

SEARCH FOR NON-RESONANT HIGGS BOSON PAIR PRODUCTION IN THE FOUR  
BOTTOM QUARK DECAY CHANNEL WITH THE CMS EXPERIMENT

By

DANIEL FERNANDO GUERRERO IBARRA

CERN-THESIS-2021-218



A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL  
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT  
OF THE REQUIREMENTS FOR THE DEGREE OF  
DOCTORATE OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2021

© 2021 Daniel Fernando Guerrero Ibarra

To my parents and my grandparents

## ACKNOWLEDGEMENTS

First, I would like to thank my advisor, Jaco Konigsberg, for giving me his support, encouragement and advice since the very beginning of my doctoral studies. I want to thank the University of Florida CMS group members that have accompanied me throughout these years. I am especially grateful with Luca Cadamuro, the HH guru, for his constant guidance, patience, teaching, and friendship. Moreover, I would like to thank Darin Acosta for his kind help and advice while I was based at CERN. I also want to thank the members of my dissertation committee for their review of this work and insightful discussions: Darin Acosta, Andrey Korytov, Konstantin Matchev and Arunava Banerjee.

I thank the University of Florida, the Fermilab LHC Physics Center (LPC), and the U.S. Department of Energy for their support to make this research possible. I would like to thank the CMS Analysis Review members for their comments and approving the result for the public: Maurizio Pierini, Nadya Chernyavskaya, Silvio Donato, Devdatta Majumder, and Don Lincoln. There are also several brilliant people who have helped along the way while working on my Ph.D. research. Special thanks go to Caterina Vernieri, Fabio Ravera, Vladimir Rekovic, Anadi Canepa, Maria Cepeda, Chris Palmer, Mia Liu, Sergo Jindariani, Jia Fu Low, and Javier Duarte. I would like to thank friends and colleagues who enriched my time at UF, Fermilab and CERN: Vladimir Martinez, Atul Divakarla, Jake Rosenzweig, Suzanne Rosenzweig, Jorge Martinez, Osvaldo Miguel, Chad Freer, Xunwu Zuo, Antonio di Maria, Daniel Guest, Andres Delanoy, Raffaele Gerosa, Elisa Fontanesi, Deborah Pinna, and Camilla Galloni. I would also like to thank my undergraduate advisors, Edgar Carrera and Francisco Yumiceva, for their support through my career as a particle physicist.

Lastly, I would like to thank my family. I will always be grateful to my parents Fernando and Janeth and my sister Nicole, for their endless encouragement, trust, love and cheers through these years living abroad. I would also like to thank the Williams family for making me a part of the family. And Michelle, thank you for being my source of support, love and inspiration.



# TABLE OF CONTENTS

|                                                         | <u>page</u> |
|---------------------------------------------------------|-------------|
| ACKNOWLEDGEMENTS .....                                  | 4           |
| LIST OF TABLES.....                                     | 9           |
| LIST OF FIGURES.....                                    | 15          |
| LIST OF ABBREVIATIONS.....                              | 16          |
| ABSTRACT .....                                          | 17          |
| CHAPTER                                                 |             |
| 1 INTRODUCTION .....                                    | 18          |
| 1.1 Overview.....                                       | 18          |
| 1.2 Brief Review of the Standard Model .....            | 20          |
| 1.2.1 Elementary Particles and Interactions.....        | 20          |
| 1.2.2 The Fermion and Gauge Sectors .....               | 22          |
| 1.2.3 The Scalar Sector .....                           | 23          |
| 1.2.4 The Higgs Boson.....                              | 24          |
| 1.3 Higgs Boson Pair Production .....                   | 29          |
| 1.3.1 Standard Model Production .....                   | 29          |
| 1.3.2 Beyond the Standard Model Production .....        | 31          |
| 1.4 Searches for Higgs Boson Pair Production.....       | 36          |
| 1.4.1 Decay Channels.....                               | 36          |
| 1.4.2 Experimental Status at the LHC.....               | 37          |
| 1.4.3 Future Prospects at the High Luminosity LHC ..... | 40          |
| 2 EXPERIMENTAL APPARATUS.....                           | 43          |
| 2.1 The Large Hadron Collider.....                      | 43          |
| 2.2 The CMS Experiment.....                             | 47          |
| 2.2.1 Description .....                                 | 47          |
| 2.2.2 Detector Layers .....                             | 48          |
| 2.2.3 Event Reconstruction.....                         | 56          |
| 2.2.4 Trigger System.....                               | 65          |
| 2.2.5 Upgrades in Preparation for the HL-LHC.....       | 67          |
| 3 ANALYSIS STRATEGY.....                                | 70          |
| 3.1 Challenges and Methods .....                        | 70          |
| 3.2 Dataset.....                                        | 72          |
| 3.3 Trigger.....                                        | 73          |
| 3.3.1 Requirements .....                                | 73          |
| 3.3.2 Efficiency Measurement .....                      | 74          |
| 3.4 Event Preselection .....                            | 77          |
| 3.4.1 Isolated Lepton Veto.....                         | 77          |
| 3.4.2 Higgs Boson Pair Candidate .....                  | 77          |
| 3.4.3 Vector Boson Fusion Jet Pair Candidate .....      | 81          |

|          |                                                         |     |
|----------|---------------------------------------------------------|-----|
| 3.5      | Event Categorization .....                              | 83  |
| 3.5.1    | Motivation .....                                        | 83  |
| 3.5.2    | Production Mode Score: ggFKiller .....                  | 83  |
| 3.5.3    | Performance of the Production Mode Categorization ..... | 86  |
| 3.6      | Event Regions .....                                     | 88  |
| 3.6.1    | The ‘4b’ and ‘3b’ Regions .....                         | 88  |
| 3.6.2    | Analysis and Validation Regions .....                   | 89  |
| 3.7      | Event Subcategorization .....                           | 92  |
| 3.7.1    | Subcategory Boundaries .....                            | 92  |
| 3.7.2    | Signal Extraction Observables .....                     | 96  |
| 4        | MODELING OF PHYSICS PROCESSES .....                     | 101 |
| 4.1      | Monte Carlo Simulation .....                            | 101 |
| 4.1.1    | Motivation .....                                        | 101 |
| 4.1.2    | Simulated Samples .....                                 | 102 |
| 4.1.3    | HH Signal Model .....                                   | 103 |
| 4.2      | Data-driven Background Model .....                      | 106 |
| 4.2.1    | Motivation .....                                        | 106 |
| 4.2.2    | Boosted Decision Tree Reweighting .....                 | 107 |
| 4.2.3    | Implementation .....                                    | 108 |
| 4.2.4    | Validation Studies .....                                | 115 |
| 4.2.5    | Additional Checks .....                                 | 122 |
| 4.3      | Systematic Uncertainties .....                          | 126 |
| 4.3.1    | Signal Uncertainties .....                              | 126 |
| 4.3.2    | Background Model Uncertainties .....                    | 127 |
| 4.3.3    | Expected Impact of Nuisance Parameters .....            | 130 |
| 5        | RESULTS .....                                           | 132 |
| 5.1      | Statistical Analysis .....                              | 132 |
| 5.1.1    | Statistical Model .....                                 | 132 |
| 5.1.2    | Validation of the Statistical Model .....               | 135 |
| 5.1.3    | Observed Distributions .....                            | 136 |
| 5.2      | Upper Limits .....                                      | 139 |
| 5.3      | Likelihood Scans .....                                  | 142 |
| 6        | CONCLUSIONS .....                                       | 144 |
| APPENDIX |                                                         |     |
| A        | TRIGGER EFFICIENCY MEASUREMENT .....                    | 147 |
| B        | GGFKILLER SCORE .....                                   | 162 |
| C        | SIGNAL EFFICIENCIES VS HIGGS BOSON COUPLINGS .....      | 167 |
| D        | LIST OF SIMULATION SAMPLES .....                        | 169 |

|     |                                         |     |
|-----|-----------------------------------------|-----|
| E   | SIGNAL MODEL .....                      | 172 |
| E.1 | Gluon Fusion Production.....            | 172 |
| E.2 | Vector Boson Fusion Production.....     | 175 |
| F   | BACKGROUND MODELING .....               | 176 |
| F.1 | BDT Reweighting Hyperparameters .....   | 176 |
| F.2 | Analysis Control Region Results.....    | 177 |
| F.3 | Validation Control Region Results ..... | 184 |
| F.4 | Validation Signal Region Results .....  | 191 |
| G   | SIGNAL OBSERVABLES .....                | 198 |
|     | REFERENCES .....                        | 204 |
|     | BIOGRAPHICAL SKETCH .....               | 215 |

## LIST OF TABLES

| <u>Tables</u>                                                                                                                               | <u>page</u> |
|---------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| 1-1 BEH Lagrangian terms after the spontaneous electroweak symmetry breaking (EWSB)....                                                     | 24          |
| 1-2 Higgs boson production cross section and relative uncertainties .....                                                                   | 25          |
| 1-3 Higgs boson branching ratios and relative uncertainties for several Higgs boson<br>( $m_H = 125$ GeV) decay modes .....                 | 26          |
| 1-4 Higgs boson pair production cross section (in pb) and relative uncertainties .....                                                      | 29          |
| 1-5 Effective field theory benchmarks for LHC searches for non-resonant Higgs boson pair<br>production .....                                | 35          |
| 1-6 ATLAS and CMS expected significance (in standard deviation) for SM Higgs pair pro-<br>duction in different channels at the HL-LHC ..... | 41          |
| 2-1 Parameters used in the calculation of the instantaneous luminosity at the Large Hadron<br>Collider proton-proton collisions .....       | 45          |
| 3-1 Summary details of the datasets used for analysis .....                                                                                 | 72          |
| 3-2 The full Run-2 dataset per year and its integrated luminosity and uncertainty .....                                                     | 73          |
| 3-3 HLT Trigger Paths used in the analysis by year .....                                                                                    | 73          |
| 3-4 Summary of requirements for isolated electrons and muons .....                                                                          | 77          |
| 3-5 Summary of jet properties required for b-jet candidates .....                                                                           | 78          |
| 3-6 Summary of jet kinematic and quality requirements for VBF-jet candidates .....                                                          | 81          |
| 3-7 Event yields of data and MC simulation after preselection requirements .....                                                            | 82          |
| 3-8 Variables used to train the ggF Killer score .....                                                                                      | 85          |
| 3-9 Event yields of data and MC simulation after categorization into ggF and VBF categories                                                 | 86          |
| 3-10 Event yields of data and MC simulation after categorization requirements in the 3b region                                              | 88          |
| 3-11 Selection requirements used to define the event subcategorization .....                                                                | 92          |
| 3-12 BDT input variables in ggF Category 1 and ggF Category 2 .....                                                                         | 96          |
| 4-1 Couplings variations used for the ggF and VBF signal models .....                                                                       | 104         |
| 4-2 Summary of the quantities involved in the estimation of the background normalization in<br>the 2016 data analysis .....                 | 111         |
| 4-3 Summary of the quantities involved in the estimation of the background normalization in<br>the 2017-2018 data analysis .....            | 111         |
| 4-4 BDT reweighter input variables in ggF Category 1 and ggF Category 2 .....                                                               | 113         |
| 4-5 BDT reweighter input variables in VBF Category 1 .....                                                                                  | 113         |
| 4-6 Summary of the transfer factor and normalization estimation in the 2016 validation test                                                 | 116         |

|      |                                                                                                                             |     |
|------|-----------------------------------------------------------------------------------------------------------------------------|-----|
| 4-7  | Summary of the transfer factor and normalization estimation in the 2017-2018 validation test.....                           | 116 |
| 4-8  | Data-background model compatibility p-value obtained from the goodness of fit test in the validation region .....           | 119 |
| 4-9  | Summary of the background model uncertainties in the 2016 analysis.....                                                     | 130 |
| 4-10 | Summary of the background model uncertainties in the 2017-2018 analysis.....                                                | 130 |
| 5-1  | Data-background model compatibility p-value obtained from the goodness of fit test in the analysis region observables ..... | 136 |
| A-1  | Filters of the HLT_DoubleJet90_Double30_TripleBTagCSV trigger path in the 2016 dataset .....                                | 148 |
| A-2  | Filters of the HLT_QuadJet45_TripleBTagCSV trigger path in the 2016 dataset.....                                            | 148 |
| A-3  | Filters of the HLT_PFHT300PT30_QuadPFJet_75_60_45_40_TriplePFBTagCSV trigger path in 2017 dataset.....                      | 148 |
| A-4  | Filters of the HLT_PFHT330PT30_QuadPFJet_75_60_45_40_TriplePFBTagDeepCSV trigger path in the 2018 dataset .....             | 149 |
| A-5  | Physical meaning of the parameters of the CRYSTAL x ERF fitting function.....                                               | 149 |
| B-1  | Best hyperparameters of the ggF training .....                                                                              | 162 |
| D-1  | Summary of the MC samples used for the 2016 conditions .....                                                                | 169 |
| D-2  | Summary of the MC samples used for the 2017 conditions .....                                                                | 170 |
| D-3  | Summary of the MC samples used for the 2018 conditions .....                                                                | 171 |
| F-1  | Best BDT-reweighter hyperparameters in 2016 analysis control region .....                                                   | 176 |
| F-2  | Best BDT-reweighter hyperparameters in 2017-2018 analysis control region.....                                               | 176 |
| F-3  | Best BDT-reweighter hyperparameters in 2016 validation test.....                                                            | 176 |
| F-4  | Best BDT-reweighter hyperparameters in 2017-2018 validation test .....                                                      | 176 |
| G-1  | Best hyperparameters of the BDT output training in ggF categories .....                                                     | 200 |
| G-2  | Performance of the BDT output training in Train and Test samples for the classifiers C1 and C2 .....                        | 203 |

