}^H from the $H \rightarrow b\bar{b}$ Radion spin0 at 250GeV, 500 GeV, 1000GeV; 89
- 4.13 Distribution of pileup in 2016 (left), 2017 (center) and 2018 (right) Data and MC, before and after the application of PU weights. 91

Chapter 5

- 5.1 . Slice showing CMS sub-detectors and how particles interact with them as well as a detailed illustration of the experimental signature of final state particles in the CMS detector. This picture is adapted from the Ref. [142] 99
- 5.2 . The Output of the multiclassifier discriminant for prompt electrons matched to truth electrons from Z decay (blue) and for misidentified electrons (red). Events are all taken from Drell-Yan with jets MC sample. This plot is adapted from Ref. [162] 102
- 5.3 . Distribution of leptonic Z^0 -boson invariant mass m_{Z^0} 109
- 5.4 . Distribution of leptonic nominal Z_1^0 -boson invariant mass $m_{Z_1^0}$ (left) and off-shell Z_w^0 -boson invariant mass $m_{Z_w^0}$ (right) with full RunII data 137fb^{-1} at $\sqrt{s} = 13\text{TeV}$ 110
- 5.5 Distribution of the invariant mass $m_{H(4l)}$ of the leptonic Higgs H boson with full RunII data 137fb^{-1} at $\sqrt{s} = 13\text{TeV}$ 111
- 5.6 . Di-jet invariant mass built from RECO jets selected with the three different methods considered: the two highest p_T jets (blue line), the highest b-tagger score jet and the highest p_T jet (green line), the two highest b tagger score jets in the event (red line). 112
- 5.7 . Schematic view of the observation of b-tagged jets (in blue) coming from the decay of the Higgs boson, in association with a e^+e^- track (in red) coming from a Z boson (VH)[153] 112
- 5.8 Fake rates 2016: Fake rates for 2016 dataset are shown in terms of the lepton probe p_T and η . The η distributions (barrel and endcap) corresponds to the regions defined by $|\eta| < 1.479$ (1.2) for electrons (muons). The background $W^\pm Z^0$ from MC has been subtracted. 115
- 5.9 Fake rates 2017: Fake rates for 2017 dataset are shown in terms of the lepton probe p_T and η . The η distributions (barrel and endcap) corresponds to the regions defined by $|\eta| < 1.479$ (1.2) for electrons (muons). The background $W^\pm Z^0$ from MC has been subtracted. 116

- 5.10 Fake rates 2018: Fake rates for 2018 dataset are shown in terms of the lepton probe p_T and η . The η distributions (barrel and endcap) corresponds to the regions defined by $|\eta| < 1.479$ (1.2) for electrons (muons). The background $W^\pm Z^0$ from MC has been subtracted. 117
- 5.11 . Schematic view of a decision tree. Starting from the root node, a sequence of binary splits using the discriminating variables x_i is applied to the data. Each split uses the variable that at this node gives the best separation between signal and background when being cut on. The same variable may thus be used at several nodes, while others might not be used at all. The leaf nodes at the bottom end of the tree are labeled "S" for signal and "B" for background depending on the majority of events that end up in the respective nodes. For regression trees, the node splitting is performed on the variable that gives the maximum decrease in the average squared error when attributing a constant value of the target variable as output of the node, given by the average of the training events in the corresponding (leaf) node. This picture is adapted from Ref. [129] 122
- 5.12 . ROC curves after the training of the BDT for the channels : 1) 2016 (top): $b\bar{b}4e$ (top left) with an area of 0.934, $b\bar{b}4\mu$ (top center) with an area of 0.932 and $b\bar{b}2e2\mu$ (top right) with an area of 0.924 for the SM HH at NLO signal sample; 2) 2017 (center): $b\bar{b}4e$ (center left) with an area of 0.891, $b\bar{b}4\mu$ (center center) with an area of 0.912 and $b\bar{b}2e2\mu$ (center right) with an area of 0.914 for the SM HH at NLO signal sample; 3) 2018 (bottom): $b\bar{b}4e$ (bottom left) with an area of 0.930, $b\bar{b}4\mu$ (bottom center) with an area of 0.932 and $b\bar{b}2e2\mu$ (bottom right) with an area of 0.933 for the SM HH at NLO signal sample. 125
- 5.13 . Classifier output after the training of the BDT for the channels : 1) 2016 (top): $b\bar{b}4e$ (top left), $b\bar{b}4\mu$ (top center) and $b\bar{b}2e2\mu$ (top right) for the SM HH at NLO signal sample; 2) 2017 (center): $b\bar{b}4e$ (center left), $b\bar{b}4\mu$ (center center) and $b\bar{b}2e2\mu$ (center right) for the SM HH at NLO signal sample; 3) 2018 (bottom): $b\bar{b}4e$ (bottom left), $b\bar{b}4\mu$ (bottom center) and $b\bar{b}2e2\mu$ (bottom right) for the SM HH at NLO signal sample; 126
- 5.14 . BDT input variables distributions for 2018 datasets in case only of the $b\bar{b}2e2\mu$ final state as an example. 127
- 5.15 BDT correlation plots for 2018: correlation between background's variables (right), correlation between signal's variables (left) for each final state ($b\bar{b}4e$ top, $b\bar{b}4\mu$ center, $b\bar{b}2e2\mu$ bottom). 131
- 5.16 BDT score in the sideband of the 4 leptons distribution: 1) Top left: for 2016 datasets; 2) Top right: for 2017 datasets; 3) Bottom: for 2018 datasets;. 132
- 5.17 1) Top left: B tagger score of the jet with the highest value of the deepCSV discriminant for 2016 datasets; 2) Top right: B tagger score of the jet with the second-highest value of the deepCSV discriminant for 2016 datasets; 3) Center left: p_T of jet with the highest value of the deepCSV discriminant for 2016 datasets; 3) Center right: p_T of jet with the second-highest value of the deepCSV discriminant for 2016 datasets; 5) Bottom left: ΔR between the $H \rightarrow 4l$ and $H \rightarrow b\bar{b}$ for 2016 datasets; 6) Bottom right: Invariant mass of the jet built with the two jets with the highest value of the deepCSV discriminant for 2016 datasets. 133

- 5.18 1) Top left: B tagger score of the jet with the highest value of the deepCSV discriminant for 2017 datasets; 2) Top right: B tagger score of the jet with the second-highest value of the deepCSV discriminant for 2017 datasets; 3) Center left: p_T of jet with the highest value of the deepCSV discriminant for 2017 datasets; 3) Center right: p_T of jet with the second-highest value of the deepCSV discriminant for 2017 dataset; 5) Bottom left: ΔR between the $H \rightarrow 4l$ and $H \rightarrow b\bar{b}$ for 2017 datasets; 6) Bottom right: Invariant mass of the jet built with the two jets with the highest value of the deepCSV discriminant for 2017 datasets. 134
- 5.19 1) Top left: B tagger score of the jet with the highest value of the deepCSV discriminant for 2018 datasets; 2) Top right: B tagger score of the jet with the second-highest value of the deepCSV discriminant for 2018 datasets; 3) Center left: p_T of jet with the highest value of the deepCSV discriminant for 2018 datasets; 3) Center right: p_T of jet with the second-highest value of the deepCSV discriminant for 2018 dataset; 5) Bottom left: ΔR between the $H \rightarrow 4l$ and $H \rightarrow b\bar{b}$ for 2018 datasets; 6) Bottom right: Invariant mass of the jet built with the two jets with the highest value of the deepCSV discriminant for 2018 datasets. 135
- 5.20 . ROC curves after the training of the BDT for the channels or Bulk Graviton $m_{HH}^{BKGr} = 400\text{GeV}$, Radion $m_{HH}^{RD} = 500\text{GeV}$ for all three years and three channels (column left: $bb4e$, column center: $bb4\mu$, column right: $bb2e2\mu$) 136
- 5.21 . Classifier output after the training of the BDT for Bulk Graviton $m_{HH}^{BKGr} = 400\text{GeV}$, Radion $m_{HH}^{RD} = 500\text{GeV}$ for all three years and three channels (column left: $bb4e$, column center: $bb4\mu$, column right: $bb2e2\mu$). 137
- 5.22 . BDT input variables distributions for 2018 datasets of the KK Bulk Graviton spin-2 with $m_{HH}^{BK-Gr} = 400\text{ GeV}$ in case only of the $b\bar{b}2e2\mu$ final state as an example. 138
- 5.23 . BDT input variables distributions of the KK Bulk Radion spin-0 with $m_{HH}^{RD} = 500\text{ GeV}$ for 2018 datasets in case only of the $b\bar{b}2e2\mu$ final state as an example. 139
- 5.24 BDT correlation matrix plots for 2018 for KK Bulk Graviton spin-2 with $m_{HH}^{BK-Gr} = 400\text{ GeV}$ and KK Bulk Radion spin-0 with $m_{HH}^{RD} = 500\text{ GeV}$ for the correlation between background's variables (top) and correlation between signal's variables (bottom) for the final state $b\bar{b}2e2\mu$. 140
- 6.1 . Distribution of the 4-leptons invariant mass $m_{H(4l)}$ (top left) and the distribution of the invariant of the di-jet system $m_{H(b\bar{b})}$ (top right), where only events in the mass window $115\text{GeV} < m_{4l} < 135\text{ GeV}$ are considered. The distributions and the BDT score for the signal HH , which is scaled by an illustrative factor (in red dotted line), the estimated background components, as well the data are shown in the bottom plot. A good signal / background separation is observed for the three years, in particular $S\sqrt{B}$ (where S is the signal and B is the background bin per bin) is higher in the last three bins, this means that the majority of the signal is concentrated at between 0.6 and 1. 145
- 6.2 . Signal region plots Full RunII ($137fb^{-1}$): $m_{H(4l)}$ (left), p_T^{4l} (right) 146
- 6.3 . Signal region plots Full RunII ($137fb^{-1}$): $p_T^{jet_1}$ (left), $p_T^{jet_2}$ (right) 146
- 6.4 . Signal region plots Full RunII ($137fb^{-1}$): m_{HH} (right), p_T^{HH} (left) 147
- 6.5 . Upper limit on the signal strength for each dataset and for the full Run2 combination, where we obtain an observed(expected) upper limit on the signal strength to be about $r_{SM}^{HH} < 30(37)@95\% CL_s$. 151

6.6	. Expected and observed 95% CL_s upper limit on the $HH \rightarrow b\bar{b}4l$ production cross section derived as a function of $k_\lambda = \lambda_{HHH}/\lambda_{SM}$ for the full Run II, with a constraint range at $-9(-10.5) < r_{(k_\lambda)} < 14(15.5)$ at 95% CL_s where the red line represents the theoretical value of the cross section. 152
6.7	. Upper limit on the signal strength (left) and expected 95% CLs upper limit on the cross section as a function of HH invariant mass (right) for the KK Bulk Graviton spin-2 for RunII at $137.2fb^{-1}$. The solid red line shows the theoretical predictions. 154
6.8	. Upper limit on the signal strength (left) and expected 95% CLs upper limit on the cross section as a function of HH invariant mass (right) for the KK Bulk Radion spin-0 for RunII at $137.2fb^{-1}$. The solid red line shows the theoretical predictions. 154
6.9	. Upper limit at the 95% confidence level CL_s on the HH production cross section as a function of $k_\lambda = \lambda_{HHH}/\lambda_{SM}$. The red band indicated the theoretical production cross section. Right: expected likelihood scan as a function of $k_\lambda = \lambda_{HHH}/\lambda_{SM}$. In both figures the results are shown separately for the five decay channels studied and for their combination. This table is adapted from Ref. [157]. 156
6.10	. 1) Left: Minimum negative-log-likelihood as a function of k_λ , calculated by performing a conditional signal+background fit to the background and SM signal. The black line corresponds to the combined ATLAS and CMS results, while the blue and red lines correspond to the ATLAS and CMS standalone results respectively. 2) Right: The different colours correspond to the different channels, the plain lines correspond to the CMS results while the dashed lines correspond to the ATLAS results. This table is adapted from Ref. [157]. 157
A.1	Invariant mass m_{HH} for the k_λ -scan A-6
A.2	$p_T(HH)$ k_λ -scan A-7
A.3	$p_T(H1), p_T(H2)$ k_λ -scan A-7
B.1	. 1) Top left: Upper limit on the signal strength for 2016; 2) Top right: Expected 95% CLs upper limit on the cross section as a function of HH invariant mass for 2016. 3) Bottom left: Upper limit on the signal strength for 2016. 4) Bottom right: Expected 95% CLs upper limit on the cross section as a function of HH invariant mass for 2016. The solid red line shows the theoretical predictions. A-11
B.2	. 1) Top left: Upper limit on the signal strength for 2017; 2) Top right: Expected 95% CLs upper limit on the cross section as a function of HH invariant mass for 2017. 3) Bottom left: Upper limit on the signal strength for 2017. 4) Bottom right: Expected 95% CLs upper limit on the cross section as a function of HH invariant mass for 2017. The solid red line shows the theoretical predictions. A-12
B.3	. 1) Top left: Upper limit on the signal strength for 2018; 2) Top right: Expected 95% CLs upper limit on the cross section as a function of HH invariant mass for 2018. 3) Bottom left: Upper limit on the signal strength for 2018. 4) Bottom right: Expected 95% CLs upper limit on the cross section as a function of HH invariant mass for 2018. The solid red line shows the theoretical predictions. A-13

List of Tables

- 1.1 Branching fractions of the main Higgs boson decay modes for a SM Higgs boson of mass $m_H = 125.09$ GeV [167]. Theoretical uncertainties combine the uncertainties on the Higgs boson partial width, on the value of α_s , and on the quark masses. 14
- 1.2 Computations of the signal cross sections (in fb) for double Higgs (H) production at different centre-of-mass energies [36], including the available QCD corrections according to the recommendations of the LHC Higgs Cross Section Working Group [168]. The renormalisation and factorisation scales have been set to $m_{HH}/2$ for gluon gluon fusion, to the individual virtualities $Q_{1,2} = \sqrt{-q_{1,2}^2}$ of the t-channel vector-bosons for VBF, m_{HHV} ($V = W, Z$) for HHV production, $m_{t\bar{t}}/2$ for $t\bar{t}HH$, $m_{HH}/2$ for $tjHH$ production. PDF4LHC15 parton distributions have been used to obtain the results, and the corresponding $PDF + \alpha_s$ uncertainties. The cross sections for $tjHH$ involve both top and anti-top production. 18
- 1.3 NNLL matched to NNLO cross sections for $gg \rightarrow HH$ with a central scale $\mu_0 = M_{HH}/2$ with $m_H = 125.09$ GeV [48][72] computed in the HEFT approximation. The uncertainties from top quark mass effects are not included in this table. Uncertainties are evaluated using the PDF4LHC recommendation based on the PDF4LHC15_nnlo_mc set. 19
- 1.4 K -factors as defined in Eq. 1.41 for $gg \rightarrow HH$ with a central scale $\mu_0 = M_{HH}/2$ and $\mu_0 = M_{HH}$ with $m_H = 125.09$ GeV. 20
- 1.5 Signal cross section (in fb) for $gg \rightarrow HH$ at NLO QCD in the $m_t \rightarrow \inf$ limit, reweighted by the exact LO result [18], with a central scale $\mu_0 = M_{HH}/2$ with the scale variation and top quark mass effects uncertainties. 21
- 1.6 MC generations of the benchmarks shapes: Each benchmark shape corresponds to a specific choice of the 5 EFT couplings, whose choices do not have a special physical meaning, it is only meant to represent a shape, not a special point in the EFT parameter space 22
- 2.1 Nominal LHC beam parameters [68] 39
- 3.1 The six adjustable low voltages (LV) for the Front-End custom ASIC discriminators and fast timing pre-amplifiers: V_{TH} , V_{DIS} , V_{AMP} , V_{PU} , V_{LVDS} , V_{PREAMP_Left} , V_{PREAMP_Right} . In the 2nd column are the respective threshold values of those voltages given from the manufacturer. In column 3 and 5 we have the applied voltages (V_{Appl}) and current (I_{Oset}), meanwhile in 4 and six with have the measured values V_{Meas} , I_{mon} . 64
- 3.2 . Comparison of between the two resistivity regions for both read-out electronics. In all results of the tests, for the higher resistivity of the graphite ($600k\Omega$) we get better results in comparison to the lower resistivity ($50k\Omega$). 69
- 4.1 Comparison of the same NLO SM-HH with full top quark mass m_t dependence generated with two different MC generator frameworks: POWHEG-BOX-V2 and MadGraph5_aMC@NLO 77

4.2	Comparison of the same NLO SM-HH with full top quark mass m_t dependence generated with two different MC generator frameworks: POWHEG-BOX-V2 and MadGraph5_aMC@NLO	79
4.3	. Computed cross sections of the event generation $X(BK - Gr/Radion) \rightarrow HH$ [122]	83
4.4	. 2016 signal samples and cross sections for the non-resonant HH	87
4.5	. 2017 signal samples and cross sections for the non-resonant HH	87
4.6	. 2018 signal samples and cross sections for the non-resonant HH	88
4.7	. 2016 signal samples and cross sections for the Bulk Graviton spin-2 and Radion spin-0	88
4.8	. 2017 signal samples and cross sections for the Bulk Graviton spin-2 and Radion spin-0	89
4.9	. 2018 signal samples and cross sections for the Bulk Graviton spin-2 and Radion spin-0	90
4.10	. Background Monte Carlo samples and cross sections	92
4.11	. Trigger paths used in 2016 collision data	94
4.12	. 2016 datasets used in the analysis	95
4.13	. 2017 datasets used in the analysis	96
4.14	. Trigger paths used in 2017 collision data	96
4.15	. 2018 datasets used in the analysis	97
4.16	Trigger paths used in 2018 collision data	97

Chapter 5

5.1	. Minimum BDT score required for passing the electron identification for 2016, 2017 and 2018 [128].	101
5.2	. Overview of input variables to the identification classifier [128].	103
5.3	. The requirements for a muon to pass the Tracker High- p_T Identification (ID). Note that these are equivalent to the Muon POG High- p_T ID with the global track requirements removed. [128].	105
5.4	Reducible background yields for all the channels: 2016 (1 st row), 2017 (2 nd row) and 2018 (3 rd row) datasets	117
5.5	Systematic uncertainties on the reducible background estimate for the 3 years, [161].	118
5.6	. Summary of the experimental systematic uncertainties for all the tree years [161].	118
5.7	. Theoretical uncertainties of the background samples [161].	120
5.8	. Theoretical uncertainties of the signal samples	120
5.9	Configuration options used for the BDT training for all the tree years (2016-2017-2018) samples. Other options used in the algorithm are kept to their default values.	123
5.10	. BDT ranking of input variables of $b\bar{b}4e$ -channel for 2018 dataset.	124
5.11	. BDT ranking of input variables of $b\bar{b}4\mu$ -channel for 2018 dataset.	124
5.12	. BDT ranking of input variables of $b\bar{b}2e2\mu$ -channel for 2018 dataset.	124

- 5.13 . UAC of the ROC curves after the training of the BDT for the channels for the Bulk Graviton $m_{HH}^{BKGr} = 400\text{GeV}$ and Radion $m_{HH}^{RD} = 500\text{GeV}$ for all three years and each channel ($b\bar{b}4e$, $b\bar{b}4\mu$, $b\bar{b}2e2\mu$) 129
- 5.14 . BDT ranking of input variables of $b\bar{b}4e$ -channel for 2018 dataset, Bulk Graviton spin-2, 400 GeV. 129
- 5.15 . BDT ranking of input variables of $b\bar{b}4\mu$ -channel for 2018 dataset, Bulk Graviton spin-2, 400 GeV. 129
- 5.16 . BDT ranking of input variables of $b\bar{b}2e2\mu$ -channel for 2018 dataset, Bulk Graviton spin-2, 400 GeV. 129
- 5.17 . BDT ranking of input variables of $b\bar{b}4e$ -channel for 2018 dataset, Radion spin-0, 500 GeV. 130
- 5.18 . BDT ranking of input variables of $b\bar{b}4\mu$ -channel for 2018 dataset, Radion spin-0, 500GeV. 130
- 5.19 . BDT ranking of input variables of $b\bar{b}2e2\mu$ -channel for 2018 dataset, Radion spin-0, 500 GeV. 130
- 6.1 Cut flow table after signal region selection for the full RunII dataset ($137.2fb^{-1}$). The ZZ yields contain contributions from the processes $gg/q\bar{q} \rightarrow ZZ$, Higgs + VBF contains contributions from the single Higgs boson gluon fusion (ggH) and vector boson fusion (VBFH) production modes, others contains contributions from TTW, WWZ, WZZ, ZZZ and processes. 144
- 6.2 . Cut flow table after signal region selection for the Run II dataset ($137.2fb^{-1}$) for the Bulk-Graviton spin-2 and Radion spin-0 147
- 6.3 . Table of limits for all three years and the combination of the full Run II. 151
- 6.4 . Table of limits for the KK Bulk Graviton spin-2 for full RunII at a total integrated luminosity of $\mathcal{L} = 137fb^{-1}$. 153
- 6.5 . Table of limits for the KK Bulk Radion spin-0 for full RunII at a total integrated luminosity of $\mathcal{L} = 137fb^{-1}$. 154
- 6.6 . Upper limit at the 95% confidence level, significance, projected measurement at 68% confidence level of the Higgs boson self coupling λ_{HHH} for the five channels studied and their combination for the CMS-Phase2 projections. Systematic and statistical uncertainties are considered. This table is adapted from Ref. [157]. 156
- 6.7 . Significance in standard deviations of the individual channels as well as their combination. This table is adapted from Ref. [157]. 157

Appendix A

Appendix B

- B.1 Cut flow table after signal region selection for the 2016 dataset ($35.8fb^{-1}$). The ZZ yields contain contributions from the processes $gg/q\bar{q} \rightarrow ZZ$, Higgs + VBF contains contributions from the single Higgs boson gluon fusion (ggH) and vector boson fusion (VBFH) production modes, others contains contributions from TTW, WWZ, WZZ, and ZZZ processes. A-9

- B.2 Cut flow table after signal region selection for the 2017 dataset ($41.53fb^{-1}$). The ZZ yields contain contributions from the processes $gg/qq \rightarrow ZZ$, Higgs + VBF contains contributions from the single Higgs boson gluon fusion (ggH) and vector boson fusion (VBFH) production modes, others contains contributions from TTW, WWZ, WZZ and ZZZ processes. [A-9](#)
- B.3 Cut flow table after signal region selection for the 2018 dataset ($59.74fb^{-1}$). The ZZ yields contain contributions from the processes $gg/qq \rightarrow ZZ$, Higgs + VBF contains contributions from the single Higgs boson gluon fusion (ggH) and vector boson fusion (VBFH) production modes, others contains contributions from TTW, WWZ, WZZ, ZZZ and processes. [A-9](#)
- B.4 . Cut flow table after signal region selection for the 2016 dataset ($35.8fb^{-1}$) for the BulkGraviton spin-2 and Radion spin-0 [A-10](#)
- B.5 . Cut flow table after signal region selection for the 2017 dataset ($41.53fb^{-1}$) for the Bulk-Graviton spin-2 and Radion spin-0 [A-10](#)
- B.6 . Cut flow table after signal region selection for the 2018 dataset ($59.74fb^{-1}$) for the Bulk-Graviton spin-2 and Radion spin-0 [A-10](#)
- B.7 . Table of limits for the Bulk Graviton spin-2 and Radion spin-0 for 2016 at a total integrated luminosity of $\mathcal{L} = 35.8fb^{-1}$ [A-11](#)
- B.8 . Table of limits for the Bulk Graviton spin-2 and Radion spin-0 for 2017 at a total integrated luminosity of $\mathcal{L} = 41.53fb^{-1}$ [A-12](#)
- B.9 . Table of limits for the Bulk Graviton spin-2 and Radion spin-0 for 2018 at a total integrated luminosity of $\mathcal{L} = 59.74fb^{-1}$. [A-13](#)

Introduction

We might question ourselves in specific moments of our life about us, the world around us and, most importantly, what life is and how is that all we see, feel and sense, so real? How is that possible at all?

Well, even though those questions are so old, so deep in meaning, but at the same time so genuine and so right to be formulated, their answers are much more complicated than the humanity has thought since its existence on planet Earth.

We still don't really know how old our "intelligent" civilizations are, specially after the incredible discovery of Gobekli Tepe in Turkey, an ancient site, which according to archaeologists it might be even older than 10.000 years old, even older than the Sumerian or the Egyptian civilizations, with archaeological artifacts of einkorn wheat used for agriculture of an age of more than twelve thousands years, making everything questionable, even how old we actually as humans might have been as an intelligent society, being able to produce agriculture and, generally speaking, thinking about scientific rationality.

Until the moment we are speaking, we already know that at least the theory of a universe built out of elementary particles, the atom ("atomos"), was formulated in ancient Greece and India, where they thought that matter might be formed in such a way, that in its deepest only "uncuttable" particles (the atoms) are the actual bricks of the matter itself. This was of course a theory based on a philosophical reasoning of the world, a very courageous one, which was even supported by alchemists and chemistry pioneers such as John Dalton in the 1800-th century. At the very end of the day, it has nothing to do with the nowadays scientific concept of the atom, where we know today after several experimental discoveries that the atom is a composed structure of a negatively charged electron cloud with electrons pairing according to Pauli's exclusion theory in their respective energy levels surrounding a positively charged nucleus, made of baryons (protons and neutrons), themselves composite particles, where the gluons, the mediators of the strong interaction forces, bind very "strongly" the quarks together into a triplet, so robust, that their life-span might be defined as almost infinite, in case of no external acting "destructive" forces.

Besides the historical phenomena and happenings, we have in certain points of our history in recorded documents, from the Sumerian tablets, to the Egyptian papyrus rolls or to manuscripts and the handwritten books and personal writings, of the scientific evolution of our thinking and our discoveries, which opened to our modern society several doors to the secrets of the Universe and Life.

But, what about the Physics? How did it appear to the human society? And maybe, when? Could we even call it that time as such?

Interesting enough, the concept of Physics as we know it, is a very recent one. During the "explosion" of the Mesopotamian, Egyptian, Indian and European civilizations, Physics did not exist as a separate science as we know it nowadays. The ancient Greeks called it as "episteme" (knowledge of nature), a terminology used for a philosophical rational thinking of a natural event and the explanation of it by only logic but no experimental proof of its verity. This concept of just accepting an explanation of a natural phenomenon, was drastically changed by Galileo Galilei, due to his implementation of the scientific method, which integrates the experiment as the only way to an ultimate proof of a theoretical hypothesis, written with the language of

the universe: Mathematics, the only language to explain and describe "objectively" the laws of nature.

This simple idea of combining Mathematics and the experiment for the explanation and prove of the laws of Nature, opened a "Pandora" box for the secrets of the Universe, where a long list of brilliant minds started a long way to multiple discoveries and better understandings of the Universe and how it works. One of those beautiful minds was Isaac Newton, who within his master piece "Philosophiæ Naturalis Principia Mathematica" implemented the laws of motion and the laws of the universal gravitation. This incredible contribution of Isaac Newton to Physics was also possible thanks to the formulation of the mathematical theory of infinitesimal calculus, also independently formulated by Gottfried Wilhelm Leibniz at the same time, a theory which made possible the intimate boundary between Physics parameters and their dependencies as a function f changing with a parameter x , infinitesimally. As an example, it serves to our scope to remind the second axiom of Isaac Newton, where an external force is applied over a body with mass $m > 0[kg]$ (a scalar), which will feel a proportional acceleration, in other words, the body's velocity with mass $m > 0[kg]$ will change with time as a result of an external applied force over it.

The contribution of Isaac Newton in Optics, but also the studies on electric current from André-Marie Ampère and Michael Faraday, as best known for his studies on electromagnetic induction, and more contributions done during the 19th century on electromagnetism and electrodynamics in general, they were all elegantly incorporated in the system of the integral/differential equations of James Clerk Maxwell. This system of equations is so powerful, that they can explain at high precision most of the known phenomena of electromagnetism and electrodynamics. Furthermore they predicted the existence of the electromagnetic waves, which they propagate in the space-time at the velocity of light at $\approx 3 \times 10^5 kms^{-1}$. An incredible discovery, which led Albert Einstein to the formulation of his theory of special relativity, introducing him to the eternity of science history in the early 20th century.

The 20th century, was a gold mine for physicists, which started with an incredible milestone: The birth of quantum mechanics. Max Planck solved the Ultraviolet paradox of black body radiation with the magnificent concepts of the "quanta". The energy is emitted by bodies in portions ("Kuantitaeten"), a concept completed by Albert Einstein and priced in 1921 with the delivery of the Nobel price due to the discovery of the photoelectric effect. Einstein added that, not only energy is emitted in a discrete number of quanta, but at the same time, it is also absorbed in a discrete number of quantities, where the mediators where the photons: They are absorbed and emitted by bodies (like atoms), only if their kinetic energy is equal to a specific energy level of the atom/molecule, which is eligible of absorbing/emitting those photons, causing the jump to higher/lower energy levels of the electrons within the atom/molecule and in extreme case also their expulsion (ionization of the atom).

The discovery of the electron at the end of the 19th century, triggered a deeper understanding of the structure of the atom. The Rutherford experiment led to the atomic model structure due to α -Particles bombardment over a thin gold plate revealing the emptiness of the atom and the existence of a positive charged nucleus and a negative charged electron cloud. The discovery of the protons and the neutrons closed the circle of the atom's structure. Nevertheless, the great stability of the atom's nucleus remained a mystery until the discovery of the strong interaction forces. The discover of the muons μ through the cosmic radiation gave an important sign to the existence of further particles, which they might complete the puzzle of particle Physics.

The development of the particle detector methods and experiments, such as the Gargamelle chambers (a bubble chamber) at CERN led to the evidence of the weak neutral current, requiring the existence of a neutral particle (Z-boson) to carry the weak fundamental force. This was predicted by the electroweak theory, according to which the weak force and the electromagnetic force are different versions of the same force. This evidence led to the understanding of the electroweak interaction forces and their carriers: $W^\pm, Z^0$

bosons. Indeed, the potential and the credence of the Standard Model (SM) theory grew by time due to several experimental evidences during the 20th century. The SM is a theory developed during the second half of the past century. It reached its peak in 2012 with the discovery of the Higgs boson at the large hadron collider (LHC) experiment at CERN, almost six decades after the theoretical formulation of the spontaneous symmetry breaking on the scalar sector of the SM.

The SM model seems to be the most complete theory we have till now, which can explain the world we see. Nevertheless, due to some important experimental observations, there might be some evidence of Physics beyond the SM. One of those evidences of BSM phenomena, would be the discovery of the neutrino oscillation, which led to the prove of the neutrino mass, even though it has a relative small value. Furthermore, the anomaly of the rotational velocity of the galaxies, which instead of decreasing while going away from its center, it stays constant, enlightening physicists to idea of the existence of "dark stuff" (dark matter and dark energy). What is this dark energy and dark matter? How does regular matter interact with this dark stuff? According to theoretical estimations, the regular visible matter, which we are made of, is only about 4-5 % of all the universe composition, a fact, which implies a long journey towards the determination of a more complete and detailed description of the structure of our universe.

Objectives and Scope

This thesis is based on the analysis of the double Higgs production with the final state decay of the di-Higgs system into a pair of b-tagged jets and a combination of four leptons, electrons $e^\pm$ and $\mu^\pm$ muons. The analysis is presented for the full Run II datasets at the CMS experiment at CERN from proton-proton collisions pp with a center of mass energy of $\sqrt{s} = 13$ TeV at an integrated luminosity of $\mathcal{L} = 137.2 \text{ fb}^{-1}$. This thesis will give an outline for two variants of the double Higgs production:

1. The non-resonant HH production: The SM case of the di-Higgs phenomenology is reported, with the implementation of possible BSM outline regarding the deviation of the Higgs self coupling constant λ_{HHH} value due to a so called k_λ -scan. This analysis was the first HH approved analysis by the CMS collaboration on July 2020 under the following credentials: PAS-PUB, CMS PAS HIG-20-004, Analysis Note: AN-2019/117 and is part of the full Run II CMS HH combination. To whom it may concern, my personal contribution to this work is the detailed study on the non-resonant HH signal regarding the SM mode ($k_\lambda = 1$) and the BSM considerations regarding the Higgs self-coupling constant $\lambda_{3H/4H}$ incorporated into a k_λ -scan. This contribution contains the MC generations of the signal samples with different models, both SM and BSM considerations, with respective comparisons between the MC generations at leading order (LO) and Next-to-Leading order (NLO) as well as between the MC generators, such as MadGraph5_aMC@NLO and the POWHEG-BOX-V2. Furthermore, this contributions is enforced with the MC generations at NLO with full top-quark mass dependence of the SM HH signal samples and a full k_λ -scan with the range $k_\lambda = [-10, +10]$. I also calculated the theoretical uncertainties such as the up/down scale uncertainties, PDF and $\alpha_s(Q^2)$ uncertainties comparing those values for two different Parton Distribution Functions (PDF) sets. The results of these important studies were subsequently presented and approved within the CMS collaboration. The values of the full PDF sets uncertainties as well a the full total cross sections were then extrapolated from NLO into NNLO values through the k -factors. At last, not at least, this work was concluded with the full CMS detector response simulations of the SM and BSM MC samples of the full Runn II dataset, which made possible the study of the non-resonant $HH \rightarrow b\bar{b}ZZ(4l)$ and its approval.
2. The resonant HH production: The Kaluza-Klein theory and the Randall-Sundrum model provide possible massive gravitational particles in the scalar sector, where the decay of the KK Bulk Graviton

spin-2 and Radion spin-0 decaying into a pair of Higgs bosons is presented. This analysis is described within the CMS Analysis Note: AN-2021/115 and will follow the requested steps for pre-approval and final approval in the near future. To whom it may concern, this analysis is a full personal contribution of mine regarding all the steps and preliminary results presented in this Thesis.

In terms of the Phase2 upgrade for the High Luminosity LHC, my contribution regarding the Resistive Plate Chambers (RPC) system is presented. This study includes the performance effects of different resistivity values of the graphite, coated on the new improved RPC chambers bakelite plates regarding the Phase2 CMS upgrade of the RPC system. The result of this study is to certify the new resistivity value of the graphite for the RPC chambers system towards the HL-LHC.

Outlines: Contents of the PhD thesis

The PhD thesis is composed as following:

1) The double Higgs (H) production:

In Chapter 1 are introduced the basic concepts and theoretical annotations of the SM of particle physics as well as the the Brout-Englert-Higgs mechanism, which predicted the existence of the Higgs boson as well as the experimental evidence of it at the LHC. In addition to that, the EFT theory for non-resonant HH processes is implemented and a discussion over non-resonant HH phenomenology is reported. Furthermore, a "light" introduction over the Kaluza-Klein and Randall-Sundrum model and the Graviton particle is presented.

2) The CMS Experiment at the LHC (CERN):

In Chapter 2 are reported the details of the experimental apparatus regarding the LHC and its components, specially the CMS experiment. In addition to that, in Chapter 3, a study related the graphite resistivity for the improved RPC is presented. This study is relevant for the RPC system upgrade of the Phase 2 at HL-LHC.

3) $HH \rightarrow b\bar{b}ZZ(4l)$ Analysis:

The Monte Carlo (MC) generations of the signal sample for both resonant and non-resonant HH are presented in Chapter 4. Particle identification, isolation and reconstruction of muons, electron and jets are described in Chapter 5. In addition to that, the analysis strategy within the multivariate analysis with the implementation of a Boosted Decision Tree as a machine learning technique, are reported in the same chapter.

4) RESULTS:

The final results on both resonant and non-resonant HH are presented in Chapter 6. In this chapter signal yields of the full Run II analysis for signal, background and data are listed as well as upper limits are set on the signal strength modifier μ of each year and full Run II combination of the $b\bar{b}4l$ final state. The limits are set for the SM HH case, for the k_λ -scan and for both gravitational resonances KK Bulk Graviton spin-2 and Radion spin-0. A projection for the HL-LHC of the non-resonant HH mode for the CMS Phase 2 and with the combination of ATLAS and CMS is also described.



Part I

Double Higgs production

Chapter 1

The double Higgs (H) production

The Standard Model (SM) of particle Physics is the most complete theory upon experimental confirmations during the 20th and beginning of 21st centuries, which explains the Physics we know today. It is a renormalizable chiral gauge quantum field theory [20] that incorporates relativity and quantum mechanics and describes phenomena at the nuclear and sub-nuclear scales. It describes the weak, the strong and the electromagnetic forces and interactions and incorporates a minimal scalar sector, which originates from the spontaneous breaking of the electroweak symmetry and of the masses of the fermions. The SM has gained strong credence due to experimental observations at particle physics detectors and collider experiments from the discovery of the weak $W^{\pm}, Z^0$ bosons, the discovery of the top quark upon the confirmation of the Higgs boson in 2012. Nevertheless, several observations from the sub-nuclear range (such as the neutrino oscillations) to the astrophysical scales (such as dark matter and dark energy), are motivated from theoretical assumptions to further consider the formulation of a more complete theory than the SM. The formulation of a theory beyond the SM (BSM) may incorporate the gravitational force, unifying at last all the four elementary interaction forces. The scalar sector of the SM it would be a further extension of it, specially after the Higgs boson discovery. The production of Higgs boson pairs (HH) is one the products of the scalar sector, which it could be an interesting possibility to be tested at the LHC. It allows the determination of the Higgs boson self-interaction λ_{HHH} , and it may provide possible joint signs of BSM physics in case of a deviation related to the measurements of the $\lambda_{3H/4H}$ value. This chapter will introduce the main theoretical concepts of the SM Physics and it will discuss the relevance of the study of double Higgs production (HH) in the context of both SM and new BSM physics. A focus on the Higgs boson properties and couplings, HH production and further BSM considerations will be introduced. The Kaluza-Klein theory and the Randall-Sundrum model on possible gravitational particles such as the KK-Bulk Graviton spin-2 and Radion spin-0 will be discussed in the context of the resonant mode of the double Higgs production for the resonant HH signatures. Furthermore, a discussion is introduced related the phenomenology in collider experiments and results previously obtained at the LHC for the discovery and measurement of the Higgs boson as well as the respective cross section productions and branching ratios.

1.1 The standard model (SM) of particle physics.

The Standard Model of particle Physics is a theory, which describes the Physics behind the three fundamental interactions: the electromagnetic, the weak and the strong interactions. For the moment, gravity is not part of the SM. Nevertheless, the "Theory of Everything (TOE)" may also include gravity [51]. The SM is based on two main sectors: the gauge sector and the scalar sector. Both sectors derive from the theoretical concepts of gauge invariance and spontaneous symmetry breaking, which will be discussed in the subsections below. In the figure below, we have an overview on the SM elementary particles: the twelve fundamental fermions (six quarks and six leptons) and five fundamental bosons (four gauge bosons and the scalar Boson: the Higgs boson).

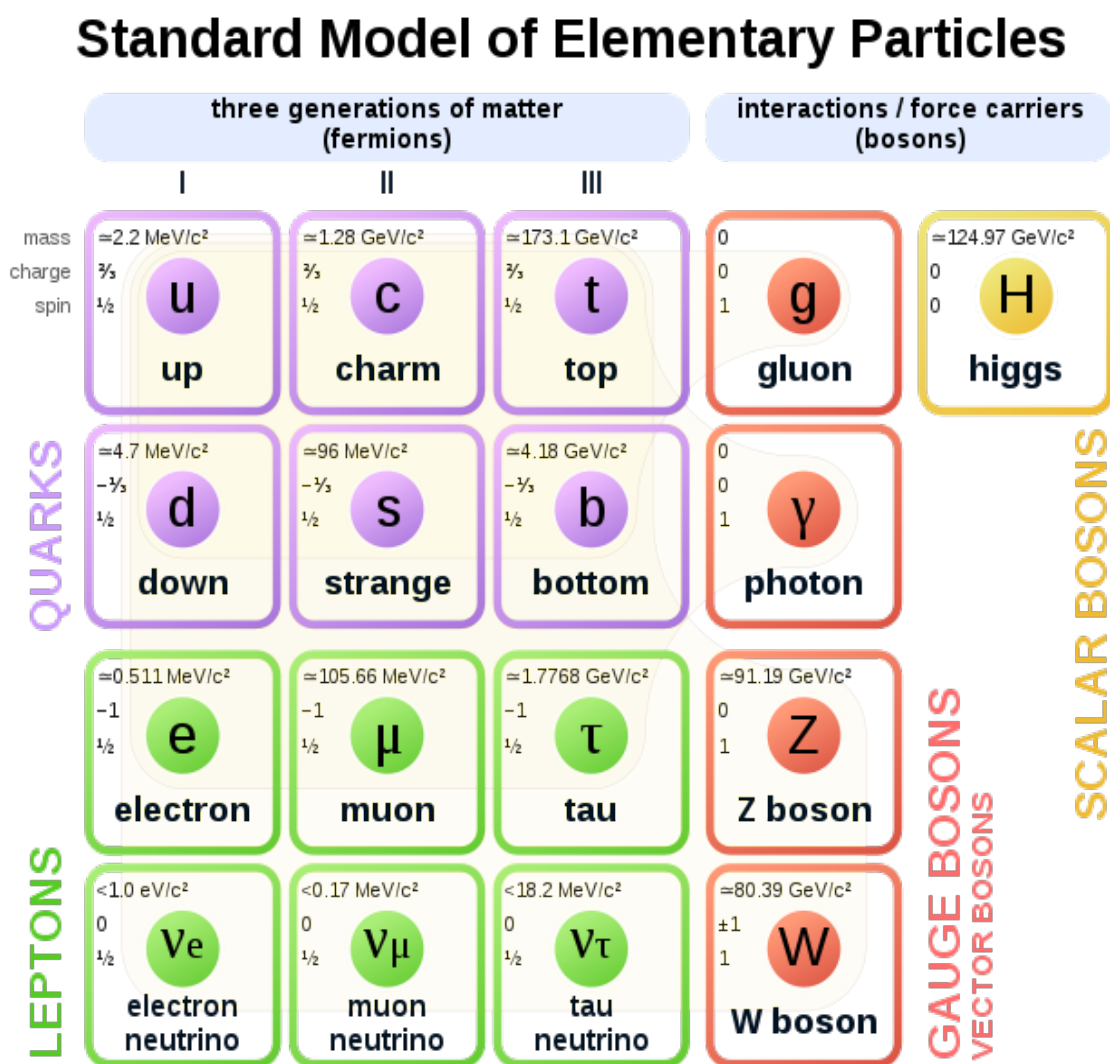


Figure 1.1. Standard model of elementary particles and their Physics properties as invariant mass, spin and electrical charge: 1) Top left in violet: quarks; 2) Bottom left in green: fermions; 3) The brown column: gauge bosons; 4) top right: The scalar boson, the Higgs particle

1.2 Structure of the standard model (SM)

The SM is based on the mathematical formulation of the local gauge invariance of its Lagrangian under the gauge group, from which we derive the respective phenomena of the strong, weak, and electromagnetic interactions, represented from the gauge interactions:

$$G_{SM} = SU(3)_C \times SU(2)_W \times U(1)_Y \quad [20] \quad (1.1)$$

- $SU(3)_C$ invariance [20]: Exploits the existence of gluons (g) as the mediators of the strong force, a concepts integrated within the Quantum Chromodynamics (QCD) theory.
- $SU(2)_W \times U(1)_Y$ invariance [20]: Describes the weak and electromagnetic forces, mediated by the $W^\pm$ and Z^0 bosons and the photon (γ), respectively.

where the subscripts C , W , and Y denote colour, weak isospin and hypercharge, respectively. As mentioned above, the SM formulation does not contain the gravitational interaction, which is negligible at the subnuclear scales. The components (particles) of the matter is described in the SM by the fermion fields. They have spin $\frac{1}{2}$ and their interactions are mediated by spin -1 boson fields. Experimental observations show that twelve physical fermion fields exist and they are structured into two main families:

- Six "quark" fields: Three quarks of electric charge $+\frac{2}{3}$ (up, charm and top quark) and three of electric charge $-\frac{1}{3}$ (down, strange and bottom quark)
- Six "lepton" fields: Three lepton leptons of electric charge -1 (electron e , muon μ and tau τ lepton) and three with charge 0 (electron neutrino ν_e , muon neutrino ν_μ and tau neutrino ν_τ)

To each fermion corresponds an antiparticle with identical properties but opposite quantum numbers.

Quarks

Quarks are subject to all the three forces and, in particular, are the only fermions to possess a "colour" charge to which QCD owes its name. The first family of quarks is composed of the up (u) and down (d) quarks, with a mass in order of MeV. The former has a positive electric charge of $+\frac{2}{3}$ while the latter has a negative electric charge of $-\frac{1}{3}$. Being the lightest quarks, they are stable and compose the ordinary matter. Their counterparts in the second family are the charm (c) and the strange (s) quarks, of masses of about 1.28 GeV and 95 MeV respectively. Finally, the third family is composed of the top (t) and bottom (b) quarks, which masses are about 173 and 4.2 GeV respectively. Because of the QCD colour confinement properties, quarks do not exist as free states but can be experimentally observed only as bound states. Collider experiments thus detect "mesons", that are composed by a quark-antiquark pair, or "baryons", composed by three quarks. The proton and neutron composing the ordinary matter are stable examples of the latter. Mesons and baryons are collectively denoted as hadrons. The creation of hadrons from a single quark produced in a collider experiment is a complex process that takes the name of "hadronization". As its timescales, related to the QCD energy scales, are of the order of 10^{-24} s, hard scatter and hadronization phenomena can be treated separately thanks to a factorization of their effects. The top quark represents an exception in this sense, as its lifetime is so short ($\approx 0.5 \times 10^{-24}$ s) that it decays before bound states can be formed. Quark flavour is conserved in electromagnetic and strong interactions but not in weak ones, as quark mass eigenstates do not correspond to the weak interaction eigenstates. Their mixing is described by the Cabibbo-Kobayashi-Maskawa (CKM) matrix [80].

Leptons

Leptons have no colour charge and are subject only to the electromagnetic and weak forces. The charged leptons of the three families are respectively denoted as the electron (e), muon (μ) and tau lepton (τ). The electron is stable, being the lightest lepton with a mass of 511keV. The muon has a mass of 105.7 MeV and a lifetime of $2.2\mu\text{s}$, that is sufficiently long to consider it as a stable particle at an LHC experiment given the detector size and the typical muon momentum. Finally, the tau lepton has a mass of 1.8GeV and a lifetime of $2.9 \times 10^{-13}\text{s}$, that is instead short enough for observing it only through its decay products. Most notably, the tau is the only lepton that has a sufficient mass to decay semi-leptonically. Each lepton has a corresponding a neutrino, respectively denoted as ν_e , ν_μ and ν_τ . Neutrinos are electrically neutral and they interact with the matter only via the weak force and consequently they are not directly detectable at collider experiments. Little is known about their masses, but the observation of their flavor oscillations prove that they are not zero. The mixing of weak and mass eigenstates is represented by the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix [22].

Strong interaction

QCD is built on the local gauge invariance under the $SU(3)_C$ group, the subscript denoting the relation to the colour charge arising from this symmetry. The free-field Dirac Lagrangian density of a massless spin fermion [53, 54, 55] (the quark fields in this case) is:

$$\mathcal{L} = \bar{\psi}(x) (i\gamma^\mu \partial_\mu) \psi(x) \quad (1.2)$$

where ψ is the fermion field at the space-time coordinate x and γ^μ are the Dirac matrices. The notation $\gamma^\mu \partial_\mu$ is also used in some cases. The following discussion on QCD is valid in presence of a $m\bar{\psi}\psi$ mass term. The reason to consider massless fermions is explained in the context of the electroweak interactions described in the next section. The fermion field transforms in the following way under the $SU(3)_C$ group:

$$\psi(x) \rightarrow e^{ig\frac{\lambda^a}{2}\theta_a(x)} \psi(x) \quad (1.3)$$

where $\frac{\lambda^a}{2}$ are the eight Gell-Mann matrices [79] that generate the group. An important remark is that the derivatives $\partial_\mu \psi(x)$ don't transform in the same way. Hence, the only way for the Lagrangian density (Eq. 1.2) to be invariant under the transformation (1.3) is to redefine the derivative ∂_μ as covariant derivatives:

$$D_\mu = \partial_\mu - igA_\mu^a(x)\frac{\lambda^a}{2} \quad (1.4)$$

where the gauge vector fields $A_\mu^a(x)$ correspond to the eight gluons that mediate the strong force. To satisfy the local gauge invariance of the Lagrangian, the gluon fields must transform as:

$$A_\mu^a \rightarrow A_\mu^a + \partial_\mu \theta^a + gf^{abc}A_\mu^b \theta^c \quad (1.5)$$

The f^{abc} symbols denote the structure constants of the group that are defined from the commutation rules:

$$\left[\frac{\lambda^a}{2}, \frac{\lambda^b}{2} \right] = if^{abc} \frac{\lambda^c}{2}$$

The introduction of the vector fields ensures that the covariant derivative transforms in the same way as the fermion fields and that, consequently, the Lagrangian density is invariant under the local gauge transformation.

The Lagrangian density can be completed with a kinetic term for the gluon fields in the form $-\frac{1}{4}F_a^{\nu\mu}F_{\mu\nu}^a$, where:

$$F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + gf^{abc}A_\mu^b A_\nu^c \quad (1.6)$$

The complete QCD Lagrangian density thus reads:

$$\mathcal{L}_{QCD} = \bar{\psi}(x) (i\gamma^\mu \partial_\mu) \psi(x) - g\bar{\psi}(x)\gamma^\mu \frac{\lambda^a}{2} \psi A_\mu^a - \frac{1}{4}F_a^{\nu\mu}F_{\mu\nu}^a \quad (1.7)$$

with a summation over all quark fields being implied. The first term is the same as in the original Lagrangian density and represents the free-field propagation of the quark. The second one stems from the introduction of the covariant derivative and represents the interaction of the quark with the vector field A_μ . The strength of the interaction is parameterized by the constant g , usually redefined as the strong coupling constant:

$$\alpha_s = \frac{g^2}{4\pi} \quad (1.8)$$

The third term has been introduced as the kinetic term of the vector field. The generators of the $SU(3)_C$ group do not commute and therefore the structure constants f^{abc} are not all zero. As a consequence, the $gf^{abc}A_\mu^b A_\nu^c$ terms in Eq. 1.6, when inserted in the kinetic term of Eq. 1.7, result in cubic and quartic self-interactions of the gluon fields. Such interactions between the force mediators are a general property of non-abelian gauge theories. Requiring the local gauge invariance led to the introduction of gauge bosons (the gluons) and to the description of their interactions with the fermion fields (the quarks). Choosing the $SU(3)_C$ group implies the presence of eight generators, the eight gluons, that are mathematically described by the adjoint representation of the group (8), and differ by the colour charge that they carry. Quarks and anti-quark are instead described in the simplest non-trivial representations of $SU(3)_C$, 3 and $\bar{3}$, which explains the three colour quantum numbers of quarks. It is important to remark that explicit mass terms in the form $A_\mu^a A_\mu^a$ for the gauge bosons in Eq. 1.7 would break the gauge invariance of the Lagrangian.

Electroweak interaction

Electroweak interactions are explained in the SM with the same local gauge invariance mechanism as strong interactions by imposing a symmetry under the $SU(2)_L \times U(1)_Y$ group. Experimental observations show that parity is violated by weak interactions, which is accounted for in the theoretical description by assigning different interactions to fermions of opposite chiralities. The left and right chiral components of a field are defined from the Dirac's (gamma) matrices [53]:

$$\gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3\gamma^4 \quad (1.9)$$

This is used to define left and right chirality projection operators as $\frac{1-\gamma^5}{2}$ and $\frac{1+\gamma^5}{2}$, respectively. In the limit of a massless particle, the chirality corresponds to the helicity, that is defined as the normalized projection of the spin vector onto the spatial momentum vector. The $SU(2)_L$ gauge group is a non-abelian group to which the weak isospin quantum number (I_3) is associated. The gauge invariance under this group results in the presence of 3 gauge fields W_μ^i ($i = 1, 2, 3$). Fermion fields of left chirality are represented by $SU(2)_L$ doublets while fermions of right chirality are $SU(2)_L$ singlets and do not interact with the W_μ^i fields. The $U(1)_Y$ gauge group is abelian and is associated to the weak hypercharge Y . A single gauge field, denoted as B_μ , results from the $U(1)_Y$ local gauge invariance, and interacts with both ψ_L and ψ_R . The $U(1)_Y$ group can be seen in close analogy with the electromagnetic $U(1)_{em}$ group associated to the electric charge Q , but is distinct from it. The Gell-Mann-Nishijima formula [47] determines the relation with the electric charge:

$$Q = I_3 + \frac{Y}{2} \quad (1.10)$$

Fields can therefore be represented as one doublet and two singlets:

$$\Psi_L = \frac{1 - \gamma^5}{2} \begin{pmatrix} \psi \\ \psi' \end{pmatrix} \quad (1.11)$$

The fields ψ and ψ' represent either the neutrino and charged lepton fields or the up and down-type quark fields. The two sectors are however separated and neither the strong nor the electroweak interactions can transform quark fields into lepton fields or vice-versa: a direct lepton-quark coupling is not predicted in the SM. Under this notation, the Lagrangian density can be written as a sum of three terms:

$$\mathcal{L} = i\bar{\Psi}_L \not{D} \Psi_L + i\bar{\psi}_R \not{D} \psi_R + i\bar{\psi}'_R \not{D} \psi'_R \quad (1.12)$$

where the covariant derivative, that is introduced to ensure the gauge invariance, is defined as:

$$D_\mu = \partial_\mu - igW_\mu^i T_i - ig' \frac{Y_\psi}{2} B_\mu \quad (1.13)$$

$$T_i = \frac{\sigma_i}{2}$$

where σ_i are the Pauli matrices, the generators of the $SU(2)_L$ group for the Ψ_L field and 0 for ψ_R and ψ'_R . The coupling constants g and g' define the strength of the interactions. More interestingly, the Lagrangian density Eq. 1.12 can be rewritten as:

$$\mathcal{L} = \mathcal{L}_{kin} + \mathcal{L}_{CC} + \mathcal{L}_{NC} \quad (1.14)$$

with the three parts defined respectively as:

$$\begin{aligned} \mathcal{L}_{kin} &= i\bar{\Psi}_L \not{D} \Psi_L + i\bar{\psi}_R \not{D} \psi_R + i\bar{\psi}'_R \not{D} \psi'_R \\ \mathcal{L}_{CC} &= \frac{g}{\sqrt{2}} W^+ \bar{\psi}_L \gamma^\mu \psi'_L + \frac{g}{\sqrt{2}} W^- \bar{\psi}'_L \gamma^\mu \psi_L \\ \mathcal{L}_{NC} &= \frac{g}{\sqrt{2}} W_\mu^3 \left[\bar{\psi}_L \gamma^\mu \psi_L - \bar{\psi}'_L \gamma^\mu \psi'_L \right] \\ &\quad + \frac{g'}{\sqrt{2}} B_\mu \left[Y_{\Psi_L} \left(\bar{\psi}_L \gamma^\mu \psi_L + \bar{\psi}'_L \gamma^\mu \psi'_L \right) + Y_{\psi_R} \bar{\psi}_R \gamma^\mu \psi_R + Y_{\psi'_R} \bar{\psi}'_R \gamma^\mu \psi'_R \right] \end{aligned} \quad (1.15)$$

and

$$\begin{aligned} W_\mu^\pm &= \frac{1}{\sqrt{2}} \left(W_\mu^1 \mp iW_\mu^2 \right) \\ \sigma_\mu^\pm &= \frac{1}{\sqrt{2}} \left(\sigma_\mu^1 \mp i\sigma_\mu^2 \right) \end{aligned} \quad (1.16)$$

where the the annotation CC stand for charged current, while NC stands for neutral current. Therefore, there is a charged current interaction that couples the fields ψ_L and ψ'_L and is mediated by the $W^\pm$ fields, that correspond to the charged weak bosons interacting with fermions. A neutral current interaction also exists, although neither the W_μ^3 nor the B_μ fields can be interpreted as the photon field since they couple to neutral fields. However, we can express them in terms of the physical Z_μ field (the neutral Z boson field) and the A_μ field (the photon field) through a linear superposition parameterized with the Weinberg angle θ_W :

$$\begin{aligned} B_\mu &= A_\mu \cos\theta_W - Z_\mu \sin\theta_W \\ W_\mu^3 &= A_\mu \sin\theta_W + Z_\mu \cos\theta_W \end{aligned} \quad (1.17)$$

Upon substituting this expression in Eq. 1.16, two neutral current interactions appear, the one with the A_μ field being determined by a coupling of strength:

$$g \sin \theta_W I_3 + g' \cos \theta_W \frac{Y}{2} \quad (1.18)$$

The unification of the weak and electromagnetic forces is completed if we require this expression to be equal to the coupling constant of the photon field eQ . We can arbitrarily set $Y_{\Psi_L} = -1$ as the hypercharge only appears multiplied by g' , and solve the equation by using $Q = 0$ for the neutrino fields and $Q = -1$ for the lepton fields. Upon substitution, Eq. 1.18 is obtained together with the following relation between the coupling constants:

$$g \sin \theta_W = g' \cos \theta_W = e \quad (1.19)$$

The Lagrangian density of Eq. 1.16 only needs to be completed with the kinetic terms of the gauge fields. The full electroweak Lagrangian can be expressed in a compact form as:

$$\mathcal{L}_{EWK} = i\bar{\Psi}_L \not{D} \Psi_L + i\bar{\psi}_R \not{D} \psi_R + i\bar{\psi}'_R \not{D} \psi'_R - \frac{1}{4} B^{\mu\nu} B_{\nu\mu} - \frac{1}{4} W_i^{\mu\nu} W_{\nu\mu}^i \quad (1.20)$$

The field strength tensors are:

$$\begin{aligned} B_{\nu\mu} &= \partial_\nu B_\mu - \partial_\mu B_\nu \\ W_{\nu\mu}^i &= \partial_\nu W_\mu^i - \partial_\mu W_\nu^i + g \varepsilon^{abc} W_\mu^b W_\nu^c \end{aligned} \quad (1.21)$$

Equation 1.20 contains the free fermion Dirac Lagrangian and the charged and neutral current interactions of the fermions previously discussed. The kinetic terms of the W_μ^i field, once developed, show a rich structure of self-interactions of the gauge bosons. Trilinear (ZWW , γWW) and quadrilinear ($ZZWW$, $\gamma\gamma WW$, γZWW , $WWWW$) interactions are predicted by the theory. As already observed for the strong interactions, explicit mass terms of the gauge fields would break the gauge invariance. Direct fermion mass terms are also not allowed, because they are not invariant under the gauge transformation: the left and right chiralities of the fields transform differently under $SU(2)_L \times U(1)_Y$, while the mass can be decomposed as:

$$m\psi\bar{\psi}' = m(\bar{\psi}_R\psi_L + \bar{\psi}_L\psi_R) \quad (1.22)$$

1.3 The Brout-Englert-Higgs mechanism

All began in the early 1960-s from a proposal of Philip Warren Anderson [7] in 1962, where he proposed a mechanism to explain symmetry breaking in superconductivity in the late 1950s and a 1960 paper by Yoichiro Nambu that discussed its application within particle physics. Then, the Brout-Englert-Higgs mechanism was formulated in 1964 independently and almost simultaneously by three independent groups: Robert Brout and François Englert (Universite Libre de Bruxelles) [59], Peter Higgs (University of Edinburgh) [92], Gerald Guralnik, Carl Richard Hagen, and Tom Kibble (Imperial College) [85], as a solution to explain the mass generation without "breaking" the gauge theory and the fermion masses. The mechanism is based on the concept of spontaneous symmetry breaking, a phenomenon observed in Nature, whenever individual ground states of a system do not satisfy the symmetries of the system itself. The spontaneous symmetry breaking is realized through the introduction of a complex scalar doublet of fields:

$$\Phi = \begin{pmatrix} \phi_+ \\ \phi_0 \end{pmatrix}$$

where the field must be scalar to satisfy space isotropy. The vacuum expectation value (VEV) v would be frame-independent and constant to satisfy space homogeneity, meanwhile the hypercharge of the field is $Y_\Phi = 1$. The covariant derivative is:

$$D_\mu = \partial_\mu - igW_\mu^i \frac{\sigma}{2} - \frac{1}{2}ig' B_\mu \quad (1.23)$$

and thus, the The Brout-Englert-Higgs Langrangian is:

$$\begin{aligned} \mathcal{L}_{BEH} &= (D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\Phi^\dagger \Phi) \\ V(\Phi^\dagger \Phi) &= -\mu \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \quad \text{where } \mu, \lambda > 0 \end{aligned} \quad (1.24)$$

All the doublets that satisfy the condition: $|\Phi|^2 = \frac{\mu^2}{2\lambda} \equiv \frac{v^2}{2}$ are minima of these potential, and are connected through gauge transformations that change the phase of the field Φ but not its modulus.

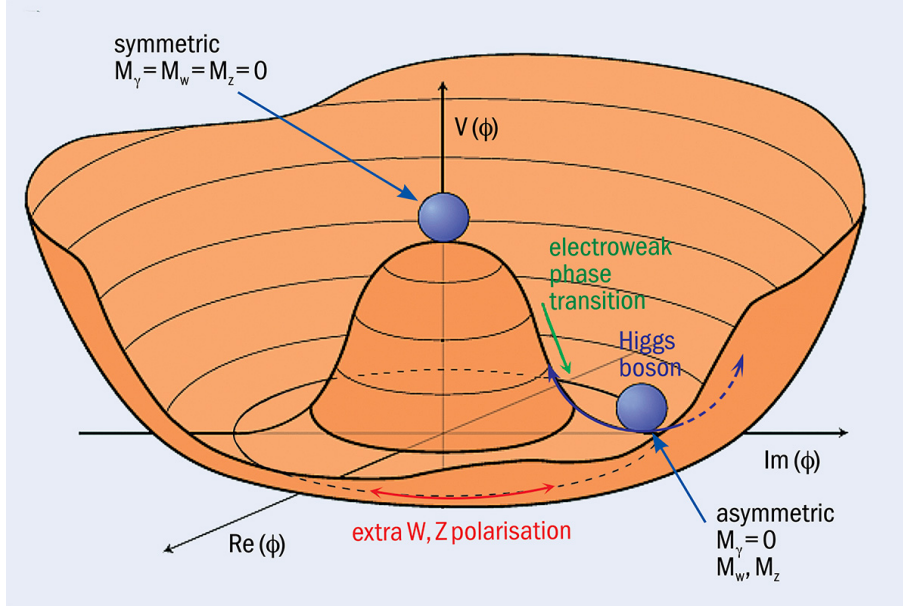


Figure 1.2. . Representation of the Higgs potential: The top of the "mexican hat" represents the vacuum state, describing the symmetry of the Higgs potential, while the ground state is not symmetric. In the Higgs mechanism, the rotational degree of freedom along the rim becomes the longitudinal polarization of the $W^\pm$ and Z^0 bosons, which thereby acquire mass. This image is adapted from Ref. [105]

If the symmetry is spontaneously broken into the ground state parallel to the ϕ_0 component of the doublet, than, this specific ground state is still invariant under the $U(1)_{em}$ symmetry group. As a consequence, the field expansion around this minimum is written as:

$$\Phi(x) = \frac{1}{\sqrt{2}} \exp \left[\frac{i\sigma_i \theta^i(x)}{v} \right] \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \quad (1.25)$$

This corresponds to the presence of a scalar real massive field H and three massless fields $\theta^i(x)$. The latter are expected as consequence of the Goldstone theorem [84]. Such massless bosons are not observed in Nature. They can be removed with an $SU(2)_L$ transformation that consists in the choice of a specific gauge called "unitary gauge":

$$\Phi(x) \rightarrow \Phi'(x) = \exp \left[-\frac{i\sigma_i \theta^i(x)}{v} \right] \Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \quad (1.26)$$

After the transformation, only the real scalar field $H(x)$ remains and its quanta correspond to a new physical massive particle, the Higgs boson (H):

The Brout-Englert-Higgs Lagrangian

$$\begin{aligned} \mathcal{L}_{BEH} = & \frac{1}{2} \partial_\mu H \partial^\mu H - \frac{1}{2} (2\lambda v^2) H^2 \\ & + \left[\left(\frac{gv}{2} \right)^2 W^{\mu+} W_\mu^- + \frac{1}{2} \frac{(g^2 + g'^2) v^2}{4} Z^\mu Z_\mu \right] \left(1 + \frac{H}{v} \right)^2 \\ & + \lambda v H^3 + \frac{\lambda}{4} H^4 \end{aligned} \quad (1.27)$$

The Lagrangian Eq. 1.27 has three important terms, where the first one represents the evolution of the Higgs scalar field with a mass:

$$m_H = \sqrt{2\lambda v^2} = \sqrt{2\mu^2} \quad (1.28)$$

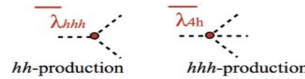
meanwhile the second part includes the mass terms of the weak bosons:

$$\begin{aligned} m_W^2 &= \frac{g^2 v^2}{4} \\ m_Z^2 &= \frac{(g^2 + g'^2) v^2}{4} = \frac{m_W^2}{\cos^2 \theta_W} \end{aligned} \quad (1.29)$$

The massless Goldstone bosons due to the Brout-Englert-Higgs mechanism are absorbed within the Higgs field as additional degree of freedoms of $W^\pm, Z^0$, after they are removed with the transformation in Eq. 1.25, respective to their longitudinal polarizations: the mechanism confers mass to the weak bosons.

The third line implements prediction of cubic and quartic self-interaction coupling $\lambda_{3H,4H}$ of the Higgs boson. In the SM, the Higgs self-couplings are uniquely determined by the structure of the scalar potential, where in contrast to the weak boson self-interactions that have a gauge nature. The BEH potential can be rewritten in terms of a trilinear and a quadrilinear coupling [70]. The Higgs field potential energy function, whose minimum gives the nonzero expectation value that triggers spontaneous symmetry breaking, is almost completely unconstrained and requires a significant non-linear self-interaction of the Higgs field, with the potential taking the form:

$$V = \frac{1}{2} m_H^2 H^2 + \lambda_{HHH} v H^3 + \frac{1}{4} \lambda_{HHHH} H^4 - \frac{\lambda}{4v^4} \quad (1.30)$$



$$\lambda_{HHH} = \lambda_{HHHH} = \frac{m_H^2}{2v^2} \approx 0.13$$

Experimentally measuring λ_3, λ_4 is thus a crucial test of the mechanism of electroweak symmetry breaking, thus are responsible for the mass of the Higgs boson itself. A measurement of λ_3 requires double Higgs boson production, while λ_4 will be only proved first in the tri-linear production of Higgs bosons.

An important part of Eq. 1.27 is a constant term in the Lagrangian density, which contributes to the vacuum energy and it is related to the cosmological constant that determines the curvature of the Universe. The value of this constant is predicted in the SM but it is not compatible with astronomical observations, a paradox that needs further understanding in the quantum theory of gravity, to specifically understand why the Higgs field vacuum energy density is so high. In the Lagrangian density term of Eq. 1.27 the energy scale of the electroweak symmetry breaking is represented the vacuum expectation value (VEV) v and can be determined from the Fermi constant G_F that is precisely estimated for the μ (muon) lifetime:

$$\frac{G_F}{\sqrt{2}} = \left(\frac{g}{2\sqrt{2}} \right)^2 \frac{1}{m_W^2} \Leftrightarrow v = \sqrt{\frac{1}{\sqrt{2}G_F}} \approx 246 \text{ GeV} \quad (1.31)$$

The Higgs field generates the mass terms for the fermions through a Yukawa interaction that couples the left and right chiral fields, which arise from the Yukawa interactions described by the Yukawa Lagrangian density [4] as following:

$$-\mathcal{L}_{Yukawa} = \sum_{a,b} \left[h_{a,b}^{(u)} \bar{q}_{aL} \hat{\phi} u_{bR} + h_{a,b}^{(d)} \bar{q}_{aL} \phi d_{bR} + h_{a,b}^{(l)} \bar{\Psi}_{aL} \phi l_{bR} \right] + h.c \quad (1.32)$$

where $\hat{\phi} = i\tau_2 \phi^*$ are the charge-conjugated Higgs field, the $h^{(u,d,l)}$ matrices are called Yukawa couplings, which can be diagonalized, meanwhile the indices a, b stand for the generations indices. The Yukawa Lagrangian density is Lorentz and gauge invariant and renormalizable, so that it can be included in the SM Lagrangian density. If we substitute the non-zero vacuum expectation values:

$$\phi_0 = \sqrt{\frac{\mu^2}{2\lambda}} = \frac{v}{\sqrt{2}} \quad (1.33)$$

the following mass terms for up and down quarks as well as the charge leptons are generated, as such, we can express the Lagrangian in Eq. 1.32 as shown below:

$$\mathcal{L}_{mass} = \sum_{a,b} \left[\bar{u}_{aL} M_{ab}^{(u)} u_{bR} + \bar{d}_{bR} M_{a,b}^{(d)} d_{bR} + \bar{l}_{bL} M_{a,b}^{(l)} l_{bR} \right] + h.c \quad (1.34)$$

where the sum iterates over both types of fermions and the mass terms, thus would be:

$$M_{a,b}^{(f)} = h_{a,b}^{(f)} \frac{v}{\sqrt{2}} \quad f = u, d, l \quad (1.35)$$

According to the quark and/or lepton basis we would like to inspect, the coupling matrices $h^{(u,d,l)}$ can be diagonalized so that we have for instance $u_a^0 = u_a$ and $l_a^0 = l_a$. The $M^{(d)}$ is however, a complex non-diagonal matrix in this basis and can be diagonalized by the following bi-unitary transformation:

$$V_L M^{(d)} V_R^\dagger = D^{(d)} \quad (1.36)$$

The Brout-Englert-Higgs mechanism gives an important contribute to the electroweak theory thus explaining how massless Goldstone bosons becomes massive after they are absorbed as degree of freedoms of the vector boson fields. The Higgs boson couples the left and right chiral components of the fermion fields in a Yukawa interaction, determining the fermion masses with a purely quantum-relativistic mechanism. Though, the SM still does not explain the origin of these couplings and the hierarchy of the three fermion families (electron $e^\mp$, muons $\mu^\mp$ and taus $\tau^\mp$).

1.3.1 The experimental observation of the single Higgs (H) boson event at the LHC

The experimental observation and final prove of the BEH mechanism was public in July 2012 with the discovery of a new scalar boson of a mass of approximately 125 GeV announced by the ATLAS and CMS collaboration at CERN [135][119][12]. The discovery was performed with the data collected at $\sqrt{s} = 8$ TeV (Run I period) and led by the high resolution of two main decay channels:

$$\begin{aligned} (CMS) H &\rightarrow ZZ^* \rightarrow l^+ l^- l'^+ l'^- \\ (ATLAS) H &\rightarrow \gamma\gamma \end{aligned} \quad (1.37)$$

where: $l^\pm = e^\pm, \mu^\pm$. The existence of this scalar particle was consistently established and further confirmed also with the new data collected at $\sqrt{s} = 13$ TeV during the Run II, as shown in Figure 1.3.

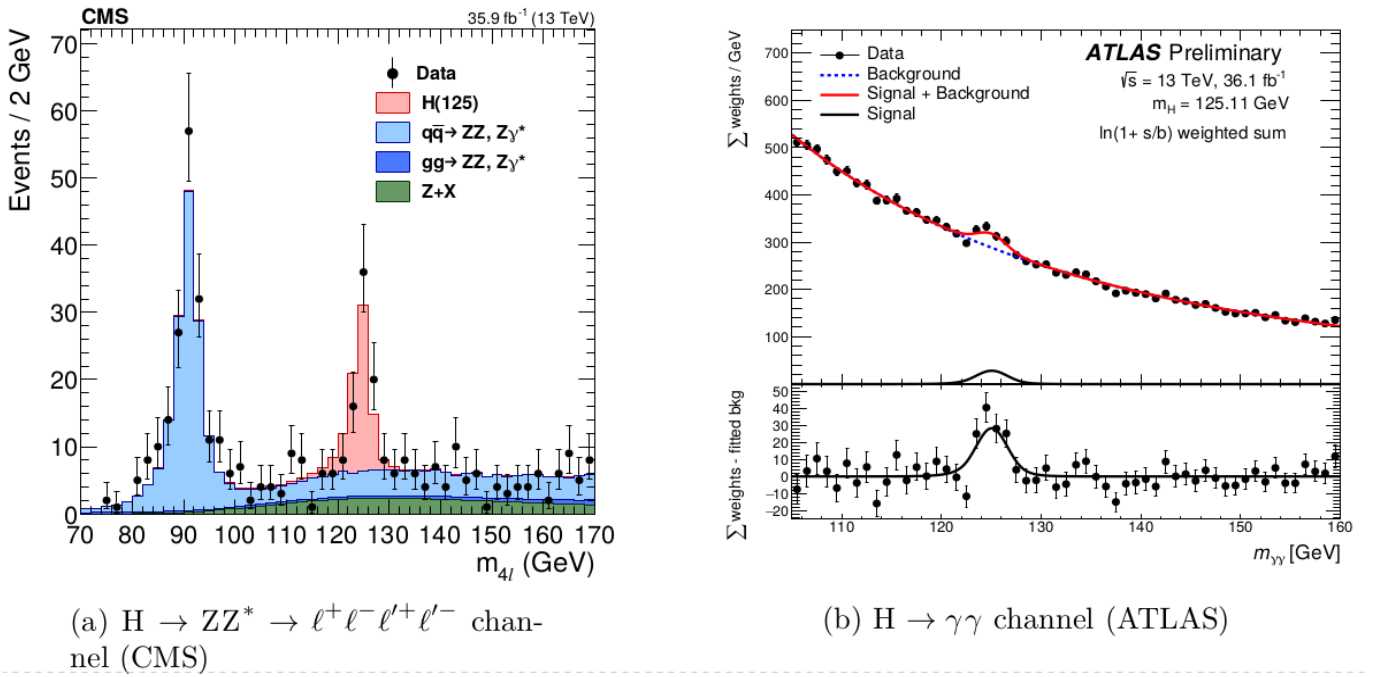


Figure 1.3. Four-lepton invariant mass (left)[146] and photon pair invariant mass (right)[121] distributions in events collected at $\sqrt{s} = 13$ TeV

Several mechanisms contribute to the production of a Higgs boson at a pp collider:

- The dominant one, with a cross section of about 49 pb, is the gluon fusion (ggF) production, that proceeds through a heavy quark loop;
- The vector boson fusion (VBF) mechanism is the second most frequent mechanisms, about a order smaller than ggF . The Higgs boson is produced in association with a jet pair of large invariant mass;
- The production in association with a single vector boson (VH, $V = W^\pm, Z^0$) is the third mechanism in terms of.
- The production of the Higgs boson in association with a pair of top quark ($t\bar{t}H$) or a single top quark (tH). Both mechanisms, $t\bar{t}H$ and tH are very important because they allow a direct determination of the magnitude of the top Yukawa coupling y_t

In Table 1.1 and in Figure 1.4 are reported the total cross sections of the Higgs boson production via the gluon fusion, vector-boson fusion, double Higgs-strahlung and double Higgs bremsstrahlung off top quark mode as well as the branching ratios of the the possible decay modes of the Higgs boson.

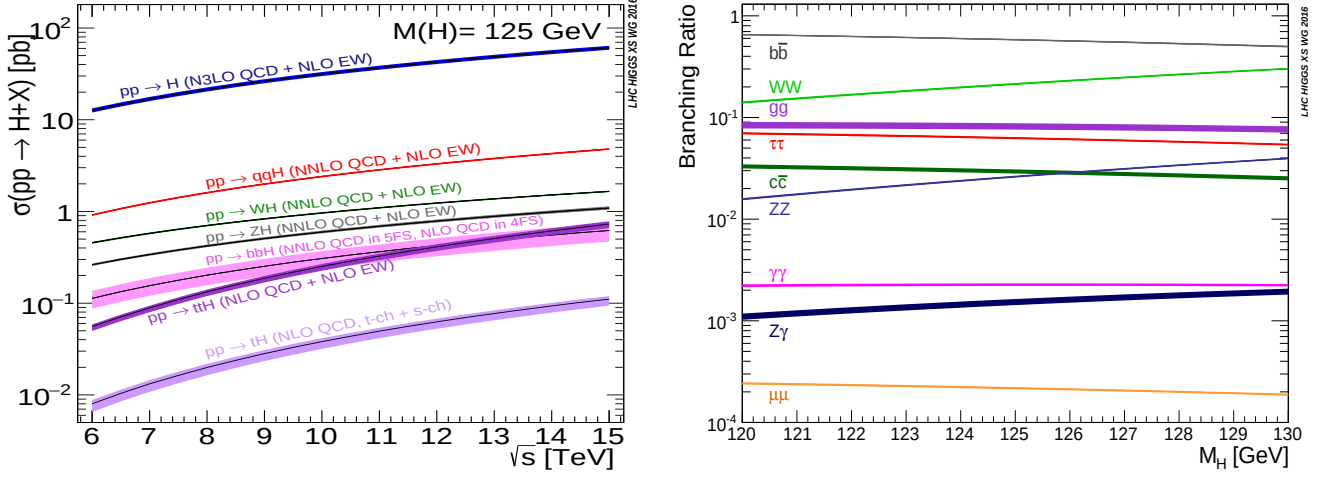


Figure 1.4. Left: Total production cross sections for the single Higgs (H) within the SM via gluon fusion, vector-boson fusion, double Higgs-strahlung and double Higgs bremsstrahlung off top quarks. The size of the bands shows the total uncertainties originating from the scale dependence and the $PDF + \alpha_s$ uncertainties [66]. Right: Standard Model Higgs boson decay branching ratios (CERN Yellow Report 4) [36]

Decay Mode	$\mathcal{B}$ [%]
$H \rightarrow b\bar{b}$	$58.09^{+0.72}_{-0.73}$
$H \rightarrow W^{\pm}W^{\mp*}$	$21.52^{+0.33}_{-0.33}$
$H \rightarrow gg$	$8.16^{+0.42}_{-0.42}$
$H \rightarrow \tau^+\tau^-$	$6.27^{+0.10}_{-0.10}$
$H \rightarrow c\bar{c}$	$2.88^{+0.16}_{-0.06}$
$H \rightarrow ZZ^*$	$2.641^{+0.040}_{-0.040}$
$H \rightarrow \gamma\gamma$	$0.227^{+0.0047}_{-0.0047}$
$H \rightarrow Z\gamma$	$0.1541^{+0.0090}_{-0.0090}$
$H \rightarrow \mu^+\mu^-$	$0.02171^{+0.00036}_{-0.00037}$

Table 1.1. Branching fractions of the main Higgs boson decay modes for a SM Higgs boson of mass $m_H = 125.09$ GeV [167]. Theoretical uncertainties combine the uncertainties on the Higgs boson partial width, on the value of α_s , and on the quark masses.

The Run I Higgs boson discovery was performed inclusively for all the production mechanisms. The combination of the high-resolution channels between the ATLAS and CMS experiments resulted in a precise determination of m_H of:

$$m_H = 125.09 \pm 0.21(\text{stat.}) \pm 0.11(\text{syst.})\text{GeV} \quad (1.38)$$

With the observation of m_H and the measurement of its mass, the last important parameter of the SM has been determined. However, the observation of this new particle only represents the first step in the exploration of the EWSB. The efforts of the experimental community are thus focusing on the characterization of this boson. The measurements of its couplings (Figure 1.5) via exclusive production modes and decay channels, of its spin-parity, and of its differential production cross section, need to be thoroughly investigated to verify that they correspond precisely to the SM predictions.

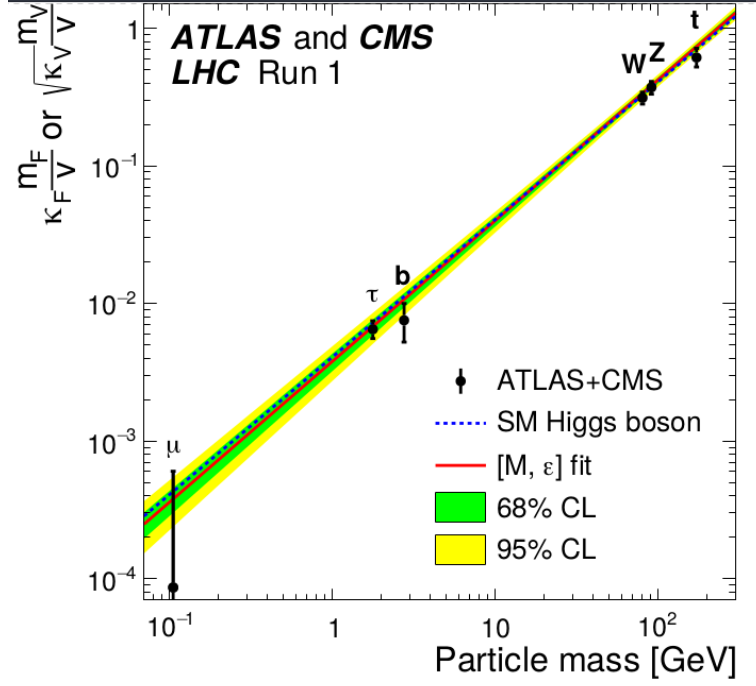


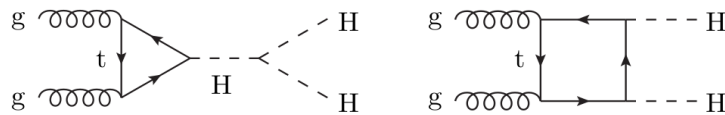
Figure 1.5. Normalized Higgs boson coupling constant as a function of the boson or fermion mass: Best fit values as a function of particle mass for the combination of ATLAS and CMS data, with parameters defined as $k_F m_F/v$ for the fermions, and as $\sqrt{k_V} m_V/v$ for the weak vector bosons, where $v = 246$ GeV is the vacuum expectation value of the Higgs field. The dashed (blue) line indicates the predicted dependence on the particle mass in the case of the SM Higgs boson. The solid (red) line indicates the best fit result to the $[M, \epsilon]$ phenomenological model [120] with the corresponding 68% and 95% CL bands.

1.4 The double Higgs production

The main Higgs pair production mechanisms predicted from the SM are five [75][13] and they are listed below in the decreasing order of their cross sections [36] with the respective Feynman's diagrams at the leading order (LO) accuracy production mode:

1) Gluon fusion production $gg \rightarrow HH$ (ggF-HH):

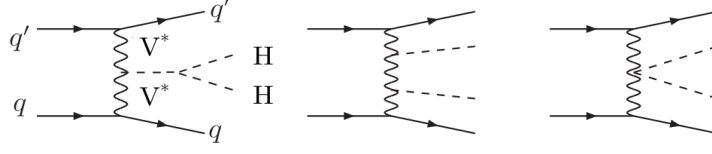
It is the dominant HH generation mode [36] in terms of cross section σ_{HH} . This production contains the trilinear Higgs boson self-coupling λ_{HHH} and the radiation of two on-shell Higgs bosons from a heavy quark loop, where we have to take in consideration that the contribution of the b-quark loop is smaller than 1% at leading order and can be neglected, given current accuracy of theoretical computations and experimental sensitivity. The total cross section at leading order for the gluon gluon fusion depends from the trilinear Higgs boson self-coupling λ_{HHH} and the Yukawa top quark couplings y_t .



$$\sigma_{14TeV}^{NNLL,NNLO}[fb] = 36.69^{+2.1\%}_{-3.9\%}(scale) \pm 3.1\%(PDF) \pm 2.1\%(\alpha_s) \pm 2.1\%(top)$$

2) Vector boson fusion (VBF) production $q\bar{q} \rightarrow jjHH$ (VBF-HH):

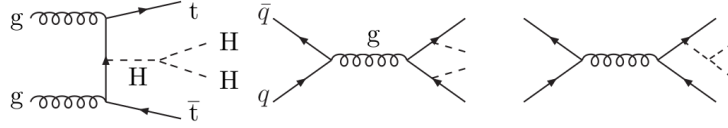
For the VBF-HH production, the total cross section at leading order of it depends from the trilinear and quadrilinear Higgs boson self-coupling λ_{HHH} , λ_{HHHH} of a Higgs pair to a vector boson pair as well as on the single Higgs boson coupling to vector bosons [36]. The two final state jets provide a clear signature of this production in terms of discrimination signal-background for the event selection, despite the one order of magnitude smaller cross section in comparison to $ggF - HH$ production.



$$\sigma_{14TeV}^{N3LO}[fb] = 2.055^{+0.03\%}_{-0.04\%}(scale) \pm 2.1\%(PDF + \alpha_s)$$

3) Top quark pair associated production $q\bar{q}/gg \rightarrow t\bar{t}HH$ ($t\bar{t}HH$):

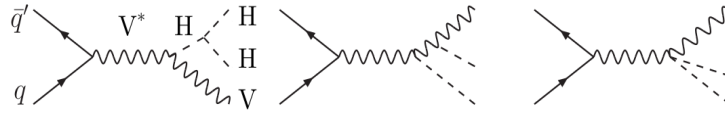
This possible production at LO-accuracy is an extension of the 2018 observed Physics event $t\bar{t}H$ [118], where either we have a radiation or a production of a pair of Higgs bosons from a pair of top quarks. Computations calculated higher cross sections of $t\bar{t}HH$ at high transverse momenta and high center-of-mass energies, exceeding those of the $VBF - HH$ cross sections.



$$\sigma_{14TeV}^{NLO}[fb] = 0.949^{+1.7\%}_{-4.5\%}(scale) \pm 3.1\%(PDF + \alpha_s)$$

4) Vector boson associated production $q\bar{q} \rightarrow VHH$, $V = W^\pm, Z$:

This production mode contains the same Physics generation and couplings as the VBF-HH plus an on-shell vector boson in the final state.