## LIST OF FIGURES

| <u>Figures</u>                                                                                                                                                              | <u>page</u> |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| 1-1 Standard Model of elementary particles and interactions .....                                                                                                           | 21          |
| 1-2 Predicted interactions of the Higgs boson with fermions and vector bosons.....                                                                                          | 25          |
| 1-3 Examples of leading order diagrams contributing to the Higgs boson production .....                                                                                     | 26          |
| 1-4 Selected CMS Higgs physics results in Run-2 .....                                                                                                                       | 28          |
| 1-5 Differential cross section distribution as a function of the Higgs boson invariant mass ....                                                                            | 30          |
| 1-6 Examples of leading order (LO) diagrams contributing to the Higgs boson pair production<br>mode mechanisms .....                                                        | 31          |
| 1-7 Diagrams of production of a new resonances decaying to a Higgs boson pair at the LHC..                                                                                  | 32          |
| 1-8 Higgs boson pair production cross section at $\sqrt{s} = 14$ TeV LHC proton-proton collisions<br>for various modes.....                                                 | 33          |
| 1-9 Normalized Higgs boson pair invariant mass for different values of the coupling modifiers                                                                               | 34          |
| 1-10 Leading order diagrams of contact interactions .....                                                                                                                   | 34          |
| 1-11 Higgs boson pair invariant mass distribution for different leading order effective field<br>theory (EFT) shape benchmarks: A) Nodes 1 to 6, and B) Nodes 7 to 12. .... | 35          |
| 1-12 Higgs boson pair branching fractions assuming a Higgs mass $m_H = 125$ GeV .....                                                                                       | 36          |
| 1-13 95% CL observed and expected upper limits on the Higgs boson pair production cross<br>section as function of the coupling modifiers .....                              | 39          |
| 1-14 Signal observables used in searches for HH production .....                                                                                                            | 40          |
| 1-15 Minimum negative-log-likelihood as a function of the $\kappa_\lambda$ modifier obtained by perform-<br>ing a fit to the expected background and SM signal .....        | 42          |
| 2-1 The CERN Accelerator Complex .....                                                                                                                                      | 43          |
| 2-2 The Large Hadron Collider schedule starting from Run-1 operations (2011) to the<br>next decades and beyond.....                                                         | 45          |
| 2-3 CMS operation conditions in proton-proton collisions during the LHC Run-2 data-taking.                                                                                  | 46          |
| 2-4 The CMS conventional coordinate system.....                                                                                                                             | 48          |
| 2-5 Three-dimensional view of the subdetectors of the CMS experiment.....                                                                                                   | 49          |
| 2-6 Layout of the CMS tracking system in the longitudinal direction .....                                                                                                   | 50          |
| 2-7 The CMS pixel detector .....                                                                                                                                            | 51          |
| 2-8 The CMS electrognetic calorimeter (ECAL).....                                                                                                                           | 52          |
| 2-9 Longitudinal layout of the CMS hadronic calorimeter (HCAL) .....                                                                                                        | 53          |
| 2-10 Three-dimensional artistic view of the CMS magnet .....                                                                                                                | 54          |

|                                                                                                                                                               |    |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 2-11 One quadrant of the CMS muon system in the longitudinal direction.....                                                                                   | 55 |
| 2-12 A transverse slice of the CMS detector representing the interaction of stable particles.....                                                             | 57 |
| 2-13 Reconstruction of a di-jet in a simulated event without pile-up.....                                                                                     | 61 |
| 2-14 Heavy-flavour (b quark or c quark) jet kinematic properties .....                                                                                        | 62 |
| 2-15 Performance of the DeepJet algorithm with respect to the DeepCSV algorithm.....                                                                          | 63 |
| 2-16 The CMS Level-1 trigger architecture during the 2016-2018 operations.....                                                                                | 67 |
| 2-17 The Phase-2 CMS muon system .....                                                                                                                        | 69 |
| 2-18 The Phase-2 Upgrade of the CMS Level-1 trigger architecture.....                                                                                         | 69 |
| 3-1 Examples of the measured trigger filter efficiency in single muon ( $\mu$ ) data and top-quark pair ( $t\bar{t}$ ) MC simulation events .....             | 75 |
| 3-2 Fraction of SM ggF signal events that are attribute a specific value of the trigger scale factor as function of the Higgs boson pair mass. ....           | 76 |
| 3-3 Closure test of the trigger simulated efficiency in the SM ggF signal events .....                                                                        | 76 |
| 3-4 Sketch of the pairing method using the two-dimensional plane of the two reconstructed Higgs boson masses.....                                             | 79 |
| 3-5 Invariant mass of the two reconstructed Higgs bosons in SM ggF HH simulated events ...                                                                    | 80 |
| 3-6 Purity of the lowest transverse momentum VBF jet candidate .....                                                                                          | 82 |
| 3-7 Distribution of the ggFKiller input variables in 2016 simulation .....                                                                                    | 84 |
| 3-8 The ggFKiller output distributions .....                                                                                                                  | 85 |
| 3-9 Performance of the ggFKiller in the production mode classification as a function of the couplings.....                                                    | 87 |
| 3-10 Definition of signal and control regions for the analysis and validation regions .....                                                                   | 89 |
| 3-11 Distribution of the preselected signal events in the $m_{H_1}$ - $m_{H_2}$ plane.....                                                                    | 90 |
| 3-12 Selection efficiency ( $\epsilon$ ) of the 2018 selections versus the ggF and VBF couplings .....                                                        | 91 |
| 3-13 Variables used to create ggF and VBF subcategories in the ggF and VBF signal region ...                                                                  | 93 |
| 3-14 Distribution of data events in the ggF category 1 in the $(m_{H_1}, m_{H_2})$ plane for the $V_{CR}(3b)$ region using different $m_{HH}$ thresholds..... | 94 |
| 3-15 Performance of the production mode subcategorization as function of the ggF and VBF couplings.....                                                       | 95 |
| 3-16 Input variables for the BDT output training in the 2017-2018 ggF Category 1.....                                                                         | 97 |

|      |                                                                                                                                                               |     |
|------|---------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 3-17 | Input variables for the BDT output training in the 2017-2018 ggF Category 2 .....                                                                             | 98  |
| 3-18 | Compatibility checks on the classifiers used for the 2017-2018 ggF observables .....                                                                          | 100 |
| 4-1  | Sketch of processes involved in a proton-proton collision .....                                                                                               | 102 |
| 4-2  | Higgs boson pair production cross section times the branching fraction into four bottom quarks as a function of the couplings using the signal modeling ..... | 105 |
| 4-3  | Dependence of the normalization transfer factor on the $m_{\parallel}$ variable.....                                                                          | 110 |
| 4-4  | Example of the discriminant check step in the model for analysis control region data ....                                                                     | 114 |
| 4-5  | Example of the data versus background modeling agreement in 20172018 ggF category 2115                                                                        |     |
| 4-6  | Example of the discriminant check step in the model for validation control region data ..                                                                     | 117 |
| 4-7  | Example of the data versus background modeling agreement in 20172018 ggF category 2 validation control region.....                                            | 118 |
| 4-8  | Example of the data versus background modeling agreement in 20172018 ggF category 2 validation control region.....                                            | 118 |
| 4-9  | Distributions of the observables in all validation signal regions for 2016 data.....                                                                          | 120 |
| 4-10 | Distributions of the observables in all validation signal regions for 2017-2018 data .....                                                                    | 121 |
| 4-11 | Example of the self-bias impact on fitted signal strength .....                                                                                               | 123 |
| 4-12 | Example of the self-bias impact on the fitted signal strength pull.....                                                                                       | 124 |
| 4-13 | Impact of the Higgs boson production with associate top-quark pairs in the four observables of the ggF $A_{SR}(4b)$ regions .....                             | 125 |
| 4-14 | Linear fits used to derive a shape uncertainty for VBF category 1 .....                                                                                       | 129 |
| 4-15 | Regions used to determine the background shape uncertainty in ggF categories.....                                                                             | 129 |
| 4-16 | Expected impact of the highest systematic uncertainties on the measurement of the signal strength .....                                                       | 131 |
| 5-1  | Observed impact of the highest systematic uncertainties on the measurement of the signal strength .....                                                       | 135 |
| 5-2  | Observed distribution in the analysis signal regions for the 2016 data .....                                                                                  | 137 |
| 5-3  | Distributions of the observables in the analysis signal regions for the 2017-2018 data....                                                                    | 138 |
| 5-4  | Distribution of the expected signal-to-background ratio for events selected in all categories and datasets .....                                              | 139 |
| 5-5  | Observed and expected 95% CL limits on the standard model Higgs boson pair (HH) production signal strength breakdown by year .....                            | 140 |



|      |                                                                                                                                                          |     |
|------|----------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 5-6  | Observed and expected 95% CL limits on the standard model Higgs boson pair (HH) production signal strength breakdown by category .....                   | 140 |
| 5-7  | Observed and expected 95% CL upper limits on Higgs boson pair cross section as a function of the coupling modifiers .....                                | 141 |
| 5-8  | Observed and expected 95% CL upper limits on the gluon fusion Higgs boson pair production cross section for the 12 benchmarks .....                      | 141 |
| 5-9  | Observed and expected negative log-likelihood scan as a function of the couplings .....                                                                  | 142 |
| 5-10 | Negative log-likelihood contours at 68% and 95% CL as a function of the $(\kappa_\lambda, \kappa_{2V})$ couplings .....                                  | 143 |
| 5-11 | Negative log-likelihood contours at 68% and 95% CL as a function of the $(\kappa_{2V}, \kappa_V)$ couplings .....                                        | 143 |
| A-1  | Efficiency measured in single muon data and $t\bar{t}$ MC for filters of HLT_DoubleJet90_Double30_TripleBTagCSV path in 2016 .....                       | 150 |
| A-2  | Efficiency measured in single muon data and $t\bar{t}$ MC for filters of HLT_QuadJet45_TripleBTagCSV path in 2016 .....                                  | 151 |
| A-3  | Efficiency measured in single muon data and $t\bar{t}$ MC for filters of the path overlap in 2016 .....                                                  | 152 |
| A-4  | Efficiency measured in single muon data and $t\bar{t}$ MC for filters of the HLT_PFHT300PT30_QuadPFJet_75_60_45_40_TriplePFBTagCSV path in 2017 .....    | 153 |
| A-5  | Efficiency measured in single muon data and $t\bar{t}$ MC for filters of the HLT_PFHT330PT30_QuadPFJet_75_60_45_40_TriplePFBTagDeepCSV path in 2018 .... | 154 |
| A-6  | Efficiency fits in SingleMuon filters of HLT_DoubleJet90_Double30_TripleBTagCSV in 2016 .....                                                            | 155 |
| A-7  | Efficiency fits in SingleMuon for filters of HLT_QuadJet45_TripleBTagCSV in 2016 ....                                                                    | 155 |
| A-8  | Efficiency fits in SingleMuon for filters in paths overlap in 2016 .....                                                                                 | 156 |
| A-9  | Efficiency fits in $t\bar{t}$ for filters of HLT_DoubleJet90_Double30_TripleBTagCSV in 2016 ...                                                          | 156 |
| A-10 | Efficiency fits in $t\bar{t}$ for filters of HLT_QuadJet45_TripleBTagCSV in 2016 .....                                                                   | 157 |
| A-11 | Efficiency fits in $t\bar{t}$ for filters in paths overlap in 2016 .....                                                                                 | 157 |
| A-12 | Efficiency fits in SingleMuon for filters of the trigger path in 2017 (1/2) .....                                                                        | 158 |
| A-13 | Efficiency fits in SingleMuon for filters of the trigger path in 2017 (2/2) .....                                                                        | 158 |
| A-14 | Efficiency fits in $t\bar{t}$ for filters of the trigger path in 2017 (1/2) .....                                                                        | 159 |
| A-15 | Efficiency fits in $t\bar{t}$ for filters of the trigger path in 2017 (2/2) .....                                                                        | 159 |
| A-16 | Efficiency fits in SingleMuon for filters of the trigger path in 2018 (1/2) .....                                                                        | 160 |

|      |                                                                                                                                                                       |     |
|------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| A-17 | Efficiency fits in SingleMuon for filters of the trigger path in 2018 (2/2) .....                                                                                     | 160 |
| A-18 | Efficiency fits in $t\bar{t}$ for filters of the trigger path in 2018 (1/2) .....                                                                                     | 161 |
| A-19 | Efficiency fits in $t\bar{t}$ for filters of the trigger path in 2018 (2/2) .....                                                                                     | 161 |
| B-1  | Distribution of the ggFKiller input variables in 2017 simulation .....                                                                                                | 163 |
| B-2  | Distribution of the ggFKiller input variables in 2018 simulation .....                                                                                                | 164 |
| B-3  | The relative importance of the ggFKiller training variables .....                                                                                                     | 165 |
| B-4  | Over-training check of the ggFKiller score .....                                                                                                                      | 166 |
| C-1  | Selection efficiency ( $\epsilon$ ) of the 2016 selections versus the ggF and VBF couplings .....                                                                     | 167 |
| C-2  | Selection efficiency ( $\epsilon$ ) of the 2016 selections versus the ggF and VBF couplings .....                                                                     | 168 |
| F-1  | Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 2016 ggF category 1. ....                         | 178 |
| F-2  | Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 2016 ggF category 2. ....                         | 179 |
| F-3  | Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 2017/2018 ggF category 1.....                     | 180 |
| F-4  | Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 2017/2018 ggF category 2.....                     | 181 |
| F-5  | Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 2016 VBF category 1. ....                         | 182 |
| F-6  | Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 2017/2018 VBF category 1.....                     | 183 |
| F-7  | Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 2016 ggF category 1 in the validation study.....  | 185 |
| F-8  | Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 2016 ggF category 2 in the validation study.....  | 186 |
| F-9  | Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 2017/2018 ggF category 1 in the validation study. | 187 |
| F-10 | Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 2017/2018 ggF category 2 in the validation study. | 188 |
| F-11 | Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 2016 VBF category 1 in the validation study.....  | 189 |

|      |                                                                                                                                                                                        |     |
|------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| F-12 | Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 20172018 VBF category 1 in the validation study.                   | 190 |
| F-13 | Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 2016 ggF category 1 in the validation signal region study. ....    | 192 |
| F-14 | Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 2016 ggF category 2 in the validation signal region study. ....    | 193 |
| F-15 | Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 20172018 ggF category 1 in the validation signal region study..... | 194 |
| F-16 | Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 20172018 ggF category 2 in the validation signal region study..... | 195 |
| F-17 | Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 2016 VBF category 1 in the validation signal region study. ....    | 196 |
| F-18 | Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 20172018 VBF category 1 in the validation signal region study..... | 197 |
| G-1  | Input variables for the BDT output training in the 2016 ggF Category 1. ....                                                                                                           | 199 |
| G-2  | Input variables for the BDT output training in the 2016 ggF Category 2.....                                                                                                            | 200 |
| G-3  | Training performance of the classifier C1 used for building the ggF observables.....                                                                                                   | 201 |
| G-4  | Training performance of the classifier C2 used for building the ggF observables.....                                                                                                   | 202 |
| G-5  | The relative importance of the training variables for classifier C1 .....                                                                                                              | 203 |
| G-6  | Compatibility checks on the classifiers used for the 2016 ggF observables .....                                                                                                        | 204 |

## LIST OF ABBREVIATIONS

|        |                                       |
|--------|---------------------------------------|
| ALICE  | A Large Ion Collider Experiment       |
| ATLAS  | A Toroidal LHC ApparatuS              |
| BDT    | Boosted Decision Tree                 |
| BEH    | Brout Englert Higgs                   |
| BSM    | Beyond the Standard Model             |
| CMS    | Compact Muon Solenoid                 |
| CL     | Confidence Level                      |
| DNN    | Deep Neural Network                   |
| EWSB   | Electroweak Symmetry Breaking         |
| HL-LHC | High Luminosity Large Hadron Collider |
| LHC    | Large Hadron Collider                 |
| LHCb   | Large Hadron Collider beauty          |
| LO     | Leading order                         |
| ML     | Machine learning                      |
| MC     | Monte Carlo                           |
| NLO    | Next to Leading Order                 |
| NNLO   | Next to Next to Leading Order         |
| QCD    | Quantum Chromodynamics                |
| QED    | Quantum Electrodynamics               |
| QFT    | Quantum Field Theory                  |
| SM     | Standard Model                        |

Abstract of Dissertation Presented to the Graduate School  
of the University of Florida in Partial Fulfillment of the  
Requirements for the Degree of Doctorate of Philosophy

SEARCH FOR NON-RESONANT HIGGS BOSON PAIR PRODUCTION IN THE FOUR  
BOTTOM QUARK DECAY CHANNEL WITH THE CMS EXPERIMENT

By

Daniel Fernando Guerrero Ibarra

December 2021

Chair: Jacobo Konigsberg

Cochair: Darin Acosta

Major: Physics

This dissertation presents a search for non-resonant Higgs boson pair production, focusing on the four bottom quark decay channel. It explores the gluon fusion and vector boson fusion production mechanisms. The analysis is performed with a dataset of proton-proton collisions at a center-of-mass energy of 13 TeV, collected by the CMS detector at the LHC, corresponding to an integrated luminosity of around 138 inverse femtobarns. Innovative techniques in the areas of object identification, event categorization, signal identification, and background modeling are used to maximize the analysis sensitivity. No excess of signal events are observed relative to the background-only expectation, and 95% CL upper limits on the production cross section are set. The observed upper limit on the standard model production cross section is set at 3.6 times the theoretical expectation. At the time of writing this dissertation, it is the most stringent constraint at the LHC from an individual channel. Furthermore, the observed constraint on the coupling modifier of the Higgs boson self-interaction is set between -2.3 and 9.4. The observed constraint on the coupling modifier of the di-vector-boson-di-Higgs-boson interaction is set between -0.1 and 2.2.

## CHAPTER 1 INTRODUCTION

### 1.1 Overview

The Standard Model of particle physics (SM) is a theory formalized in the 1960s to describe the properties and interaction forces associated to the fundamental constituents of the Universe's matter. According to the SM,  $10^{-12}$  seconds after the Big Bang, the Universe transitioned into the lowest energy state of the Brout-Englert-Higgs (BEH) energy potential inducing the 'spontaneous' breaking of the electroweak symmetry (EWSB), generating physical masses to the weak force carriers (the W and Z vector bosons). This process is known as the BEH mechanism, which also postulates the existence of a fundamental scalar, the Higgs boson (H). Quarks and leptons, the building blocks of matter, obtained their masses from their interactions with it. Moreover, the Higgs boson interaction with itself is responsible for its own mass ( $m_H$ ).

Experiments at LEP and Tevatron colliders searched for the Higgs boson and excluded the existence of this particle over a broad range of  $m_H$  hypotheses. The CERN Large Hadron Collider (LHC), designed to produce proton-proton (pp) collisions at a center-of-mass energy ( $\sqrt{s}$ ) of 14 TeV, was built with the goal of finding it. In July 4th 2012, the ATLAS and Compact Muon Solenoid (CMS) experiments at the LHC announced its discovery by analyzing the data collected during the LHC Run-1 operation from 2010 to 2012. This ground-breaking milestone found the last predicted SM particle. It is the first confirmation of the existence of the BEH mechanism.

The SM is a beautiful and successful description of nature. The SM predictions have been corroborated at particle colliders such as LEP, Tevatron, and LHC. However, it does not provide explanations for some SM features. For instance, in the scalar sector, the Higgs boson mass does not have a fundamental symmetry that projects it from quadratically divergent corrections, known as the Hierarchy problem. It does not explain why there are three fermion families. Furthermore, it does not describe some astrophysical and cosmological observations, such as the gravitational effects from non-visible matter or 'dark' matter, and the matter-antimatter asymmetry in the Universe. These theoretical aspects and observations suggest that the SM model is not a complete theory, and physics beyond the Standard Model (BSM) exists with undiscovered particles and

new interactions. The predictions of BSM physics scenarios that may explain the aforementioned SM shortcomings are currently being tested at the LHC.

Since the Higgs boson discovery, a successful physics program has been carried out by the CMS and ATLAS experiments. Within the current experimental precision, physicists have established that this particle behaves as in the SM by measuring its properties (e.g.  $m_H \sim 125$  GeV) and couplings to vector bosons and the heaviest fermions (bottom ( $b$ ) quark, top ( $t$ ) quark, tau ( $\tau$ ) lepton, and muon). The Higgs self-interaction coupling or ‘self-coupling’  $\lambda_{HHH}$  is one of the most important parameters of the SM because it is intrinsically connected with the shape of the BEH potential. As the Higgs boson mass is known, then the SM value of  $\lambda_{HHH}$  is fully determined ( $\lambda_{HHH}^{SM} \sim 0.13$ ). Any measured deviations could have profound consequences in our understanding of the EWSB, and demonstrate new physics hidden in the scalar sector.

The value of  $\lambda_{HHH}$  can be extracted from the measurement of the Higgs boson pair (HH) production cross section at the LHC. Unfortunately, this cross section is very small, and thus it is a very elusive process to measure with the current amount of data. However, BSM physics scenarios predict modifications of the Higgs couplings values or even new couplings, which may enhance the production cross section and kinematics properties, increasing our chances to discover them in the current datasets. With this goal in mind, CMS and ATLAS researchers have performed searches for HH production via the gluon fusion mode in various decay channels using the data collected during the LHC Run-1 operation at  $\sqrt{s} = 8$  TeV and the LHC 2016 operations at  $\sqrt{s} = 13$  TeV. Thus far, these searches found no evidence of new physics and set constraints on the values of the couplings. The analysis of the data collected during LHC 2016-2018 operations, known as Run-2, is ongoing in multiple channels and production modes.

This dissertation presents the work on the search for HH events produced by the gluon fusion and vector fusion production mechanisms using the data collected by the CMS detector during Run-2. The analysis focuses on the four bottom quark decay channel ( $HH \rightarrow b\bar{b}b\bar{b}$ ). Although this is the most frequent HH decay channel, its experimental study is affected by the overwhelming production multijet background processes, which exceed that of the signal by

several orders of magnitude. Therefore, the development of dedicated analysis techniques is crucial for this search. It implements novel techniques in the areas of object identification, event categorization, signal identification, and background modeling. The results are interpreted in terms of both SM and anomalous production cross sections and couplings. They are currently public in the form of a CMS Physics Analysis Summary (PAS) [1], and internally documented in the CMS Analysis Note 2019-250. At the time of writing this dissertation, this work has the best observed limit on the SM production cross section from an individual channel at the LHC. A paper draft is under internal review for journal submission.

This dissertation is organized in six chapters. Chapter 1 introduces the theoretical concepts, phenomenology, and experimental status and prospects of the Higgs boson pair production at the LHC. Chapter 2 discusses the experimental apparatus used to collect the data for this investigation. Chapter 3 describes the strategy followed to analyze events and optimize the search for the HH signature in data. Chapter 4 presents the description of the Monte Carlo (MC) simulation used to model signal events and the data-driven method used to model the background events. The results obtained in the observed collision data are presented in Chapter 5. Lastly, Chapter 6 summarizes the concluding remarks and future prospects of this work.

## 1.2 Brief Review of the Standard Model

The SM is a relativistic quantum field theory (QFT) that describes how elementary fermion or ‘matter’ fields interact among themselves through the electromagnetic, weak and strong interactions mediated by gauge bosons fields. It provides with a mechanism for mass generation through a scalar field sector. The underlying foundation of the SM is built on the concepts of local gauge symmetry and spontaneous symmetry breaking. It does not include a description of the gravitational interaction, which has negligible impact at the subnuclear scales.

### 1.2.1 Elementary Particles and Interactions

Before introducing the SM formalism, we briefly review the observed properties of the elementary particles. A particle is a localized excitation or quanta associated to the quantum field. The elementary fermions are quarks and leptons. There are six quarks (q): up (u), down (d),



charm (c), strange (s), top (t) and bottom (b). There are three electrically charged leptons ( $\ell^-$ ) and their corresponding neutrinos ( $\nu_\ell$ ): electron ( $e^-$ ), muon ( $\mu^-$ ), tau ( $\tau^-$ ), e-neutrino ( $\nu_e$ ),  $\mu$ -neutrino ( $\nu_\mu$ ) and  $\tau$ -neutrino ( $\nu_\tau$ ). The strong, weak and electromagnetic force carriers are the gluons (g), massive vector bosons ( $W^\pm$  and Z) and the photon ( $\gamma$ ), respectively. The last piece is the Higgs boson (H), which is the only elementary scalar and associated to the generation of mass.

A summary of the properties and interactions of the SM elementary particles are presented in Figure 1-1. Quarks interact through the strong, weak and electromagnetic interactions, whereas leptons only participate in weak and electromagnetic (if electrically charged) interactions. Fermions can be grouped in three families or generations, which have carbon-copy properties and interactions except for their masses. Charged fermions have their respective anti-matter counterpart, but properties such as the electric charge are opposite.

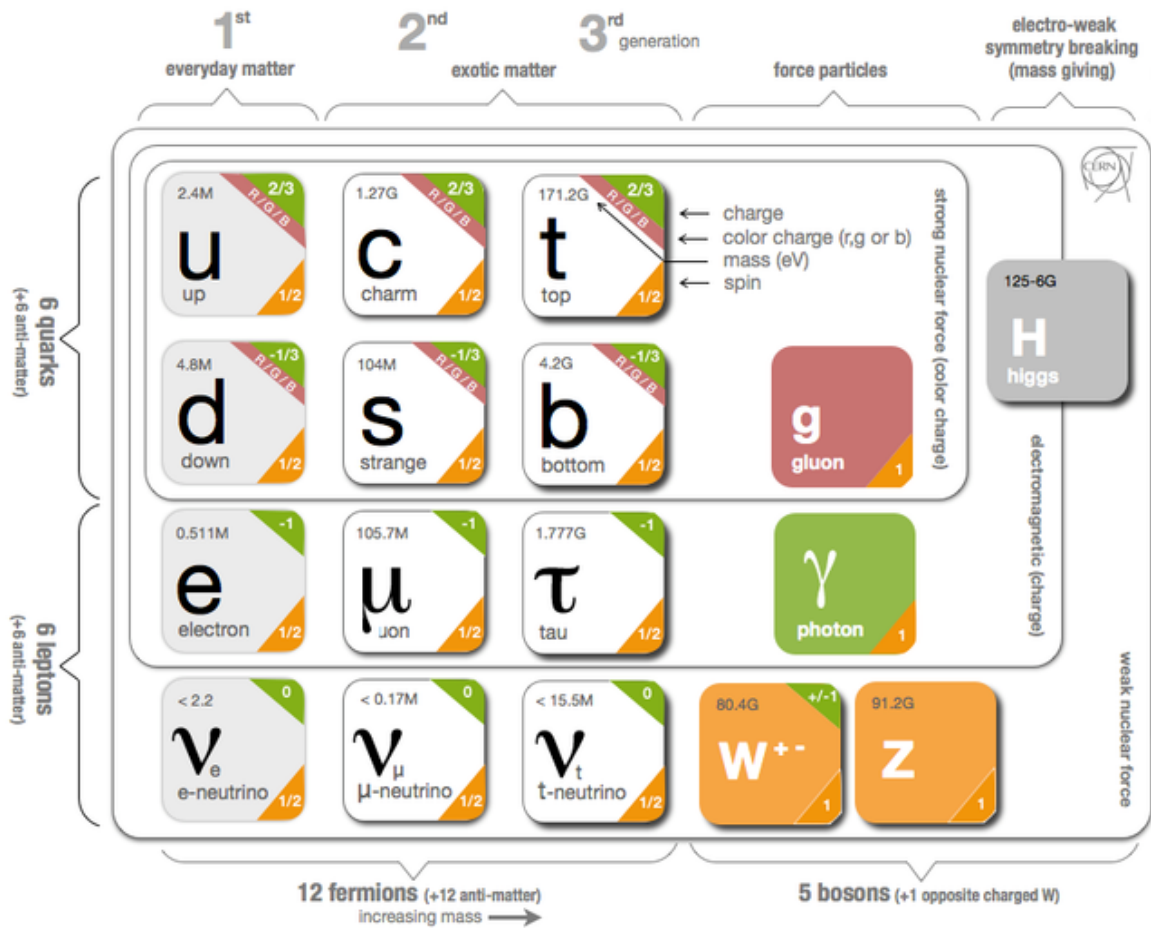


Figure 1-1. Standard model of elementary particles and interactions [2].

### 1.2.2 The Fermion and Gauge Sectors

In QFT, the requirement of gauge invariance or symmetry of the Lagrangian density under local phase transformations of a mathematical group introduces the presence of massless boson fields, which are interpreted as the quanta mediating the interaction. One example of a gauge QFT is quantum electrodynamics (QED). The predictions of QED have been successfully tested by very high-precision experiments [3]. The Lagrangian density of the electron field ( $\Psi$ ) interacting with an electromagnetic field is defined as  $\mathcal{L}_{\text{QED}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + i\bar{\Psi}D_\mu\gamma^\mu\Psi - m_e\bar{\Psi}\Psi$ , where  $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$  is the electromagnetic field strength,  $A_\mu$  is the photon field introduced by gauge invariance,  $D_\mu = \partial_\mu - iqA_\mu$  is the covariant derivative, and  $q$  is the absolute value of electron electric charge. The  $\mathcal{L}_{\text{QED}}$  Lagrangian density is invariant under the  $U(1)_Q$  group local transformations of the electron and photon fields:  $\Psi'(x) = e^{iq\alpha(x)}\Psi(x)$  and  $A'_\mu = A_\mu - \frac{1}{q}\partial_\mu\alpha(x)$ , respectively. Note that  $\mathcal{L}_{\text{QED}}$  is not invariant if there is mass term associated to the photon field is imposed (e.g.  $\frac{1}{2}m_A^2 A^\mu A_\mu$ ) [4, 5, 6].

The SM is a generalization of the fundamental interactions under the  $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$  group. The strong interaction is studied by quantum chromodynamics (QCD) based on the  $SU(3)_C$  color group. In QCD, colored fermion fields (quarks) represented by  $SU(3)$  color triplets  $q = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix}$  interact via 8 massless gauge boson fields (gluons) corresponding to the  $SU(3)_C$  generators. Both quarks and gluons have three possible color charges: red (r), blue (b) and green (g). The strength of the interaction is quantified by the strong coupling constant  $g_s$  (or  $\alpha_s = \frac{g_s^2}{4\pi}$ ). An important result of the QCD formulation is a phenomenon known as asymptotic freedom. At low-energy scales ( $Q^2$ ), the value of the  $\alpha_s$  strength increases, dictating the confinement of quarks and gluons within color-neutral composite particles named as hadrons (e.g. the proton). At high energies scales, the  $\alpha_s$  strength decreases ( $\alpha_s \ll 1$ ), allowing the experimental study of the hadron constituents often named partons. The inner structure of hadrons are made of valence quarks, gluons and quark-antiquark pairs, which are described by the Parton Distribution Function (PDF). There are no free quarks observed in nature because they almost instantaneously ( $10^{-24} - 10^{-23}$ s) become part of hadrons within a

collimated shower of particles (hadronic jets) through a process named hadronization (except the top quark which decays at  $\sim 10^{-25}$  s) [5, 6].