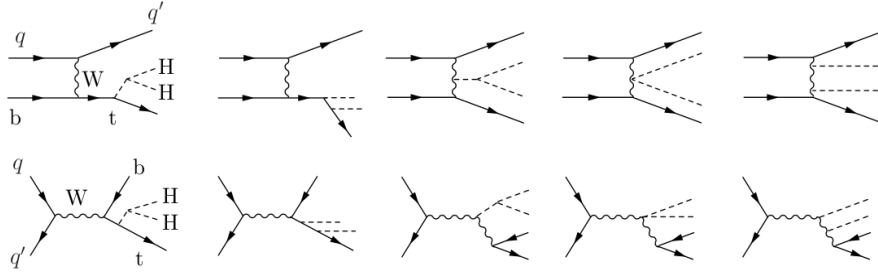


$$\sigma_{14TeV}^{NNLO}[fb] = (\sigma_{W^+}, \sigma_{W^-}) = 0.369(0.173)^{+0.33\%(1.2)}_{-0.39\%(-1.3)}(scale) \pm 2.1\%(2.7)(PDF + \alpha_s)$$

5) Single top quark associated production $q\bar{q} \rightarrow t j HH$:

This production mode is rich in term of theoretical computations. It is associated either by a t - or a s -channel, as shown in the diagrams Feynman' diagrams below and the t -channel diagrams are also called as the 5F scheme [97]. The total cross sections of this generation mode is too small for the current experimental sensitivity at the LHC.

$$\sigma_{14TeV}^{NLO}[fb] = 0.0367^{+4.2\%}_{-1.8\%}(scale) \pm 4.6\%(PDF + \alpha_s)$$



The gluon fusion $gg \rightarrow HH$ is the dominant Higgs (H) pair production mode. Therefore, SM HH MC production will be considered to be used for both SM and BSM variants. There are two different variants of double Higgs (H) production:

1. Non-resonant HH production
2. Resonant HH production

1.4.1 The SM non-resonant production mechanism of the HH

The non-resonant HH production is derived from the extension of the SM Lagrangian with dimension 4 and 6 operators is one of the BSM implementation in order to enhance the HH cross section, which allows and requires anomalous Higgs (H) couplings constants as shown in Eq. 1.2[71][117] (for more details, check [Appendix A](#)), meanwhile the BSM couplings highlighted in red are shown in [Figure 1.6](#).

The EFT Lagrangian for HH processes

$$\begin{aligned}
 \mathcal{L}_{HH} = & \frac{1}{2} \partial_\mu \partial^{\mu\nu} h - \frac{1}{2} m_h^2 h^2 - \mathbf{k}_\lambda \lambda_{SM} v h^3 \\
 & - \frac{m_t}{v} \left(v + \mathbf{k}_t h + \frac{\mathbf{c}_2}{v} hh \right) (\bar{t}_L t_R + h.c.) \\
 & + \frac{\alpha_s}{12\pi v} \left(\mathbf{c}_{1g} h - \frac{\mathbf{c}_{2g}}{2v} hh \right) G_{\mu\nu}^A G^{A\mu\nu}
 \end{aligned} \tag{1.39}$$

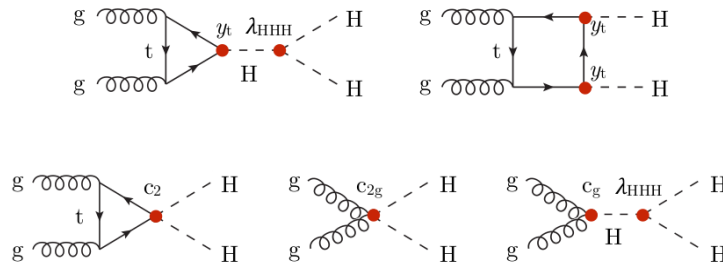


Figure 1.6. The Feynman diagrams of the HH production via gluon fusion at LO, where the BSM couplings are highlighted in red.

In [Table 1.2](#), we have listed the computations of the signal cross sections (in fb) for double Higgs (H) production at different centre-of-mass energies, including the available QCD corrections according to the

recommendations of the LHC Higgs Cross Section Working Group [168]. The cross section are computed at the highest theoretical accuracies at the current time. In case of the $ggF - HH$ the non-resonant double Higgs production cross section σ_{HH} , which is our main production mode of interest, is computed at next-to-next-leading-order (NNLO) of the theoretical perturbative QCD calculations including next-to-next-leading-logarithm (NNLL) corrections and finite top quark mass effects at next-to-leading-order (NLO).

The theoretical uncertainties are calculated by taking in consideration the uncertainties related to Parton Distributions Functions (PDF), strong coupling interaction constant $\alpha_s(Q^2)$, the QCD factorization and renormalization scales (μ_f, μ_r) and the unknown effects from the finite top quark mass. The production cross sections [112] are shown in Figure 1.7.

$\sqrt{s}$	13 TeV	14 TeV	27 TeV	100 TeV
ggF HH	$31.05^{+2.2\%}_{-5.0\%} \pm 3.0\%$	$36.69^{+2.1\%}_{-4.9\%} \pm 3.0\%$	$139.9^{+1.3\%}_{-3.9\%} \pm 2.5\%$	$1224^{+0.9\%}_{-3.2\%} \pm 2.4\%$
VBF HH	$1.73^{+0.03\%}_{-0.04\%} \pm 2.1\%$	$2.05^{+0.03\%}_{-0.04\%} \pm 2.1\%$	$8.40^{+0.11\%}_{-0.04\%} \pm 2.1\%$	$82.8^{+0.13\%}_{-0.04\%} \pm 2.1\%$
ZHH	$0.363^{+3.4\%}_{-2.7\%} \pm 1.9\%$	$0.415^{+3.5\%}_{-2.7\%} \pm 1.8\%$	$1.23^{+4.1\%}_{-3.3\%} \pm 1.5\%$	$8.23^{+5.9\%}_{-4.6\%} \pm 1.7\%$
W^+HH	$0.329^{+0.32\%}_{-0.41\%} \pm 2.2\%$	$0.369^{+0.33\%}_{-0.39\%} \pm 2.1\%$	$0.941^{+0.52\%}_{-0.53\%} \pm 1.8\%$	$4.70^{+0.90\%}_{-0.96\%} \pm 1.8\%$
W^-HH	$0.173^{+1.2\%}_{-1.3\%} \pm 2.8\%$	$0.198^{+1.2\%}_{-1.3\%} \pm 2.7\%$	$0.568^{+1.9\%}_{-2.0\%} \pm 2.1\%$	$3.30^{+3.5\%}_{-4.3\%} \pm 1.9\%$
$t\bar{t}HH$	$0.775^{+1.5\%}_{-4.3\%} \pm 3.2\%$	$0.949^{+1.7\%}_{-3.1\%} \pm 3.1\%$	$5.24^{+2.9\%}_{-6.4\%} \pm 2.5\%$	$82.1^{+7.9\%}_{-7.4\%} \pm 1.6\%$
$tjHH$	$0.0289^{+5.5\%}_{-3.6\%} \pm 4.7\%$	$0.0367^{+4.2\%}_{-1.8\%} \pm 4.6\%$	$0.254^{+3.8\%}_{-2.8\%} \pm 3.6\%$	$4.44^{+2.2\%}_{-2.8\%} \pm 2.4\%$

Table 1.2. Computations of the signal cross sections (in fb) for double Higgs (H) production at different centre-of-mass energies [36], including the available QCD corrections according to the recommendations of the LHC Higgs Cross Section Working Group [168]. The renormalisation and factorisation scales have been set to $m_{HH}/2$ for gluon gluon fusion, to the individual virtualities $Q_{1,2} = \sqrt{-q_{1,2}^2}$ of the t-channel vector-bosons for VBF, m_{HHV} ($V = W, Z$) for HHV production, $m_{t\bar{t}}/2$ for $t\bar{t}HH$, $m_{HH}/2$ for $tjHH$ production. PDF4LHC15 parton distributions have been used to obtain the results, and the corresponding PDF + α_s uncertainties. The cross sections for $tjHH$ involve both top and anti-top production.

In order to understand more related to the non-resonant double Higgs production, we have to go deeper into the theory of such production. Our main production mode of interest is the gluon-gluon fusion (ggF- HH), since in term of cross section is the most dominant one. Theoretical computation of $ggF - HH$ have been extended up to the next-to-next o leading order (NNLO)[72] and next-to-next-to-leading logarithmic (NNLL)[73] accuracy for the QCD correction and up to the NLO correction related to finite top quark mass effects on the HH production and/or radiation[91][89][90]. The QCD corrections improve the ggF- HH cross section almost by double from LO to NLO, with another $\sim 20\%$ increase going from NLO to NNLO. The threshold resummation corrections for Higgs boson pair production at NNLL [72][73] further increase the rate. The NNLL threshold resummed cross sections are combined with the fixed order NNLO results as shown in Table 1.3, with the rate being weighted by the exact LO finite m_t result normalized to the $m_t \rightarrow \infty$ LO result. This procedure is called as the HEFT approximation. The scale is set to be $\mu_0 = M_{hh}/2$, meanwhile the scale variation is:

$$\mu_0/2 < \mu_R, \mu_F < 2\mu_0 \quad \text{with the restriction :} \quad 1/2 < \mu_R/\mu_F < 2 \quad (1.40)$$

The numerical importance of the threshold resummation is minimized for $\mu_0 = M_{hh}/2$. The scale uncertainties are $\sim 4 - 6\%$ and the PDF uncertainties are $\sim 2 - 3\%$ at $\sqrt{s} = 13\text{TeV}$, as shown in Table 1.4. For convenience, two K factors are defined and computed in $m_t \rightarrow \infty$ limit:

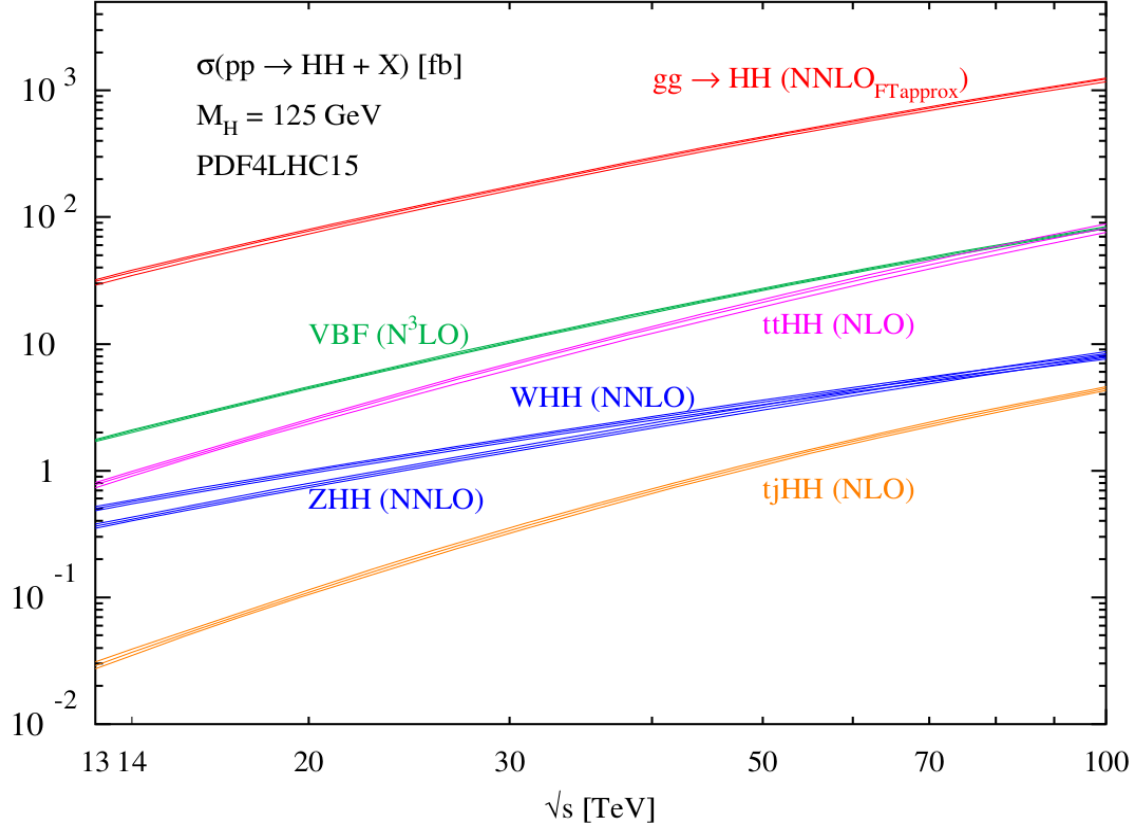


Figure 1.7. Total production cross sections for Higgs pairs within the SM via gluon fusion, vector-boson fusion, double Higgs-strahlung and double Higgs bremsstrahlung off top quarks [112]. The size of the bands shows the total uncertainties originating from the scale dependence and the $PDF + \alpha_s$ uncertainties.

$m_H = 125.09 \text{ GeV}$	$\sigma_{NNLL} (fb)$	Scale Unc. (%)	PDF Unc. (%)	α_s Unc. (%)
$\sqrt{s} = 7 \text{ TeV}$	7.61	+5.6 -6.0	± 3.3	± 2.8
$\sqrt{s} = 8 \text{ TeV}$	11.00	+5.5 -6.0	± 3.0	± 2.6
$\sqrt{s} = 13 \text{ TeV}$	37.30	+5.1 -6.1	± 2.1	± 2.3
$\sqrt{s} = 14 \text{ TeV}$	44.20	+5.2 -6.1	± 2.1	± 2.2
$\sqrt{s} = 100 \text{ TeV}$	1712.00	+5.2 -6.2	± 1.7	± 2.0

Table 1.3. NNLL matched to NNLO cross sections for $gg \rightarrow HH$ with a central scale $\mu_0 = M_{HH}/2$ with $m_H = 125.09 \text{ GeV}$ [48][72] computed in the HEFT approximation. The uncertainties from top quark mass effects are not included in this table. Uncertainties are evaluated using the PDF4LHC recommendation based on the PDF4LHC15_nnlo_mc set.

$$K \equiv \frac{\sigma_{NNLL}}{\sigma_{NLO}}, \quad K_0 \equiv \frac{\sigma_{NNLL}}{\sigma_{LO}} \quad (1.41)$$

1.4.2 Invariant mass $M_{inv}(HH)$ of the non-resonant double Higgs production

The invariant mass distribution has a peak at $m_{hh} \sim 400 \text{ GeV}$ [38], while the p_t distribution is maximum at $p_T \sim 150 \text{ GeV}$. The majority of events have an invariant mass that is far above the kinematic threshold at $2m_h$. This observation has two important implications:

1. The invalidity of the Higgs low-energy theorem in $gg \rightarrow HH$: It is known that the $m_t \rightarrow \infty$ limit works well in F_Δ (loop function of the triangle Feynman diagram) and terribly in $F_\square, G_\square$ (loop functions

$\mu_0 = M_{HH}/2$	$\sqrt{s} = 7 \text{ TeV}$	$\sqrt{s} = 8 \text{ TeV}$	$\sqrt{s} = 13 \text{ TeV}$	$\sqrt{s} = 14 \text{ TeV}$	$\sqrt{s} = 100 \text{ TeV}$
K	1.203	1.200	1.193	1.192	1.195
K_0	2.299	2.296	2.301	2.304	2.472
$\mu_0 = M_{HH}$	$\sqrt{s} = 7 \text{ TeV}$	$\sqrt{s} = 8 \text{ TeV}$	$\sqrt{s} = 13 \text{ TeV}$	$\sqrt{s} = 14 \text{ TeV}$	$\sqrt{s} = 100 \text{ TeV}$
K	1.426	1.413	1.378	1.373	1.305
K_0	2.987	2.949	2.847	2.835	2.699

Table 1.4. K -factors as defined in Eq. 1.41 for $gg \rightarrow HH$ with a central scale $\mu_0 = M_{HH}/2$ and $\mu_0 = M_{HH}$ with $m_H = 125.09 \text{ GeV}$.

of the box diagram with association of the top quarks and without)[83][49]. The partonic center-of-mass energy in the double Higgs production must always be above the kinematic threshold: $\hat{s} \geq 4m_h^2$ (partonic Mandelstam variable $\hat{s}$), while the low-energy theorems require $\hat{s} \ll 4m_t^2$ [107].

2. The second has to do with the relative weight between c_{tri} and c_{nl} , where the loop function F_Δ (Appendix A) has the coefficient:

$$c_{tri} \frac{3m_h^2}{\hat{s} - m_h^2} + c_{nl}$$

The coupling c_{tri} becomes more important at small invariant mass, near the kinematic threshold $m_{hh} \sim 2m_h$, while c_{nl} could dominate over c_{tri} at large m_{hh} . In fact, since most of events have $m_{hh} \gg 2m_h$, the contribution from c_{tri} will be suppressed in the total cross-section [42]. The individual contribution from c_{tri} , c_{nl} , and c_{box} , respectively, in the m_{hh} distribution and compare them with the SM expectation. Indeed, when $c_{tri} = c_{nl}$ the contribution to the total cross section from the Higgs trilinear coupling is quite small. As a result, turning on a small c_{nl} would have a significant impact on the measurement of c_{tri} .

An important feature is of the $ggF - HH$ production mode is the relative small total cross section, a feature, which is explained from the destructive interference between the box and triangle diagram, as shown in Figure 1.8.

1.4.3 Top Quark Mass Uncertainties

The inclusion of the top quark mass effects decreases the NLO cross section from the HEFT result [18], while the FT -approx is a better approximation to the total rate. The inclusion of the mass effects reduces the NLO rate by an energy dependent factor which can be parameterized as:

$$\sigma(gg \rightarrow hh)_{NLO}^{exact} = \sigma(gg \rightarrow hh)_{NLO}^{HEFT} (1 + \delta_t) \quad (1.42)$$

where, using $m_t = 173 \text{ GeV}$:

$$\begin{aligned} \delta_t(7 \text{ TeV}) &= -9.94\% \\ \delta_t(8 \text{ TeV}) &= -10.88\% \\ \delta_t(13 \text{ TeV}) &= -13.72\% \\ \delta_t(14 \text{ TeV}) &= -14.11\% \end{aligned} \quad (1.43)$$

The top mass effects can be included as shown below:

$$\sigma'_{NNLL} = \sigma_{NNLL} + \delta_t \sigma_{NLO}^{HEFT} \quad (1.44)$$

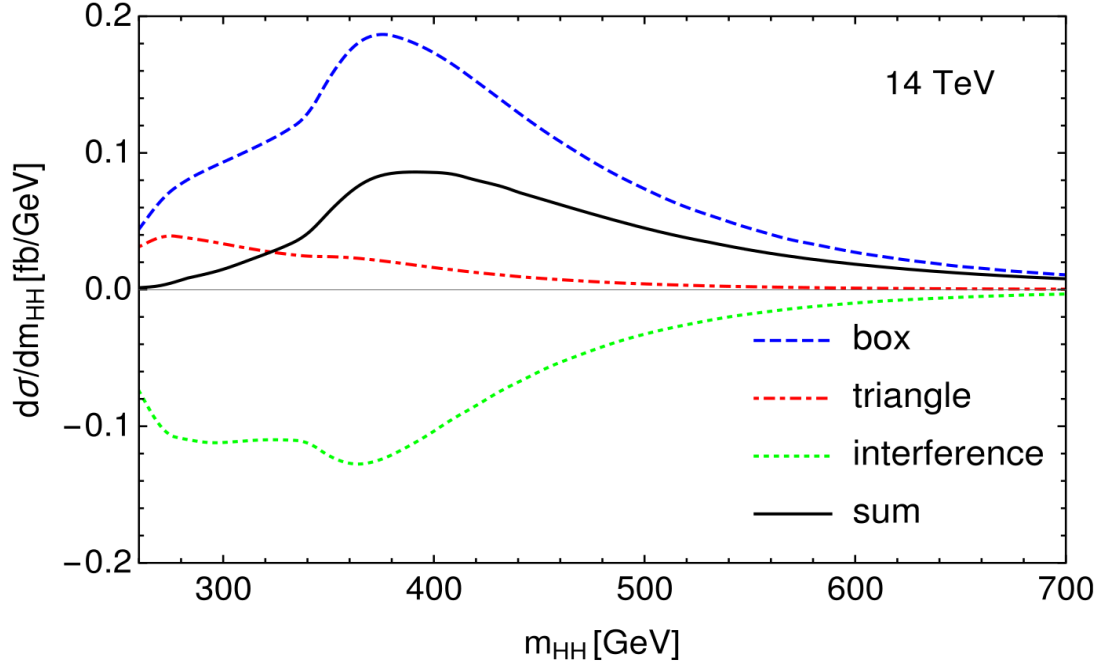


Figure 1.8. Higgs pair invariant mass distribution at leading order for the different contributions to the gluon fusion production mechanism and their interference [36]

m_H GeV	$\sqrt{s} = 7$ TeV	$\sqrt{s} = 8$ TeV	$\sqrt{s} = 13$ TeV	$\sqrt{s} = 14$ TeV	$\sqrt{s} = 100$ TeV
125 GeV	$6.415^{+20\%}_{-16.8\%} \pm 4.0\%$	$9.318^{+19.5\%}_{-16.4\%} \pm 3.7\%$	$31.81^{+18.2\%}_{-15.0\%} \pm 2.7\%$	$37.79^{+18\%}_{-14.8\%} \pm 2.6\%$	$1464^{+16-1\%}_{-13-8\%} \pm 2.1\%$

Table 1.5. Signal cross section (in fb) for $gg \rightarrow HH$ at NLO QCD in the $m_t \rightarrow \text{inf}$ limit, reweighted by the exact LO result [18], with a central scale $\mu_0 = M_{HH}/2$ with the scale variation and top quark mass effects uncertainties.

The differential distribution after the MC generations at different Lo,NLO orders are shown in Figure 1.9, where the implementation of the m_t -mass effect has been taken in consideration.

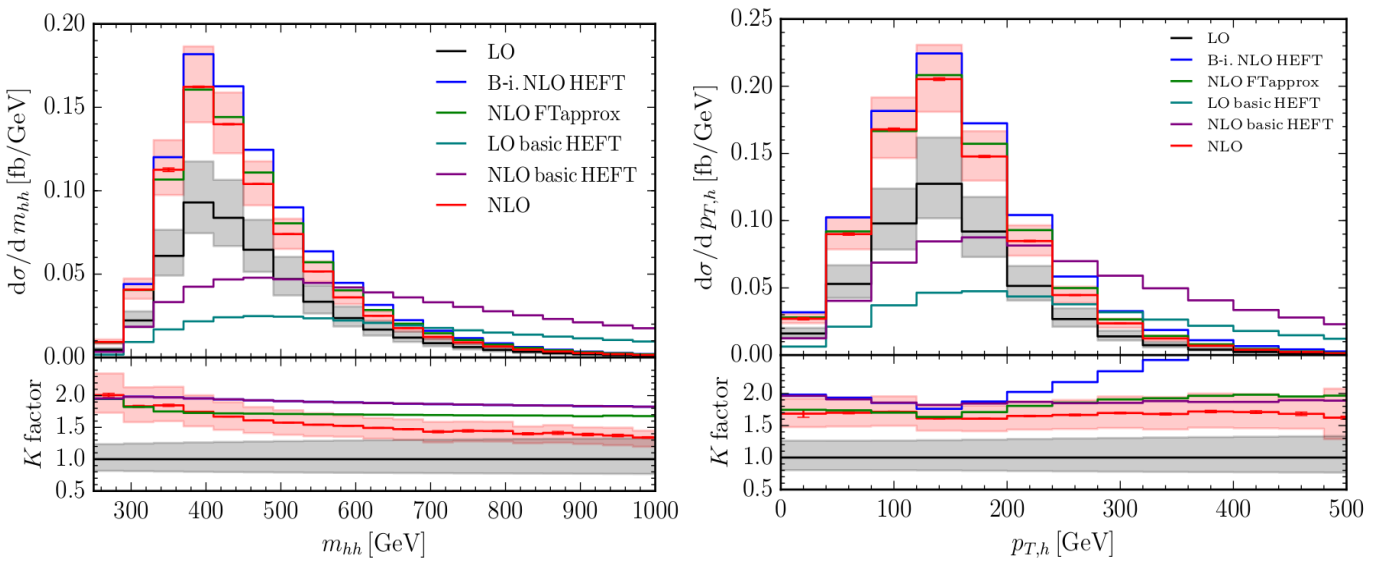


Figure 1.9. Higgs pair invariant mass and transverse-momentum distributions for a center-of-mass energy of 14 TeV in various approximations. The full NLO results are shown in red. The red bands show the renormalization and factorization scale dependence obtained from a 7-point scale variation around the central scales $\mu_f = \mu_r = m_{HH}/2$ [18].

1.4.4 Non-resonant HH production mode beyond the SM

BSM physics signatures in HH final states can also be probed in nonresonant HH production. If the scale of BSM physics is sufficiently high that resonances cannot be directly produced, its effects can be observed through the contributions in the quantum loops responsible for HH production. Experimental signatures are in this case an enhancement of the HH production rate and a sizeable modification of the Higgs boson pair kinematic properties. In the following, the HH phenomenology in presence of an anomalous trilinear Higgs boson coupling is first discussed, and serves as an introduction to a more general approach based on an effective SM field theory.

Benchmark nr.	k_λ	k_t	c_2	c_g	c_{2g}
1	7.5	1.0	-1.0	0.0	0.0
2	1.0	1.0	0.5	-0.8	0.6
3	1.0	1.0	-1.5	0.0	-0.8
4	-3.5	1.5	-3.0	0.0	0.0
5	1.0	1.0	0.0	0.8	-1.0
6	2.4	1.0	0.0	0.2	-0.2
7	5.0	1.0	0.0	0.2	-0.2
8	15.0	1.0	0.0	-1.0	1.0
9	1.0	1.0	1.0	-0.6	0.6
10	10.0	1.5	-1.0	0.0	0.0
11	2.4	1.0	0.0	1.0	-1.0
12	15.0	1.0	1.0	0.0	0.0
SM	1.0	1.0	0.0	0.0	0.0

Table 1.6. MC generations of the benchmarks shapes: Each benchmark shape corresponds to a specific choice of the 5 EFT couplings, whose choices do not have a special physical meaning, it is only meant to represent a shape, not a special point in the EFT parameter space

A "visual" representation of the BSM benchmark distributions is to be seen in the [Figure 1.10](#), where the histograms are generated from the MC files at LO with MadGraph5_aMC@NLO.

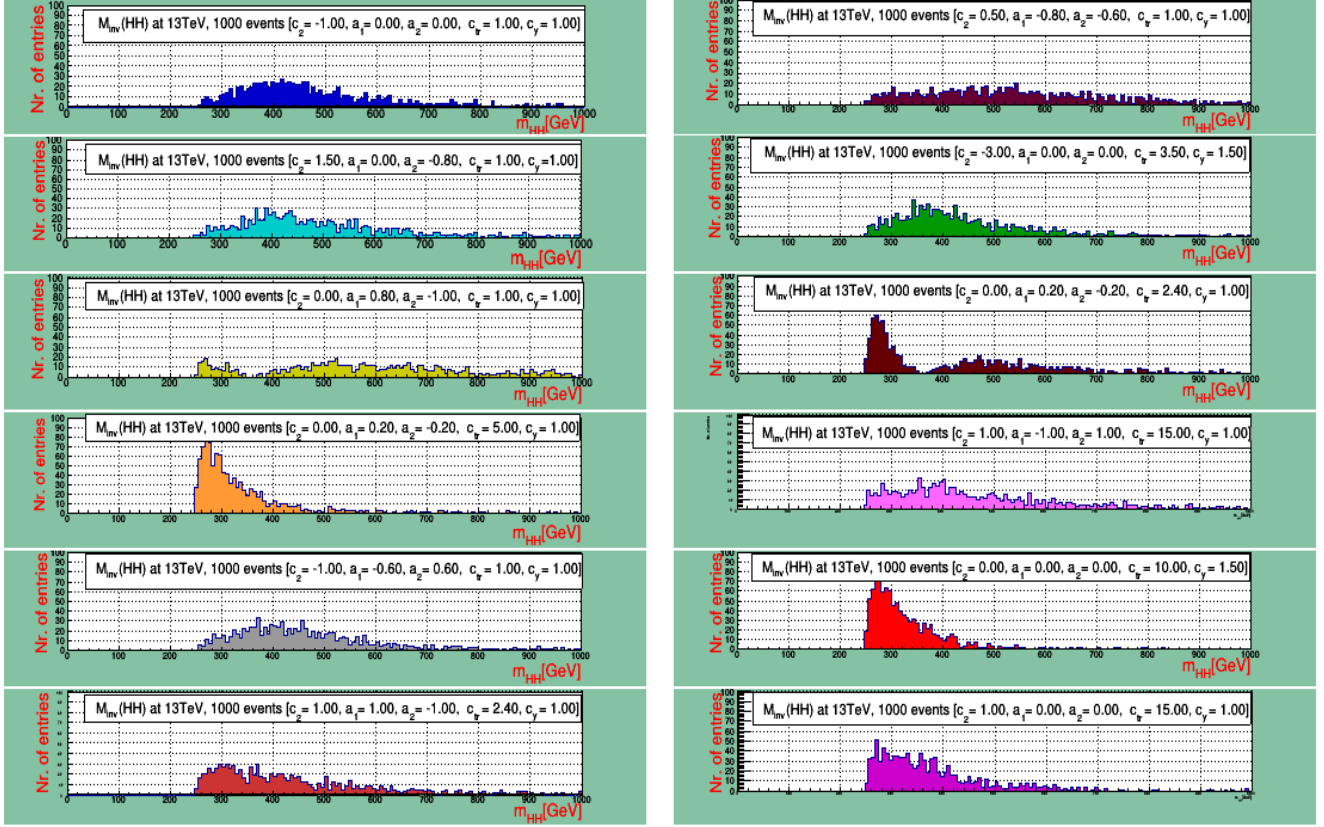


Figure 1.10. The benchmark signal shapes, generating several distributions (1000 events each) of double Higgs boson invariant mass $M_{inv}(HH)$, in LO accuracy using as MC generator MadGraph5_aMC@NLO (MG5_aMC_v2.6.3.2), combining the values of the coupling coefficients from the SM extended Lagrangian in the framework of the EFT

1.5 Resonant HH production

Different from the sections above, where we spoke about the non-resonant double Higgs (H) production, this section will implement possible variants of a di-Higgs production from different theoretical and MC computations models as a result of a decay from a possible massive resonances X .

1.5.1 Higgs Singlet Model

The Higgs singlet model[17][115] is one of several theoretical models of massive resonances decaying into a pair of Higgs bosons. The model contains a Higgs doublet ϕ and Higgs singlet S :

$$\begin{aligned}\Phi^T &= (\phi^+, \tilde{\phi}_0 = \frac{\phi_0 + v}{\sqrt{2}}) \\ S &= \frac{s + \langle S \rangle}{\sqrt{2}}\end{aligned}\tag{1.45}$$

and is described by 5-parameters in the potential:

$$V = -m^2 \Phi^\dagger \Phi - \mu^2 S^2 + \lambda_1 (\Phi^\dagger \Phi)^2 + \lambda_2 S^4 + \lambda_3 \Phi^\dagger \Phi S^2\tag{1.46}$$

where a Z_2 symmetry $S \rightarrow -S$ and $\Phi \rightarrow -\Phi$ has been imposed for simplicity. After electroweak symmetry breaking, both $\tilde{\phi}_0$ and S get vacuum expectation values and the physical fields h, H are mixtures of the original fields:

$$\begin{aligned} h &= \cos\alpha\phi_0 - \sin\alpha s \\ H &= \sin\alpha\phi_0 - \cos\alpha s \end{aligned} \quad (1.47)$$

and we assume $M_H > m_h$ in the following. The LO trilinear Higgs boson couplings are:

$$\begin{aligned} \lambda_{hhh} &= -\frac{3m_h^2}{3} (\cos^3\alpha - \tan\beta \sin^3\alpha) \\ \lambda_{Hhh} &= -\frac{m_h^2}{3} \sin^3(2\alpha) (\cos\alpha + \sin\alpha \tan\beta) \left(1 + \frac{M_H^2}{2m_h^2}\right) \end{aligned} \quad (1.48)$$

where $\tan\beta = v/\langle s \rangle$ is the ratio of the VEV of the two fields. The input parameters can be taken as:

$$m_h = 125\text{GeV}, \quad M_H, \quad \cos\alpha, \quad v, \quad \tan\beta = v/\langle s \rangle$$

and the Higgs boson branching ratios to SM particles, X_{SM} , are:

$$\begin{aligned} \Gamma(h \rightarrow X_{SM}X_{SM}) &= \cos^2\alpha \Gamma(h \rightarrow X_{SM}X_{SM})_{SM} \\ \Gamma(H \rightarrow X_{SM}X_{SM}) &= \sin^2\alpha \Gamma(H \rightarrow X_{SM}X_{SM})_{SM} \\ \Gamma_H &= \sin^2\alpha \Gamma_{H,SM}(M_H) + \Gamma(H \rightarrow hh) \\ \Gamma_h &= \cos^2\alpha \Gamma_{h,SM}(m_h) \end{aligned} \quad (1.49)$$

where $\Gamma_{H,SM}(M_H)$ are the Standard Model Higgs boson widths evaluated at M_H and are completely fixed in terms of $\tan\beta$, M_H and $\cos\alpha$. The branching ratio, $H \rightarrow hh$, can be quite large, $O \sim 20 - 30\%$ [113], leading to large effects in di-Higgs boson production.

1.5.2 2-Higgs-Doublet Model (2HDM)

The 2-Higgs-Doublet model (2HDM) [19] is an other example of the extension of the SM, which can contain resonance effects. The 2HDM predicts 5 physical Higgs bosons:

- 2 neutral scalars, h^0, H^0
- One pseudo-scalar, A
- Two charged scalars H^+, H^-

The 2HDMs Higgs bosons mediate tree level flavour changing neutral currents (FCNCs), which must be suppressed. This suppression is imposed by a Z_2 discrete symmetry, where the fermions of a given charge only couple to one of the Higgs doublets. There are two versions of the 2HDM model:

1. Type I model: All of the fermions couple to the same Higgs doublet
2. Type II model: The $Q = 2/3$ quarks couple to one doublet and the $Q = -1/3$ quarks and leptons couple to the other

The type II model is the Higgs sector of the MSSM (Minimal Supersymmetric Standard Model Physics). Two additional versions interchange the lepton assignments. In the "lepton-specific" model, all of the quarks couple to one doublet, while the leptons couple to the other, and in the "flipped" model, the $Q = 2/3$ quarks and leptons couple to one doublet and the $Q = -1/3$ quarks couple to the other. The couplings of the Higgs

bosons to fermions are described by two free parameters. The ratio of the vacuum expectation values of the two Higgs doublets is :

$$\tan\beta \equiv \frac{v_2}{v_1} \quad (1.50)$$

and the mixing angle, which diagonalizes the neutral scalar mass matrix is α . The couplings of the light (heavy) CP even Higgs boson, h^0 (H^0), to fermions and gauge bosons relative to the Standard Model couplings are given for all four 2HDMs considered. 2HDMs are significantly limited by experimental data. Higgs boson coupling measurements restrict $\cos(\alpha - \beta)$ to be close to the SM limit:

$$\cos(\alpha - \beta) \sim 0$$

while heavy Higgs searches restrict M_{H^0} as a function of $\cos(\alpha - \beta) \sim 0$. The benchmarks were further chosen such that the total rate for di-Higgs boson production is similar to that of the SM. These benchmarks can exhibit significant resonance effects (B1 and B2), while B7 is almost indistinguishable from the SM.

1.5.3 Minimal supersymmetric extension of the SM (MSSM)

In the minimal supersymmetric extension of the SM (MSSM) [56], the addition of a complex Higgs doublet implies the existence of five new particles:

- Two CP-even (i.e scalars) neutral h and H
- One CP-odd (i.e pseudoscalar) A
- Two charged scalars H^+ , H^-

As for the singlet Higgs model, it is assumed that $m_H > m_h$ and one of the two particles (usually the lightest one) is interpreted as the scalar boson observed at the LHC. At the tree level, the model can be completely described in terms of two parameters, that are usually chosen as the mass of the pseudoscalar m_A and the ratio of the VEV of the two fields:

$$\tan\beta = \frac{v_u}{v_d}$$

The mass of the charged scalar correspond to:

$$m_{H^\pm}^2 = m_A^2 + m_W^2$$

while the mass of the neutral scalars are given from the mass matrix:

$$\delta M_{tree}^2 = m_Z^2 \begin{pmatrix} \cos^2\beta & -\sin\beta\cos\beta \\ -\sin\beta\cos\beta & \sin^2\beta \end{pmatrix} + m_A^2 \begin{pmatrix} \sin^2\beta & -\sin\beta\cos\beta \\ -\sin\beta\cos\beta & \cos^2\beta \end{pmatrix} \quad (1.51)$$

The introduction of an additional doublet produces deviations of the couplings of the Higgs bosons and, more importantly for this context, implies the presence of a Hhh coupling that produces a resonant $H \rightarrow hh$ signature.

1.5.4 Kaluza-Klein Bulk Graviton spin-2 and Radion spin-0

The first ideas and theories related to a multi-dimensional spacetime with more than three spatial dimensions were not generated recently, but they date back to the 1920's, where Theodor Kaluza and Oscar Klein put an effort to unify the forces of Nature. The scope was to unify gravitation and electromagnetism around the idea of a fifth dimension beyond the common 4D-spacetime. The idea of extra dimensions coming from the Kaluza-Klein theory (also called the KK theory) has become an important part of quantum gravity theories and a key role in string theories nowadays. The addition of extra dimensions was strongly supported by Nima Arkani-Hamed, Savas Dimopoulos and Gia Dvali in 1998 [8], where they suggested that the weakness of the gravitational force can be explained by its propagation through these additional dimensions.

The model of Lisa Randall and Raman Sundrum [9] related to the "5-dimensional warped geometry theory" gives a mathematical option about extra dimensions, which are compactified ("warped extra dimensions") between two points of the space referred to as "branes". One brane corresponds to the Planck scale (M_{Pl}) and the other to the electroweak scale, and the region separating them, the "bulk", is controlled by an exponential metric. The gap between these two fundamental scales of the Nature is parametrized by a warp factor k . The consequence of these models is the existence of new particles: Bulk KK-Graviton of spin-2 and Radion of spin-0. This model is very important specially to detector and collision Physics, where it found applications for possible studies at the LHC at CERN [2] and which implies that those possible massive resonances such as the Bulk KK-Graviton of spin-2 and Radion of spin-0 could possibly decay to a pair of Higgs bosons. The former is the mediator of the gravitational force itself, while the latter is required to stabilize the size of the extra dimension l .

The phenomenology varies whether only the gravitational fields or the SM fields as well are allowed to propagate in the bulk. In the former case, denoted as RS1 model, couplings of the Graviton to matter fields are determined by $\tilde{k} = k/M_{Pl}$, with $\bar{M}_{Pl} = M_{Pl}/8\pi$. In the latter case, the bulk RS-model, the couplings depend on the localization of the SM fields in the bulk. These models make the studies on the Bulk KK-Graviton and Radion even more interesting, since their invariant masses can extent up to the order of TeV, increasing the importance of the Higgs Physics also in those energy regions. From the other side, this could possibly allow high-energy unification of gauge couplings and provides a natural hierarchy of masses. This means that light quarks can be localized close to the Planck brane, while the top quarks are close to the TeV brane (elementary top hypothesis), providing a large top quark mass.

The properties of the Radion resonance are parameterized in terms of $\tilde{k}$ and of $\Lambda_R = \sqrt{6}e^{(-kl)}M_{Pl}$, the latter usually interpreted as the ultraviolet cutoff scale of the theory. The Graviton production cross section is proportional to $\tilde{k}^2$ and the radion cross section is proportional to $1/\Lambda_R^2$ [9].

The solution form to solve the classical Einstein motion equations, maintaining 4D Poincaré invariance, is:

$$ds^2 = e^{-2\sigma(\phi)}\eta_{\mu\nu}dx^\mu dx^\nu + r_c^2 d\phi^2, \quad \mu, \nu = 1, 2, 3, 4 \quad (1.52)$$

where μ, ν are the 4D indexes ϕ the coordinate for the fifth dimension [108]. The full classical action is as following:

$$S = S_{Gravity} + S_{TeV} + S_{Planck} + S_{Matter} \quad (1.53)$$

where $S_{Gravity}$ is the bulk gravitational action, S_{Matter} is the action matter fields. We consider that the parts of the gravitational action $S_{Gravity}$ is confined on the branes and therefore we denote them by $S_{TeV/Planck}$. The gravitational part of the action can be so written as following:

$$\begin{aligned}
S_{i=TeV/Planck} &= - \int d^4 \sqrt{g(\phi=0, \pi)} \Lambda_{i=TeV/Planck} \\
S_{Gravity} &= \int d^4 \int_{-\pi}^{\pi} d\phi \sqrt{g} (\Lambda_{bulk} + 2M_5^3 R)
\end{aligned} \tag{1.54}$$

where the Λ 's are vacuum energy densities on the branes/bulk, R is the Ricci tensor, M_5 is the 5D-Planck mass and g_{XY} is the 5D-metric. If we denote the trace of the space-time metric as $g = -g_X^X$, the unity of volume would be $-\int d^5 x \sqrt{g}$. If we take in consideration the condition:

$$\Lambda_{bulk} = \Lambda_{Planck} = \Lambda_{TeV} = \Lambda \tag{1.55}$$

then we obtain [108]:

$$\begin{aligned}
\sigma(\phi) &= r_c \mid \phi \mid \sqrt{\frac{-\Lambda}{24M_5^2}} \equiv r_c \mid \phi \mid k \\
\text{with } k &\equiv \sqrt{\frac{-\Lambda}{24M_5^2}}
\end{aligned} \tag{1.56}$$

where k is the curvature factor, which depends on the relation between the negative cosmological constant and the brane tensions. Integrating out the extra dimension, we find the 4D-Planck mass to be:

$$\bar{M}_{Pl}^2 = \frac{M_5^3}{k} \left(1 - e^{2\pi k r_c} \right) \tag{1.57}$$

where M_5 is the 5D-Planck mass. The curvature factor k is assumed to be of the order of 5D-Planck scale M_5 . If we confine the pair of the Higgs (H) bosons to propagate only on the TeV brane, as we can observe in Figure 1.11 its four dimensional vacuum expectation value would be given by:

$$v \equiv e^{-\pi k r_c} v_0 \tag{1.58}$$

where r_c is the compactification radius.

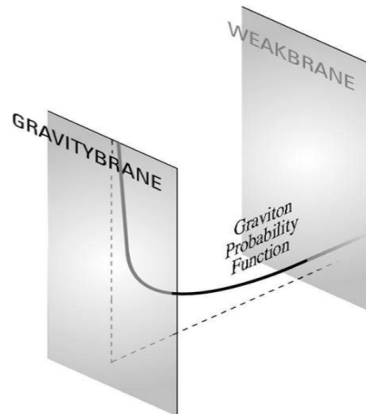


Figure 1.11. Scheme of dimensions on RS theory. The Gravity (Planck) and Weak (TeV) branes are the 4 dimensional boundaries of the extra dimension ϕ compactified on a interval $(0, \pi)$. The figure also illustrate the metric behavior along the extra dimension

The five dimensional Higgs vacuum expectation value v_0 is of the order of the 5D fundamental mass scale M_5 . The separation between Planck mass and EW scales is produced by the metric, when $kr_c \sim 11$ and it is called as "Planck-EW hierarchy reduction". Despite this theoretical value, both k and r_c are free parameters

of the theory. Probes on the range of validity of classical gravity constrain the value of the compactification radius r_c . The combinations $\tilde{k} \equiv \frac{k}{\bar{M}_{Pl}}$ and kr_c are used as the main parameters of the theory.

1.5.5 Gravity particles

We refer as gravity particles such particles, those which get created as a result of the quantum fluctuations around the classical metric solution Eq.1.56. The fluctuation modes can be decomposed into a 4D tensor $\otimes$ 4D vector $\otimes$ scalar components:

$$\delta g_{MN}(x, \phi) = \begin{pmatrix} h_{\mu\nu}(x, \phi) & h_{\mu,5}(x, \phi) \\ h_{\mu,5}(x, \phi) & h_{5,5}(x, \phi) \end{pmatrix} \quad (1.59)$$

The motion's relations for the independent fields of the equation above, form a over constrained system of equations. The axial gauge, defined by $h_{\mu 5} = 0$, it decouples the dynamics of tensor and scalar perturbations. The perturbed space-time metric on this gauge has the form:

$$ds^2 = e^{-2(\epsilon_r h_{55} + \sigma(\phi, r_c))} (\eta_{\mu\nu} + \epsilon_g h_{\mu\nu}(x, \phi)) dx^\mu dx^\nu + r_c^2 (2\epsilon_r h_{55}(x, \phi) + 1)^2 d\phi^2 \quad (1.60)$$

ϵ_g and ϵ_r are the expansion parameters on the perturbed metric, with dimensions $[mass]^{-3/2}$ and $[mass]^{-1}$ respectively. The tensor fluctuations ($h_{\mu\nu}$) correspond to Graviton modes and the scalar fluctuation (h_{55}) correspond to the Radion mode [45][37]. The dynamics and interactions of the gravity particles is found on the gravitational action as following:

$$S_{Gravity} = S_{Graviton} + S_{radion} \quad (1.61)$$

The five dimensional motion's equations are solved performing a Fourier expansion on the field 5D wave function that separates the ϕ, x_μ variables as shown below:

$$X(x_\mu, \phi) = \sum_{n=0}^{\infty} X^n(x^\mu) f_X^{(n)}(\phi) \quad (1.62)$$

Every single four dimensional expansion mode is known as the n -th Kaluza-Klein (KK) mode and it acts as a particle in the effective 4D-theory. The 5th dimensional part of KK-modes wave functions $f_X^{(n)}(\phi)$ is commonly called as profiles. The profiles $f_X^{(n)}$ obey a differential Sturm-Liouville problem defined by the equations of motion, spin of the field and boundary conditions. The profiles for all the known fields are described by a combination of exponential and Bessel functions. We denote the Graviton four dimensional wave functions (profiles) as $h_{\mu\nu}^{(n)}(x_\mu) f_h^{(n)}(\phi)$. The zero mode of the Graviton field correspond to the mediator of gravitational force. The first bulk KK-mode graviton has a profile that is localized on the TeV region, and its 4D effective mass is:

$$m_{Gr} \equiv m_{Gr}^{(1)} = x_1 \tilde{k} e^{kL} \bar{M}_{Pl} \sim TeV \quad (1.63)$$

where $x_1 = 3.83$ is the first zero of the Bessel function J_1 . If we use as the free parameter of theory m_{Gr} and $\tilde{k}$, then the relation for that would be as following:

$$kr_c = \frac{1}{\pi} Ln \left(x_1 \tilde{k} \frac{\bar{M}_{Pl}}{M_{Gr}} \right) \quad (1.64)$$

Following the above relation, the kr_c value varies between $\sim 10 - 12$ in case if we define $\tilde{k}$ to fluctuate between 0.01 and 1 and M_{Gr} between 100 GeV and 1.5 TeV. A similar procedure defines the physics of the

scalar fluctuation of the metric, known as Radion. The 5D radion field $r(x, \phi)$ has canonical kinetic term and relates to $h_{55}(x, \phi)$ as [58]:

$$r(x, \phi) = \sqrt{\frac{24M_5^3}{k}} e^{-k\epsilon_r h_{55}(x, \phi)\pi} \quad (1.65)$$

If we consider terms with only two insertions of the r field, ϵ_r becomes a neglectable parameter to the theory. In the simplest case, the zero mode of the Radion would behave as a massless field [57]. The Radion however couples with the SM particles through the expansions of the metric on SM matter action, including photons [95], slightly modifying propagation of light, which would be experimentally disfavored [86]. In addition to that, the distance between the branes (r_c) was introduced in the non-perturbed metric (Eq. 1.52) as a free parameter. The value of r_c is important to define the hierarchy of gravity and weak scales.

KK-Bulk Graviton

The two main benchmarks to be considered are the Bulk and the RS1 scenario. Those two scenarios are important for the couplings of the gravity particles with matter fields and they imply how the SM matter fields behave in the extra dimension. The SM model variant implies that the Higgs pair might be localized on the TeV brane. Meanwhile, the implication of an extra dimension, gives the possibility for the matter fields to propagate in the bulk "through" that extra dimension[45]. The first variant is referred as the RS1-model and the second as the Bulk model, as shown in the Figure 1.12 below.

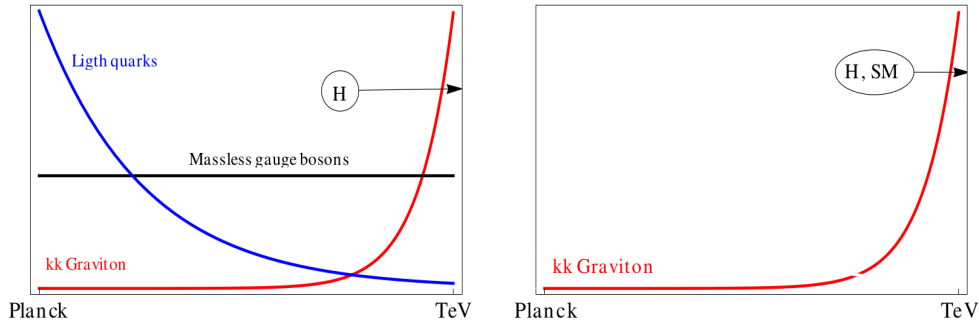


Figure 1.12. Representation of the two main benchmarks WED scenarios for the localization of the matter, where left we have Bulk scenario and right we have the RS1 scenario.

The KK expansion for the bulk scenario is applied to each field in a similar way as is Eq. 1.62. The zero mode of each KK tower correspond to the SM correspondent SM particle, assumed to be originally massless. The effective 4D-strength of the interaction between any fields are proportional to the integral of their profiles in the 5th dimension and the interactions of the KK-graviton with the matter are suppressed by:

$$\Lambda_G \equiv \frac{m_G}{x_1 \bar{k}} \quad (1.66)$$

and so we obtain the following Lagrangian:

$$\mathcal{L} = -\frac{x_1 \bar{k}}{m_G} h^{\mu\nu(1)} \times d_i T_{\mu\nu}^i \quad (1.67)$$

with $T_{\mu\nu}^i \equiv -\frac{2}{\sqrt{g}} \frac{\delta \mathcal{L}}{\delta g^{\mu\nu}}$ and $d_i \equiv \int d\phi f_G f_i^2$

where $T_{\mu\nu}^i$ is the four dimensional canonical energy-momentum tensor for the field i [74] and d_i is the integral overlap between the profiles of the fields i with KK-graviton. The RS1 scenario implies that all the particles are localized on the TeV region. The Bulk scenario implies that the profile functions are normalized

and the parameter $d_i \leq 1$. The parameter f_{Gr} is on the TeV region, so d_i indicates by how much the profile of the field i is away from TeV brane. In case the Higgs pair is confined to TeV brane, the volume suppression parameter for the coupling strength would be as following:

$$d_H = f_h(\phi = \pi) \sim 1 \quad (1.68)$$

KK-Bulk Radion

The scalar resonance is the most promising candidate to be the lightest particle on the particle spectra, as it is in WED [46]. If we accept that the narrow width approximation is valid, the phenomenology of a radion and a heavy Higgs is identical. The radion production cross sections for the GF production mode are calculated at NNLO + NNLL order in QCD including soft-gluon resummation [52]. The PDF used is PDF4LHC15@NNLO with 30 replicas, and the renormalization and factorization scales choices follows the description from the beginning of the chapter. The Born level couplings between radion and massless bosons are:

$$\frac{c_{iir}}{\Lambda_R} (p_1 p_2 g_{\mu\nu} - p_{1,\mu} p_{2,\nu}) \quad (1.69)$$

where $i = g, \gamma$ and:

$$\begin{aligned} c_{ggr} &= \frac{\alpha_s}{4\pi} \sum_{fermions} Q_f F_{1/2}(\tau_f) - a_g \\ c_{\gamma\gamma r} &= \frac{\alpha_{QED}}{4\pi} \left(\sum_{fermions} Q_f F_{1/2}(\tau_f) - F_1(\tau_\omega) \right) - a_\gamma \end{aligned} \quad (1.70)$$

where $\tau_i = m_{R^*}^2 / 4m_i^2$, $F_{1/2}$, F_1 are the one loop level functions to Higgs-like scalar coupling to massless bosons mediated by fermion and weak boson loops[93], and Q_f the charge of the fermion on the loop. The dominant contribution comes from the top quark in the loop function in comparison to the contributions of the bottom, charm and tau running in the loop. The production cross sections of a_X particle is matched with the one of R using the relation:

$$\begin{aligned} \sigma_{GF}^{R^*}(m_{R^*}) &= \frac{1}{\Lambda_R^2} \frac{|c_{ggR^*}|^2}{|c_{SM}|^2} \times \sigma_{GF}^X(m_X = m_R) \\ \text{where } c_{SM} &= \frac{\alpha_s}{4\pi} \sum Q_f F_{1/2}(\tau_f) \end{aligned} \quad (1.71)$$

The strength of the radion coupling with the gluons is dominated by the point-like interaction induced by the β_{QCD} , that can be interpreted as take the infinite top approximation in the calculation on the GF of a Higgs-like boson. When $m_X > 1\text{TeV}$ the infinite top approximation is valid in the calculation of the k-factor. To m_X closer to $2 \times m_T$ the infinite top approximation is good up to 10% in a NLO calculation [28][98]. The point-like interaction proportional to β_{QED} is dominant in the coupling strength. We add 10% uncertainty in the ggF production to include finite top mass effects in the soft gluon resummation. The radion branching ratios can be calculated with eHDecay [43], that is originally a tool to calculate decay rates and branching fractions of a Higgs-like particle using an effective lagrangian with dimension 6-operators. We use the parametrization of the dimension 6 operators of the couplings in the SILH scheme[41]. The calculations are done in NLO in QCD. The decay of the radion to a pair of 125 GeV Higgs bosons is added to the calculation at LO, following the equation bellow:

$$\Gamma_{R' \rightarrow HH} = \frac{1}{32m_R^2} \sqrt{m_R^2 - 4m_H^2} \frac{(m_R^2 - 2m_H^2)^2}{\Lambda_R^2} \quad (1.72)$$

In Figure 1.13 we can see the production cross section for the Radion, with $k_l = 35$ and $\Lambda_R = 3\text{TeV}$ and the effect of the bulk term in radion production. The continuous lines stand for the bulk scenario with $kr_c\pi = 35$ while the dashed ones for the RS1 scenario. The decay rate to bosons is proportional to m_R^4 while to fermions is proportional to m_R^2 making the branching ratio to a top quark pair to be suppressed to $m_R > 1\text{TeV}$. The sensitivity of the radion branching ratios to heavy particles the parameter kr_c is negligible.

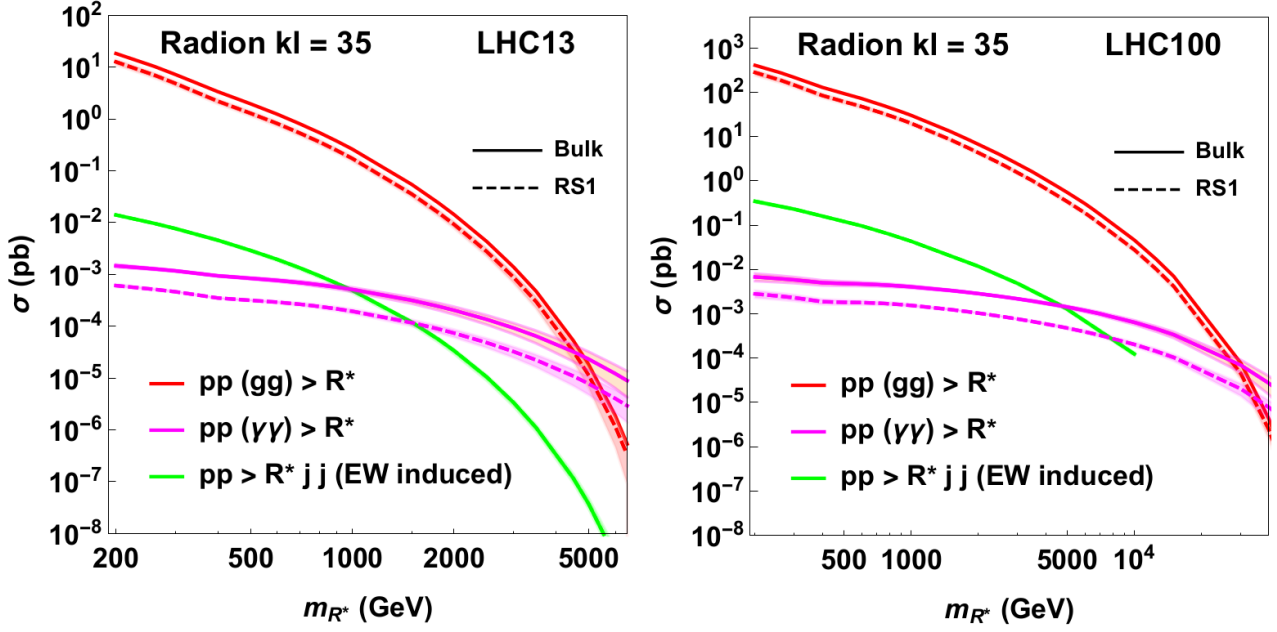


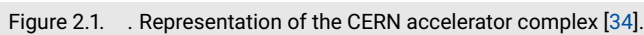
Figure 1.13. Production cross sections in pb for the Radion, with $k_l = 35$ and $\Lambda_R = 3\text{TeV}$ [27]. The GF (red) and VBF (Green) production modes for both to bulk (continuous lines) and RS1 (dashed lines) scenario. Right: 13TeV. Left: 100TeV.

Part II

The CMS Experiment at the LHC (CERN)

The Large Hadron Collider Experimental Apparatus

What is the LHC? The Large Hadron Collider (LHC) [31], by definition, is a proton-proton and heavy-ion superconducting accelerator and collider, built at CERN between 1998 and 2008. The purpose of LHC and



its experiments is to study the physics phenomena at high energy scale, and test the predictions of the physics models based on theories, such as the Standard Model (SM). The CERN facilities and laboratories complex host the Large Hadron Collider (LHC), designed to accelerate protons inside its 26.7 km long tunnel up to a center-of-mass energy of 14 TeV.

How does the LHC work at all? Well, the very funny thing of the story is that everything starts just from a bottle of liquid Hydrogen gas. The Hydrogen atoms are stripped from their electrons and the collected protons are accelerated in bunches every $100\mu\text{s}$ from an energy of $\sim 100\text{keV}$ to an energy of 50 MeV within the 30m long Linear Accelerator (LINAC2), which later injects the accelerated protons to the Proton Synchrotron Booster (PSB). The PSB further accelerates the protons up to 1.4GeV. The beam is then injected into the Proton Synchrotron (PS), the oldest CERN accelerator still in activity, for a further acceleration to 25 GeV. Subsequently, the last step of pre-acceleration happens within the 6.9km long Super Proton Synchrotron (SPS), where protons reach an energy of 450 GeV. The proton beams are finally transferred to the two LHC beam pipes, where the beams are accelerated and shaped into proton bunches thanks to radio-frequency cavities operating at 400 MHz. Once the protons reach the nominal energy and the beam dynamics is stabilized, protons are brought to collide in four points along the LHC ring: ATLAS, CMS, ALICE, and LHCb.

2.1 The Large Hadron Collider (LHC) at CERN

The Large Hadron Collider (LHC) is particle accelerator complex (Figure 2.1), which is able to accelerate protons in two beam of opposite directions within a distance of 26.7km . The protons are accelerated in two separate, counter-rotating particle beam lines thanks to the magnetic field generated by super-conducting niobium-titanium coils. The particles are steered by a magnetic field of 8.3T generated by a current of about 11kA in the 1232 dipole magnets, each measuring 15 meters of length and 35 tonnes of weight. The stability of the beam dynamics is ensured by 392 quadru-poles magnets measuring $5 - 7\text{m}$ of length, that focus the particles and keep them in a narrow beam. Special quadrupoles are installed in front of the collision points to squeeze the beams and increase the proton density in the collisions. Superconducting magnets are cooled with superfluid Helium and kept to a working temperature of 1.9K . The LHC complex of pre-acceleration and the system of magnets within the LHC tunnel achieves a center-of-mass energy of the proton beams up to:

$$\sqrt{s} = 6.5\text{TeV} + 6.5\text{TeV} = 13\text{TeV} \quad (2.1)$$

Since only a fraction of the protons energy contributes to the hard scattering, the center-of-mass energy of the beam must be much higher than the mass of the particle that has to be produced. In fact, when two protons collide, two of its partons (quarks and gluons) can take part in a hard interaction with high transferred p_T . The effective centre-of-mass energy of the hard scattering ($\sqrt{\hat{s}}$) is smaller than the energy of the machine, and it is proportional to the fractional energies x_a, x_b carried by the two interacting partons:

$$\sqrt{\hat{s}} = \sqrt{x_a x_b s} \quad (2.2)$$

The interaction of two protons forms a primary vertex, from which the particles created in the interaction originate. Several independent proton-proton interactions can take place in a bunch crossing at the interaction point, this is known as pile-up. The number of interactions on average depends on the beam parameters, e.g. number of particles in a bunch and the focusing area size. The nominal average pile-up in LHC is ≈ 25 , meanwhile the mean values of interactions per crossing (pileup) and the minimum bias cross sections for 2011-2012 and 2015-2018 can be seen in Figure 2.2.

The instantaneous luminosity is defined as the number of collisions per unit time and cross-sectional area of the beams:

$$\begin{aligned} A &= \frac{4\pi\epsilon_n\beta^*}{\gamma F}, \quad F = \left(1 + \frac{\theta_c\sigma_z}{2\sigma_{xy}}\right)^{-\frac{1}{2}} \\ \rightarrow \mathcal{L} &= \frac{N_1 N_2 n_b f \gamma}{4\pi\epsilon_n\beta^*} \left(1 + \frac{\theta_c\sigma_z}{2\sigma_{xy}}\right)^{-\frac{1}{2}} \end{aligned} \quad (2.3)$$

where N_1, N_2 are the number of particles in the two colliding bunches, n_b is the number of bunches in one beam, f is the revolution frequency of one bunch (with a design value of 11245Hz), A is the overlap area of the two bunches transverse to the beam, F is the geometric luminosity reduction factor of the instantaneous luminosity, ϵ_n is the normalized transverse beam emittance, β^* is the betatron function at the interaction point, γ is the relativistic Lorentz factor, θ_c is the crossing angle at the interaction point, σ_z and σ_{xy} are the transverse and longitudinal r.m.s. bunch sizes at the interaction point. In Table 2.1 are listed the summarized nominal values of the LHC machine parameters.

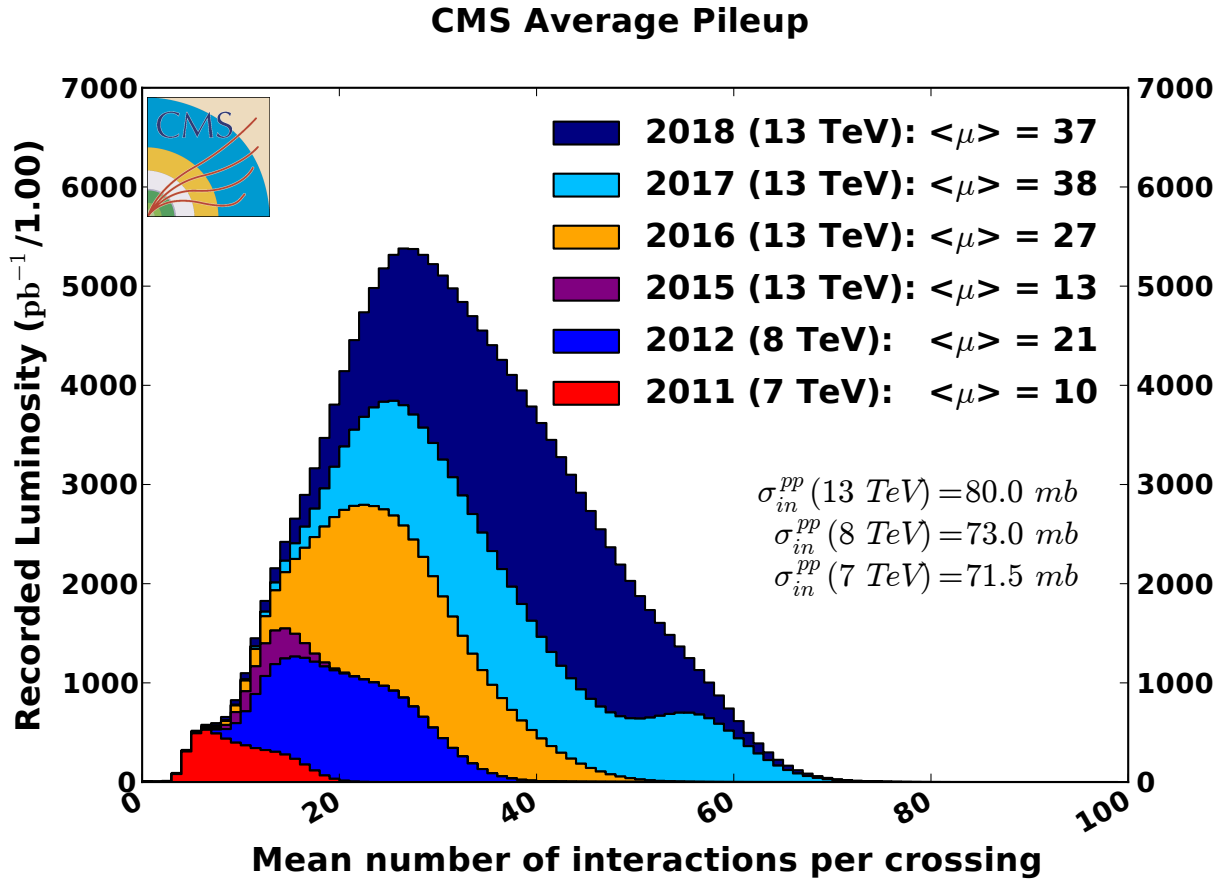


Figure 2.2. . The mean values of interactions per crossing (pileup) and the minimum bias cross sections for 2011-2012 and 2015-2018, where the distribution of each year is stacked above the previous year [124].

The nominal LHC instantaneous luminosity is $\mathcal{L} \approx 1 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ [124]. The integrated luminosity $\mathcal{L}_{int}$ is derived by integrating the instantaneous luminosity over a given period of time, as in Eq. 2.3 and the values over the time period of 2015-2018 can be seen in Figure 2.3. The integral is taken over the sensitive time, i.e., excluding possible dead time of LHC. It is expressed in inverse femtobarn (fb^{-1}):

$$\mathcal{L}_{int} = \int_{t_0}^{t_1} \mathcal{L} dt \quad (2.4)$$

The integrated luminosity multiplied by the interaction cross-section (σ_p) for a process, gives the total number of events (N) produced during a time period. This parameter characterizes the efficiency of obtaining physical information from LHC.

$$N = \mathcal{L}_{int} \times \sigma_p \quad (2.5)$$

On 20 th November 2009, LHC started its first proton-proton collisions. At this moment the LHC machine officially became the world's most powerful particle accelerator and entered in its Physics Run-I that lasted until February 2013. During Run-I of the LHC program, the center-of-mass energy was $\sqrt{s} = 7\text{ TeV}$, half of the nominal energy. Nevertheless, the energy and luminosity displayed during Run-I were enough for both CMS and ATLAS experiments to discover the Higgs boson [69]. After the first shutdown (LS1) for upgrades, the LHC started again in 2015 for Run-II with center-of-mass energy $\sqrt{s} = 13\text{ TeV}$, and on December 3rd 2018 ended with a last heavy ion run [32], as we can see in Figure 2.4. During the operations, the instantaneous luminosity was successfully brought to a value of $\approx 2 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, exceeding the design value. Run-II has

Parameters		Units	LHC Design
Center of mass energy	$\sqrt{s}$	TeV	13
Bunch Spacing	Δt_b	ns	25
Peak Luminosity	$\mathcal{L}$	$cm^{-2}s^{-1}$	2.16×10^{34}
Number of bunches per beam	n_b	N/A	2808
Proton per Bunch	N_b	N/A	1.1×10^{11}
Pile-up	PU	N/A	25
Revolution frequency	f	kHz	11.2
Transverse beam emittance	ϵ_n	μm	3.75
Beta function	β^*	m	0.55
Crossing angle at i.p.	θ_c	μrad	285
Transverse r.m.s. bunch size at i.p.	σ_{xy}	μm	16.7
Longitudinal r.m.s. bunch size	σ_z	cm	7.55

Table 2.1. Nominal LHC beam parameters [68]

been the occasion to acquire more data to study the properties of the Higgs boson with more precision. From the end of 2018 to early 2022 the LS2 will take place as we can see in Figure 2.4. This second maintenance stop will be the occasion to boost once again the beam energy to finally reach the design energy of LHC, $13.6 TeV$. The Run-III is scheduled to operate at $\sqrt{s} = 13.6 TeV$ and at $\approx 2 \times 10^{34} cm^{-2}s^{-1}$ as instantaneous luminosity. During these 3 years of run, the LHC will deliver as much integrated luminosity as what brought during the almost 7 years of both Run-I and Run-II of data taking. Phase-I will end with an overall $300 fb^{-1}$ delivered.

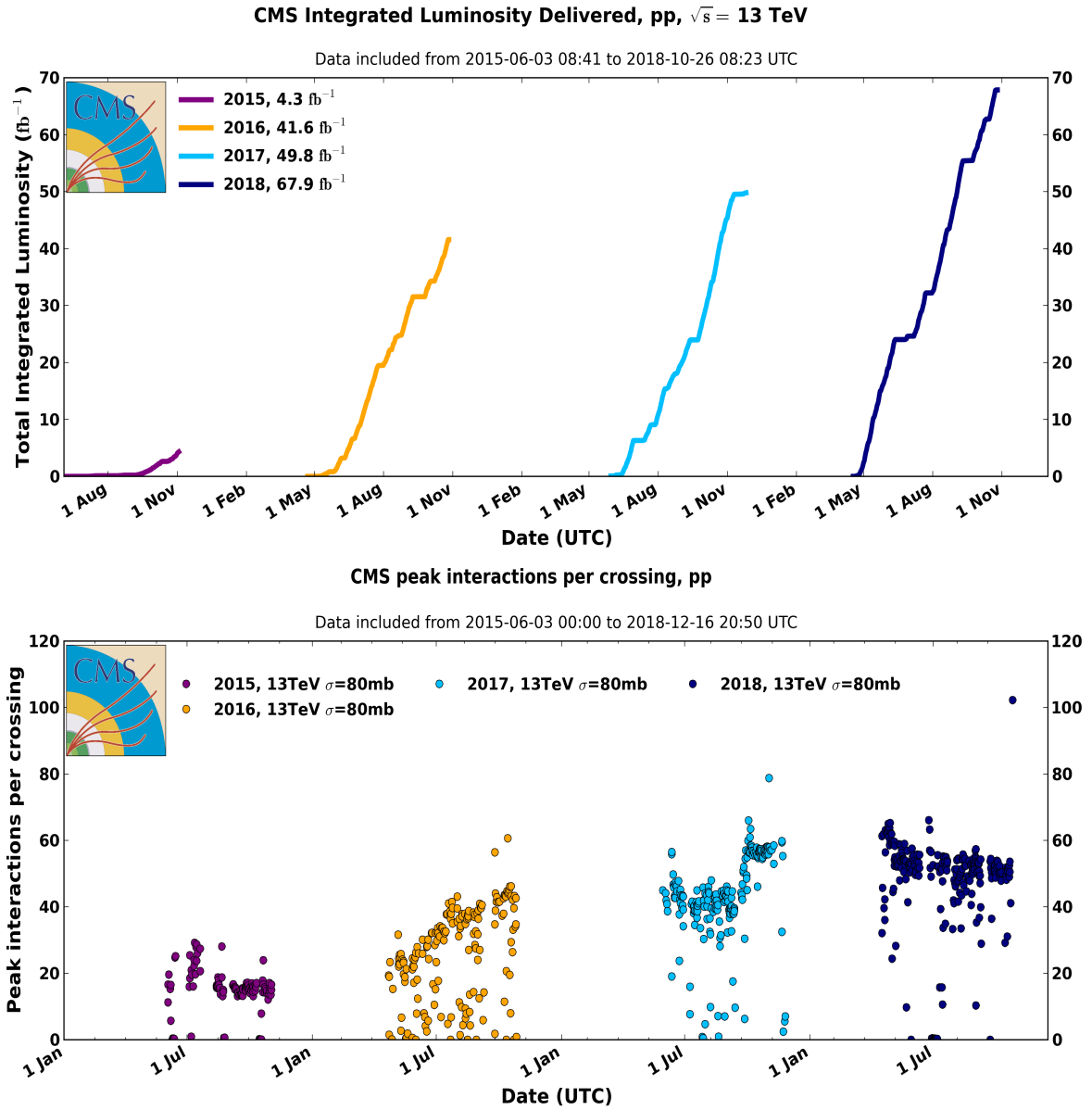


Figure 2.3. . Total integrated luminosity (top) and peak instantaneous luminosity (bottom) of the LHC as a function of the year, as measured by the CMS experiment [124]

2.2 LHC experiments

Along the accelerator ring there are four main interaction points (IP) where the two proton beams, traveling in separate vacuum pipes, are made to collide. The LHC detectors are located in the interaction points to allow physicists to test the predictions of different theories of particle physics.

The four main experiments can be seen in Figure 2.5 above and are briefly described below:

- A Large Ion Collider Experiment (ALICE) [29]: a particle detector located at IP2, dedicated to the measurements of the nuclear matter properties at high energy densities. The focus of ALICE detector is the study of Quark Gluon Plasma (QGP) and different features of heavy-ion physics by means of lead-lead and proton-lead collisions.
- Large Hadron Collider beauty (LHCb) [35]: a single-sided spectrometer located at IP8, built to investigate matter-antimatter asymmetry, focused on the physics with hadrons originating from bottom quarks



Figure 2.4. . Baseline schedule of LHC and HL-LHC operations [123]

and the measure of CP violation parameters. The biggest discovery of LHCb collaboration at CERN has observed the matter-antimatter asymmetry, known as CP violation, in the decays of a D_0 meson for the first time [170].

- A Toroidal LHC ApparatuS (ATLAS) [30]: a general-purpose particle detector located at IP1, designed to cover wide range of precise measurements of the possible physics processes, such as electroweak and strong physics. The detectors were formulated especially to provide the best possible sensitivity to Higgs boson and new physics such as Supersymmetry (SUSY), dark matter, extra-dimensions, etc. The main feature of the ATLAS detector is its enormous doughnut-shaped magnet system. This consists of eight 25m long superconducting magnet coils, arranged to form a cylinder around the beam pipe through the centre of the detector. ATLAS is the largest volume collider detector ever constructed.
- Compact Muon Solenoid (CMS) [33]: a multi-purpose particle detector located at IP5, with the same physics research program as ATLAS, but different technical solutions and design. CMS is discussed in details in the sections below.

2.3 The Compact Muon Solenoid (CMS) Experiment

The Compact Muon Solenoid (CMS) is a general-purpose detector at the LHC, having a broad physics programme, including studies such as phenomena of the Standard Model like the Higgs boson Physics to searches for physics beyond the SM (BSM) such as dark matter particles or even gravity related particles.

The CMS and ATLAS experiments have been built as complementary experiments, in order to compare and Physics phenomena, eventhough the technical design of detectors and of the geometry of the magnetic

LHC experiments and Detectors

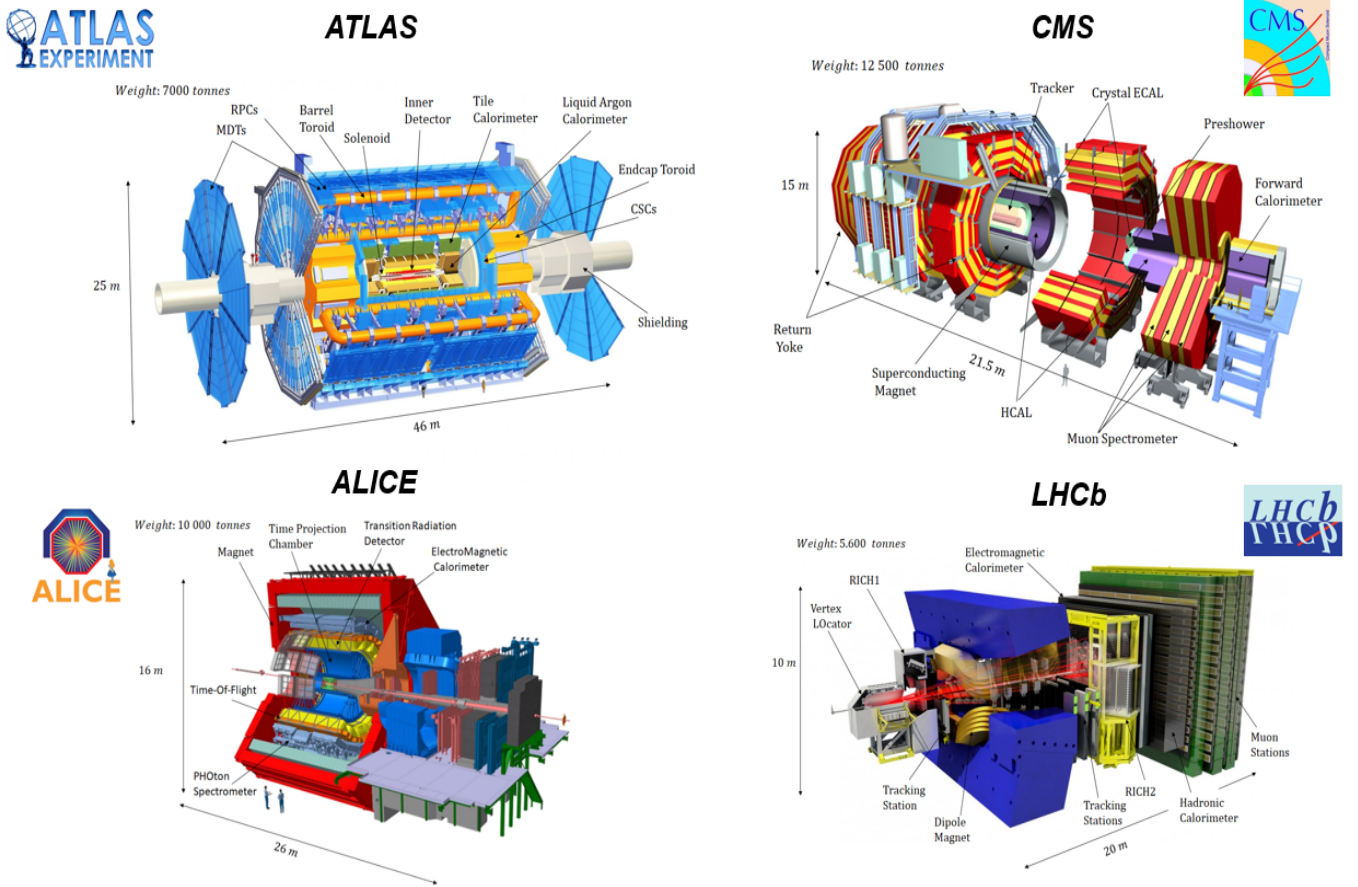


Figure 2.5. The LHC main experiments: ATLAS, CMS, ALICE and LHCb

field of both experiments differ. The CMS detector [164] has a cylindrical barrel shape closed by two endcap disks, with a full length of $\approx 21\text{m}$, a diameter of $\approx 15\text{m}$, and a total weight of about 14,000 tonnes. The CMS detector uses an onion like layout around the interaction point in order to maximize the acceptance. Each layer constitutes a sub-detector designed to stop, track or measure a different type of particle emerging from LHC collisions. The innermost layer is a silicon-based tracker. Surrounding it is a scintillating crystal electromagnetic calorimeter, which is itself surrounded by a sampling calorimeter for hadrons. The tracker and the calorimeters are compact enough to fit inside the CMS solenoid. Outside the magnet are located the large muon detectors, which are inside the return yoke of the magnet. The overall view of the CMS detector is shown in Figure 2.6. A Cartesian 0_{xyz} -coordinate system is chosen as the reference frame for the measurement of the physical quantities at the CMS detector, which is represented in Figure 2.7. Its origin is the interaction point, the 0_z -axis is collinear along the beam, the x -axis points radially toward the center of the LHC ring, and the y -axis is perpendicular to the $x - z$ plane and points upward. Reconstruction algorithms, however, use a spherical coordinate system based on the radial distance:

$$R = \sqrt{x^2 + y^2} \quad (2.6)$$

from the z -axis, the azimuthal angle ϕ with respect to the y -axis, and polar angle θ with respect to the beam line.

However, taking into account that the particles produced by proton collisions are strongly squeezed along to the collision axis and distributed over the angle θ very unevenly, it is customary to use a different, more convenient, kinematic variable called pseudorapidity (η).