The electromagnetic and weak interactions are studied in a unified manner by the electroweak (EW) theory based on the  $SU(2)_L \otimes U(1)_Y$  group [7, 8, 9]. Experimental measurements show that only fermions with left-handed chirality interact via charged weak interactions. The right-handed (R) and left-handed (L) chirality components of a fermion field  $\Psi$  are obtained using projection operators as  $\Psi_{R/L} = \frac{1 \pm \gamma^5}{2} \Psi$ . In the EW formalism, the massless gauge bosons are  $W_\mu^i$  ( $i = 1, 2, 3$ ) and  $B_\mu$ . They mediate interactions for fermions with non-zero isospin  $I$  and hypercharge  $Y$ , respectively. The relation between electric charge ( $Q$ ), third component of isospin ( $I_3$ ) and hypercharge ( $Y$ ) is  $Q = I_3 + \frac{Y}{2}$ . The left-handed fermions ( $I \neq 0$ ) are represented by the  $SU(2)$  doublets  $\begin{pmatrix} \nu_\ell \\ \ell^- \end{pmatrix}_L$  and  $\begin{pmatrix} u \\ d \end{pmatrix}_L$ , respectively. The right-handed fermions ( $I = 0$ ) are represented by  $SU(2)$  singlets ( $\ell_R^-$ ,  $u_R$  and  $d_R$ ).

The experimental short range nature of weak interactions implies the existence of massive vector bosons. Furthermore, observed fermions like the electron are massive particles. However, explicit mass terms associated to the gauge boson and fermion fields are forbidden in the EW Lagrangian density due to  $SU(2)_L \otimes U(1)_Y$  gauge symmetry. Therefore, the formulation needs a gauge invariant mechanism to generate these masses [5, 6].

### 1.2.3 The Scalar Sector

The Brout-Englert-Higgs (BEH) mechanism solves elegantly the mass-generation paradigm without violating gauge symmetry [10, 11, 12]. It adds a complex scalar doublet  $\Phi$  with a Lagrangian density  $\mathcal{L}_{\text{BEH}} = \mathcal{D}_\mu^\dagger \Phi \mathcal{D}^\mu \Phi - V(\Phi)$ , where  $V(\Phi) = -\mu^2(\Phi^\dagger \Phi) + \lambda(\Phi^\dagger \Phi)^2$  and  $\mu^2, \lambda > 0$  [5, 6].

In the BEH mechanism,  $\Phi$  develops a vacuum expectation state ( $\langle \Phi \rangle_0$ ) that satisfies the condition:  $\langle \Phi^\dagger \Phi \rangle_0 = \frac{\mu^2}{2\lambda} = \frac{v^2}{2}$ , where  $v$  is known as the vacuum expectation value (vev). This leads to a ‘spontaneous’ breaking of the electroweak symmetry group  $SU(2)_L \otimes U(1)_Y$  (or EWSB) to the electromagnetic group  $U(1)_Q$  with a massless photon, and generating the masses to the charged- and neutral-current weak bosons  $W^\pm$  and  $Z$ , respectively. Furthermore, a real scalar

field named the Higgs field  $H$  is introduced with its massive particle, popularly known as the Higgs boson. Table 1-1 summarizes the BEH Lagrangian terms after EWSB. Finally, charged fermion fields acquire masses through Higgs-fermion Yukawa interactions seen Figure 1-2 E), where fermion masses are given by  $m_f = \frac{y_f v}{\sqrt{2}}$ , and  $y_f$  is the associated Yukawa coupling. Neutrino fields remain massless in the theory [5, 6].

One important prediction is the structure of the Higgs boson potential (See Table 1-1) with the associated cubic and quartic Higgs self-interactions via the self-couplings  $\lambda_{HHH}$  and  $\lambda_{HHHH}$ , respectively. These self-interactions are responsible for the Higgs boson mass  $m_H$ . The Higgs self-couplings are connected to the Higgs boson mass and the vev by the relation  $\lambda_{HHH} = \lambda_{HHHH} = \frac{m_H^2}{2v^2}$ . The experimental confirmation of this relation is crucial to our understanding of the BEH mechanism and a test of the internal consistency of the SM formulation. The Higgs boson interactions predicted by the SM are illustrated in Figure 1-2 [5, 6].

Table 1-1. BEH Lagrangian terms after the spontaneous eletroweak symmetry breaking (EWSB). The following variables are used where  $V = W, Z$ : Higgs boson mass ( $m_H$ ),  $W^\pm$  boson mass ( $m_W$ ),  $Z$  boson mass ( $m_Z$ ), coupling  $g_{HVV} = \frac{2m_V^2}{v}$ , coupling  $g_{VVHH} = \frac{m_V^2}{v^2}$ , and coupling  $\lambda = \lambda_{HHH} = \lambda_{HHHH} = \frac{m_H^2}{2v^2}$ .

| $\mathcal{L}_{\text{BEH}}$ terms after EWSB                                                         | Meaning                                    |
|-----------------------------------------------------------------------------------------------------|--------------------------------------------|
| $\frac{1}{2}(\partial_\mu H)(\partial^\mu H)$                                                       | Higgs kinematic term                       |
| $-\frac{1}{2}m_H^2 H^2 - v\lambda_{HHH}H^3 - \frac{1}{4}\lambda_{HHHH}H^4 + \frac{1}{4}\lambda v^4$ | Higgs potential term and self-interactions |
| $+m_W^2 HW_\mu^- W^{+, \mu} + g_{HWW}HW_\mu^- W^{+, \mu}$                                           | $W^\pm$ mass term and HWW interaction      |
| $+\frac{1}{2}m_Z^2 Z_\mu Z^\mu + \frac{1}{2}g_{HZZ}HZ_\mu Z^\mu$                                    | $Z$ mass term and HZZ interaction          |
| $+g_{WWHH}H^2 W_\mu^- W^{+, \mu}$                                                                   | WWHH quartic interaction                   |
| $+\frac{1}{2}g_{ZZHH}H^2 Z_\mu Z^\mu$                                                               | ZZHH quartic interaction                   |

### 1.2.4 The Higgs Boson

On July 4<sup>th</sup> 2012, the CMS and ATLAS Collaborations at the LHC announced the observation of a new resonance with mass around  $m_H = 125$  GeV and properties compatible with the SM Higgs boson [13, 14]. Since then, both collaborations have performed associated precision measurements using 7, 8 and 13 TeV pp collision data, which continue to establish it as the fundamental scalar predicted by the SM. In what follows is presented a brief summary of the Higgs boson production phenomenology, decay channels, and experimental status.

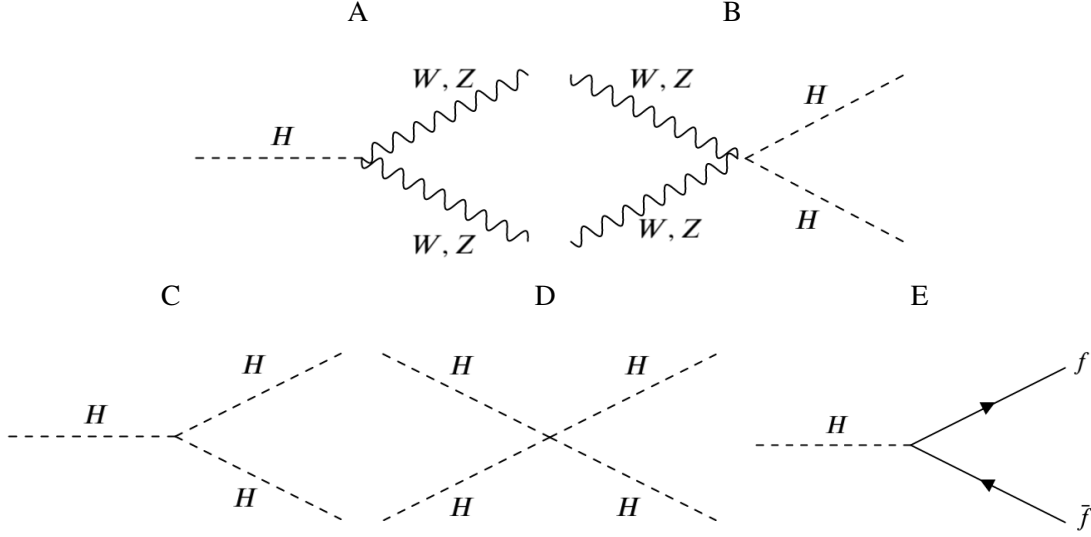


Figure 1-2. Predicted interactions of the Higgs boson (H) with fermions (f) and vector bosons (V): A) HVV interaction, B) HHVV interaction, C) HHH self-interaction, D) HHHH self-interaction, and E) Higgs-fermion Yukawa interaction.

#### 1.2.4.1 Production mechanisms and decay channels

The value of  $m_H$  was the only free parameter left to predict the Higgs boson phenomenology. Given that  $m_H$  is known, then the production mechanisms and decays are fully determined. The Higgs boson has several production mechanisms. Gluon fusion (ggF-H) is the main production mode, and then the vector-boson fusion (VBF-H) follows it, with a production cross section of approximately one order of magnitude smaller. Then, the subdominant production modes are the following: associated production with a W or a Z boson (VH) and associated production with a top quark pair ( $t\bar{t}H$ ), among others. Examples of representative LO diagrams are shown in Figure 1-3. In addition, Table 1-2 presents the theoretical production cross section for several modes.

Table 1-2. Higgs boson production cross section (in pb) and relative uncertainties (theory, PDF,  $\alpha_S$ ) for several Higgs boson ( $m_H = 125$  GeV) production modes at  $\sqrt{s} = 13$  and 14 TeV proton-proton collisions [6]. The order of QCD (EW) calculation is N3LO (NLO) for ggF and NNLO (NLO) for the others.

| $\sqrt{s}$ [TeV] | ggF-H                    | VBF-H                    | WH                       | ZH                       | $t\bar{t}H$              | total |
|------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|-------|
| 13               | $48.6^{+4.4\%}_{-7.0\%}$ | $3.78^{+2.2\%}_{-2.2\%}$ | $1.37^{+2.6\%}_{-2.6\%}$ | $0.88^{+4.1\%}_{-3.5\%}$ | $0.50^{+6.8\%}_{-9.9\%}$ | 55.1  |
| 14               | $54.7^{+4.4\%}_{-7.0\%}$ | $4.28^{+2.2\%}_{-2.2\%}$ | $1.51^{+1.9\%}_{-2.0\%}$ | $0.99^{+4.1\%}_{-3.7\%}$ | $0.60^{+6.9\%}_{-9.8\%}$ | 62.1  |

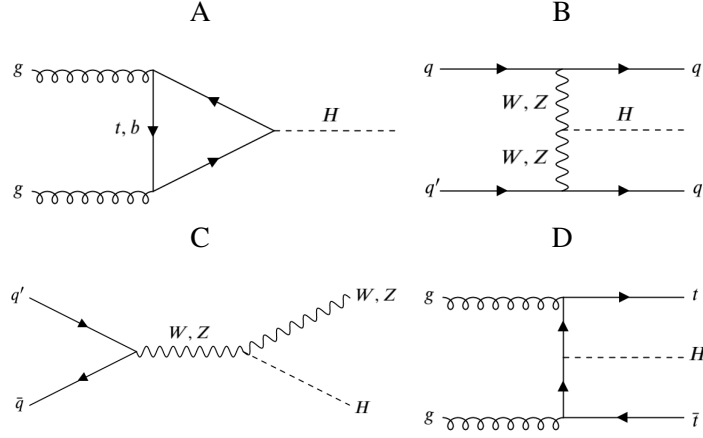


Figure 1-3. Examples of leading order diagrams contributing to the Higgs boson production: A) Gluon fusion, B) Vector boson fusion, (C) Associate production with a W or Z boson, (D) Associate production with a top-quark pair.

The theoretical Higgs boson decay branching ratios to gauge bosons and to some fermions are presented in Table 1-3. The Higgs boson decay to a bottom quark pair ( $H \rightarrow b\bar{b}$ ) is the dominant decay channel, with around 58.2% branching ratio. It is important to note that the Higgs boson can indirectly decay into massless particles via intermediate fermion loops: two photons ( $H \rightarrow \gamma\gamma$ ), a Z boson and a photon ( $H \rightarrow Z\gamma$ ), and two gluons ( $H \rightarrow gg$ ).

Table 1-3. Higgs boson branching ratios and relative uncertainties for several Higgs boson ( $m_H = 125$  GeV) decay modes [6].

| Decay mode                   | Branching ratio       | Relative uncertainty |
|------------------------------|-----------------------|----------------------|
| $H \rightarrow \gamma\gamma$ | $2.22 \times 10^{-3}$ | $\pm 2.1\%$          |
| $H \rightarrow ZZ$           | $2.62 \times 10^{-2}$ | $\pm 1.5\%$          |
| $H \rightarrow W^+W^-$       | $2.14 \times 10^{-1}$ | $\pm 1.5\%$          |
| $H \rightarrow \tau^+\tau^-$ | $6.27 \times 10^{-2}$ | $\pm 1.6\%$          |
| $H \rightarrow b\bar{b}$     | $5.82 \times 10^{-1}$ | $+1.2\%$<br>$-1.3\%$ |
| $H \rightarrow c\bar{c}$     | $2.89 \times 10^{-2}$ | $+5.5\%$<br>$-2.0\%$ |
| $H \rightarrow Z\gamma$      | $1.53 \times 10^{-3}$ | $\pm 5.8\%$          |
| $H \rightarrow \mu^+\mu^-$   | $2.18 \times 10^{-4}$ | $\pm 1.7\%$          |

#### 1.2.4.2 Experimental Status

The LHC exploration of the Higgs boson is carried out analyzing a broad range of production mechanisms and decay channels. The most sensitive decay channels for studying the Higgs boson are as follows. The  $H \rightarrow ZZ \rightarrow 4\ell$  ( $\ell = e, \mu$ ) and  $H \rightarrow \gamma\gamma$  channels have low

expected rates but benefit from small background processes and have the best mass resolution ( $\sim 1\text{-}2\%$ ) due to well-reconstructed final state products. The  $H \rightarrow W^+W^- \rightarrow \ell^+ \nu_{\ell} \ell'^- \bar{\nu}_{\ell'}$  channel has relatively better rates, but the unmeasured energy from neutrinos has an impact on the resolution of the mass-related observables, e.g. the transverse mass resolution is  $\sim 20\%$ . The  $b\bar{b}$  ( $\tau^+\tau^-$ ) decay channel benefits from high rates and intermediate mass resolutions  $10\%$  ( $15\%$ ), but is affected by large and irreducible background contamination [6]. Challenging channels due to low very rates or large background levels are  $H \rightarrow \mu^+\mu^-$  and  $H \rightarrow c\bar{c}$ , respectively.

In the LHC Run-1 (2009-2012), the ATLAS and CMS collaborations analyzed the pp collision datasets at  $\sqrt{s} = 7$  and 8 TeV to discover the Higgs boson [13, 14]. The decay channels involving gauge bosons,  $H \rightarrow ZZ \rightarrow 4\ell$  ( $\ell = e, \mu$ ) and  $H \rightarrow \gamma\gamma$ , contributed with the most significant excesses. The  $H \rightarrow W^+W^- \rightarrow \ell^+ \nu_{\ell} \ell'^- \bar{\nu}_{\ell'}$ ,  $H \rightarrow b\bar{b}$  and  $H \rightarrow \tau^+\tau^-$  channels were also included but contributed with lower significance [13, 14]. After that, studies of the spin-parity ( $J^P$ ) quantum numbers revealed the scalar and even-parity nature ( $J^P = 0^+$ ) of the discovered particle as predicted by the SM [15, 16]. The Run-1 Combined ATLAS and CMS Higgs boson mass measurement yields a value of  $m_H = 125.09 \pm 0.24$  GeV using  $H \rightarrow ZZ \rightarrow 4\ell$  ( $\ell = e, \mu$ ) and  $H \rightarrow \gamma\gamma$ , benefiting from their excellent mass resolution [17].

In the LHC Run-2 (2015-2018), ATLAS and CMS experiments had an excellent operation performance collecting high-quality data at  $\sqrt{s} = 13$  TeV, corresponding to approximately 6 times the LHC Run-1 dataset integrated luminosity. Measurements of the mass, couplings, differential and fiducial cross sections have been performed with the partial and full dataset by ATLAS [18, 19] and CMS [20, 21]. The Higgs boson mass measurement has been performed using the  $H \rightarrow ZZ \rightarrow 4\ell$  ( $\ell = e, \mu$ ) and  $H \rightarrow \gamma\gamma$  decay channels. The distribution of the di-photon invariant mass in the 2016 data is presented in Figure 1-4 A). The best mass measurement yields a value of  $m_H = 125.38 \pm 0.14$  GeV from the CMS combination of Run-1 and 2016 datasets [21].

Both experiments have presented the observation of the Higgs coupling to third generation fermions studying the  $H \rightarrow b\bar{b}$  [22, 23] and  $H \rightarrow \tau^+\tau^-$  [24, 25] decays, and the  $t\bar{t}H$  production mode [26, 27]. In particular, the measurement of  $H \rightarrow b\bar{b}$  decays was believed to be impossible

to achieve due to the large multi-jet background and intermediate mass resolution. However, novel analysis methods based on machine learning (ML) were developed for object identification, reconstruction and signal identification, to maximize the search sensitivity. Figure 1-4 B) illustrates the distribution signal-to-background ratio bins in the  $H \rightarrow b\bar{b}$  measurement. The CMS experiment found the first evidence for the Higgs coupling to the second generation fermions in the analysis of the  $H \rightarrow \mu^+\mu^-$  decay channel [28]. ATLAS show the first evidence for the  $H \rightarrow (\gamma/Z)^*\gamma \rightarrow \ell\ell\gamma$  decay channel [29]. Lastly, both experiments are currently developing methods to measure the  $H \rightarrow c\bar{c}$  decay but the development of more sophisticated methods and more data are needed to measure it in the future [30, 31]. The Figure 1-4 C) shows the Higgs coupling measurements using CMS Run-2 data.

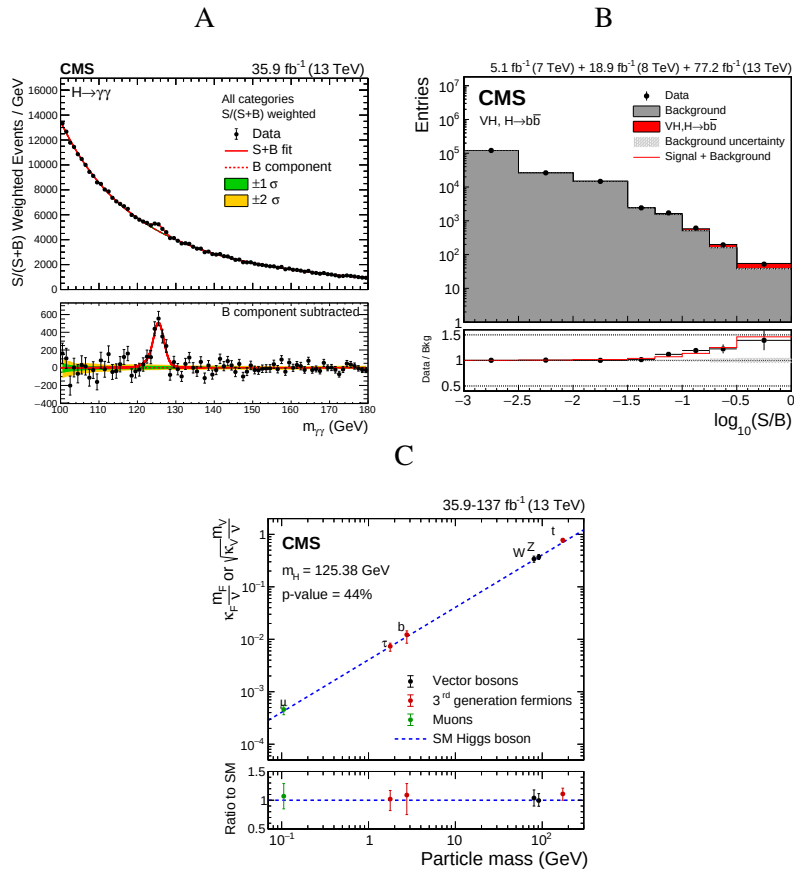


Figure 1-4. Selected CMS Higgs physics results in Run-2. A) Di-photon mass distribution from the 2016 dataset mass measurement [21], B) Distribution of event yields sorted into bins of similar signal-to-background ratio in the observation of Higgs coupling to two bottom-quarks [23], C) Measurement of the Higgs couplings [28].



### 1.3 Higgs Boson Pair Production

This section is a review of the theoretical aspects and phenomenology associated to the Higgs boson pair (HH) production. First, the physics of SM HH production is discussed. Then, the motivation and characteristics of the BSM HH production are summarized.

#### 1.3.1 Standard Model Production

From the precise measurement of the Higgs boson mass, the value of the Higgs self-couplings are fully predicted. The Higgs trilinear coupling or self-coupling  $\lambda_{HHH} = \frac{m_H^2}{2v^2} \sim 0.13$ , where,  $v$  is around 246 GeV from measurements of the Fermi constant  $G_F$ . Non-resonant HH production is the only SM process with direct access to measure  $\lambda_{HHH}$  at the LHC. The main production mechanisms are the following, ordered by their cross section (see Table 1-4): Gluon fusion production (ggF), vector boson fusion production (VBF), associated production with a Z or W boson (VHH) and top quark pairs associated production ( $t\bar{t}HH$ ). Some representative LO Feynman diagrams for each mode are illustrated in Figure 1-6. Note that single Higgs production has indirect access to the  $\lambda_{HHH}$  coupling via NLO EW corrections [32, 33].

Table 1-4. Higgs boson pair production cross section (in fb) and relative uncertainties for several Higgs boson ( $m_H = 125$  GeV) production modes at  $\sqrt{s} = 13$  and 14 TeV proton-proton collisions [34].

| $\sqrt{s}$   | 13 TeV                                | 14 TeV                                |
|--------------|---------------------------------------|---------------------------------------|
| ggF          | $31.05^{+6.0\%}_{-23.0\%} \pm 3.0\%$  | $36.69^{+6.0\%}_{-23.0\%} \pm 3.0\%$  |
| VBF          | $1.73^{+0.03\%}_{-0.04\%} \pm 2.1\%$  | $2.05^{+0.03\%}_{-0.04\%} \pm 2.1\%$  |
| ZHH          | $0.363^{+3.4\%}_{-2.7\%} \pm 1.9\%$   | $0.415^{+3.5\%}_{-2.7\%} \pm 1.8\%$   |
| $W^+HH$      | $0.329^{+0.32\%}_{-0.41\%} \pm 2.2\%$ | $0.369^{+0.33\%}_{-0.39\%} \pm 2.1\%$ |
| $W^-HH$      | $0.173^{+1.2\%}_{-1.3\%} \pm 2.8\%$   | $0.198^{+1.2\%}_{-1.3\%} \pm 2.7\%$   |
| $t\bar{t}HH$ | $0.773^{+1.5\%}_{-4.3\%} \pm 3.2\%$   | $0.949^{+1.7\%}_{-4.5\%} \pm 3.1\%$   |

In the dominant mode, gluon fusion, Higgs boson pairs proceed from quantum loops involving mainly top-quarks (and a small b-quark contribution). At LO, a triangle-loop diagram and a box-loop diagram contribute to the ggF mode physics (see Figure 1-6 A)). The triangle-loop diagram contains the Higgs self-coupling  $\lambda_{HHH}$ , while the box diagram contains the top quark Yukawa coupling ( $y_t$ ). There is a large negative interference between the two loop diagrams, resulting in a very small production cross section. The individual contributions to the cross

section from the box, triangle and their interference are illustrated as function HH invariant mass in Figure 1-5. The triangle or self-interaction diagram contribution is mostly located at low HH mass values (i.e. soft spectrum) and is largely suppressed by the interference contribution.

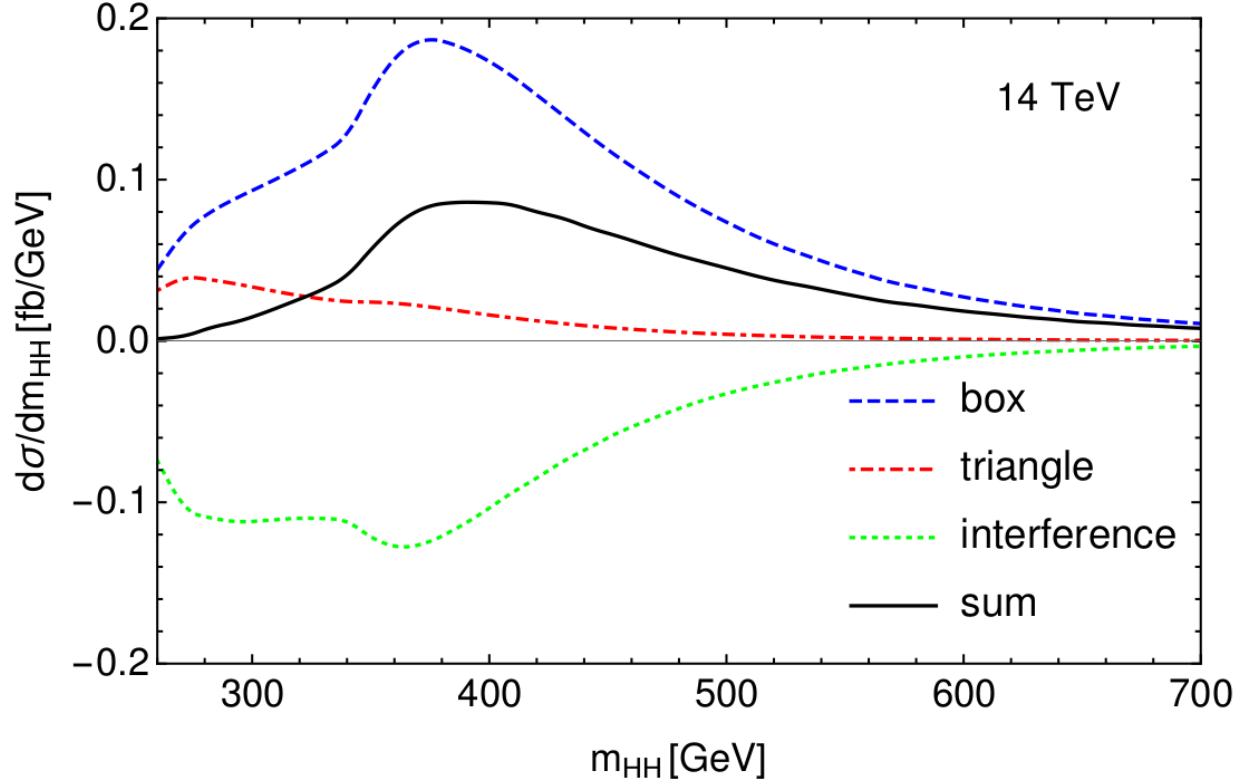


Figure 1-5. Differential cross section distribution as a function of the Higgs boson invariant mass at leading order [34]. The contributions to the gluon fusion production cross section (black line) are presented: box (dashed blue), triangle (dashed red), interference (dashed green).

The secondary production mode, vector boson fusion (VBF), has three LO contributing diagrams shown in Figure 1-6 B). The VBF production cross section is approximately one order of magnitude smaller than the ggF one but has a very characteristic signature with two quark hadronic jets with large invariant mass and separation in the final state. Besides having an extra handle on the  $\lambda_{HHH}$  coupling, this process has a unique access to study the interaction between two vector bosons and two Higgs bosons through the  $g_{HHVV}$  coupling. Moreover, this mode has access to the interaction between the Higgs boson and two vector bosons via the  $g_{HVV}$  coupling.

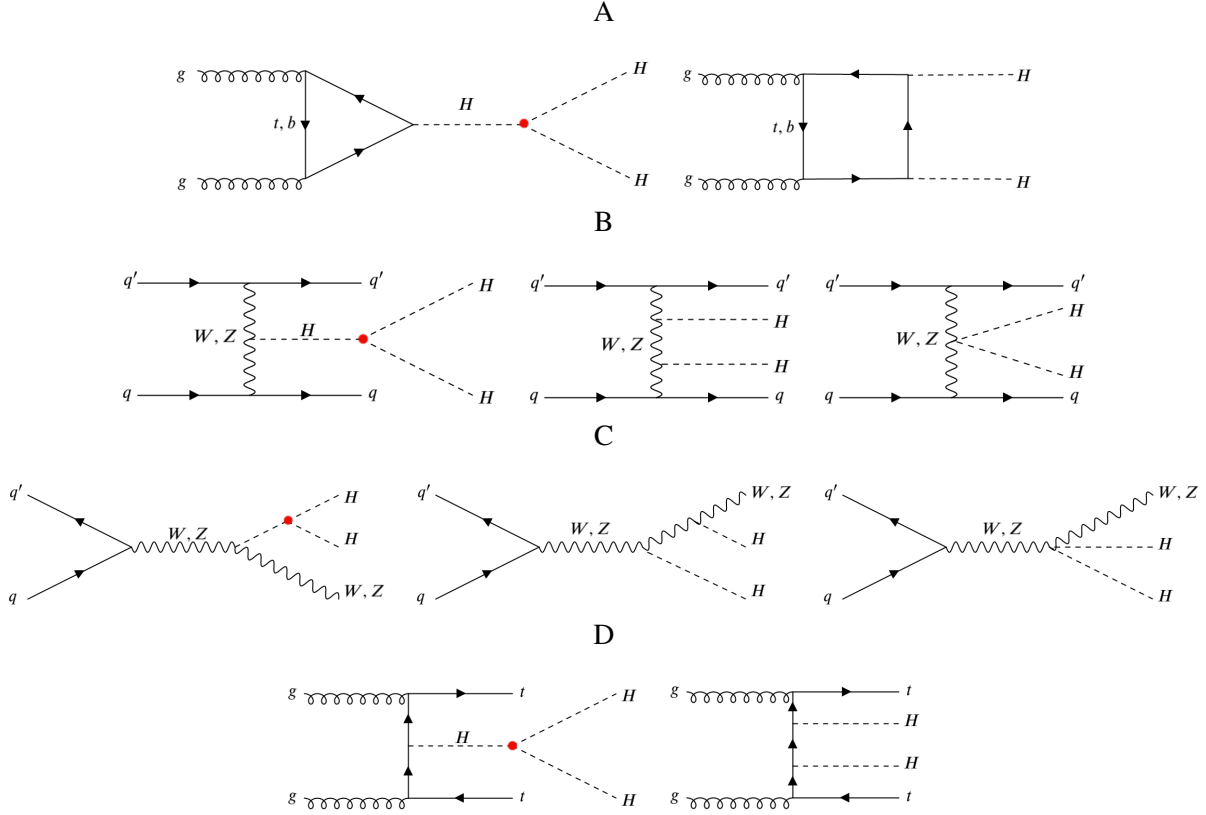


Figure 1-6. Examples of leading order (LO) diagrams contributing to the Higgs boson pair production mode mechanisms. A) Gluon fusion, B) Vector boson fusion C) Associate production with a W or Z boson, D) Associate production with top-quark pair. The red circle represents the Higgs self-coupling  $\lambda_{HHH}$ .