CMS DETECTOR

Total weight : 14,000 tonnes
 Overall diameter : 15.0 m
 Overall length : 28.7 m
 Magnetic field : 3.8 T

STEEL RETURN YOKE
 12,500 tonnes

SILICON TRACKERS
 Pixel ($100 \times 150 \mu\text{m}$) $\sim 16\text{m}^2 \sim 66\text{M}$ channels
 Microstrips ($80 \times 180 \mu\text{m}$) $\sim 200\text{m}^2 \sim 9.6\text{M}$ channels

SUPERCONDUCTING SOLENOID
 Niobium titanium coil carrying $\sim 18,000\text{A}$

MUON CHAMBERS
 Barrel: 250 Drift Tube, 480 Resistive Plate Chambers
 Endcaps: 468 Cathode Strip, 432 Resistive Plate Chambers

PRESHOWER
 Silicon strips $\sim 16\text{m}^2 \sim 137,000$ channels

FORWARD CALORIMETER
 Steel + Quartz fibres $\sim 2,000$ Channels

CRYSTAL
 ELECTROMAGNETIC
 CALORIMETER (ECAL)
 $\sim 76,000$ scintillating PbWO_4 crystals

HADRON CALORIMETER (HCAL)
 Brass + Plastic scintillator $\sim 7,000$ channels

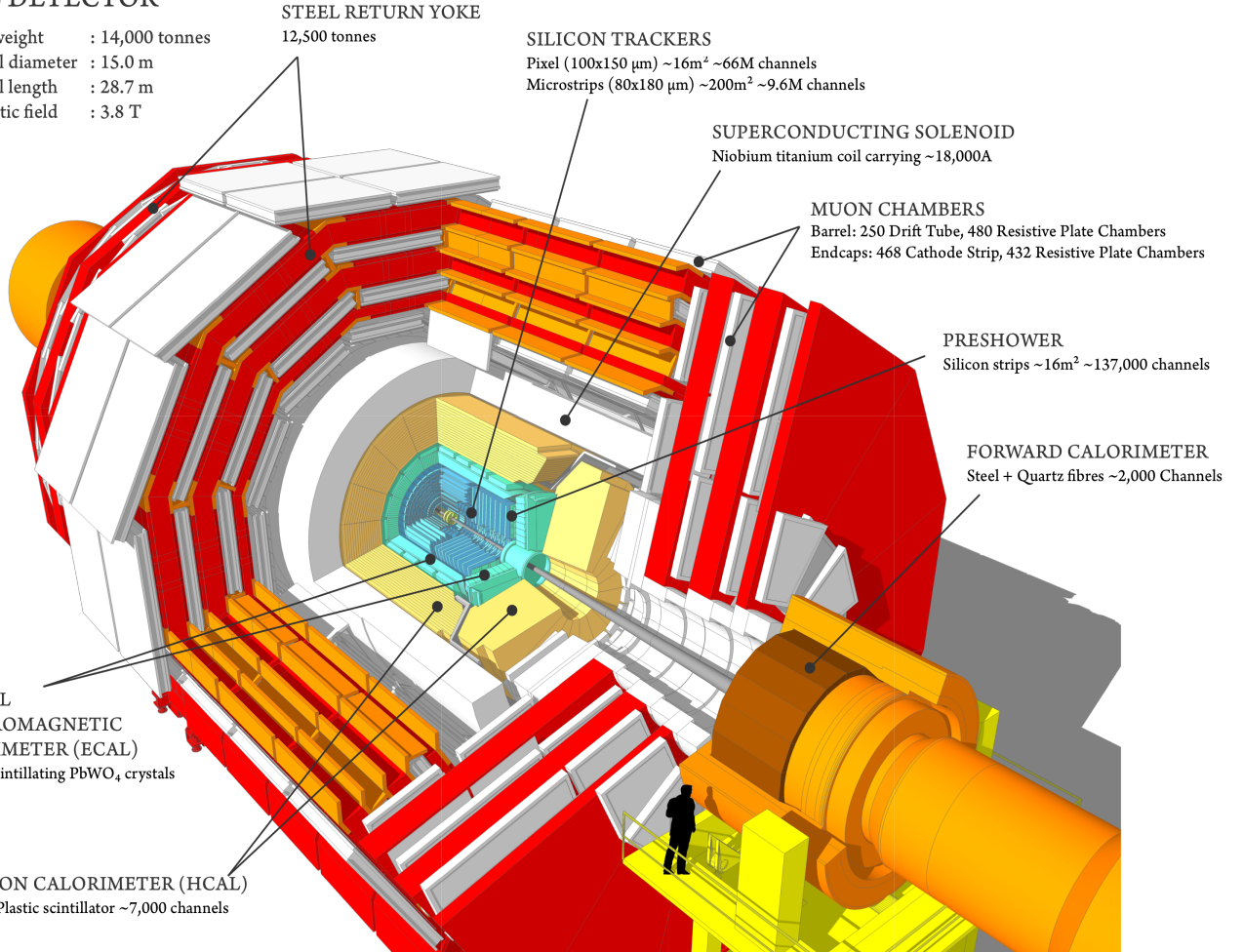


Figure 2.6. . Perspective view of the CMS detector layout

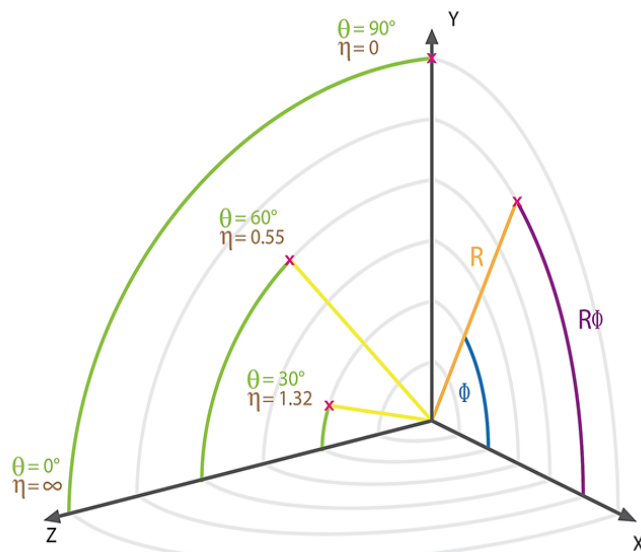


Figure 2.7. . The CMS coordinate system

The pseudorapidity η value is determined by the formula:

$$\eta = -\ln \left[\tan \left(\frac{\theta}{2} \right) \right] \quad (2.7)$$

2.4 The CMS Solenoid Magnet

A magnetic field is essential for identifying the particles emerging from collisions: it curves their trajectory allowing to calculate their momentum and to establish whether they have a positive or negative charge. Stronger magnetic field allows to obtain higher momentum measurements resolution. The CMS superconducting solenoid magnet has an internal diameter of $6m$, length of $12.5m$, and it is designed to produce a uniform field in the axial direction of $4T$. The flux return is assured by an external iron yoke with three layers, capable of stopping almost all particles (except muons and neutrinos), in between which the muon system is installed.

2.5 The CMS Tracker

The innermost sub-detector of CMS, surrounding the interaction point, is the silicon tracker [149]. Its aim is the accurate, robust and efficient tracking of the charged particles produced by the collisions. Of fundamental importance is the precise measurement of secondary vertices with a resolution of $10\mu m$.

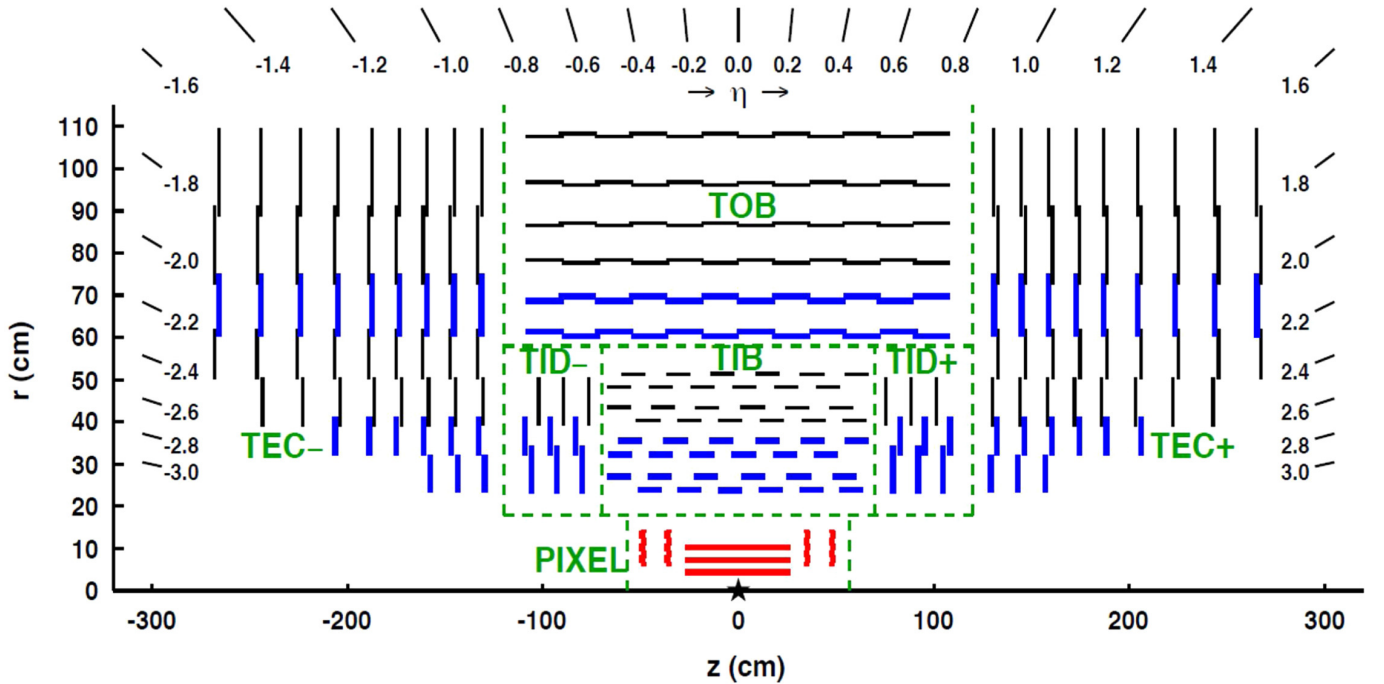


Figure 2.8. Overview of the CMS tracker system

At the LHC design, each bunch-crossing produces about one thousand particles crossing the tracker every $25ns$. This large number of tracks, coming from different primary vertices, needs a fast, high-granularity and radiation-resistant detector. The full volume of the tracker has a length of $5.8m$, a diameter of $2.5m$, and can provide tracking coverage up to $|\eta| < 2.5$. The tracker consists of two subsystems: the pixel tracker and the strip tracker. A general overview of the CMS tracker system is given in Figure 2.8, with its subsystems: Pixel, Tracker Inner Barrel (TIB) and Tracker Inner Disks (TID), Tracker Outer Barrel (TOB) and Tracker Endcaps

(TEC).

Pixel detector: the inner pixel detector is installed within a distance of 20cm from the beam pipe. It provides high-resolution 3-D measurements for the charged track reconstruction. In overall, the pixel detector consists of three barrel layers (BPix) with two endcap disks (EPix) on each side of the barrel.

Strips detector: in addition to the pixel detectors, the inner tracker is composed of several layers of silicon microstrip detectors: four-layer Tracker Inner Barrel (TIB), six-layer tracker outer barrel (TOB) and on each side three-disk Tracker Inner Disks (TID) and nine-disk Tracker Endcaps (TEC). The full tracker consists of 15148 microstrip detectors, with a pitch size ranging 80 – 180 μm .

2.6 The CMS Calorimeter

Two types of calorimeters are installed in CMS: the internal electromagnetic calorimeter (ECAL) and the external hadron calorimeter (HCAL). Electrons, photons and hadrons are stopped by the calorimeters allowing their energy to be measured. Neutrinos, as well as all the non interacting particles, escape direct detection but their presence can be inferred as an apparent energy imbalance in a collision.

2.6.1 The Electromagnetic Calorimeter (ECAL)

The CMS electromagnetic calorimeter [136] covers the pseudorapidity range $|\eta| < 3$ and its schematic structure is represented in ?? It is an hermetic homogeneous calorimeter consisting of 75848 lead-tungstate (PbWO₄) scintillating crystals, designed to fully absorb the energy of electromagnetically interacting particles, mainly photons and electrons.

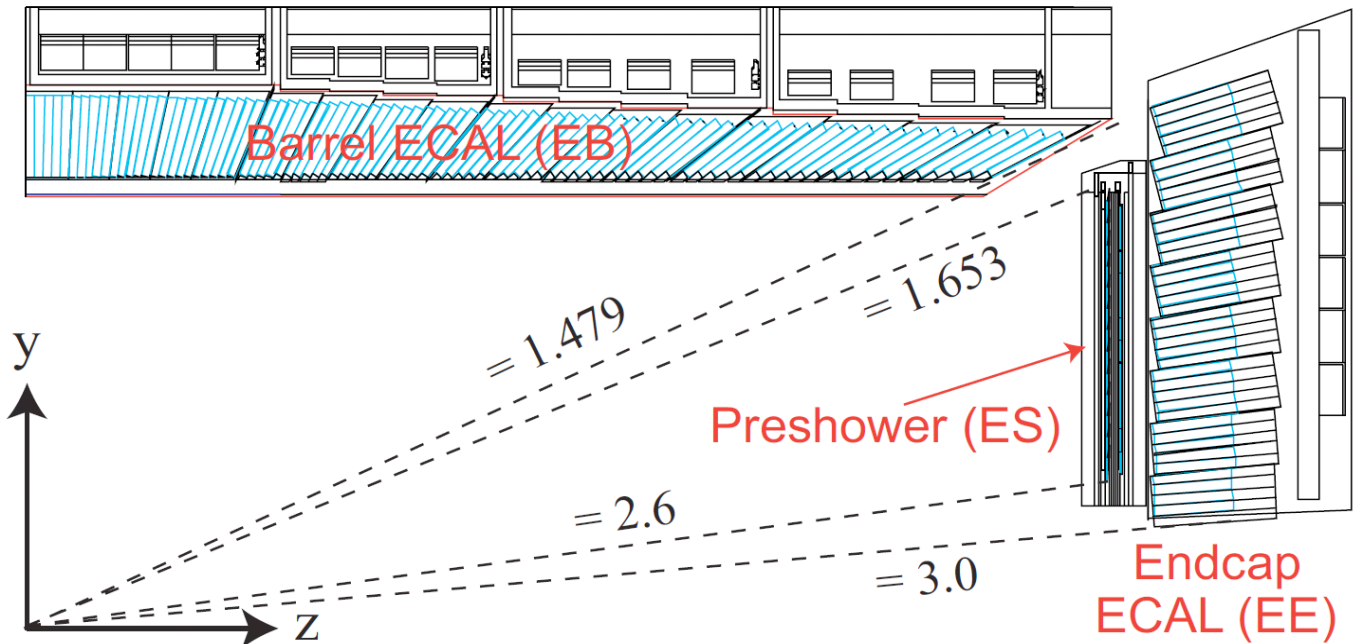


Figure 2.9. Longitudinal view of the ECAL layout, representing one quarter of the detector. The barrel and endcaps sections and the preshower detector are shown.

One of the driving criteria in the design was the capability to detect the decay into two photons of the Higgs boson ($H \rightarrow \gamma\gamma$). The reason to choose $PbWO_4$ material is due to its radiation resistance, high density of 8.28 g/cm^3 , small Molière length of 0.89 cm and a short radiation length 2.2 cm which results in high granularity and high energy resolution of ECAL. The ionized atoms in the $PbWO_4$ crystal induce the blue scintillating light ($\lambda \approx 420 \text{ nm}$), which is then measured in photodetectors. The total amount of scintillating light produced is proportional to the energy deposited in the crystals by the electrons and photons.

2.6.2 The Hadronic Calorimeter (HCAL)

The hadronic calorimeter [131] plays an essential role in the identification of quarks, gluons, and neutrinos by measuring the energy and direction of the jets and of the missing transverse energy flow in events.

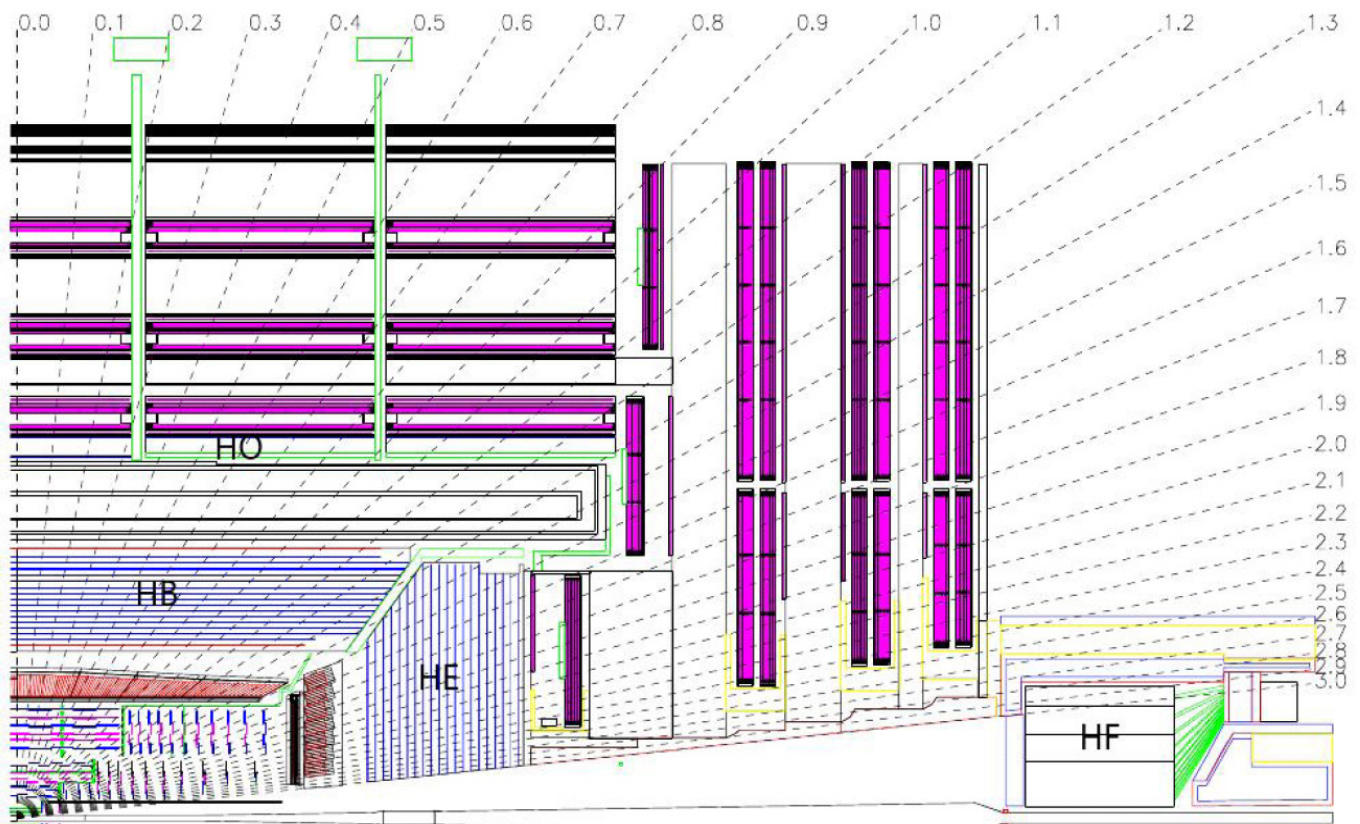


Figure 2.10. Longitudinal view of the HCAL layout, where are visible the barrel (HB) and endcaps (HE) detectors located inside the solenoid, the outer detector (HO) outside the solenoid, and the forward calorimeter (HF)

The reconstruction of jets and missing transverse energy (MET) is crucial for many analysis which involves neutrinos and neutralinos. The HCAL is a sampling calorimeter with brass as absorber, plastic scintillator tiles as active medium, and wavelength shifting fibers to transfer the light to the detector. The HCAL consists of about 70000 scintillating tiles and is divided into a barrel region (HB) up to $|\eta| < 1.3$, an endcap region (HE) up to $|\eta| < 3$ and a forward region (HF) further down the beam pipe at 11.2 m that increases the coverage up to $|\eta| < 5.2$, as we can see in Figure 2.10. The limited space for stopping power in the barrel region is overcome with an additional calorimeter outside (HO) the solenoid up to $|\eta| < 1.3$ that uses the coil and a 19.5 cm layer of steel yoke as the absorber. Muons and neutrinos produced in the interactions go unregistered in the HCAL except for a small muon MIP trace. The statistical fluctuations in the sampling and leakage degrade the performance of the HCAL in comparison to the ECAL.

2.7 The Muon Spectrometer

In order to provide an independent muon identification, robust trigger and accurate momentum and charge measurements, the muon spectrometer [127] is placed outside the magnet and it consists of 4 stations of detectors integrated into the iron return yokes.

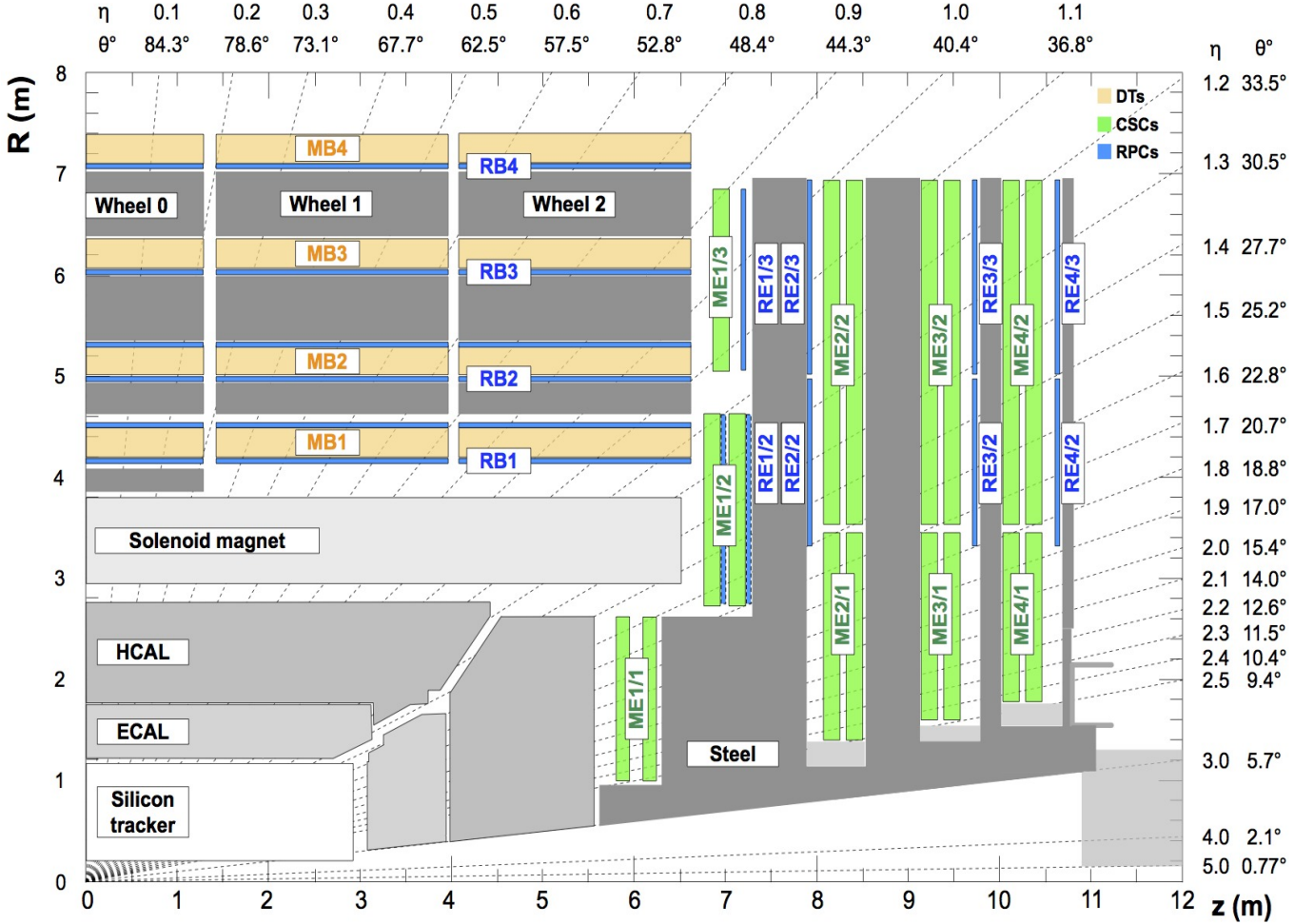


Figure 2.11. Schematic view of the CMS Muon spectrometer. The interaction point is at the lower left corner. The locations of the various muon stations are shown in color (MB = DT = Drift Tubes, ME = CSC = Cathode Strip Chambers, RB and RE = RPC = Resistive Plate Chambers). M denotes Muon, B stands for Barrel and E for Endcap. The magnet yoke is represented by the dark gray areas.

The muon identification is guaranteed by the amount of material in front of the chambers and in the return yoke of the magnet which shields the spectrometer from charged particles other than muons. The robustness and redundancy of the spectrometer is guaranteed by the presence of three types of gaseous detector: Drift Tubes (DT) in the barrel region, Cathode Strip Chambers (CSC) in the endcap and Resistive Plate Chambers (RPC) in both barrel and endcap regions. A schematic view of the CMS Muon spectrometer is reported in Figure 2.11. The barrel region is divided into five wheels (Wheel 0, Wheel ± 1 , Wheel ± 2). The muon chambers are installed on the outer and inner sides of the yokes, in the pockets between layers, arranged in four stations at different radius. Each station consists of 12 chambers, one per sector, except for the fourth one where 14 chambers are present. The two endcaps are made of 3 iron disks and 4 stations, divided in 2 or 3 rings. The rings are formed by 18 or 36 trapezoidal chambers that overlap in ϕ giving geometrical coverage close to 100%.

2.7.1 The Drift Tubes

In the barrel the muon detector does not operate in particularly demanding conditions since the background rate in this region is low and the magnetic field is well contained in the iron plates of the magnet return yokes [139]. For this reason, the DT were chosen to cover the pseudorapidity region up to $|\eta| < 1.2$. The basic detector unit is a drift cell: a gas-filled tube with rectangular cross-section showed in Figure 2.12.

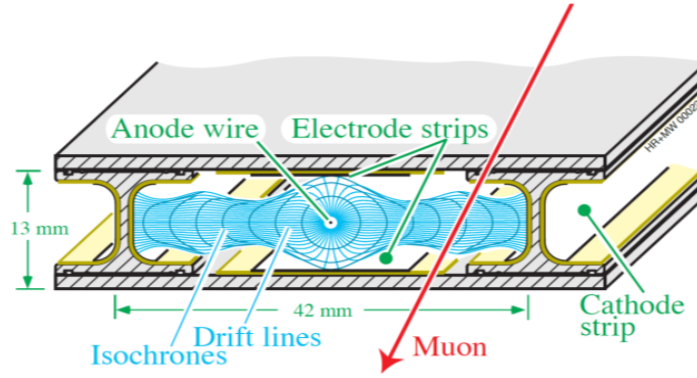


Figure 2.12. . Structure of a Drift Tube Chamber with three superlayers composed of 4 layers each of drift cells. Right: Layout of a drift cell showing the electric field lines in the gas volume.

The two shorter sides of the rectangle form the cathodes, while an anode wire is strung through the middle. A charged particle passing through the detector volume ionizes the gas, producing a cloud of electrons that drift toward the wire. The drift time is measured and converted to distance using the knowledge of drift velocity. A single drift cell has a cross-section of $42 \times 13 \text{ mm}^2$ and wire length of $2 - 3 \text{ m}$. It is filled with a 85% – 15% mixture of Ar/CO_2 , giving a 350 ns maximum drift time. Single wire measurement resolution is of the order of $200 \mu\text{m}$. The drift tubes in a chamber are grouped into SuperLayers (SL) consisting of four layers of tubes, staggered by half a tube. In each chamber there are two SLs with wires parallel to the beam direction, measuring muon position in the bending plane of the magnetic field. These are separated by a 128 mm thick aluminium honeycomb spacer, providing good angular resolution within one chamber. Additional SL measuring the η coordinate of the muon is present in the three inner stations.

2.7.2 The Cathode Strip Chambers

The CSC cover the endcap muon region. This type of detector has been chosen because of its capability to provide precise time and position measurement in the presence of high and inhomogeneous magnetic field, and high particle rate [155]. The CSC detector is a multi-wire proportional chamber with one of the cathode planes being segmented in strips running orthogonally to the wires. The principle of operation is shown in Figure 3.5 (right): a muon crossing the chamber produces an avalanche in the gas (40% – 50% – 10% mixture of $\text{Ar}/\text{CO}_2/\text{CF}_4$) that is collected by the wire. This induces an electrical charge on several adjacent cathode strips. The measured distribution of charge picked up by the strips gives an estimate of the position of the muon along the wire. The CSC technical construction is shown in Figure 2.13.

Each individual chamber has a trapezoidal shape and is made of seven cathode panels stacked together, forming six gas-gaps, each containing an array of anode wires. The gaps are 9.5 mm thick and one of the two cathode planes for each gap is segmented into radial strips orthogonal to the wires. The strips cover a constant area in ϕ ($2.33 - 4.65 \text{ mrad}$, depending on the disk). The orthogonal coordinate (R) is measured by the wires, which, to reduce the number of channels, are read out in groups of 5 to 16.

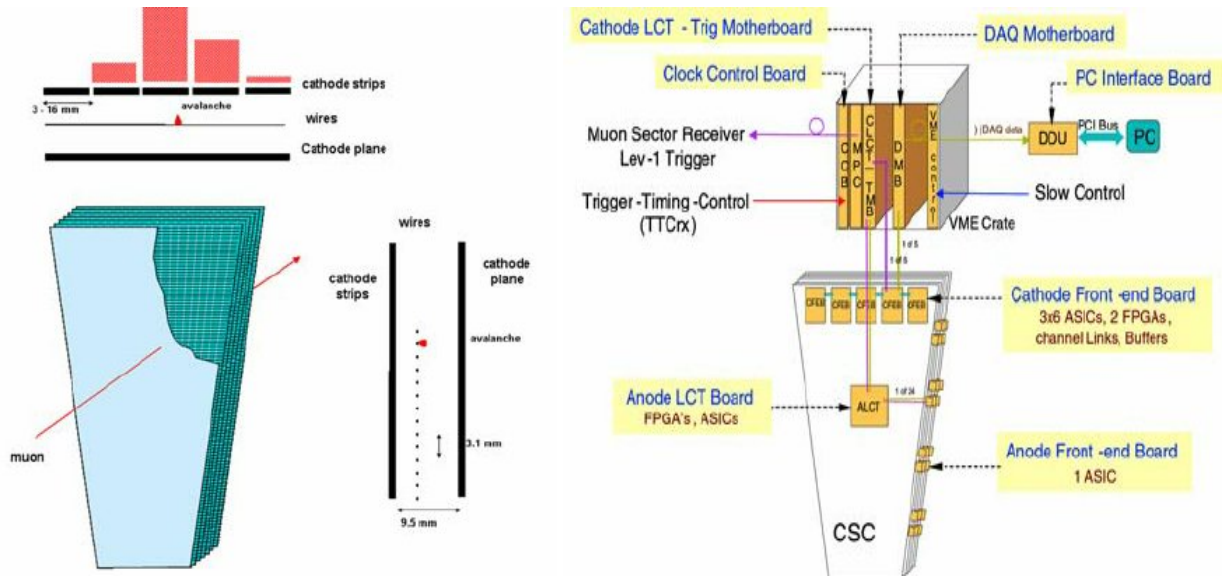


Figure 2.13. . 1) Left: Diagram and principle of operation of a CSC endcap muon chamber in CMS; 2) Right: Schematic of the on-chamber and peripheral crate electronics system.

2.7.3 Resistive Plate Chambers

The RPC detectors are used in CMS in the barrel region along with CMS Drift Tubes, and in the end-cap region, along with CMS Cathode Strip Chambers. The RPCs have an excellent time resolution about 1.5ns, this time is shorter than the 25ns LHC bunch spacing, therefore RPC detectors are used for muon (μ) triggering purpose. A total of 1056 RPC detectors are present in CMS muon system covering $|\eta| < 1.9$ region. The RPCs geometry depends on their position in the detector. There are six layers of RPC chambers in the barrel region, four of them are attached to each side of DT chambers (MB1 and MB2), the other two layers are attached to the inner side of MB3 and MB4. In end-caps region; there are four disks of trapezoidal RPCs attached to the CSCs chambers. The CMS RPC chambers consist of two layers of double gap geometry with a sheet of segmented copper readout strips sandwiched between them. The chambers operate with a humidified gas mixture composed of 95.2% $C_2H_2F_4$ (Freon-R134a), 4.5% $i - C_4H_{10}$ (isobutane) and 0.3% SF_6 (Sulphur Hexafluoride). In the Chapter [3], the RPCs detectors are discussed in more detail.

2.8 The Trigger System

The data collected for each interaction, at nominal LHC condition, would require 40 TB of storage per second which is not feasible for CERN main computing farm, as its processing is limited by CPU performance and storage capacity.

Most of the events, from soft interaction between the protons known as “minimum bias” events, are not interesting. The required reduction of the event rate, to save the interesting information, is achieved by the so-called trigger system [150]. The aim of the trigger is the selection of potentially interesting events and the reduction of the rate to a manageable value of a few hundred Hz. While most high energy physics experiments use three trigger levels, the CMS trigger system implements a two-stage, rate reduction which provides additional flexibility. The first is the Level-1 (L1) trigger, largely based on Field Programmable Gate Arrays (FPGAs) and Application Specific Integrated Circuits (ASICs), that reduce the rate from 40 M Hz to 100 kHz with the latency 1 of $3.2\mu s$ [165]. The L1 trigger uses the raw information from Calorimeters and Muon system, and it searches for key signatures of interesting events: leptons, photons, hadronic jets, and p miss, known as trigger objects. The L1 trigger from muon and calorimeter are then combined in the L1

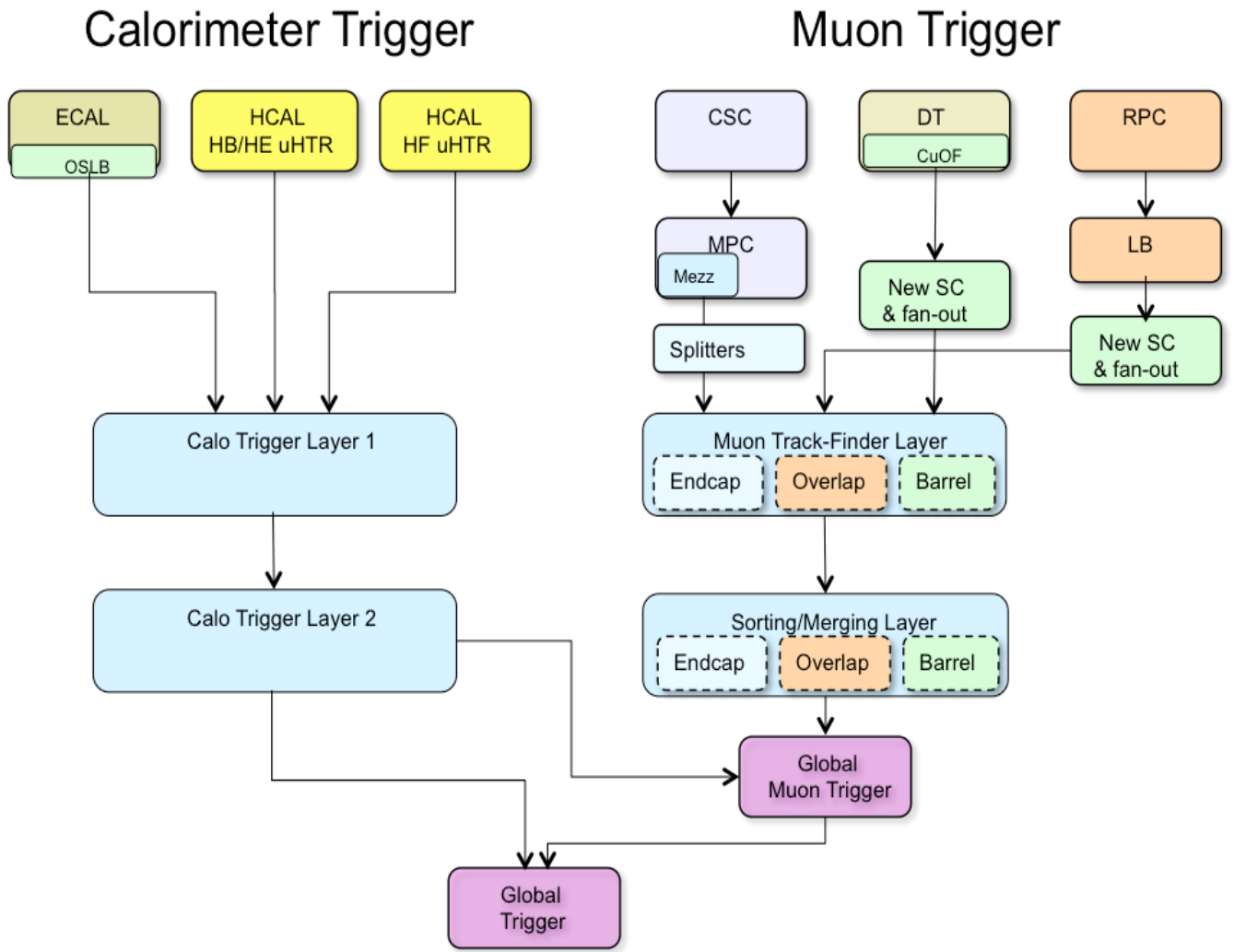


Figure 2.14. . Achitecture of the L1 system [?]

Global Trigger (GT) which decides whether to pass or reject the event. The data are temporarily stored in pipelines of processing elements while waiting for the decision. Figure 2.14 shows the schematics of L1 trigger system. The L1 muon trigger is described in more details in chapter [4]. The events which pass the L1-trigger requirements are selected and passed to the High-Level Trigger (HLT) [130]. The HLT is a software-based system, implemented in computing facility with approximately 16000 CPU cores, that reduces the L1 trigger rate to a viable level for storage of about 100Hz, with a processing delay time of about 100 ms. HLT can be used to reconstruct basic physics objects like jets, leptons and p_T , MET etc. To pass the HLT, an event needs to satisfy the requirements of at least one of its paths, which are in HLT menu. The events which passed the HLT are recorded permanently on the disk by the data acquisition (DAQ) system for further physics analysis and transferred to CERN Tier0 (T0) in one or more data streams.

2.9 The CMS upgrade for HL-LHC

The increase in energy and in luminosity foreseen during HL-LHC has many implications on the CMS subdetectors, a needed upgrade with challenges. The possible workflow for the future of the LHC is shown in Figure 2.15. The higher pileup, up to 140 to 200 interactions per bunch crossing, will lead to worse background conditions, higher trigger rates, more difficult event reconstruction, and accelerated aging of components. To cope with the much higher collision rate, and to fully profit from it, the CMS experiment,

including the muon detectors, need major upgrades. The basic goal of the CMS experiment upgrade is to maintain the excellent performance of the detector subsystems in terms of efficiency, resolution, and optimize the reconstruction of the particles at higher luminosities. A brief description of the upgrade program for each subdetector is given in the following, with a detailed section for the muon system upgrade.

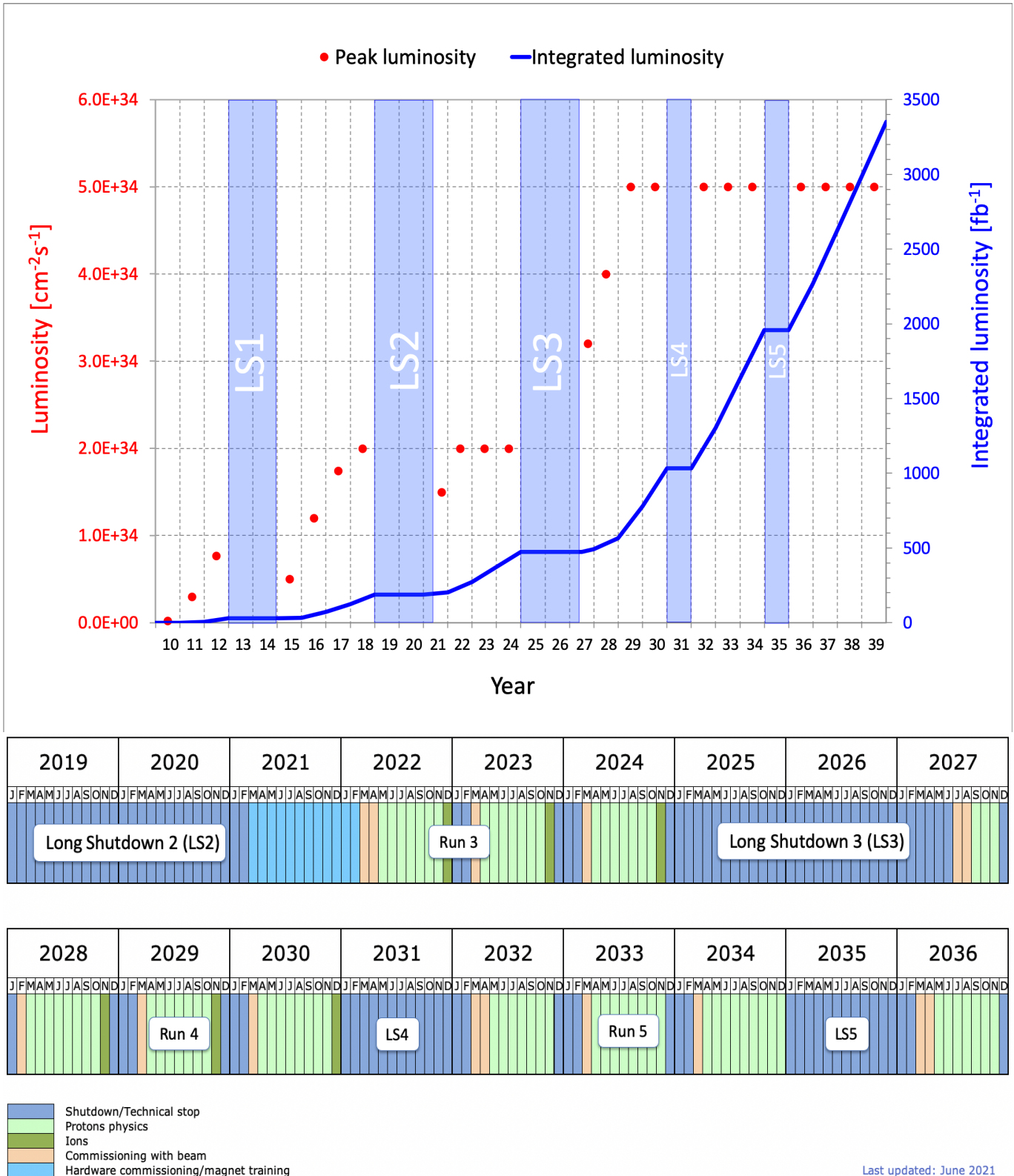


Figure 2.15. Expected LHC instantaneous and integrated luminosity as a function of time up to 2037 before the pandemic (top) and after the pandemic (bottom). Dates for long-shutdowns (LS) periods are also marked [159]

- Tracker: the CMS pixel will not be able to operate at the HL-LHC conditions. The pixel upgrade already started during the year technical stop of 2016/2017 with the replacement of the present pixel detector [149]. Nevertheless, because the entire tracker will suffer significant radiation damage, by LS3 the tracker must be completely replaced for Phase-II [149]. The granularity of both the outer tracker and the pixel systems will be increased by a factor four. Design improvements will lead to a much lighter outer tracker, improving, therefore, the p_T resolution and providing a lower rate of γ -conversions. Moreover, the module design will provide track-stub information to the L1 trigger at 40MHz for tracks with $p_T > 2\text{GeV}/c$, ensuring meaningful power rejection at the earliest stage of the event selection. The pixel system will implement smaller pixels and thinner sensors for improved impact parameter resolution and better two-track separation. The system coverage will be extended to almost $|\eta| = 4$, thanks to the addition of up to 10 additional pixel disks in each of the forward regions.
- Calorimeters: for Phase-II the replacement of the electromagnetic and hadronic endcap calorimeters due to the possibility of the radiation damage has been planned by LS3 [126]. The new calorimeter is called the High Granularity Calorimeter (HG-Cal). Its electromagnetic sensor consists of 28 tungsten plates interleaved with silicon sensors as the active material.
- Trigger: concerning the Level-1 trigger, the electronics for the calorimeter trigger, muon trigger, and the global trigger are planned to be improved [150]. For this aim, during the Phase-I the high bandwidth optical links for most of the data communication between trigger cards were installed, and modern, large FPGAs and large memory resources for the trigger logic were used. Using optical links allows changing the architecture very fast, while large FPGAs allow algorithms to evolve as needed. For Phase-II, the higher L1 accept rate is necessary to cope with the HL-LHC luminosity without raising the trigger thresholds. The increase in latency from 3.6 to 12.5 μs is a direct consequence of the new Level-1 track trigger envisioned for the upgraded silicon tracker. Thus, the front-end electronics in some present sub-detectors must be upgraded.
- The minimum ionizing particles (MIP) Timing Detector (MTD) [169]: The MTD is a new detector planned for CMS during the High Luminosity LHC (HL-LHC) era. Its function is to measure with high precision the production time of minimum ionizing particles (MIP). The MTD goal is to disentangle the large number of expected pile-up interactions (around 200) that will occur in each bunch crossing of the LHC. It also provides new capabilities for charged hadron identification and the search for long-lived particles. The MTD will give timing information for MIPs with 30 – 40 ps resolution at the beginning of HL-LHC operation and will slowly fall to lower time resolution as a result of expected radiation damage to 50 – 60 ps by the end of HL-LHC operations. This time resolution will help to assign with high precision charged tracks to the correct interaction vertices in bunch crossings with an average of 200 collisions. The working principle of the MTD is to assign associating tracks from a vertex to hits and their corresponding times, such that the vertexes can be reconstructed at the time the vertex collision occurred. This high time resolution of the MTD will help to distinguish between the tracks of very close vertices in space. The MTD will provide timing in the barrel and endcap regions, with a hermetic angular coverage up to a pseudorapidity $\eta \pm 3$. The time of electromagnetic showers will also be determined to a precision of 30 – 50 ps above a transverse momentum of few GeV, so that we can associate photons to the correct charged particle vertex, based on compatibility with the time obtained from the MTD.

The CMS Muon system upgrade

- Background rate [148]: the HL-LHC will produce well above 100 proton-proton interactions for every bunch crossing, taking place every 25ns, yielding billions of secondary particles each second. Thus, detectors and electronics have to be upgraded to handle the rate.

- Longevity [148]: detector and electronics aging is strongly affected by the radiation. In general the performance of gaseous detectors deteriorates with an increase of the integrated charge released in the gas volume. All present gaseous detectors of the CMS muon system, and a large fraction of their electronics, are expected to stay also for HL-LHC, so their longevity must be validated.
- Redundancy and resolution [148]: a major challenge at the HL-LHC is the reliable triggering on muons. Both the muon trigger and the offline muon identification and reconstruction can only work well if a sufficient number of detector hits is measured for each muon track. currently, the high η -region is not fully equipped with all the muon chambers foreseen for Phase-I, and in addition the two endcap regions suffer from high muon and background rates and reduced magnetic bending. The number of hits measured per track must be increased to enhance the redundancy of the muon system and to obtain a robust track reconstruction. The Phase-II muon detector upgrade foreseen the installation of new forward muon detectors covering the pseudorapidity range up to 2.4. The first ring of the disks 1 and 2 will be equipped with Gas Electron Multiplier (GEM), GE1/1 and GE2/1. While, the first rings of the third and fourth disks will be equipped with improved RPC (iRPC) detectors, RE3/1 and RE4/1. In addition to GE1/1 and GE2/1 station, the ME0 station will be also instrumented with GEM detectors, extending the acceptance of the muon system further to $|\eta| < 2.8$. The technologies of the new detectors have to handle the particle hit rates at maximum HL-LHC luminosity. The new detectors provide further high resolution measurement points and significantly increase the lever arm for track reconstruction.
- New ecogas mixture [148]: recent European regulations demand the reduction of the greenhouse gases (GHG) emissions [50]. The EU regulation on fluorinated gases is aimed to limit the total amount of F-gases that can be sold in EU from 2015 onwards with a reduction to one-fifth from 2014 to 2030, banning the use of F-gases where less harmful alternatives are available. Despite GHGs would remain available for research applications, CERN is committed to follow the regulations, and so is the CMS Experiment. In order to comply with these constraints, the gas mixtures of some CMS muon subdetectors, such as CSC and RPC, must be adapted, or the release of greenhouse gases into the atmosphere must be substantially reduced.

The upgrade plans for each subdetector are briefly described below, and the following chapter will be dedicated to the RPC upgrade project.

- Drift Tube [148]: the expected hit rates in the DT chambers at the HL-LHC will be up to about 50kHz per drift cell. The highest rates are expected for the MB1 and for the MB4 chambers. The highest integrated charge in DTs per unit of wire length is estimated to be of the order of 20mCcm^{-1} . Measurements performed on a spare detector show that the DT chambers can handle the hit rate expected during Phase-II. Longevity tests demonstrate that all MB2 and MB3 chambers, as well as the MB1 chambers of the three low- z wheels will maintain their current level of performance during Phase-II. Only in the MB1 stations in the outer wheels and in the MB4 detectors, a significant drop of the hit efficiency is predicted. The loss of efficiency is due to a decrease of the gas gain caused by deposits forming on the anode wires. The main culprit is believed to be outgassing of DT materials. Furthermore, possible mitigation measures of the aging effects are being actively studied. They include a decrease of the high voltage and the front-end electronic thresholds. For MB4 chambers, the background rates can be reduced by improving the neutron shielding of these detectors. The DT electronics will be replaced since the survival is not guaranteed, and it will not be able to cope with the increase of the Level 1 trigger acceptance rate beyond 500kHz .
- Cathode Strip Chambers [148]: the charge corresponding to 10 years of HL-LHC running, is expected to be $\approx 0.2\text{Ccm}^{-1}$ per unit of wire length for the CSCs closest to the beam line. Longevity tests of spare

chambers demonstrate that any significant deterioration of chamber performance, such as gas gain, detection efficiency, spurious signal rates, strip-to-strip resistance, or dark currents has been observed after having integrated around twice the amount of charge expected. Thus, it is expected that the CSCs themselves can be used until the end of the HL-LHC operation. While the CSC modules are expected to provide satisfactory performance throughout the HL-LHC program, some CSC read-out electronic boards will need to be replaced in order to handle the more stringent trigger requirements. In general, the new boards will cope with higher L1 trigger rates and chamber occupancies by employing high-speed output optical links and faster processors. Similarly, the extended trigger latency requirement can be met with deeper (digital) buffers

- Gas Electron Multiplier [148]: GEMs are gas detectors, characterized by a spacial resolution of order $\approx 100\mu m$, and time resolution of order $\approx 8ns$ at the efficiency plateau for Ar/CO₂ (70/30) as gas mixture [45]. By combining three foils to a triple-GEM chamber, the gas gain reaches a value of the order of 10^4 . The gas gain is observed to be stable up to $150kHzcm^{-2}$, which is the maximum rate expected during HL-LHC. The GE1/1 and GE2/1 detectors are "super-chambers", each made of a double layer of trapezoidal triple-GEM detector. They will cover 10° in GE1/1 station and 20° in GE2/1 station, overlapping in ϕ just like the corresponding CSC detectors. Therefore, 36 super-chambers will be installed for the GE1/1 station and 18 for the GE2/1 station, in each muon endcap in order to ensure the full azimuthal coverage. The ME0 detector station comprises 36 module stacks (18 per endcap), each composed of six layers of triple-GEM detectors, compared to the two-layer design of GE1/1 and GE2/1, covering the region $2 < |\eta| < 2.8$, in order to allow proper rejection of neutron background, particularly intense in that region. The large η coverage foreseen would lead to an improvement of the acceptance, with a consequent boost in the signal over noise ratio in the signature with muons, especially with the multiple-muon final states, and better hermeticity for signature without muons.

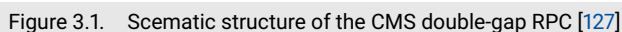
Part III

The Resistive Plate Chambers

Graphite resistivity studies on Phase 2 improved RPC detectors

3.1 The RPC chamber: Working principle

They are composed of four high pressure laminate (HPL) planes with a thickness of 2mm and with a resistivity in order of $1 - 6 \times 10^{10} \Omega\text{cm}$ [132], which serve as electrodes. The HPL plates in each gap are separated by a $10 \times 10\text{cm}^2$ grid of polycarbonate spacers in order to maintain thickness and space between the plates and as well as mechanical stability. The inner electrodes surfaces are covered with a $200\mu\text{m}$ conductive graphite plate, thus forming the high voltage and ground electrodes. The surface of the gaps is coated with a $5 - 10\mu\text{m}$ thick linseed oil, in order to smoothen the surface and quench the Ultraviolet (UV) photons [67]. The two gaps are placed between the readout copper strips, isolated from the graphite by mylar sheets. A schematic layout of a CMS double-gap RPC chamber is shown in Figure 3.1. They have a positively-charged anode and a negatively-charged cathode.



57

- 95.2% freon ($C_2H_2F_4$): Ensures the electrons drift velocity, important for high timing properties ($\sim ns$) of RPCs;
- 4.5% isobutane ($i - C_4H_{10}$): They are polyatomic molecules and are used as quencher, since they absorb UV photons and thus, they contain the avalanche development;
- 0.3% sulphur hexafluoride (SF_6): It is used to control the excess electrons through its high electron affinity (electro-negativity);

The gas mixture $C_2H_2F_4 + i - C_4H_{10} + SF_6$ configuration requires a working point at 9.6 kV. Since the electrode resistivity is relatively high, a property which could increase the HPL to dry out, this situation is solved by ensuring the gas mixture with a relative humidity of $\approx 40\%$.

The RPCs are implemented in barrel and end-caps regions of the CMS detector, thus changing respectively the geometry of the readout strips, as we can see in (Figure 3.2):

- Barrel: The readout strips adapt a rectangular shape
- End-cap: The readout strips follow a trapezoidal shape

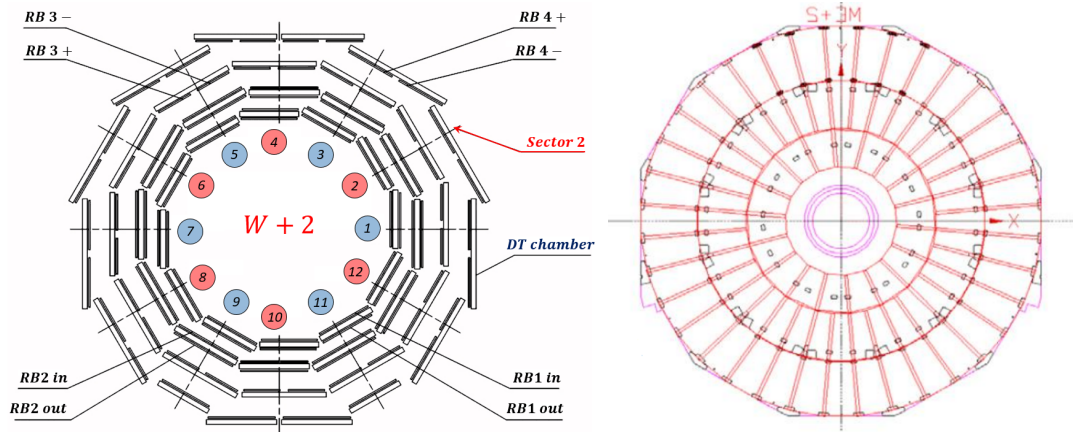


Figure 3.2. . Schematic view of the barrel wheels (left) and schematic view of the RPC endcap station (right)

The working principle of the RPC detector, as a gaseous detector, is the ionization of the gas molecule from an incoming charged particle (e.g. charged muons $\mu^\pm$) with a kinetic energy E_{kin} , which is higher than the potential energy of the electrons of the gas molecules in their respective energy levels. In case of interaction, the incoming free charged particle ejects the electron of the molecule, ionizing it and creating a pair of a positive charged ion and a negative charged electron (primary ionization). In case the incoming particle and the ejected electron/electrons have further E_{kin} , high enough to further ionize other gas molecules, they say, a secondary ionization has occurred. An external electric field is applied to accelerate the electrons, which triggers further ionization in the gas, thus creating an avalanche of electrons in the gaseous medium. A fast charge signal in the read-out copper strip is induced from the drift of the charge in the gas gap towards the anode. The electromagnetic induction of the copper read-out strips represents the signal of the RPC. Since the CMS RPC detectors are double gap gaseous detectors, the total induced signal is the sum of the two single-gap signals. This configuration together with the gas mixtures, makes possible operate single-gaps at lower high voltage with an improved effective detector efficiency, thanks also to the OR logic of the single-gap efficiencies, as we can notice in the graph below (Figure 3.3):

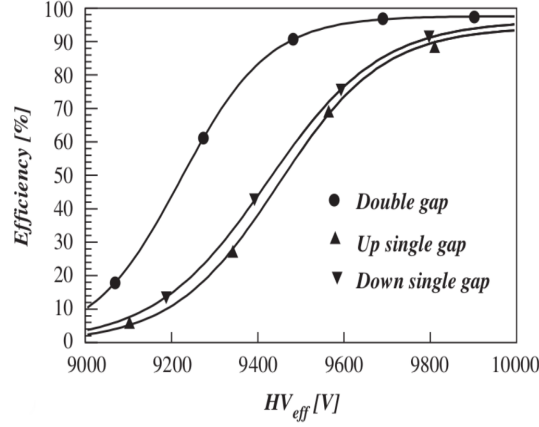


Figure 3.3. Comparison of efficiency curves of CMS RPC double and single-gap, measured with cosmic muons

The performance parameters of the present (Run II) RPC systems are as following [127] [148]:

- High detector efficiency: $> 95\%$;
- High rate capability: $\sim 300\text{Hz}/\text{cm}^2$;
- Average cluster size: ~ 2 strips;
- Time resolution: $< 1.6\text{ns}$;
- Intrinsic Noise: $< 5\text{Hz}/\text{cm}^2$;
- Spatial resolution: ~ 10 mm.

3.1.1 The RPC system upgrade for HL-LHC

The CMS RPC system is contributing to the muon system trigger due to the RPC's time resolution, which is about 1.5ns and it allows the capability of tagging an ionizing event in time shorter than the 25ns between two LHC successive bunch crossing. The performance of the RPCs is crucial for the muons system trigger in order to assign the measured muon to the right bunch crossing.

The LHC will be upgraded for the High Luminosity LHC Project (see section [2.9]), which will require to resist high radiations for a period of almost ten years, a stress test for the hardware and the electronics of the whole Large Hadron Collider . The total integrated luminosities between will achieve values about 3000fb^{-1} up to 4000fb^{-1} [160], and an instantaneous luminosity up to $7.5 \times 10^{34}\text{cm}^{-2}\text{s}^{-1}$ [160] followed by up to 200 collisions per bunch crossing (25ns to 1.6ns) as well as a sampling frequency from 40MHz (present system) to 640MHz (upgraded) [148]. Furthermore, due to the relative high instantaneous luminosity, also the frequency of hits per unit area (rate) will increase, requesting an improvement on the rate capability of the new RPC chambers. The current rate for the RPC modules is $180\text{Hz}/\text{cm}^2$ [160] along with the safety factor of three, meanwhile in case of the HL-LHC will be up to $3 \times 700 \approx 2000\text{Hz}/\text{cm}^2$ [160] for the new RPC detectors in the high η region of the endcap ring sectors RE3/1 and RE4/1. In terms of total integrated charge, the new modules and read-out electronics must sustain up to $3 \times 0.28 = 0.84\text{Ccm}^{-2}$ [148] for the full Phase-2 running period.

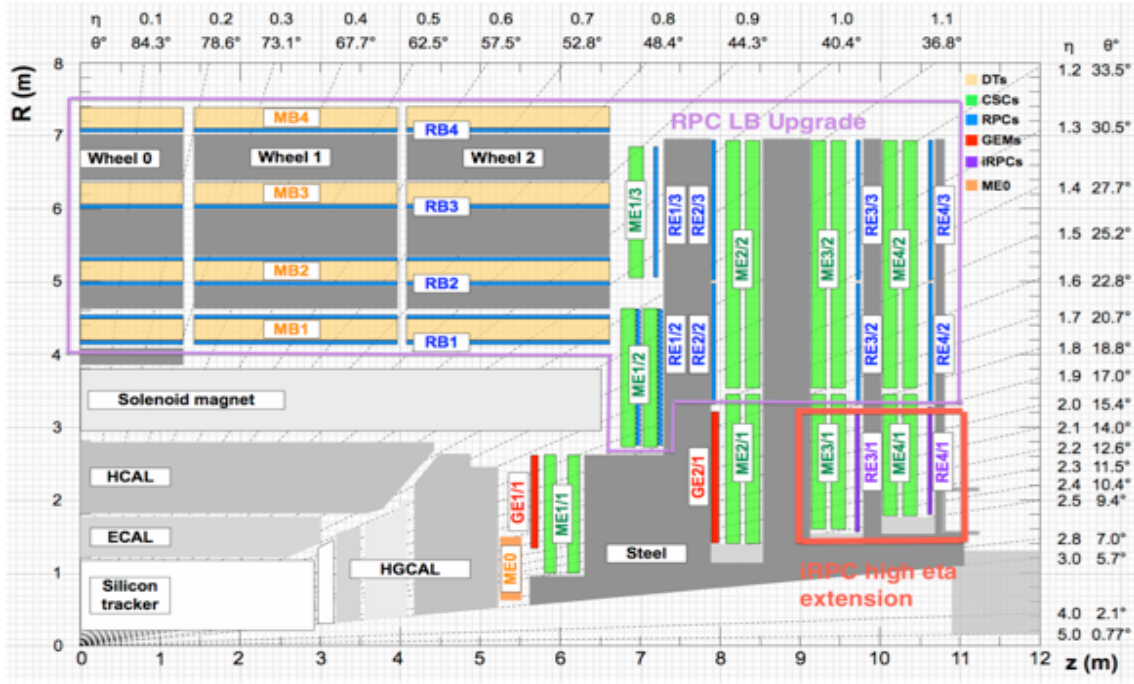


Figure 3.4. Extention the inner ring of the third (RE3/1) and fourth disk (RE4/1) of the endcap stations from $|\eta| < 1.8$ up to $|\eta| < 2.4$ shown bottom right in the orange box.

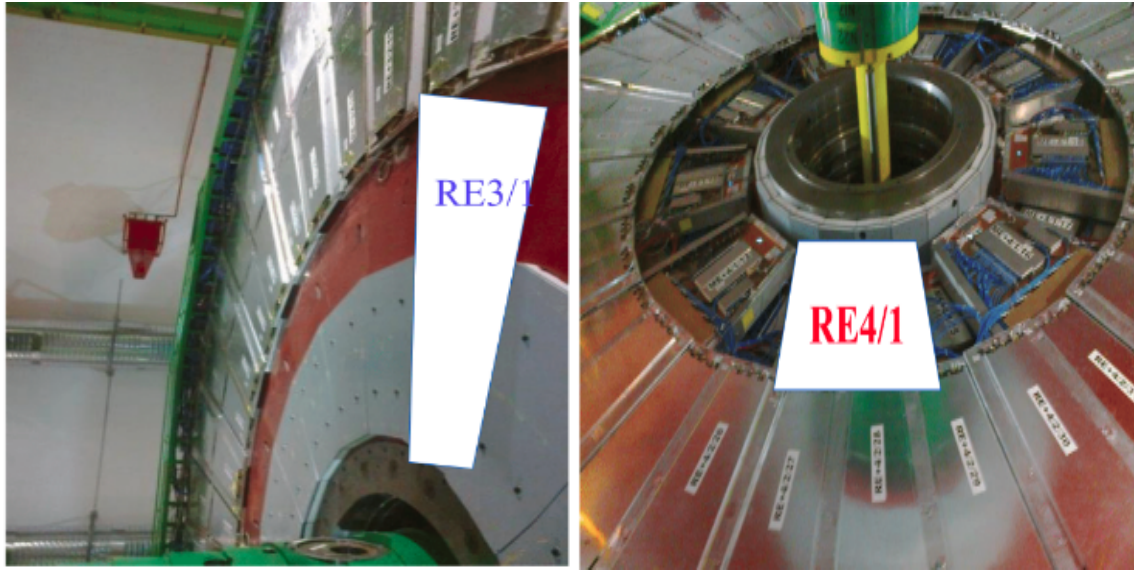


Figure 3.5. . Scematic representation of the planned extension of the RE3/1 (left) and RE4/1 (right) endcap.

For the RPC system two major upgrade are foreseen. The current RPC chambers can operate until the end of Phase-2, but the link system, which connects the front-end board to the trigger processors, must be exchanged, because is made of electronic boards using electronics components of the front-end boards, which are not certified for the full HL-LHC life cycle. This problem will be solved by the implementation of an optical link system, where the present link boards, which use ASICs, will be replaced by fast FPGAs [148], processing several signals from many RPC detectors in parallel. The data will be send in parallel to the main computer via Ethernet. In addition, new detectors for the forward region, RE3/1 and RE4/1, 72 RPCs are proposed to an extension up to $|\eta| = 2.4$, in order to increase the muon system redundancy and to improve the muon tracking and triggering in the forward region.

3.2 The Performance Studies of the double coated 1/2 size i-RPC

The double coated iRPC graphite chamber will be tested with two different resistivity values of the hot-coated and melted graphite on top gas gap planes. The values used for this test were $600\text{k}\Omega$ and $50\text{k}\Omega$. The goal is to understand how these values will effect on the suppression towards the dispersion of the avalanche charges and will then be the results on muon cluster size, efficiencies and working point WP (in kV) of the improved RPC. The schematic structure of the 1/2 size graphite iRPC can be seen in Figure 3.6, meanwhile below the list of the technical details [156] are as following:

- Double coated, hot-melted graphite chamber
- The chamber is equipped with PCB strip plane in both graphite regions
- Two different resistivity region of $50\text{k}\Omega/600\text{k}\Omega$ respectively
- The strips are 5.25mm wide and have 1mm space between them
- The chambers' dimensions are $100\text{cm}\times 70\text{cm}$.
- The gas gap and electrode thickness is 1.4mm
- Gas volume of 2l

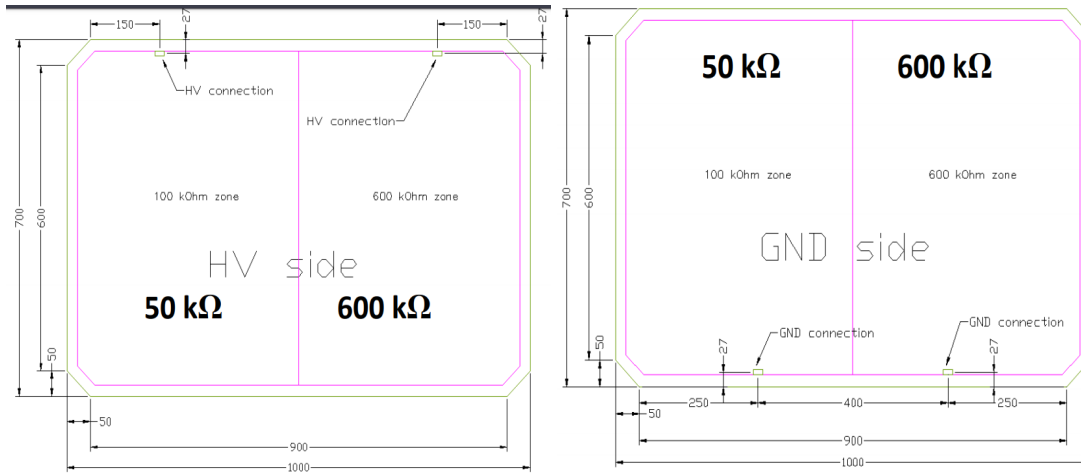


Figure 3.6. . Schematic structure of the Kodel 1/2 size graphite iRPC at the ground site (left) and high voltage HV site (right), where the two different graphite resistivity $50\text{k}\Omega/600\text{k}\Omega$ are shown.

The chamber is equipped with two different read-out electronics (Figure 3.7) to test them on the two graphite resistivity regions, they are:

1. Fast charge preamplifier: The new preamplifier is made in Silicon Bipolar Junction Transistor (BJT) technology [65] and it workson the concept of a fast charge integration with the possibility to match the input impedance to a transmission line
2. Full-custom ASIC discriminator: The new discriminator ASIC designed for iRPCs is implemented in SiGe BiCMOS technology [64] and the working priciple the discriminator is based on the limit amplifier: If the signal passes the threshold, it will be amplified until saturation.

The main features of the discriminator are [61]:

- Relative easy to operate with threshold values up to a few mV ;
- Small transition region around $300 \mu V$;
- Time-Over-Threshold measurement with the discriminator;
- Minimum pulse width of $3ns$.

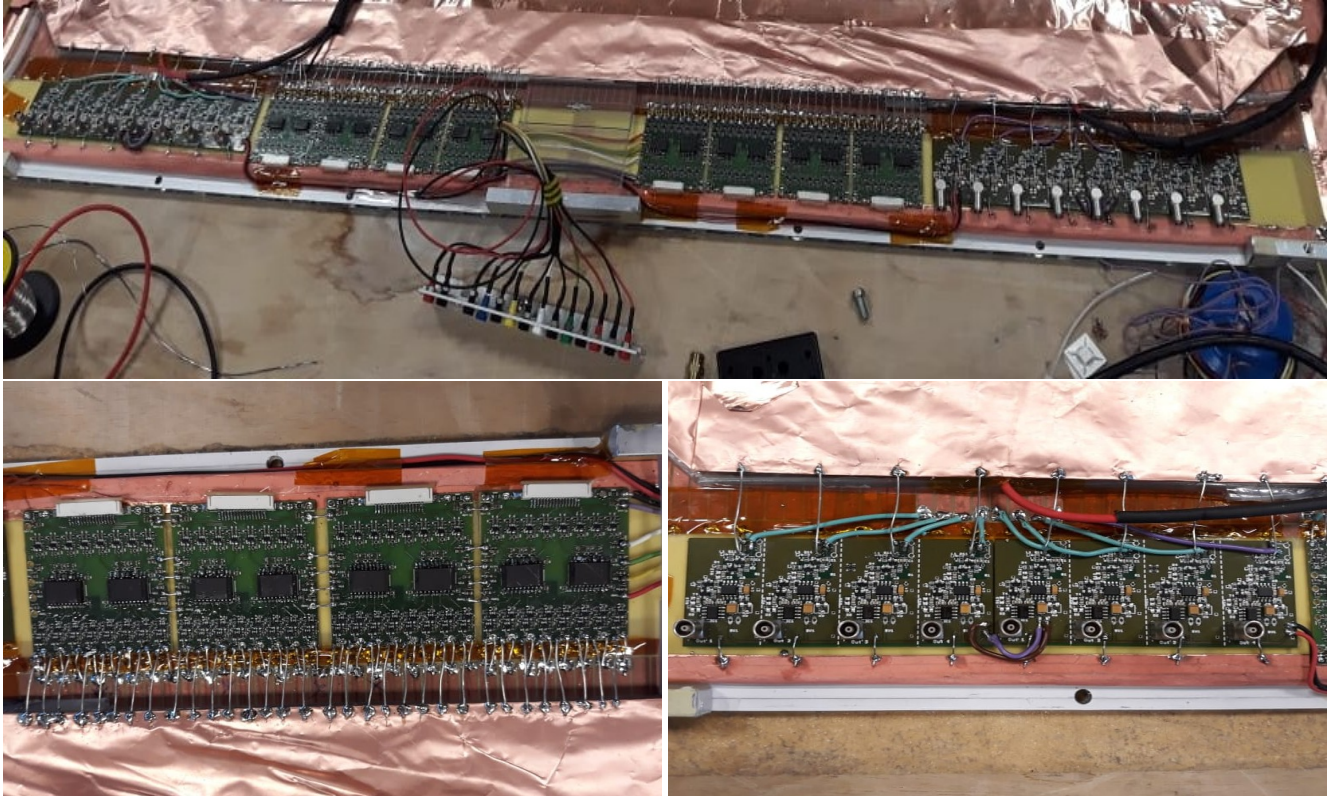


Figure 3.7. . The read-out electronics on the strips PCB of the double coated 1/2 size graphite iRPC chamber, where we have in the center eight custom ASIC discriminators FEBs and and in both sides left and right in total sixteen fast charge analog preamplifier.

The full-custom ASIC discriminator RPC FEBs can read eight strips each, and for each channel there are new discrete-components preamplifiers based on Silicon BJT, as well as eight discriminators in two ASIC chips based on SiGe BiCMOS technology. This technology provides better performance in terms of noise, bandwidth, and power consumption [62]. The dimensions of the FEBs are: $4.8cm \times 5cm$. Furthermore, they have a low power consumption (less $0.5W/channel$) and there is no need of active cooling of FEBs components.

3.3 Performance studies of the two resistivity region of the improved RPC module

The chamber was flushed at $5l/h$ with the gas mixture of 95.2% freon ($C_2H_2F_4$), 4.5% isobutane ($i-C_4H_{10}$) and 0.3% sulphur hexafluoride (SF_6) with a relative humidity of $\sim 40\%$. Scintillators have been used as the trigger system during the tests with cosmic radiation, as we can see in Figure 3.9. The read-out electronics was mounted onto the PCB strip plane.



Figure 3.8. . Full assembled double coated graphite iRPC chamber at the B. 904 at CERN

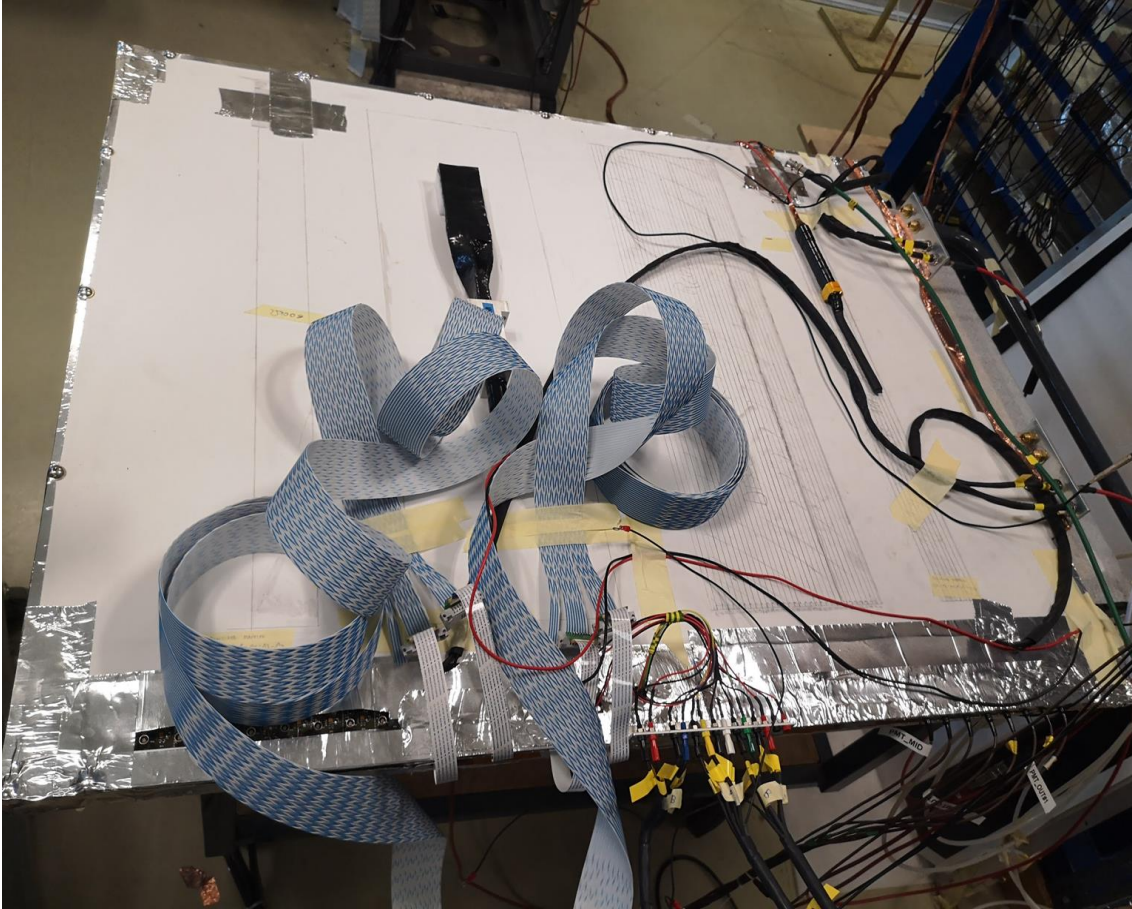


Figure 3.9. . Cosmic radiation tests with the scintillators (covered in black tape) as a trigger system on the improved RPC module at B. 904.

The double coated graphite iRPC chamber was tested in 904 (cosmics) and in the GIF++ facilities (cosmics + background) with two electronics:

1. 2×8 strips fast charge pre-amplifiers + oscilloscope: Study induced signal wave forms, muon cluster size and efficiency plots;
2. The custom ASIC discriminators FEBs (2×32 strips): Study induced muon cluster size and efficiency plots.

The performance tests listed above are performed for the two different graphite resistivity regions: $600k\Omega$ and $50k\Omega$, an important feature to understand the suppression towards the dispersion of the avalanche

charges $I_{mon} = f(HV_{eff})$ and noise scan. From the tests we will learn also more about important properties of both electronics in both resistivity regions in term of performance.

We defined the six adjustable low voltages (LV) for the Front-End custom ASIC discriminators. We have a list of treshold values for the voltages, as shown in the [Table 3.1](#), where we also can notice that the given values V_{Appl} , I_{0set} are not the same as the voltages and currents V_{Meas} , I_{mon} measured from the CAEN TDS and DAQ system.

CHANNELS	Threshold values [V]	V_{Appl} [V]	V_{Meas} [V]	I_{0set} [A]	I_{mon} [A]
V_{TH}	~ 1.5 – 2.0	2.00	1.980	0.10	0.038
V_{DIS}	~ 2.5	2.57	2.506	0.30	0.286
V_{AMP}	~ 1.0 – 2.0	2.05	2.004	0.20	0.197
V_{PU}	~ 0.3	0.30	0.298	0.10	0.000
V_{LVDS}	~ 2.4	2.61	2.503	0.50	0.481
$V_{PREAMP_{left}}$	~ 1.0 – 3.5	3.00	3.023	0.10	0.046
$V_{PREAMP_{right}}$	~ 1.0 – 3.5	3.00	3.004	0.10	0.048
$(-)V_{(TH,DIS,AMP,PU,LVDS)}$	0.0 (common grounded)	—	—	—	—
$(-)V_{(PRE_AMP)}$	0.0 (separate grounded)	—	—	—	—

Table 3.1. The six adjustable low voltages (LV) for the Front-End custom ASIC discriminators and fast timing pre-amplifiers: V_{TH} , V_{DIS} , V_{AMP} , V_{PU} , V_{LVDS} , V_{PREAMP_Left} , V_{PREAMP_Right} . In the 2nd column are the respective threshold values of those voltages given from the manufacturer. In column 3 and 5 we have the applied voltages (V_{Appl}) and current (I_{0set}), meanwhile in 4 and six with have the measured values V_{Meas} , I_{mon} .

The low voltage (LV) system with cables was plug in to the patch-panel of the read-out electronics for both FEBs as shown in [Figure 3.10](#), where the cables are labeled and colored respectively.



Figure 3.10. The patch-panel with the labeled 6 adjustable low voltages (LV) cables for the the powering of the front-end FEBs: V_{TH} , V_{DIS} , V_{AMP} , V_{PU} , V_{LVDS} , V_{PREAMP_Left} , V_{PREAMP_Right}

Two main features are reported:

1. Muon cluster size as a function of the High Voltage (HV): The number of strips fired in the same bunch crossing window with response to the passing particles.

2. Efficiency curves on the working point.

The sigmoid function of the efficiency for the RPC chamber is given as following:

$$\varepsilon = \frac{\varepsilon_{max}}{1 + e^{-\lambda(HV_{eff.} - HV_{50\%})}} \quad (3.1)$$

where $HV_{eff.}$ is the high voltage corrected for temperature and pressure value according to the following environment conditions and it is given as below:

$$HV_{eff.} = \beta HV_{app} = \frac{HV_{app}}{\left[(1 - \alpha) + \alpha \frac{P}{P_0} \frac{T_0}{T} \right]} \quad (3.2)$$

where λ is proportional to the slope at half maximum and $HV_{50\%}$ is the voltage value at which the efficiency reaches half of the maximum. In general, the conventional values used for λ , P_0 , T_0 are: $\lambda = 0.8$, $T_0 = 293K$, $P_0 = 990mbar$. Meanwhile, the pressure P and temperature T are the corresponding environmental values during the test. The working point is defined as $WP = "Knee" + 150V$. The "Knee" is the high voltage point on the sigmoid, where the efficiency is for the RunII RPCs are $> 95\%$ of the efficiency plateau.

3.3.1 Fast timing pre-amplifiers

The fast timing pre-amplifiers are mounted on strips (in [Figure 3.7](#) bottom right), eight on low resistivity graphite region ($50k\Omega$) and eight on high resistivity graphite region ($600k\Omega$), where the amplification voltage applied here is a fixed value of $2.5V$. The analog signals directly read by CAEN digitizer from the direct analysis of pulse shape and the offline threshold is adjustable. It is also possible to evaluate the rate of the noise. The tests are based on a high voltage scan within the interval $[6.0kV, 7.6KeV]$ with a step of $200V$. In the configuration below ([Figure 3.11](#)), we can see six strips fired from a passing muon. The red box shows the muon signal region on the oscilloscope, meanwhile the blue box shows the noise window. The configuration below is a very good result, seance it shows that the strips induce the signal to the FEBs and the signal (in this case the passing muon) is measured from the oscilloscope properly.

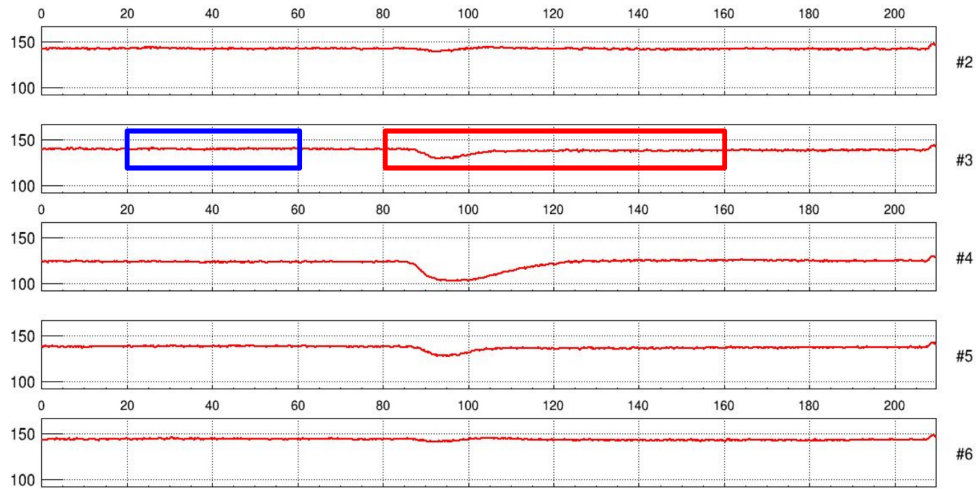


Figure 3.11. Blue box: Noise window; Red Box: Muon window

The histograms in Figure 3.12 are showing the muon efficiency curves as a result of the high voltage scan, where the values have been fitted with the sigmoid function of Eq. 3.1. The data have been analyzed for the fast charge pre-amplifier in both resistivity regions. It can be noticed that at despite the similar working point, in the high resistivity region, we get a slightly higher efficiency with a value of $\varepsilon(WP) = 97.2\%$.

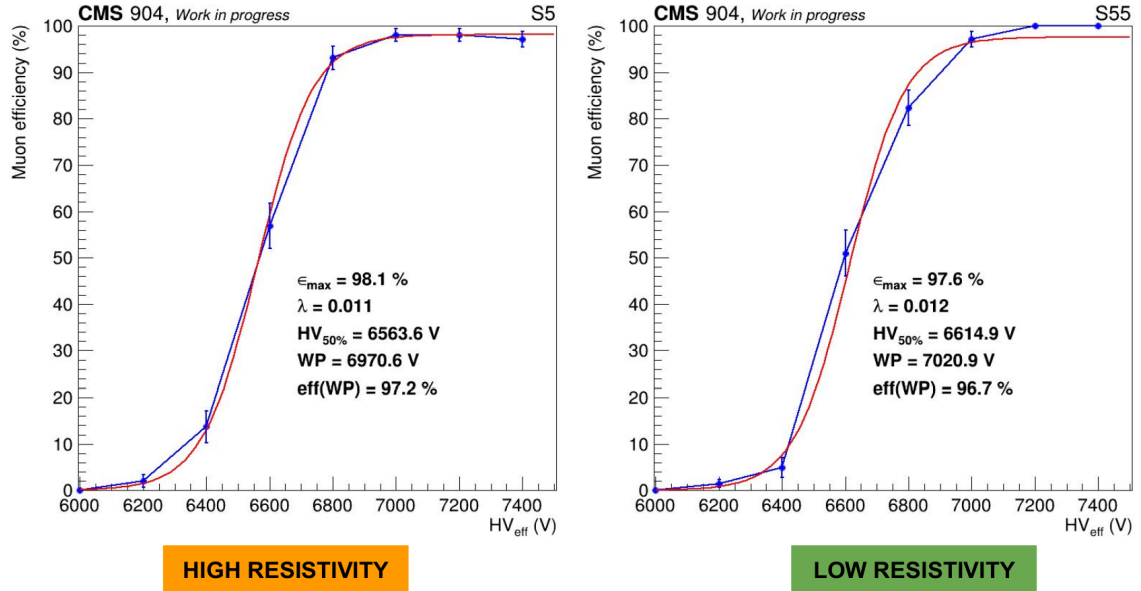


Figure 3.12. Efficiency curves on working point $\sim 7\text{ kV}$ with offline threshold of $10 * \sigma$ of the noise, where similar efficiency for both high and low resistivity regions

The histograms in Figure 3.13 are representing the cluster size at the working point of 7 kV in both resistivity regions, where we can notice that in the high resistivity region the muon cluster size is smaller in comparison to the low resistivity region with a value of 4.5 vs 5.13 .