### 1.3.2 Beyond the Standard Model Production

Although the SM have been successfully in predicting the measurements at particle colliders, it does not provide explanations to many experimental and theoretical aspects about nature. Many of them are connected to the scalar sector, which hints that BSM physics is hidden there. From the theoretical point of view, there is no fundamental symmetry to protect the Higgs boson mass from quadratically divergent radiative corrections [6]. Furthermore, a metastable scalar potential appears to be favored based on the measured Higgs boson and top-quark masses, challenging the long-term stability of the Universe [35].

Many BSM physics scenarios are able to regulate the shortcomings of the SM. Some of these models have extra singlets [36, 37], extra doublets inspired by the next-to-minimal

supersymmetric extension of the SM(NMSSM) [38, 39], or warped extra-dimensions with a new scalar or a graviton [40, 41]. A broad class of these models predict the production of new particles ( $X$ ) that can subsequently decay in a sizeable branching fraction to a pair of Higgs boson ( $X \rightarrow HH$ ) as illustrated in the diagrams in Figure 1-7. At the LHC, one can probe these BSM models using a common experimental signature consisting of an enhancement of the cross section at the  $X$  mass ( $m_X$ ). The studied masses range from 250 GeV ( $m_X > 2 m_H$ ) to a few TeVs. Recent theoretical work has also highlighted the possibility of spectacular multi-scalar resonant decays into a new scalar ( $Y$ ) and a Higgs boson ( $X \rightarrow YH$ ), and a Higgs boson triplet ( $X \rightarrow YH \rightarrow HHH$ ) [42, 43].

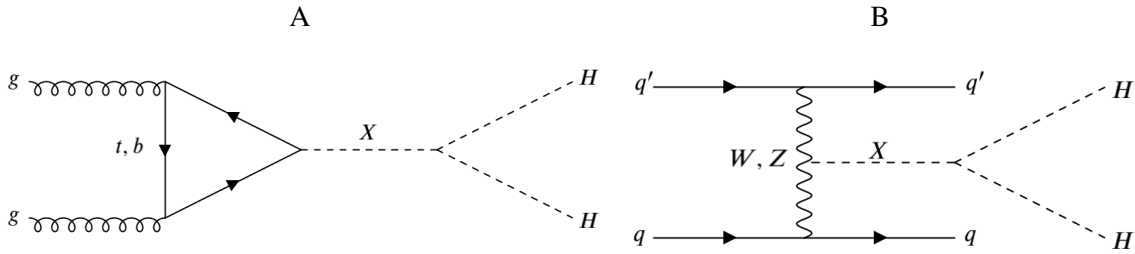


Figure 1-7. Diagrams of production of a new resonances decaying to a Higgs boson pair at the LHC: A) Gluon fusion production mode, and B) Vector boson production mode.

If the scale of the new physics is beyond the direct reach of the LHC, the effects of the new particles and their interactions within quantum loops may result in modifications of the  $HH$  production cross section and kinematics. These modifications can be parametrized with respect to the SM value through coupling modifiers( $\kappa$ ). For instance, anomalous Higgs self-coupling,  $g_{HVV}$  coupling and  $g_{VVHH}$  coupling have the modifiers  $\kappa_\lambda = \lambda_{HHH}/\lambda_{HHH}^{SM}$ ,  $\kappa_V$ , and  $\kappa_{2V}$ , respectively. Figure 1-8 presents the dependence of the cross section for different production modes as a function of the  $\kappa_\lambda$  modifier (assuming other couplings are set to their SM values). The ggF mode cross section has a quadratic  $\kappa_\lambda$ -dependence, with a minimum at around 2.45 (maximal interference effect).

BSM  $\kappa_\lambda$  values ( $\kappa_\lambda \neq 1$ ) modify the kinematic properties of the Higgs boson pair. This effect is illustrated on Figure 1-9 A), where the  $HH$  invariant mass spectrum for different values

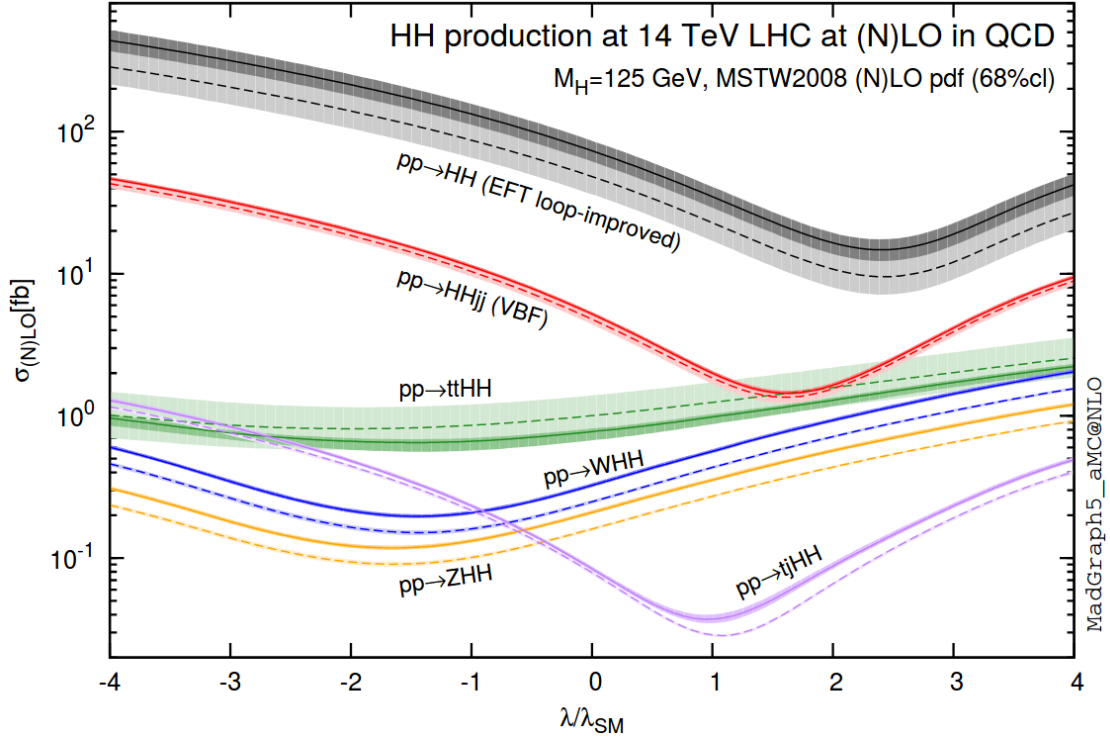


Figure 1-8. Higgs boson pair production cross section at  $\sqrt{s} = 14$  TeV LHC proton-proton collisions for various modes calculated at the LO and NLO precision in QCD, as a function of the Higgs self-coupling with respect to the SM ( $\lambda/\lambda_{\text{SM}}$ ) [44]. The dashed (solid) lines and light- (dark-)color bands correspond to the LO (NLO) results and to the scale and PDF uncertainties added linearly.

of the  $\kappa_\lambda$  modifier in the ggF mode is shown. The softest spectrum (experimentally challenging due to event selection thresholds), corresponds to  $\kappa_\lambda = 5$ . Note that the case where mostly the triangle diagram contributes to the HH production cross section ( $\kappa_\lambda = 20$ ) has a soft spectrum, whereas the case of box diagram ( $\kappa_\lambda = 20$ ) has an intermediate mass spectrum. For the  $\kappa_\lambda = 2.45$  case, a spectrum with a low mass peak and intermediate mass peak is seen.

In the VBF mode, the amplitude of the longitudinal scattering of vector bosons contributing to the production cross section depends on the center-of-mass energy ( $\hat{s}$ ) and  $\kappa_V$ , and  $\kappa_{2V}$  as  $\mathcal{A}(V_L V_L \rightarrow HH) \simeq \frac{\hat{s}}{v^2} (\kappa_{2V} - \kappa_V^2)$ . Note that this amplitude is suppressed in the SM ( $\kappa_{2V} = \kappa_V = 1$ ), but is present for cases when  $\kappa_{2V} \neq \kappa_V^2$ . Consequently, the effect of BSM physics may manifest as an enhancement at high HH masses, illustrated in Figure 1-9 B) for anomalous  $\kappa_{2V}$  values 0 (no VVHH interaction) and 2 (enhanced VVHH interaction).

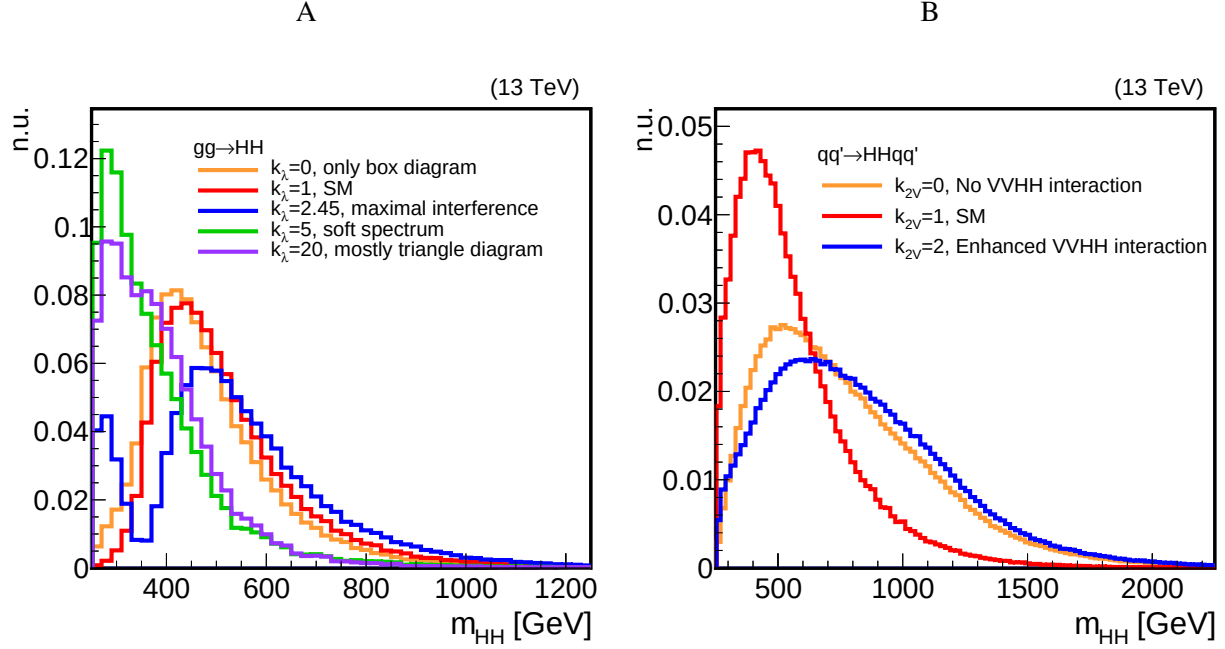


Figure 1-9. Normalized Higgs boson pair invariant mass for different values of the coupling modifiers: A)  $\kappa_\lambda$  variations in the ggF mode, and B)  $\kappa_{2V}$  in the VBF mode. Other Higgs couplings are set to their SM values.

A generalization of the BSM effects on the ggF HH production is achieved in the context of an effective field theory (EFT), where the SM Lagrangian (dimension  $\leq 4$ ) is extended with new terms associated to dimension-6 operators [45]. The effects of these new operators represent modifications of the SM  $y_t$  and  $\lambda_{HHH}$  couplings with strength  $\kappa_\lambda$  and  $\kappa_t = y_t/y_t^{\text{SM}}$ , and three new vertices associated to contact interactions gHH, ggHH and ttHH with coupling strengths denoted as  $c_g$ ,  $c_{2g}$  and  $c_2$ , respectively. Figure 1-10 illustrates the LO diagrams of the three EFT contact interactions. The EFT NLO accuracy model is presented in Ref. [46].

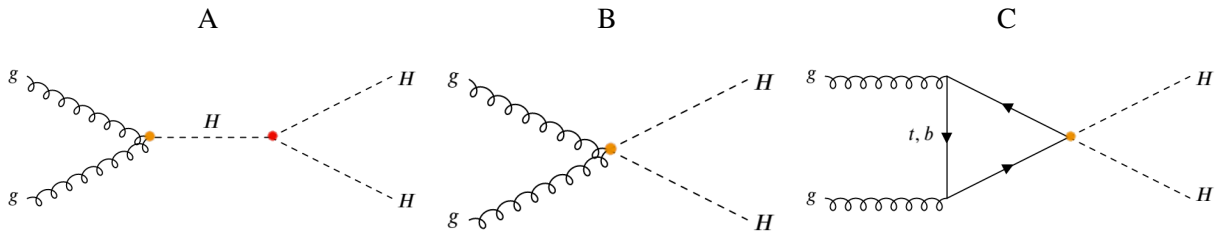


Figure 1-10. Leading order diagrams of contact interactions: A) gHH, B) ggHH and C) ttHH. The orange (red) circle represents the vertices absent in the SM (Higgs self-coupling).

Given that it is not feasible to experimentally study all possible combinations of the five EFT couplings, 12 shape benchmarks are chosen to target characteristic kinematic properties of large regions of the five-dimensional phase space by the clustering method presented in Ref. [47]. The coupling combination of the benchmarks are listed in Table 1-5 and their  $m_{HH}$  shapes are illustrated in Figure 1-11. Note that the softest spectrum corresponds to benchmark 7, whereas the hardest spectrum corresponds to benchmark 2.

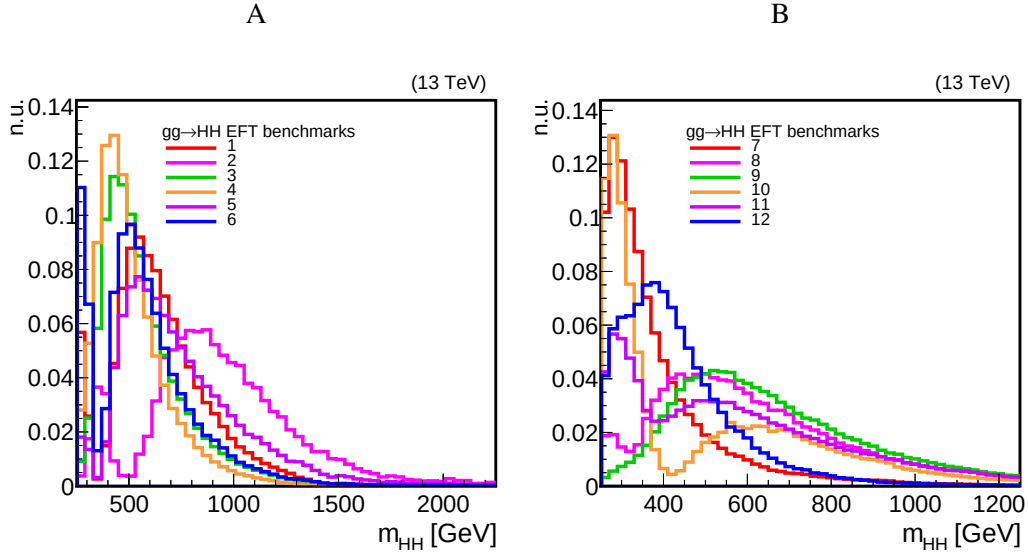


Figure 1-11. Higgs boson pair invariant mass distribution for different leading order effective field theory (EFT) shape benchmarks: A) Nodes 1 to 6, and B) Nodes 7 to 12.

Table 1-5. Effective field theory benchmarks for LHC searches for non-resonant Higgs boson pair production via the gluon-fusion mode [47].

| Benchmark | $\kappa_\lambda$ | $\kappa_t$ | $c_2$ | $c_g$ | $c_g$ |
|-----------|------------------|------------|-------|-------|-------|
| 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   |

## 1.4 Searches for Higgs Boson Pair Production

This section reviews the phenomenology and experimental aspects of the ongoing HH program at the LHC. First, the HH decay channels and their properties are discussed. Then, the results of the searches performed using LHC pp collision data are summarized. It ends with the most recent projections for the SM HH production at the High Luminosity LHC (HL-LHC), which were prepared in the context of the European Strategy for Particle Physics before the development of this dissertation [48].

### 1.4.1 Decay Channels

SM HH production is an elusive process to measure due to the low expected rates at the LHC. However, as mentioned before, rates and kinematics can be modified by BSM effects and potentially discovered with the current datasets. The HH signature has a rich variety of decay channels with different branching fractions depending on the two Higgs boson decays as seen in Figure 1-12. The  $b\bar{b}b\bar{b}$  is the dominant decay channel with  $\sim 33\%$  branching fraction. Decay channels with at least one  $H \rightarrow b\bar{b}$  are used in searches to maintain a sizeable branching fraction.

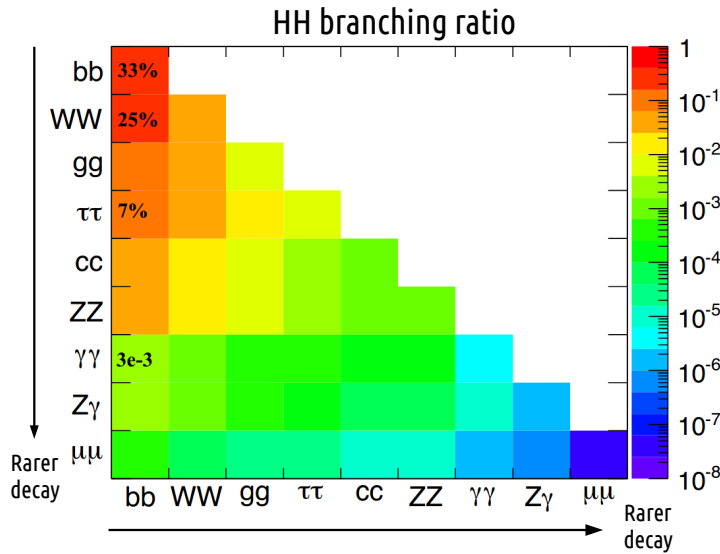


Figure 1-12. Higgs boson pair branching fractions assuming a Higgs mass  $m_H = 125$  GeV. The Higgs decays considered in the figure are: two bottom quarks ( $b\bar{b}$ ), two W bosons ( $W^+W^-$ ), two gluons (gg), two tau leptons ( $\tau^+\tau^-$ ), two charm quarks (cc), two Z bosons (ZZ), two photons ( $\gamma\gamma$ ), a Z boson and a photon ( $Z\gamma$ ), and two muons ( $\mu\mu$ ).



### 1.4.2 Experimental Status at the LHC

The LHC has a broad experimental program exploring resonant and non-resonant HH production in multiple channels and production modes. The first results were obtained analyzing the Run-1 collision data at 8 TeV, corresponding to around  $20\text{fb}^{-1}$  of integrated luminosity. ATLAS (CMS) studied the gluon fusion HH production in the  $b\bar{b}\gamma\gamma$ ,  $b\bar{b}\tau^+\tau^-$ ,  $b\bar{b}b\bar{b}$  and  $\gamma\gamma W^+W^-$  ( $b\bar{b}\gamma\gamma$ ,  $b\bar{b}\tau^+\tau^-$  and  $b\bar{b}b\bar{b}$ ). The Run-1 ATLAS (CMS) observed combined upper limits on the SM-like HH production cross section corresponds to 70 (43) times the SM prediction, respectively. Other results in the context of BSM  $\kappa_\lambda$  couplings and resonances were presented, but no significant excess of signal was observed [49, 50].

At the present time, CMS and ATLAS experts are exploring searches with more production modes and new decay channels by analyzing the full Run-2 dataset collected at 13 TeV. While those results are being finished, both experiments already released a very comprehensive study of direct searches via the ggF mode by combining the results obtained in several channels using partial Run-2 dataset [51, 52]. The ATLAS (CMS) dataset was collected during the 2015-2016 (2016) operations with an integrated luminosity of 27.1–36.1 (35.9)  $\text{fb}^{-1}$ . The ATLAS (CMS) channels used in the combination are the following:  $b\bar{b}\gamma\gamma$  [53],  $b\bar{b}\tau^+\tau^-$  [54],  $b\bar{b}b\bar{b}$  [55],  $W^+W^-W^+W^-$  [56] and  $\gamma\gamma W^+W^-$  [57] ( $b\bar{b}b\bar{b}$  boosted [58] and resolved [59, 60],  $b\bar{b}\gamma\gamma$  [61],  $b\bar{b}\tau^+\tau^-$  [62, 63],  $b\bar{b}VV$  [64]). In addition, both experiments studied indirect  $\kappa_\lambda$  constraints from single Higgs searches [20, 65].

The ATLAS (CMS) combination set an observed limit on the SM production cross section at 6.9 (22) times the SM prediction. The observed allowed  $\kappa_\lambda$  interval at 95% CL ranges from -5.0 (-11.8) to 12.0 (18.8). Results on EFT benchmarks were also presented. Furthermore, both experiments presented results on resonant searches for new particles with Spin-0 and Spin-2 in the mass range between 250 GeV and 3 TeV. However, no statistically significant excess of resonant events were observed, and 95% CL limits on the production cross section of new resonances decaying into Higgs pairs were provided. In addition, the ATLAS (CMS) results from single Higgs searches presented the 95% CL allowed  $\kappa_\lambda$  interval from -3.2 (-3.5) to 11.9 (14.5).

Excluding the results of this dissertation, several Run-2 HH non-resonant analyses have been performed. The ATLAS experiment has presented the study of the ggF production mode in the  $b\bar{b}\gamma\gamma$  [66],  $b\bar{b}\tau^+\tau^-$  [67], and  $b\bar{b}W^+W^-(\ell\nu\ell\nu)$  [68] channels, whereas the CMS experiment has explored the same mode in the  $b\bar{b}\gamma\gamma$  [69] and  $b\bar{b}ZZ(4\ell)$  [70] channels. Moreover, the VBF production mode has been explored by the ATLAS experiment in the  $b\bar{b}b\bar{b}$  [71], whereas CMS experiment has done it in the  $b\bar{b}\gamma\gamma$  [69] channel, and the  $b\bar{b}b\bar{b}$  channel studying highly Lorentz-boosted Higgs bosons [72]. Once the main decay channels are analyzed, then a full Run-2 HH combination will be performed. In addition, full Run-2 resonant searches have also been performed in ATLAS [66, 67, 71] and CMS [73, 74].

The best LHC results on non-resonant HH production are obtained from the currently finalized full Run-2 analyses. From the ATLAS combination of the results of the  $b\bar{b}\gamma\gamma$  and  $b\bar{b}\tau^+\tau^-$  channels, the current observed limit on the SM-like HH production cross section is 3.1 times the SM expectation, and the observed allowed  $\kappa_\lambda$  interval at 95% CL is from -1.0 to 6.6 as presented in Figure 1-13 A) [75]. Furthermore, the current observed limit on VBF production cross section is 225 times the SM prediction by the CMS search in the  $b\bar{b}\gamma\gamma$  channel. Finally, the observed allowed interval of the  $\kappa_{2V}$  modifier is from 0.6 to 1.4, corresponding to the search in the  $b\bar{b}b\bar{b}$  channel with highly Lorentz-boosted Higgs bosons (illustrated in Figure 1-13 B)) [72].

The search sensitivity of the HH searches depends on the rates (production cross section and branching ratio), selection efficiencies, level of background contamination and mass resolution. The HH signal is reconstructed by the ATLAS and detectors from objects associated to the final state particles. These objects are b-quark hadronic jets (b) or ‘b-jets’, light-quark jets (j), hadronic or leptonic  $\tau$  leptons, muons and electrons ( $\ell$ ), photons ( $\gamma$ ), and missing transverse energy (MET) if neutrinos ( $\nu$ ) are present. As the quality of the object reconstruction has a direct impact on the HH signal identification, ATLAS and CMS have developed algorithms to improve it. For instance, there are dedicated jet identification algorithms based on machine learning to improve the reconstruction of the  $H \rightarrow b\bar{b}$  decay from two small-radius b-jets or a large-radius jet (if H has a large Lorentz boost) in CMS [76, 77] and ATLAS [78, 79]. Moreover, there are multivariate

regression algorithms to correct the reconstructed b-jet energy from the unmeasured energy associated neutrinos inside the b-jet from semileptonic B hadron decays (e.g. see Ref.[80]).

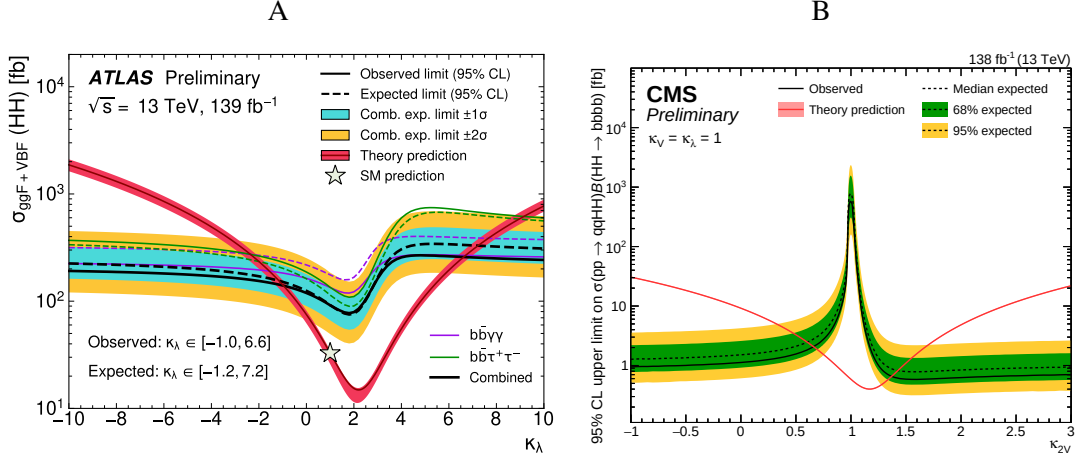


Figure 1-13. 95 % CL observed and expected upper limits on the Higgs boson pair production cross section as function of the coupling modifiers. A) Gluon fusion and vector boson fusion production cross section as a function of  $\kappa_\lambda$  [75], B) Vector boson fusion cross section as a function of  $\kappa_{2V}$  [72].

There are three crucial HH decay channels based on the experience gained in the Run-1 and 2016 analyses:  $b\bar{b}\gamma\gamma$ ,  $b\bar{b}\tau^+\tau^-$  and  $b\bar{b}b\bar{b}$ . Their properties and experimental aspects are summarized in the following points:

- $b\bar{b}\gamma\gamma$ : This channel has a small branching ratio ( $\sim 0.3\%$ ) but a resonant signature over smoothly falling small background from multi-jet + photons ( $N\gamma + \text{jets}$ ,  $N = 1, 2$ ) processes. There is also a background contribution from single Higgs production ( $H \rightarrow \gamma\gamma + \text{jets}$ ) production. It is very sensitive for searches at low  $m_{\text{HH}}$  mass. The observables used for signal identification are based on the reconstructed masses of the two Higgs bosons taking advantage of the excellent di-photon mass resolution, as illustrated in Figure 1-14 A).
- $b\bar{b}\tau^+\tau^-$ : This decay channel has an intermediate branching ratio but challenging background from top-quark pair, Drell-Yan ( $Z/\gamma + \text{jets}$ ) and multi-jet production. Searches are performed using the  $H \rightarrow \tau^+\tau^-$  final states with two hadronic decays ( $\tau_{\text{had}}\tau_{\text{had}}$ ), one hadronic and one leptonic decays ( $\tau_{\text{had}}\tau_{\ell=e,\mu}$ ). The signal identification is performed using observables based on Higgs pair mass and machine learning algorithms, as seen in Figure 1-14 B).
- $b\bar{b}b\bar{b}$ : This channel is the focus of this dissertation. It benefits from the largest branching fraction. However, there is an overwhelming amount of multi-jet background processes, which make it very hard to measure it. Consequently, these searches relies on the good

performance of the jet flavor tagging algorithms and a precise measurement of the background from data. The signal identification discriminants are based on machine learning and the Higgs pair mass, e.g. in Figure 1-14 C). This is the most sensitive channel for searches at high  $m_{HH}$  mass.