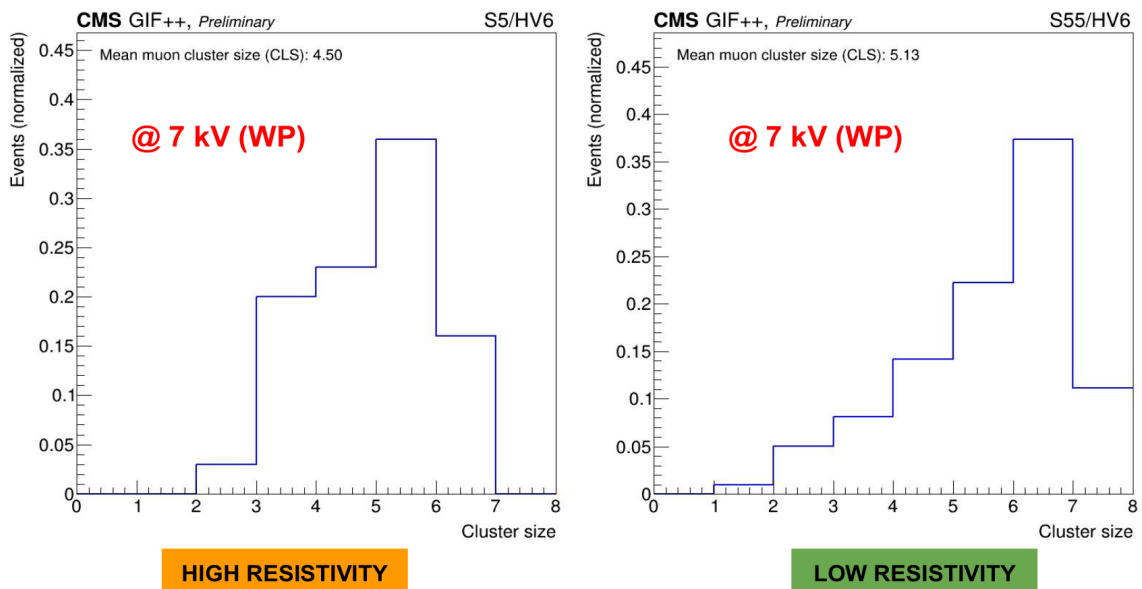


Figure 3.13. Cluster size distributions at WP, where we can notice that the average cluster size is slightly lower for high resistivity region and it has a different shape in cluster size, where it is broader for the low ($50\text{ k}\Omega$) resistivity region

3.3.2 The custom ASIC discriminators FEBs

Four custom ASIC discriminators FEBs are mounted on PCB strips on both the resistivity regions of the graphite chamber, where the read-out is extended in 32 CHANNELS on both resistivity regions ($50k\Omega$ and $600k\Omega$). The read-out of the signal is done thank to a conventional CAEN-TDC modules. The tests are be based on a high voltage scan within the interval $[6.0kV, 7.6KeV]$ with a step of 200 V. The histograms in Figure 3.14 are showing the muon efficiency curves, where the values have been fitted with the sigmoid function of Eq. 3.1 for the custom ASIC discriminators FEBs in both resistivity regions. It can be noticed that at despite the similar working point, in the high resistivity region, we get a slightly higher efficiency at $\varepsilon(WP) = 99.2\%$.

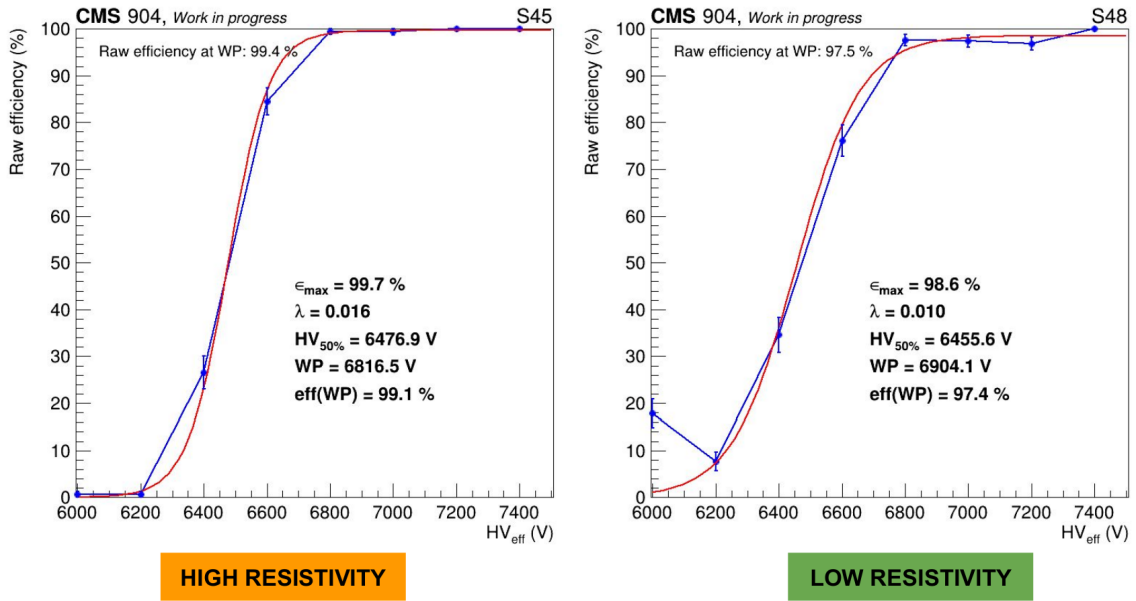


Figure 3.14. Efficiency curves on working point $\sim 7kV$ with offline threshold of $10 * \sigma$ of the noise, where similar efficiency for both high and low resistivity regions with the FEBs

The histograms in Figure 3.15 are representing the cluster size at the working point of $7kV$ in both resistivity regions, where we can notice that in the high resistivity region the muon cluster size is smaller in comparison to the low resistivity region with a value of 4.82 vs 5.75.

In the high resistivity region a scan was performed for three different values of amplification voltage $V_{Amp} = [2.0, 2.25, 2.5]V$, as we see in Figure 3.16. The tests are be based on a high voltage scan within the interval $[6.0kV, 7.6KeV]$ with a step of 200 V. The interest of this scan was of technical and performance interest, where we can notice that the working points and the efficiencies changes accordingly to the amplification voltage of the discriminators. The best performance is obtained at $V_{Amp} = 2.25V$, where we get $HV_{WP} = 7KV$ and $\varepsilon_{eff} = 98.3\%$.

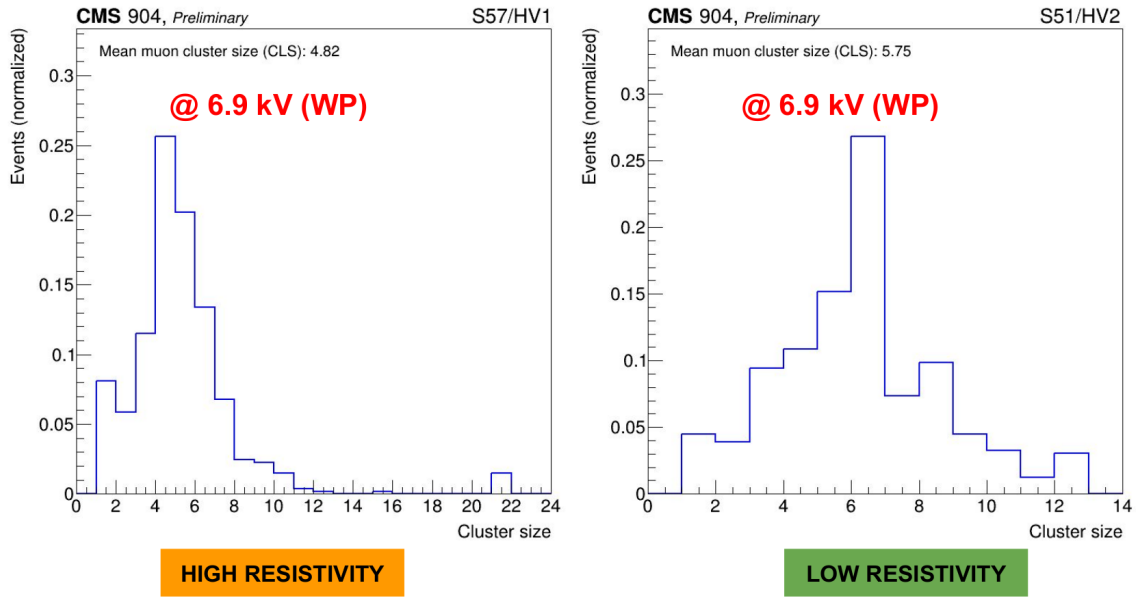


Figure 3.15. Average muon cluster size vs. HV, where the cluster size is lightly lower for high resistivity region with the FEBs

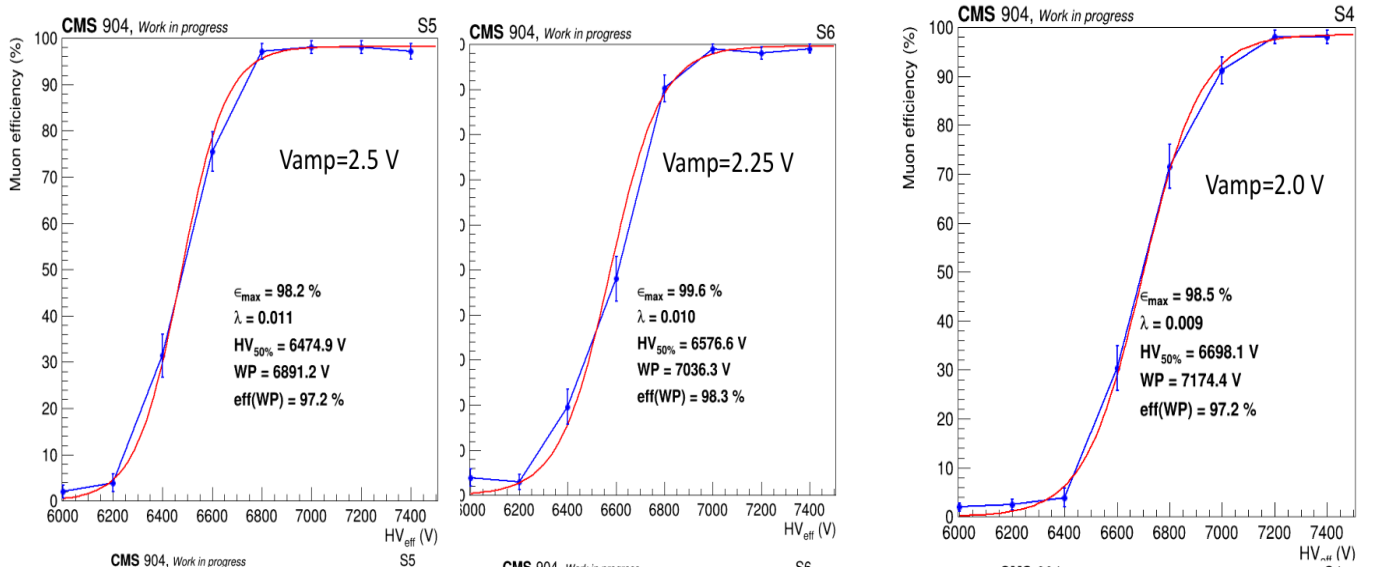


Figure 3.16. Efficiency curves on three different value of the amplification power-up value of the voltages for the custom ASIC discriminators FEBs in the high resistivity region on the graphite (600 kΩ): 1) Left: $V_{amp} = 2.5$ V (left); 2) Center: $V_{amp} = 2.25$ V; 3) Right: $V_{amp} = 2.0$ V

3.4 Conclusions

The goal of this study is the effects on the performance of the improved RPC modules for different resistivity values of the graphite, coated on the new improved RPC chambers for the RPC system. The result of this study is to certify the new resistivity value of the graphite for the RPC chambers system for the HL-LHC. In Table 3.2 we listed the results of the test explained in the section above comparing both resistivity regions, 50 kΩ and 600 kΩ for both read-out electronics. The comparison shows that at the higher resistivity value of the graphite at 600 kΩ, we get a lower working point of the double gap improved RPC module for both read-out electronics, the fast charge amplifier and the ASIC custom digitizer, followed by higher efficiency and lower muon cluster size.

	PREAMP		DISCR.	
Parameters	50kΩ	600kΩ	50kΩ	600kΩ
Working Point [V]	7020.9	6970.6	6904.1	6816.5
Muon Cluster Size	5.13 (at 7kV)	4.5 (at 7kV)	5.75 (at 6.9 kV)	4.82 (at 6.9 kV)
Efficiency [%]	96.7%	97.2 %	97.4 %	99.1 %

Table 3.2. . Comparison of between the two resistivity regions for both read-out electronics. In all results of the tests, for the higher resistivity of the graphite (600kΩ) we get better results in comparison to the lower resistivity (50kΩ).

As a conclusion to the reported results, the value of the hot-coated and melted graphite over the high pressure laminate bakalite gaps of the improved RPC modules, was decided to use the value of 600kΩ instead of 50kΩ for the HL-LHC.

Part IV

$HH \rightarrow b\bar{b}ZZ(4l)$ Analysis

Chapter 4

Monte Carlo simulation

In this chapter the generation of the MC signal samples of both resonant and non-resonant HH signal samples is presented. The MC generation is a complex but very important step for the data analysis work of every Physics we would like to perform. In order to perform a MC generation of the signal samples, we need to have a theoretical model to be then computed through a MC generator framework such as MadGraph5_aMC@NLO[6][5] or POWHEG-BOX-V2[106][104]. Different models at different accuracy levels are presented in this chapter in case of the non-resonant di-Higgs production and are compared to each other. Furthermore, theoretical uncertainties of the different studies PDF sets are calculated and compared. A full MC generation for both resonant and non-resonant HH has been performed as well as a full CMS detector response has been obtained from the MC signal samples thanks to the CMS full simulation framework for all the Run II datasets.

4.1 MC generations studies related to the Monte Carlo generators di-Higgs (H) signal models

The phenomenology of multi-Higgs boson final states will provide complementary information on the observation of the single Higgs boson at the LHC. Due to generically small inclusive cross sections and a difficult signal vs. background discrimination, the best motivated multi-Higgs final states at the Large Hadron Collider are Higgs boson pair final states, of which gluon-gluon fusion $ggF\ gg \rightarrow HH$ is the dominant production mode, as explained in section [1.4].

Many Monte Carlo models of physics beyond the Standard Model with SM-compatible single Higgs boson signal strengths can exhibit a di-Higgs phenomenology vastly different from the SM expectation [82][81]. In this sense, a successful discovery of Higgs boson pair production at the LHC and the subsequent measurement of potential deviations from the SM constitute an important path for in the search for physics beyond the SM. In particular, modifications of the Higgs trilinear couplings (e.g. via a modified Higgs self interaction) can only be directly observed in Higgs boson pair production. In the gluon fusion process this occurs via the interference of the box and triangle diagrams, as already shown in Figure 1.8 (chapter [1]). To facilitate such a measurement, it is crucial to establish the Higgs boson pair production cross section in the SM to the best theoretical accuracy possible and to provide BSM benchmarks that reflect the phenomenology of Higgs boson pairs at the LHC in a consistent and concise fashion.

Two are the main MC generators used for our purpose, also implemented within the CMS Simulation framework:

1. MadGraph5_aMC@NLO[6][5]: It is a Monte Carlo generator framework used for the automatic generation of matrix elements for High Energy Physics processes. It generates all Feynman diagrams for the chosen process, and outputs the computer code necessary to evaluate the matrix element at a given phase space point. The matrix element evaluation is done using calls to helicity wavefunctions and amplitudes (with the HELAS package [99]). This implementation is efficient because it naturally allows helicity wavefunctions corresponding to identical subdiagrams to be reused across diagrams. The computer

code produced by MadGraph can then be used for cross section or decay width calculations and event generation (the MadEvent package [96]);

2. POWHEG-BOX-V2[106][104]: It is a Monte Carlo generator framework that automatizes the construction of NLO+PS[102] (Parton Shower) generators, once the matrix elements are available. It uses the POWHEG method that allows to generate events with positive weights and, because of this, it has become the best choice when large samples of events are needed. Furthermore, the POWHEG method implements the inclusion of processes with decaying coloured resonances, so enabling the handling of the interference of the emitted radiation generated in production and decay.

Both of the two above described MC generators can be interfaced with a very important framework, PYTHIA (latest version: PYTHIA 8.306), used for the parton shower and hadronization of a generated MC event, and interfaced to the CMS Simulation framework. The official description is as following:

- "PYTHIA is a program for the generation of high-energy physics collision events, i.e. for the description of collisions at high energies between electrons, protons, photons and heavy nuclei. It contains theory and models for a number of physics aspects, including hard and soft interactions, parton distributions, initial- and final-state parton showers, multiparton interactions, fragmentation and decay. It is largely based on original research, but also borrows many formulae and other knowledge from the literature. As such it is categorized as a general purpose Monte Carlo event generator" [116].

4.1.1 Non-resonant HH MC generator models

There are several models used within the CMS collaboration for the di-Higgs production: leading order (LO) and next-to-leading order (NLO). The Figure 4.1 below gives a clear picture between the LO-MC and NLO-MC[102] models on the HH MC production:

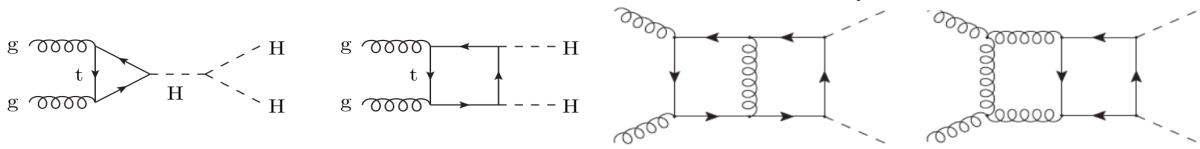


Figure 4.1. a) Left: Feynman diagram at LO of the HH -box and triangle production used in MadGraph5_aMC@NLO (`loop_sm_hh` with MG5_aMC_v2_5_5 and `BSM_gg_hh` ($k_\lambda = k_t = 1$ with MG5_aMC_v2_6_1)[82]; b) Right: Feynman diagram at NLO of the HH production used in MadGraph5_aMC@NLO and POWHEG-BOX-V2 (`SMEFT_UFO-no_b_mass` with MG5_aMC_v2_6_0 and `ggHH_EWChL` with POWHEG-BOX-V2) [82][81]

We focus on the two models for SM di-Higgs production with full top quark mass dependency at NLO:

1. SMEFT_UFO-no_b_mass [82]: at NLO accuracy with the MadGraph5_aMC@NLO MC generator;
2. ggHH_EWChL [82][81]: at NLO accuracy with the POWHEG-BOX-V2 MC generator;

The main difference between LO and NLO can be noticed in Figure 4.1 with the Feynman diagrams, where at next leading order (NLO) and beyond, we have infrared divergences (collinear and soft) in the Feynman diagrams and consequently in the cross section calculations. To obtain a finite results one needs to sum together, in a given order of perturbation, diagrams involving loops with real emissions[104][102]. Furthermore, the calculation of the NLO QCD corrections is reduced to a one-loop calculation for the virtual corrections and a tree-level calculation for the matrix elements of the real corrections. The NLO final result for the total gluon-fusion cross section can be decomposed as:

$$\sigma_{NLO}(pp \rightarrow HH) = \sigma_{NLO} + \Delta\sigma_{virt} + \Delta\sigma_{gg} + \Delta\sigma_{gq} + \Delta\sigma_{q\bar{q}} \quad (4.1)$$

where these extra corrections contribute to the total cross section by increasing the total computed NLO $\sigma_{NLO}(pp \rightarrow HH)$ cross section. More details are discussed in Appendix A.

4.1.2 Comparisons

1) LO vs NLO for the SM-HH with full top quark mass m_t dependence model

The study related to the MC generation of the non resonant double Higgs production signal samples with full top quark mass m_t dependence required first a comparison with the existing MC LO and NLO models:

- `loop_sm_hh` model within `MadGraph5_aMC@NLO`: It is based on the SM Effective Field Theory (SMEFT theory, see [Appendix A](#)) with the loop implementation for the box and triangle Feynman diagram. Instead of the top quark loop, this model uses a b-quark virtual loop diagram at LO accuracy;
- `BSM_gg_hh` model within `MadGraph5_aMC@NLO`: It is based on the SM Effective Field Theory (SMEFT theory) with the loop implementation for the box and triangle Feynman diagram with full top quark mass dependence at LO accuracy.
- `ggHH_EWChL` model within `POWHEG-BOX-V2`: It is based on the SM Effective Field Theory (SMEFT theory) with the loop implementation for the box and triangle Feynman diagram with full top quark mass dependence at NLO accuracy.

The results of the computed cross sections of each above described model and its respective accuracy, are as following:

$$\begin{aligned}\sigma_{NLO}^{PWG}(pp \rightarrow HH)[fb] &= 28.06 \pm 0.133 & NLO - PWG \\ \sigma_{LO}^{MG5}(pp \rightarrow HH)[fb] &= 19.18 \pm 0.138 & LO - MG5 - BSM_gg_HH \\ \sigma_{LO}^{MG5}(pp \rightarrow HH)[fb] &= 11.03 \pm 0.096 & LO - MG5 - loop_sm_hh\end{aligned}\quad (4.2)$$

where we can notice that the NLO SM-HH model due to the extra contributions at NLO accuracy from Eq. 4.1 has an increase of almost 50% on the total computed cross sections in comparison with the LO MC generator model. In [Figure 4.2](#) we can see a comparison at GEN-SIM level (MC generation plus Parton Shower level only) for the kinematic distributions such as the invariant mass M_{HH} of the Higgs pair, their transversal momentum p_T^{HH} as well as the transversal momentum of both Higgs bosons $p_T^{H_1}, p_T^{H_2}$ and the angular separation ΔR_{HH} of the di-Higgs. In all four diagrams we can notice that in case of the NLO for the `POWHEG-BOX-V2` model `ggHH_EWChL` more statistics is concentrated around the peaks, meanwhile in the tails in case of the LO models still seem to have much more events than the NLO model. It is also important to notice the smaller statistical errors of the NLO model in comparison to both LO models.

2) `POWHEG-BOX-V2` vs. `MadGraph5_aMC@NLO` for NLO SM-HH with full top quark mass m_t dependence model

Meanwhile, another very important comparison is between the same theoretical model at NLO, but generated within two different MC generator frameworks. The next comparison will involve the NLO MC models:

- `SMEFT_UF0-no_b_mass` model within `MadGraph5_aMC@NLO`: It is based on the SM Effective Field Theory (SMEFT theory) with the loop implementation for the box and triangle Feynman diagram with full top quark mass dependence at NLO accuracy.
- `POWHEG-BOX-V2` model `ggHH_EWChL`: It is based on the SM Effective Field Theory (SMEFT theory) with the loop implementation for the box and triangle Feynman diagram with full top quark mass dependence at NLO accuracy.

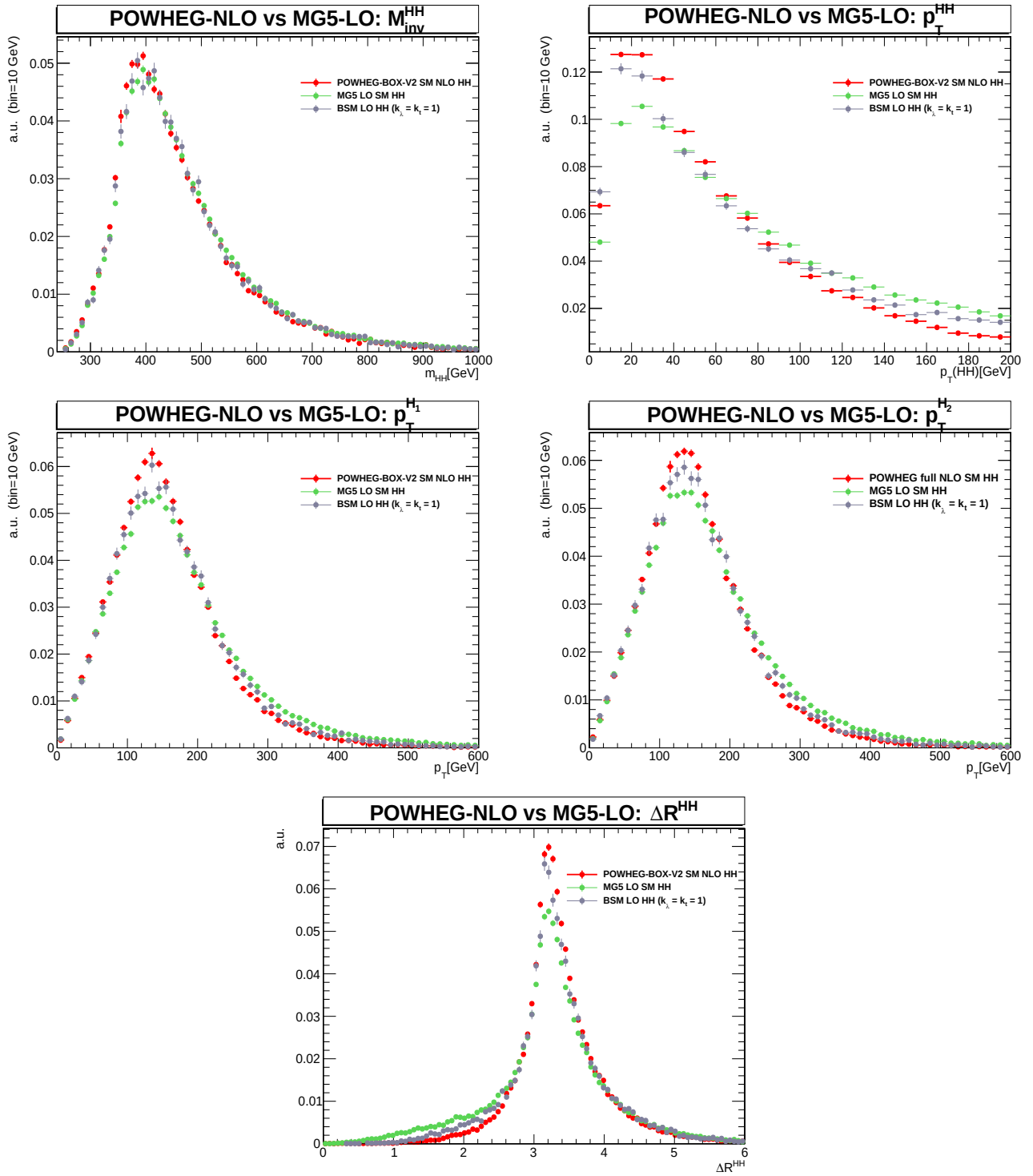


Figure 4.2. Comparison between the at NLO-MC model ggHH_EWchL within the POWHEG-BOX-V2 (in red) and the LO-MC model `loop_sm_hh` (green) and `BSM_gg_hh` (in grey) within the MadGraph5_aMC@NLO. a) Top left: Distribution of the M_{inv}^{HH} peaking around 400GeV; b) Top right: Distribution of the p_T^{HH} peaking around 20GeV; c) Center left: Distribution of the $p_T^{H_1}$ of the leading H -boson peaking around 150GeV; d) Center right: Distribution of the $p_T^{H_2}$ of the second H -boson peaking around 150GeV; e) Bottom: Distribution of the angular separation ΔR_{HH}

The comparison is done with sample of 100.000 generated events of both sides. This comparison is of major importance, since they are the only model at highest computational accuracy (NLO) available. In Table 4.2 we can see that in terms of total cross sections at 13 and 14 TeV we have very close values, meanwhile the POWHEG-BOX-V2 delivers a smaller percentage on negative weighted events as MadGraph5_aMC@NLO does (6.68% PWG vs 9.08% MG5), even though the POWHEG-BOX-V2 is a much more power consumption MC

generator, since it takes in average 7.1s to generate one single event.

Models	Time/Event [s]	Negative weights(%)	σ_{NLO}^{HH} (14TeV)[fb]	σ_{NLO}^{HH} (13TeV)[fb]
ggHH_EWChL (POWHEG-BOX-V2)	7.1s	6.68%	33.00[fb]	28.08 [fb]
SMEFT_UFO-no_b_mass (MadGraph5_aMC@NLO)	0.3s	9.08%	32.97[fb]	27.86 [fb]

Table 4.1. Comparison of the same NLO SM-HH with full top quark mass m_t dependence generated with two different MC generator frameworks: POWHEG-BOX-V2 and MadGraph5_aMC@NLO

In Figure 4.3 we can see a comparison at GEN-SIM level (MC generation plus Parton Shower level only) of the same NLO SM-HH with full top quark mass m_t dependence between the two different MC generator frameworks: POWHEG-BOX-V2 and MadGraph5_aMC@NLO. There are plotted the kinematic distributions such as the invariant mass M_{HH} of the Higgs pair, their transversal momentum p_T^{HH} as well as the transversal momentum of both Higgs bosons $p_T^{H_1}, p_T^{H_2}$ and the angular separation ΔR_{HH} of the di-Higgs. In all four diagrams we can notice that in case of the NLO SM-HH ggHH_EWChL model implemented within the POWHEG-BOX-V2, more statistics is concentrated around the peaks, meanwhile in the tails in case of the NLO SM-HH SMEFT_UFO-no_b_mass within the MadGraph5_aMC@NLO still seem to have much more events than its competitor model. Furthermore, the SMEFT_UFO-no_b_mass model shows much larger statistical errors than the ggHH_EWChL compiled by the POWHEG-BOX-V2. In addition to the technical details of both of the models of the same NLO SM-HH with full top quark mass m_t uncertainty, with the ggHH_EWChL model we can also perform a k_λ -scan, which we explain in comre details in the following subsections .

As a conclusion, since the NLO SM-HH ggHH_EWChL model implemented within the POWHEG-BOX-V2 has a better performance as explained above and also can computationally perform a k_λ -scan, an important task needed to be done, in order to understand if the experimental data will deliver to us in Run II or in further runs in the future at the LHC, the ultimate prove of the tri-linear interaction of the Higgs bosons, or we will be facing a deviation of the SM model expectations, the decision was set to use it for the final official MC generation within the CMS framework, a result of major importance within the CMS collaboration.

4.1.3 Theoretical uncertainties

The Parton density functions (PDFs) are used to calculate cross-section at hadron colliders. They encode the process-independent momentum structure of partons within hadrons, with which partonic cross-sections must be involved to obtain results that can be compared to experimental data. By definition the PDFs are incorporated within the hadron collider cross-section calculations in the form [21], also called as the QCD Factorization Theorem:

$$\sigma_{tot} = \sum_{i,j} \int_0^1 dx_1 dx_2 f_i(x_1, \mu_f) f_j(x_2, \mu_f) \hat{\sigma}_{ij \rightarrow X}(\hat{s}, \mu_f, \mu_r) \quad (4.3)$$

where $f_{i/j}(x_{1/2}, Q^2)$ is the parton density function for parton i/j (first / second beam of protons in the case of the LHC) at a factorization scale μ_f and renormalization scale μ_r , while $\hat{\sigma}_{ij \rightarrow X}$ is the partonic cross section.

The estimation of the theoretical uncertainties is part of the systematic uncertainties studies for the analysis at CMS. The study related to the signal model is relatively new and has to be fully implemented within the analysis. The estimation of theoretical uncertainties on the PDF sets (Parton Distribution Functions),

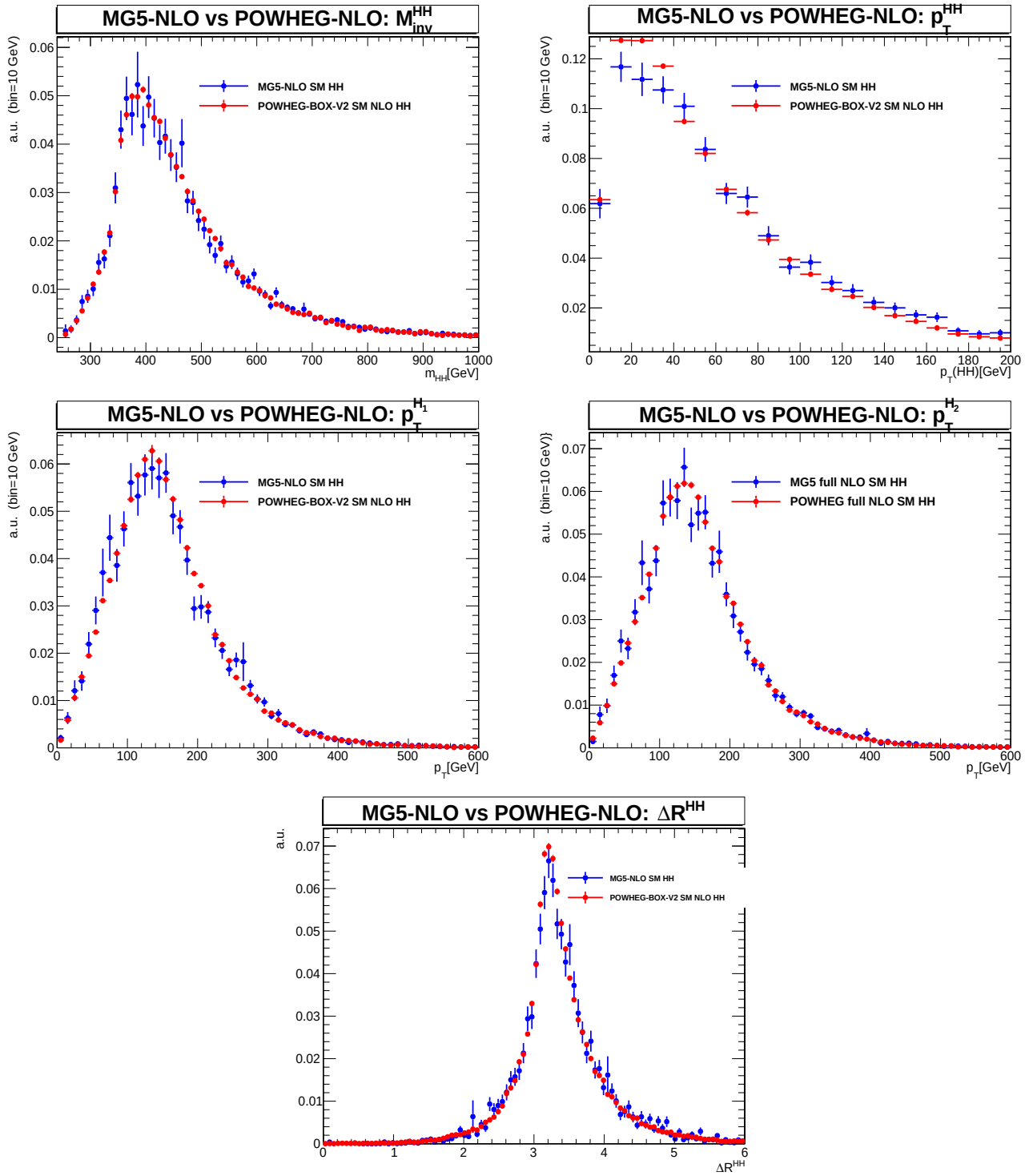


Figure 4.3. Comparison between the at NLO-MC model ggHH_EWChL within the POWHEG-BOX-V2 (in red) and the NLO-MC model SMEFT_UF0-no_b_mass in blue within the MadGraph5_aMC@NLO; a) Top left: Distribution of the M_{inv}^{HH} peaking around 400GeV; b) Top right: Distribution of the p_T^{HH} peaking around 20GeV; c) Center left: Distribution of the $p_T^{H_1}$ of the leading H-boson peaking around 150GeV; d) Center right: Distribution of the $p_T^{H_2}$ of the second H-boson peaking around 150GeV; e) Bottom: Distribution of the angular separation ΔR_{HH}

used for the MC generation, was performed. The theoretical uncertainties have been estimated with the Hessian method, which implies the Hessian matrix with a certain number of the eigenvectors, also called for convenience "PDF replica". The symmetric Hessian formula in Eq. 4.4, also denoted as the "one-sigma PDF uncertainty" [26], is shown below:

$$(HH)\sigma_{H,\alpha}^{PDF}(x, Q^2) = \sqrt{\sum_{n=1}^{N_{eig}} \left(\tilde{f}_{\alpha}^{(i)}(x, Q^2) - f_{\alpha}^{(0)}(x, Q^2) \right)^2} \quad (4.4)$$

Two different PDF sets are used for the estimation of the theoretical uncertainties and they have been compared (Figure 4.4):

1. PDF4LHC15_nlo_30[11] (LHAPDF_ID: 306000): It is used in Ref. [81] and it contains in total one central PDF, 30 PDF replica, and two $\alpha_s(Q^2) = 0.118 \rightarrow (0.16, 0.20)$ replica
2. NNPDF31_nnlo_hessian_pdfas[26][21][163](LHAPDF_ID: 306000): It is generally used by the MC generator group at CMS and it contains in total one central PDF, 100 PDF replica, two $\alpha_s(Q^2) = 0.118 \rightarrow (0.16, 0.20)$ replica

According to Eq. 4.4, the procedure of the theoretical uncertainties would be as following:

- 1 generation of the central PDF set (with $\alpha_s(Q^2) = 0.118$);
- Generations of 30 (in case of PDF4LHC15_nlo_30) and 100 PDF (in case of NNPDF31_nnlo_hessian_pdfas) replicas related to the PDF uncertainties only;
- 2 generations related to $\alpha_s(Q^2) = 0.116, 0.120$ values $\Leftrightarrow \alpha_s(Q^2)$ and $PDF + \alpha_s(Q^2)$ uncertainties from the PDF sets;
- 6 generations related to the scale uncertainties:

$$\{\mu_f : \mu_r\} : \{(0.5; 0.5), (0.5; 1), (2; 1), (1; 2), (1; 0.5), (2; 2)\}, \text{ besides } \{(1; 1)\} \text{ (SM case)}$$

Meanwhile, it is important to remind that the top mass scheme uncertainty does not depend from the PDFs.

In the table below, we list the final result of the procedure above, related to the total theoretical uncertainties estimations for scale $\{\mu_f : \mu_r\}$, PDF and $\alpha_s(Q^2)$ at NLO for 13TeV.

PDF sets	Time/Event [s]	Negative weights(%)	σ_{NLO}^{HH} (13TeV)[fb]	Scale up	Scale down	PDF	$\alpha_s(Q^2)$	m_t uncert.
PDF4LHC15_nlo_30	7.1s	6.68%	28.08 ± 0.186 [fb]	+ 14.19 %	- 10.56 %	$\pm 3.27\%$	-	$\pm 2.7\%$
NNPDF31_nnlo_hessian_pdfas	6.3s	5.17%	26.64 ± 0.146 [fb]	+ 13.74 %	- 10.78 %	$\pm 1.47\%$	$\pm 2.51\%$	$\pm 2.7\%$

Table 4.2. Comparison of the same NLO SM-HH with full top quark mass m_t dependence generated with two different MC generator frameworks: POWHEG-BOX-V2 and MadGraph5_aMC@NLO

In the Figure 4.4, are represented the kinetic distributions of the invariant mass M_{HH} of the Higgs pair, their transversal momentum p_T^{HH} as well as the transversal momentum of both Higgs bosons $p_T^{H_1}, p_T^{H_2}$ after the generation using the two different PDF sets PDF4LHC15_nlo_30 and NNPDF31_nnlo_hessian_pdfas, where we can notice that the kinematic distributions are actually the same within the range of the error bars.

From the Table 4.2 above, we can notice that both PDF sets do deliver different results in term of total cross section but also in term of theoretical uncertainties, which we can sum them up in the following way:

$$\begin{aligned} \text{PDF4LHC15: } \sigma(13\text{TeV})[fb] &= 28.078_{-10.56\%}^{+14.19\%}(\text{scale}) \pm 3.276\%(\text{PDF}) \pm 2.7\%(\text{top}) \\ \text{NNPDF31: } \sigma(13\text{TeV})[fb] &= 26.642_{-10.78\%}^{+13.74\%}(\text{scale}) \pm 1.47\%(\text{PDF}) \pm 2.509\%(\alpha_s(Q^2)) \pm 2.7\%(\text{top}) \end{aligned} \quad (4.5)$$

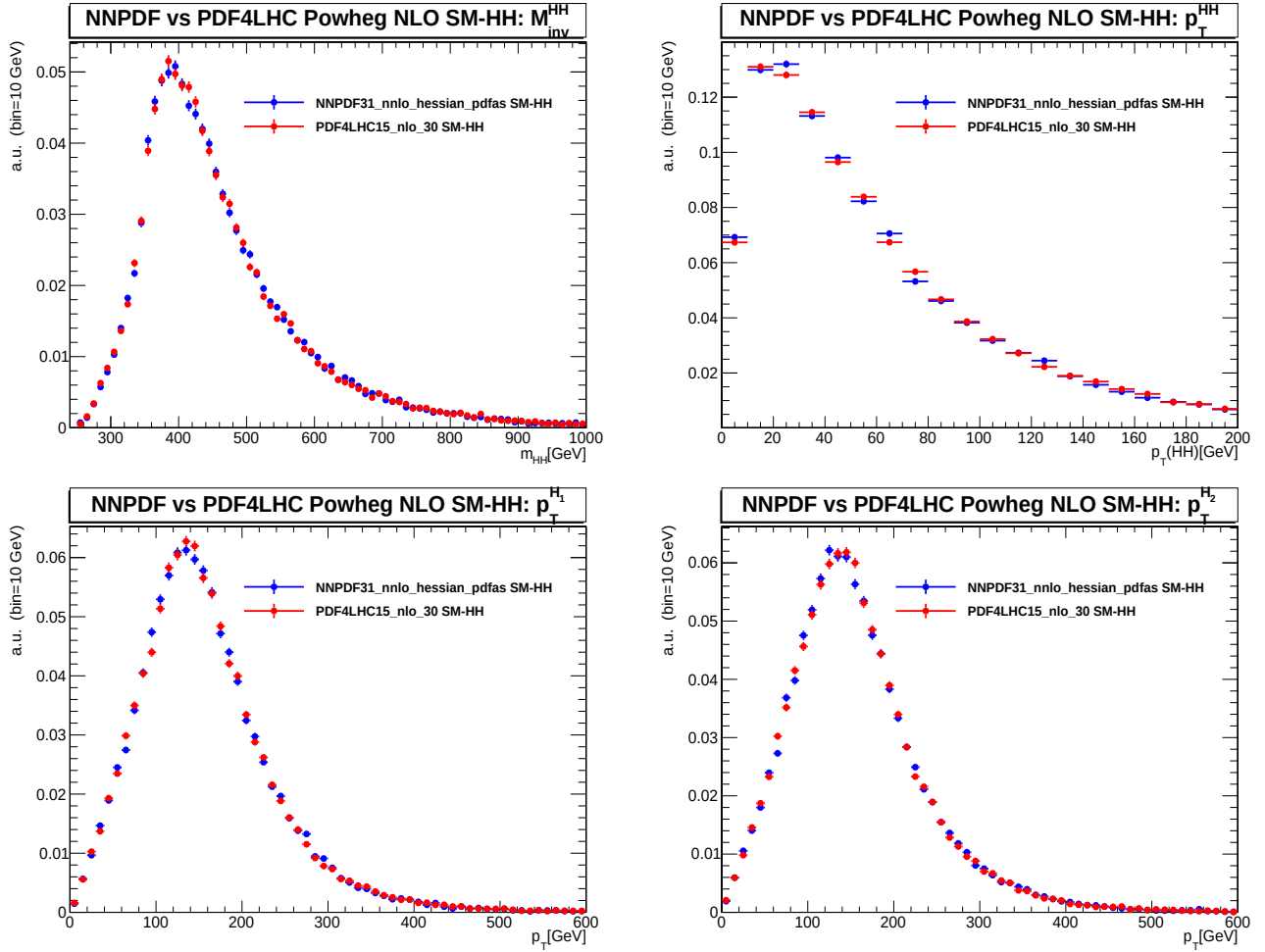


Figure 4.4. Comparison between two different PDF-sets usage; a) Top left: Distribution of the M_{inv}^{HH} peaking around 400GeV; b) Top right: Distribution of the p_T^{HH} peaking around 20GeV; c) Bottom left: Distribution of the p_T^{H1} of the leading H -boson peaking around 150GeV; d) Bottom right: Distribution of the p_T^{H2} of the second H -boson peaking around 150GeV

The PDF4LHC15_nlo_30 sets are used from the theoretical group developer of the NLO SM-HH ggHH_EWChL model and they deliver a slightly higher total cross section as after the usage of NNPDF31_nnlo_hessian_pdfas suggest from the CMS MC generator group for official MC geneations for the same NLO SM-HH ggHH_EWChL model, but it also delivers sglightly higher scale-up and PDF uncertainties. Furthermore, they can't deliver an estimation of the $\alpha_s(Q^2)$ -uncertainties, which the NNPDF31_nnlo_hessian_pdfas sets do deliver this extra information.

4.1.4 The k_λ -factor scan for the λ_{hhh} at 13TeV

As mentioned above, the study related to the trilinear and quartic Higgs self-coupling constant $\lambda_{3H,4H}$ is very important for the Higgs Physics for two main reasons:

1. In case the SM prediction of the value of $\lambda_{3H,4H} = \frac{m_H^2}{2v^2} \approx 0.13$ is true: The verification of this prediction woud imply that the Higgs boson discovered in the period 2011-2012 is the one predicted from the SM;
2. In case of deviation from $\lambda_{3H,4H} = \frac{m_H^2}{2v^2} \approx 0.13$ woud imply existence of new Physics related to BSM.

Since the selected MC model ggHH_EWChL from the POWHEG-BOX-V2 was selected and also gives the possibility to simulate a k_λ -scan of the $\lambda_{3H,4H}$ value, a scan was generated and simulated within the CMS framework of for 2016,2017,2018 modes. The k_λ -scan and its quadratic behaviour is shown in Figure 4.5, where we plotted the cross sections of each k_λ -scan value from $[-10; 10; 1]$ with the theoretical uncertainties of each point generated with the POWHEG-BOX-V2 for scale ($\mu_f : \mu_r$), PDF and $\alpha_s(Q^2)$ at NLO for 13TeV as shown in the section above.

As we can see from the Figure 4.5, the HH-cross section has a minimum at $k_\lambda = 2.45$, where this point is called as the point with the maximal (destructive) interference between the triangle and box Feynman diagrams, as mentioned above and we can observe this in the Figure 4.6.

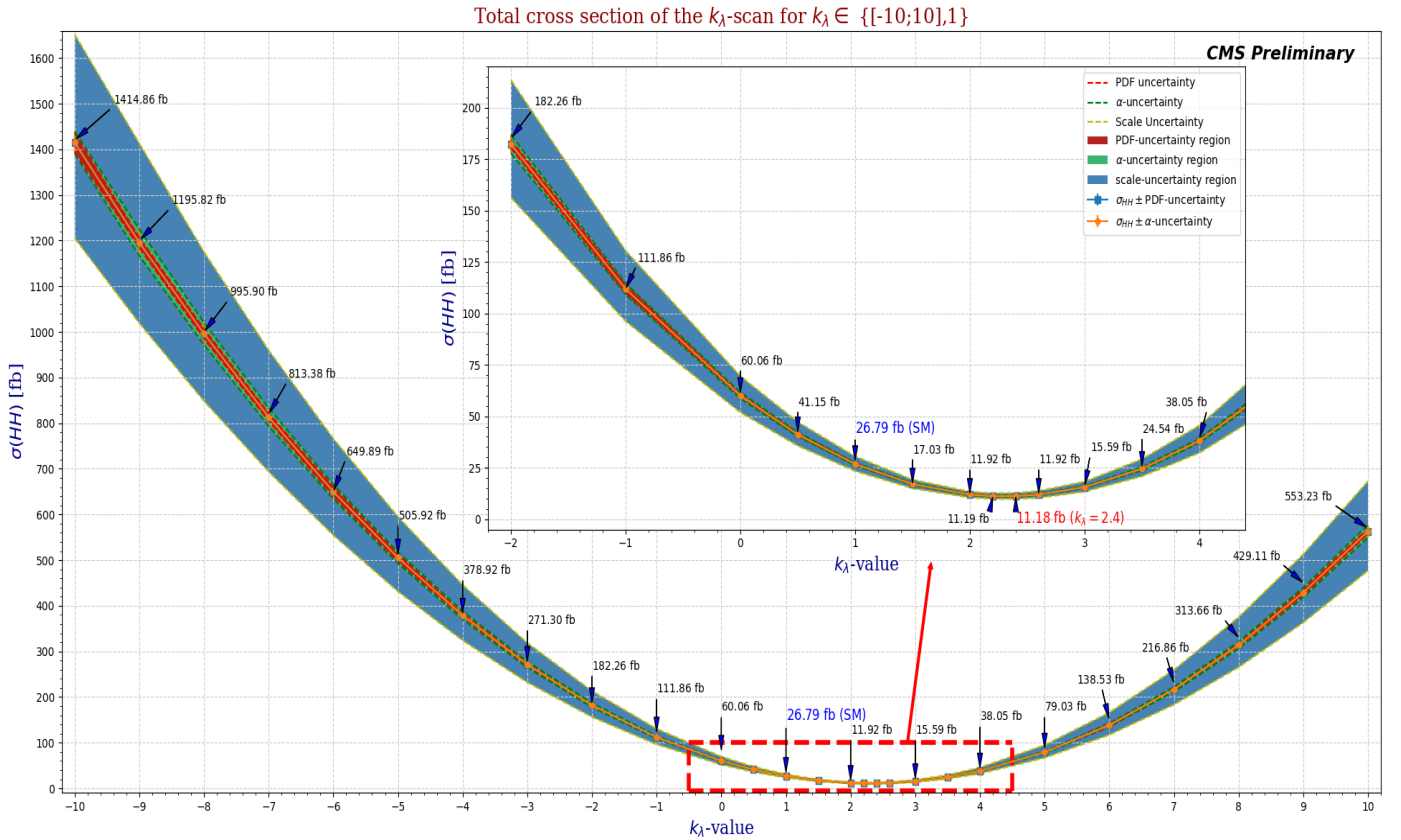


Figure 4.5. The k_λ -scan $[-10; 10; 1]$ of the trilinear and quartic Higgs self coupling constant $\lambda_{3H,4H}$ within the PDF, $\alpha(Q_s^2)$, scale and top quark mass uncertainties

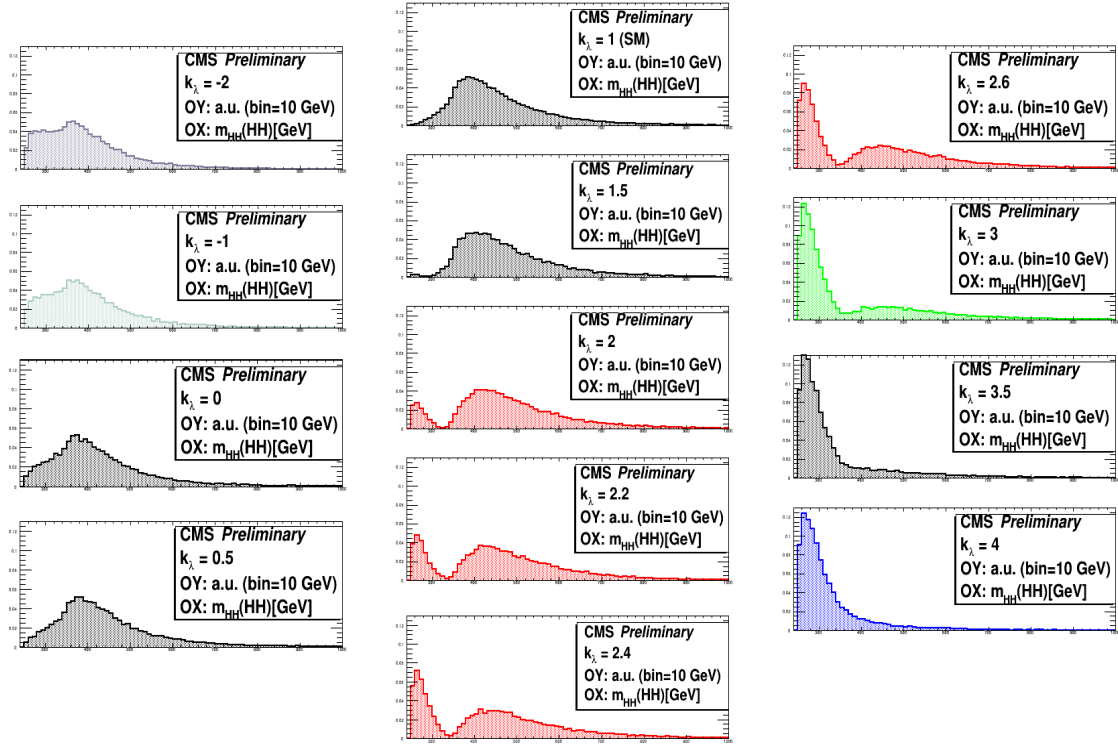


Figure 4.6. Mass distributions of the k_λ -scan within the restricted range $[-2; 4]$ of the trilinear and quartic Higgs self coupling constant $\lambda_{3H,4H}$, where we can notice the maximum interference evolution between the triangle and box Feynman diagrams from $k_\lambda = 2.2$ and $k_\lambda = 2.6$

4.1.5 The resonant HH production

The Kaluza-Klein KK-theory and the Randall-Sundrum model of "5-dimensional warped geometry theory" are our main field of interest for our phenomenology studies at the LHC and part of our group analysis here at Bari and at the LHC in general [27]. The Bulk KK-Graviton spin-2 and the Radion spin-0 are very important signatures of those models.

The resonant HH production is driven from a narrow resonance X with a mass m_X using the invariant mass spectrum m_{HH} , where the cross section is significantly enhanced on resonances up to pb . The Randall-Sundrum warped extra dimensions from the Kaluza-Klein theory [109][110], predict massive resonances such as the Radion spin-0 or KK Bulk Graviton spin-2, with mass ranges up to several TeV.

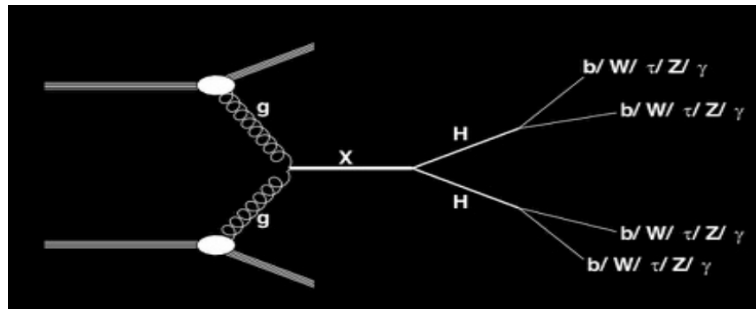


Figure 4.7. The Feynman diagrams of the resonant HH production via gluon fusion at LO, where a massive X resonance decays into a pair of Higgs bosons.

The calculation of the QCD induced G^* production cross section is done by a computational model at

Next to Leading Order (NLO) QCD corrections through the glon fusion (ggf) production of a spin-2 and spin-0 particle [10, 50]. The LO calculations are done with the framework Feynrules [39] as part of the MC generator MG5_aMC@NLO, as well with the WED implementation default on Pythia8 [116] for the particle decay. Part of the WED scenario was also the framework CalChep [15] framework.

The samples of resonant HH signals events are produced by Monte Carlo simulation. Those MC generations are used for the purpose of estimating signal and background yields in the analysis and to train machine learning algorithms. The samples are produced at leading-order (LO) matrix elements and generated with MadGraph5_aMC@NLO. The models used for the MC signal generations are from the Kaluza-Klein theory and from the warped extra dimensions models (Randall-Sundrum). The generated Physics contains the Bulk KK graviton with spin-2 and the Bulk KK Radion with spin-0. The samples are generated with different mass points, ranging from 260 GeV up to 1 TeV, as aused and analyzed in our analysis. The computed cross sections of the event generation $X(BK - Gr/Radion) \rightarrow HH$ are listed in Table 4.3.

m_{HH}	$\sigma(Graviton \rightarrow HH)[pb]$	$\sigma(Radion \rightarrow HH)[pb]$
200	11.4	18.3
260	3.40	10.3
300	1.78	7.2
400	0.483	3.37
500	0.175	1.96
600	0.073	1.24
700	0.035	0.804
800	1.82×10^{-2}	0.543
900	9.8×10^{-3}	0.374
1000	05.67×10^{-3}	0.263
1100	3.3×10^{-3}	0.187
1200	2.01×10^{-3}	0.135
1300	1.25×10^{-3}	9.86×10^{-2}
1400	8.48×10^{-4}	7.31×10^{-2}
1500	5.73×10^{-4}	5.42×10^{-2}
2000	9.06×10^{-5}	1.43×10^{-2}
2500	1.87×10^{-5}	4.29×10^{-3}
3000	4.41×10^{-6}	1.38×10^{-3}
3500	1.15×10^{-6}	4.70×10^{-4}
4500	8.93×10^{-8}	5.34×10^{-5}
5000	2.49×10^{-8}	1.83×10^{-5}
6000	1.83×10^{-9}	1.86×10^{-6}
6500	4.61×10^{-10}	5.35×10^{-7}

Table 4.3. . Computed cross sections of the event generation $X(BK - Gr/Radion) \rightarrow HH$ [122]

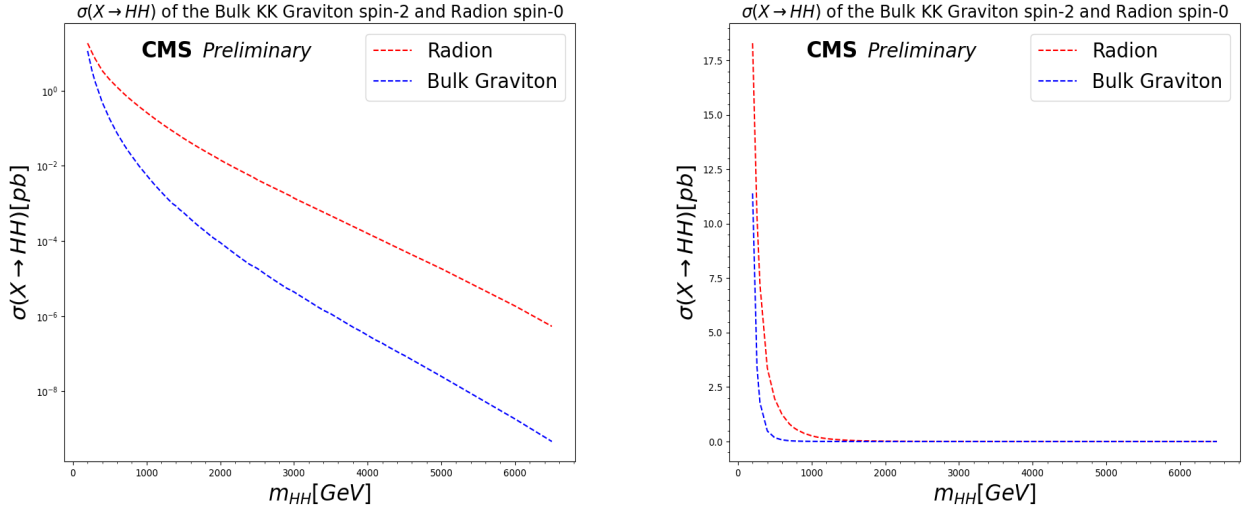


Figure 4.8. Plots of the computed cross sections in linear (right) and log (left) mode of the event generation $X(BK - Gr/Radion) \rightarrow HH$.

4.2 Simulations

In the section [4.2] above, the generation of Monte Carlo generated signal events have been thoroughly discussed and explained. Anyway, this is not the final format and stage we have achieved to perform our $HH \rightarrow b\bar{b}ZZ(4l)$ analysis. The goal is to simulate the interaction of generated particles with the detector, to reconstruct their four momenta from the signal detected and so reconstruct the full event. This final data format containing all the information needed is called MiniAOD (Mini Analysis Object Data) [147]. The detector simulation step is an essential part of each high energy experiment as much as the construction of the detectors themselves. The detector simulation is the way to mimic the interaction of particles with detector material such as it realistically happens in the detector. The detector simulation is based on a full description of the CMS detector geometry which is implemented in the GEANT4 (GEometry ANd Tracking) [60] simulation toolkit. An accurate description of the full geometry of the detector and the materials of the detecting devices are provided to GEANT4 to simulate the particle response. The particles from the generator level are propagated through the detector materials taking into account the measured magnetic field map (for charged particles) and the interactions processes between the particle and the detector material such as the bremsstrahlung, the multiple scattering and the photon conversions. As a result of the interactions, GEANT4 produces a set of simulated hits in the active material such as the energy loss of a particle in the sensitive volume of a detector. Then those set of information hits, energy deposit, etc. from various sub-detectors are used to reconstruct high-level objects such as jets and leptons in a process known as "reconstruction".

At the CMS experiment, the physics event consists of the signals from all particles of an interaction, or possibly even several interactions, joined together. The process of pattern recognition serves to sort out which bits of information are related to the same particle, meanwhile the process of reconstruction is crucial, since we need to reconstruct the kinematic properties of each particle, in order to reveal the physical nature of the whole event.

The logical sorting of the events from the simulation to reconstruction relies on several data formats:

1. DAQ-RAW: It is the primary record of physics event, which comes from the detector response of the CMS front-end electronics together with the Level-1 (L1) trigger results.
2. RAW: It is the primary archive of the events of CERN tiers (Tier-0), where the data from the L1 triggers

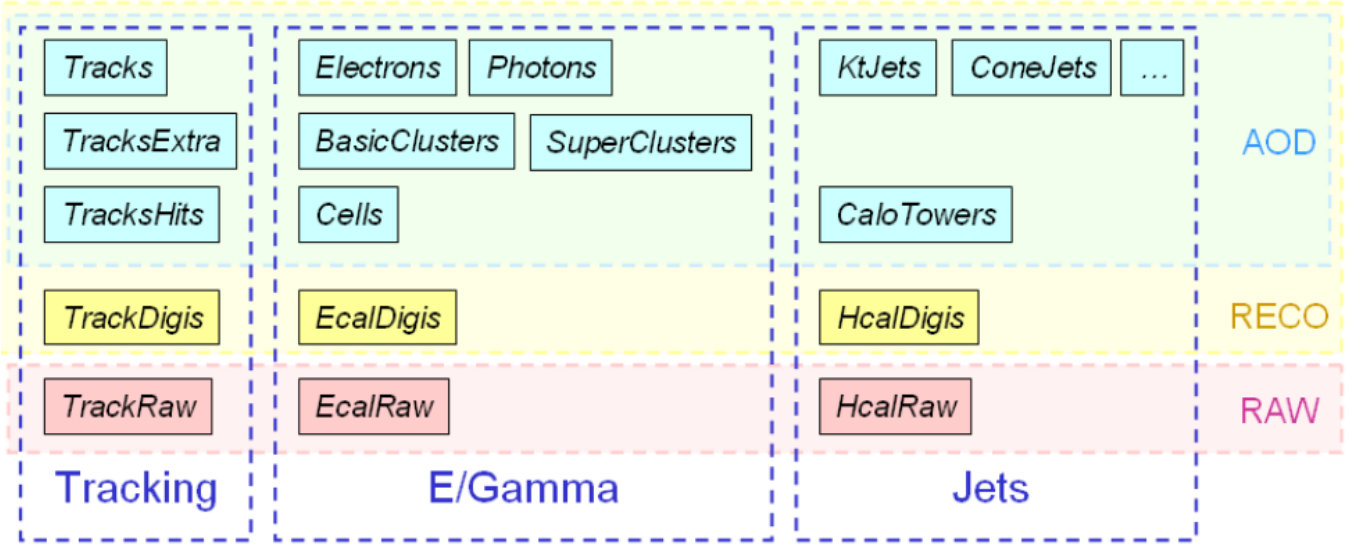


Figure 4.9. Schematic representation of the data tiers: Reconstructed (RECO) Data and Analysis Object Data (AOD). The RECO data contains objects from all stages of reconstruction coming from the hits of Physics objects (muons, electrons, etc.) of the particle shower during the pp -collisions, which were converted into detector response digital information from the CMS trigger system (RAW), meanwhile the AODs are derived from the RECO information to provide data for physics analyses in a convenient, compact format.

are formatted online to the High Level Trigger (HLT) selections data.

3. RECO: The archived HLT data from Tier-0 at CERN are analyzed and subsequently reconstructed, building the physics objects (e.g. tracks, vertices, jets, electrons, muons, etc.)
4. GEN: Data format including generated particle from the MC generator program (section [4.2])
5. SIM: Data format including the energy depositions of MC particles into the detector (sim hits) with the interfacing of a particle hadronizer (PHYTHIA . 8).
6. DIGI RAW: Data format including the Sim hits that are converted into the detector response. Basically the same as the RAW output of the detector of the experimental data.
7. MinAOD: Compact data format as a subset of RECO data tier, containing all the information needed of the reconstructed objects for the analysis, such as tracks, vertices, jets, electrons, muons, etc..

After finalizing the MC generations of the signal samples, for both: non-resonant HH production and the k_λ -scan, as well as the resonant HH production regarding the KK-Bulk Graviton spin-2 and Radion spin-0, we also performed a full detector simulation of all the three years (2016-2017-2018) for both cases and each mass point and k_λ value.

4.2.1 Simulation of the non-resonant double Higgs (H) 2016, 2017 and 2018 Signal Samples

Concerning the non-resonant signal HH production, a full CMS simulation with the production of the MiniAOD samples for each year (2016-2017-2018) was made and the samples are listed in [Table 4.4](#), [Table 4.5](#) and [Table 4.6](#); some kinematic distributions of the di-Higgs generation after the reconstruction within the CMS framework CMSSW_10_2_15 at the analysis step $|m_{4l} - 125.09|$ GeV and at least two jets can be seen in [Figure 4.10](#).

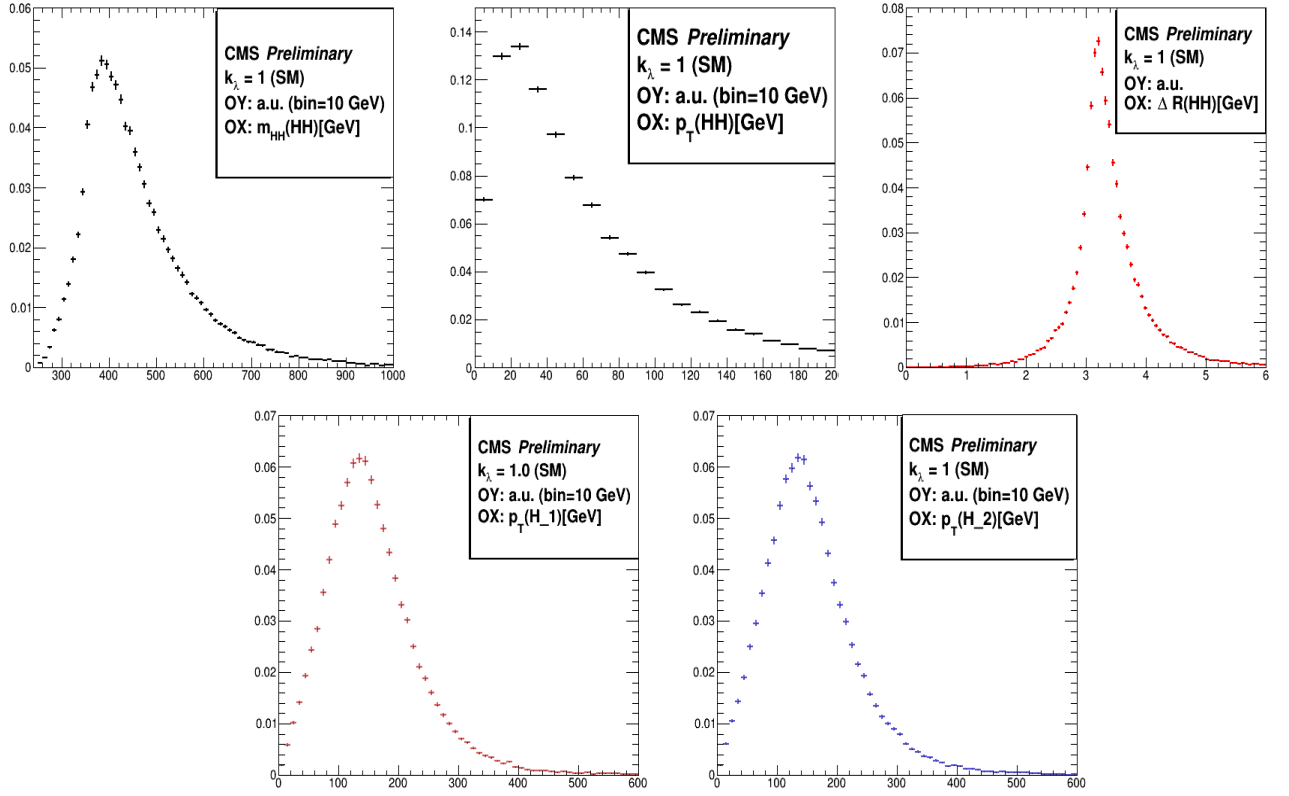


Figure 4.10. a) Top-left: Invariant mass of the di-Higgs M_{inv}^{HH} ; b) Top-center: Transverse momentum of the di-Higgs p_T^{HH} ; c) Top-right: Isolation of the two Higgs bosons within the cone with radius ΔR_{HH} ; d) Bottom-left: transverse momentum of the first Higgs boson $p_T^{H_1}$; e) Bottom-right: transverse momentum of the second Higgs boson $p_T^{H_2}$

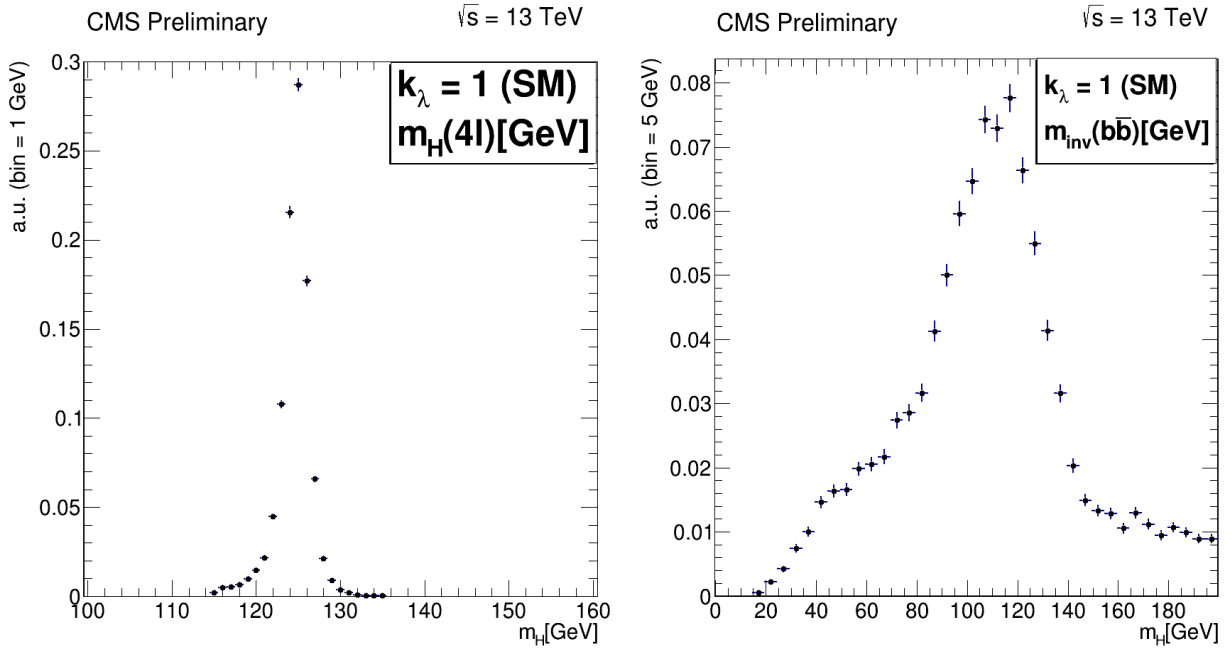


Figure 4.11. f) Left: Mass distribution of the Higgs boson reconstructed at GEN-SIM level from the two ZZ -bosons decaying into leptons ($l = e, \mu$); f) Right: Mass distribution of the Higgs boson reconstructed at GEN-SIM level from two b -tagged jets

k_λ value	MC signal sample	$\sigma_{gg \rightarrow HH}$ [fb]
0	/HH_bb4l_2016_NLO_k_lambda_0/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/USER	60.06
2	/HH_bb4l_2016_NLO_k_lambda_2_GEN-SIM/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/USER	11.92
3	/HH_bb4l_2016_NLO_k_lambda_3_GEN-SIM/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/USER	15.59
4	/HH_bb4l_2016_NLO_k_lambda_4_GEN-SIM/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/USER	38.05
5	/HH_bb4l_2016_NLO_k_lambda_5_GEN-SIM/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/USER	79.03
6	/HH_bb4l_2016_NLO_k_lambda_6_GEN-SIM/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/USER	138.53
7	/HH_bb4l_2016_NLO_k_lambda_7_GEN-SIM/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/USER	216.86
8	/HH_bb4l_2016_NLO_k_lambda_8_GEN-SIM/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/USER	313.66
9	/HH_bb4l_2016_NLO_k_lambda_9_GEN-SIM/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/USER	429.11
10	/HH_bb4l_2016_NLO_k_lambda_10_GEN-SIM/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/USER	553.23
-1	/HH_bb4l_2016_NLO_k_lambda_-1_GEN-SIM/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/USER	111.86
-2	/HH_bb4l_2016_NLO_k_lambda_-2_GEN-SIM/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/USER	182.26
-3	/HH_bb4l_2016_NLO_k_lambda_-3_GEN-SIM/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/USER	271.30
-4	/HH_bb4l_2016_NLO_k_lambda_-4_GEN-SIM/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/USER	378.92
-5	/HH_bb4l_2016_NLO_k_lambda_-5_GEN-SIM/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/USER	505.92
-6	/HH_bb4l_2016_NLO_k_lambda_-6_GEN-SIM/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/USER	649.89
-7	/HH_bb4l_2016_NLO_k_lambda_-7_GEN-SIM/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/USER	813.38
-8	/HH_bb4l_2016_NLO_k_lambda_-8_GEN-SIM/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/USER	995.90
-9	/HH_bb4l_2016_NLO_k_lambda_-9_GEN-SIM/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/USER	1195.82
-10	/HH_bb4l_2016_NLO_k_lambda_-10_GEN-SIM/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/USER	1414.86
SM case:1	/SM-HH-NLO_POWHEG_GEN-SIM/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/USER	26.79

Table 4.4. . 2016 signal samples and cross sections for the non-resonant HH

k_λ value	MC signal sample	$\sigma_{gg \rightarrow HH}$ [fb]
0	/HH_bb4l_2017_NLO_k_lambda_0_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/USER	60.06
2	/HH_bb4l_2017_NLO_k_lambda_2_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/USER	11.92
3	/HH_bb4l_2017_NLO_k_lambda_3_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/USER	15.59
4	/HH_bb4l_2017_NLO_k_lambda_4_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/USER	38.05
5	/HH_bb4l_2017_NLO_k_lambda_5_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/USER	79.03
6	/HH_bb4l_2017_NLO_k_lambda_6_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/USER	138.53
7	/HH_bb4l_2017_NLO_k_lambda_7_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/USER	216.86
8	/HH_bb4l_2017_NLO_k_lambda_8_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/USER	313.66
9	/HH_bb4l_2017_NLO_k_lambda_9_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/USER	429.11
10	/HH_bb4l_2017_NLO_k_lambda_10_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/USER	553.23
-1	/HH_bb4l_2017_NLO_k_lambda_-1_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/USER	111.86
-2	/HH_bb4l_2017_NLO_k_lambda_-2_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/USER	182.26
-3	/HH_bb4l_2017_NLO_k_lambda_-3_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/USER	271.30
-4	/HH_bb4l_2017_NLO_k_lambda_-4_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/USER	378.92
-5	/HH_bb4l_2017_NLO_k_lambda_-5_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/USER	505.92
-6	/HH_bb4l_2017_NLO_k_lambda_-6_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/USER	649.89
-7	/HH_bb4l_2017_NLO_k_lambda_-7_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/USER	813.38
-8	/HH_bb4l_2017_NLO_k_lambda_-8_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/USER	995.90
-9	/HH_bb4l_2017_NLO_k_lambda_-9_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/USER	1195.82
-10	/HH_bb4l_2017_NLO_k_lambda_-10_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/USER	1414.86
SM case:1	/SM-NLO-HH-POWHEG/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/USER	26.79

Table 4.5. . 2017 signal samples and cross sections for the non-resonant HH

4.2.2 Simulation of the resonant KK-Bulk Graviton spin-2 and Radion spin-0 2016, 2017 and 2018 Signal Samples

Concerning the resonant signal HH production, a full CMS simulation with the production of the MiniAOD sample for each year (2016-2017-2018) was performed and the samples are listed in Table 4.7, Table 4.8 and

k_λ value	MC signal sample	$\sigma_{gg \rightarrow HH}$ [fb]
0	/HH_bb4l_2018_NLO_k_lambda_0_GEN-SIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/USER	60.06
2	/HH_bb4l_2018_NLO_k_lambda_2_GEN-SIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/USER	11.92
3	/HH_bb4l_2018_NLO_k_lambda_3_GEN-SIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/USER	15.59
4	/HH_bb4l_2018_NLO_k_lambda_4_GEN-SIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/USER	38.05
5	/HH_bb4l_2018_NLO_k_lambda_5_GEN-SIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/USER	79.03
6	/HH_bb4l_2018_NLO_k_lambda_6_GEN-SIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/USER	138.53
7	/HH_bb4l_2018_NLO_k_lambda_7_GEN-SIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/USER	216.86
8	/HH_bb4l_2018_NLO_k_lambda_8_GEN-SIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/USER	313.66
9	/HH_bb4l_2018_NLO_k_lambda_9_GEN-SIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/USER	429.11
10	/HH_bb4l_2018_NLO_k_lambda_10_GEN-SIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/USER	553.23
-1	/HH_bb4l_2018_NLO_k_lambda_-1_GEN-SIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/USER	111.86
-2	/HH_bb4l_2018_NLO_k_lambda_-2_GEN-SIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/USER	182.26
-3	/HH_bb4l_2018_NLO_k_lambda_-3_GEN-SIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/USER	271.30
-4	/HH_bb4l_2018_NLO_k_lambda_-4_GEN-SIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/USER	378.92
-5	/HH_bb4l_2018_NLO_k_lambda_-5_GEN-SIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/USER	505.92
-6	/HH_bb4l_2018_NLO_k_lambda_-6_GEN-SIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/USER	649.89
-7	/HH_bb4l_2018_NLO_k_lambda_-7_GEN-SIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/USER	813.38
-8	/HH_bb4l_2018_NLO_k_lambda_-8_GEN-SIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/USER	995.90
-9	/HH_bb4l_2018_NLO_k_lambda_-9_GEN-SIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/USER	1195.82
-10	/HH_bb4l_2018_NLO_k_lambda_-10_GEN-SIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/USER	1414.86
SM case:1	/SM_NLO_HH_POWHEG/ilmargje-MiniAOD-bf1c2c60985de2425c87e47a33fef7/USER	26.79

Table 4.6. . 2018 signal samples and cross sections for the non-resonant HH

Table 4.9. In the [Figure 4.12](#) below, we can observe some kinematical distributions of the di-Higgs generation coming from the decay of Radion spin-0, after the reconstruction within the CMS framework CMSSW_10_2_22 at the analysis step $|m_{4l} - 125.09|$ GeV and at least two jets:

Process	MC signal sample (*: USER)	$\sigma_{X \rightarrow HH}$
$X \rightarrow HH \rightarrow b\bar{b}4l$ Graviton spin-2	/BulkGraviton_M260/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/*	3.4000 pb
	/BulkGraviton_M300/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/*	1.7800 pb
	/BulkGraviton_M400/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/*	0.4830 pb
	/BulkGraviton_M500/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/*	0.1750 pb
	/BulkGraviton_M600/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/*	0.0735 pb
	/BulkGraviton_M700/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/*	0.0350 pb
	/BulkGraviton_M800/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/*	0.0182 pb
	/BulkGraviton_M900/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/*	0.0098 pb
	/BulkGraviton_M1000/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/*	0.0057 pb
$X \rightarrow HH \rightarrow b\bar{b}4l$ Radion spin-0	/Radion_M260/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/*	10.300 pb
	/Radion_M300/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/*	7.200 pb
	/Radion_M400/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/*	3.370 pb
	/Radion_M500/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/*	1.960 pb
	/Radion_M600/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/*	1.240 pb
	/Radion_M700/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/*	0.804 pb
	/Radion_M800/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/*	0.543 pb
	/Radion_M900/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/*	0.374 pb
	/Radion_M1000/ilmargje-MiniAOD-53f8667ba4b240d5eafd36e71bf34742/*	0.263 pb

Table 4.7. . 2016 signal samples and cross sections for the Bulk Graviton spin-2 and Radion spin-0

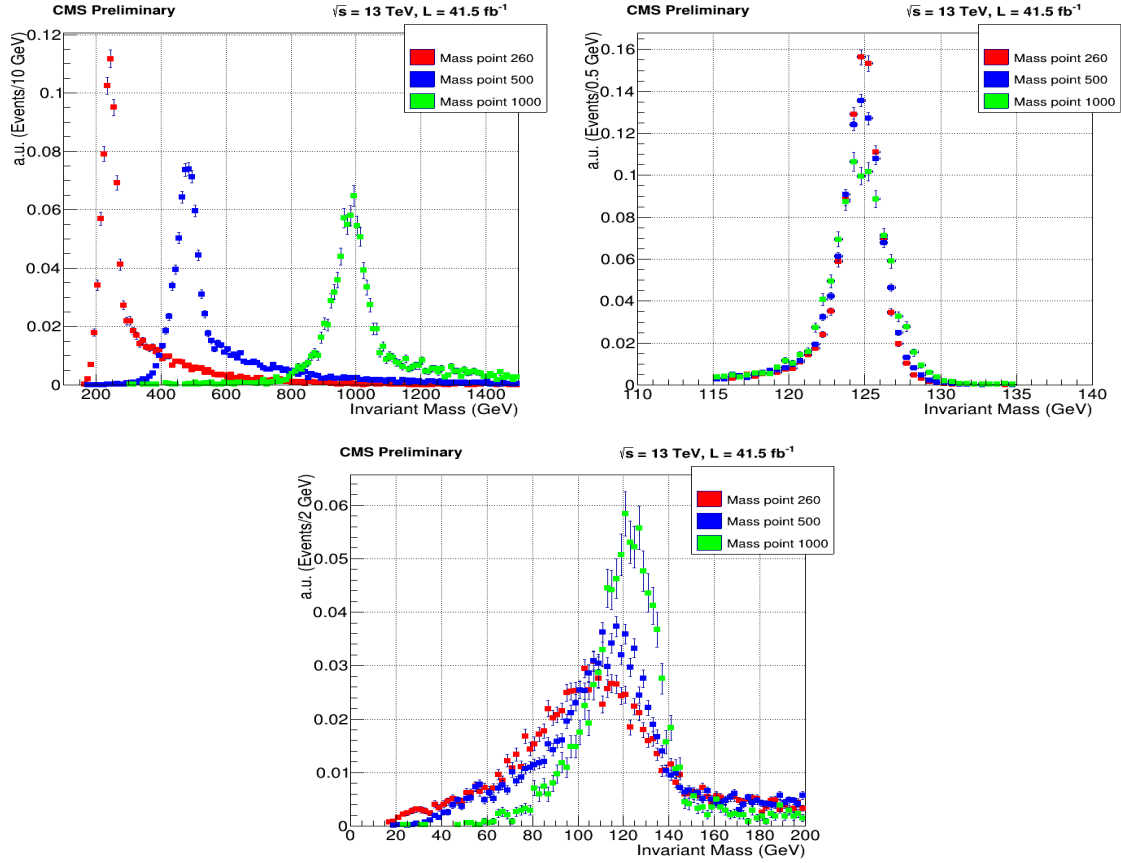


Figure 4.12. Kinematical distributions of the Radion spin-0 at 250 GeV, 500 GeV, 1000 GeV at the construction step within the signal region of $|m_{4l} - 125.09| - 10 \text{ GeV} \&\& \geq 2 \text{ jets}$: a) Left: invariant mass distribution M_{inv}^{HH} Radion spin-0 at 250 GeV, 500 GeV, 1000 GeV; b) Centre: invariant mass distribution M_{inv}^H from the $H \rightarrow ZZ(4l)$ Radion spin-0 at 250 GeV, 500 GeV, 1000 GeV; c) Right: invariant mass distribution M_{inv}^H from the $H \rightarrow b\bar{b}$ Radion spin-0 at 250 GeV, 500 GeV, 1000 GeV;

Process	MC signal sample (*: USER)	$\sigma_{X \rightarrow HH}$
$X \rightarrow HH \rightarrow b\bar{b}4l$ Graviton spin-2	/BKGr_260GeV_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/*	3.4000 pb
	/BKGr_300GeV_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/*	1.7800 pb
	/BKGr_400GeV_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/*	0.4830 pb
	/BKGr_500GeV_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/*	0.1750 pb
	/BKGr_600GeV_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/*	0.0735 pb
	/BKGr_700GeV_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/*	0.0350 pb
	/BKGr_800GeV_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/*	0.0182 pb
	/BKGr_900GeV_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/*	0.0098 pb
	/BKGr_1000GeV_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/*	0.0057 pb
$X \rightarrow HH \rightarrow b\bar{b}4l$ Radion spin-0	/Radion_260GeV_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/*	10.300 pb
	/Radion_300GeV_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/*	7.200 pb
	/Radion_400GeV_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/*	3.370 pb
	/Radion_500GeV_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/*	1.960 pb
	/Radion_600GeV_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/*	1.240 pb
	/Radion_700GeV_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/*	0.804 pb
	/Radion_800GeV_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/*	0.543 pb
	/Radion_900GeV_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/*	0.374 pb
	/Radion_1000GeV_GEN-SIM/ilmargje-MiniAOD-320a10544d366c5c2cbf9f153fc62ac6/*	0.263 pb

Table 4.8. . 2017 signal samples and cross sections for the Bulk Graviton spin-2 and Radion spin-0

Process	MC signal sample	$\sigma_{X \rightarrow HH}$
$X \rightarrow HH \rightarrow b\bar{b}4l$ Graviton spin-2	/BulkGravitonM260-GENSIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/*	3.4000 pb
	/BulkGravitonM300-GENSIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/*	1.7800 pb
	/BulkGravitonM400-GENSIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/*	0.4830 pb
	/BulkGravitonM500-GENSIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/*	0.1750 pb
	/BulkGravitonM600-GENSIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/*	0.0735 pb
	/BulkGravitonM700-GENSIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/*	0.0350 pb
	/BulkGravitonM800-GENSIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/*	0.0182 pb
	/BulkGravitonM900-GENSIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/*	0.0098 pb
	/BulkGravitonM1000-GENSIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/*	0.0057 pb
$X \rightarrow HH \rightarrow b\bar{b}4l$ Radion spin-0	/Radion_M260-GENSIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/*	10.300 pb
	/Radion_M300-GENSIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/*	7.200 pb
	/Radion_M400-GENSIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/*	3.370 pb
	/Radion_M500-GENSIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/*	1.960 pb
	/Radion_M600-GENSIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/*	1.240 pb
	/Radion_M700-GENSIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/*	0.804 pb
	/Radion_M800-GENSIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/*	0.543 pb
	/Radion_M900-GENSIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/*	0.374 pb
	/Radion_M1000-GENSIM/ilmargje-MiniAOD-0bd58594e6ade05f64e0c3a8301c3139/*	0.263 pb

Table 4.9. . 2018 signal samples and cross sections for the Bulk Graviton spin-2 and Radion spin-0

4.2.3 Background samples

The MC generated samples of Physics events for the background, were not generated by us, instead it was done by the CMS official production team in connection with the generator subgroup of the MC generations of such samples. The list of these samples used in both analysis, non-resonant SM $HH \rightarrow b\bar{b}ZZ(4l)$, where $l = e, \mu$, is shown in Table 4.10. The SM single Higgs boson processes constitute a background for this analysis. Descriptions of the SM Higgs boson to $1Z^0Z^0$ production are obtained by using the POWHEG V2 [76] [100][3] generator for the main production modes: gluon fusion ($gg \rightarrow H$) including quark mass effects [14], vector boson fusion (VBF) [101], and associated production ($W^\pm H$, $Z^0 H$, $b\bar{b}bH$ and $t\bar{t}H$ [87]). In the case of $W^\pm H$ and $Z^0 H$, the MINLO HVJ extension of POWHEG is used [94]. The description of the decay of the Higgs boson to four leptons is obtained using the JHUGEN generator [78]. In the case of $W^\pm H$, $Z^0 H$ and $t\bar{t}H$, the Higgs boson is allowed to decay to $H \rightarrow ZZ \rightarrow 2l2X$ such that 4-lepton events where two leptons originate from the decay of associated $W^\pm$, Z^0 bosons or top quarks are also taken into account in the simulation. Showering of parton-level events is done using PYTHIA-8.209, and in all cases matching is performed by allowing QCD emissions at all energies in the shower and vetoing them afterwards according to the POWHEG internal scale. All samples are generated with the NNPDF3.1 NLO parton distribution functions (PDFs). Production of Z^0Z^0 via quark-antiquark annihilation is generated at next-to-leading order (NLO) using POWHEG-BOX-V2 [103] and PYTHIA-8.2, with the same settings as for the Higgs signal. As this simulation covers a large range of Z^0Z^0 invariant masses, dynamical QCD factorization and renormalization scales have been chosen, equal to $m_{Z^0Z^0}$. The $gg \rightarrow Z^0Z^0$ process is simulated at leading order (LO) with MCFM [24][25]. In order to match the $gg \rightarrow HH \rightarrow Z^0Z^0$ transverse momentum spectra predicted by POWHEG at NLO, the showering for MCFM samples is performed with different PYTHIA8 settings, allowing only emissions up to the parton-level scale. Table 4.10 summarizes these MC simulation datasets used for this analysis.

4.2.4 The Pile-Up reweighting and MC weights

An important feature to ensure that the simulation can mimik the collisions in real data, is the procedure of the pile-up (PU) re-weighting, where the PU is the total number of pp interactions per bunch. The Monte Carlo samples are generated with distributions implementing a certain theoretical number of pileup inter-

actions according to the expectations of preliminary studies from the "Lumi POG" [124], which is meant to cover approximately the real PU value. In MC generation, the "pileup" refers to the number of additional interactions added to the primary hard-scatter process. This value is not expected to exactly match the PU present in the data, since only the experimental data collection will then deliver to us the present number of PU.

The PU is calculated from three parameter: the instantaneous luminosity, the total pp -inelastic cross section and LHC orbit frequency of 11.246 Hz [166]:

$$\mu_{PU} = \frac{\mathcal{L}_{inst} \cdot \sigma_{pp}}{f_{rev}} \quad [166] \quad (4.6)$$

The PU value is then computed on a "per luminosity section" basis (lumi-section), where a lumi-section is CMS luminosity unit (about 23.3 s). The number of reconstructed primary vertices is sensitive to primary vertex reconstruction and to any differences in the underlying event in data vs MC. Furthermore, the number of reconstructed primary vertices can be biased by the offline event selection criteria and even by the trigger. In order to factorize these effects, we reweight the number of pileup interactions from the simulation truth (stored in the MC samples as PileupInfo collection). The target pileup distribution for data is derived by using the lumi-section instantaneous luminosity per bunch crossing together with the total pp -inelastic cross-section to generate an expected pileup distribution, correctly weighted by integrating the luminosity per bunching crossing lumi-section over the entire data-taking period. These values then are stored in the JSON [166] data format files.

The distribution of pileup for individual events will therefore be a Poisson distribution around this average. The difference is corrected by re-weighting simulated events to match the PU distribution in data. An example of re-weighting procedure for 2016-2017-2018 data is shown in the Figure 4.13 below.

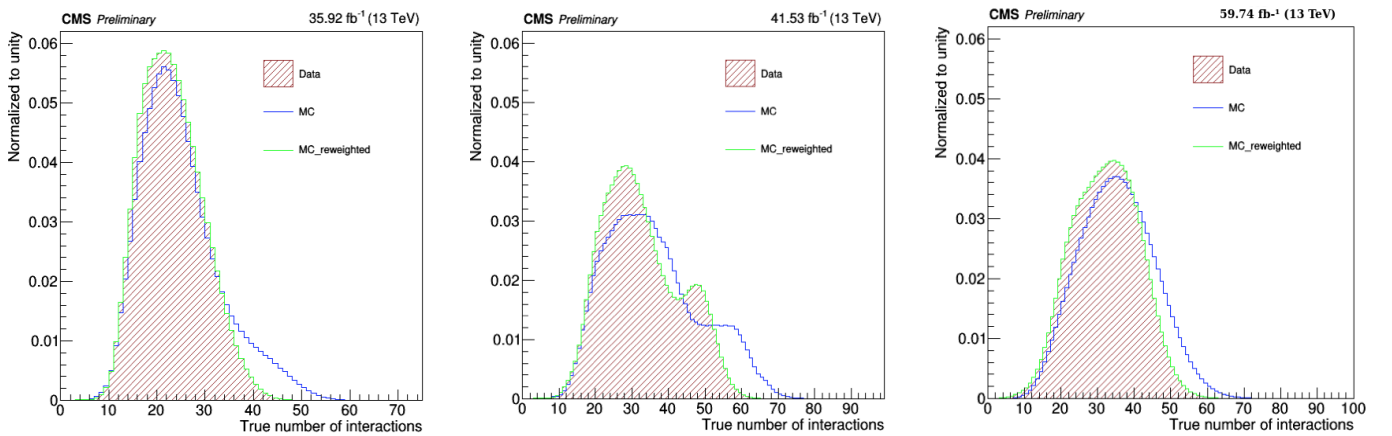


Figure 4.13. Distribution of pileup in 2016 (left), 2017 (center) and 2018 (right) Data and MC, before and after the application of PU weights.

In addition to the pileup weight, each MC event needs to be assigned a weight, which will then effect the final event selection. The MC event is weighted by a factor w , as following:

$$w = \frac{\mathcal{L} \cdot \sigma(Z \rightarrow XY) \cdot BR \cdot \mu_{PU} \cdot SF_{Al}}{N} \cdot \epsilon_{Trig}. \quad (4.7)$$

where $\sigma(Z \rightarrow XY)$ is the process production cross section, $\mathcal{L}$ is the integrated luminosity, N is the total number of MC simulated events for the corresponding process, ϵ_{Trig} is the efficiency of the trigger (check section [4.3]) selection for the MC samples, given as the ration between the number of events, which survived the Trigger selection conditions, over the total number of the events of that MC samples.

Process	Dataset Name	$\sigma \times BR(\times \epsilon_{filter})$
$gg \rightarrow H \rightarrow ZZ \rightarrow 4l$	/GluGluHToZZTo4L_M125_13TeV_powheg2_JHUGenV709_pythia8/[1]	12.18 fb
$q\bar{q} \rightarrow q\bar{q}H \rightarrow q\bar{q}ZZ \rightarrow 4lq\bar{q}$	/VBF_HToZZTo4L_M125_13TeV_powheg2_JHUGenV709_pythia8/[1]	1.044 fb
$q\bar{q} \rightarrow W^+H \rightarrow W^+ZZ \rightarrow 4l + X$	/WplusH_HToZZTo4L_M125_13TeV_powheg2-minlo-HWJ_JHUGenV709_pythia8/[1]	0.232 fb
$q\bar{q} \rightarrow W^-H \rightarrow W^-ZZ \rightarrow 4l + X$	/WminusH_HToZZTo4L_M125_13TeV_powheg2-minlo-HWJ_JHUGenV709_pythia8/[1]	0.147 fb
$q\bar{q} \rightarrow ZH \rightarrow ZZZ \rightarrow 4l + X$	/ZH_HToZZ_4LFilter_M125_13TeV_powheg2-minlo-HZJ_JHUGenV709_pythia8/[1]	0.668 fb
$gg \rightarrow t\bar{t}H \rightarrow t\bar{t}ZZ \rightarrow 4l + X$	/ttH_HToZZ_4LFilter_M125_13TeV_powheg_JHUGenV709_pythia8/[1]	0.393 fb
$gg \rightarrow b\bar{b}H \rightarrow b\bar{b}ZZ \rightarrow 4l + X$	/bbH_HToZZTo4L_M125_13TeV_JHUGenV7011_pythia8/[1]	0.135 fb
$q\bar{q}/qg \rightarrow tHq \rightarrow tqZZ \rightarrow 4l + X$	/tqH_HToZZTo4L_M125_13TeV_JHUGenV7011_pythia8/[1]	0.0213 fb
$q\bar{q} \rightarrow ZZ \rightarrow 4l$	/ZZTo4L_13TeV_powheg_pythia8/[1]	1.256 pb
$q\bar{q} \rightarrow ZZ \rightarrow 4l$	/ZZTo4L_13TeV-amcatnloFXFX-pythia8/[1]	1.212 pb
$gg \rightarrow ZZ \rightarrow 4e$	/GluGluToContinToZZTo4e_13TeV_MCFM701/[1]	0.00159 pb
$gg \rightarrow ZZ \rightarrow 4\mu$	/GluGluToContinToZZTo4mu_13TeV_MCFM701/[1]	0.00159 pb
$gg \rightarrow ZZ \rightarrow 4\tau$	/GluGluToContinToZZTo4tau_13TeV_MCFM701/[1]	0.00159 pb
$gg \rightarrow ZZ \rightarrow 2e2\mu$	/GluGluToContinToZZTo2e2mu_13TeV_MCFM701/[1]	0.00319 pb
$gg \rightarrow ZZ \rightarrow 2e2\mu$	/GluGluToContinToZZTo2e2tau_13TeV_MCFM701/[1]	0.00319 pb
$gg \rightarrow ZZ \rightarrow 2\mu2\tau$	/GluGluToContinToZZTo2mu2tau_13TeV_MCFM701/[1]	0.00319 pb
$ZZ \rightarrow ll + \text{jets}$	/DYJetsToLL_M-50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/[1]	6104 pb
$ZZ \rightarrow ll + \text{jets}$	/DYJetsToLL_M-10to50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/[1]	18610 pb
$WZ \rightarrow ZZ \rightarrow 3l\nu$	/WZTo3LNu_TuneCUETP8M1_13TeV_powheg-pythia8/[1]	4.430 pb
WWZ	/WWZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8/[1]	0.1651 pb
WZZ	/WZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8/[1]	0.05565 pb
ZZZ	/ZZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8/[1]	0.01398 pb
$t\bar{t} + \text{jets}$	/TTJets_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/[1]	815.96 pb
$t\bar{t} + ZZ$	/TTZZ_TuneCUETP8M2T4_13TeV-madgraph-pythia8/[1]	0.001572 pb
$t\bar{t} + WW$	/TTWW_TuneCUETP8M2T4_13TeV-madgraph-pythia8/[1]	0.007883 pb
$t\bar{t} + Z$	/ttZJets_13TeV_madgraphMLM/[1]	0.259 pb
[1] RunIISummer16MiniAODv3-PUMoriond17_94X_mcRun2_asymptotic for 2016 [1] RunIIFall17MiniAODv2-PU2017_12Apr2018_94X_mc2017_realistic_v* for 2017 [1] RunIIAutumn18MiniAOD-102X_upgrade2018_realistic_v* for 2018		

Table 4.10. Background Monte Carlo samples and cross sections

4.3 Trigger selection and real data samples

This analysis uses a data sample recorded by the CMS experiment during the time period of 2016-2017-2018 corresponding to a total integrated luminosity of 137 fb^{-1} .

4.3.1 The 2016, 2017 and 2018 datasets

Our full Run II analysis is based on electron, muons and jets in the final state of the di-Higgs system, therefore it relies on single and multilepton HLT paths, exploiting the five different primary datasets (PDs), DoubleEG, DoubleMuon, MuEG, SingleElectron and SingleMuon for 2016 and 2017, meanwhile in case of 2018 they are four PDs: DoubleMuon, MuEG, EGamma and SingleMuon, where each of them combines a certain collection of high-level trigger (HLT), as we can see in [Table 4.12](#), [Table 4.13](#) and [Table 4.15](#).

Year	Luminosity fb^{-1}	ReReco	Dataset
2016	35.8	17Jul2018-v1	DoubleEG, DoubleMuon, MuEG, SingleMuon, SingleElectron
2017	41.5	31Mar2018-v1	DoubleEG, DoubleMuon, MuEG, SingleMuon, SingleElectron
2018	59.7	17Sep2018-v1	DoubleMuon, MuEG, Egamma, SingleMuon

The HLT paths used for 2016, 2017 and 2018 collisions data are listed with the associated primary datasets in [Table 4.11](#), [Table 4.14](#) and [Table 4.16](#). In order to avoid duplicate events from different PDs, events are taken as following:

1) 2016, 2017:

- DoubleEG, if they pass the diEle (at least 2 electrons and more per event) or triEle (at least 3 electrons and more per event) triggers,
- DoubleMuon, if they pass the diMuon (at least 2 muons and more per event) or triMuon (at least 2 muons and more per event) triggers and fail the diEle and triEle triggers,
- MuEG, if they pass the MuEle or MuDiEle or DiMuEle triggers and fail the diEle, triEle, diMuon, and triMuon triggers,
- SingleElectron, if they pass the singleElectron trigger and fail all the above triggers
- SingleMuon, if they pass the singleMuon trigger and fail all the above triggers

2) 2018:

- from EGamma, if they pass the diEle or triEle or singleElectron triggers,
- from DoubleMuon, if they pass the diMuon or triMuon triggers and fail the diEle and triEle triggers,
- from MuEG, if they pass the MuEle or MuDiEle or DiMuEle triggers and fail the diEle, triEle, singleElectron, diMuon, and triMuon triggers,
- from SingleMuon, if they pass the singleMuon trigger and fail all the above triggers.

Datasets	Triggers
DoubleEG	PASS: diEle triEle triggers
DoubleMU	PASS: diMu triMu triggers FAIL: all the above triggers
MuEG	PASS: MuEle MuDiEle DiMuEle FAIL: all the above triggers
SingleElectron	PASS: SingleElectrons FAIL: all the above triggers
SingleMuon	PASS: SingleMuons FAIL: all the above triggers

The efficiency in data of the combination of triggers used in the analysis with respect to the offline reconstruction and selection is measured by considering 4l events triggered by single lepton triggers. One of the four reconstructed leptons (the “tag”) is geometrically matched to a trigger object passing the final filter of one of the single muon or single electron triggers. The other three leptons are used as “probes”. In each 4l event there are up to 4 possible tag-probe combinations, and all possible combinations are counted in the denominator of the efficiency. For each of the three probe leptons all matching trigger filter objects are collected. Then the matched trigger filter objects of the three probe leptons are combined in attempt to reconstruct any of the triggers used in the analysis. If any of the analysis triggers can be formed using the probe leptons, the set of probes is also counted in the numerator of the efficiency. This method does not have a perfect closure in MC events due to the fact that the presence of a fourth lepton increases the trigger efficiency, and this effect is not accounted for. Also, in the $2e2\mu$ final state, the three probe leptons cannot be combined to form all possible triggers which can collect events with two electrons and two muons (e.g. if the tag lepton is an electron, the three remaining leptons cannot pass a double electron trigger). Therefore the method is also exercised on MC and the difference between data and MC is used to determine the reliability of the simulation.

HLT path	primary dataset
HLT_Ele17_Ele12_CaloIdL_TrackIdL_IsoVL_DZ	DoubleEG
HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ	DoubleEG
HLT_DoubleEle33_CaloIdL_GsfTrkIdVL	DoubleEG
HLT_Ele16_Ele12_Ele8_CaloIdL_TrackIdL	DoubleEG
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL	DoubleMuon
HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL	DoubleMuon
HLT_TripleMu_12_10_5	DoubleMuon
HLT_Mu8_TrkIsoVVL_Ele17_CaloIdL_TrackIdL_IsoVL	MuonEG
HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL	MuonEG
HLT_Mu17_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL	MuonEG
HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL	MuonEG
HLT_Mu23_TrkIsoVVL_Ele8_CaloIdL_TrackIdL_IsoVL	MuonEG
HLT_Mu8_DiEle12_CaloIdL_TrackIdL	MuonEG
HLT_DiMu9_Ele9_CaloIdL_TrackIdL	MuonEG
HLT_Ele25_eta2p1_WPTight	SingleElectron
HLT_Ele27_WPTight	SingleElectron
HLT_Ele27_eta2p1_WPLoose_Gsf	SingleElectron
HLT_IsoMu20 OR HLT_IsoTkMu20	SingleMuon
HLT_IsoMu22 OR HLT_IsoTkMu22	SingleMuon

Table 4.11.. Trigger paths used in 2016 collision data

Run-Range	Datasets	Integrated luminosity $\mathcal{L}$
273150-275376	/DoubleMuon/Run2016B-17Jul2018-ver2-v1/MINIAOD /DoubleEG/Run2016B-17Jul2018-ver2-v1/MINIAOD /MuonEG/Run2016B-17Jul2018-ver2-v1/MINIAOD /SingleElectron/Run2016B-17Jul2018-ver2-v1/MINIAOD /SingleMuon/Run2016B-17Jul2018-ver2-v1/MINIAOD	5.892 fb^{-1}
276315-276811	/DoubleMuon/Run2016C-17Jul2018-v1/MINIAOD /DoubleEG/Run2016C-17Jul2018-v1/MINIAOD /MuonEG/Run2016C-17Jul2018-v1/MINIAOD /SingleElectron/Run2016C-17Jul2018-v1/MINIAOD /SingleMuon/Run2016C-17Jul2018-v1/MINIAOD	2.646 fb^{-1}
276315-276811	/DoubleMuon/Run2016D-17Jul2018-v1/MINIAOD /DoubleEG/Run2016D-17Jul2018-v1/MINIAOD /MuonEG/Run2016D-17Jul2018-v1/MINIAOD /SingleElectron/Run2016D-17Jul2018-v1/MINIAOD /SingleMuon/Run2016D-17Jul2018-v1/MINIAOD	4.353 fb^{-1}
276831-277420	/DoubleMuon/Run2016E-17Jul2018-v1/MINIAOD /DoubleEG/Run2016E-17Jul2018-v1/MINIAOD /MuonEG/Run2016E-17Jul2018-v1/MINIAOD /SingleElectron/Run2016E-17Jul2018-v1/MINIAOD /SingleMuon/Run2016E-17Jul2018-v1/MINIAOD	4.117 fb^{-1}
277932-278808	/DoubleMuon/Run2016F-17Jul2018-v1/MINIAOD /DoubleEG/Run2016F-17Jul2018-v1/MINIAOD /MuonEG/Run2016F-17Jul2018-v1/MINIAOD /SingleElectron/Run2016F-17Jul2018-v1/MINIAOD /SingleMuon/Run2016F-17Jul2018-v1/MINIAOD	3.186 fb^{-1}
278820-280385	/DoubleMuon/Run2016G-17Jul2018-v1/MINIAOD /DoubleEG/Run2016G-17Jul2018-v1/MINIAOD /MuonEG/Run2016G-17Jul2018-v1/MINIAOD /SingleElectron/Run2016G-17Jul2018-v1/MINIAOD /SingleMuon/Run2016G-17Jul2018-v1/MINIAOD	7.721 fb^{-1}
281207-284068	/DoubleMuon/Run2016H-17Jul2018-v1/MINIAOD /DoubleEG/Run2016H-17Jul2018-v1/MINIAOD /MuonEG/Run2016H-17Jul2018-v1/MINIAOD /SingleElectron/Run2016H-17Jul2018-v1/MINIAOD /SingleMuon/Run2016H-17Jul2018-v1/MINIAOD	8.857 fb^{-1}

Table 4.12. 2016 datasets used in the analysis

Run-Range	Datasets	Integrated luminosity $\mathcal{L}$
297046-299329	/DoubleMuon/Run2017B-31Mar2018-v1/MINIAOD /DoubleEG/Run2017B-31Mar2018-v1/MINIAOD /MuonEG/Run2017B-31Mar2018-v1/MINIAOD /SingleElectron/Run2017B-31Mar2018-v1/MINIAOD /SingleMuon/Run2017B-31Mar2018-v1/MINIAOD	4.792 fb^{-1}
299368-300676	/DoubleMuon/Run2017C-31Mar2018-v1/MINIAOD /DoubleEG/Run2017C-31Mar2018-v1/MINIAOD /MuonEG/Run2017C-31Mar2018-v1/MINIAOD /SingleElectron/Run2017C-31Mar2018-v1/MINIAOD /SingleMuon/Run2017C-31Mar2018-v1/MINIAOD	9.755 fb^{-1}
302030-303434	/DoubleMuon/Run2017D-31Mar2018-v1/MINIAOD /DoubleEG/Run2017D-31Mar2018-v1/MINIAOD /MuonEG/Run2017D-31Mar2018-v1/MINIAOD /SingleElectron/Run2017D-31Mar2018-v1/MINIAOD /SingleMuon/Run2017D-31Mar2018-v1/MINIAOD	4.319 fb^{-1}
303824-304797	/DoubleMuon/Run2017E-31Mar2018-v1/MINIAOD /DoubleEG/Run2017E-31Mar2018-v1/MINIAOD /MuonEG/Run2017E-31Mar2018-v1/MINIAOD /SingleElectron/Run2017E-31Mar2018-v1/MINIAOD /SingleMuon/Run2017E-31Mar2018-v1/MINIAOD	9.424 fb^{-1}
305040-306462	/DoubleMuon/Run2017F-31Mar2018-v1/MINIAOD /DoubleEG/Run2017F-31Mar2018-v1/MINIAOD /MuonEG/Run2017F-31Mar2018-v1/MINIAOD /SingleElectron/Run2017F-31Mar2018-v1/MINIAOD /SingleMuon/Run2017F-31Mar2018-v1/MINIAOD	13.50 fb^{-1}

Table 4.13. 2017 datasets used in the analysis

HLT path	primary dataset
HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL	DoubleEG
HLT_DoubleEle33_CaloIdL_GsfTrkIdVL	DoubleEG 0
HLT_Ele16_Ele1E_le8_CaloIdL_TrackIdL	DoubleEG
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8	DoubleMuon
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass8	DoubleMuon
HLT_TripleMu_12_10_5	DoubleMuon
HLT_TripleMu_10_5_5_D2	DoubleMuon
HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL	MuonEG
HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ	MuonEG
HLT_Mu12_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ	MuonEG
HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ	MuonEG
HLT_DiMu9_Ele9_CaloIdL_TrackIdL_DZ	MuonEG
HLT_Mu8_DiEle12_CaloIdL_TrackIdL	MuonEG
HLT_Mu8_DiEle12_CaloIdL_TrackIdL_DZ	MuonEG
HLT_Ele35_WPTightGsf_v	SingleElectron
HLT_Ele38_WPTight_Gsf_v	SingleElectron
HLT_Ele40_WPTight_Gsf_v	SingleElectron
HLT_IsoMu27	SingleMuon

Table 4.14. Trigger paths used in 2017 collision data

Run-Range	Datasets	Integrated luminosity $\mathcal{L}$
315252-316995	/DoubleMuon/Run2018A-17Sep2018-v2/MINIAOD /MuonEG/Run2018A-17Sep2018-v1/MINIAOD /SingleMuon/Run2018A-17Sep2018-v2/MINIAOD /EGamma/Run2018A-17Sep2018-v2/MINIAOD	14.00 fb^{-1}
317080-319310	/DoubleMuon/Run2018B-17Sep2018-v1/MINIAOD /MuonEG/Run2018B-17Sep2018-v1/MINIAOD /SingleMuon/Run2018B-17Sep2018-v1/MINIAOD /EGamma/Run2018B-17Sep2018-v1/MINIAOD	7.10 fb^{-1}
319337-320065	/DoubleMuon/Run2018C-17Sep2018-v1/MINIAOD /MuonEG/Run2018C-17Sep2018-v1/MINIAOD /SingleMuon/Run2018C-17Sep2018-v1/MINIAOD /EGamma/Run2018C-17Sep2018-v1/MINIAOD	6.94 fb^{-1}
320673-325175	/DoubleMuon/Run2018D-PromptReco-v2/MINIAOD /MuonEG/Run2018D-PromptReco-v2/MINIAOD /SingleMuon/Run2018D-PromptReco-v2/MINIAOD /EGamma/Run2018D-PromptReco-v2/MINIAOD	31.93 fb^{-1}

Table 4.15. 2018 datasets used in the analysis

HLT path	primary dataset
HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_v*	EGamma
HLT_DoubleEle25_CaloIdL_MW_v*	EGamma
HLT_Ele32_WPTight_Gsf_v*	EGamma
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8_v*	DoubleMuon
HLT_TripleMu_10_5_5_DZ_v*	DoubleMuon
HLT_TripleMu_12_10_5_v*	DoubleMuon
HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_v*	MuEG
HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v*	MuEG
HLT_Mu12_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v*	MuEG
HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v*	MuEG
HLT_DiMu9_Ele9_CaloIdL_TrackIdL_DZ_v*	MuEG
HLT_Mu8_DiEle12_CaloIdL_TrackIdL_DZ_v*	MuEG
HLT_IsoMu24_v*	SingleMuon

Table 4.16. Trigger paths used in 2018 collision data

Chapter 5

Analysis Workflow

The present analysis is based on the following physics objects: electrons, muons, and jets. The reconstruction process has several steps before we can have the final format files to be used for the analysis. As it was already explained in section [4.3], the reconstruction process starts from the raw data made of the particle hits within the tracker and muon system and the energy deposit of the hadrons within then CMS calorimeter, as we can see in [Figure 5.1](#). This "bombardment" of the particle shower coming from the proton-proton collisions, implies a relatively high number of physics parameters to be reconstructed and assigned properly to the all particles and physics events accounted for.

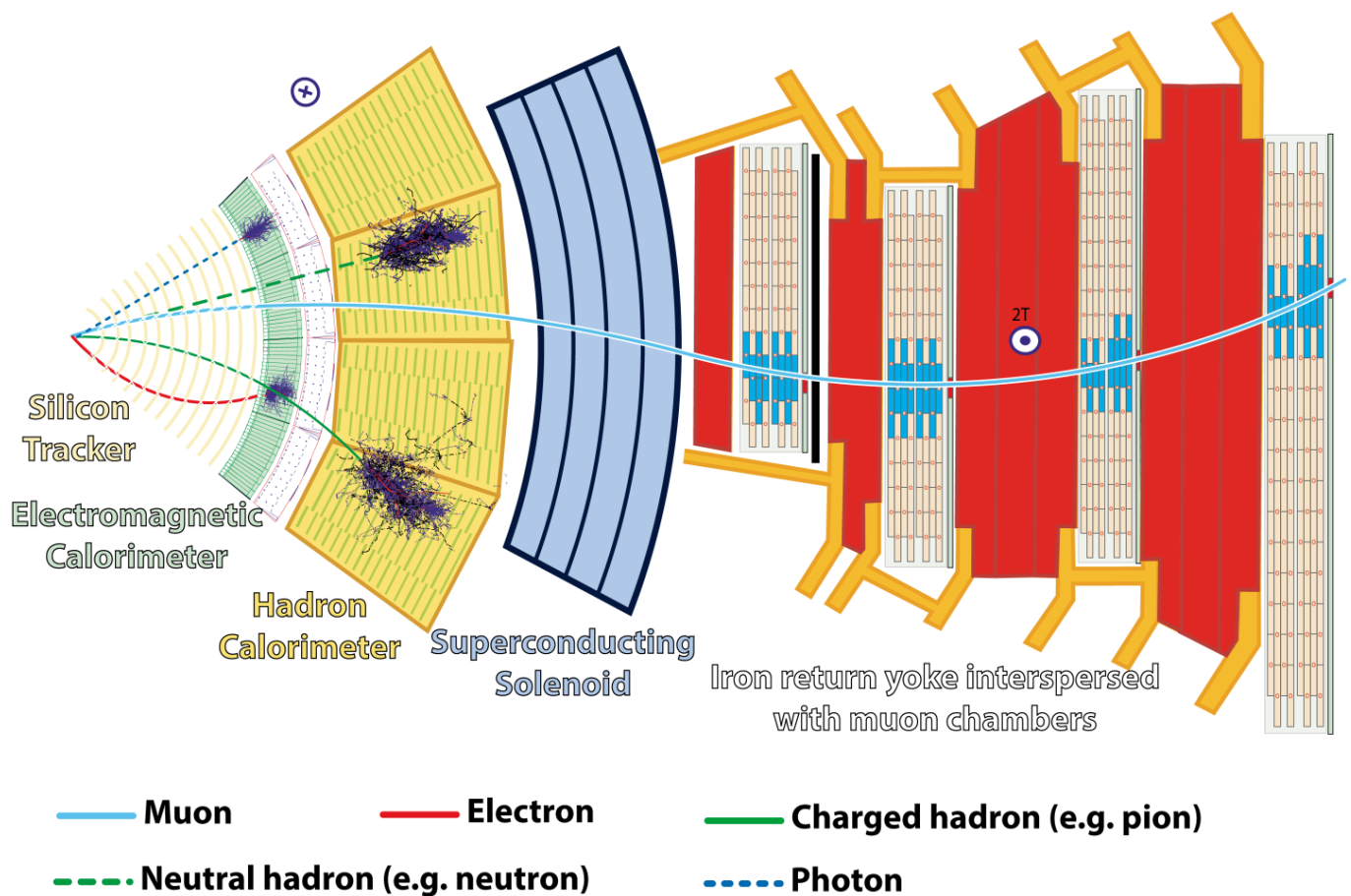


Figure 5.1. . Slice showing CMS sub-detectors and how particles interact with them as well as a detailed illustration of the experimental signature of final state particles in the CMS detector. This picture is adapted from the Ref. [142]

5.1 Physics Objects Selection

5.1.1 Particle Flow Algorithm

The particle flow (PF) algorithm [143] is used for the CMS experiment to reconstruct and identify all the particles from the redundant measurements from the CMS subsystems and reconstruct physics objects from raw detector data, performing a global event reconstruction. It collects all the information coming from all sub-detectors of the CMS experiment in order to reconstruct the full "physical" particles, which released their signal through the various layers. The source of all physics data comes from the first interaction point between two colliding protons, the primary vertex (PV) [137], where the hard scattering or proton-proton collisions happens, while the secondary vertices arise from the decay of long lived particles. The PF muon is reconstructed if the momentum of the global muon track matches the momentum of the muon track in the tracker within three standard deviations. If the tracks pass the criteria for PF muon they are removed from the PF block, meanwhile electron candidates are identified and reconstructed in case if a track in the tracker matches an ECAL cluster, then it will be identified as an electron, otherwise, if there is no track, in the ECAL cluster, this will correspond to a photon. The remaining tracks in the PF block that are linked to energy deposit in the HCAL, are used to create the charged hadron candidates. The left-over cluster in the ECAL without tracks in the tracker gets assigned to photons and similarly to the neutral hadrons in the HCAL.

5.1.2 Tracking and Clustering

As it is shown in Figure 5.1, after a pp -collision, the generated particles within the primary interaction vertex, propagate within the detector complex of CMS, leaving their signatures. The charged particles leave hits in the silicon detector of the tracker system and they deposit an amount of energy in the Electromagnetic Calorimeter (ECAL), meanwhile the neutral particles deposit energy in the Hadron calorimeter (HCAL), without hits in the tracker system. The muons are the only particles that can reach the muon system leaving hits with the muon spectrometer, where their identification and the reconstruction can occur.

The reconstruction of the hits and their momenta within the tracking system is performed by the Combinatorial Track Finder algorithm (CTF) based on Kalman Filtering (KF) [16, 77] to reconstruct the tracks across a wide range of particle momentum (100 MeV to 1 TeV). The CTF algorithm aims to build tracks from hits inside the silicon detectors and the momentum measurement by following the fired cells through the layers of the tracker system. The reconstructed tracks are required to have at least eight hits in total and in a distance of few mm from the beam axis. Finally, the reconstructed tracks should pass quality requirements based on number of hits, the track fit χ^2 and the track compatibility to originate from primary vertex.

Depending to how many cells are fired from the passing charged particles, we can define a cluster size in the calorimeter. The cluster size is formed from the cells of the calorimeter, where the energy deposit of each cell measured is either higher than that of a threshold value or higher of the neighboring cell. The combination of the clusters is called "superclusters". The clustering process is done separately for ECAL and HCAL and in both, barrel and endcap regions. The energy response of the calorimeter clusters is calibrated from test beam data, radioactive sources, and early collision data events. The goal of clustering process in the calorimeters is to define the following parameters: energy of electrons with their bremsstrahlung photons, energy of photons and neutral hadrons and energy of charged hadrons.

5.2 ELECTRONS $e^\mp$

Electron identification

Reconstructed electrons are identified and isolated by means of a Gradient Boosted Decision Tree (GBDT) multivariate classifier algorithm, which exploits observables from the electromagnetic cluster, the matching between the cluster and the electron track, observables based exclusively on tracking measurements as well as particle-flow (PF) isolation sums. The advantage of this technique over cut-based electron identification is the ability to reduce the amount of fake electrons especially in low p_T region without large impact on the signal efficiency. The classifier is trained on a Drell-Yan + jets MC sample for both signal and background, where the real electrons are reconstructed electrons geometrically matched with generated electrons from a Z boson, meanwhile all other electrons are considered as background or fake electrons. The result of the discriminant between true and fake electrons after the training can be seen in [Figure 5.2](#). The training is performed for two electron's p_T variant:

$$1) \quad 7\text{GeV} < p_T < 10\text{GeV} \quad \text{and} \quad 2) \quad p_T > 10\text{GeV}$$

The classification of the electron coming from the barrel or the endcap region or the transition between them, is also taken in consideration, where we build three η regions:

$$1) \quad |\eta| < 0.8, \quad 2) \quad 0.8 > |\eta| > 1.479, \quad 3) \quad |\eta| > 1.479$$

Finally, a cut on the classifier score has been chosen to distinguish between the signal and the fake electrons while preserving a high signal electron efficiency, as shown in the [Table 5.1](#).

Year	p_T -bins	η -bins		
		$ \eta < 0.8$	$0.8 > \eta > 1.479$	$ \eta > 1.479$
2016	$7\text{GeV} < p_T < 10\text{GeV}$	0.9503	0.9461	0.9387
	$p_T > 10\text{GeV}$	0.3782	0.3587	-0.5745
2017	$7\text{GeV} < p_T < 10\text{GeV}$	0.8521	0.8268	0.8694
	$p_T > 10\text{GeV}$	0.9825	0.9692	0.7935
2018	$7\text{GeV} < p_T < 10\text{GeV}$	0.8956	0.9111	0.9401
	$p_T > 10\text{GeV}$	0.0424	0.0047	-0.6042

Table 5.1. . Minimum BDT score required for passing the electron identification for 2016, 2017 and 2018 [\[128\]](#).

The classified "loose" electrons passing the identification cut are now classified as "tight" electrons for both of the analysis presented in this thesis. The full list of observables used can be seen in the [Table 5.2](#).

After the reconstruction process of the electrons, we end up with set of electrons that could be real electrons or fake objects that pass the electron reconstruction process and hence are considered as electron candidates. The electron reconstruction efficiency [\[158\]](#) is very high, which means that we only miss a small fraction of real electrons but we could have an amount of fake electrons. Therefore, a set of selection cuts are applied to the reconstructed electrons in order to decrease the amount of fake objects while keeping as much as possible the real electrons.

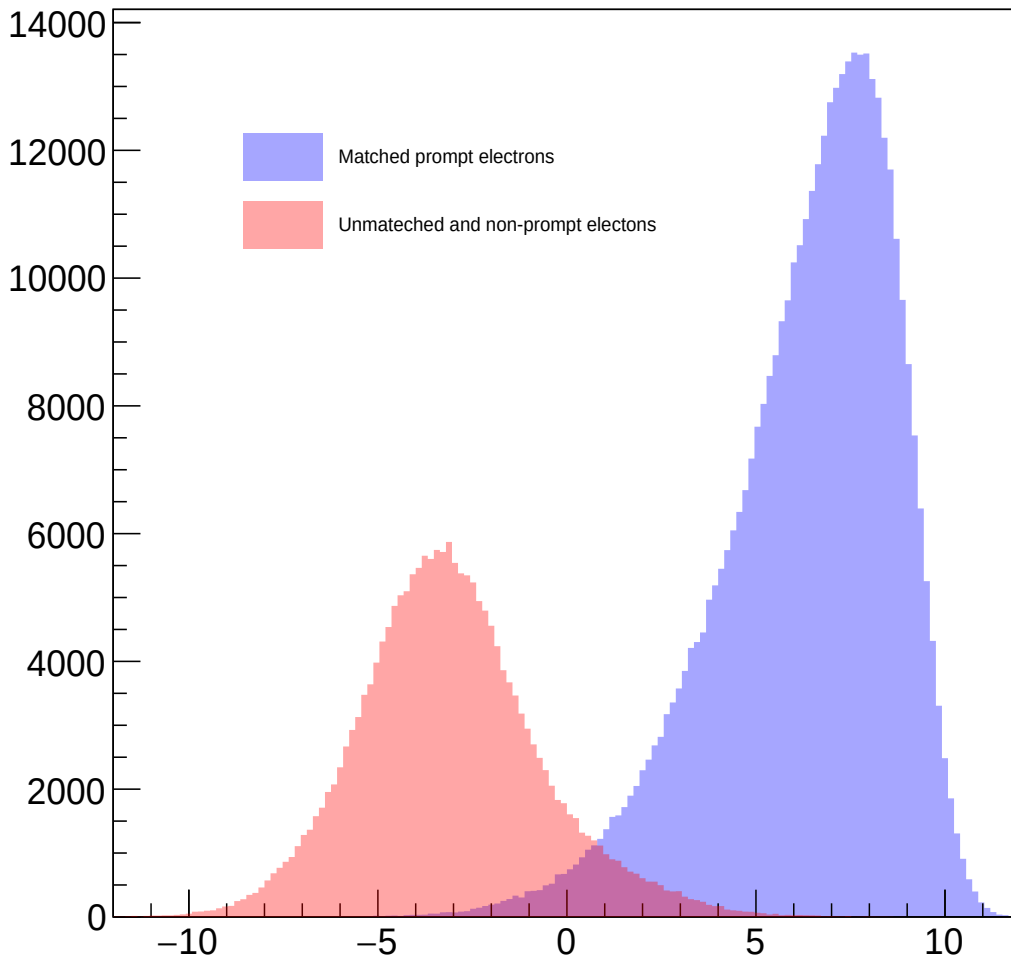


Figure 5.2. . The Output of the multiclassifier discriminant for prompt electrons matched to truth electrons from Z decay (blue) and for misidentified electrons (red). Events are all taken from Drell-Yan with jets MC sample. This plot is adapted from Ref. [162]

Electron kinematic selection criteria

In our analysis, electrons $e^\pm$ are required to have a transverse momentum $p_T^{e^\pm} > 7\text{GeV}$ and a pseudorapidity of $|\eta_{e^\pm}| < 2.5$, in order to reduce the contribution of the QCD background while preserving the highest possible signal efficiency. They may satisfy a loose primary vertex (PV) constraint, based on the impact parameter, in order to identify the electrons coming from the primary vertex or from a secondary vertex. The impact parameter cuts are defined as following:

$$|d_{xy}| < 0.5\text{cm} \quad \text{and} \quad |dz| < 1\text{cm} \quad (5.1)$$

Electrons passing the cuts listed above are referred to as "loose" electrons. The data-MC discrepancy is corrected using scale factors as it is done for the electron selection with efficiencies measured in data using the same tag-and-probe technique [138].

In order to ensure that the leptons are consistent with a common primary vertex (PV), we require that they have an associated track with a small impact parameter with respect to the event primary vertex. We use the three-dimensional significance of the impact parameter to the event vertex, defined as:

Observable type	Observable name
Cluster shape	RMS of the energy-crystal number spectrum along η and ϕ ; $\sigma_{i\eta i\eta}$, $\sigma_{i\phi i\phi}$ super cluster width
	along η and ϕ ratio of the hadronic energy behind the electron supercluster to the supercluster
	energy, H/E circularity $\frac{(E_{5\times 5} - E_{5\times 1})}{E_{5\times 5}}$ sum of the seed and adjacent crystal over
	the super cluster energy R_9 for endcap traing bins: energy fraction in pre-shower $\frac{E_{PS}}{E_{raw}}$
Track-cluster matching	energy-momentum agreement $\frac{E_{tot}}{p_{in}}$, $\frac{E_{ele}}{p_{out}}$, $\frac{1}{E_{tot}} - \frac{1}{p_{in}}$
	position matching $\Delta\eta_{in}$, $\Delta\phi_{in}$, $\Delta\eta_{seed}$
Tracking	fractional momentum loss $f_{brem} = 1 - \frac{p_{out}}{p_{in}}$ number of hits of the KF and GSF
	track N_{KF} , N_{GSF} reduced χ^2 of the KF and GSF track χ^2_{KF} , χ^2_{GSF} number
	of expected but missing inner hits probability transform of conversion vertex fit χ^2
Isolation	particle-flow photon isolation sum
	particle-flow charged hadrons isolation sum
	particle-flow neutral hadrons isolation sum
For PU-resilience	mean energy density in the event: ρ

Table 5.2. . Overview of input variables to the identification classifier [128].

$$SIP_{3D} = \frac{|IP_{3D}|}{\sigma_{IP_{3D}}} < 4.0 \quad (5.2)$$

where IP_{3D} is the lepton impact parameter in three dimensions at the point of closest approach with respect to the primary interaction vertex, and $\sigma_{IP_{3D}}$ is the associated uncertainty. Therefore, a "primary lepton" is a lepton satisfying $|SIP_{3D}| < 4$.

Electron isolation

The isolation variable is performed in order to identify the real electron from fake electron (mostly coming from hadronic jet) by requiring that a real electron is not surrounded by other activity within a cone. The PF isolation computes the contribution of all charged and neutral particles energy surrounding the electron in a cone of $\Delta R < 0.3$ and compare their sum to the electron's energy as following:

$$I_{iso}^{PF} = \frac{\sum p_T^{charged} + \sum E_T^\gamma + \sum E_T^{neutral}}{p_T^l} \quad (5.3)$$

where $p_T^{charged}$ is the charged hadron transverse momentum, E_T^γ and $E_T^{neutral}$ are the photon and neutral hadron transverse energies respectively and p_T^l represents the transverse momentum of the lepton. The

isolation variable is sensitive to the pileup since the pileup events give rise to extra energy deposits that affect this quantity. The effect of pile-up in isolation variable can be decreased by associating the charged particle candidates to the primary vertex while discarding the charged particle candidates associated with the other vertices. The assignment of charged particles to a vertex is more reliable than for neutral particles (photons, neutral hadrons) which makes their contributions highly affected by the amount of pileup in the event. So the neutral component of the isolation variable needs to be corrected to take into account the pileup effect:

$$I_{iso}^{PF} = \frac{\sum p_T^{charged} + \max [\sum E_T^{\gamma} + \sum E_T^{neutral} - p_T^{PU}]}{p_T^l} < 0.35 \quad (5.4)$$

The pileup contribution is defined as product of the geometric area of the isolation cone, called also as the effective area A_{eff} and the mean of the energy density distribution ρ of the for the neutral particles in the event:

$$p_T^{PU} = \rho A_{eff} \quad (5.5)$$

5.3 MUONS $\mu^\mp$

Muons are the only charged particles that can reach to the muon system leaving signature in the Muon Spectrometer chambers. The combination of information between the inner tracker and the muon system ensures reconstructed muon tracks with good resolution for momentum measurement and high muons selection efficiency [134]. The muon reconstruction [134, 154] chain starts with the local reconstruction of the muon tracks independently in the inner tracking (tracker track) and in the muon systems (standalone muon track) as follow:

- Stand-alone tracks: Those are muons reconstructed using the information from muon system only. The hits from DT and CSC chambers are clustered to form track segment "track stubs" then the segments are matched together to form seeds. Then by matching the seeds with the RPC hits to form the muon track a standalone muon track is built.
- Tracker-tracks: Those are muons reconstructed using the informations from inner tracker only. The hits in the silicon tracker are clustered to form a tracker-track called as inner track or a silicon track.
- Global Muon Reconstruction (outside-in): For each standalone muon track, a search for matching tracker track reconstructed in the inner tracking system is performed, and the best-matching tracker track is selected. For each "standalone muon" - "tracker track" pair, the track fit using all hits in both tracks is performed based on the Kalman filter technique. For muons with $p_T > 200$ GeV, the global muon reconstruction provides better momentum resolution than tracker muon reconstruction.
- Tracker Muon Reconstruction (inside-out): In this approach, all tracker tracks with $p_T > 0.5$ GeV and total momentum $p > 2.5$ GeV are considered to be potential muon candidates by searching for at least one compatible segment in the muon stations. Tracker tracks identified as muons by this method are referred to as "tracker muons". This method has a better energy resolution for low momentum muons having $p_T < 5$ GeV, those muons do not always able to traverse the CMS detector up to the muon system. This criteria used to define a tracker track as "tracker muon" are very loose and should not be used without further requirements.

Muon Identification

The CMS (Compact Muon Solenoid) detector, as its name reminds us, is optimized for muon identification and reconstruction, thanks to the Muon Spectrometer [134, 154, 162]. Therefore for the muons a multivariate classifier is not needed to be used, as it happen for the electrons. The muon has to be reconstructed as a "Global Muon" or "Tracker Muon", as described above. If a muon is classified as a "Standalone Muon Tracks", which are only reconstructed in the muon system, must be rejected. Below, we list a summary of the requirements need to identify and select the "proper" muons for our analysis:

Requirement	Description
Muon station matching	Muons are matched to segments in at least two muon stations "arbitrated tracker muon"
p_T measurement	$\frac{p_t}{\sigma_{p_t}} < 0.3$
Vertex compatibility (x-y)	$dxy < 0.2 \text{ cm}$
Vertex compatibility (z)	$dz < 0.5 \text{ cm}$
Pixel hits	At least one pixel hit
Tracker hits	Hits in at least six tracker layers

Table 5.3. . The requirements for a muon to pass the Tracker High- p_T Identification (ID). Note that these are equivalent to the Muon POG High- p_T ID with the global track requirements removed. [128].

Muon Reconstruction

In our analysis, we select those muons that satisfies the kinematic conditions:

$$p_T > 5\text{GeV}, \quad |\eta| < 2.4, \quad |dxy| < 0.5\text{cm}, \quad |dz| < 1\text{cm} \quad SIP_{3D} < 4$$

where dxy and dz are defined with respect to the primary vertex (PV). The $p_T^\mp$ -cut is slightly lower than for electrons because the low p_T reconstruction of muons is slightly more efficient, allowing to reach lower values of momentum.

Muon Isolation

Muons are required to be isolated similarly as described for electrons in section [5.2]. The pileup contribution subtraction is performed in a different way for muons, where a $\Delta\beta$ correction is implemented in form shown below:

$$\Delta\beta = \frac{1}{2} \left(\sum_{PU}^{\text{charged had.}} p_T \right) \quad (5.6)$$

where the summation is over all the transverse momentum of the PF charged hadrons not originating from the primary vertex and gives an estimate of the energy deposit from neutral particles coming from pileup vertices, which is only of the charged particles contribution associated to pileup. The relative muon isolation is then defined as:

$$I_{iso}^{PF} = \frac{\sum^{\text{charged hadron}} p_T + \max \left[\sum^{\text{neutral had.}} E_T + \sum E_T^Y - \Delta\beta, 0 \right]}{p_T^{\text{lepton}}} \quad (5.7)$$

The isolation cut for the muons has the same value of that used for the electrons:

$$I_{iso}^{PF} < 0.35 \quad (5.8)$$

5.4 PHOTONS γ

Leptons from coming from the Z -bosons decays can radiate a high-energy photon due to final state radiation (FSR). This can cause accuracy loss in the information extracted from the $4l$ finale state selection of the first Higgs H boson in our analysis. In order to fully reconstruct the Higgs boson decay system, we need to identify and collect such FSR photons, associating them to their parent leptons. The FSR recovery algorithms [140, 144, 162] is designed to discriminate FSR photons from the background, due to pileup interactions or initial state radiation. This is done exploiting the particular kinematics of FSR photons, that are irradiated in a direction collinear to the one of their parent lepton, and tend to be isolated from other particles. The selection of FSR photons is performed per-lepton and does not depend on any Z -mass criteria, which significantly simplifies the building and selection of ZZ candidates.

In the association of photons and leptons the rectangular cuts on $\Delta R(\gamma, l)$ and E_T, γ have been replaced by a cut on $\Delta R(\gamma, l)/E_T^2, \gamma$. Starting from the collection of photons provided by the PF algorithm (PF photons), the selection of photons and their association to a lepton proceeds as follows:

1. The pre-selection of PF photons is done by requiring $p_T^\gamma > 2\text{GeV}$, and relative isolation of the PF photon candidates:

$$I_{iso}^{PF\gamma} = \frac{\sum p_T^\gamma + \sum p_T^{neutral} + \sum p_T^{charged}}{p_T^\gamma} < 1.8 \quad (5.9)$$

where $p_T^\gamma, \sum p_T^{neutral}, p_T^{charged}$ are the scalar sums of the transverse momenta of photons, neutral and charged hadrons contained inside a cone of radius $\Delta R = 0.3$. The I_{iso}^{PF} - computation uses a threshold of 0.2GeV on charged hadrons, with a veto of cone of $\Delta R > 0.0001$, and 0.5GeV on neutral hadrons and photons, with a veto cone of $\Delta R > 0.01$, also including the contribution from pileup vertices (with the same radius and threshold as per-charged isolation);

2. Supercluster veto: We remove all PF photons that match with any electron passing both the loose ID and SIP cuts. The matching is performed by directly associating the two PF candidates;
3. Photons are associated to the closest lepton in the event among all those that pass both the loose ID and SIP cuts;
4. Photons that do not satisfy $\Delta R(\gamma, l)/E_T^2, \gamma < 0.012$ and $\Delta R(\gamma, l) < 0.5$ are discarded;
5. If more than one photon is associated to the same lepton, we select the one with lowest $\Delta R(\gamma, l)/E_T^2, \gamma$
6. Each selected FSR photon is removed from the isolation sum of all the leptons in the event that pass both the loose ID and SIP cuts. This concerns the photons that are in the isolation cone and outside the isolation veto of said leptons ($\Delta R < 0.4$ AND $\Delta R > 0.01$ for muons, and $\Delta R < 0.4$ AND ($\eta_{SC} < 1.479$ OR $\Delta R > 0.08$) for electrons).

5.5 JETS

Jets are the product of the hadronization of quarks and gluons and their reconstruction is enhanced with the anti- k_T clusering algorithm [23]. This algorithm iteratively combines PF candidates close to each other, clustering them around the hardest particle in the event in order to produce jets of an approximate conic shape. The size of the jet cone is defined by the distance parameter R , at which the algorithm is operated. Jets considered in this analysis have $\Delta R = 0.4$. The anti- k_T algorithm is resilient against infrared and collinear effects. The jet momentum is computed as the vectorial sum of all particle momenta in the jet, and a set of

corrections are applied to calibrate the jet response using the information of generated particles in simulations.

In order to ensure a good reconstruction efficiency and to reduce the instrumental background as well as the contamination from pileup, tight identification criteria based on the multiplicities and energy fractions carried by charged and neutral hadrons are imposed on jets [151]. The kinetics requirements for the jets in our analysis are as following:

$$p_T^{jets} > 20\text{GeV} \quad \text{and} \quad |\eta| < 2.4$$

In order to avoid lepton-jet contamination [88], jets are also required to be separated from identified leptons with an angular separation of:

$$\Delta R(jet, l) > 0.3$$

Jet energy corrections are applied to account for the effects of pileup, uniformity of the detector response, and residual differences between the jet energy scale in the data and in the simulation. The jet energy scale calibration [133] relies on corrections parameterized in terms of the jet p_T , η . For the 2016 [141] and 2017 [145] data taking years, the ECAL gradual timing shift was not propagated to the L1 trigger primitives properly and this resulted in an incorrect association of a large fraction of high- η trigger primitives to the previous bunch crossing. Since the L1 trigger is protected against firing over two consecutive bunch crossings, this resulted in events that self vetoed if a large deposit of ECAL energy was found in the pseudorapidity region:

$$2 < |\eta| < 3$$

Since these large energy deposits are left by jets with a large electromagnetic component, and since this effect is not described in the MC simulation, L1 pre-firing event weights are computed to correct for this effect. These corrections are obtained considering the contributions of all the jets in the event and are applied to correct the MC simulations for the 2016 and 2017 years. In order to select the $4l b \bar{b}$ final state, it is necessary to identify b-tagged quark jets. For this purpose, the DeepCSV algorithm is employed as b tagging algorithm. It combines impact parameter significance, secondary vertex and jet kinematics information to provide a final output discriminator whose score is computed with a Deep Neural Network technique. Data to simulation b tagging scale factors are applied as a function of the jet p_T , η and flavour to correct possible discrepancies of b tagging efficiencies in data and MC. They are applied as event weights and they are computed following the formula [152]:

$$SF_{weight} = \prod_i^{N_{jets}} SF(D_i, p_{T_i}, \eta_i, h_{flavour}) \quad (5.10)$$

where D_i is the b-tagging discriminant value and $h_{flavour}$ is the hadron flavour of the considered jet. The scale factors are applied and computed separately for all three years.

5.6 Analysis Strategy

In this thesis, the analysis strategy will be performed in the same way for both variants: resonant and non-resonant di-Higgs (HH) production. The two Higgs bosons coming from the ggF generation are reconstructed from a leptonic final state of 4 leptons, where a pair of Z^0 -bosons will be their source: $ZZ \rightarrow 4l$, where $4l = (4e, 4\mu \text{ or } 2e2\mu)$. From the other side, the second Higgs boson will be reconstructed from two b-tagged jets. The branching ratios $\mathcal{BR}$ of this combination is as following:

1. $H \rightarrow ZZ \rightarrow 4l(l = e^\pm, \mu^\pm)$: The branching ratio for the ggF-H into a leptonic final state is $\mathcal{BR}(ggF - H \rightarrow 4l(e, \mu)) = 1.251 \times 10^{-4}$ at $m_H = 125.09\text{GeV}$ [167];
2. $H \rightarrow b\bar{b}$: The branching ratio the ggF-H into a pair of b-tagged jets final state is $\mathcal{BR}(ggF - H \rightarrow b\bar{b}) = 5.809 \times 10^{-1}$ at $m_H = 125.09\text{GeV}$ [167]

The total cross section of the $ggF \rightarrow HH \rightarrow b\bar{b}ZZ(4l)$ will be calculated as following:

$$\sigma_{13\text{TeV}}^{HH} \{HH \rightarrow b\bar{b}ZZ(4l)\} [fb] = 2 \cdot \sigma_{13\text{TeV}}^{HH}(HH) \cdot \mathcal{BR}(H \rightarrow 4l(e, \mu)) \cdot \mathcal{BR}(H \rightarrow b\bar{b}) \quad (5.11)$$

The analysis is divided into several steps until we can constrain into the phase space, where we can perform the final HH reconstruction:

1. HLT Trigger selection;
2. Vertex selection;
3. Physics objects selection ($e^\pm, \mu^\pm$ and jets);
4. $Z(2l)$ candidates selection;
5. Z_1, Z_2 candidates selection;
6. $H \rightarrow ZZ \rightarrow 4l$ candidates selection: First Higgs boson;
7. $H \rightarrow b\bar{b}$ candidates selection: Second Higgs boson;
8. Resonant and non-resonant double Higgs HH selection.

The steps listed above will be explained in details in the sub-sections below.

5.6.1 Building of the $H \rightarrow ZZ \rightarrow 4l$ System

The reconstruction and identification of the electrons and muons was explained in the sections [5.2] and [5.3] respectively. Such leptons have $SIP_{3D} < 4$ as a vertex constraint and isolation cuts, where the final state radiation FSR photons (section[5.4]) are removed from the isolation cone. A lepton cross cleaning, which discards electrons with $\Delta R \leq 0.05$ from tight muons, is also applied. The building of a ZZ system from four candidate leptons proceeds according to the following sequential steps:

1) Z candidates: They are built from a pair of selected leptons of opposite charge and same flavor (e^+e^- , $\mu^+\mu^-$), having invariant mass satisfying $12 < m_{ll}(\gamma) < 120\text{GeV}$, where the Z candidate mass takes into account a selected FSR photon (if any) as we can see in Figure 5.3;

2) ZZ candidates: They are built from a pair of Z -candidates which do not have common leptons (non-overlapping). The Z -boson with closest mass $m_{ll(\gamma)}$ to the nominal Z -boson mass, which is $m_{Z^0} = 91,1876 \pm 0,0021\text{GeV}/c^2$, is denoted as Z_1 (on-mass shell) and the second one is the Z_2 , which is reconstructed from the remaining two leptons. In case there more than two remaining lepton, which pass all the conditions, then the Z_2 will be reconstructed from the the two leptons, which the sum of their p_T^l has the highest value. The building of the ZZ system must satisfy the following requirements:

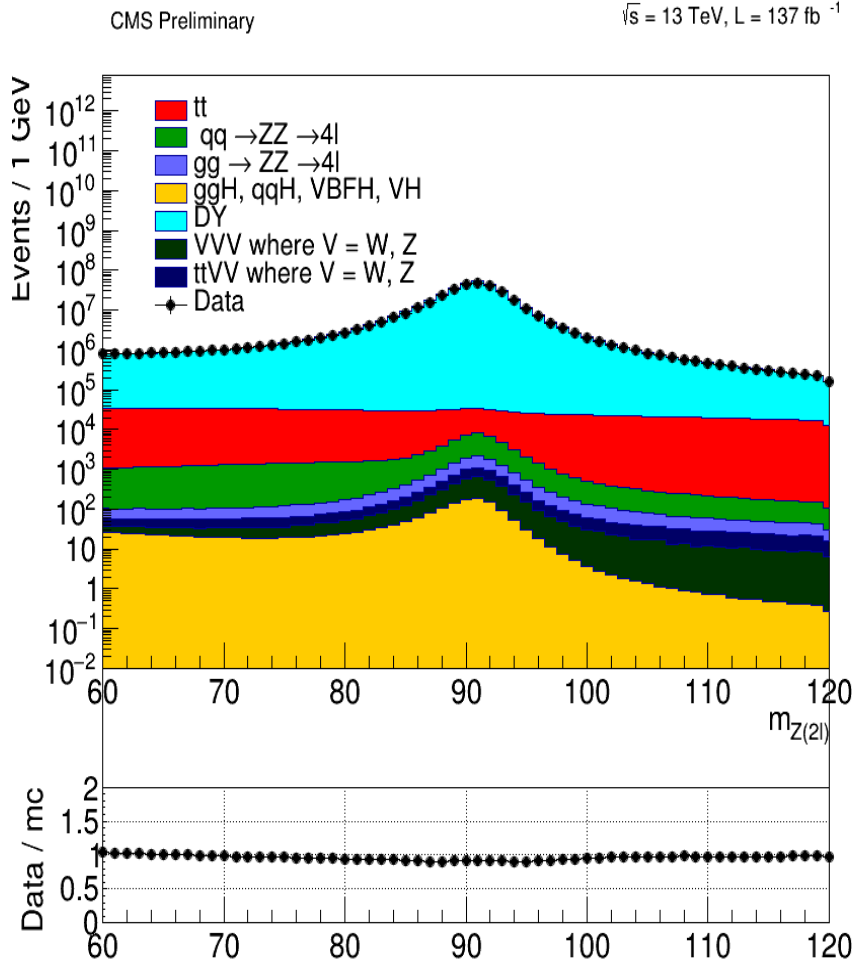


Figure 5.3. . Distribution of leptonic Z^0 -boson invariant mass m_{Z^0}

- Ghost removal: Leptons must have an angular separation of $\Delta R(\eta, \phi) > 0.02$ between them, in order to be identified as separated objects;
- Lepton p_T : at least two out of the four leptons must have $p_T^i > 10 \text{ GeV}/c$ and $p_T^i > 20 \text{ GeV}/c$;
- QCD suppression: All opposite-sign lepton pair that can be built out of the four leptons (regardless of lepton flavor) must satisfy $m_{ll} > 4 \text{ GeV}$. Here, the selected FSR photons are not included in the mass computation, since QCD-induced low mass di-leptons (e.g. J/Ψ) may have photons nearby (e.g. from π^0);
- Z_1 selection: We select the first Z -boson with the condition $m_{Z_1} > 40 \text{ GeV}$;
- Alternative pairing check (Smart cut): Z_a and Z_b are defined as the mass-sorted alternative pairing Z -candidates, where Z_a is defined to be the closest one to the nominal Z -boson mass. The ZZ candidate is not selected, in case the invariant mass of $m_{Z_b} < 12 \text{ GeV}/c^2$ and the m_{Z_a} is closer to the nominal Z -boson invariant mass than the Z_1 candidate. Selected FSR photons are included in the m_Z 's computation. In other words, this cut discards 4μ and $4e$ candidates which have an alternative pairing similar to an on-shell Z boson together with a low-mass pair l^+l^- ;
- Four-lepton invariant mass: $m_{4l} > 70 \text{ GeV}/c^2$;

- Choice of the best ZZ candidate: If there are more than one ZZ -candidate, we will select only those candidates with the highest four leptons p_T scalar sum.

The two reconstructed Z^0 bosons Z_1, Z_2 can be seen in Figure 5.4.

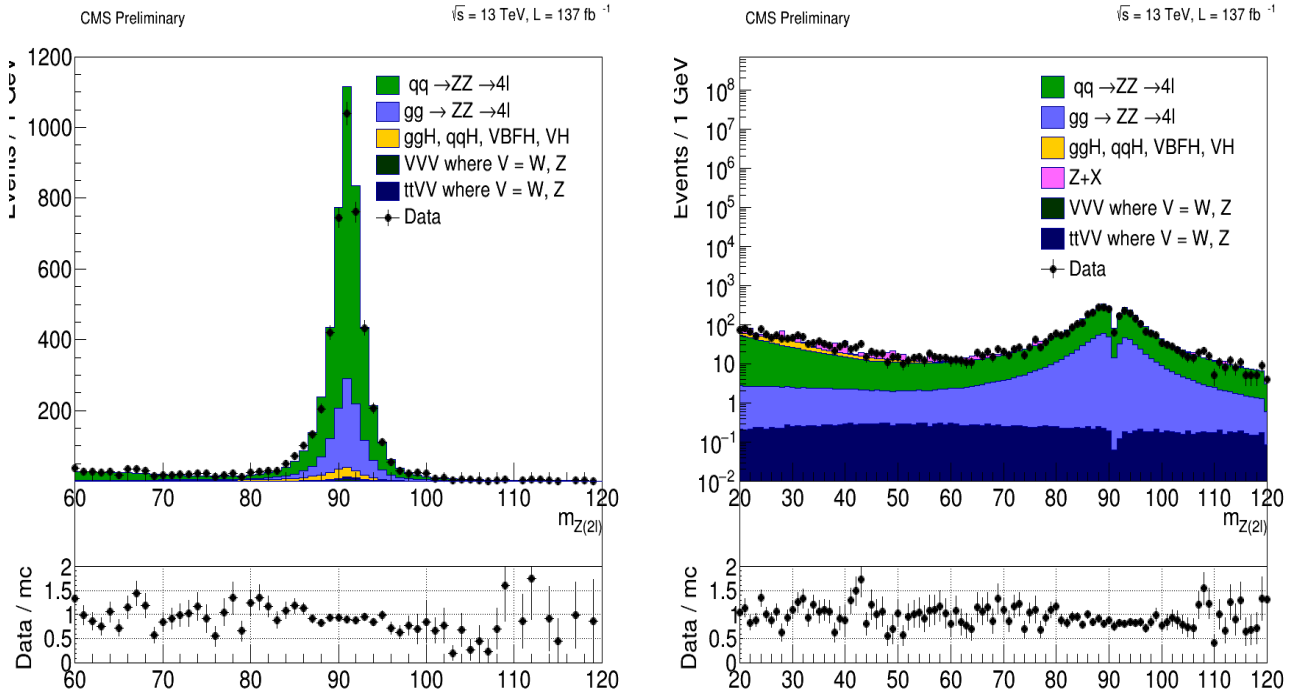


Figure 5.4. . Distribution of leptonic nominal Z_1^0 -boson invariant mass $m_{Z_1^0}$ (left) and off-shell ZW^0 -boson invariant mass m_{ZW^0} (right) with full RunII data 137fb^{-1} at $\sqrt{s} = 13\text{TeV}$

3) SM Higgs $4l$ -selection: events containing at least one ZZ system satisfying all the previous selections are then used to tag the Higgs decaying into four lepton final state, as it is done for the SM Higgs analysis [36], which we can see in Figure 5.5;

5.6.2 Building the $H \rightarrow b\bar{b}$ system

In order to reconstruct the $H \rightarrow b\bar{b}$, we need to select all the events with at least two jets or more, otherwise all the other events with less than 2 jets will be discarded. The pairs of jets candidates ($b\bar{b}$ candidates) are built from the jets that pass all criteria described in section [5.5] with $|\eta| < 2.4$ and $p_T > 20\text{GeV}/c$ in order to reduce the background coming from the vector boson fusion Higgs boson (VBF-H) production mode. A cross cleaning check is done with all the jets passing p_T, η condition, where each jet must have an angular separation of $\Delta R(\text{jet}, \text{lepton}/\gamma_{\text{FSR}}) > 0.3$ toward all the electrons, muons or FSR photons, that "survived" the cuts till now. If there are at least two jets in the event, the $b\bar{b}$ candidate is formed with these two jets. If there are more than two jets in the event, the $b\bar{b}$ candidate is formed taking the two jets with the highest b -tagger score in the event.

A study was performed on the non-resonant SM HH -signal Monte Carlo sample, in order to verify the most efficient way of the jets selection. A cone of radius $\Delta R = 0.4$ is built around each RECO (see section [4.2]) jet selected, where we search for a b -jet at the MC generation level (GEN, (see section [4.2])), coming from the decay of the Higgs boson (within the Pythia8-HADRONIZER, check section [4.2]), and we check if the

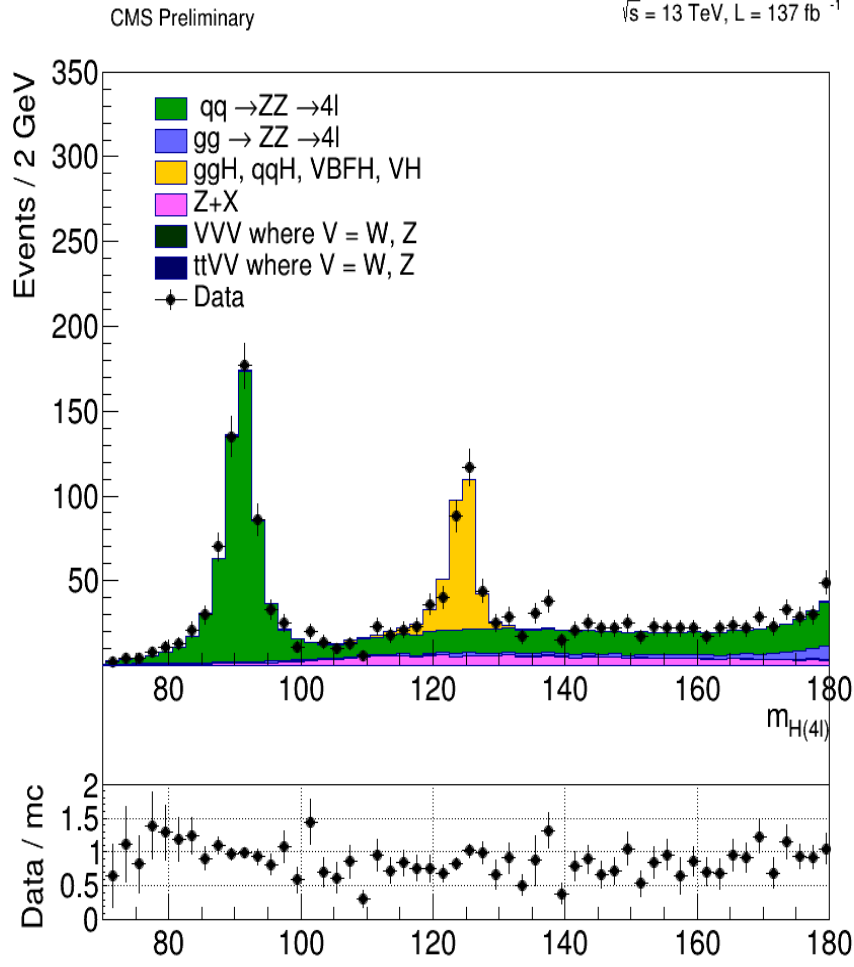


Figure 5.5. Distribution of the invariant mass $m_{H(4l)}$ of the leptonic Higgs H boson with full RunII data 137fb^{-1} at $\sqrt{s} = 13\text{TeV}$

considered RECO jet is matching with the GEN jet, inside the cone. Three different selection methods of the di-jet candidate are considered and has been compared, as we can see in Figure 5.6:

1. Selection of the event with two highest p_T jets;
2. Selection of the event with two highest b -tagger score jets;
3. Selection of the event with highest b -tagger score jet and the highest p_T jet (verifying that they are not the same jet).

As we can see in the Figure 5.6 above, the results of this study has shown that we get the best efficiency on the selection of the di-jet candidate with the condition of the two highest b -tagger score jets (left plot), which propagates to the $m_{H \rightarrow b\bar{b}}$ (right plot), where more statistics is accumulated around the Higgs boson invariant mass and less in tails (red line), in comparison to the other two selections modes.

In Figure 5.7, we have the first observation of the $H \rightarrow b\bar{b}$ in association with a W or Z boson (VH) [153].

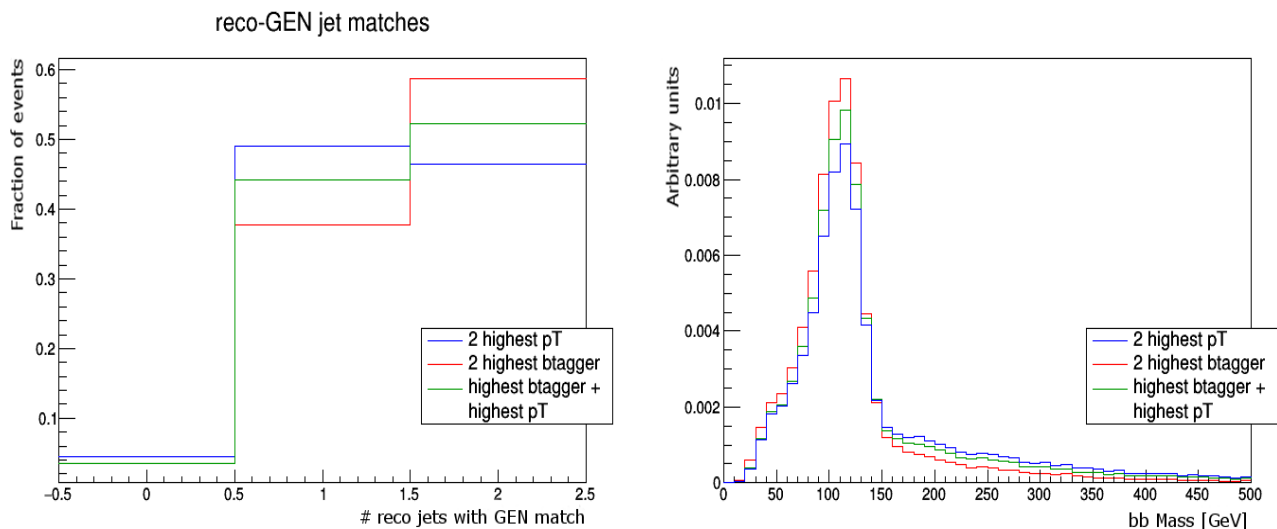


Figure 5.6. . Di-jet invariant mass built from RECO jets selected with the three different methods considered: the two highest p_T jets (blue line), the highest b-tagger score jet and the highest p_T jet (green line), the two highest b tagger score jets in the event (red line).

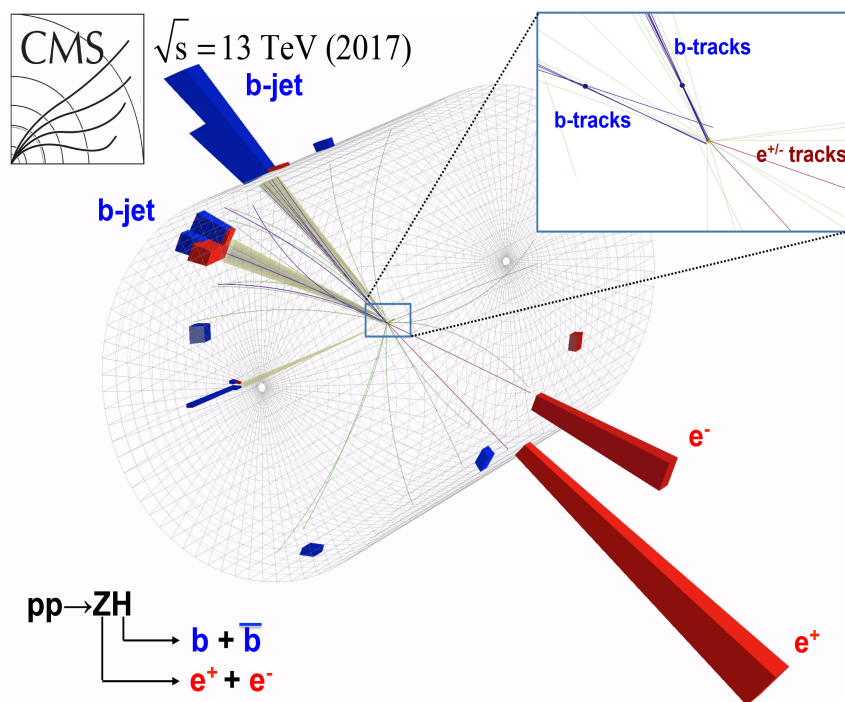


Figure 5.7. . Schematic view of the observation of b-tagged jets (in blue) coming from the decay of the Higgs boson, in association with a e^+e^- track (in red) coming from a Z boson (VH)[153]

5.6.3 Signal region

Our phase space will be then build from all events containing one selected ZZ candidate and one $b\bar{b}$ candidate, passing the final requirement:

$$115\text{GeV}/c^2 < m_{H(4l)} < 135\text{GeV}/c^2$$

This final condition of selection of our $HH \rightarrow b\bar{b}ZZ(4l)$ candidate, is defined as the signal region.

5.7 Background Estimation

In this analysis the backgrounds are composed by the three SM Higgs production modes, gluon fusion (ggH) and associated vector ($W^\pm H$, ZH and $t\bar{t}H$), and the SM backgrounds $q\bar{q} \rightarrow ZZ$, $gg \rightarrow ZZ$ and $Z + X$ (which will be explained in the subsections below).

5.7.1 Irreducible Backgrounds

Estimation of the SM Higgs Production Modes

The single SM Higgs boson process is the main background source for this analysis, for both of the two Higgs boson final states studied in this thesis. Among the different single Higgs boson production modes, the gluon fusion $ggF - H$ and the vector boson fusion $VBF - H$ mechanisms represent the main sources of background due to the relatively large cross section and due to the presence of two jets in the final state (especially in the $VBF - H$ mode). The description of these single SM Higgs boson processes is obtained using the POWHEG-BOX-V2 generator. The SM Higgs production modes are estimated from MC by normalizing the yields according to the cross sections times the branching ratio fractions ($\sigma \times BR$) as computed by the LHC Higgs Cross Section Working Group [168].

ZZ^* production via quark-antiquark annihilation $q\bar{q} \rightarrow ZZ$ Background

The ZZ^* production via quark-antiquark annihilation $q\bar{q} \rightarrow ZZ$ background events are generated at NLO and it is an other important background to be estimated, since it contains four leptons in the final state of this background samples. The fully differential cross sections, which is not yet available in a partonic level event generator, is computed at NNLO. NNLO/NLO k-factors are then applied to the $q\bar{q} \rightarrow ZZ$ NLO POWHEG sample. The inclusive cross sections obtained using the same PDF, as well the renormalization and factorization scales, as the POWHEG LO, NLO and NNLO are used. The NNLO/NLO k-factors are applied as function of m_{ZZ} [161]. Additional NLO electroweak corrections, which depend on the initial state quark flavor and kinematics, are also applied in the mass range of $m_{ZZ} > 2m_Z$, where the corrections have been computed [161].

ZZ^* production via ggF mode $gg \rightarrow ZZ$ Background

The ZZ^* production via ggF mode $gg \rightarrow ZZ$ background is an other important background, where again we have to deal with four leptons in the final states with a non distinguishable signature with respect to the single SM Higgs boson production, and jets passing the selection (section [5.6]) can arise from the hadronization of initial state gluons. This background is estimated from the MC simulation, where k-factors are applied to the MC samples as a function of $m(ZZ^*)$ to account for the different order between the MC simulations and theoretical calculations for this process. This sample is generated at LO with the generator MCFM 7.0 [24]. This background does not have exact calculation beyond LO.

VVV and ttV backgrounds

Other background sources for the analysis are represented by tri-boson production with at least one Z boson with leptonic decays, and by $t\bar{t}Z^0$ and $t\bar{t}W^\mp$ production. These processes, in fact, can present four leptons and two jets in the final state, passing the event selection. Leptons can derive from the leptonic decays of the

$W^\mp$ or Z bosons, or from the leptonic decay of the $W^\mp$ boson deriving from the top quark decay, while jets can derive from hadronic decays of the same objects. In particular, b -jets are produced in almost every decay of the top quark. The contribution of these background processes is estimated from the MC simulation.

5.7.2 Reducible Backgrounds

Estimation of $Z + X$ Background

The $Z + X$ background is a reducible background in the $H \rightarrow ZZ \rightarrow 4l$ final [161] and originates from processes which contain one or more non-prompt leptons in the four-lepton final state. Non-prompt leptons are mainly not isolated leptons which can originate from the heavy-flavor mesons decay, mis-reconstructed jets (usually coming from light-flavor quarks) and electrons from γ conversions. Such non-prompt leptons are usually called "fake" leptons. The estimation of such background in the $H \rightarrow ZZ \rightarrow 4l$ analysis is done by measuring the $f_e(f_\mu)$ probability of fake electrons (fake muons) passing the loose selection criteria (see section [5.2,5.3]) to also pass the final selection criteria (the tight selection). These probabilities are called as fake ratios or fake rates and are applied into two defined control regions (CRs) in order to extract the expected $Z + X$ yield in the signal region. In the following sections, we will describe the steps needed for the estimation of $Z + X$ yields and shape.

Measuring the Fake Rates

Since in this analysis we require to have at least 2 isolated jets, we estimate the fake rate in $Z(ll) + e+$ at least 2 isolated jets and $Z(ll) + \mu+$ at least 2 isolated jets events that are expected to be completely dominated by final states which include a Z boson, a fake lepton and at least 2 jets isolated with respect to the leptons. The leptons forming the Z must be opposite sign, same flavor and have $p_T > 20(10)$ GeV. The third lepton (e/μ) has to pass the loose selection and is used as the probe to compute the f_e, f_μ probabilities. In addition, we require at least two isolated jets. The invariant mass given by the fake lepton and the opposite-sign tight lepton composing the Z is required to be $m_{ll} > 4\text{GeV}$ to reduce QCD contamination. Also, in order to suppress contamination from γ conversions to electrons the invariant mass of the two leptons composing the Z must satisfy $|m_{Z(ll)} - m_Z| < 7\text{GeV}$. The remaining events are expected to also contain contribution from $W^\pm Z^0$ and $t\bar{t}$, which are reduced by requiring $E_T^{miss} < 25\text{GeV}$. The fake rate is parameterized in terms of the probe lepton p_T, η . Fig. 4 - Fig. 6 show the fake rate estimation separately for electrons and muons in terms of p_T, η . There are two FR distributions in Fig. 4 - Fig. 6. The points with solid error bars are the fake rates computed directly from data through the procedure described above. But, since $W^\pm Z^0$ background potentially contributes with three real leptons and such background is already included in the analysis via MC, it is then subtracted from data (separately in the numerator and in the denominator used to compute the FR's). The resulting FR after this subtraction is shown by the points with dotted error bars in Fig. 4 - Fig. 6. With the requirement of 2 isolated jets we force the probe lepton to be far away from jets, we discharge all the events with a fakeable object coming from the surroundings of the jets (mostly bjets misreconstructed as leptons); this lowering the fake rate probability especially at low p_T , where most of the bjets are misreconstructed as leptons. Now, as one can see, there is a significant dependence between the FR's and the (p_T, η) of the loose leptons. For this reason the final parameterization of the FR's is made as a function of these two variables and it is shown in the [Figure 5.8](#), [Figure 5.9](#) and [Figure 5.10](#).

Building Control Regions

In order to apply the fake rates described previously, two control regions are defined requiring events with two leptons passing the selection of the first Z (tight leptons) and additionally a pair of loose leptons of same

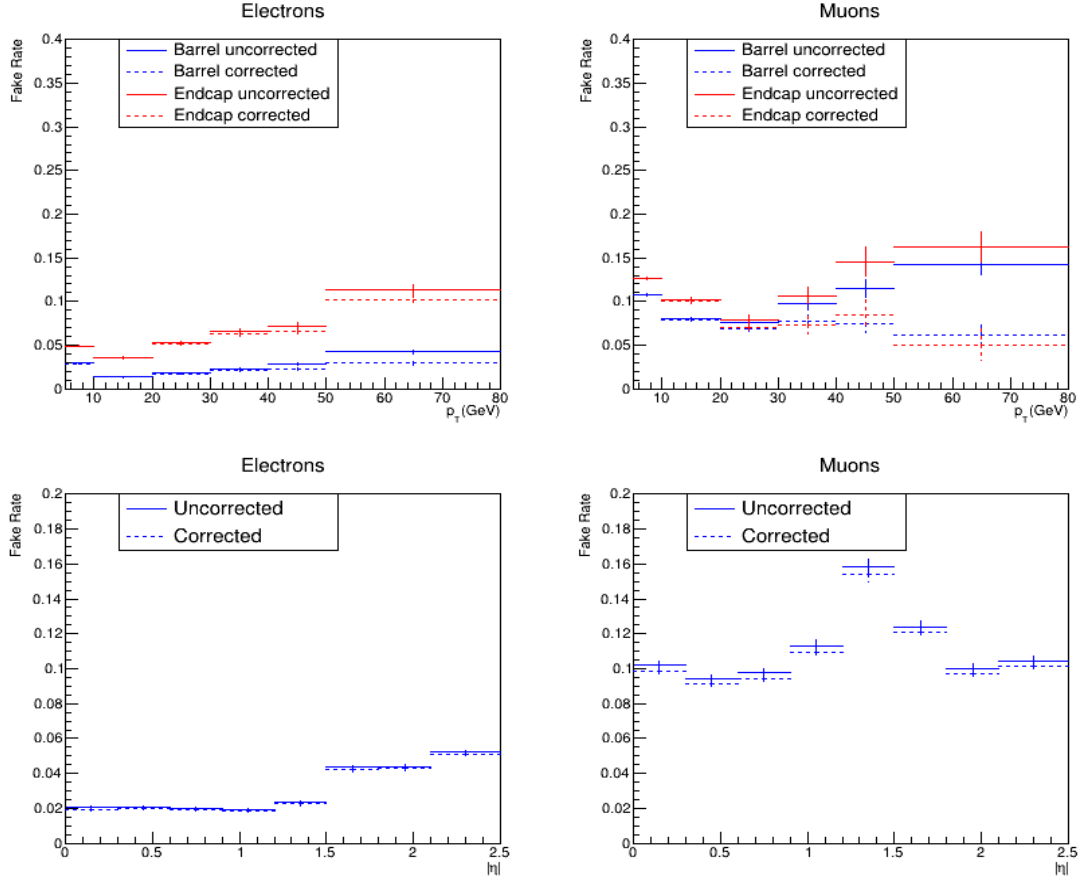


Figure 5.8. Fake rates 2016: Fake rates for 2016 dataset are shown in terms of the lepton probe p_T and η . The η distributions (barrel and endcap) corresponds to the regions defined by $|\eta| < 1.479$ (1.2) for electrons (muons). The background $W^\pm Z^0$ from MC has been subtracted.

flavor, opposite charge and passing the SIP_{3D} cut. These events also must satisfy all the kinematic cuts applied for the Higgs phase space selection (check section [5.6]). The first control region is then built by requiring that the two loose leptons do not pass the final identification and isolation criteria. This control sample is nominated as "2 Prompt + 2 Fail + 2 jets" (being referred as 2P2F from now on) and it is expected to be populated by events that have only two prompt leptons: mostly DY (Drell Yan) with a small fraction of $t\bar{t}$ and $Z\gamma$. The second control region is obtained by requiring that just one of the four leptons does not pass the final identification and isolation criteria. The remaining three leptons in the event should pass those selections. This control region is then nominated "3Prompt + 1Fail + 2jets" (being referred as 3P1F from now on). The events appearing in this control region are the same present in the 2P2F, with different relative proportions though, plus WZ events that are expected to have three prompt leptons. These two control regions, which are enriched by fake leptons, are orthogonal to the Higgs signal region and are used to estimate the yields and shape of the Z+X background in the signal region. The expected number of reducible background events in the 3P1F control region can be computed by weighting each event observed in the 2P2F control region, as shown below [161]:

$$N_{3P1F}^{bkg} = \sum \left(\frac{f_i}{1 - f_i} + \frac{f_j}{1 - f_j} \right) N_{2P2F} \quad (5.12)$$

which the f 's correspond to fake rates of each loose lepton. If the fake rates was measured in a sample that has exactly the same background composition as the $2P + 2F$ sample, the difference between the observed number of events in the $3P + 1F$ sample and the expected background predicted from the $2P + 2F$ sample would solely amount to the (small) $W^\pm Z^0$ and $Z\gamma$ contribution. Large differences arise because the fake rates

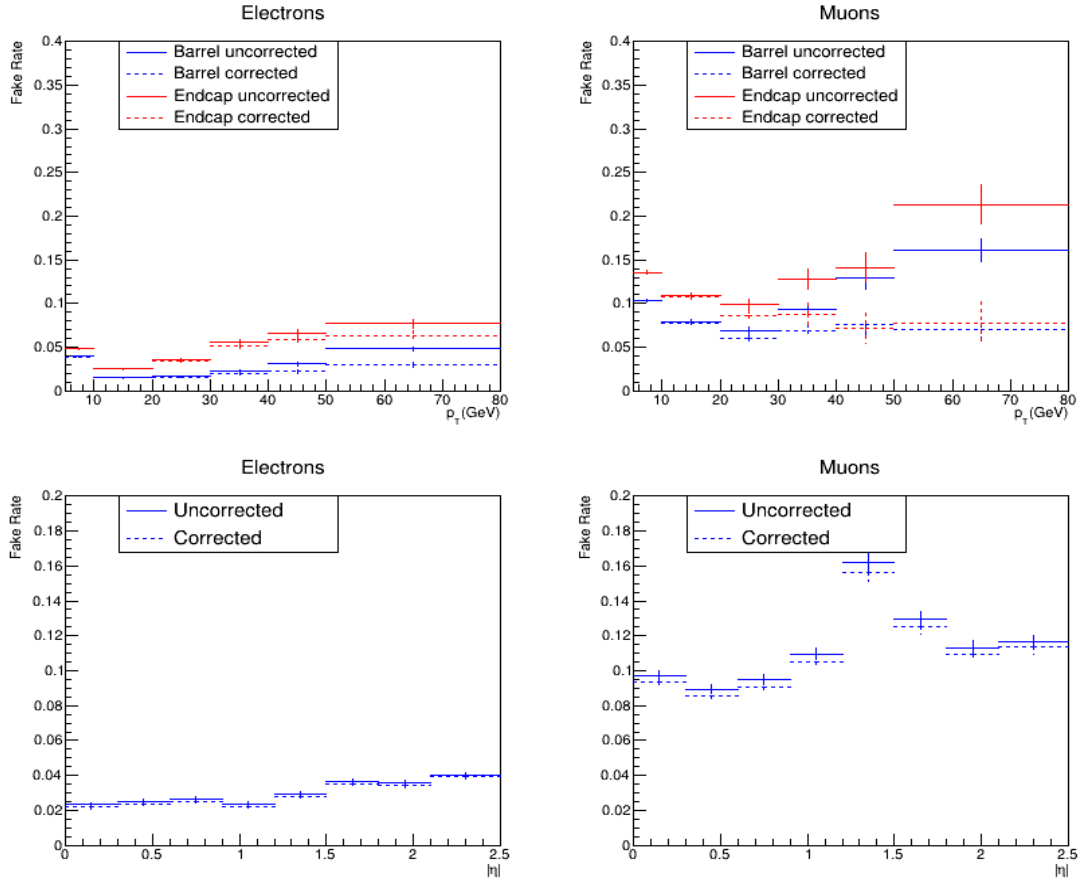


Figure 5.9. Fake rates 2017: Fake rates for 2017 dataset are shown in terms of the lepton probe p_T and η . The η distributions (barrel and endcap) corresponds to the regions defined by $|\eta| < 1.479$ (1.2) for electrons (muons). The background $W^\pm Z^0$ from MC has been subtracted.

used in Eq. 5.12 do not properly account for the background composition of the $2P + 2F$ control sample. The difference between the $3P + 1F$ observation and the prediction from $2P + 2F$ to recover the missing contribution from photon conversion. The expected reducible background in the signal region is given by the sum of two terms. In order to correct those discrepancies and correctly estimate the $Z + X$ background, one needs two components:

- a component directly from $2P2F$, computed by weighting each observed event in the $2P2F$ control region using the factor $\left(\frac{f_i}{1 - f_i} + \frac{f_j}{1 - f_j} \right)$, in which the f 's are the fake rates of loose leptons as explained before;
- a component from $3P1F$, obtained by the difference between the observed events in $3P1F(N_{3P1F})$, the expected contribution from $2P2F$ and from ZZ in the signal region $N_{3P1F}^{ZZ} + N_{3P1F}^{bkg}$. The ZZ contribution is taken from the MC simulated events selected in the $3P1F$ control region. The N_{3P1F}^{bkg} is the $3P1F$ extrapolation from $2P2F$ as discussed above.

The final expression for the estimation of $Z + X$ can be then symbolically represented as [161]:

$$N_{SR}^{bkg} = \sum \left(\frac{f_a^i}{1 - f_a^i} \right) (N_{3P1F} - N_{3P1F}^{bkg} - N_{3P1F}^{ZZ}) + \sum \left(\frac{f_b^i}{1 - f_b^i} + \frac{f_c^j}{1 - f_c^j} \right) N_{2P2F} \quad (5.13)$$

which is more conveniently expressed as [161]:

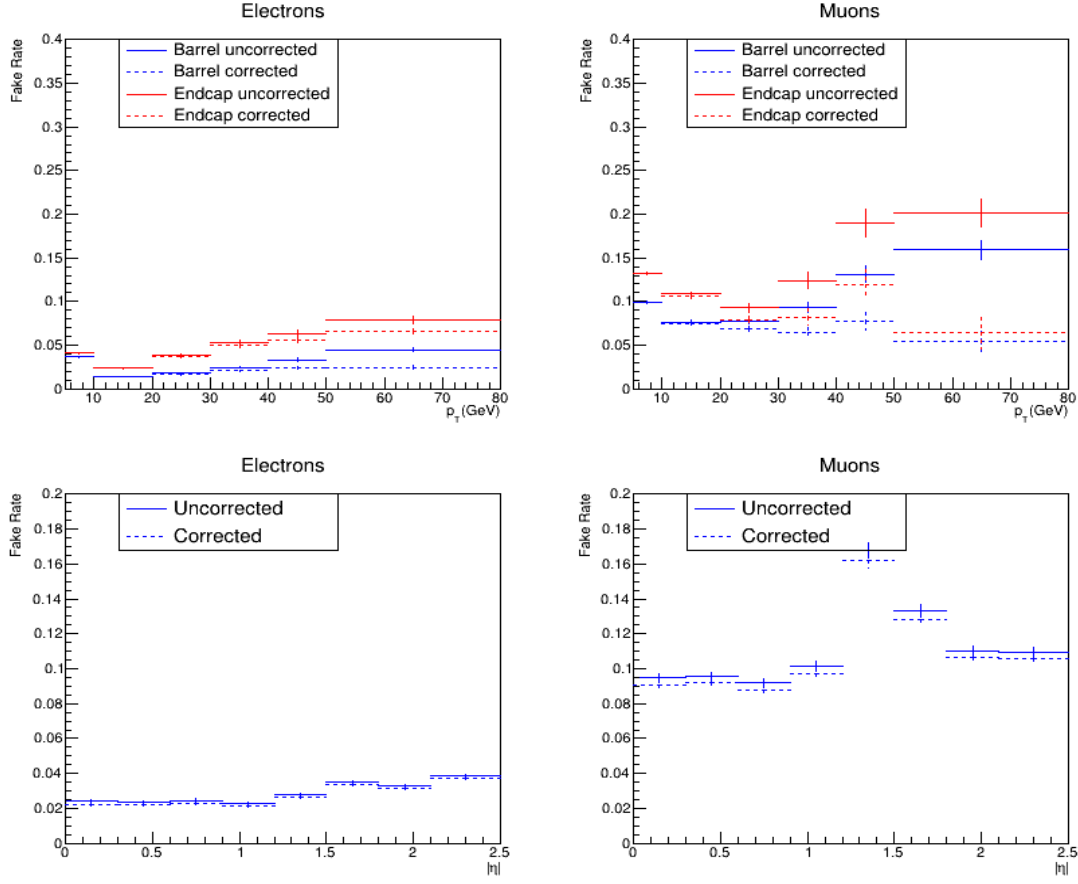


Figure 5.10. Fake rates 2018: Fake rates for 2018 dataset are shown in terms of the lepton probe p_T and η . The η distributions (barrel and endcap) corresponds to the regions defined by $|\eta| < 1.479$ (1.2) for electrons (muons). The background $W^\pm Z^0$ from MC has been subtracted.

$$N_{SR}^{bkg} = \left(1 - \frac{N_{3P1F}^{ZZ}}{N_{3P1F}}\right) \sum_i \frac{f_a^i}{1 - f_a^i} - \sum_j \left(\frac{f_b^j}{1 - f_b^j} + \frac{f_c^j}{1 - f_c^j} \right) \quad (5.14)$$

Equation 7 represents what has been done to estimate the $Z + X$ background contribution in this analysis. The $Z+X$ event yield in the signal region is obtained after applying the additional cuts of the selection; yields obtained are shown in Table 5.4.

	4μ	$4e$	$2e2\mu$	$2\mu 2e$
Z+X+2jets	1.85 ± 0.59 (stat only)	0.18 ± 0.07 (stat only)	1.90 ± 0.58 (stat only)	0.19 ± 0.09 (stat only)
Z+X+2jets	1.51 ± 0.52 (stat only)	0.07 ± 0.04 (stat only)	2.23 ± 0.65 (stat only)	0.19 ± 0.08 (stat only)
Z+X+2jets	3.01 ± 0.77 (stat only)	0.32 ± 0.10 (stat only)	2.07 ± 0.64 (stat only)	0.20 ± 0.08 (stat only)

Table 5.4. Reducible background yields for all the channels: 2016 (1st row), 2017 (2nd row) and 2018 (3rd row) datasets

Uncertainties on the Estimation of $Z + X$

The statistical uncertainty on the $Z + X$ estimation comes from the limited number of events selected in the control regions where one measures and applies the fake ratio methods. The systematic uncertainty arises because the composition of reducible backgrounds in the control regions where one measures and applies

the fake ratios are usually not the same. This is the main source of systematic uncertainty of the fake ratio method. This uncertainty is estimated in three steps. First, one measures the fake rates for each background process (DY , $t\bar{t}$, $W^\pm Z^0$, ZZ , $Z\gamma$) using MC and applying the selections explained above, to obtain samples of $Z + l$. Then, one defines the fake rates for MC (simulation) by computing the weighted average of those individual fake rates. Second, one reweights the individual fake rates using the expected yields of the reducible backgrounds selected in the $2P2F$ control region. Then, one applies the average and the re-weighted fake rates in order to get the final yields of $Z + X$ and finally, the difference between them is used as an estimation of the uncertainty on the measurement of fake rates. The systematic uncertainty on $Z + X$ has been taken from the $H \rightarrow ZZ \rightarrow 4l$ analysis and it's effect is propagated to the final estimates, and it amounts to about 30% – 41% [161].

	4μ	$4e$	$2e2\mu$
2016	30%	41%	35%
2017	30%	38%	33%
2018	30%	37%	33%

Table 5.5. Systematic uncertainties on the reducible background estimate for the 3 years, [161].

5.7.3 The Experimental Uncertainties

The first source of experimental uncertainties, which affect both signal and background, are the uncertainty on the integrated luminosity, equals to (from to 2.3% to 2.6%, depending on the data taking period) and the uncertainty on the lepton identification and reconstruction efficiency (variates from 1 to 2.5% and from 11 to 15.5% on the overall yields, in the 4μ , $4e$ channels, respectively). Experimental uncertainties in the fake rate method are the ones used to estimate the $Z + X$ background uncertainty, as described in Sec. 5.4 vary between 6% to 32%. The summary of experimental systematic uncertainty is reported in Table 5.6.

Source	2016	2017	2018
Luminosity	2.6%	2.3%	2.5%
Leptons ID and reco eff	1.6-15.5%	1.1-12.1%	1.0-11%
b tagging SF	shape	shape	shape
JES	shape	shape	shape
JER	shape	shape	shape
Z+X uncertainties	30-41%	30-38%	30-37%

Table 5.6. . Summary of the experimental systematic uncertainties for all the tree years [161].

5.7.4 The Theoretical Uncertainties

The theoretical uncertainties, affecting both signal and background estimation, include the renormalization and factorization scales and the choice of the Parton Density Function (PDF) set. The variation of these scales between 0.5 and 2 times their nominal value, while keeping their ratio between 0.5 and 2, provides the

uncertainty on the normalization and factorization scales. The uncertainty from the PDF set is determined by taking the root mean square of the variation when using different replicas of the default NNPDF set. An additional uncertainty of 10% on the k-factor used for the prediction of $gg \rightarrow ZZ$ is applied as described in Sec. 5.3. Associated to the $H \rightarrow ZZ \rightarrow 4lBR$ there is a 2% systematic uncertainty, which affects only the Higgs samples yields (VBF and the others). All experimental uncertainties are considered uncorrelated while all theoretical uncertainties are considered correlated among the three years. The summary of theory systematic uncertainty is reported in [Table 5.7](#), [Table 5.8](#).

Background MC	PDF + α_s unc.[%]
ggH	1.8%
VBFH	2.59-2.62%
ZH	2.1%
WH	1.6%
WH	1.3%
bbH	3.2%
ttH	3.6%
qqZZ	3.1-3.4%
ttW	2.5-3.75%
ttZ	7-14%
VVV	2-17%
ggZZ	3.2%

Background MC	QCD Scale unc. %
ggH	4.27-6.49%
VBFH	0.3-0.4%
ZH	2.7-3.5%
WH	0.5%
bbH	4.6-6.7%
ttH	6.0-9.2%
qqZZ	3.2-4.2%
ttW	3-4%
ttZ	2-3%
VVV	3%
ggZZ	4.6-6.7%

Electroweak corrections	Uncertainty
qqZZ	0.1%
ggZZ	10.0%

Table 5.7. . Theoretical uncertainties of the background samples [161].

SM NLO HH signal [168]			
SM HH	PDF+ α_s	Scale +/−	m_{top} unc.
HH	3.0%	2.2-5%	2.6 %

KK Bulk Graviton spin-2 [122]			
m_{Gr} [GeV]	PDF	α_s	Scale +/−
260	11.8%	11.3%	7.8-5.3%
300	12.7%	12.2%	6.6-4.5%
400	15.0%	15.0%	4.0-2.9%
500	17.3%	17.5%	3.1-2.6%
600	19.5%	19.8%	2.1-1.7%
700	22.2%	22.5%	2.7-1.9%
800	24.4%	24.6%	1.8-1.4%
900	27.2%	26.9%	1.9-1.5%
1000	29.6%	29.0%	1.9-1.5%

KK Bulk Radion spin-0 [122]			
m_{RD} [GeV]	PDF	Scale +/−	α_s
260	1.9%	11.4%	2.8%
300	2.1%	10.9%	3.2%
400	2.6%	10.3%	3.9%
500	3.1%	10.0%	4.3%
600	3.6%	9.8%	4.5%
700	4.1%	9.7%	4.8%
750	4.4%	9.6%	4.8%
800	4.6%	9.5%	4.9%
900	5.0%	9.5%	5.0%
1000	5.5%	9.5%	5.1%

Table 5.8. . Theoretical uncertainties of the signal samples

5.8 Multivariate Analysis

The goal of any analysis in particle physics is to enhance the signal while minimizing the background, in order to improve the signal to background ratio, exploiting some appropriate cuts on specific discriminating variables. The specific cut values depend upon comparison between the MC signal distribution and those for the MC background processes, which were listed in section [4.2.3]. The discriminant abilities of Multi-Variate techniques are exploited by combining some variables, which help us to reduce the complexity of the problem. The Toolkit for Multi-Variate Analysis (TMVA) is a framework within the ROOT-integrated environment, containing several multivariate classification techniques [129], like the Boosted Decision Trees (BDT), which we are going to explain below) and other techniques. The TMVA analysis consists of a training, testing and evaluation of classifiers using data samples with unknown signal and background composition. The chosen classifier computes optimal values of some parameters w (weights), which can be tuned, and the mapping function $y(\vec{x})$ in order to maximize Signal-Background separation. Moreover, the selected trained classifiers are then used to classify the data samples. Such an analysis is based on input variables, which are provided from the previous cut-based analysis, and will be used according to the method we want to deal with (e.g. the BDT).

5.8.1 The Boosted Decision Trees (BDT)

The Boosted Decision Tree (BDT) is a machine learning technique, where a decision (regression) tree is a binary tree structured into classifier (regressor). Yes or no decisions are taken on one single variable at a time until a stop criterion is fulfilled, as we can see in Figure 5.11.

The phase space is split into regions classified as signal or background, depending on the majority of training events that end up in the final leaf node. In case of regression trees, each output node represents a specific value of the target variable. The classification and regression performance of the decision tree can be enhanced due to boosting: The trees are derived from the same training ensemble by reweighting events and are finally combined into a single classifier, given by a weighted average of the individual decision trees. This method helps to stabilize the response of the decision trees with respect to fluctuations in the training sample, thus enhancing the performance of a single tree. There are four different boosting algorithms available for decision trees in TMVA:

1) Ada(ptive)Boost: It is a MVA boosting algorithm, where misclassified events during the training of a decision tree are given a higher event weight in the training of the following one. Starting with the original event weights when training the first decision tree, the subsequent one is trained using a modified event sample where the weights of previously misclassified events are multiplied by a common boost weight α . This weight is derived from the mis-classification rate E of the previous tree [40]:

$$\alpha = \beta \times \frac{1 - E}{E} \quad (5.15)$$

where β is a free boosting parameter to adjust the strength of boosting. The result of an individual classifier is $h(\vec{x})$, with $\vec{x}$ being the n-tuple of input variables, encoded for Signal and Background :

$$h(\vec{x}) = \begin{cases} +1 & \text{SIGNAL} \\ -1 & \text{BACKGROUND} \end{cases} \quad (5.16)$$

The boosted event classification $y_{Boost}(\vec{x})$ is then given as following:

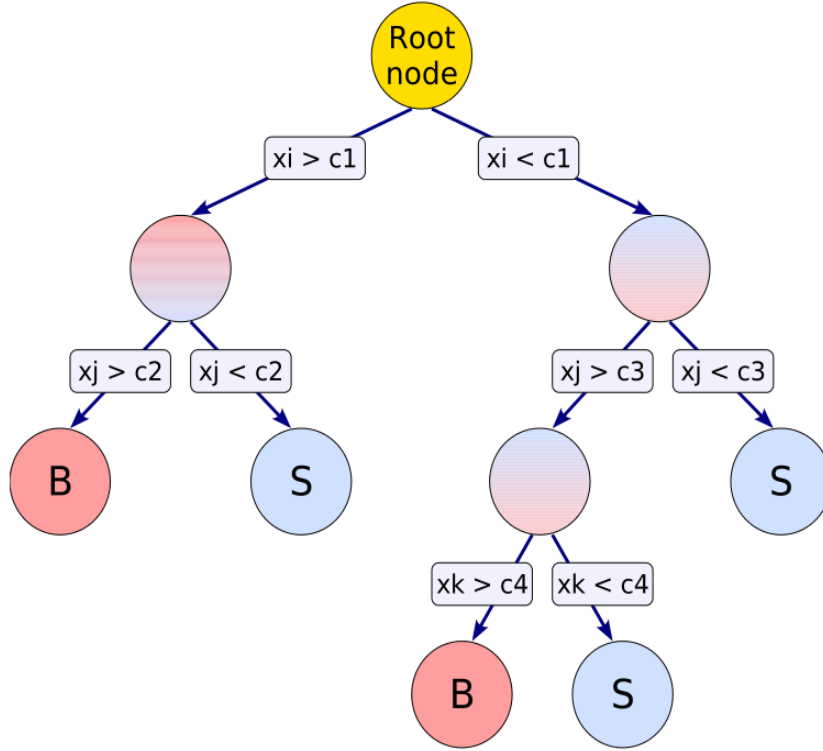


Figure 5.11. . Schematic view of a decision tree. Starting from the root node, a sequence of binary splits using the discriminating variables x_i is applied to the data. Each split uses the variable that at this node gives the best separation between signal and background when being cut on. The same variable may thus be used at several nodes, while others might not be used at all. The leaf nodes at the bottom end of the tree are labeled "S" for signal and "B" for background depending on the majority of events that end up in the respective nodes. For regression trees, the node splitting is performed on the variable that gives the maximum decrease in the average squared error when attributing a constant value of the target variable as output of the node, given by the average of the training events in the corresponding (leaf) node. This picture is adapted from Ref. [129]

$$y_{Boost}(\vec{x}) = \frac{i}{N_{collection}} \sum_i^{N_{collection}} \ln(\alpha_i) \cdot h_i(\vec{x}) \quad (5.17)$$

where the sum is over all classifiers in the collection. Small (large) values for $y_{Boost}(\vec{x})$ is an indication for a background-like (signal-like) event.

2) Gradient Boost: The Gradient Boost algorithm is a MVA algorithm, which aims to enhance the performance of a decision tree through the idea of additive expansion [129]. For that, let a function $F(\vec{x})$ be the one to be estimated, assuming to be a weighted sum of parameterized base functions $f(x; a_m)$, so-called "weak-learners" [88]. So, each base function in this expansion corresponds to a decision tree:

$$F(\vec{x}|P) = \sum_{m=0}^M \beta_m f(x; a_m), \quad P \in \{\beta_m, a_m\} \quad (5.18)$$

where the parameter a_m is the splitting nodes of tree branches, P is the consecutive boosting step. The boosting procedure is employed to adjust the parameters P in such a way that the deviation between the model response $F(\vec{x})$ and the true value y obtained from the training sample, is minimized. The deviation is measured by the loss-function $L(F, y)$:

$$L(F, y) = [F(\vec{x}) - y]^2 \quad (5.19)$$

3) Bagging: It is a resampling technique, where a classifier is repeatedly trained using re-sampled training events such that the combined classifier represents an average of the individual classifiers.

4) Randomised Trees: Each tree is grown in such a way that at each split only a random subset of all variables is considered. Moreover, each tree in the forest is grown using only a (resampled) subset of the original training events.

Booking and building of the BDT

The machine learning technique we are using within the TMVA framework, is the BDT and as a boosting method, we use the gradient boosting. Further details on our machine learning algorithm are listed in the [Table 5.9](#) below.

Parameters	Values	Description
NTrees	400	nr. of decision trees in the forest
MinNodeSize	2.5%	Minimum percentage of training events required in a leaf node
BoostType	Gradient	Boosting method used
Shrinkage	0.10	Learning rate for GradBoost algorithm
UseBaggedBoost	true	Boost the forest of decision trees
BaggedSampleFraction	0.5	Relative size of bagged event sample to original size of the data sample
nCuts	20	Number of cuts performed on one decision tree
MaxDepth	2	Max depth of the decision tree

Table 5.9. Configuration options used for the BDT training for all the tree years (2016-2017-2018) samples. Other options used in the algorithm are kept to their default values.

In order to better discriminate between signal and background events and to improve the sensitivity of the analysis, a Boosted Decision Tree is trained exploiting different kinematic variables. The network is trained in a region enriched with signal events passing the following requirements:

- $H \rightarrow 4l$ full selection
- $|m_{4l} - 125| < 10$ GeV
- at least 2 jets in the event

After this selection, the two jets with the highest b -discriminator are selected to reconstruct the $H \rightarrow b\bar{b}$. The signal is trained against all the backgrounds which remain in the signal region: $t\bar{t}H$, $t\bar{t}Z^0$, $t\bar{t}W^\pm$, SM-H, VBF, $W^\pm H$, $Z^0 H$, $Z^0 Z^0$. For each year, the strategy is to train three times the BDT, depending on the different Higgs decay mode":

1. $HH \rightarrow b\bar{b}4\mu$
2. $HH \rightarrow b\bar{b}4e$ and
3. $HH \rightarrow b\bar{b}2e2\mu$;

The shape of the BDT for all this three combination will be used as input in the statistical analysis. A large number of studies were performed to choose the best set of variables. Looking at both the area under the ROC curve and the overtraining check as a figure of merit, the final choice for the set of variables has been chosen distinguishing between the non-resonant and the resonant HH signal samples for the full RunII datasets.

BDT performance results of the non-resonant HH studies.

The input variables used for the non-resonant analysis are as following:

- p_T of each of the four leptons;
- ΔR between the $H \rightarrow 4l$ and $H \rightarrow b\bar{b}$;
- Value of the b-tagging algorithm score of the two jets with the highest b-tagger score;
- p_T of the two jets with the highest b tag discriminator score;
- two jets invariant mass;

After the training of the BDT using the input variables listed above, we look for the performance of the BDT, by studying the output of the Receiver Operating Characteristic metrics, also called as ROC curves, where the area under the curve tells us "how good" the output of the classifier is. In the [Figure 5.12](#) the ROC curves for each year and each channel with their respective area under the ROC curves for the non-resonant HH signal samples. The BDT input variables distributions for signal and background samples, obtained in case of the $b\bar{b}2e2\mu$ final state for the 2018 datasets, are shown in [Figure 5.14](#). In these plots it is possible to observe the differences between signal and background distributions. These distributions are obtained by counting how often the variables are used to split the decision tree nodes, and by weighting each split occurrence by the signal-background separation gain-squared achieved and by the number of events in each node [129]. The rankings of the BDT input variables for three possible final states $b\bar{b}4e$, $b\bar{b}4\mu$ and $b\bar{b}2e2\mu$ for the 2018 datasets, are reported in Tab. [Table 5.10](#), [Table 5.11](#) and [Table 5.12](#).

Rank	Variable	Value	Rank	Variable	Value	Rank	Variable	Value
1	ΔR_{HH}	13.90 %	1	ΔR_{HH}	15.12 %	1	ΔR_{HH}	15.47 %
2	b-disc. jet1	12.26 %	2	b-disc. jet1	12.43 %	2	b-disc. jet1	13.19 %
3	$m_{H(jj)}$	11.69 %	3	$m_{H(jj)}$	12.34 %	3	$m_{H(jj)}$	12.22 %
4	$p_{T(jet_2)}$	9.716 %	4	b-disc. jet2	10.54 %	4	b-disc. jet2	9.882 %
5	b-disc. jet2	9.590 %	5	$p_{T(l_1)}$	10.03 %	5	$p_{T(jet_2)}$	9.618 %
6	$p_{T(jet_1)}$	9.573 %	6	$p_{T(jet_2)}$	9.900 %	6	$p_{T(l_1)}$	9.299 %
7	$p_{T(l_1)}$	9.411 %	7	$p_{T(jet_1)}$	8.370 %	7	$p_{T(jet_1)}$	9.102 %
8	$p_{T(l_2)}$	8.590 %	8	$p_{T(l_3)}$	7.317 %	8	$p_{T(l_2)}$	7.717 %
9	$p_{T(l_3)}$	7.779 %	9	$p_{T(l_2)}$	7.288 %	9	$p_{T(l_3)}$	7.006 %
10	$p_{T(l_4)}$	7.487 %	10	$p_{T(l_4)}$	6.659 %	10	$p_{T(l_4)}$	6.491 %

Table 5.10. . BDT ranking of input variables of $b\bar{b}4e$ -channel for 2018 dataset.

Table 5.11. . BDT ranking of input variables of $b\bar{b}4\mu$ -channel for 2018 dataset.

Table 5.12.. BDT ranking of input variables of $b\bar{b}2e2\mu$ -channel for 2018 dataset.

The ranking of the variables changes by few percentages accordingly to the final state. Anyway, from the rankings we can notice that the angular separation ΔR_{HH} between the two Higgs bosons, the invariant mass of the Higgs boson into a pair of jets $m_{H(jj)}$, the b-tagger scores and the jets $p_{T(jet_1/jet_2)}$ contribute at most to the discrimination between background and signal.

The distributions of the discrimination between signal and background events can be seen in [Figure 5.13](#) for all three final states ($b\bar{b}4e$, $b\bar{b}4\mu$ and $b\bar{b}2e2\mu$) for each year (2016-2017-2018 period). From these

distributions we can conclude that the training of the BDT with the input variables listed above was able to discriminate between the signal and background events for the non-resonant SM HH analysis. Figure 5.14 and Figure 5.14 shows the distributions of the input variables used in the training for both signal and background processes. The majority of the input variables have a lower correlation for both signal and background as shown in Figure 5.15. More details about the final results with the signal sidebands will be described in chapter [Results].

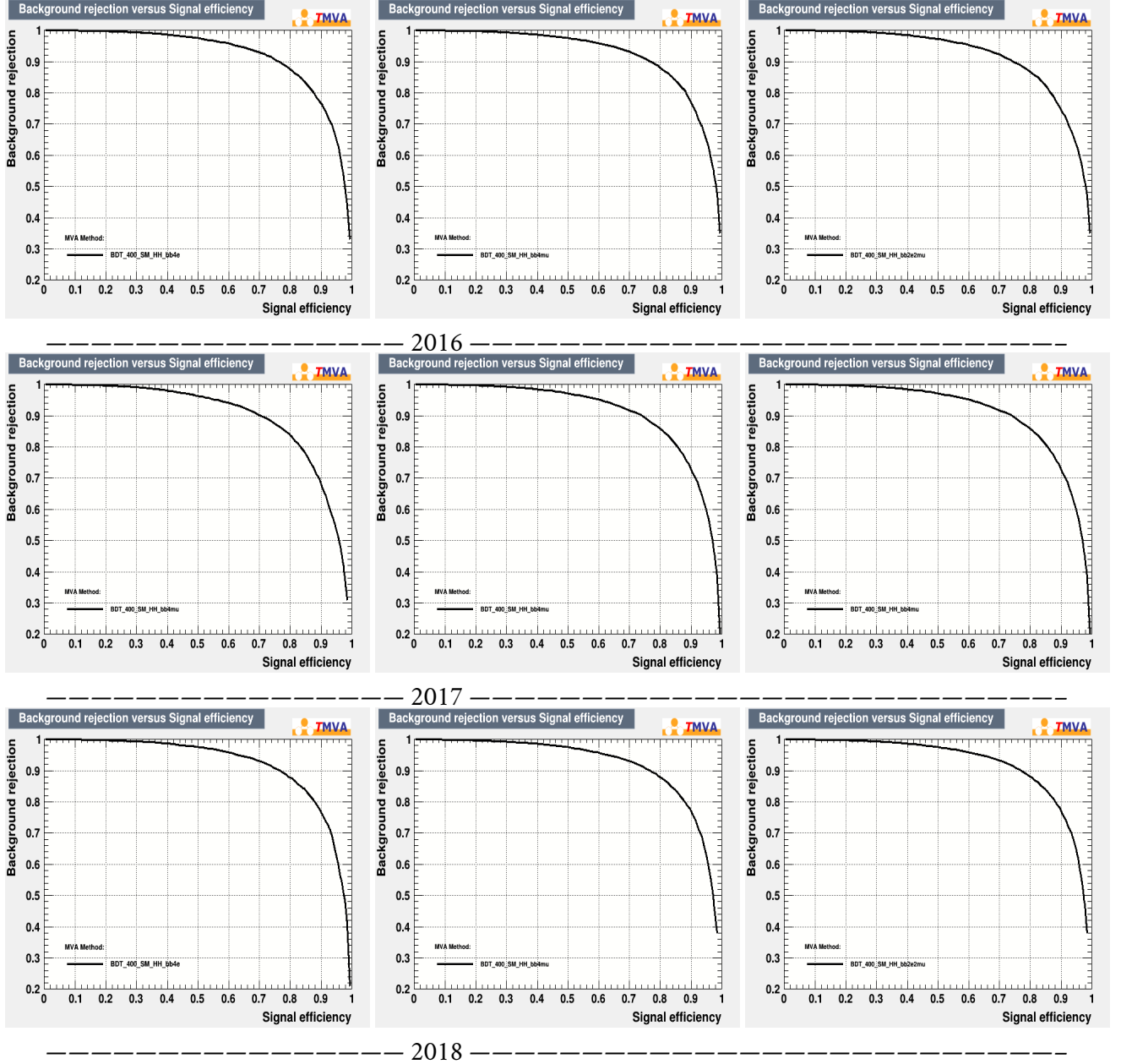


Figure 5.12. ROC curves after the training of the BDT for the channels : 1) 2016 (top): $b\bar{b}4e$ (top left) with an area of 0.934, $b\bar{b}4\mu$ (top center) with an area of 0.932 and $b\bar{b}2e2\mu$ (top right) with an area of 0.924 for the SM HH at NLO signal sample; 2) 2017 (center): $b\bar{b}4e$ (center left) with an area of 0.891, $b\bar{b}4\mu$ (center center) with an area of 0.912 and $b\bar{b}2e2\mu$ (center right) with an area of 0.914 for the SM HH at NLO signal sample; 3) 2018 (bottom): $b\bar{b}4e$ (bottom left) with an area of 0.930, $b\bar{b}4\mu$ (bottom center) with an area of 0.932 and $b\bar{b}2e2\mu$ (bottom right) with an area of 0.933 for the SM HH at NLO signal sample.

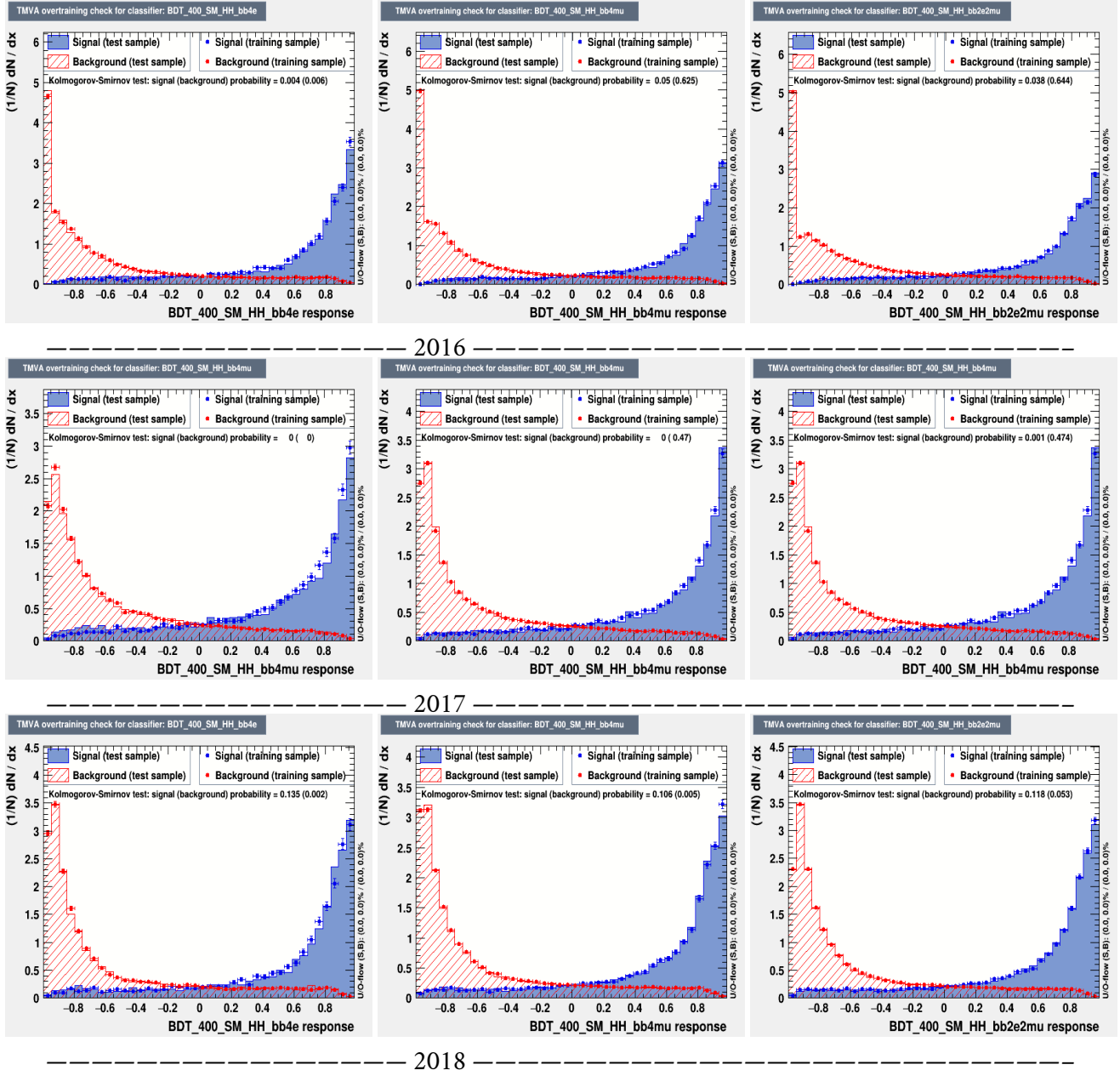


Figure 5.13. . Classifier output after the training of the BDT for the channels : 1) 2016 (top): $bb4e$ (top left), $bb4\mu$ (top center) and $bb2e2\mu$ (top right) for the SM HH at NLO signal sample; 2) 2017 (center): $bb4e$ (center left), $bb4\mu$ (center center) and $bb2e2\mu$ (center right) for the SM HH at NLO signal sample; 3) 2018 (bottom): $bb4e$ (bottom left), $bb4\mu$ (bottom center) and $bb2e2\mu$ (bottom right) for the SM HH at NLO signal sample;

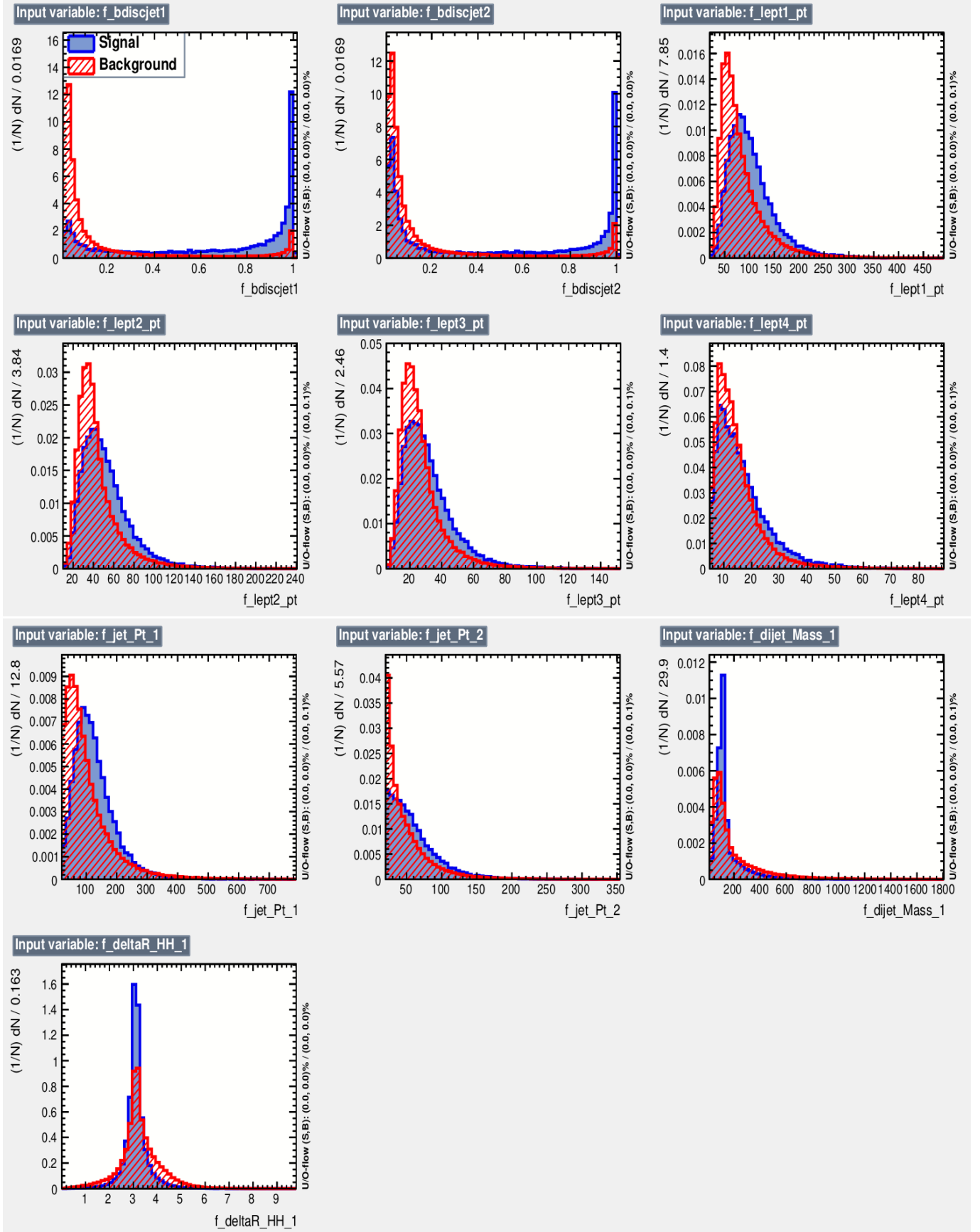


Figure 5.14. . BDT input variables distributions for 2018 datasets in case only of the $b\bar{b}2e2\mu$ final state as an example.

The non-resonant analysis $HH \rightarrow b\bar{b}ZZ(4l)$, $l = e, \mu$ is already approved and so we could perform a data to MC comparison for the BDT input variables listed above. The distributions of these kinematic variables within the signal region, also called as the sideband of the four leptons, where at least two jets with the highest b-tagger score are also implemented, were reported with all the three final states ($b\bar{b}4e$, $b\bar{b}4\mu$ and $b\bar{b}2e2\mu$) summed together. The BDT response is obtained applying the BDT to data, MC samples, and Z+X data-driven background, for each of the three final states and thus, for each year during he Run II (2016-2017-2018). The response distributions of data, signal, and all background contributions are then used as input for the statistical analysis, which will be reported in the chapter [Results], performed to extract the final results. The distribution of the BDT response showing the data, signal, and all backgrounds contributions ranges within the interval $[-1, 1]$, where the events that present background-like characteristics are distributed towards -1 , while events more signal-like are concentrated around 1.

Even if the statistics in our signal region is very low, a good agreement is observed for all the variables used in the BDT training. In order to understand if the BDT response is well described by data, the score of the BDT is plotted in the 4-leptons sideband:

- $H \rightarrow 4l$ selection;
- $|m_{4l} - 125.09| > 10 \text{ GeV}$.

The BDT response overall for all the three years in the sideband of the 4-leptons distribution shows a good data / mc agreement, as we can seen in [Figure 5.17](#), [Figure 5.18](#), [Figure 5.19](#) and [Figure 5.16](#). This means that the BDT is under control and thus the shape of the BDT in signal region can be used to extract the upper limit on the signal strength.

BDT performance results of the resonant KK Bulk Graviton spin-2 and Radion spin-0 HH studies.

The resonant HH analysis was subsequently performed after the final approve of the non-resonant $HH \rightarrow b\bar{b}ZZ(4l)$, $l = e, \mu$. Thus, the BDT training for this analysis was set-up differently, where more input variables were implemented. The input variables used for the HH resonant analysis are as following:

- p_T of each of the four leptons;
- Transverse momentum $p_T^{H_{4l}}$ and invariant mass $m_{H_{4l}}$ of the Higgs boson into a four lepton final state;
- Invariant mass m_{HH} of each mass point, the transverse momentum p_T^{HH} of the di-Higgs and the angular separation ΔR between the $H \rightarrow 4l$ and $H \rightarrow b\bar{b}$;
- Value of the b-tagging algorithm score of the two jets with the highest b-tagger score;
- $p_T^{H_{jj}}$ of the two jets with the highest b tag discriminator score the di-jets invariant mass $m_{H_{jj}}$ and the angular separation ΔR_{jj} between the two jets;

After the training of the BDT using the input variables listed above, we look for the performance of the BDT, as we did in the subsection above. The BDT was set-up with the same configuration options of the classifier shown in [Table 5.9](#) and this for all the tree years (2016-2017-2018) and for the three possible final states $b\bar{b}4e$, $b\bar{b}4\mu$ and $b\bar{b}2e2\mu$ of the resonant signal samples. We are going to show:

- ROC curves;
- BDT discriminant distribution of the classifier between the signal and background;

- BDT input variables distributions for signal and background samples;
- BDT input variables rankings.

For convenience, we are going to show this exercise with the above listed processes for two mass points: the KK Bulk Graviton spin-2 of invariant mass at $m_{HH}^{BkGr} = 400\text{GeV}$ and Radion spin-0 of invariant mass at $m_{HH}^{RD} = 500\text{GeV}$. The respective ROC for each channel and resonance at each year can be seen in [Figure 5.20](#), meanwhile the area under the curve (AUC) of the respective ROC are listed in [Table 5.13](#).

Signal	Year	AUC $2e2\mu$	AUC $4e$	AUC 4μ
Radion $m_{HH}^{RD} = 500\text{GeV}$	2016	0.946	0.953	0.951
Bulk Graviton $m_{HH}^{BKG} = 400\text{GeV}$	2016	0.959	0.956	0.965
Radion $m_{HH}^{RD} = 500\text{GeV}$	2017	0.962	0.954	0.955
Bulk Graviton $m_{HH}^{BKG} = 400\text{GeV}$	2017	0.966	0.946	0.965
Radion $m_{HH}^{RD} = 500\text{GeV}$	2018	0.967	0.961	0.963
Bulk Graviton $m_{HH}^{BKG} = 400\text{GeV}$	2018	0.974	0.969	0.975

Table 5.13. UAC of the ROC curves after the training of the BDT for the channels for the Bulk Graviton $m_{HH}^{BkGr} = 400\text{GeV}$ and Radion $m_{HH}^{RD} = 500\text{GeV}$ for all three years and each channel ($bb4e$, $bb4\mu$, $bb2e2\mu$)

Rank	Variable	Value	Rank	Variable	Value	Rank	Variable	Value
1	b-disc. jet1	8.771 %	1	b-disc. jet1	8.967 %	1	m_{HH}	9.932 %
2	m_{HH}	8.337 %	2	m_{HH}	8.870 %	2	b-disc. jet1	9.520 %
3	$p_T^{H(4l)}$	7.734 %	3	$p_T^{H(4l)}$	8.408 %	3	$p_T^{H(4l)}$	9.301 %
4	p_T^{HH}	7.662 %	4	ΔR_{jj}	7.183 %	4	$p_{T(jet_1)}$	7.493 %
5	$p_{T(jet_1)}$	6.732 %	5	$p_{T(jet_1)}$	7.032 %	5	p_T^{HH}	6.956 %
6	ΔR_{jj}	6.724 %	6	p_T^{HH}	7.006 %	6	ΔR_{jj}	6.807 %
7	b-disc. jet2	6.643 %	7	$m_{H(jj)}$	6.706 %	7	$m_{H(jj)}$	6.178 %
8	ΔR_{HH}	6.134 %	8	ΔR_{HH}	6.663 %	8	$p_{T(jet_2)}$	6.115 %
9	p_t^{jj}	6.119 %	9	b-disc. jet2	6.277 %	9	p_t^{jj}	6.081 %
10	$m_{H(4l)}$	5.874 %	10	$p_{T(jet_2)}$	6.179 %	10	b-disc. jet2	5.988 %
11	$p_{T(jet_2)}$	5.683 %	11	p_t^{jj}	6.034 %	11	ΔR_{HH}	5.771 %
12	$m_{H(jj)}$	5.312 %	12	$m_{H(4l)}$	5.331 %	12	$m_{H(4l)}$	5.481 %
13	$p_{T(l_2)}$	5.175 %	13	$p_{T(l_1)}$	4.547 %	13	$p_{T(l_1)}$	3.694 %
14	$p_{T(l_1)}$	4.937 %	14	$p_{T(l_2)}$	4.031 %	14	$p_{T(l_3)}$	3.666 %
15	$p_{T(l_3)}$	4.123 %	15	$p_{T(l_3)}$	3.452 %	15	$p_{T(l_2)}$	3.623 %
16	$p_{T(l_4)}$	4.041 %	16	$p_{T(l_4)}$	3.312 %	16	$p_{T(l_1)}$	3.395 %

Table 5.14. . BDT ranking of input variables of $bb4e$ -channel for 2018 dataset, Bulk Graviton spin-2, 400 GeV.

Table 5.15. . BDT ranking of input variables of $bb4\mu$ -channel for 2018 dataset, Bulk Graviton spin-2, 400 GeV.

Table 5.16. . BDT ranking of input variables of $bb2e2\mu$ -channel for 2018 dataset, Bulk Graviton spin-2, 400 GeV.

Rank	Variable	Value	Rank	Variable	Value	Rank	Variable	Value
1	m_{HH}	9.552 %	1	m_{HH}	9.889 %	1	m_{HH}	11.01 %
2	b-disc. jet1	8.314 %	2	$p_T^{H(4l)}$	8.285 %	2	b-disc. jet1	9.103 %
3	$p_T^{H(4l)}$	7.610 %	3	ΔR_{jj}	8.189 %	3	$p_T^{H(4l)}$	8.641 %
4	ΔR_{HH}	7.289 %	4	b-disc. jet1	7.789 %	4	p_T^{HH}	7.999 %
5	ΔR_{jj}	6.729 %	5	b-disc. jet2	7.274 %	5	ΔR_{jj}	7.332 %
6	p_T^{HH}	6.691 %	6	$m_{H(jj)}$	7.130 %	6	b-disc. jet2	7.304 %
7	ΔR_{jj}	6.559 %	7	ΔR_{HH}	6.991 %	7	ΔR_{HH}	7.203 %
8	b-disc. jet2	6.317 %	8	p_T^{HH}	6.956 %	8	$m_{H(4l)}$	6.322 %
9	$p_{T(jet_1)}$	6.240 %	9	$p_{T(jet_1)}$	6.775 %	9	$p_{T(jet_1)}$	6.043 %
10	$p_{T(jet_2)}$	6.011 %	11	$p_{T(jet_2)}$	5.096 %	10	$m_{H(jj)}$	6.026 %
11	p_t^{jj}	5.496 %	10	$m_{H(4l)}$	6.704 %	11	$p_{T(jet_2)}$	5.637 %
12	$m_{H(4l)}$	5.486 %	12	p_t^{jj}	4.645 %	12	p_t^{jj}	4.990 %
13	$p_{T(l_2)}$	4.980 %	13	$p_{T(l_1)}$	3.871 %	13	$p_{T(l_1)}$	3.439 %
14	$p_{T(l_3)}$	4.642 %	14	$p_{T(l_4)}$	3.739 %	14	$p_{T(l_3)}$	3.260 %
15	$p_{T(l_1)}$	4.399 %	15	$p_{T(l_2)}$	3.515 %	15	$p_{T(l_4)}$	3.218 %
16	$p_{T(l_4)}$	3.687 %	16	$p_{T(l_3)}$	3.152 %	16	$p_{T(l_2)}$	2.475 %

Table 5.17. . BDT ranking of input variables of $bb4e$ -channel for 2018 dataset, Radion spin-0, 500 GeV.

Table 5.18. . BDT ranking of input variables of $bb4\mu$ -channel for 2018 dataset, Radion spin-0, 500GeV.

Table 5.19.. BDT ranking of input variables of $bb2e2\mu$ -channel for 2018 dataset, Radion spin-0, 500 GeV.

The ranking of the variables are listed in the Tables [Table 5.14](#), [Table 5.15](#) and [Table 5.16](#) KK Bulk Graviton spin-2 at $m_{HH}^{BKGr} = 400\text{GeV}$ as well as for Tables [Table 5.17](#), [Table 5.18](#) and [Table 5.19](#) above, for the Bulk Radion spin-0 at $m_{HH}^{RD} = 500\text{GeV}$ for the 2018 datasets. As we can notice, the listed variables rank between $\sim 2.4 - 11\%$ for both events, Radion and Graviton. We can notice that in all three final states the invariant mass m_{HH}^{BKGr} and m_{HH}^{RD} of the Graviton and Radion respectively, is the variable, which contributes at most on the discrimination of signal and background samples, followed by the b-tagger score of the two jets with the highest b-tagger score values, transverse momentum of both Higgs bosons ($p_T^{H(4l)}$, p_t^{jj}), angular separation ΔR_{jj} of the two jets, which reconstruct the Higgs boson $H \rightarrow jj(b\bar{b})$ and further kinematic variables of the di-Higgs system like the angular separation ΔR_{HH} and transverse momentum p_T^{HH} . [Figure 5.22](#) and [Figure 5.23](#) shows the distributions of the input variables used in the training for both signal and background processes.. The majority of the input variables have a lower correlation for both signal and background as shown in [Figure 5.24](#).

The distributions of the discrimination between signal and background events for KK Bulk Graviton spin-2 with $m_{HH}^{BK-Gr} = 400\text{ GeV}$ and KK Bulk Radion spin-0 with $m_{HH}^{RD} = 500\text{ GeV}$ can be seen in [Figure 5.21](#) for all three final states ($b\bar{b}4e$, $b\bar{b}4\mu$ and $b\bar{b}2e2\mu$) for each year (2016-2017-2018 period). From these distributions we can conclude that the training of the BDT with the input variables listed above was able to discriminate between the signal and background events for the resonant HH analysis. More details about the final results with the signal sidebands will be describes in chapter [Results].

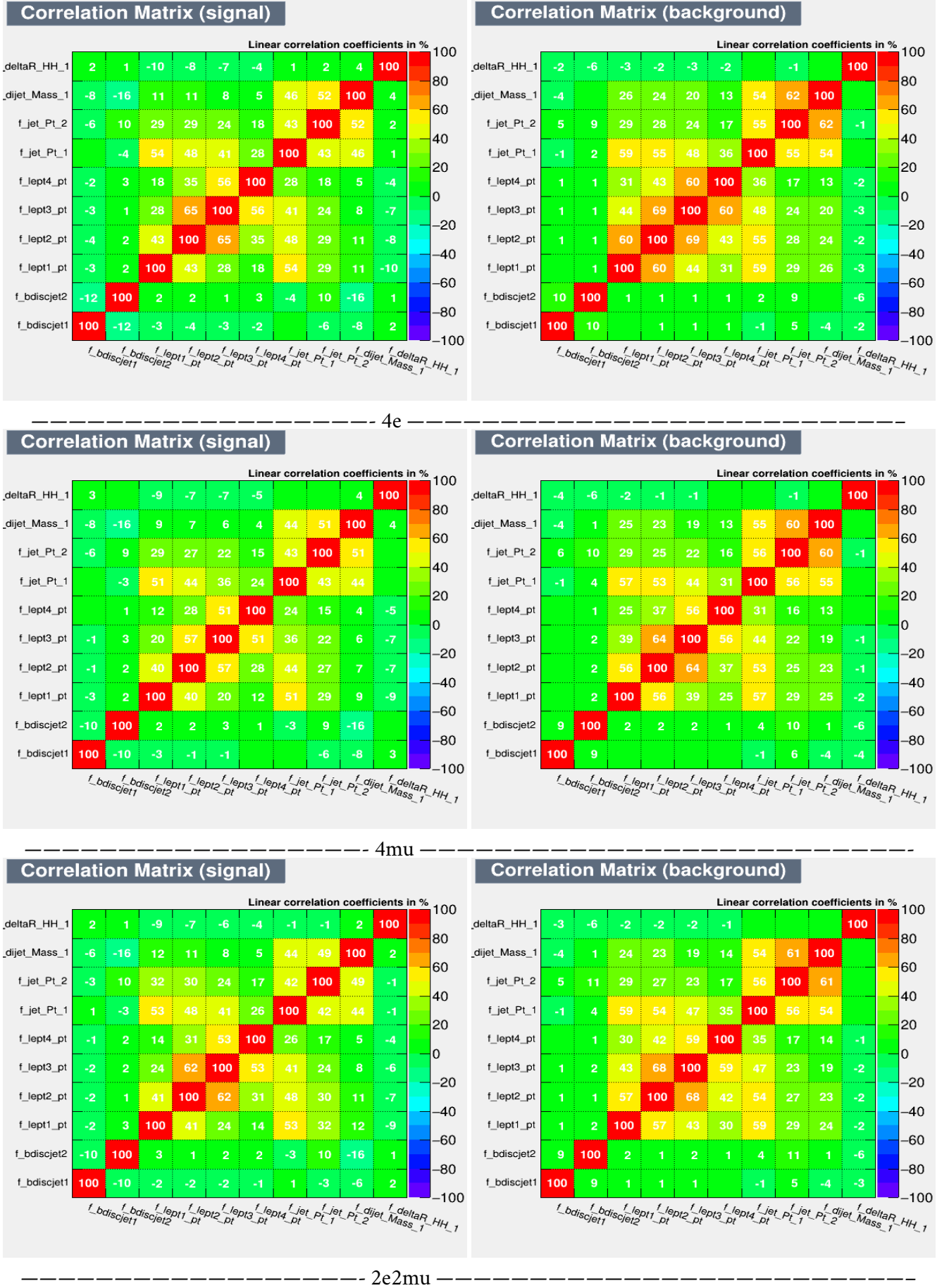


Figure 5.15. BDT correlation plots for 2018: correlation between background's variables (right), correlation between signal's variables (left) for each final state ($bb4e$ top, $bb4\mu$ center, $bb2e2\mu$ bottom).

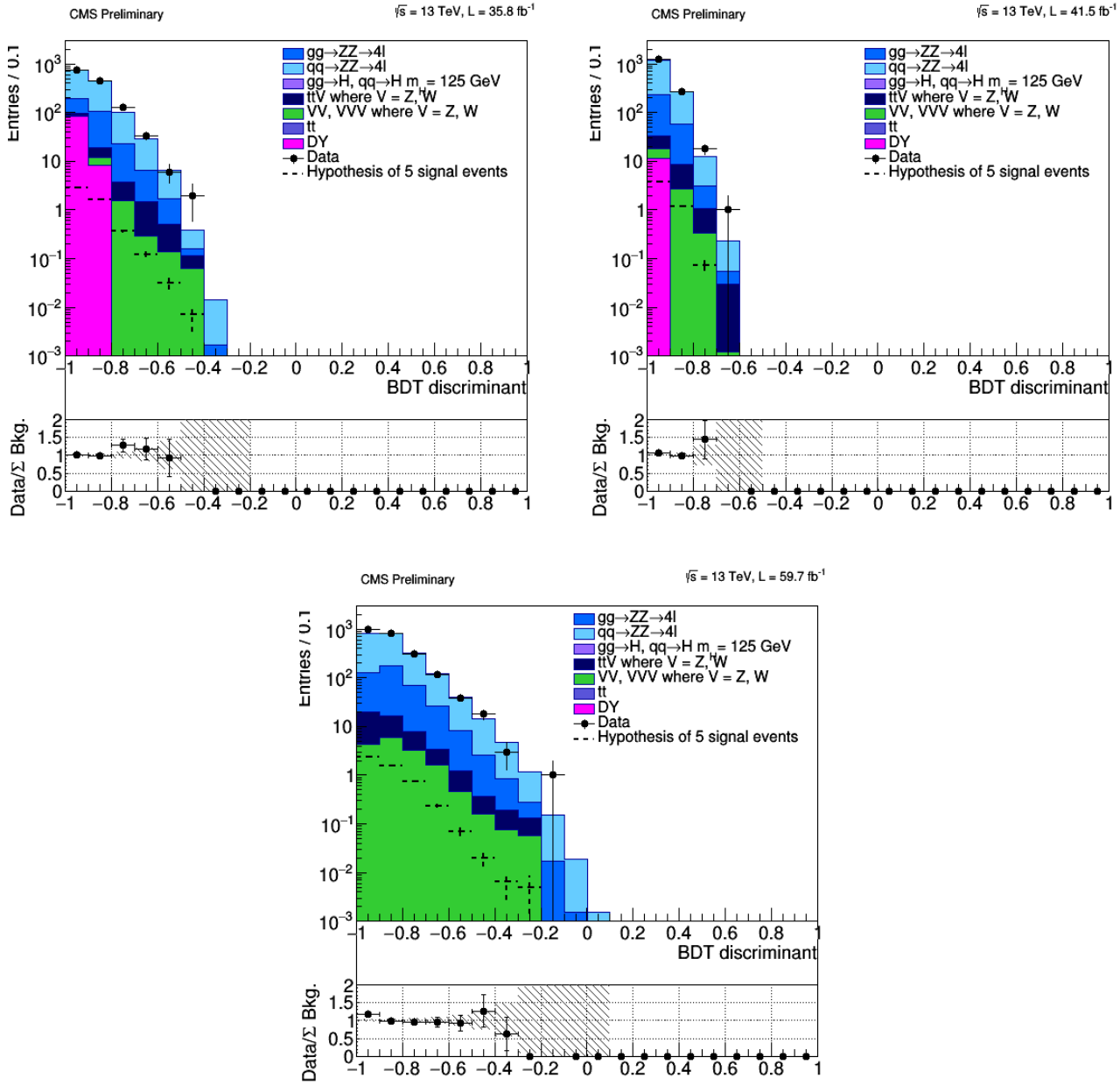


Figure 5.16. BDT score in the sideband of the 4 leptons distribution: 1) Top left: for 2016 datasets; 2) Top right: for 2017 datasets; 3) Bottom: for 2018 datasets;

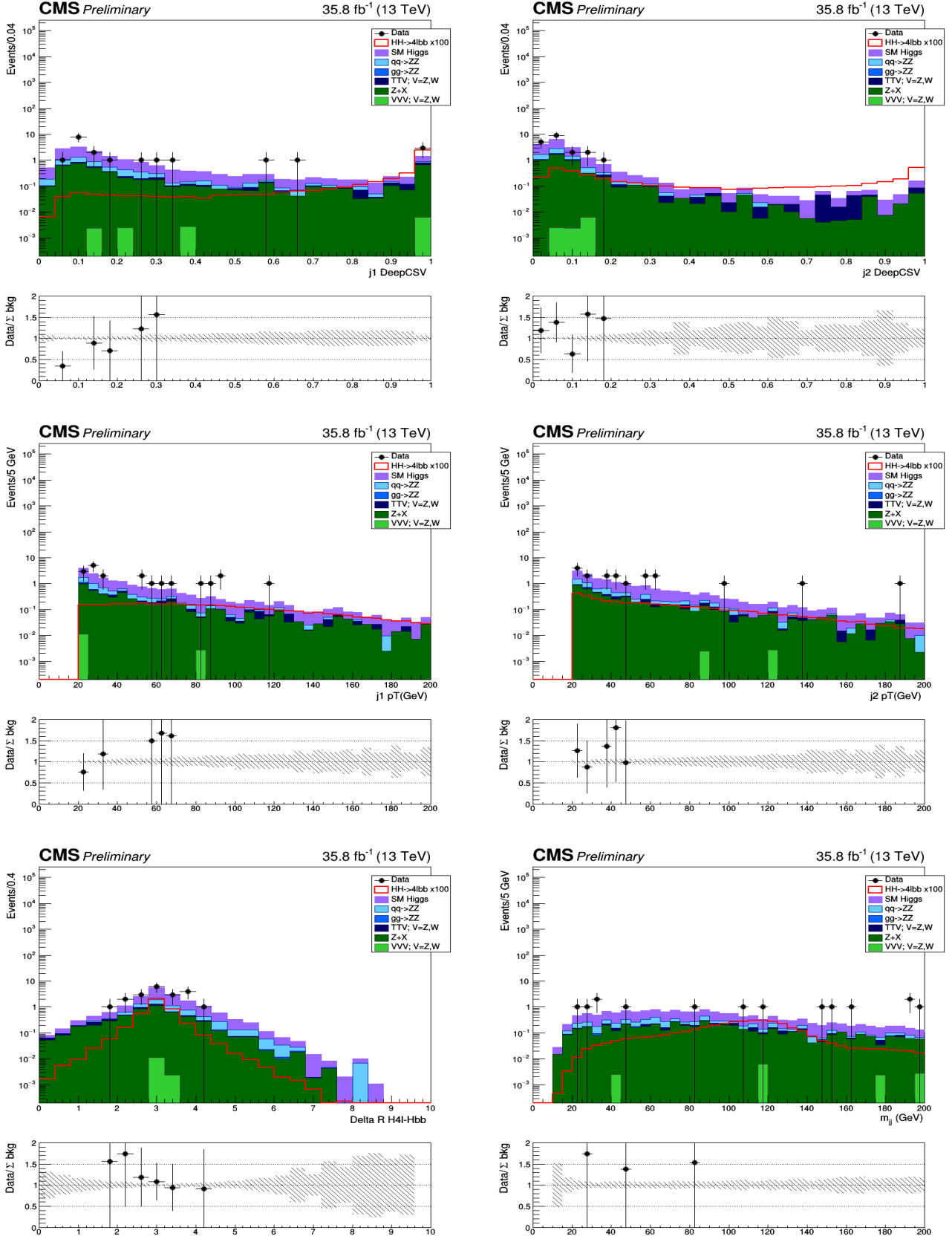


Figure 5.17. 1) Top left: B tagger score of the jet with the highest value of the deepCSV discriminant for 2016 datasets; 2) Top right: B tagger score of the jet with the second-highest value of the deepCSV discriminant for 2016 datasets; 3) Center left: p_T of jet with the highest value of the deepCSV discriminant for 2016 datasets; 4) Center right: p_T of jet with the second-highest value of the deepCSV discriminant for 2016 datasets; 5) Bottom left: ΔR between the $H \rightarrow 4l$ and $H \rightarrow bb$ for 2016 datasets; 6) Bottom right: Invariant mass of the jet built with the two jets with the highest value of the deepCSV discriminant for 2016 datasets.

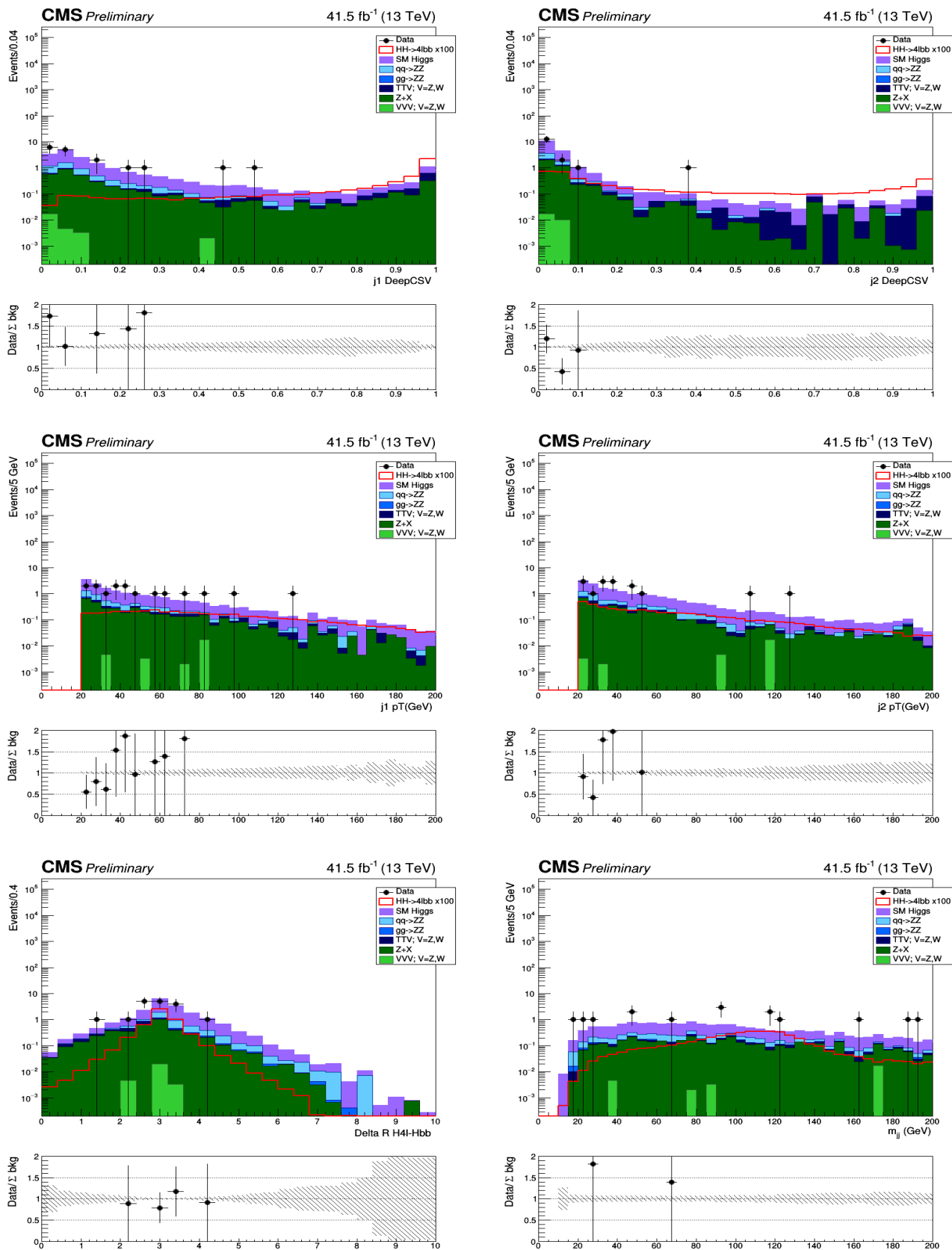


Figure 5.18. 1) Top left: B tagger score of the jet with the highest value of the deepCSV discriminant for 2017 datasets; 2) Top right: B tagger score of the jet with the second-highest value of the deepCSV discriminant for 2017 datasets; 3) Center left: p_T of jet with the highest value of the deepCSV discriminant for 2017 datasets; 3) Center right: p_T of jet with the second-highest value of the deepCSV discriminant for 2017 dataset; 5) Bottom left: ΔR between the $H \rightarrow 4l$ and $H \rightarrow b\bar{b}$ for 2017 datasets; 6) Bottom right: Invariant mass of the jet built with the two jets with the highest value of the deepCSV discriminant for 2017 datasets.

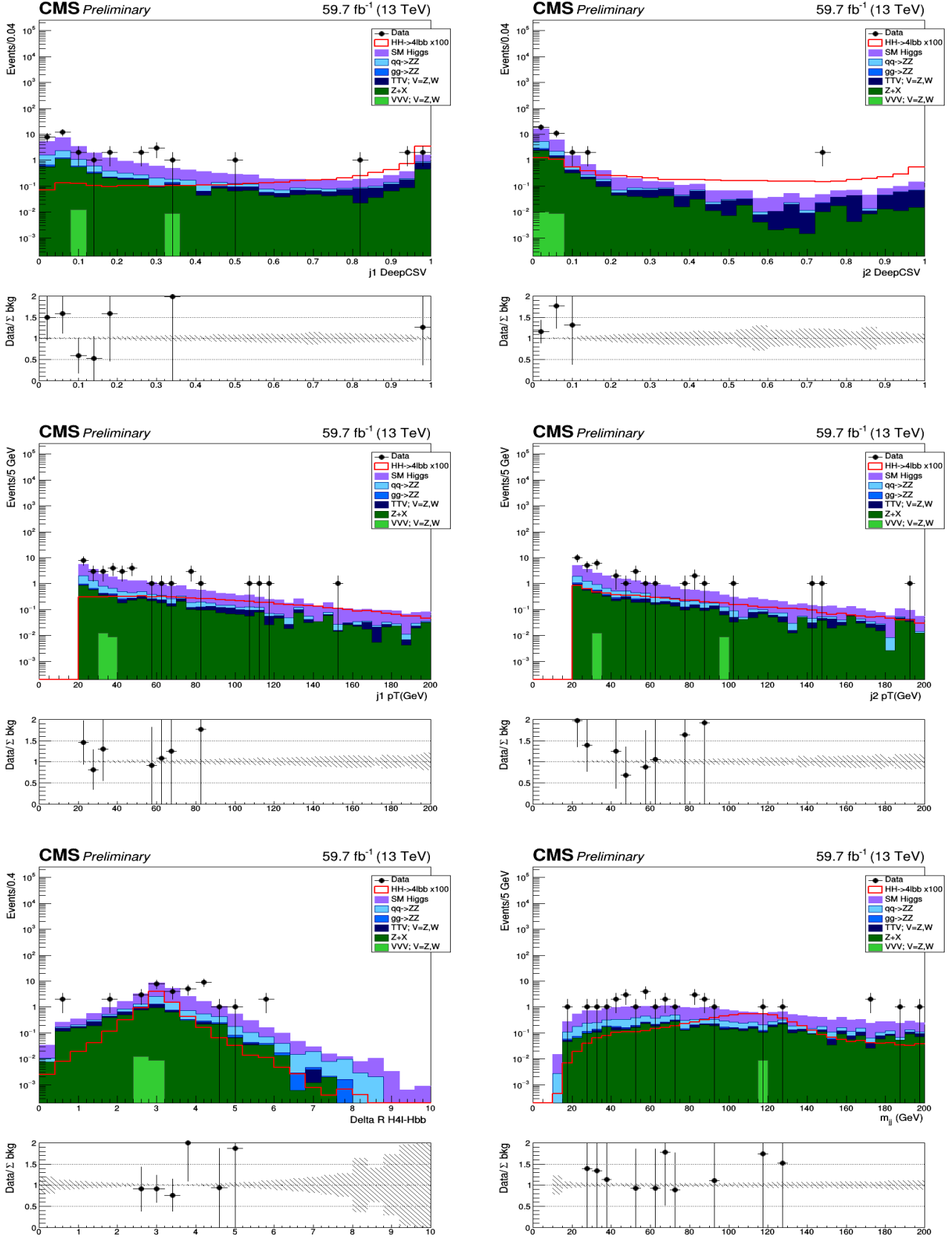


Figure 5.19. 1) Top left: B tagger score of the jet with the highest value of the deepCSV discriminant for 2018 datasets; 2) Top right: B tagger score of the jet with the second-highest value of the deepCSV discriminant for 2018 datasets; 3) Center left: p_T of jet with the highest value of the deepCSV discriminant for 2018 datasets; 4) Center right: p_T of jet with the second-highest value of the deepCSV discriminant for 2018 dataset; 5) Bottom left: ΔR between the $H \rightarrow 4l$ and $H \rightarrow b\bar{b}$ for 2018 datasets; 6) Bottom right: Invariant mass of the jet built with the two jets with the highest value of the deepCSV discriminant for 2018 datasets.

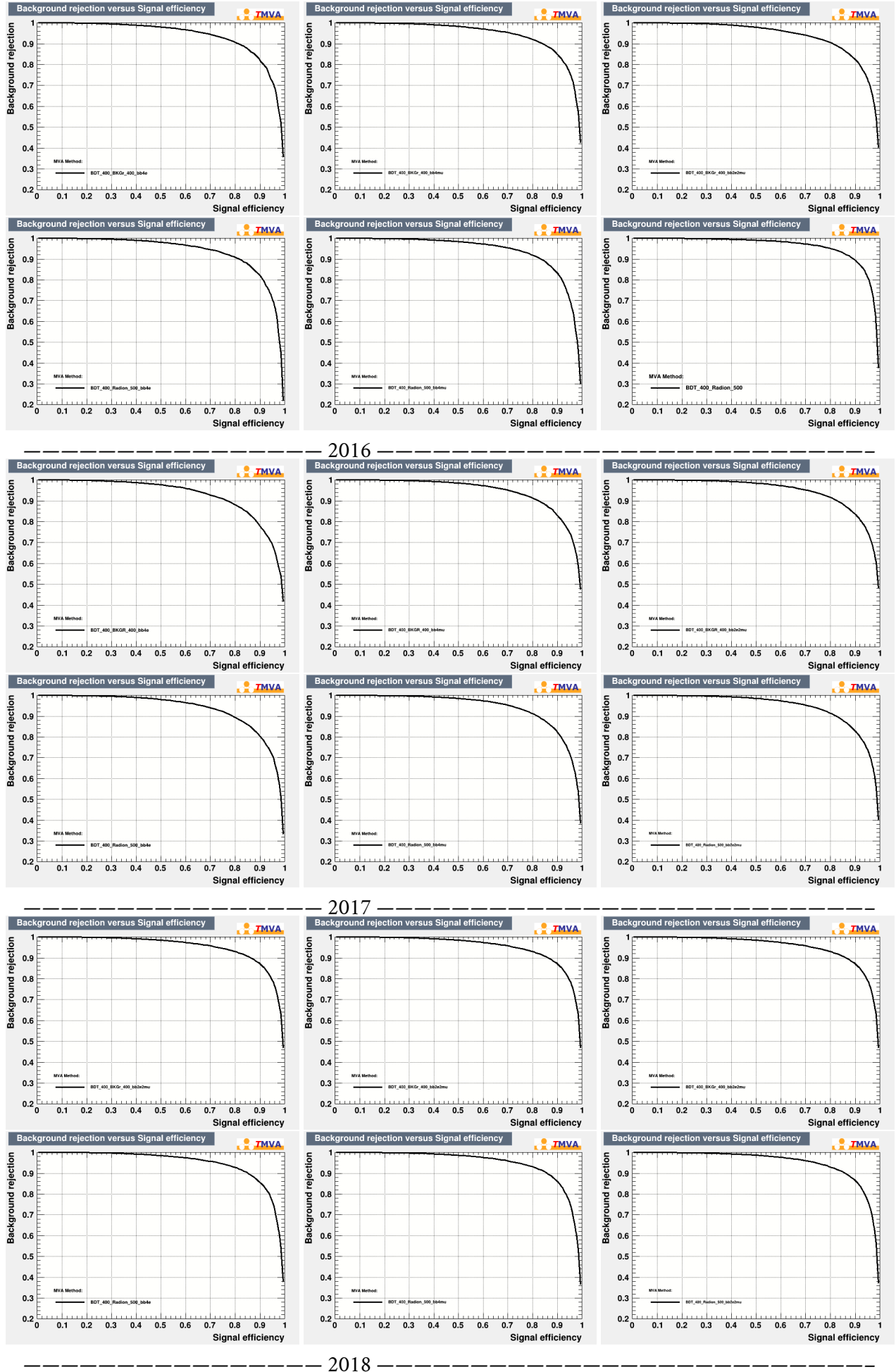


Figure 5.20. . ROC curves after the training of the BDT for the channels or Bulk Graviton $m_{HH}^{BKGr} = 400\text{GeV}$, Radion $m_{HH}^{RD} = 500\text{GeV}$ for all three years and three channels (column left: $bb4e$, column center: $bb4\mu$, column right: $bb2e2\mu$)

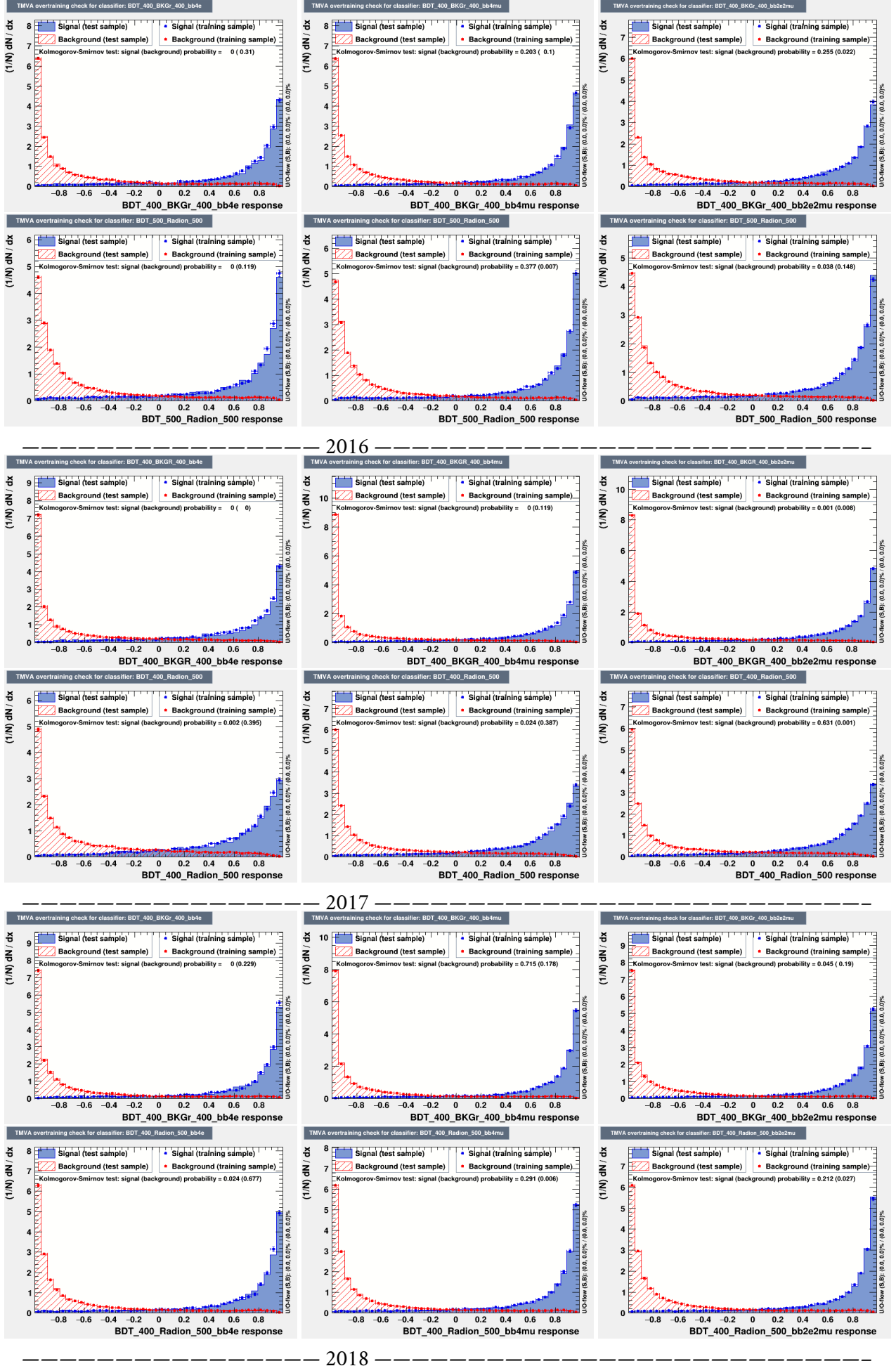


Figure 5.21. Classifier output after the training of the BDT for Bulk Graviton $m_{HH}^{BKGr} = 400\text{GeV}$, Radion $m_{HH}^{RD} = 500\text{GeV}$ for all three years and three channels (column left: $bb4e$, column center: $bb4\mu$, column right: $bb2e2\mu$).

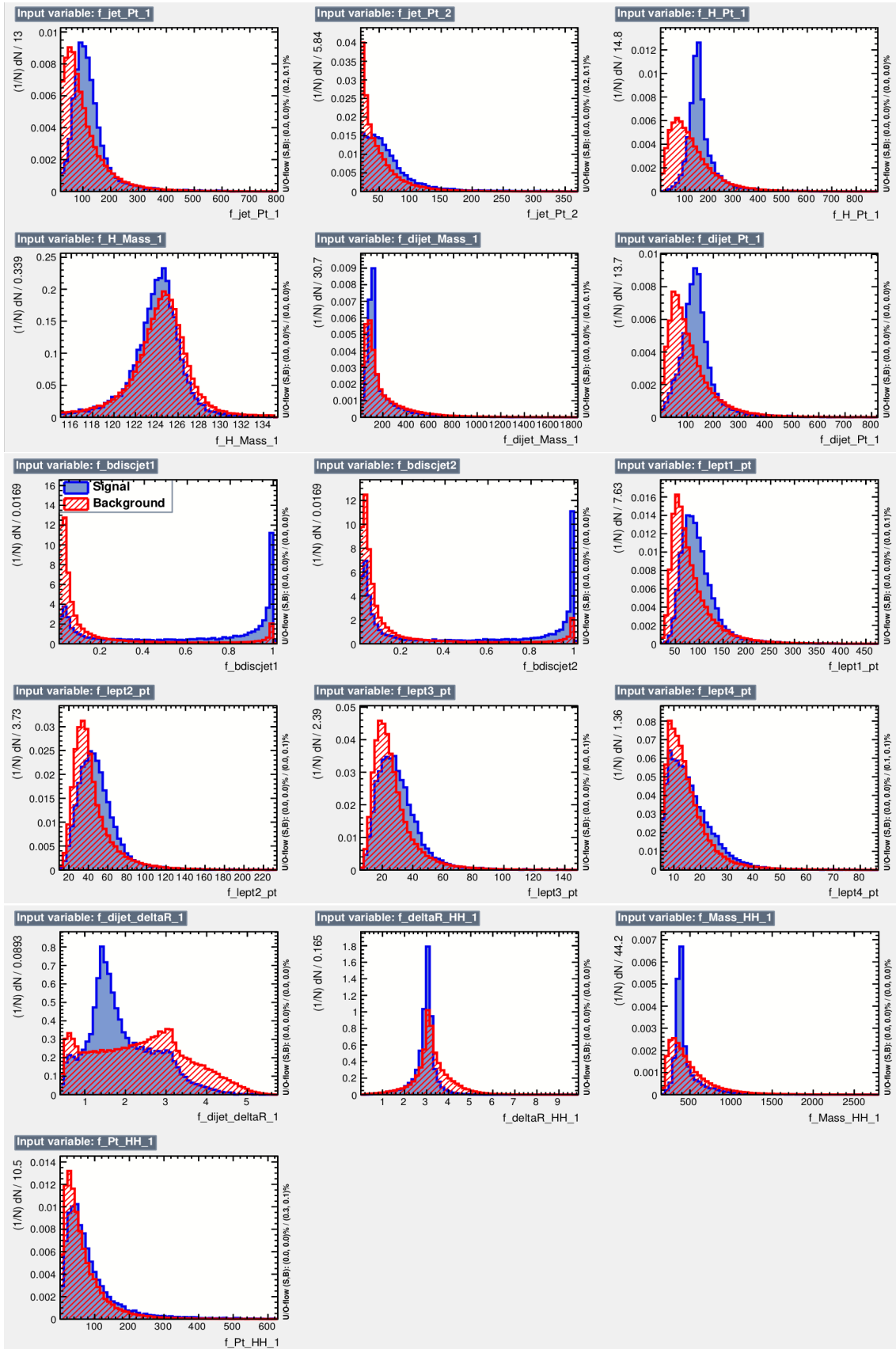


Figure 5.22. . BDT input variables distributions for 2018 datasets of the KK Bulk Graviton spin-2 with $m_{HH}^{BK-Gr} = 400$ GeV in case only of the $b\bar{b}2e2\mu$ final state as an example.

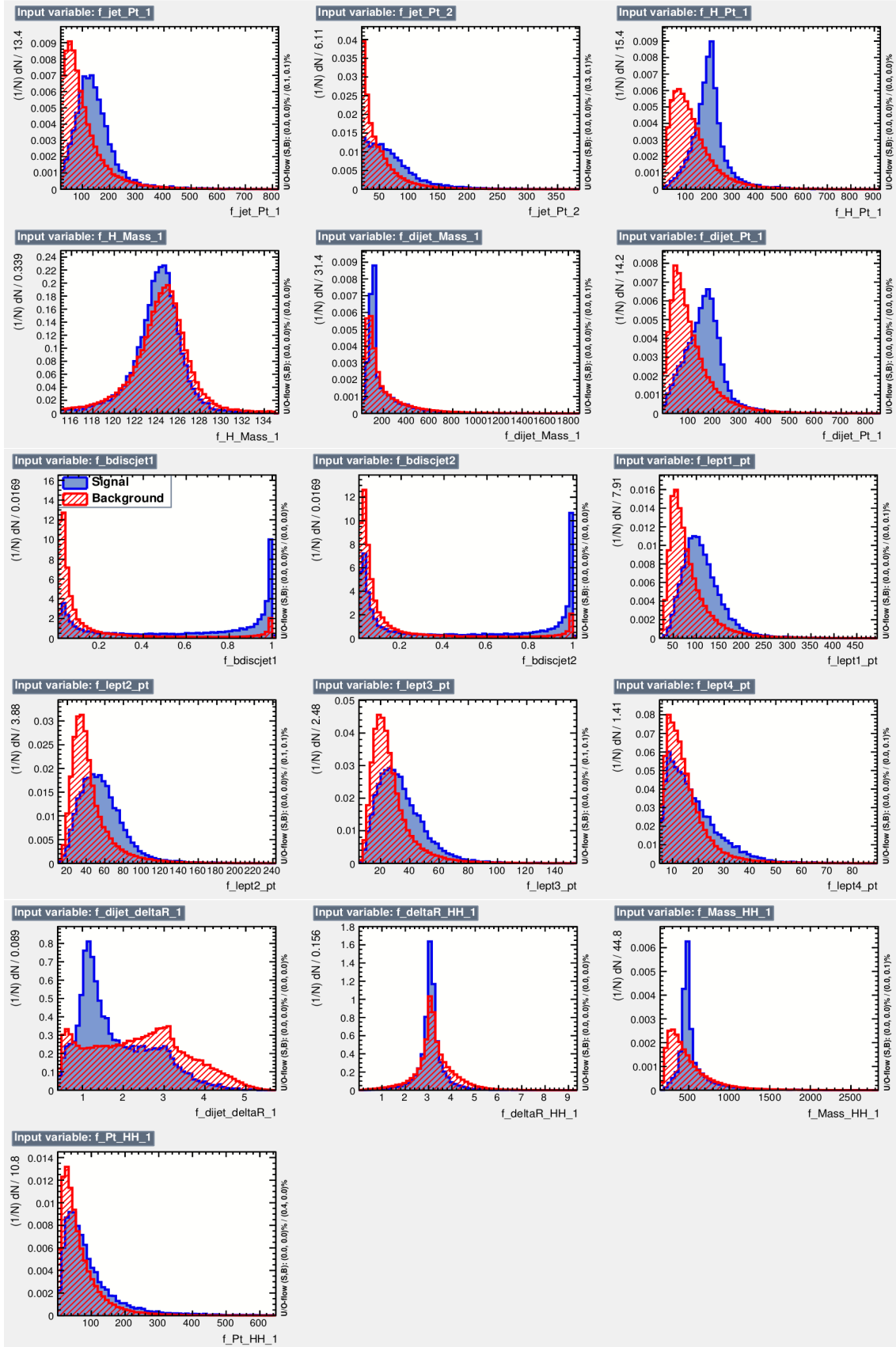
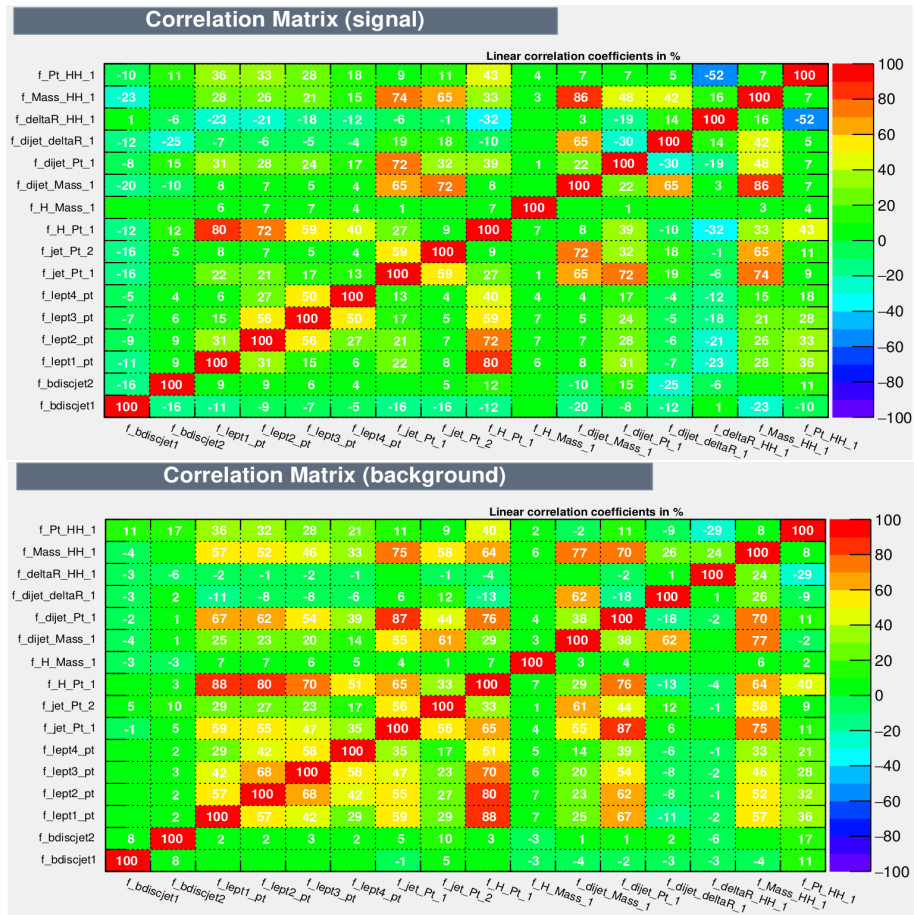
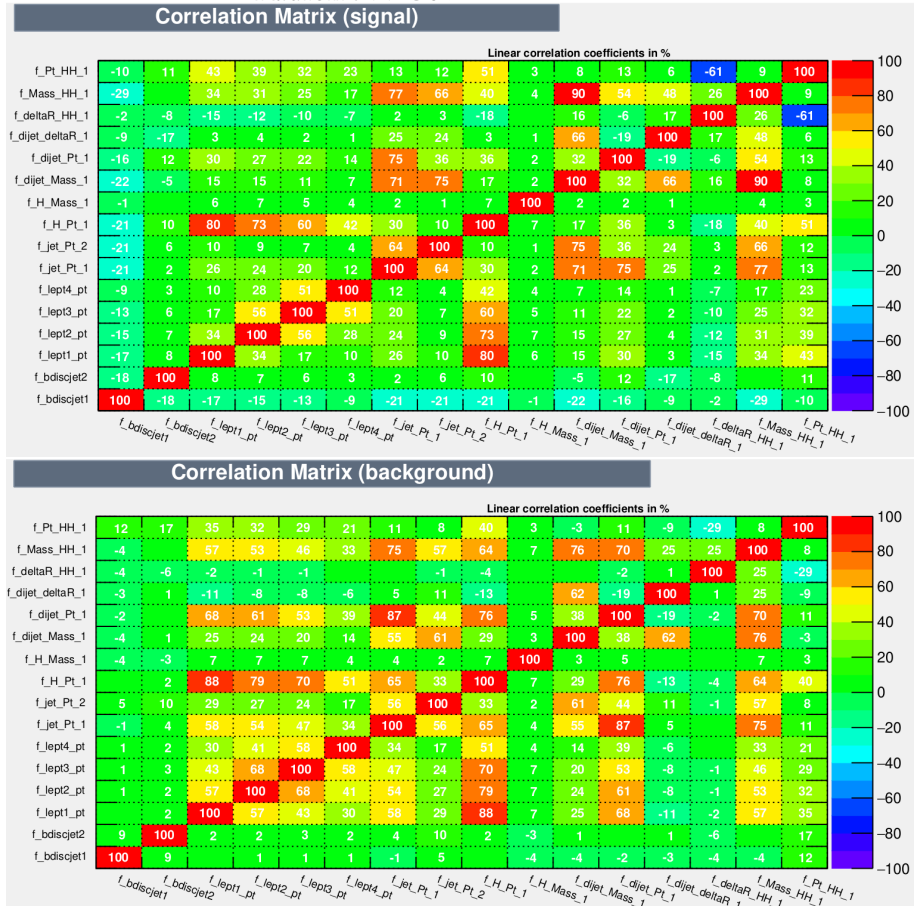


Figure 5.23. . BDT input variables distributions of the KK Bulk Radion spin-0 with $m_{HH}^{RD} = 500$ GeV for 2018 datasets in case only of the $b\bar{b}2e2\mu$ final state as an example.



Radion 500 GeV



Bulk Graviton 400 GeV

Figure 5.24. BDT correlation matrix plots for 2018 for KK Bulk Graviton spin-2 with $m_{HH}^{BK-Gr} = 400$ GeV and KK Bulk Radion spin-0 with $m_{HH}^{RD} = 500$ GeV for the correlation between background's variables (top) and correlation between signal's variables (bottom) for the final state $b\bar{b}2e2\mu$.

Part V

Results

Chapter 6

Results

The moment has finally come to give an outlook towards the essence of this analysis after the long journey we had to be part of, from the MC studies in chapter [4] and the full analysis Workflow explained in chapter [5], which starts from the physics objects identification and reconstruction to the training of a machine learning technique, such as the BDT within the TMVA framework. Thus, the goal of this chapter is to collect the final results extracted from the chapters before.

The final yields of each year for both variants, non-resonant and resonant $HH \rightarrow b\bar{b}ZZ(4l)$ for the full Run II datasets at $\sqrt{s} = 13\text{TeV}$ will be presented. A shape based analysis and a statistical analysis is performed, where the expected and observed limits are set on the signal strength for the full Run II datasets.

6.1 Signal yields

As mentioned in the introduction of this chapter, after the application of all analysis cut flow (explained in details in chapter [4,5]), we can now evaluate and report the yields within our signal region for signal and background processes:

- $H \rightarrow 4l$ selection;
- $|m_{4l} - 125.09| > 10 \text{ GeV}$;
- at least two isolated jets (with respect to the leptons) in the event.

We distinguish in the coming sections and subsections between the variants of the double Higgs production analysis: the non-resonant SM HH and the resonant *Graviton/Radion* $\rightarrow HH$ analysis; the reasons behind this decision are two:

1. The non-resonant SM HH analysis into a final state of two b-tagged jets and 4-leptons, considering the k_λ -scan (in BSM models) with full Run II data, is already approved, meanwhile the resonant *Graviton/Radion* $\rightarrow HH$ analysis was subsequently performed and it has still to go through all the steps up the final approval;
2. Both the analysis are based on different Physics evaluations regarding the Theoretical basis, explained for both of them in chapter [1].

In the following sections, signal yields and the statistical analysis regarding both variants will be reported.

6.1.1 Non-resonant $HH \rightarrow b\bar{b}ZZ(4l)$ signal region yields

In case of the non-resonant and resonant $HH \rightarrow b\bar{b}ZZ(4l)$ for the full Run II datasets at $\sqrt{s} = 13\text{TeV}$ the yields observed in data are extracted in the signal region and are reported in Table B.1, Table B.2 and Table B.2 for the 2016, 2017, and 2018 datasets, respectively, meanwhile for the full Run II datasets at $\sqrt{s} = 13\text{TeV}$ for the non-resonant HH analysis, the signal yields are listed in Table 6.1.

Topology	SM-HH	ttZ	ttH	bbH	ZZ	H+VBF	WH	ZH	Others	Z+X	Total bkg.	Total Data.
4μ	0.054	0.56	0.67	0.13	3.72	11.89	0.88	0.70	0.04	3.87	22.38	29
$4e$	0.028	0.35	0.38	0.07	1.43	5.93	0.45	0.37	0.02	2.64	11.60	12
$2e2\mu$	0.079	0.92	0.88	0.19	4.93	15.07	1.12	0.94	0.11	7.22	31.36	33

Table 6.1. Cut flow table after signal region selection for the full RunII dataset ($137.2fb^{-1}$). The ZZ yields contain contributions from the processes $gg/qq \rightarrow ZZ$, Higgs + VBF contains contributions from the single Higgs boson gluon fusion (ggH) and vector boson fusion (VBFH) production modes, others contains contributions from TTW, WWZ, WZZ, ZZZ and processes.

As we can notice in the tables listed above, the biggest contributions within our signal regions in all the three years is related to the single Higg boson plus the VBF-H contribution, followed by the Z+X and ZZ background samples, since all those samples contain either a four lepton final state or a multi-jet (at least two) final state. Nevertheless, all those contributions, besides the bbH and the VVV ("Others" = WWW, WWZ,ZZZ), are higher than the ones for the SM $HH \rightarrow b\bar{b}ZZ(4l)$ signal, contaminating heavily our signal region. The distribution of the BDT discriminant in the signal region for the three years is shown in Figure 6.1. The binning of the BDT is chosen in order to have (almost) the same statistical uncertainties in the last 4-5 bins.

In Fig. Figure 6.1 the $m_{H(4l)}$, $m_{H(bb)}$ and BDT score for the full Run II datasets at a $\sqrt{s} = 13\text{TeV}$ and a total integrated luminosity of $\mathcal{L} = 137fb^{-1}$ are shown, where we can notice that the data / mc agreement is relatively good. A good signal / background separation is observed for the three years, in particular $S/\sqrt{B}$ (where S is the signal and B is the background bin per bin) is higher in the last three bins, this means that the majority of the signal is concentrated at between 0.6 and 1, as expected for a well trained network.

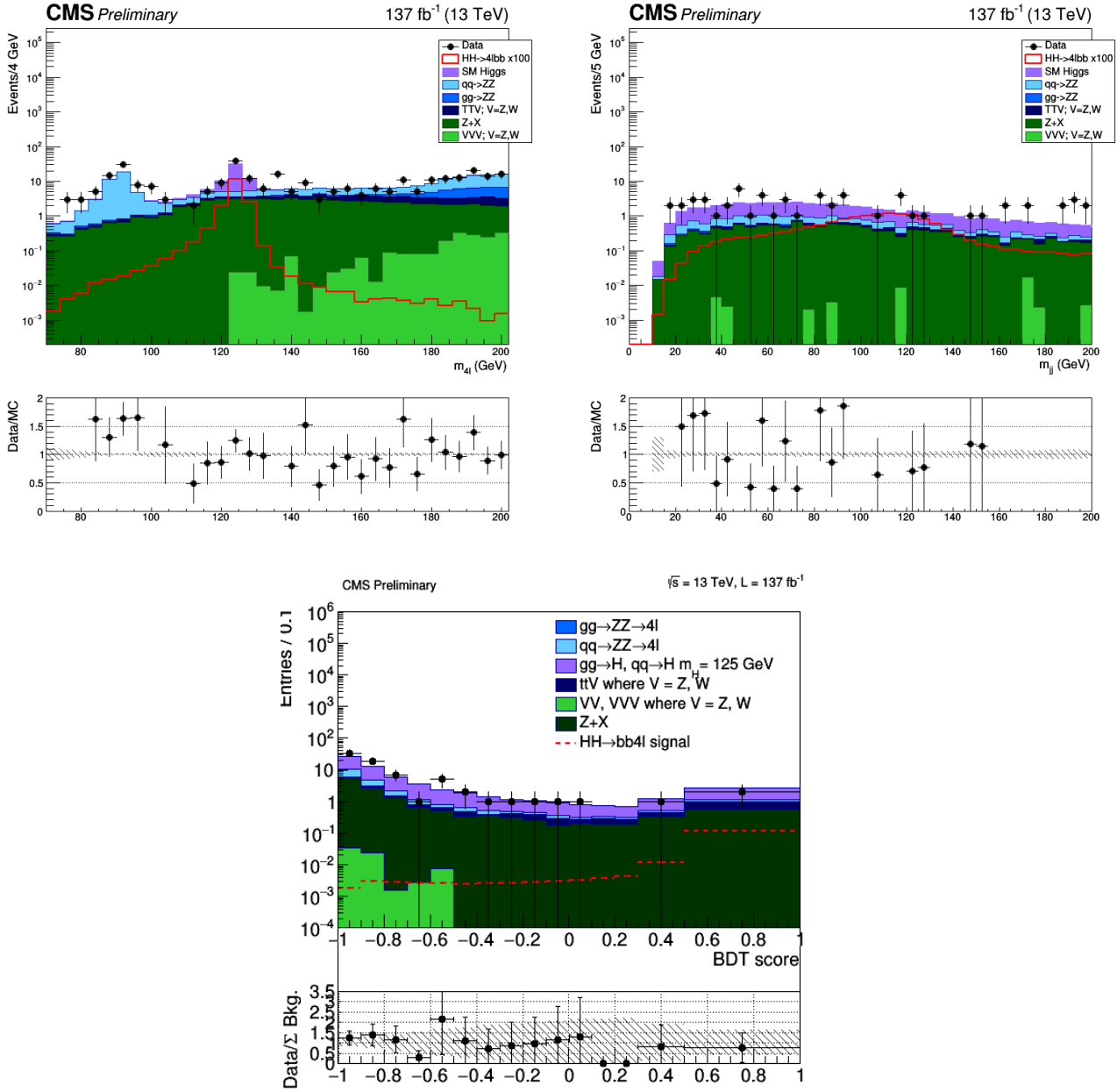


Figure 6.1. . Distribution of the 4-leptons invariant mass $m_{H(4l)}$ (top left) and the distribution of the invariant of the di-jet system $m_{H(bb)}$ (top right), where only events in the mass window $115\text{GeV} < m_{4l} < 135\text{GeV}$ are considered. The distributions and the BDT score for the signal HH , which is scaled by an illustrative factor (in red dotted line), the estimated background components, as well the data are shown in the bottom plot. A good signal / background separation is observed for the three years, in particular $S\sqrt{B}$ (where S is the signal and B is the background bin per bin) is higher in the last three bins, this means that the majority of the signal is concentrated at between 0.6 and 1.

6.1.2 Resonant Bulk Graviton spin-2 and Radion spin-0 $HH \rightarrow b\bar{b}ZZ(4l)$ signal region yields

In case of the resonant $HH \rightarrow b\bar{b}ZZ(4l)$ for the full Run II datasets at $\sqrt{s} = 13\text{TeV}$, the situation differs from the non-resonant HH analysis, because we are aiming for official approval within the collaboration; therefore this analysis for the moment is still blind. It is also important to remind that, in case of massive resonances such as the KK Bulk Graviton spin-2 and Radion spin-0, the strategy we are following is a mass-point like scan, where several mass points of each resonance were generated and studied. The mass scan was performed in the following values:

1. KK Bulk Graviton spin-2: $m_{HH}^{Graviton} = [260, 300, 400, 500, 600, 700, 800, 900, 1000] \text{ GeV}$
2. KK Bulk Radion spin-0: $m_{HH}^{Radion} = [260, 300, 400, 500, 600, 700, 800, 900, 1000] \text{ GeV}$

The distributions of signal region plots with full RunII datasets (137 fb^{-1}) of the kinematic variables: $m_{4l}, p_T^{4l}, p_T^{jj}, p_T^{jet_1}, p_T^{jet_2}, m_{HH}, p_T^{HH}$ is reported in Figure 6.4.

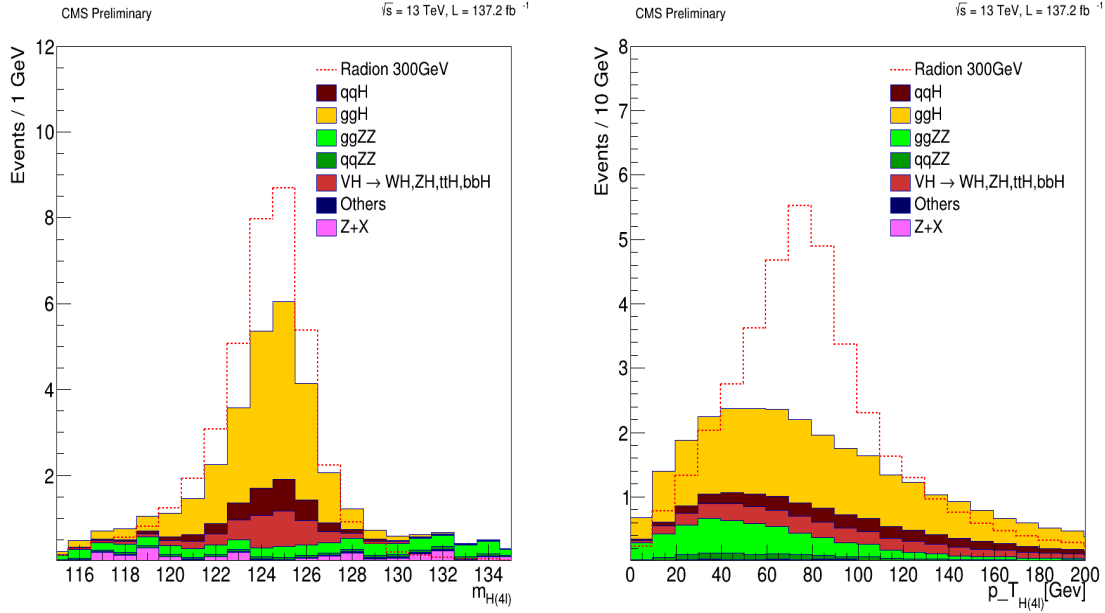


Figure 6.2. . Signal region plots Full RunII (137 fb^{-1}): $m_{H(4l)}$ (left), p_T^{4l} (right)

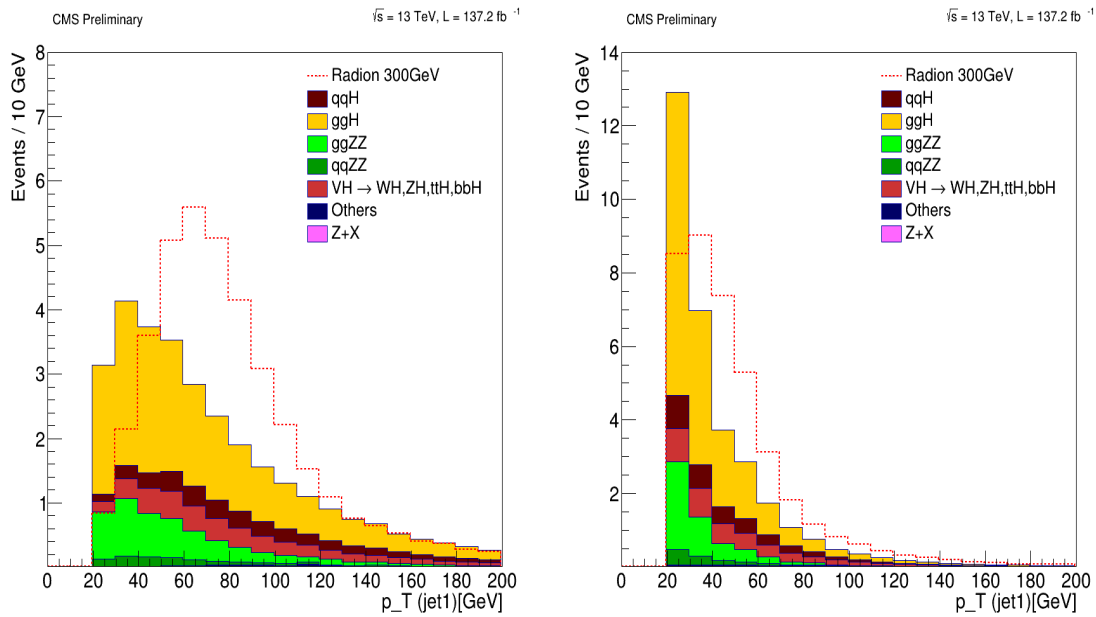


Figure 6.3. . Signal region plots Full RunII (137 fb^{-1}): $p_T^{jet_1}$ (left), $p_T^{jet_2}$ (right)

The yields observed in data are extracted in the signal region of each resonance and its respective mass points for every final state and year of data taking. After the full event selection in our signal region in the

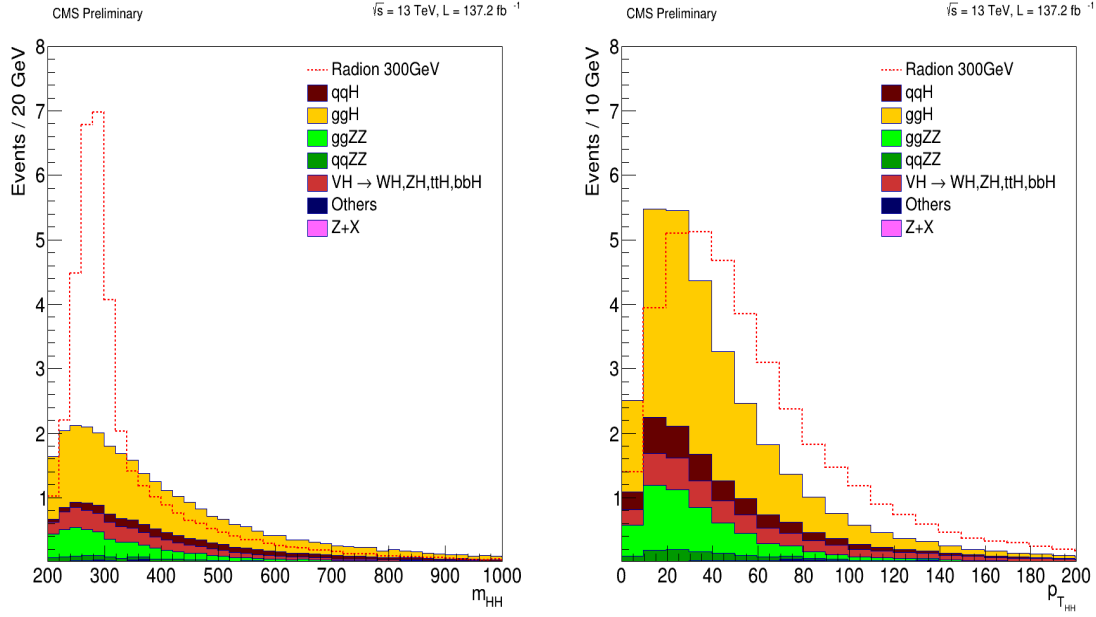


Figure 6.4. . Signal region plots Full RunII ($137 fb^{-1}$): m_{HH} (right), p_T^{HH} (left)

same way done in the previous analysis, the yields then are reported in [Table B.4](#), [Table B.5](#) and [Table B.6](#) for the 2016, 2017, and 2018 datasets, respectively, meanwhile for the full Run II datasets at $\sqrt{s} = 13\text{TeV}$ for the resonant HH analysis, the signal yields are listed in [Table 6.2](#).

Bulk Graviton RunII	260	300	400	500	600	700	800	900	1000
4μ	5.27	3.47	1.15	0.51	0.21	0.11	0.06	0.03	0.02
$4e$	1.92	1.32	0.52	0.24	0.11	0.06	0.03	0.02	0.01
$2e2\mu$	7.45	4.68	1.71	0.78	0.33	0.17	0.09	0.05	0.03
Radion Full RunII	260	300	400	500	600	700	800	900	1000
4μ	17.01	14.10	8.29	5.30	3.57	2.39	1.58	1.13	0.80
$4e$	6.24	5.52	3.63	2.56	1.84	1.29	1.02	0.67	0.52
$2e2\mu$	24.10	19.97	12.20	7.93	5.48	3.80	2.85	1.92	1.37

Table 6.2. . Cut flow table after signal region selection for the Run II dataset ($137.2 fb^{-1}$) for the BulkGraviton spin-2 and Radion spin-0

6.2 Statistical analysis

A statistical analysis is performed in order to evaluate the presence of HH signal in the observed data, or to set an upper limit on its production cross section if no excess over the background predictions is observed. The statistical analysis also aims at setting a constraint on the anomalous Higgs boson self coupling for the non-resonant HH analysis as well as to define an upper limit for both Bulk Graviton and Radion cross section production for the $b\bar{b}ZZ(4l)$ final state with the full Run II datasets.

The statistical analysis reported in this thesis is based on the likelihood-based statistical tests, as described in Ref. [44] and following the prescriptions of the CMS Statistics Committee [125]. The most important steps

used to set our constraints and evaluate upper limits, will be explained below.

6.2.1 Likelihood function $\mathcal{L}$ and Nuisance parameter θ

In order to perform a statistical analysis, a Likelihood function[44] has to be defined as the Probability Density Function (PDF) that characterizes the set of experimental observables considered in the analysis, given the parameters of the model. The model adopted relies on a signal strength modifier parameter μ , that modifies the SM HH cross section by the scale μ and leaves the $HH \rightarrow b\bar{b}4l$ decay branching fraction unchanged. It can be thus defined as the double-Higgs boson rate in the $b\bar{b}4l$ channel over the SM expectation. Each independent source of systematic uncertainty described in section [4,5] is assigned a nuisance parameter θ_i [44], and the full set is denoted as θ . These parameters are not of direct interest to the analysis, but they need to be considered in order to correctly extract the results. The uncertainty on their determination, in fact, will reflect on the parameters of interest of the model. To take into account the probability to measure a value $\tilde{\theta}_i$ for the i -th parameter, given its true value θ_i , the PDF $p_i(\tilde{\theta}_i | \theta_i)$ associated to each nuisance parameter is considered. The PDF for all the nuisance parameters is denoted as $p(\tilde{\theta} | \theta)$. The expected background yields b and the expected signal yield s are functions of the nuisance parameters θ . The likelihood function can thus be defined as:

$$\mathcal{L}(data | \mu, \theta) = \prod_c \mathcal{L}_c(data | \mu \cdot s(\theta) + b(\theta)) \cdot \prod_i p_i(\tilde{\theta}_i | \theta_i) \quad [44] \quad (6.1)$$

where the data represents either the measured data or values from simulated datasets, i.e. sets of random numbers sampled from the expected distributions given by the likelihood function, under certain hypotheses. The first product in Eq. 6.1 runs over all the channels considered in the analysis, i.e. the nine configurations obtained from the three final states ($b\bar{b}4e$, $b\bar{b}4\mu$, $b\bar{b}2e2\mu$) and the three data taking years (2016, 2017, 2018). The $\mathcal{L}_c$ function is the PDF of the events that for binned distributions, like those used in the presented analysis, is given by the product of the Poisson probabilities for every bin j considered to observe n_j events:

$$\mathcal{L}_c(data | \mu \cdot s(\theta) + b(\theta)) = \prod_j \frac{(\mu \cdot s_j(\theta) + b_j(\theta))^{n_j}}{n_j!} \exp \{ -(\mu \cdot s_j(\theta) + b_j(\theta)) \} \quad [44] \quad (6.2)$$

The PDF associated to each nuisance parameter $p_i(\tilde{\theta}_i | \theta_i)$ assumes a different functional form according to the systematic uncertainty type described in section [4]. Normalization uncertainties are associated to log-normal PDFs. Shape uncertainty PDFs, instead, vary according to the different source of uncertainty and are modeled as alternate distributions of the BDT response predicted for each dedicated shape variation. Once the likelihood function is defined, it is possible to test different hypotheses on the basis of the observed data. When two different hypotheses are considered, the first one is called null hypothesis H_0 and the second one is the alternative hypothesis H_1 . Hypothesis testing is based on building a test statistic, a function of the observed data that is derived from the likelihood function and encodes the information about expected signal and background, and uncertainties. The test statistic is a random variable that must be defined in such a way to have different distributions under the two considered hypotheses H_0 , H_1 , in order to discriminate between them. The test statistic is usually built as a ratio of two likelihood functions, one evaluated for the observed data sample under the hypothesis H_0 , and the other under H_1 . The expected distributions followed by the test statistic under the two hypotheses are produced by generating datasets from the PDFs included in the likelihood function. The test statistic is evaluated on the basis of the observed data and its value is compared with the expected distributions to quantify the compatibility of data with the alternative hypothesis.

6.2.2 Quantifying an excess

The test statistic used in our analysis, is based on the statistical significance of an excess q_0 over the background-only expectation, defined as following:

$$q_0 = -2\ln \left[\frac{\mathcal{L}(\text{data} \mid 0, \hat{\theta}_0)}{\mathcal{L}(\text{data} \mid \hat{\mu}, \hat{\theta})} \right], \quad \hat{\mu} \geq 0 \quad (6.3)$$

The numerator is evaluated under the background-only hypothesis ($\mu = 0$) and $\hat{\theta}_0$ is the set of values of nuisance parameters that maximizes it under this null hypothesis. The denominator is evaluated under the alternative signal+background hypothesis, and the values $\hat{\mu}$ and $\hat{\theta}$ are those that maximize the likelihood in the denominator. With this definition q_0 is positive for a signal-like excess ($\mu > 0$), while in absence of an excess $q_0 = 0$ ($\mu = 0$). The significance of an excess can be quantified in terms of the local p -value, defined as the probability to obtain a value of the test statistic q_0 as large as the one observed in experimental data under the background-only hypothesis:

$$p_0 = \mathcal{P} \left(q_0 \geq q_0^{obs} \mid b \right) \quad [44] \quad (6.4)$$

In other words, p_0 characterizes the probability for a local background fluctuation to resemble the hypothesized signal at least as much as observed data do. The p -value is usually converted in the significance Z of the excess through the Gaussian one-sided tail integral:

$$p_0 = \frac{1}{\sqrt{2\pi}} \int_Z^\infty e^{-\frac{x^2}{2}} dx \quad [44] \quad (6.5)$$

The conventional $Z = 5\sigma$ threshold for claiming a discovery corresponds to a p -value of value of $p(5\sigma) = 2.8 \times 10^{-7}$.

6.2.3 Setting an upper limit

If no significant excess is observed, an upper limit is set on the considered signal hypothesis. To analyse this situation another test statistic is used, this time using the signal+background hypothesis in the numerator:

$$q_\mu = -2\ln \left[\frac{\mathcal{L}(\text{data} \mid \mu, \hat{\theta}_\mu)}{\mathcal{L}(\text{data} \mid \hat{\mu}, \hat{\theta})} \right], \quad 0 \leq \hat{\mu} \leq \mu \quad (6.6)$$

where $\hat{\theta}_\mu$ maximizes the numerator under the hypothesis of a signal with strength μ , $\hat{\mu}$, $\hat{\theta}$ maximize the likelihood in the denominator, as before, and correspond to the global maximum of the likelihood. The lower constraint $0 \leq \hat{\mu}$ is required by physics, to have a positive signal rate; the upper constraint $\hat{\mu} \leq \mu$ is imposed to guarantee a one-sided confidence interval, this way upward fluctuations of the data are not considered as evidence against the hypothesis of a signal of strength μ . This definition of the test statistic q_μ grants the possibility of deriving analytically the expected distributions of q_μ under the signal+background and background-only hypotheses, in the asymptotic limit of a large number of background events [44], instead of generating datasets. Exclusion limits in this analysis are computed with a modified frequentist approach, referred to as *CLs* [111], taking the profile likelihood q_μ as test statistic, in the asymptotic approximation. Given an observed value of the test statistic q_μ^{obs} , obtained evaluating Eq. [?] with the observed data, the probability for q_μ to be equal or larger than q_μ^{obs} , under the signal+background and background-only hypotheses, are defined, respectively, as:

$$\begin{aligned}
 CL_{s+b} &= \mathcal{P} \left(q_\mu \geq q_{\mu}^{obs} \mid \mu \cdot s + b \right) \\
 CL_b &= \mathcal{P} \left(q_\mu \geq q_{\mu}^{obs} \mid b \right)
 \end{aligned} \quad [44] \tag{6.7}$$

The quantity CL_s is defined as their ratio:

$$CL_s(\mu) = \frac{CL_{s+b}(\mu)}{CL_b(\mu)} \tag{6.8}$$

A signal of strength μ is excluded at a confidence level of σ , if:

$$CL_s(\mu) < 1 - \sigma \tag{6.9}$$

Usually, and also in this analysis, exclusion limits are computed for $\sigma = 95\%$.

6.2.4 Non-resonant SM HH limits for full Run II

Following the "instructions" regarding the theory explained above, a multi-dimensional fit is performed, using the shape of the BDT response, in order to extract the significance for the SM HH signal and the upper limit at the 95% confidence level CL_s on the signal strength:

$$\mu = \frac{\sigma_{HH}}{\sigma_{HH}^{SM}} \tag{6.10}$$

The BDT response is taken in the signal region:

- $H \rightarrow 4l$ full selection
- $|m_{4l} - 125.09| < 10\text{GeV}$
- at least two isolated jets (with respect to the leptons) in the event

The post-fit distribution of the BDT is shown in [Figure 6.1](#) (bottom) for the full Run II combination. The analysis of the signal $HH \rightarrow b\bar{b}4l$ done with a shape analysis using the full Run2 datasets leads to an expected upper limit on the signal strength:

$$r < 37@95\% \quad CL_s \tag{6.11}$$

and an observed upper limit on the signal strength

$$r < 30@95\% \quad CL_s \tag{6.12}$$

as shown in [Figure 6.5](#) and [Table 6.3](#). The observed and the expected results are compatible in within 2σ uncertainty.

6.2.5 BSM considerations on the SM HH signal: THE k_λ -SCAN limits for full Run II

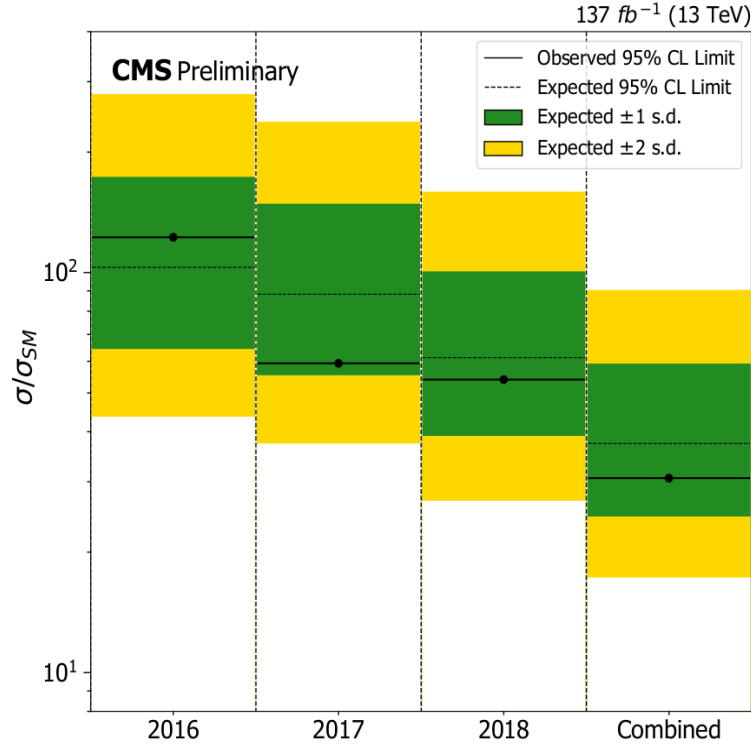
Let's assume that the SM's prediction related to the Higgs self coupling constant λ_{SM} does not fit to the theoretical expectation:

$$\lambda_{SM} \neq \frac{m_H^2}{v^2} \approx 0.13 \tag{6.13}$$

and a hypothetical deviation would be the new reality of the Higgs Physics: What's next? A possible deviation from the SM expectation of the Higgs self coupling constant $\lambda_{SM} \neq 0.13$ would imply to assumptions:

	2016	2017	2018	Combination
Expected	$\mu + 2\sigma$	43	37	27
	$\mu + 1\sigma$	64	55	30
	μ (central)	102	88	61
	$\mu - 1\sigma$	171	146	100
	$\mu - 2\sigma$	276	234	157
Observed	μ	122	59	53

Table 6.3. . Table of limits for all three years and the combination of the full Run II.

Figure 6.5. . Upper limit on the signal strength for each dataset and for the full Run2 combination, where we obtain an observed(expected) upper limit on the signal strength to be about $\mu_{SM}^{HH} < 30(37)@95\% CL_s$.

1. The Higgs boson discovered during Run I would non be the SM Higgs boson, but one the Higgs bosons theoretical predicted from several theories such as the 2HDM theory and more, where those theories might foresee the existence up to 5 (!) Higgs bosons with different Physics properties such spin, electrical charge and more.
2. It would lead to new Physics in general, opening a Pandora box beyond the SM prediction and expectations. The SM already has proven itself not to be a complete theory after the discovery of the Neutrino oscillation, acknowledging that neutrinos indeed do have mass, even though they are very "light".

So, under the assumption that no SM HH signal exists, $HH \rightarrow b\bar{b}4l$, CL_s upper limits on the HH production cross section are derived as a function of $k_\lambda = \lambda_{HHH}/\lambda_{SM}$, where λ_{SM} denotes the SM prediction and λ_{HHH} denotes the measured value. The k_λ -scan is performed for the full RunII dataset in the range $k_\lambda = [-20, 20]$, where:

1. The BSM signal samples within the range $k_\lambda = [-10, 10]$ in step of 1, were produced according to the

detailed procedure explained in chapter [4] for all the full Run II;

2. The samples with the k_λ values beyond the range $k_\lambda = [-10, 10]$, were extrapolated.

The result for full Run II on the k_λ -scan is reported in Figure 6.6. The red line in these plots represents the theoretical value of the cross section. All these results are obtained in the hypothesis where the $k_t = 1$, $c_2 = c_{2g} = c_g = 0$ couplings are assumed to be with the SM values. The analysis of the signal $HH \rightarrow b\bar{b}4l$ done performing a shape analysis using the full Run II dataset is able to constraint k_λ to the observed (expected) range:

$$-9(-10.5) < r(k_\lambda) < 14(15.5) \quad \text{at} \quad 95\%CL_s$$

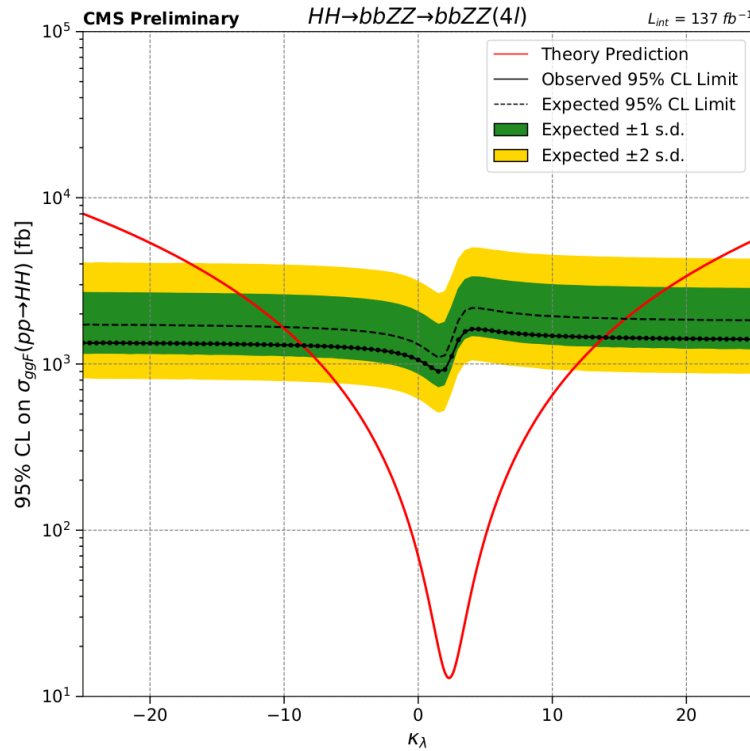


Figure 6.6. . Expected and observed 95%CL_s upper limit on the $HH \rightarrow b\bar{b}4l$ production cross section derived as a function of $k_\lambda = \lambda_{HHH}/\lambda_{SM}$ for the full Run II, with a constraint range at $-9(-10.5) < r(k_\lambda) < 14(15.5)$ at 95%CL_s where the red line represents the theoretical value of the cross section.

6.2.6 Resonant HH : limits for full Run II

Our BSM considerations expand further to more exotic assumptions related to the Higgs Physics, which would lead us to an important question:

"What if the Higgs bosons are the gateway to explain the Gravitation in terms of particle Physics, where they might be a possible final state from a decay of a massive gravitational resonance such as KK Bulk Graviton spin-2 and/or KK Bulk Radion spin-0 ?"

Of course all hypotheses are legitimate to be formulated until the final experimental observation signs the existence of the "correct" theory and all the other assumptions not matching the experimental results will fall to a "dark" region in our memory, where we will forget them forever. Nevertheless, every idea, as long as it

is based on a solid theory and as well as solid experimental infrastructure, such as the CERN experimental facilities do, it is worth it looking for them.

As we already successfully did our exercise above in case of the non-resonant SM HH case and its BSM HH extension due to k_λ -scan, a multi-dimesional fit for the resonant HH is performed. The shape of the BDT response is used in order to extract the significance for the $X(RD/GR) \rightarrow HH$ signal and the upper limit at the 95% confidence level CL_s are set on the signal strength modifier:

$$\mu_{Graviton} = \frac{\sigma_{HH}^{GR}}{\sigma_{HH}^{SM}}, \quad \mu_{Radion} = \frac{\sigma_{HH}^{RD}}{\sigma_{HH}^{SM}} \quad (6.14)$$

The BDT response is taken in the signal region:

- $H \rightarrow 4l$ full selection
- $|m_{4l} - 125.09| < 10\text{GeV}$
- at least two isolated jets (with respect to the leptons) in the event

The analysis of the signal $X(RD/GR) \rightarrow HH \rightarrow b\bar{b}4l$ done with a shape analysis using the full Run II datasets, which leads to preliminary expected upper limits on the signal strength, where the plots of the expected limits during the time period 2016-2017-2018 are shown in [Figure B.1](#) [Figure B.2](#) and [Figure B.3](#), meanwhile after the combination, the full Run II expected limits are plotted in [Figure 6.7](#), [Figure 6.8](#). The respective values are reported in [Table B.7](#), [Table B.8](#) and [Table B.9](#) for each year and in [Table 6.4](#), and [Table 6.5](#) for the Run II combination.

A preliminary result on the expected limits the KK Bul Graviton spin-2 and Radion spin-0 takes form as following:

$$\begin{aligned} r_{HH}^{GR} &\geq 400\text{GeV} \quad \text{at} \quad 95\%CL_s \\ r_{HH}^{RD} &\geq 900\text{GeV} \quad \text{at} \quad 95\%CL_s \end{aligned} \quad (6.15)$$

RunII	260	300	400	600	700	800	900	1000
$\mu + 2\sigma$	0.1683	0.2850	0.6682	2.6646	4.8496	8.3174	14.4053	23.4448
$\mu + 1\sigma$	0.2420	0.4059	0.9703	3.9400	7.2307	12.5360	21.9012	36.0547
μ (central)	0.3779	0.6289	1.5273	6.3750	11.9375	20.8750	37.2500	61.8750
$\mu - 1\sigma$	0.6129	1.0150	2.5136	10.9490	20.8832	37.3503	67.8369	115.6417
$\mu - 2\sigma$	0.9649	1.5829	4.0285	18.3968	36.0593	63.2737	112.9168	187.5863

Table 6.4. . Table of limits for the KK Bulk Graviton spin-2 for full RunII at a total integrated luminosity of $\mathcal{L} = 137fb^{-1}$.

RunII	260	300	400	500	600	700	800	900	1000
$\mu + 2\sigma$	0.0513	0.0699	0.1057	0.1384	0.1777	0.2383	0.2539	0.4285	0.5751
$\mu + 1\sigma$	0.0738	0.1002	0.1523	0.1999	0.2581	0.3496	0.3775	0.6303	0.8461
μ (central)	0.1152	0.1543	0.2354	0.3135	0.4062	0.5547	0.6133	1.0156	1.3633
$\mu - 1\sigma$	0.1869	0.2466	0.3798	0.5109	0.6686	0.9173	1.0386	1.6957	2.2871
$\mu - 2\sigma$	0.2929	0.3805	0.5897	0.8046	1.0671	1.4763	1.7100	2.7511	3.7245

Table 6.5. . Table of limits for the KK Bulk Radion spin-0 for full RunII at a total integrated luminosity of $\mathcal{L} = 137fb^{-1}$.

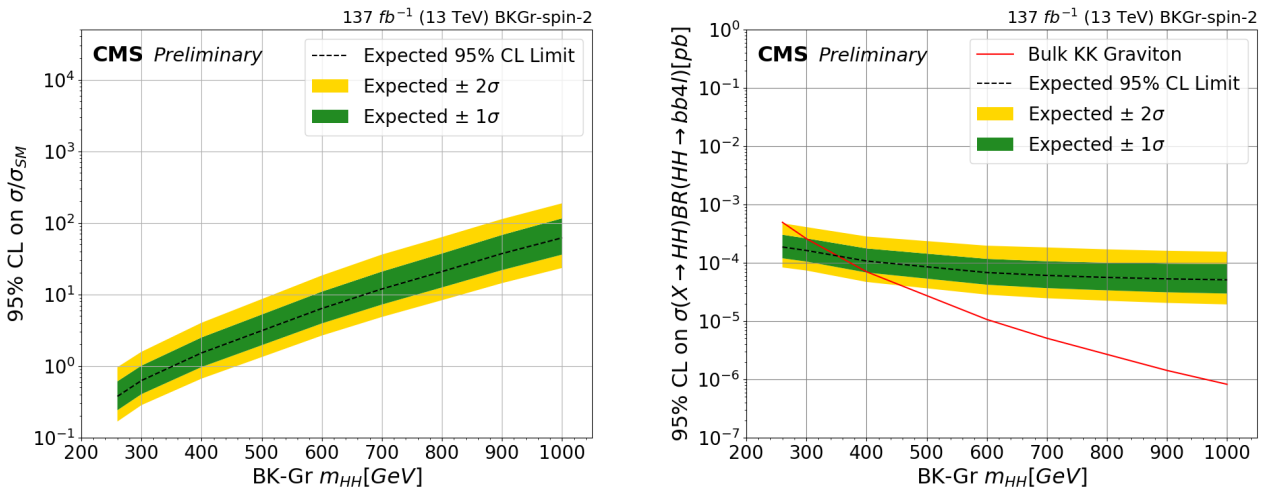


Figure 6.7. . Upper limit on the signal strength (left) and expected 95% CLs upper limit on the cross section as a function of HH invariant mass (right) for the KK Bulk Graviton spin-2 for RunII at $137.2fb^{-1}$. The solid red line shows the theoretical predictions.

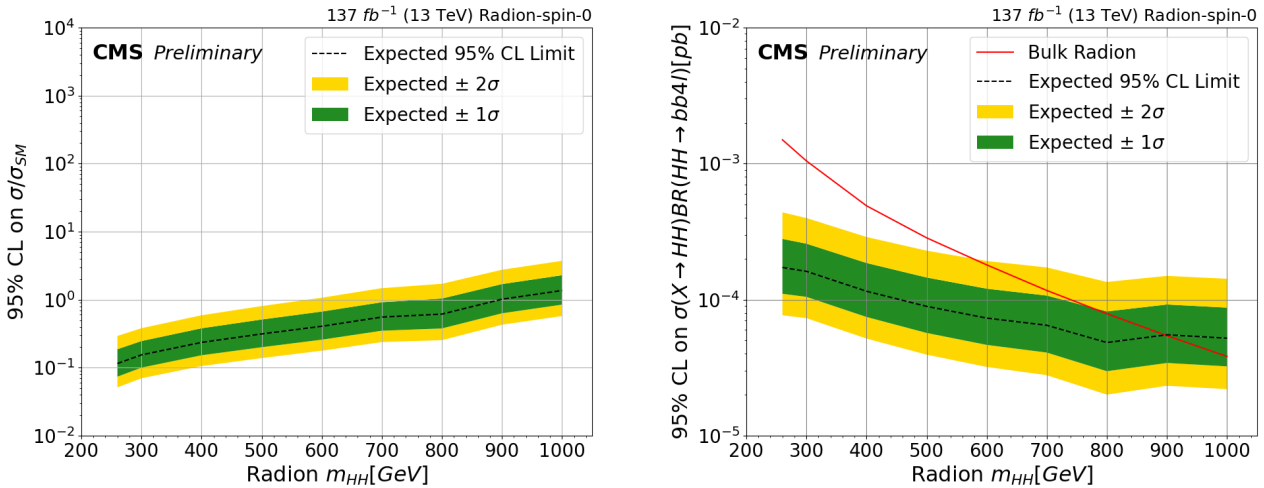


Figure 6.8. . Upper limit on the signal strength (left) and expected 95% CLs upper limit on the cross section as a function of HH invariant mass (right) for the KK Bulk Radion spin-0 for RunII at $137.2fb^{-1}$. The solid red line shows the theoretical predictions.

The expected results are compatible in within 2σ uncertainty. Only in the approval procedure, we will obtain the observed upper limits on the signal strength modifier $\mu_{Graviton}$, μ_{Radion} , to be later compared with the expected limits.

6.3 Projected HL-LHC reach for the HH studies

After the discovery of the Higgs boson in 2012 at CERN, more opportunities of further SM Physics events and maybe beyond, came out in the day-light and people wondered what it could be done with this incredible potential that the LHC is giving us for discovering new particles. One of these main topics was the double Higgs production, the basis of this PhD thesis.

6.3.1 Projected non-resonant HH studies for the HL-LHC

As mentioned in past chapters, the importance of the observation is closely related to the precise measurements to the Higgs self-coupling constant $\lambda_{3H/4H}$, a crucial factor for the final prove related to nature of the Higgs boson itself, since we still have to finally convince ourself that the discovered Higgs boson particle in 2012 is the SM Higgs boson.

The non-resonant $HH \rightarrow b\bar{b}4l$, the basis of this thesis, officially gave its contribution on the HH production to the CMS community as well as the CERN community, by being the first HH search in the CMS collaboration to produce results with the full Run II datasets. Even though the sensitivity of the $b\bar{b}4l$ channel alone is relatively limited with the current available data and statistics of the Run II, that contribution starts to take more importance with the combination of all the other available channels together like the $4b$, $bbWW$, $bb\gamma\gamma$ and more (an effort, which all the research groups of the CMS and ATLAS collaborations are putting on finalizing the full Run II HH combination).

An important question rises from the effort of everybody involved in the HH studies: What about the future? What will we do with those studies? Can we improve them?

The CMS and ATLAS collaboration have published studies related to the possible projections regarding the HH possible studies and expected results for the Phase II during the HL-LHC long data-taking period [157]. These projections are based on two main factors of the HL-LHC:

1. The expected total integrated luminosity $\mathcal{L}$ during the 10 years data-taking should reach the $3000fb^{-1}$ and maybe also it may surpass this value;
2. The LHC should run at a center of mass energy $\sqrt{s} = 14$ TeV.

The results of those projections for the searches of Higgs boson pair production and the measurement of Higgs boson self coupling at the High-Luminosity HL-LHC are obtained considering five decay channels of the HH system:

1. $HH \rightarrow bbbb$,
2. $HH \rightarrow bb\tau\tau$
3. $HH \rightarrow bbWW(l\nu l\nu)$ (with the leptonic decay of the W boson)
4. $HH \rightarrow bb\gamma\gamma$
5. And finally also the contribution $HH \rightarrow bbZZ$ (with Z bosons decaying into pairs of electrons or muons), reported in this thesis.

The combination of these final states is shown to lead to an upper limit on the HH production cross section reaching 0.77 times the SM expectation, and a significance for the signal of 2.6σ for the CMS Phase 2 upgrade, slightly below the threshold usually quoted for an evidence (at 3σ), as reported in Table 6.6. Prospects for the measurement of the Higgs boson self coupling were also studied and upper limits on the HH production cross section derived as a function of $k_\lambda = \lambda_{HHH}/\lambda_{SM}$ are shown in Figure 6.9.

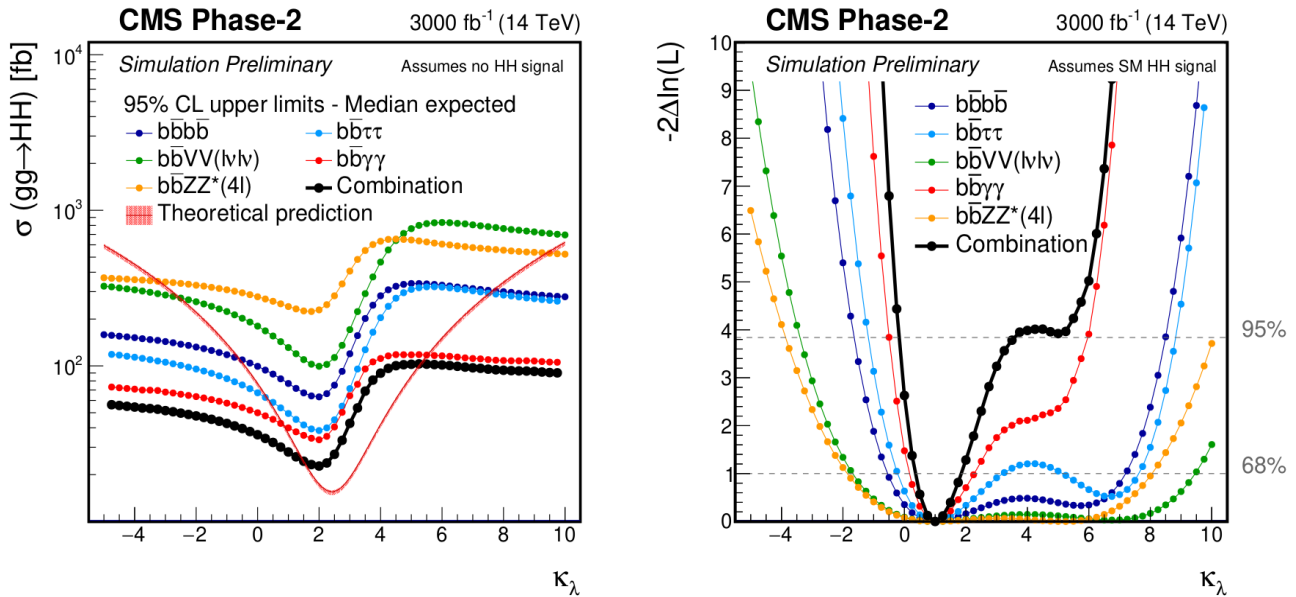


Figure 6.9. . Upper limit at the 95% confidence level CL_s on the HH production cross section as a function of $k_\lambda = \lambda_{HHH}/\lambda_{SM}$. The red band indicated the theoretical production cross section. Right: expected likelihood scan as a function of $k_\lambda = \lambda_{HHH}/\lambda_{SM}$. In both figures the results are shown separately for the five decay channels studied and for their combination. This table is adapted from Ref. [157].

Channel	Significance		95% CL_s limit on $\sigma_{HH}/\sigma_{HH}^{SM}$		Measured signal strength μ	
	Stat. + syst.	Stat. only	Stat. + syst.	Stat. only	Stat. + syst.	Stat. only
$HH \rightarrow b\bar{b}b\bar{b}$	0.95	1.2	2.1	1.6	$1.00^{+1.08}_{-1.04}$	$1.00^{+0.84}_{-0.83}$
$HH \rightarrow b\bar{b}\tau\tau$	1.4	1.6	1.4	1.3	$1.00^{+0.73}_{-0.71}$	$1.00^{+0.65}_{-0.64}$
$HH \rightarrow b\bar{b}WW(l\nu l\nu)$	0.56	0.59	3.5	3.3	$1.00^{+1.8}_{-1.8}$	$1.00^{+1.7}_{-1.7}$
$HH \rightarrow b\bar{b}\gamma\gamma$	1.8	1.8	1.1	1.1	$1.00^{+0.61}_{-0.56}$	$1.00^{+0.56}_{-0.53}$
$HH \rightarrow b\bar{b}ZZ(4l)$	0.37	0.37	6.6	6.5	$1.00^{+3.2}_{-2.5}$	$1.00^{+3.2}_{-2.5}$
Combination	2.6	2.8	0.77	0.71	$1.00^{+0.41}_{-0.39}$	$1.00^{+0.36}_{-0.36}$

Table 6.6. . Upper limit at the 95% confidence level, significance, projected measurement at 68% confidence level of the Higgs boson self coupling λ_{HHH} for the five channels studied and their combination for the CMS-Phase2 projections. Systematic and statistical uncertainties are considered. This table is adapted from Ref. [157].

The interpretation of this result brings us to realize how challenging this task is and that a possible observation of HH at LHC will most require all the contributions possible from all decay channels available and the cooperation of both CMS and ATLAS, by combining the results. A first official preliminary result of the combination between three available channels at the ATLAS collaboration:

1. $HH \rightarrow b\bar{b}b\bar{b}$,

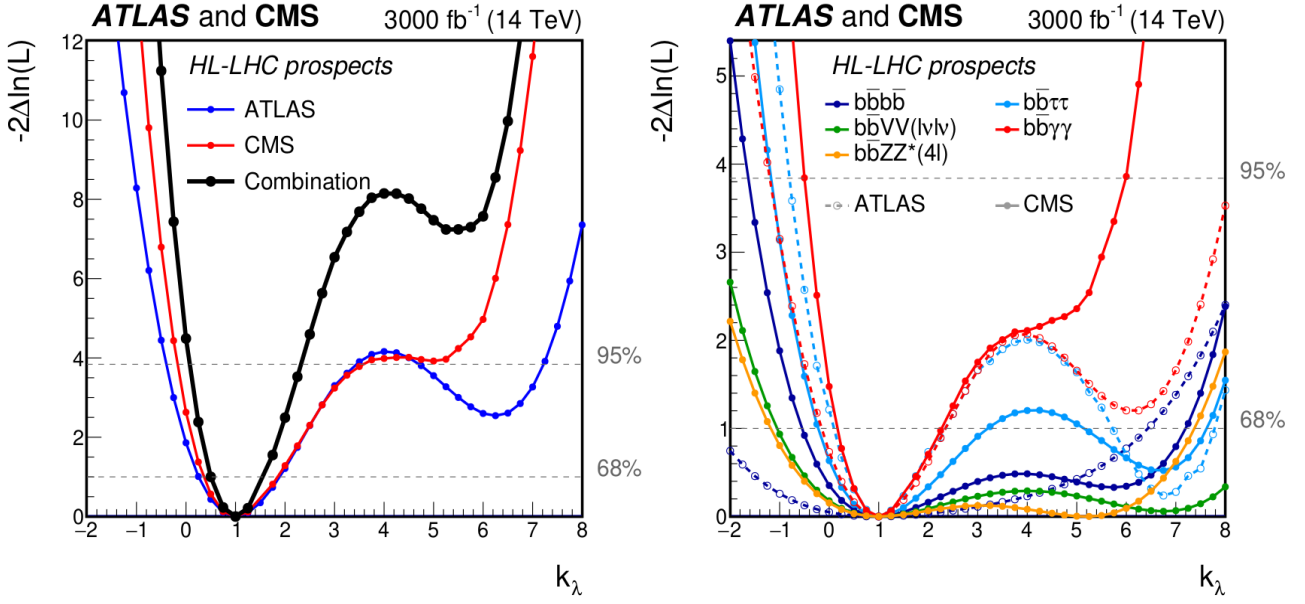


Figure 6.10. . 1) Left: Minimum negative-log-likelihood as a function of k_λ , calculated by performing a conditional signal+background fit to the background and SM signal. The black line corresponds to the combined ATLAS and CMS results, while the blue and red lines correspond to the ATLAS and CMS standalone results respectively. 2) Right: The different colours correspond to the different channels, the plain lines correspond to the CMS results while the dashed lines correspond to the ATLAS results. This table is adapted from Ref. [157].

2. $HH \rightarrow bb\tau\tau$

3. $HH \rightarrow bb\gamma\gamma$

combined with the CMS results, were published [157]. The projections for the HC-LHC at $3000fb^{-1}$ and $\sqrt{s} = 14$ TeV show that we might be close to an observation with a significance for the signal about 4σ , as reported in Table 6.7, meanwhile the prospects for the measurement of the Higgs boson self-coupling constant $\lambda_{3H/4H}$ with the respective upper limits on the HH production cross section derived as a function of $k_\lambda = \lambda_{HHH}/\lambda_{SM}$ of the full ATLAS+CMS combination can be seen in Figure 6.10.

Channel	Statistical-only		Statistical + Systematic	
	ATLAS.	CMS	ATLAS.	CMS
$HH \rightarrow bbbb$	1.4	1.2	0.61	0.95
$HH \rightarrow bb\tau\tau$	2.5	1.6	2.1	1.4
$HH \rightarrow bb\gamma\gamma$	2.1	1.8	2.0	1.8
$HH \rightarrow bbVV(l\nu\nu)$	-	0.59	-	0.56
$HH \rightarrow bbZZ(4l)$	-	0.37	-	0.37
Combination	3.5	2.8	3.0	2.6
ATLAS + CMS combined	4.5		4.0	

Table 6.7. . Significance in standard deviations of the individual channels as well as their combination. This table is adapted from Ref. [157].

6.3.2 Thoughts on resonant HH studies for the HL-LHC

The idea of the existence of Physics beyond the SM expectation is motivated due to several experimental results like the Neutrino oscillation proving the mass of such particles, or cosmic observations of rotating galaxies. The rotational velocity of stars even in the very end of a galaxy's tail was measured to remain constant, instead of decreasing, while moving away from its center. This constant rotational velocity even in the far ends of the tails of galaxies, while moving from the center, it led to the idea of the presence of dark "stuff", called as dark energy and dark matter.

The motivation of finding more particles responsible for the dark matter, SUSY (SUpEr SYmmetry) particles or even the discovery of a gravitational particle, is based of the possibility to unify quantum mechanics and the general relativity theory to one unique unified quantum gravitational theory, such as String theory or the Grand Unified Theory. This possibility is motivating a large number of research groups from everywhere to join their forces and to profit from the LHC experimental potentials with the hope of discovering new particles.

The variant of a di-Higgs system being proliferated as a decay mode of a massive resonance, has strong theoretical basis, supported by the Kaluza-Klein theory and the implementation of L. Randall and R. Sundrum, but also from the experimental observation of gravitational waves coming from the collision of two massive black holes, happened thirteen billion years ago, which shook the space-time. The solutions of this theory bring us to gravitational particles like the KK Bulk Graviton spin-2 and Radion spin-0.

In this thesis, we gave preliminary results on the expected limits on production cross section related to both variants in our $b\bar{b}4l$ channel with the full Run II datasets, where we might be able from this preliminary results to exclude an invariant mass of 400 GeV for the Bulk Graviton and much more for the Radion, where the we may start looking from 900 GeV and more. Since our channel in general has a limited sensitivity due to a small branching ratio of the Higgs boson itself into a leptonic final states the expectations would imply that these numbers could reach higher values, even the TeV region. In order to be exploited this study to its full potential This study, it needs more support from theoretical research groups to contribute for MC models at higher accuracy levels such as NLO or even NNLO with including different MC generators. Both ATLAS and CMS research groups have to collaborate for a combined analysis on different decay channels, in order to get a better experimental evaluation of possible gravitational particles, an effort, which was not saturated yet. Maybe, a task to be considered with more passion during Run III and specially during the HL-LHC.

Part VI

Summary and Conclusions

Conclusions

The discovery of the Higgs boson and the measurement of its invariant mass, has triggered a chain of hypothesis and new opportunities in the scalar sector of the SM theory. The lack of the ability of the SM to predict the invariant mass of the Higgs boson can be considered as an invitation to possible BSM topologies regarding the HH production such as the anomalous coupling constant and an hypothetical deviation of the Higgs boson self-coupling λ_{HHH} value.

In this thesis the double Higgs production has been thoroughly discussed and analysis results have been reported for the full Run II data taking at $\sqrt{s} = 13$ TeV at a total integrated luminosity of $\mathcal{L} = 137.2 \text{ fb}^{-1}$. We have considered the SM HH production at NLO accuracy with the cross sections and signal sample generated with the MC generator POWHEG-BOX-V2 with full top quark mass m_t dependence and we had courage enough to even consider a extensive scan over the k_λ value for the full Run II period at the CMS experiment.

Furthermore, we have also looked for possible new massive gravitational resonances such as the KK Bulk Graviton spin-2 and Radion spin-0, motivated from the Kaluza-Klein theory and the Randall-Sundrum model of a fifth extra dimension in the space time. For this type of analysis, we generated a full mass point scan from 260 GeV up to 1 TeV for each year and both resonances.

The search for the HH production in the $b\bar{b}ZZ(4l)$ decay channel has been developed and optimized for the full Run II conditions. A dedicated event categorization was put in place to simultaneously probe the resonant and non-resonant production mechanisms for a variety of signal hypotheses. In case of the non-resonant HH studies, the results have shown that at the 95% CL_s the upper limits on the signal strength modifier $\mu = \frac{\sigma_{HH}}{\sigma_{HH}^{SM}}$ for the observed upper limit are about $30 \times \sigma_{HH}$, while for the expected upper limit, we measured about $37 \times \sigma_{HH}$.

In addition to that, upper limits were also set for the BSM topology of the non-resonant HH production, which belong to the hypothetical deviation of the value of the Higgs self-coupling constant λ_{HHH} . Under the assumption that no SM HH signal exists, $HH \rightarrow b\bar{b}4l$ CL_s upper limits on the HH production cross section are derived as a function of $k_\lambda = \lambda_{HHH}/\lambda_{SM}$, where λ_{SM} denotes the SM prediction and λ_{HHH} denotes the measured value. The k_λ -scan is performed for the full RunII datasets in the range $k_\lambda = [-20, 20]$, where the we were able to exclude k_λ values regarding $k_\lambda > -9$ and $k_\lambda < 14$ for the observed case, meanwhile, while the expected upper limits were set between $k_\lambda > -10.5$ and $k_\lambda < 15.5$.

These results are regarded in the broader context of CMS HH searches due to the CMS full combination with other channels at 13 TeV.

The results of the resonant variant of the $HHb\bar{b}ZZ(4l)$ for both KK Bulk Graviton spin-2 and Radion spin-0 have been preliminary presented as a function of the invariant mass of both gravitational resonances from 260 GeV up to 1 TeV. The expected upper limit were set and these values are:

$$\begin{aligned}
r_{HH}^{GR} &\geq 400 GeV \quad at \quad 95\%CL_s \\
r_{HH}^{RD} &\geq 900 GeV \quad at \quad 95\%CL_s
\end{aligned}$$

As this thesis comes to touch its end after three years, the exploration of the TeV scale has just started during the Run II. The data, which will be collected at the LHC during Run III and the subsequent operations of the HL-LHC, might give us a closer look to possible answers of open questions on the SM and on its possible BSM extensions, hopefully providing us a better idea of possible BSM considerations regarding the anomalous coupling constants for non-resonant HH production, but also the possible existence of gravitational resonances, which would decay into a pair of Higgs bosons.



Appendix A

Supporting information for chapter 1

This theoretical part is adapted from Ref. [82], [81] and [66], where all the information needed related the LO, NLO, NNLO and further accuracies on the non-resonant SM HH production with full top quark mass m_s dependence from the SM Effective Field Theory can be found, in order to understand the theory behind the MC generations of the signal samples.

A.1 Effective Field Theory and QCD corrections for ggF-HH production

The EFT approach for a non-resonant Higgs pair production is well motivated from the theoretical point of view. For the double Higgs boson production in gluon fusion, in case we also take in consideration CP-violation, the full EFT Lagrangian would be as following:

$$\begin{aligned} \mathcal{L}_{SM} = & \left(c_g \frac{h}{v} + c_{gg} \frac{h^2}{2v^2} \right) \frac{g_s^2}{4} G_{\mu\nu}^A G^{A,\mu\nu} - \frac{h}{v} \sum_f \Sigma_{ij} \sqrt{m_{f_i} m_{f_j}} [\delta y_f]_{ij} \bar{f}_i f_j \\ & - \frac{h^2}{2v^2} \sum_f \Sigma_{ij} \sqrt{m_{f_i} m_{f_j}} [\delta y_f^{(2)}]_{ij} [\bar{f}_{iR} f_{jR} + hc] + \delta\lambda h^3 + \mathcal{L}_{CPV} \end{aligned} \quad (16)$$

where i, j are generation indices, the $\sum_f$ is over all charged fermion species and $\mathcal{L}_{CPV}$:

$$\begin{aligned} \mathcal{L}_{CPV} = & \left(\tilde{c}_g \frac{h}{v} + \tilde{c}_{gg} \frac{h^2}{2v^2} \right) \frac{g_s^2}{4} G_{\mu\nu}^A \tilde{G}^{A,\mu\nu} \\ & - \frac{h}{v} \sum_f \Sigma_{ij} \sqrt{m_{f_i} m_{f_j}} [\delta y_f]_{ij} \bar{f}_i f_j \left\{ \left(\cos\phi_{ij}^f - 1 \right) \bar{f}_i f_j - i \sin\phi_{ij}^f \bar{f}_i \gamma_5 f_j \right\} \end{aligned} \quad (17)$$

The hierarchy between CP-violating and CP-conserving interactions has already been established. For simplicity, we assume CP conservation and flavour diagonal Higgs boson couplings, obtaining the relevant Lagrangian for di-Higgs boson production:

$$\begin{aligned} \mathcal{L} = & \mathcal{L}_{SM} + \left(c_g \frac{h}{v} + c_{gg} \frac{h^2}{2v^2} \right) \frac{g_s^2}{4} G_{\mu\nu}^A G^{A,\mu\nu} - \frac{h}{v} \sum_f \Sigma_{ij} \sqrt{m_{f_i} m_{f_j}} [\delta y_f]_{ij} \bar{f}_i f_j \\ & - \frac{h^2}{2v^2} \sum_f \Sigma_{ij} \sqrt{m_{f_i} m_{f_j}} + \delta\lambda h^3 \end{aligned} \quad (18)$$

We do not consider enhanced b -couplings, since the contribution for ggF-HH radiation and/or generation from b -quark loops is less than 1%. The NLO QCD corrections to the gluon-fusion cross section $\sigma(gg \rightarrow HH)$ are obtained in the heavy top quark limit (HTL). The top quark loop contributions are reduced to effective couplings between the Higgs boson and gluons, described by the effective Lagrangian:

$$\mathcal{L}_{eff} = \frac{\alpha_s}{12\pi} G_{\mu\nu}^A G^{A,\mu\nu} \left(C_1 \frac{h}{v} + C_2 \frac{h^2}{2v^2} \right) \quad (19)$$

where:

$$\begin{aligned}
C_1 &= 1 + \frac{11}{4} \frac{\alpha_s}{\pi} + \left\{ \frac{2777}{288} + \frac{19}{16} L_t + N_F \left(\frac{L_t}{3} - \frac{67}{69} \right) \right\} \left(\frac{\alpha_s}{\pi} \right)^2 \\
C_2 &= C_1 + \left(\frac{35}{24} + \frac{2}{3} N_F \right) \left(\frac{\alpha_s}{\pi} \right)^2 + \mathcal{O}(\alpha^3) \\
L_t &= \log \left(\frac{\mu_R^2}{m_t^2} \right) \quad \text{Wilson coefficients}
\end{aligned} \tag{20}$$

that are presented up to NNLO, but are known up to N⁴LO. Since the top quark is integrated, the number of active flavors has to be taken as $N_F = 5$. Using these effective Higgs couplings to gluons, the calculation of the NLO QCD corrections is reduced to a one-loop calculation for the virtual corrections and a tree-level calculation for the matrix elements of the real corrections. The NLO final result for the total gluon-fusion cross section can be decomposed as :

$$\sigma_{NLO}(pp \rightarrow HH + X) = \sigma_{LO} + \Delta\sigma_{virt} + \Delta\sigma_{gg} + \Delta\sigma_{gq} + \Delta\sigma_{q\bar{q}} \tag{21}$$

where the different contributions take the following form:

$$\begin{aligned}
\sigma_{LO} &= \int_{\tau_0}^1 d\tau \frac{d\mathcal{L}^{gg}}{d\tau} \hat{\sigma}_{LO}(Q^2 = \tau s) \\
\Delta\sigma_{virt} &= \frac{\alpha_s(\mu_R)}{\pi} \int_{\tau_0}^1 d\tau \frac{d\mathcal{L}^{gg}}{d\tau} \hat{\sigma}_{LO}(Q^2 = \tau s) C_{virt} \\
\Delta\sigma_{gg} &= \frac{\alpha_s(\mu_R)}{\pi} \int_{\tau_0}^1 d\tau \frac{d\mathcal{L}^{gg}}{d\tau} \int_{\tau_0/\tau}^1 \frac{dz}{z} \hat{\sigma}_{LO}(Q^2 = z\tau s) \\
&\quad \left\{ -z P_{gg} \log \frac{\mu_F^2}{\tau s} + d_{gg}(z) + 6 \left[1 + z^4 + (1-z)^4 \right] \left(\frac{\log(1-z)}{1-z} \right) \right\} \\
\Delta\sigma_{gq} &= \frac{\alpha_s(\mu_R)}{\pi} \int_{\tau_0}^1 d\tau \sum_{q, \bar{q}} \frac{d\mathcal{L}^{gq}}{d\tau} \int_{\tau_0/\tau}^1 \frac{dz}{z} \hat{\sigma}_{LO}(Q^2 = z\tau s) \\
&\quad \left\{ -\frac{z}{2} P_{gq}(z) \log \frac{\mu_F^2}{\tau s(1-z)^2} + d_{gq}(z) \right\} \\
\Delta\sigma_{q\bar{q}} &= \frac{\alpha_s(\mu_R)}{\pi} \int_{\tau_0}^1 d\tau \sum_q \frac{d\mathcal{L}^{q\bar{q}}}{d\tau} \int_{\tau_0/\tau}^1 \frac{dz}{z} \hat{\sigma}_{LO}(Q^2 = z\tau s) d_{q\bar{q}}(z)
\end{aligned} \tag{22}$$

Here $\hat{\sigma}_{LO}(Q^2)$ denotes the leading-order partonic cross section involving the squared invariant mass Q^2 of the Higgs boson pair, $\alpha_s(\mu_R)$ the strong coupling constant at the renormalization scale μ_R , $d\mathcal{L}^{ij}/d\tau$, ($i, j = g, q, \bar{q}$) corresponding parton-parton luminosities at the factorization scale μ_F , and $P_{ij}(z)$ ($i, j = g, q, \bar{q}$) the individual Altarelli–Parisi splitting functions. The integration regions are bound by $\tau_0 = 4m_H^2/s$, with m_H being the Higgs boson mass and s the square of the hadronic center-of-mass energy. The quark-mass dependence is in general encoded in the green factors $\hat{\sigma}_{LO}(Q^2)$ for the LO cross section and the red factors C_{virt} , $d_{ij}(z)$ for the virtual and real corrections, respectively. In the HTL, the latter simplify to:

$$\begin{aligned}
C_{virt} &\rightarrow \frac{11}{2} + \pi^2 + C_{\Delta\Delta}^\infty + \left(\frac{33 - 2N_F}{6} \right) \log \left(\frac{\mu_R^2}{Q^2} \right) \\
C_{\Delta\Delta} &= \text{Re} \frac{\int_{\hat{t}_-}^{\hat{t}_+} d\hat{t} \left\{ c_1 \left[(C_\Delta F_\Delta + F_\square) + \frac{p_T^2}{\hat{t}} G_\square \right] + (\hat{t} \leftrightarrow \hat{u}) \right\}}{\int_{\hat{t}_-}^{\hat{t}_+} d\hat{t} \{ |C_\Delta F_\Delta + F_\square|^2 + |G_\square|^2 \}} \\
C_{\Delta\Delta}^\infty &= C_{\Delta\Delta} |_{c_1=2/9} \\
d_{gg}(z) &\rightarrow \frac{11}{2} (1-z)^3 \\
d_{gq}(z) &\rightarrow \frac{2}{3} z^2 - (1-z)^2 \\
d_{q\bar{q}}(z) &\rightarrow \frac{32}{27} (1-z)^3
\end{aligned} \tag{23}$$

where p_T denotes the Higgs transverse momentum, $\hat{s}, \hat{t}$ the partonic Mandelstam variables and $C_{\Delta\Delta}$ is the contribution of the one-particle reducible diagrams. The integration bounds are given by:

$$\hat{t} = -\frac{1}{2} \left[Q^2 - 2m_H^2 \mp Q^2 \sqrt{1 - \frac{4m_H^2}{Q^2}} \right] \tag{24}$$

The couplings $C_\Delta, C_\square$ and the form factors $F_\Delta, F_\square$ and $G_\square$ in the HTL take the form:

$$\begin{aligned}
C_\Delta &= \lambda_{HHH} \frac{6v}{\hat{s} - m_H^2 + im_H \Gamma_H}, \quad C_\square = 1 \\
F_\Delta &\rightarrow \frac{2}{3}, \quad F_\square \rightarrow -\frac{2}{3} \\
G_\square &\rightarrow 0
\end{aligned} \tag{25}$$

with the trilinear coupling $\lambda_{HHH} = m_H^2/(2v^2)$. For the NLO QCD corrections, the full mass dependence of the LO partonic cross section has been taken into account first, while treating the virtual corrections C_{virt} and the real corrections d_{ij} in the HTL. This approach is now called "Born-improved". This leads to a reasonable approximation for invariant Higgs pair masses in the lower range and approximates the full NLO result for the total cross section within about 15%. The NLO corrections in the HTL increase the cross section by 80 – 90%. The NNLO QCD corrections have been calculated within the same approximation. The main part of these corrections can be translated from the single Higgs case, since the effective Lagrangian does not induce a change of kinematics between single Higgs production and the differential Higgs pair production with respect to the invariant Higgs pair mass. The new ingredient of the NNLO calculation is the proper treatment of the one-particle-reducible contributions $C_{\Delta\Delta}$ of the NLO corrections. These lead to additional contributions to the NNLO virtual corrections and the interference of the NLO-real and NLO-virtual corrections that contributes to the NNLO result, too. Including the full mass dependence of the LO cross section, the NNLO QCD corrections increase the total cross section by a more moderate amount of 20 – 30%. On top of these NNLO QCD corrections, the soft-gluon resummation (threshold resummation) has been performed at next-to-next-to-leading logarithmic accuracy (NNLL) for the total cross section and invariant mass distribution, resulting in a $\mathcal{O}(10\%)$ modification of the total cross section on top of the NNLO result for a central scale $\mu_R = \mu_F = m_{HH}$, while the effects are much smaller if the scale $\mu_R = \mu_F = m_{HH}/2$ is used. The calculations in the HTL have been refined by several steps including mass effects partially at NLO. The inclusion of the full mass effects in the real correction terms d_{ij} by means of incorporating the full one-loop real matrix elements for $gg \rightarrow HHg, gq \rightarrow HHq, q\bar{q} \rightarrow HHq$ reduces the Bornimproved

HTL prediction by about 10% . This improvement is denoted as “FTapprox” (for full-theory approximation). This step has been performed by using the MG5_aMC@NLO framework for the automatic generation of the matrix elements. Another improvement has been accomplished by a systematic asymptotic large-top-mass expansion of the full NLO corrections at the integral and at the integrand level. This established sizeable mass effects emerging from the virtual two-loop corrections. In addition, the large-top-mass expansion has been extended to NNLO resulting in expected 5% mass effects of the NNLO corrections on top of the NLO result . This situation necessitated the full calculation of the mass effects at NLO. The full NLO QCD corrections have been derived by two quite different methods, both, however, building on a numerical integration of the two-loop contributions that cannot be integrated analytically with present state-of-the-art methods. It can be decomposed into triangle, one-particle-reducible and box diagrams. The triangle diagrams can be obtained from the analogous calculation for $gg \rightarrow H$ with the Higgs mass replaced by the invariant Higgs pair mass $Q = m_{HH}$. The result of the one-particle-reducible diagrams is known analytically:

$$\begin{aligned} c_1 &= [I_1(\tau, \lambda_i) - I_2(\tau, \lambda_i)]^2 \\ I_1(\tau, \lambda) &= \frac{\tau\lambda}{2(\tau - \lambda)} + \frac{(\tau\lambda)^2}{2(\tau - \lambda)^2} [f(\tau) - f(\lambda)] + \frac{\tau^2\lambda}{(\tau - \lambda)^2} [g(\tau) - g(\lambda)] \\ I_2(\tau, \lambda) &= -\frac{\tau\lambda}{2(\tau - \lambda)} [f(\tau) - f(\lambda)] \end{aligned} \quad (26)$$

with $\tau = \frac{4m_t^2}{m_H^2}$, $\lambda_i = \frac{4m_t^2}{h_{att}^2}$ and the functions:

$$\begin{aligned} f(\tau) &= \begin{cases} \arcsin^2\left(\frac{1}{\sqrt{\tau}}\right) & \text{if : } \tau \geq 1 \\ -\frac{1}{4} \left[\log \frac{1 + \sqrt{1 - \tau}}{1 - \sqrt{1 - \tau}} - i\pi \right]^2 & \text{if : } \tau < 1 \end{cases} \\ g(\tau) &= \begin{cases} \sqrt{\tau - 1} \arcsin\left(\frac{1}{\sqrt{\tau}}\right) & \text{if : } \tau \geq 1 \\ \frac{\sqrt{1 - \tau}}{2} \left[\log \frac{1 + \sqrt{1 - \tau}}{1 - \sqrt{1 - \tau}} - i\pi \right] & \text{if : } \tau < 1 \end{cases} \end{aligned}$$

This expression has to be inserted in the $C_{\Delta\Delta}$ coefficient of Eq. (1.4). The inputs in this realization of the EFT can thus be taken in general as:

$$\rightarrow c_g, c_{gg}, \delta y_t, y_t^{(2)}, \delta\lambda_3$$

The combination of c_g , δy_t is fixed by the requirement that single Higgs boson production has the experimentally observed value:

$$R_h \equiv \frac{\sigma(gg \rightarrow h)}{\sigma(gg \rightarrow h)_{SM}} \sim |1 + 12\pi^2 c_g + \delta y_t|^2 \quad (27)$$

The couplings $c_g, y_t^{(2)}, \delta\lambda_3$ cannot be probed in single Higgs boson production, but they require experimental measurement of the di-Higgs rate and distributions . A fit to the total cross section in terms of the EFT coefficients shown here. Those references construct a cross section fit in terms of effective Higgs boson couplings, as:

$$\begin{aligned}\mathcal{L}^{hh} = & \frac{1}{2}\partial_\mu h\partial^\mu h - \frac{m_h^2}{2}h^2 - \kappa_\lambda\lambda_{SM}vh^3 - \frac{m_t}{v}\left(v + \kappa_t h + \frac{c_2}{v}hh\right)(\bar{t}_L t_R + h.c.) \\ & + \frac{\alpha_s}{12\pi v}\left(c_{1g}h - \frac{c_{2g}}{2v}hh\right)G_{\mu\nu}^A G^{A,\mu\nu}\end{aligned}\quad (28)$$

The SM limit is obtained when the condition $\kappa_2 = \kappa_\lambda = 1$ and $c_2 = c_{1g} = c_{2g} = 0$ is true. This fit can be straightforwardly mapped onto the EFT parameters of Eq. (*) via the identities:

- $c_g = \frac{c_{1g}}{12\pi^2} = \frac{C_1}{12\pi^2}$
- $c_{gg} = -\frac{c_{2g}}{12\pi^2} = -\frac{C_2}{12\pi^2}$
- $y_t^{(2)} = 2c_2$
- $\delta y_t = (\kappa_t - 1)$
- $\delta y_3 = -v(\kappa_\lambda - 1)\lambda_{SM}$

Further information on the EFT coefficients can be found from hh production by noting that different EFT operators have different kinematic dependences. The LO box and triangle diagram exactly cancel each other at threshold in the SM. This implies that $\frac{d\sigma}{dM_{hh}}$ is most sensitive to variations in κ_t, κ_λ at threshold, while the dependence on κ_λ is suppressed at high partonic energies. The NLO corrections to the EFT predictions for double Higgs boson production have been investigated in the large m_t limit, with the conclusion that the K factor of the EFT shows little kinematic dependence and little dependence on the effective couplings, however with the same caveats. We can take advantage of this property of the K-factors, approximating the ratio between the cross sections obtained for different EFT parameters and the SM cross section with the corresponding LO ratio:

$$\begin{aligned}R_{HH} \equiv \frac{\sigma_{HH}}{\sigma_{HH}^{SM}} = & A_1\kappa_t^4 + A_2c_2^2 + (A_3\kappa_t^2 + A_4c_g^2)\kappa_\lambda^2 + A_5c_{2g}^2 + (A_6c_2 + A_7\kappa_t\kappa_\lambda)\kappa_t^2 \\ & + (A_8\kappa_t\kappa_\lambda + A_9c_g\kappa_\lambda)c_2 + A_{10}c_2c_{2g} + (A_{11}c_g\kappa_\lambda + A_{12}c_{2g})\kappa_t^2 \\ & + (A_{13}\kappa_\lambda c_g + A_{14}c_{2g})\kappa_t\kappa_\lambda + A_{15}c_gc_{2g}\kappa_\lambda\end{aligned}\quad (29)$$

The A_i coefficients are extracted from a simultaneous fit, based on the maximization of a likelihood, to the cross sections obtained from a LO simulation and provided in Table 61. A detailed study of theoretical uncertainties was performed in Ref. . The uncertainties related to PDF and $\alpha_s(Q^2)$ variations induces less than a 2% variation in the A_i values. In general, M_{HH} is a sensitive variable for differentiating the effects of the different EFT coefficients. In the vicinity of the SM limit the utility of this approach, however, is limited by the small di-Higgs rate, although it can nevertheless provide essential additional information on the various effective couplings, in particular in the case of an enhanced cross section. There are large correlations, typically requiring a global fit rather than a fit to a single EFT coefficient . To this end, the analysis divides kinematic points into 12 clusters, scanning over a range of EFT coefficients, where within the clusters the dependence on the kinematic parameters is similar. The clusters have clear kinematic differences between them, particularly in the peak structure of the M_{hh} distributions. Scanning over a range of EFT coefficients this procedure allows us to formulate a set of benchmark choices that exhibit a particularly interesting di-Higgs phenomenology and to exploit the kinematical particularities of a large portion of the parameter space with a limited number of analyses. In summary, to obtain σ_{HH} in the EFT, we recommend the use of the relation:

$$\sigma_{HH} = R_{HH} \cdot \sigma_{HH}^{SM} \quad (30)$$

Here σ_{hh} is the most up to date calculation based on NNLO+NNLL, with the associated uncertainties. The PDF and α_s uncertainties A_i was found to be well below 1% and can safely be neglected. The missing order uncertainties on R_{HH} are covered by the ones SM assigned to σ_{hh} .

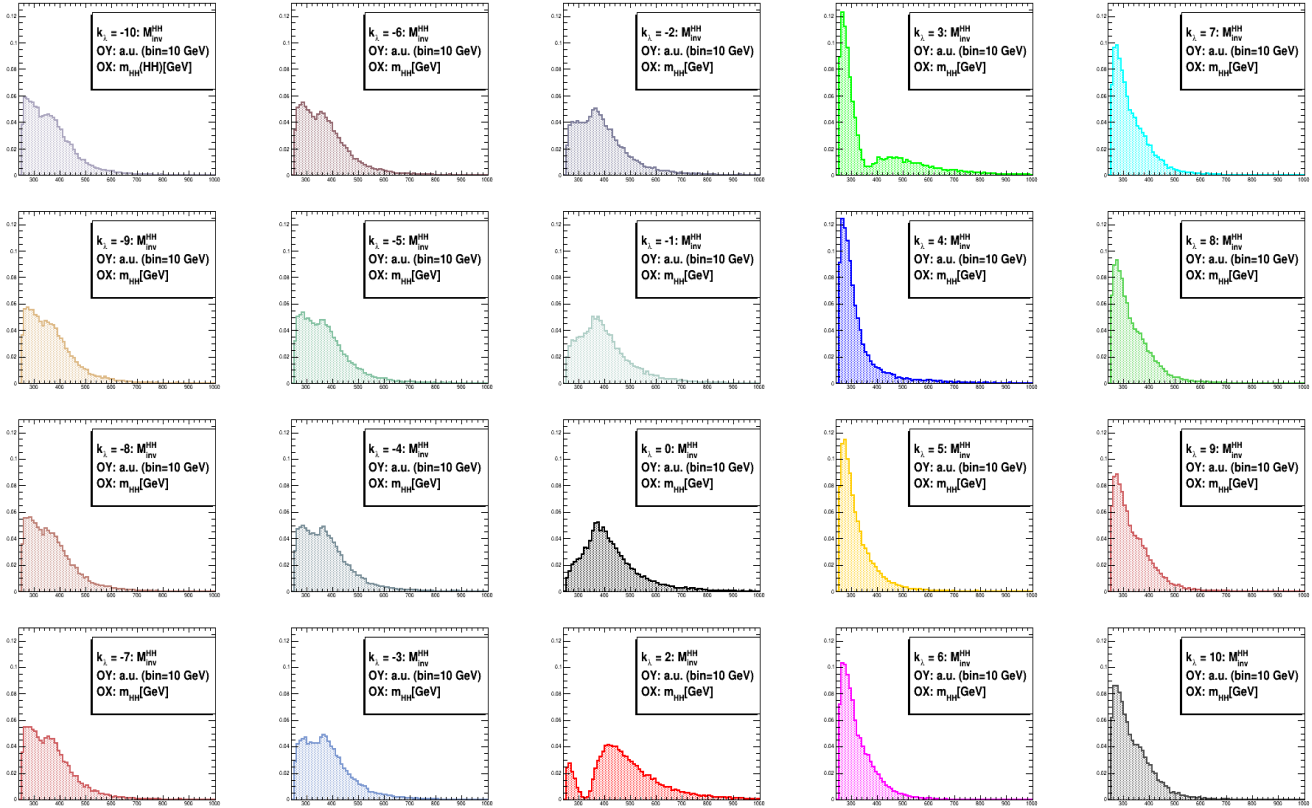
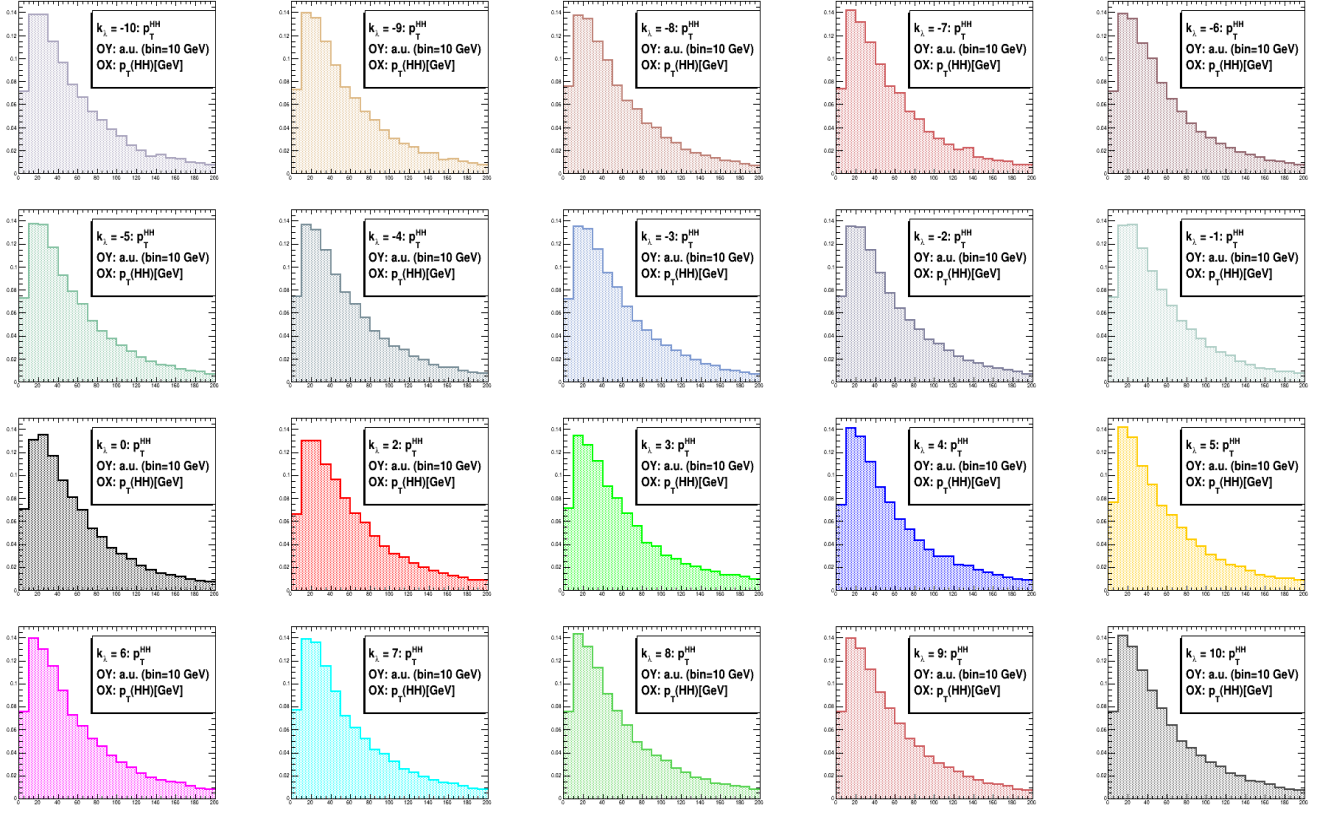
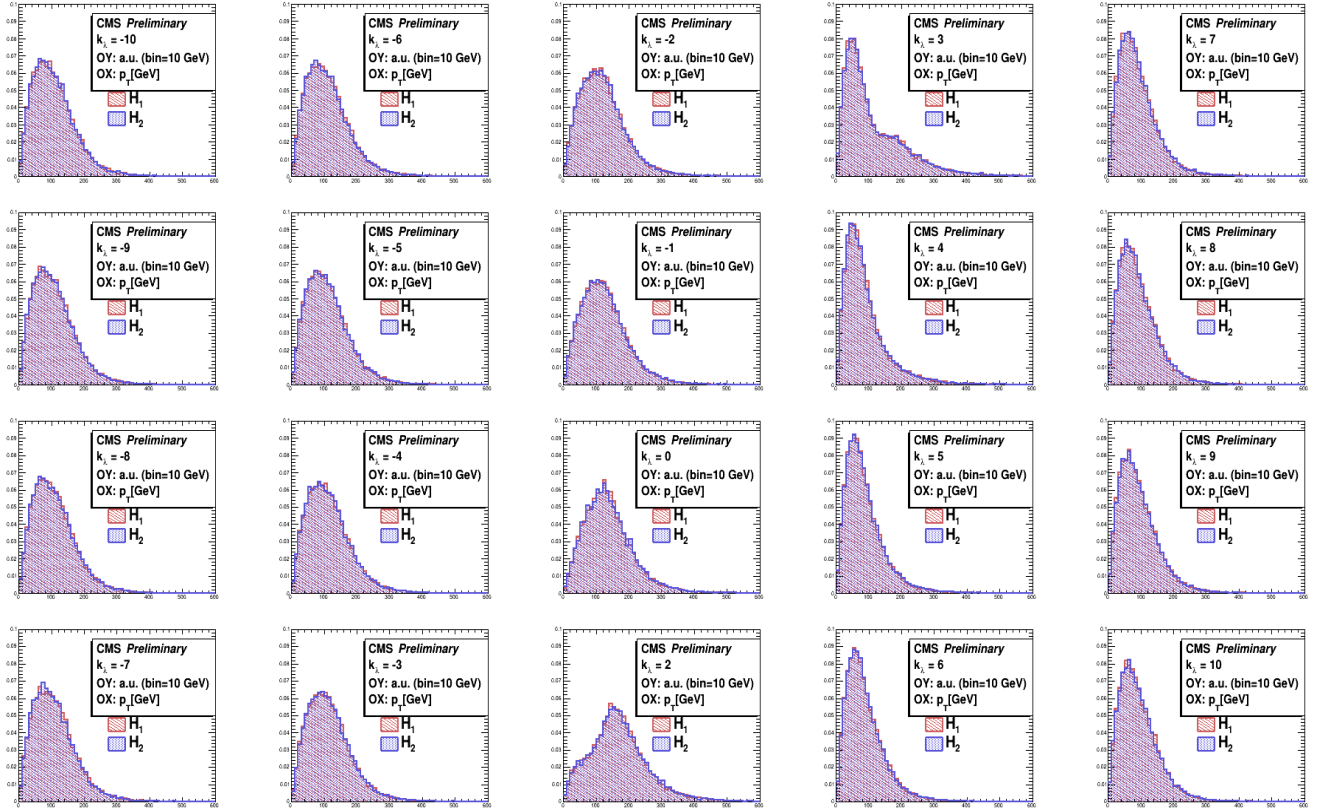


Figure A.1. Invariant mass m_{HH} for the k_λ -scan

Figure A.2. $p_T(HH)$ k_λ -scanFigure A.3. $p_T(H1), p_T(H2)$ k_λ -scan

Appendix B

Supporting information for chapter 5

B.1 Tables and Plots related to the Analysis Results

B.2 Non-resonant HH

Topology	SM-HH	ttZ	ttH	bbH	ZZ	H+VBF	WH	ZH	Others	Z+X	Total bkg.	Total bkg.
4μ	0.014	0.14	0.17	0.04	1.01	3.02	0.21	0.17	0.008	0.79	5.57	6
$4e$	0.007	0.07	0.09	0.02	0.38	1.61	0.11	0.10	0.01	1.40	3.78	3
$2e2\mu$	0.019	0.16	0.22	0.05	1.08	3.95	0.27	0.24	0.02	2.64	8.65	11

Table B.1. Cut flow table after signal region selection for the 2016 dataset ($35.8fb^{-1}$). The ZZ yields contain contributions from the processes $gg/q\bar{q} \rightarrow ZZ$, Higgs + VBF contains contributions from the single Higgs boson gluon fusion (ggH) and vector boson fusion (VBFH) production modes, others contains contributions from TTW, WWZ, WZZ, and ZZZ processes.

Topology	SM-HH	ttZ	ttH	bbH	ZZ	H+VBF	WH	ZH	Others	Z+X	Total bkg.	Total bkg.
4μ	0.016	0.18	0.20	0.05	1.06	3.73	0.27	0.21	0.01	1.48	7.21	9
$4e$	0.008	0.12	0.12	0.02	0.40	1.76	0.14	0.11	0.005	0.52	3.20	2
$2e2\mu$	0.024	0.30	0.27	0.06	1.09	4.57	0.35	0.29	0.04	2.00	9.00	6

Table B.2. Cut flow table after signal region selection for the 2017 dataset ($41.53fb^{-1}$). The ZZ yields contain contributions from the processes $gg/q\bar{q} \rightarrow ZZ$, Higgs + VBF contains contributions from the single Higgs boson gluon fusion (ggH) and vector boson fusion (VBFH) production modes, others contains contributions from TTW, WWZ, WZZ and ZZZ processes.

Topology	SM-HH	ttZ	ttH	bbH	ZZ	H+VBF	WH	ZH	Others	Z+X	Total bkg.	Total bkg.
4μ	0.025	0.24	0.30	0.07	1.65	5.14	0.40	0.32	0.02	1.60	9.76	14
$4e$	0.013	0.16	0.17	0.03	0.65	2.56	0.20	0.16	0.003	0.72	4.66	7
$2e2\mu$	0.036	0.39	0.39	0.08	1.68	6.55	0.50	0.41	0.05	2.58	12.73	16

Table B.3. Cut flow table after signal region selection for the 2018 dataset ($59.74fb^{-1}$). The ZZ yields contain contributions from the processes $gg/q\bar{q} \rightarrow ZZ$, Higgs + VBF contains contributions from the single Higgs boson gluon fusion (ggH) and vector boson fusion (VBFH) production modes, others contains contributions from TTW, WWZ, WZZ, ZZZ and processes.

B.3 Resonant HH

Graviton 2016	260	300	400	500	600	700	800	900	1000
4μ	1.07	0.84	0.24	0.10	0.0438	0.0218	0.0118	0.0063	0.0036
$4e$	0.52	0.33	0.12	0.06	0.0259	0.0132	0.0074	0.0041	0.0024
$2e2\mu$	1.47	0.94	0.35	0.15	0.0674	0.0350	0.0019	0.0010	0.0062
Radion 2016	260	300	400	500	600	700	800	900	1000
4μ	3.13	2.64	1.56	1.03	0.68	0.46	0.30	0.22	0.15
$4e$	1.40	1.20	0.79	0.54	0.38	0.27	0.21	0.14	0.10
$2e2\mu$	4.41	3.68	2.21	1.49	1.03	0.72	0.54	0.36	0.26

Table B.4. . Cut flow table after signal region selection for the 2016 dataset ($35.8fb^{-1}$) for the BulkGraviton spin-2 ans Radion spin-0

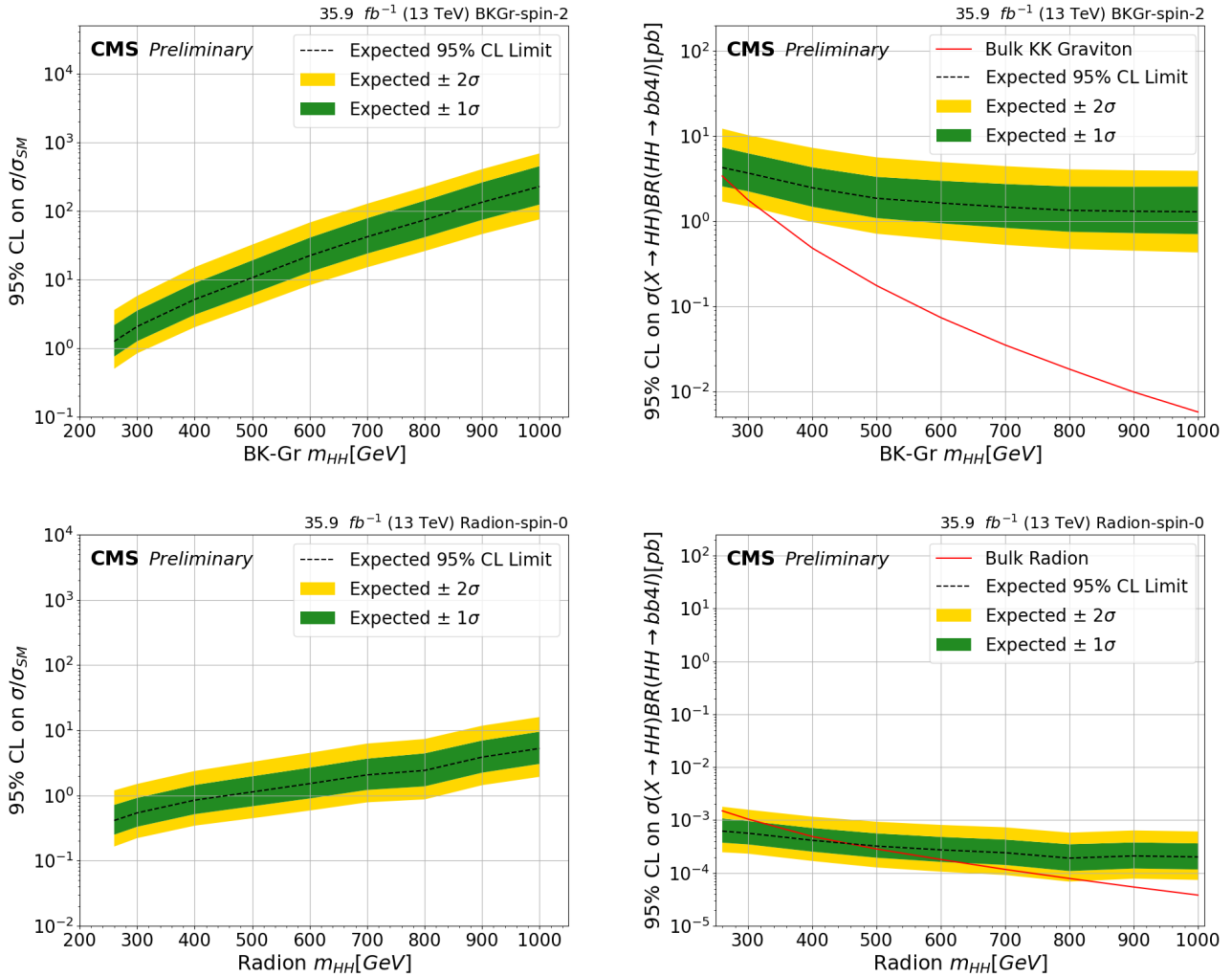
Graviton 2017	260	300	400	500	600	700	800	900	1000
4μ	1.2807	0.8154	0.2827	0.1156	0.0513	0.0258	0.0136	0.0072	0.0048
$4e$	0.5880	0.3961	0.1504	0.0651	0.0316	0.0163	0.0092	0.0051	0.0031
$2e2\mu$	1.8027	1.1236	0.4103	0.1755	0.0817	0.0417	0.0230	0.0127	0.0075
Radion 2017	260	300	400	500	600	700	800	900	1000
4μ	5.28	4.35	2.55	1.61	1.09	0.73	0.48	0.34	0.24
$4e$	2.37	2.04	1.28	0.89	0.64	0.44	0.35	0.23	0.17
$2e2\mu$	7.32	6.04	3.59	2.39	1.65	1.14	0.86	0.57	0.41

Table B.5. . Cut flow table after signal region selection for the 2017 dataset ($41.53fb^{-1}$) for the BulkGraviton spin-2 ans Radion spin-0

Graviton 2018	260	300	400	500	600	700	800	900	1000
4μ	2.9226	1.8155	0.6261	0.2945	0.1159	0.0572	0.0302	0.0163	0.0092
$4e$	0.8137	0.5963	0.2484	0.1134	0.0542	0.0290	0.016	0.0009	0.0053
$2e2\mu$	4.1735	2.6119	0.9479	0.4494	0.1844	0.0933	0.052	0.0028	0.0179
Radion 2018	260	300	400	500	600	700	800	900	1000
4μ	8.60	7.11	4.18	2.66	1.80	1.20	0.80	0.57	0.41
$4e$	2.47	2.28	1.56	1.13	0.82	0.58	0.46	0.30	0.25
$2e2\mu$	12.37	10.25	6.10	4.05	2.80	1.94	1.45	0.96	0.70

Table B.6. . Cut flow table after signal region selection for the 2018 dataset ($59.74fb^{-1}$) for the BulkGraviton spin-2 ans Radion spin-0

Graviton 2016	260	300	400	500	600	700	800	900	1000
$\mu + 2\sigma$	0.5012	0.8411	2.0220	4.0674	8.2568	15.0488	25.9277	45.8906	75.2051
$\mu + 1\sigma$	0.7553	1.2636	3.0644	6.2191	12.8483	23.8512	41.2458	73.9530	123.6667
μ (central)	1.2578	2.0703	5.125	10.625	22.25	41.875	73.75	133.5000	226.5
$\mu - 1\sigma$	2.1703	3.5392	8.9247	19.0106	40.6974	78.2626	140.7753	260.1485	448.5983
$\mu - 2\sigma$	3.6051	5.7787	15.1161	32.2052	67.4484	126.9524	223.6107	404.8149	686.8772
Radion 2016	260	300	400	500	600	700	800	900	1000
$\mu + 2\sigma$	0.1650	0.2219	0.3444	0.4469	0.5876	0.7845	0.8704	1.4380	1.9335
$\mu + 1\sigma$	0.2506	0.3316	0.5174	0.6827	0.9007	1.2164	1.3795	2.2376	3.0529
μ (central)	0.4141	0.5410	0.8477	1.1328	1.5195	2.0703	2.4219	3.8750	5.2656
$\mu - 1\sigma$	0.7177	0.9206	1.4491	1.9727	2.6703	3.6713	4.4298	6.9642	9.5054
$\mu - 2\sigma$	1.1957	1.5036	2.3748	3.2722	4.4988	6.2545	7.3416	11.7457	15.9612

Table B.7. . Table of limits for the Bulk Graviton spin-2 and Radion spin-0 for 2016 at a total integrated luminosity of $\mathcal{L} = 35.8 \text{ fb}^{-1}$ Figure B.1. . 1) Top left: Upper limit on the signal strength for 2016; 2) Top right: Expected 95% CLs upper limit on the cross section as a function of HH invariant mass for 2016. 3) Bottom left: Upper limit on the signal strength for 2016. 4) Bottom right: Expected 95% CLs upper limit on the cross section as a function of HH invariant mass for 2016. The solid red line shows the theoretical predictions.

Graviton 2017	260	300	400	500	600	700	800	900	1000
$\mu + 2\sigma$	0.4291	0.7178	1.7245	3.6013	7.1016	13.1729	23.0000	39.8066	65.2500
$\mu + 1\sigma$	0.6481	1.0726	2.5991	5.5199	10.9852	20.6220	36.4531	63.7319	105.8496
μ (central)	1.0664	1.75	4.3281	9.3125	18.9375	35.875	64.0	114.5	192.0
$\mu - 1\sigma$	1.8315	2.9638	7.5370	16.5880	34.4876	66.4768	120.6337	220.3852	377.2075
$\mu - 2\sigma$	3.0223	4.8610	12.6339	28.2263	57.4057	108.7577	194.0367	347.1794	582.2296
Radion 2017	260	300	400	500	600	700	800	900	1000
$\mu + 2\sigma$	0.1065	0.1419	0.2166	0.2942	0.3845	0.5310	0.5947	0.9780	1.3084
$\mu + 1\sigma$	0.1609	0.2110	0.3254	0.4468	0.5860	0.8093	0.9200	1.5076	2.0289
μ (central)	0.2646	0.3428	0.5332	0.7383	0.9844	1.3594	1.5859	2.5547	3.4531
$\mu - 1\sigma$	0.4524	0.5778	0.9030	1.2680	1.7142	2.3781	2.8376	4.5302	6.1234
$\mu - 2\sigma$	0.7470	0.9371	1.4811	2.1000	2.8534	3.9965	4.8071	7.6419	10.4321

Table B.8. . Table of limits for the Bulk Graviton spin-2 and Radion spin-0 for 2017 at a total integrated luminosity of $\mathcal{L} = 41.53 fb^{-1}$

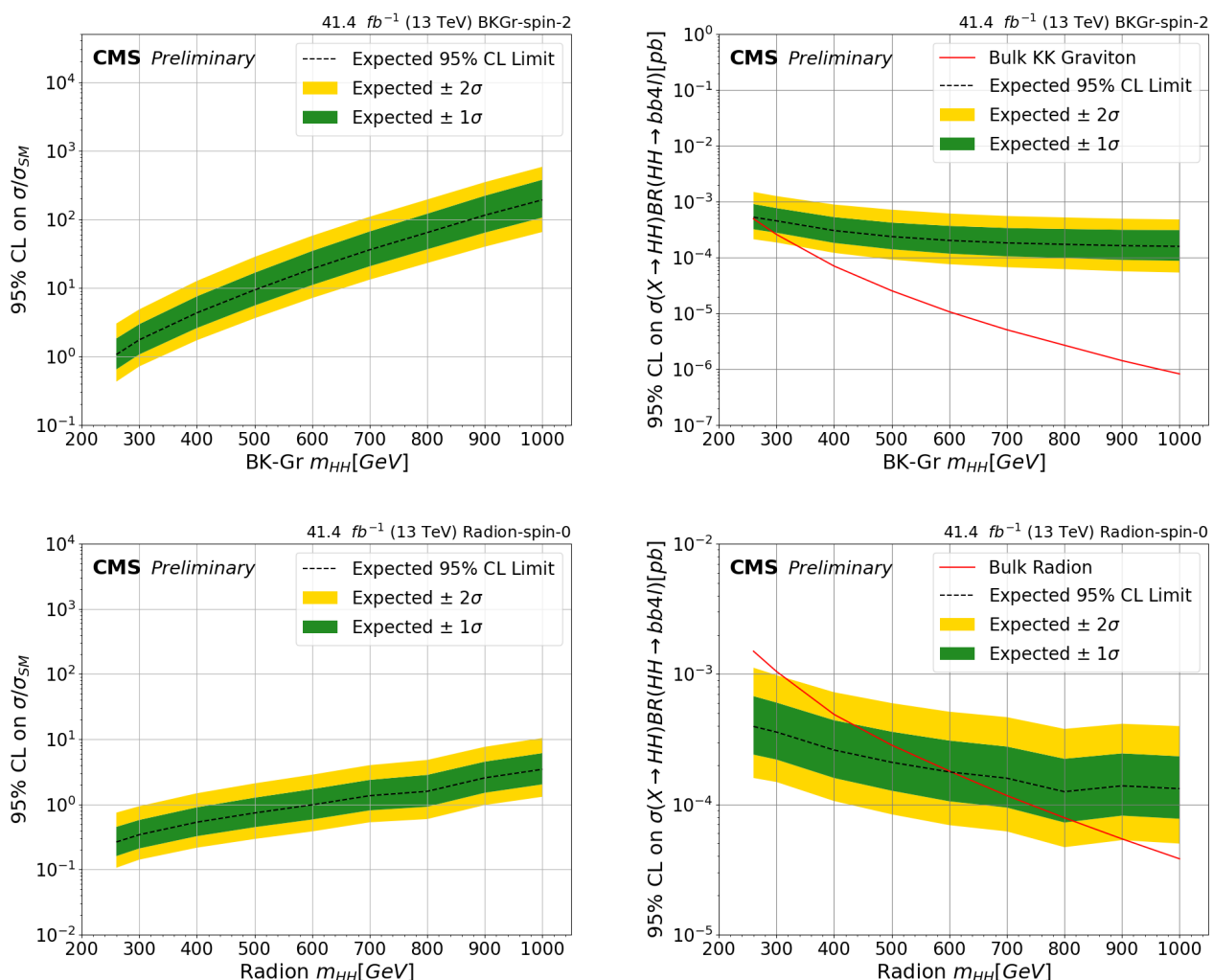
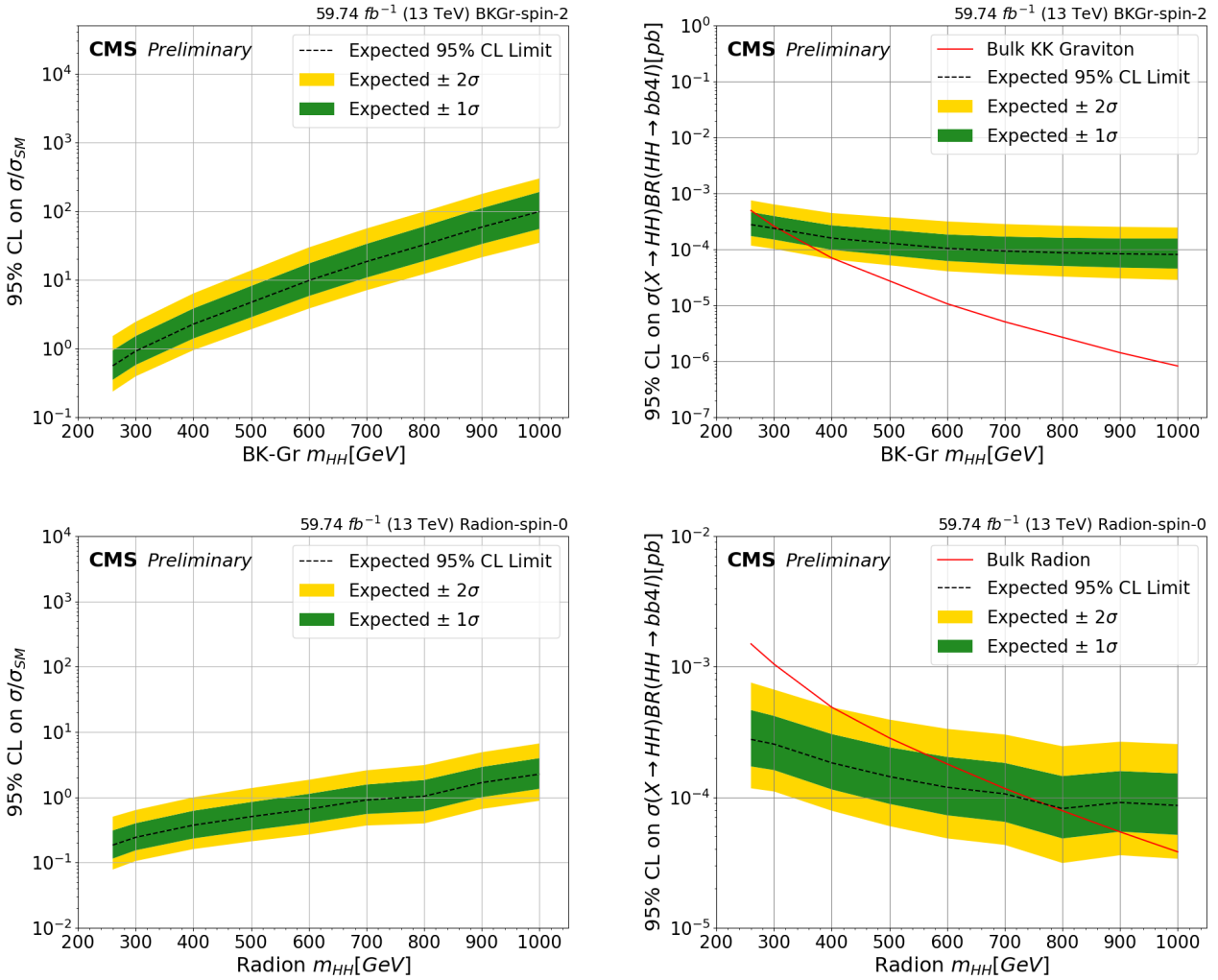


Figure B.2. . 1) Top left: Upper limit on the signal strength for 2017; 2) Top right: Expected 95% CLs upper limit on the cross section as a function of HH invariant mass for 2017. 3) Bottom left: Upper limit on the signal strength for 2017. 4) Bottom right: Expected 95% CLs upper limit on the cross section as a function of HH invariant mass for 2017. The solid red line shows the theoretical predictions.

2018	260	300	400	600	700	800	900	1000
$\mu + 2\sigma$	0.2365	0.3944	0.9470	3.8086	7.0342	12.1406	21.2520	34.5410
$\mu + 1\sigma$	0.3504	0.5785	1.4003	5.8045	10.7554	18.7800	33.1830	54.9478
μ (central)	0.5605	0.9180	2.2656	9.7500	18.3750	32.3750	58.5000	98.2500
$\mu - 1\sigma$	0.9404	1.5254	3.8370	17.2896	33.3167	59.9913	110.2667	189.8911
$\mu - 2\sigma$	1.5195	2.4551	6.3410	29.5518	55.6994	98.1472	177.3616	297.9133

2018	260	300	400	500	600	700	800	900	1000
$\mu + 2\sigma$	0.0783	0.1059	0.1611	0.2114	0.2690	0.3698	0.3978	0.6627	0.8850
$\mu + 1\sigma$	0.1152	0.1545	0.2346	0.3126	0.4041	0.5555	0.6132	1.0043	1.3488
μ (central)	0.1855	0.2441	0.3750	0.5059	0.6621	0.9102	1.0391	1.6797	2.2656
$\mu - 1\sigma$	0.3113	0.4018	0.6231	0.8486	1.1319	1.5704	1.8426	2.9250	3.9815
$\mu - 2\sigma$	0.5049	0.6386	0.9989	1.3820	1.8550	2.5899	3.1185	4.9031	6.6850

Table B.9.. Table of limits for the Bulk Graviton spin-2 and Radion spin-0 for 2018 at a total integrated luminosity of $\mathcal{L} = 59.74 \text{ fb}^{-1}$.Figure B.3. . 1) Top left: Upper limit on the signal strength for 2018; 2) Top right: Expected 95% CLs upper limit on the cross section as a function of HH invariant mass for 2018. 3) Bottom left: Upper limit on the signal strength for 2018. 4) Bottom right: Expected 95% CLs upper limit on the cross section as a function of HH invariant mass for 2018. The solid red line shows the theoretical predictions.

Bibliography

- [1] M. Abbrescia and P. F. an V. Peskov. . Wiley-VCH, 2018. Cited on page/s 57.
- [2] K. Agashe, H. Davoudiasl, G. Perez, and A. Soni. [Large Mass Hierarchy from a Small Extra Dimension](#). Phys. Rev. D 76, 036006, 2007. Cited on page/s 26.
- [3] S. Alioli, C. Oleari, E. Re, and P. Nason. [NLO vector-boson production matched with shower in POWHEG](#). JHEP0807:060,2008, 2008. Cited on page/s 90.
- [4] R. Alonso. [Dynamical Yukawa Couplings](#). arXiv:1307.1904 [hep-ph], page 140, 2013. Cited on page/s 12.
- [5] J. Alwall, F. Maltoni, O. Mattelaer, T. Stelzer, and M. Herquet. [MadGraph 5 : Going Beyond](#). arXiv:1106.0522 [hep-ph], 2011. Cited on page/s 73.
- [6] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. [The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations](#). JHEP07(2014)079, 2014. Cited on page/s 73.
- [7] P. Anderson. [Plasmons, gauge invariance, and mass](#). Physical Review. 130, page 321, 1962. Cited on page/s 9.
- [8] N. Arkani-Hamed, S. Dimopoulos, and G. Dvalic. [The hierarchy problem and new dimensions at a millimeter](#). Physics Letters B, Volume 429, Issues 3-4, 1998. Cited on page/s 26.
- [9] N. Arkani-Hamed, S. Dimopoulos, and G. Dvalic. [Large Mass Hierarchy from a Small Extra Dimension](#). Phys. Rev. Lett. 83, 3370, 1999. Cited on page/s 26.
- [10] P. Artoisenet, P. de Aquino, F. Demartin, R. Frederix, S. Frixione, F. Maltoni, M. K. Mandal, P. Mathews, K. Mawatari, V. Ravindran, S. Seth, P. Torrielli, and M. Zaro. [A framework for Higgs characterisation](#). arXiv:1306.6464 [hep-ph], 2014. Cited on page/s 83.
- [11] J. H. at al. [The PDF4LHC Working Group Interim Report](#). arXiv:1101.0536 [hep-ph], 2011. Cited on page/s 79.
- [12] ATLAS-CMS-Collaboration. [Combined Measurement of the Higgs Boson Mass in pp Collisions at \$\sqrt{s} = 7\$ and 8 TeV with the ATLAS and CMS Experiments](#). Phys. Rev. Lett. 114, 191803, 2015. Cited on page/s 13.
- [13] J. Baglio, A. Djouadi, R. Groeber, M. Muehlleitner, J. Quevillon, and M. Spira. [The measurement of the Higgs self-coupling at the LHC: theoretical status](#). Journal of High Energy Physics, 2013. Cited on page/s 15.
- [14] E. Bagnaschi, G. Degrassi, P. Slavich, and A. Vicini. [Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM](#). Journal of High Energy Physics volume 2012, Article number: 88 (2012), 2012. Cited on page/s 90.

- [15] A. Belyaev, N. D. Christensen, and A. Pukhov. [CalcHEP 3.4 for collider physics within and beyond the Standard Model](#). arXiv:1207.6082 [hep-ph], 2012. Cited on page/s [83](#).
- [16] P. Billoir. [Progressive track recognition with a Kalman-like fitting procedure](#). Computer Physics Communications, Volume 57, 1989. Cited on page/s [100](#).
- [17] T. Binoth and J. van der Bij. [Influence of strongly coupled, hidden scalars on Higgs signals](#). Z.Phys.C75:17-25,1997, 1997. Cited on page/s [23](#).
- [18] S. Borowka, N. Greiner, G. Heinrich, S. Jones, M. Kerner, J. Schlenk, U. Schubert, and T. Zirke. [Higgs boson pair production in gluon fusion at NLO with full top-quark mass dependence](#). Phys. Rev. Lett. 117, 012001 (2016), 2016. Cited on page/s [xvi](#), [xxiii](#), [20](#), [21](#).
- [19] G. C. Branco, P. M. Ferreira, L. Lavoura, M. N. Rebelo, M. Sher, and J. P. Silva. [Theory and phenomenology of two-Higgs-doublet models](#). arXiv:1106.0034 [hep-ph], 2011. Cited on page/s [24](#).
- [20] W. Buchmuller and C. Luedeling. [Field Theory and Standard Model](#). DESY-06-151, page 70, 2006. Cited on page/s [3](#), [5](#).
- [21] A. Buckley, J. Ferrando, S. Lloyd, K. Nordstrom, B. Page, M. Ruefenacht, M. Schoenherr, and G. Watt. [LHAPDF6: parton density access in the LHC precision era](#). arXiv:1412.7420 [hep-ph], 2015. Cited on page/s [77](#), [79](#).
- [22] M. Z. C. Giganti, S. Lavignac. [Neutrino oscillations: the rise of the PMNS paradigm](#). 10.1016/j.ppnp.2017.10.001, arXiv:1710.00715 [hep-ex], page 82, 2017. Cited on page/s [6](#).
- [23] M. Cacciari, G. P. Salam, and G. Soyez. [The anti- \$k_t\$ jet clustering algorithm](#). JHEP 0804:063, 2008. Cited on page/s [106](#).
- [24] J. M. Campbell and R. K. Ellis. [MCFM for the Tevatron and the LHC](#). Nucl.Phys.Proc.Suppl.205-206:10-15,2010, 2010. Cited on page/s [90](#), [113](#).
- [25] J. M. Campbell, R. K. Ellis, and C. Williams. [Bounding the Higgs width at the LHC using full analytic results for \$gg \rightarrow 2e2\mu\$](#) . rXiv:1311.3589 [hep-ph], 2013. Cited on page/s [90](#).
- [26] S. Carrazza, S. Forte, Z. Kassabov, J. I. Latorre, and J. Rojo. [An Unbiased Hessian Representation for Monte Carlo PDFs](#). arXiv:1505.06736 [hep-ph], 2015. Cited on page/s [78](#), [79](#).
- [27] A. Carvalho. [Gravity particles from Warped Extra Dimensions, predictions for LHC](#). 2018. Cited on page/s [xvi](#), [31](#), [82](#).
- [28] S. Catani, D. de Florian, M. Grazzini, and P. Nason. [Soft-gluon resummation for Higgs boson production at hadron colliders](#). JHEP 0307 (2003) 028, 2003. Cited on page/s [30](#).
- [29] CERN. [ALICE](#). CERN Public, . Cited on page/s [40](#).
- [30] CERN. [ATLAS](#). CERN Public, . Cited on page/s [41](#).
- [31] CERN. [The Large Hadron Collider](#). CERN Public, . Cited on page/s [35](#).
- [32] CERN. [A successful conclusion to Run 2](#). CERN Public, . Cited on page/s [38](#).
- [33] CERN. [CMS](#). CERN Public, . Cited on page/s [41](#).
- [34] CERN. [The CERN accelerator complex](#). CERN Public, . Cited on page/s [xvi](#), [35](#).

- [35] CERN. [LHCb](#). CERN Public, 2021. Cited on page/s 40.
- [36] CERN-Yellow-Report-Group. [Higgs boson potential at colliders: status and perspectives](#). Review in Physics (2020) 100045, 2020. Cited on page/s xv, xxiii, 14, 15, 16, 18, 21.
- [37] C. Charmousis, R. Gregory, and V. A. Rubakov. [Wave function of the radion in a brane world](#). Phys. Rev. D 62, 067505, 2000. Cited on page/s 28.
- [38] C. Chen and I. Low. [A Double Take on New Physics in Double Higgs Production](#). arXiv:1405.7040 [hep-ph], 2014. Cited on page/s 19.
- [39] N. D. Christensen and C. Duhr. [FeynRules - Feynman rules made easy](#). arXiv:0806.4194 [hep-ph], 2008. Cited on page/s 83.
- [40] Y. Coadou. [Boosted Decision Trees and Applications](#). EPJ Web of Conferences 55, 02004, 2013. Cited on page/s 121.
- [41] R. Contino, M. Ghezzi, C. Grojean, M. Muhlleitner, and M. Spira. [Effective Lagrangian for a light Higgs-like scalar](#). JHEP 1307 (2013) 035, 2013. Cited on page/s 30.
- [42] R. Contino, M. Ghezzi, M. Moretti, G. Panico, F. Piccinini, and A. Wulzer. [Anomalous Couplings in Double Higgs Production](#). arXiv:1205.5444 [hep-ph], 2013. Cited on page/s 20.
- [43] R. Contino, M. Ghezzi, C. Grojean, M. Muhlleitner, and M. Spira. [eHDECAY: an Implementation of the Higgs Effective Lagrangian into HDECAY](#). CERN-PH-TH/2013-316, KA-TP-03-2014, PSI-PR-13-20, arXiv:1403.3381 [hep-ph], 2014. Cited on page/s 30.
- [44] G. Cowan, K. Cranmer, E. Gross, and O. Vitells. [Asymptotic formulae for likelihood-based tests of new physics](#). Eur.Phys.J.C71:1554,2011, 2013. Cited on page/s 147, 148, 149, 150.
- [45] C. Csaki, M. L. Graesser, and G. D. Kribs. [Radion Dynamics and Electroweak Physics](#). Phys.Rev.D63:065002,2001, 2000. Cited on page/s 28, 29.
- [46] C. Csaki, J. Erlich, and J. Terning. [The Effective Lagrangian in the Randall-Sundrum Model and Electroweak Physics](#). Phys.Rev.D66:064021,2002, 2002. Cited on page/s 30.
- [47] H. Cui. [Derivation of Gell-Mann-Nishijima formula from the electromagnetic field modes of a hadron](#). arXiv:1001.0226 [physics.gen-ph], page 10, 2010. Cited on page/s 7.
- [48] S. Dawson, S. Dittmaier, and M. Spira. [Neutral Higgs-boson pair production at hadron colliders: QCD corrections](#). American Physical Society, 1998. Cited on page/s xxiii, 19.
- [49] S. Dawson, E. Furlan, and I. Lewis. [Unraveling an extended quark sector through multiple Higgs production?](#) Phys. Rev. D 87, 014007, 2013. Cited on page/s 20.
- [50] P. de Aquino, K. Hagiwara, Q. Li, and F. Maltoni. [Simulating graviton production at hadron colliders](#). JHEP 1106:132,2011, 2011. Cited on page/s 83.
- [51] W. de Boer. [Grand Unified Theories and Supersymmetry in Particle Physics and Cosmology](#). Prog.Part.Nucl.Phys.33:201-302, 1994. Cited on page/s 4.
- [52] D. de Florian and M. Grazzini. [Higgs production through gluon fusion: updated cross sections at the Tevatron and the LHC](#). Phys.Lett.B674:291-294,2009, 2009. Cited on page/s 30.

- [53] P. Dirac. [The quantum theory of the electron](#). Royal Society: Volume 117, Issue 778, doi/10.1098/1928.0023, page 15, 1928. Cited on page/s [6](#), [7](#).
- [54] P. Dirac. [A theory of electrons and protons](#). Royal Society: Volume 126, Issue 801, doi/10.1098/1930.0013, page 6, 1930. Cited on page/s [6](#).
- [55] P. Dirac. [Relativistic quantum mechanics](#). Royal Society: Volume 136, Issue 829, doi/10.1098/1932.0094, page 18, 1932. Cited on page/s [6](#).
- [56] A. Djouadi, L. Maiani, G. Moreau, A. Polosa, J. Quevillon, and V. Riquer. [The post-Higgs MSSM scenario: habemus MSSM?](#) The European Physical Journal C volume 73, Article number: 2650, 2013. Cited on page/s [25](#).
- [57] D. Dominici, B. Grzadkowski, J. Gunion, and M. Toharia. [The scalar sector of the Randall-Sundrum model](#). Nucl.Phys. B671 (2003) 243-292, 2002. Cited on page/s [29](#).
- [58] D. Dominici, B. Grzadkowski, J. F. Gunion, and M. Toharia. [The scalar sector of the Randall-Sundrum model](#). Nuclear Physics B Volume 671, 2003. Cited on page/s [29](#).
- [59] F. Englert and R. Brout. [Broken Symmetry and the Mass of Gauge Vector Mesons](#). Phys. Rev. Lett. 13, page 321, 1964. Cited on page/s [9](#).
- [60] J. et al. [GEANT4 – a simulation toolkit](#). Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, Volume 506, Issue 3, 2003. Cited on page/s [84](#).
- [61] L. D. et al. [Current Status and Future Trends of SiGe BiCMOSTechnology](#), IEEE Transactions on Electron Devices. IEEE Transactions on Electron Devices, Vol. 48, No. 11, 2001. Cited on page/s [62](#).
- [62] L. P. et al. [Development of a new Front End electronics in Silicon and Silicon-Germanium technology for the Resistive Plate Chamber detector for high rate experiment](#). Journal of Instrumentation, Volume 14, 2018. Cited on page/s [62](#).
- [63] M. A. et al. [Properties of C₂H₂F₄ based gas mixtures for avalanche mode operation of Resistive Plate Chambers](#). Nuclear Instruments and Methods in Physics, Volume 398, Issues 2-3,, 1997. Cited on page/s [57](#).
- [64] R. C. et al. [Performance of RPCs and diamond detectors using a new very fast low noise preamplifier](#). JINST 8 P01003, 2013. Cited on page/s [61](#).
- [65] W. S. et al. [The Theory of P-N Junctions in Semiconductors and P-N Junction Transistors](#). Bell System Technical Journal Vol. 28, No. 3, 1949. Cited on page/s [61](#).
- [66] F. M. et.al. [Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector](#). CERN Yellow Reports: Monographs Volume 2/2017 (CERN-2017-002-M), page 869, 2017. Cited on page/s [xv](#), [14](#), [A-1](#).
- [67] M. A. et.al. [Effect of the linseed oil surface treatment on the performance of resistive plate chambers](#). Nucl. Instrum. Methods Phys.A 394, 13-20, 1997. Cited on page/s [57](#).
- [68] L. Evans and P. Bryant. [LHC Machine](#). JINST3S08001, 2008. Cited on page/s [xxiii](#), [39](#).
- [69] L. Evans and P. Bryant. [Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8TeV](#). Eur. Phys. J. C 75, 212 (2015), 2015. Cited on page/s [38](#).

- [70] E.W.N.Glover and J. der Bij. [Higgs boson pair production via gluon fusion](#). Nuclear Physics B, pages 282–294, 1988. Cited on page/s 11.
- [71] A. Falkowski. [Effective field theory approach to LHC Higgs data](#). page 26, 2015. Cited on page/s 17.
- [72] D. D. Florian and J. Mazzitelli. [Higgs Boson Pair Production at Next-to-Next-to-Leading Order in QCD](#). Phys.Rev.Lett. 111 201801, 2013. Cited on page/s xxiii, 18, 19.
- [73] D. D. Florian and J. Mazzitelli. [Higgs Boson Pair Production at Next-to-Next-to-Leading Order in QCD](#). arXiv:1505.07122 [hep-ph], 2015. Cited on page/s 18.
- [74] M. Forger and H. Roemer. [Currents and the Energy-Momentum Tensor in Classical Field Theory: A fresh look at an Old Problem](#). Annals Phys. 309 (2004) 306–389, 2003. Cited on page/s 29.
- [75] R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, P. Torrielli, E. Vryonidou, and M. Zaro. [Higgs pair production at the LHC with NLO and parton-shower effects](#). Physics Letters B, Volume 732, 2014. Cited on page/s 15.
- [76] S. Frixione, P. Nason, and C. Oleari. [Matching NLO QCD computations with parton shower simulations: the POWHEG method](#). Journal of High Energy Physics, Volume 2007, JHEP11(2007), 2007. Cited on page/s 90.
- [77] R. Fruehwirth. [Application of Kalman filtering to track and vertex fitting](#). Nuclear Instruments and Methods in Physics, Volume 262, 1987. Cited on page/s 100.
- [78] Y. Gao, A. V. Gritsan, Z. Guo, K. Melnikov, M. Schulze, and N. V. Tran. [Spin determination of single-produced resonances at hadron colliders](#). Phys.Rev.D81:075022,2010, 2010. Cited on page/s 90.
- [79] M. Gell-Mann. [Symmetries of Baryons and Mesons](#). American Physical Society Phys. Rev. 125, 1067, page 18, 1962. Cited on page/s 6.
- [80] T. Gershon. [Overview of the CKM Matrix](#). 10.1007/s12043-012-0418-y, arXiv:1112.1984 [hep-ex], page 17, 2011. Cited on page/s 5.
- [81] G.Heinrich, S.P.Jones, M.Kerner, G.Luisoni, and L.Scyboz. [NLO predictions for Higgs boson pair production with full top quark mass dependence matched to parton showers](#). JHEP 1708 (2017) 088, 2017. Cited on page/s xviii, 73, 74, 79, A-1.
- [82] G.Heinrich, S.P.Jones, M.Kerner, G.Luisoni, and L.Scyboz. [Probing the trilinear Higgs boson coupling in di-Higgs production at NLO QCD including parton shower effects](#). J. High Energ. Phys. (2019) 66, 2019. Cited on page/s xviii, 73, 74, A-1.
- [83] M. Gillioz, R. Grober, C. Grojean, M. Muhlleitner, and E. Salvioni. [Higgs Low-Energy Theorem \(and its corrections\) in Composite Models](#). JHEP 1210 (2012) 004, 2012. Cited on page/s 20.
- [84] J. Goldstone. [Field theories with Superconductor solutions](#). Il Nuovo Cimento (1955-1965), Springer, 1961. Cited on page/s 10.
- [85] G. Guralnik, C. Hagen, and T. Kibble. [Global conservation laws and massless particles](#). Physical Review Letters, page 20, 2016. Cited on page/s 9.
- [86] Q. Y. H. Tam. [Production and Detection of Axion-like Particles by Interferometry](#). Physics Letters B 716 (2012), pp. 435–440, 2012. Cited on page/s 29.

- [87] H. B. Hartanto, B. Jager, L. Reina, and D. Wackeroth. [Higgs boson production in association with top quarks in the POWHEG BOX](#). Phys. Rev. D 91, 094003 (2015), 2015. Cited on page/s 90.
- [88] Z. He, D. Lin, T. Lau, and M. Wu. [Gradient Boosting Machine: A Survey](#). arXiv:1908.06951 [stat.ML], 2019. Cited on page/s 122.
- [89] G. Heinrich, S. P. Jones, M. Kerner, G. Luisoni, and E. Vryonidou. [NLO predictions for Higgs boson pair production with full top quark mass dependence matched to parton showers](#). JHEP 1708 (2017) 088, 2017. Cited on page/s 18.
- [90] G. Heinrich, S. P. Jones, M. Kerner, G. Luisoni, and L. Scyboz. [Probing the trilinear Higgs boson coupling in di-Higgs production at NLO QCD including parton shower effects](#). J. High Energ. Phys. (2019) 66, 2019. Cited on page/s 18.
- [91] G. Heinrich, S. P. Jones, M. Kerner, and L. Scyboz. [A non-linear EFT description of \$gg \rightarrow HH\$ at NLO interfaced to POWHEG](#). J. High Energ. Phys. 2020, 21 (2020), 2020. Cited on page/s 18.
- [92] P. W. Higgs. [Broken Symmetries and the Masses of Gauge Bosons](#). Phys. Rev. Lett. 13, page 508, 1964. Cited on page/s 9.
- [93] M. Kraemer and E. L. an dM. Spira. [Soft Gluon Radiation in Higgs Boson Production at the LHC](#). Nucl.Phys.B511:523-549,1998, 1998. Cited on page/s 30.
- [94] G. Luisoni, P. Nason, C. Oleari, and F. Tramontano. [HW[±]/HZ + 0 and 1 jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO](#). arXiv:1306.2542 [hep-ph], 2013. Cited on page/s 90.
- [95] L. Maiani, R. Petronzio, and E. Zavattini. [Effects of nearly massless, spin-zero particles on light propagation in a magnetic field](#). Physics Letters B, Volume 175, Issue 3, 1986. Cited on page/s 29.
- [96] F. Maltoni and T. Stelzer. [MadEvent: Automatic Event Generation with MadGraph](#). JHEP0302:027,2003, 2003. Cited on page/s 74.
- [97] F. Maltoni, G. Ridolfi, and M. Ubiali. [b-initiated processes at the LHC: a reappraisal](#). Journal of High Energy Physics, 2012. Cited on page/s 16.
- [98] F. Maltoni, K. Mawatari, and M. Zaro. [Higgs characterisation via vector-boson fusion and associated production: NLO and parton-shower effects](#). arXiv:1311.1829 [hep-ph], 2013. Cited on page/s 30.
- [99] H. Murayama, I. Watanabe, and K. Hagiwara. [HELAS: HELicity Amplitude Subroutines for Feynman diagram evaluations](#). KEK-91-11, 1992. Cited on page/s 73.
- [100] P. Nason. [A new method for combining NLO QCD with shower Monte Carlo algorithms](#). Journal of High Energy Physics, Volume 2004, JHEP11(2004), 2004. Cited on page/s 90.
- [101] P. Nason and C. Oleari. [NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG](#). JHEP 1002:037,2010, 2010. Cited on page/s 90.
- [102] P. Nason and B. Webber. [Next-to-Leading-Order Event Generators](#). Cavendish-HEP-2012-02; CERN-PH-TH-2012-028, 2012. Cited on page/s 74.
- [103] P. Nason and G. Zanderighi. [W⁺W⁻, WZ and ZZ production in the POWHEG BOX V2](#). arXiv:1311.1365 [hep-ph], 2013. Cited on page/s 90.

- [104] P. Nason, C. Oleari, M. Rocco, and M. Zaro. [An interface between the POWHEG BOX and MadGraph5_aMC@NLO](#). arXiv:2008.06364 [hep-ph], 2020. Cited on page/s [73](#), [74](#).
- [105] Y. Nir. [One Higgs, three discoveries](#). CERN-COURIER, 2020. Cited on page/s [xv](#), [10](#).
- [106] C. Oleari. [The POWHEG-BOX](#). Nucl.Phys.Proc.Suppl.205-206:36-41,2010, 2010. Cited on page/s [73](#), [74](#).
- [107] T. Plehn, M. Spira, and P. Zerwas. [Pair Production of Neutral Higgs Particles in Gluon–Gluon Collisions](#). Nucl.Phys.B479:46-64,1996; Erratum-ibid.B531:655,1998, 1996. Cited on page/s [20](#).
- [108] L. Randall and R. Sundrum. [A Large Mass Hierarchy from a Small Extra Dimension](#). Phys.Rev.Lett.83:3370-3373, 1999. Cited on page/s [26](#), [27](#).
- [109] L. Randall and R. Sundrum. [Large Mass Hierarchy from a Small Extra Dimension](#). Physical Review Letters. 83, page 17, 1999. Cited on page/s [82](#).
- [110] L. Randall and R. Sundrum. [An Alternative to Compactification](#). Physical Review Letters. 83, page 23, 1999. Cited on page/s [82](#).
- [111] A. L. Read. [Presentation of search results: the CLs technique](#). J. Phys. G: Nucl. Part. Phys. 28 2693, 2002. Cited on page/s [149](#).
- [112] R.Frederix, S.Frixione, V.Hirschi, F.Maltoni, O.Mattelaer, P.Torrielli, E.Vryonidou, and M.Zaro. [Pair Production of Neutral Higgs Particles in Gluon–Gluon Collisions](#). Physics Letters B, Volume 732, pages 142–149, 2014. Cited on page/s [xv](#), [18](#), [19](#).
- [113] T. Robens and T. Stefaniak. [Status of the Higgs Singlet Extension of the Standard Model after LHC Run 1](#). EPJC (2015) 75:105, 2015. Cited on page/s [24](#).
- [114] R.Santonico and R.Cardarelli. [Development of resistive palte counters](#). Nuclear Instruments and Methods in Physics Research 187, 1981. Cited on page/s [57](#).
- [115] R. Schabinger and J. D. Wells. [A Minimal Spontaneously Broken Hidden Sector and its Impact on Higgs Boson Physics at the Large Hadron Collider](#). Journal reference: Phys.Rev.D72:093007,2005, 2005. Cited on page/s [23](#).
- [116] T. Sjostrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands. [An Introduction to PYTHIA 8.2](#). LU TP 14-36, MCNET-14-22, CERN-PH-TH-2014-190, FERMILAB-PUB-14-316-CD, DESY 14-178, SLAC-PUB-16122, 2015. Cited on page/s [74](#), [83](#).
- [117] P. Z. T. Plehn, M. Spira. [Pair Production of Neutral Higgs Particles in Gluon–Gluon Collisions](#). Nucl.Phys.B479:46-64,1996; Erratum-ibid.B531:655,1998, page 23, 1998. Cited on page/s [17](#).
- [118] The, CMS, and Collaboration. [Observation of \$t\bar{t}H\$ production](#). Phys. Rev. Lett. 120, 231801, 2018. Cited on page/s [16](#).
- [119] The-ATLAS-Collaboration. [Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC](#). Physics Letters B, Volume 716, Issue 1, 2012. Cited on page/s [13](#).
- [120] The-ATLAS-Collaboration. [Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at \$\sqrt{s} = 7\$ and 8 TeV](#). Journal of High Energy Physics, 2016. Cited on page/s [xv](#), [15](#).

- [121] The-ATLAS-Collaboration. [Measurements of Higgs boson properties in the diphoton decay channel with \$36.1\text{fb}^{-1}\$ \$pp\$ collision data at the center-of-mass energy of 13 TeV with the ATLAS detector](#). 2017 European Physical Society Conference on High Energy Physics, 2017. Cited on page/s [xv](#), [13](#).
- [122] The-CMS-Collaboration. [Bulk Graviton and Radion XS GitHub repository CMS](#). . Cited on page/s [xxiv](#), [83](#), [120](#).
- [123] The-CMS-Collaboration. [High Luminosity LHC Project](#) . CMS Public, . Cited on page/s [xvi](#), [41](#).
- [124] The-CMS-Collaboration. [CMS Luminosity - Public Results](#). CMS Public, . Cited on page/s [xvi](#), [38](#), [40](#), [91](#).
- [125] The-CMS-Collaboration. [CMS Statistics Committee](#). CMS internal, . Cited on page/s [147](#).
- [126] The-CMS-Collaboration. [The CMS electromagnetic calorimeter project : Technical Design Report](#). CERN-LHCC-97-033 ; CMS-TDR-4, 1997-2021. Cited on page/s [52](#).
- [127] The-CMS-Collaboration. [The CMS muon project : Technical Design Report](#). CERN-LHCC-97-032 ; CMS-TDR-3, 1997-2021. Cited on page/s [xvii](#), [47](#), [57](#), [59](#).
- [128] The-CMS-Collaboration. [Electron reconstruction in CMS](#). CMS-NOTE-2006-040, 2006. Cited on page/s [xxiv](#), [101](#), [103](#), [105](#).
- [129] The-CMS-Collaboration. [TMVA 4: Toolkit for Multivariate Data Analysis with ROOT, Users Guide](#). CERN-OPEN-2007-007, 2007. Cited on page/s [xx](#), [121](#), [122](#), [124](#).
- [130] The-CMS-Collaboration. [CMS High Level Trigger](#). CERN/LHCC 2007-021 LHCC-G-134, 2007. Cited on page/s [50](#).
- [131] The-CMS-Collaboration. [The CMS Barrel Calorimeter Response to Particle Beams from 2 to 350 GeV/c](#). ssssss, 2009-2018. Cited on page/s [46](#).
- [132] The-CMS-Collaboration. [Performance of the CMS cathode strip chambers with cosmic rays](#). Journal of Instrumentation, vol. 5, T03018, 2010. Cited on page/s [57](#).
- [133] The-CMS-Collaboration. [Determination of jet energy calibration and transverse momentum resolution in CMS](#). JINST 6 P11002, 2011. Cited on page/s [107](#).
- [134] The-CMS-Collaboration. [Performance of Muon Identification in \$pp\$ collisions at \$\sqrt{s} = 7\text{TeV}\$](#) . JINST 7 P10002, 2012. Cited on page/s [104](#), [105](#).
- [135] The-CMS-Collaboration. [Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC](#). Physics Letters B, Volume 716, Issue 1, 2012. Cited on page/s [13](#).
- [136] The-CMS-Collaboration. [Energy calibration and resolution of the CMS electromagnetic calorimeter in \$pp\$ collisions at \$\sqrt{s} = 7\text{TeV}\$](#) . JINST 8 (2013) P09009, 2013. Cited on page/s [45](#).
- [137] The-CMS-Collaboration. [Description and performance of track and primary-vertex reconstruction with the CMS tracker](#). JINST 9 P10009, 2014. Cited on page/s [100](#).
- [138] The-CMS-Collaboration. [Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at \$\sqrt{s} = 8\text{TeV}\$](#) . Journal of Instrumentation 10, P06005, 2015. Cited on page/s [102](#).

- [139] The-CMS-Collaboration. [CMS Drift Tubes performance](#). CMS-CR-2015-362, 2015-2021. Cited on page/s [48](#).
- [140] The-CMS-Collaboration. [Measurement of the properties of the Higgs boson in the four lepton final state at \$\sqrt{s} = 13\text{TeV}\$](#) . CMS AN 2016/442, 2016. Cited on page/s [106](#).
- [141] The-CMS-Collaboration. [Jet energy scale and resolution performances with 13TeV data](#). CMS-DP-2016-020 ; CERN-CMS-DP-2016-020, 2016. Cited on page/s [107](#).
- [142] The-CMS-Collaboration. [CMS](#). CMS Public, 2016. Cited on page/s [xix](#), [99](#).
- [143] The-CMS-Collaboration. [Particle flow reconstruction and global event description with the CMS detector](#). JINST, 12 (2017), 2017. Cited on page/s [100](#).
- [144] The-CMS-Collaboration. [Measurement of the properties of the Higgs boson in the four lepton final state at \$\sqrt{s} = 13\text{TeV}\$ with 2017 data](#). CMS AN 2017/342, 2017. Cited on page/s [106](#).
- [145] The-CMS-Collaboration. [Jet energy scale and resolution in the CMS experiment in pp collisions at \$\sqrt{s} = 8\text{TeV}\$](#) . JINST 12 P02014, 2017. Cited on page/s [107](#).
- [146] The-CMS-Collaboration. [Measurements of properties of the Higgs boson decaying into the four-lepton final state in pp collisions at \$\sqrt{s} = 13\text{ TeV}\$ analysis of the LHC pp collision data at \$\sqrt{s} = 7\$ and \$8\text{ TeV}\$](#) . JHEP 11 (2017) 047, 2017. Cited on page/s [xv](#), [13](#).
- [147] The-CMS-Collaboration. [Mini-AOD: A New Analysis Data Format for CMS](#). Journal of Physics: Conference Series 664 (2015) 7 072052, 2017. Cited on page/s [84](#).
- [148] The-CMS-Collaboration. [The Phase-2 Upgrade of the CMS Muon Detectors](#). CERN-LHCC-2017-012 ; CMS-TDR-016, 2017-2021. Cited on page/s [52](#), [53](#), [54](#), [59](#), [60](#).
- [149] The-CMS-Collaboration. [The Phase-2 Upgrade of the CMS Tracker](#). CERN-LHCC-2017-009 ; CMS-TDR-014, 2017-2021. Cited on page/s [44](#), [52](#).
- [150] The-CMS-Collaboration. [The Phase-2 Upgrade of the CMS DAQ Interim Technical Design Report](#). CERN/LHCC 2007-021 LHCC-G-134, 2017-2021. Cited on page/s [49](#), [52](#).
- [151] The-CMS-Collaboration. [Jet algorithms performance in 13 TeV data](#). CMS-PAS-JME-16-003, 2018. Cited on page/s [107](#).
- [152] The-CMS-Collaboration. [Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV](#). JINST 13, P05011, CMS-BTV-16-002, 2018. Cited on page/s [107](#).
- [153] The-CMS-Collaboration. [Observation of Higgs Boson Decay to Bottom Quarks](#). Phys. Rev. Lett. 121, 121801, 2018. Cited on page/s [xix](#), [111](#), [112](#).
- [154] The-CMS-Collaboration. [Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at \$\sqrt{s} = 13\text{TeV}\$](#) . JINST 13 (2018) P06015, 2018. Cited on page/s [104](#), [105](#).
- [155] The-CMS-Collaboration. [Projected performance of CMS Cathode Strip Chambers at HL-LHC](#). doi:10.22323/1.321.0074, 2018. Cited on page/s [48](#).
- [156] The-CMS-Collaboration. . Journal of the Korean Physical Society volume 73, pages 1080-1087, 2018. Cited on page/s [61](#).

- [157] The-CMS-Collaboration. [Higgs Physics at the HL-LHC and HE-LHC](#). CERN-LPCC-2018-04, arXiv:1902.00134 [hep-ph], 2019. Cited on page/s [xxii](#), [xxv](#), [155](#), [156](#), [157](#).
- [158] The-CMS-Collaboration. [Performance of Electron and Photon reconstruction in the 2017 Legacy dataset of the CMS Experiment](#). CMS-DP-2020-024 ; CERN-CMS-DP-2020-024, 2020. Cited on page/s [101](#).
- [159] The-CMS-Collaboration. [CMS phase-II upgrade with GEM detector](#). Journal of Physics: Conference Series, Volume 1498, 2020. Cited on page/s [xvii](#), [51](#).
- [160] The-CMS-Collaboration. [CMS RPC Background – Studies and Measurements](#). arXiv:2005.12769 [physics.ins-det], 2020. Cited on page/s [59](#).
- [161] The-CMS-Collaboration. [Measurements of properties of the Higgs boson in the four-lepton final state at \$\sqrt{s} = 13\text{TeV}\$ with full Run II data \(paper\)](#). CMS AN-2019/139, 2021. Cited on page/s [xxiv](#), [113](#), [114](#), [115](#), [116](#), [118](#), [120](#).
- [162] The-CMS-Collaboration. [Measurements of properties of the Higgs boson in the four lepton final state at \$\sqrt{s} = 13\text{TeV}\$ with full Run II data](#). CMS AN-2019/139, 2021. Cited on page/s [xix](#), [102](#), [105](#), [106](#).
- [163] The-CMS-Collaboration. [CMS PDF Forum](#). 2021. Cited on page/s [79](#).
- [164] The-CMS-Collaboration. [CMS Physics : Technical Design Report Volume 1: Detector Performance and Software](#). CERN-LHCC-2006-001 ; CMS-TDR-8-1, 2021. Cited on page/s [42](#).
- [165] The-CMS-Collaboration. [CMS Technical Design Report for the Level-1 Trigger Upgrade](#). CMS Public, 2021. Cited on page/s [49](#).
- [166] The-CMS-Collaboration. [Utilities for Accessing Pileup Information for Data](#). CMS internal, 2021. Cited on page/s [91](#).
- [167] The-CMS-Collaboration. [SM Higgs Branching Ratios and Total Decay Widths](#). CMS internal, 2021. Cited on page/s [xxiii](#), [14](#), [108](#).
- [168] The-CMS-Collaboration. [LHC Higgs Cross Section HH Sub-group \(a.k.a LHC-HH\)](#). CMS internal, 2021. Cited on page/s [xxiii](#), [18](#), [113](#), [120](#).
- [169] The-CMS-Collaboration. [A MIP Timing Detector for the CMS Phase-2 Upgrade Technical Design Report](#). CMS Public, 2021. Cited on page/s [52](#).
- [170] The-LHCb-Collaboration. [LHCb sees a new flavour of matter-antimatter asymmetry](#). LHCb Public, 2021. Cited on page/s [41](#).