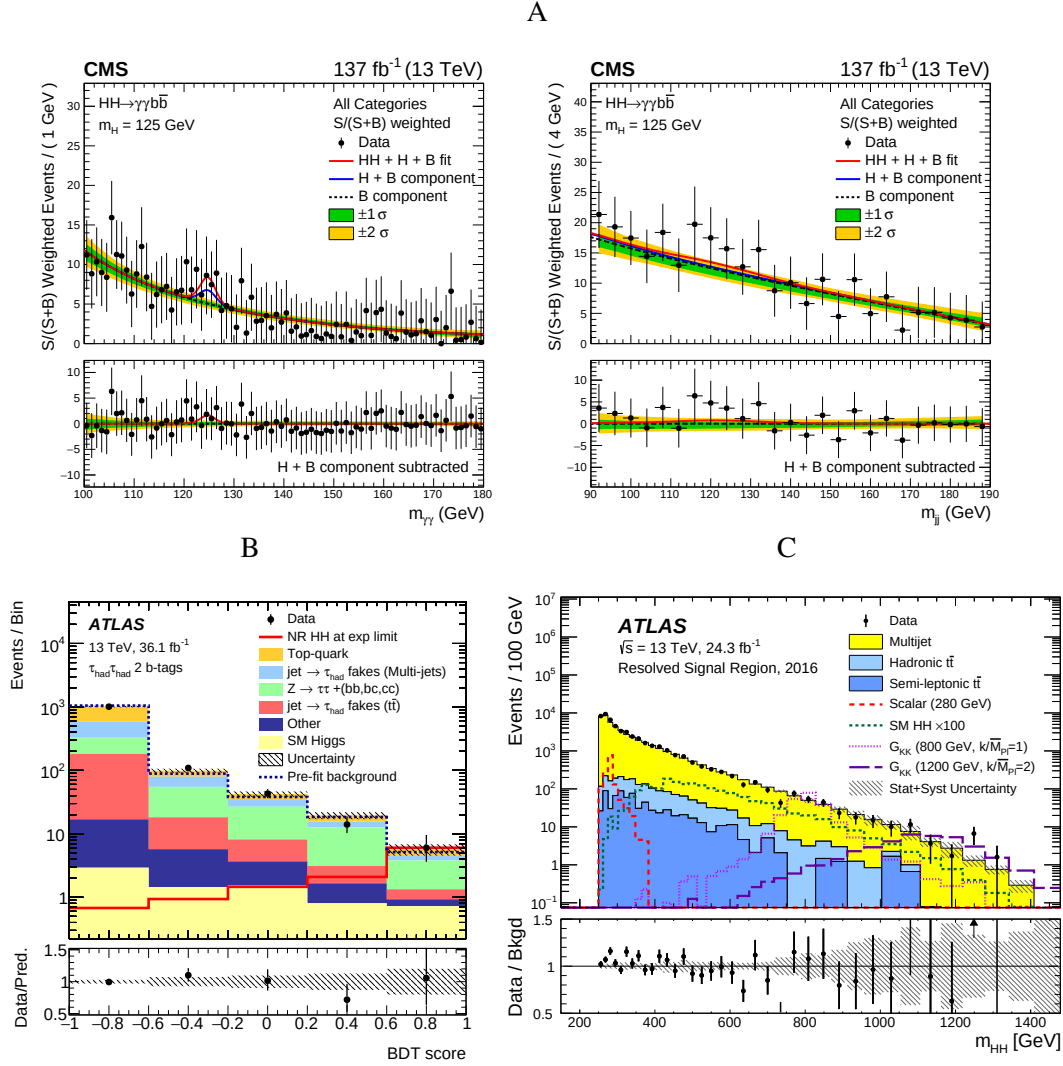


Figure 1-14. Signal observables used in searches for HH production. A) di-photon and bottom-quark pair distributions used in the last search for  $HH \rightarrow b\bar{b}\gamma\gamma$  by CMS [69], B) Boosted decision trees (BDT) distribution used in the last search for  $HH \rightarrow b\bar{b}\tau^+\tau^-$  by ATLAS [54], C) Higgs boson pair mass distribution used in the last search for  $HH \rightarrow b\bar{b}b\bar{b}$  by ATLAS [55].

### 1.4.3 Future Prospects at the High Luminosity LHC

The LHC is currently under its second long shutdown (LS2) and preparing for the Run-3 data-taking program (2022-2025). Next, CMS and ATLAS will upgrade their detectors, trigger

and data acquisition systems during a third long shutdown (LS3) to prepare themselves for the ultimate configuration of the LHC, the so-called High Luminosity LHC (HL-LHC). Each experiment is expected to collect a dataset of pp collisions at 14 TeV of center-of-mass energy within 10 years of the HL-LHC operation, corresponding to 3000–4000 fb<sup>-1</sup> of integrated luminosity ( $\sim 20$  times more than in Run-2). More details are discussed in Chapter 2.

The analysis of these large dataset will be crucial for a potential measurement of the SM HH production and the Higgs boson self-coupling  $\lambda_{HHH}$ . In consequence, ATLAS and CMS experiments have already started to study what the challenges and developments needed are to reach that goal in the future. Both have presented the most recent projections of the searches for this elusive process at the HL-LHC based on the current analysis strategies on the gluon fusion mode and assuming an expected integrated luminosity of 3000 fb<sup>-1</sup> [81, 82]. In these studies, the channels considered by ATLAS (CMS) are  $b\bar{b}\gamma\gamma$ ,  $b\bar{b}\tau^+\tau^-$  and  $b\bar{b}b\bar{b}$  ( $b\bar{b}\gamma\gamma$ ,  $b\bar{b}\tau^+\tau^-$ ,  $b\bar{b}b\bar{b}$ ,  $b\bar{b}VV(\ell\ell\nu\nu)$  and  $b\bar{b}ZZ(4\ell)$ ). In particular, the  $b\bar{b}b\bar{b}$  channel studies show that its main challenges will be our capability to have online triggers with high signal selection efficiency and model the multi-jet background events with high precision.

Table 1-6. ATLAS and CMS expected significance (in standard deviation) for SM Higgs pair production in different channels at the HL-LHC [83].

| Channel                      | ATLAS | CMS  |
|------------------------------|-------|------|
| $b\bar{b}b\bar{b}$           | 0.61  | 0.95 |
| $b\bar{b}\tau^+\tau^-$       | 2.1   | 1.4  |
| $b\bar{b}\gamma\gamma$       | 2.0   | 1.8  |
| $b\bar{b}VV(\ell\ell\nu\nu)$ | –     | 0.56 |
| $b\bar{b}ZZ(4\ell)$          | –     | 0.37 |
| Combination                  | 3.0   | 2.6  |

According to the projections, the ATLAS (CMS) expected significance yields a value of 3.0 (2.6) standard deviations from the combination of its individual channels. The expected significance per channel and per experiment are summarized in Table 1-6. In particular, for the  $b\bar{b}b\bar{b}$  channel, the ATLAS (CMS) expected significance is around 0.61 (0.95) standard deviations. Furthermore, the measurement of negative-log-likelihood as a function of the  $\kappa_\lambda$  coupling modifier, assuming that the SM process occurs, is presented by channel in Figure 1-15 A). From

the combination of the likelihood scans in different channels, the ATLAS (CMS)  $\kappa_\lambda$  measurement is expected to be  $1^{+0.90}_{-0.75}$  ( $1^{+0.90}_{-0.65}$ ) at 68% CL.

The combination of the ATLAS and CMS HL-LHC projections is reported in ref. [83]. The combined significance yields a value of 4.0 standard deviations. The combined likelihood scan is presented in Figure 1-15 B). From combined likelihood, the expected coupling measurement is  $\kappa_\lambda = 1^{+0.50}_{-0.48}$  at 68% CL. The results of these projections show that we may achieve evidence for the SM HH production process and measure the Higgs self-coupling at around 50% level precision. However, further detector and algorithm developments (e.g. triggering) are expected. These projections are conservative because larger datasets would enable more sophisticated analyses, which can boost the performance. Therefore, the actual results may be better than this projection. Dedicated future colliders will be needed to push further the precision of these measurements, e.g.  $e^+e^-$  colliders (ILC and FCC-ee) at center-of-mass energies above 0.5 TeV [48].

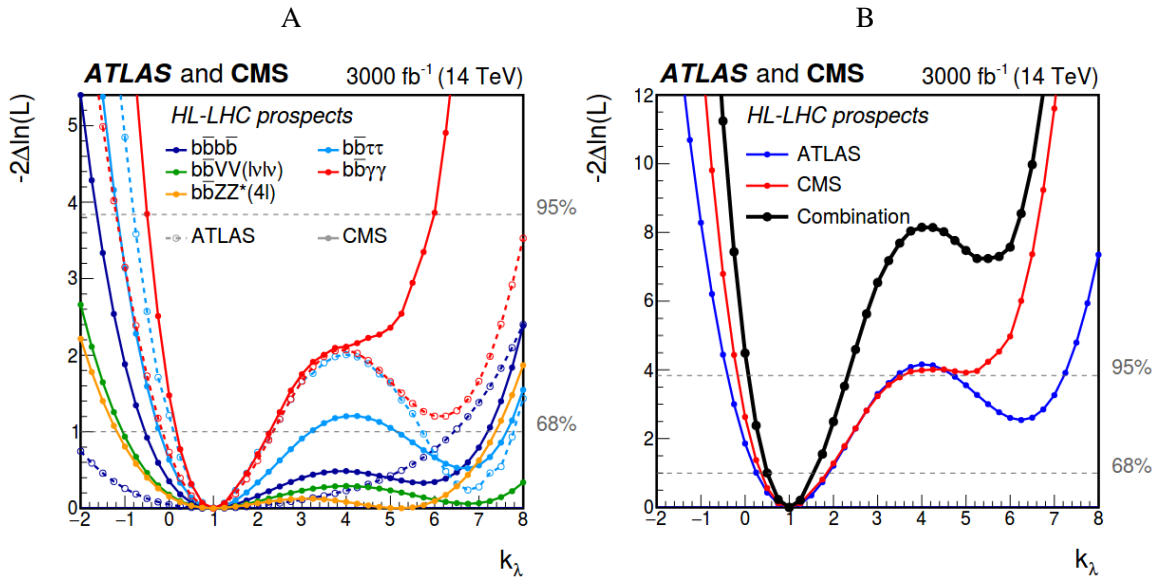


Figure 1-15. Minimum negative-log-likelihood as a function of the  $\kappa_\lambda$  modifier obtained by performing a fit to the expected background and SM signal [83]. A) Contours for different channels in the ATLAS and CMS projections, represented in dashed and solid lines, respectively. B) Combined results (black) from the CMS (red) and ATLAS (blue) projections.

## CHAPTER 2 EXPERIMENTAL APPARATUS

### 2.1 The Large Hadron Collider

European Center for Nuclear Research, or CERN, is the host of several experimental facilities. The CERN flagship experimental project is the Large Hadron Collider (LHC). The CERN accelerator complex, located at the French-Swiss border, consists of a chain of machines that accelerate particles (e.g. protons and heavy ions) in various steps until the nominal energy for collisions is achieved. The accelerator complex is illustrated in Figure 2-1.

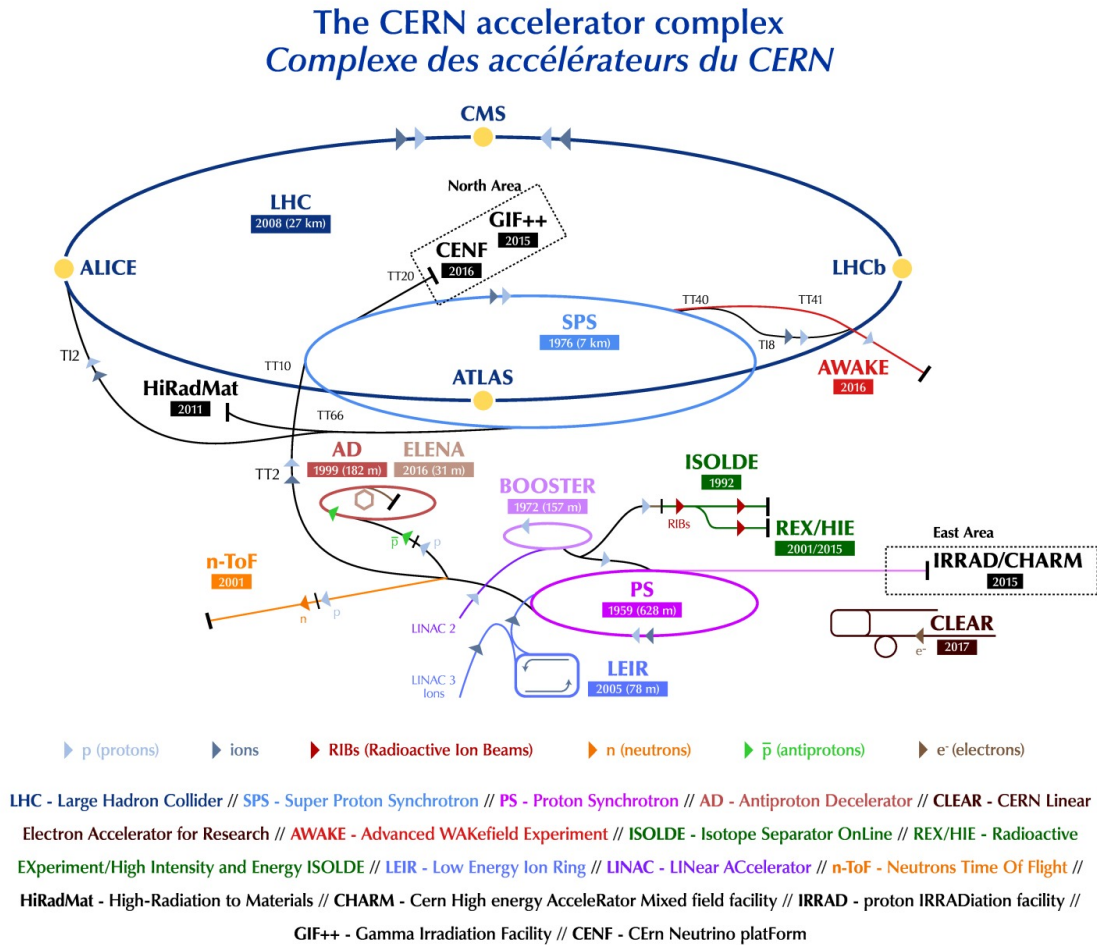


Figure 2-1. The CERN Accelerator Complex [84].

The acceleration of protons starts by stripping the electrons on hydrogen atoms originally contained in a hydrogen bottle via a duoplasmatron proton source. Next, those protons are injected and accelerated successively by the LINAC-II, Booster, PS and SPS machines. The SPS



output is a beam of protons organized in bunches with energy of 450 GeV, which is injected to the two LHC rings (one in clockwise and other in counter-clockwise direction). The two LHC beams are circulated in opposite directions and accelerated until they achieve an energy of around 6.5 TeV. Once the beams are stable, proton bunches with around  $10^{11}$  protons each are collided.

The LHC is a powerful tool that is capable to achieve high energy to probe the behavior of matter and its interactions. It is a circular superconducting accelerator and collider with 26.7 km of circumference, which is designed to produce pp collisions at up to 14 TeV of center-of-mass energy ( $\sqrt{s}$ ) by colliding proton bunches every 25 nanoseconds in four interaction points (IPs) along the ring. There are experiments located at each of these points: ATLAS (A Toroidal LHC ApparatuS) and CMS (Compact Muon Solenoid) are general-purpose experiments that concentrate on measurements of SM processes and BSM physics searches at the TeV scale, the LHCb (LHC-beauty) experiment focuses on B-hadron physics, and ALICE (A Large Ion Collider Experiment) studies heavy-ion collisions phenomena (e.g. quark-gluon plasma).

A key parameter to describe the collider operation is the instantaneous luminosity ( $\mathcal{L}$ ), which depends on the beam properties as described in equation 2-1 and listed in Table 2-1. The LHC operation was originally designed to achieve an instantaneous luminosity of around  $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ . High luminosities are important because they increase the production rate of interesting physics process, and thus our potential and sensitivity for physics discoveries. However, high luminosities imply the increase of the number of pp interactions per bunch crossing or pile-up (UP), which poses challenges to the experiments for reconstruction and radiation harness.

$$\mathcal{L} = \frac{N_b^2 n_b f \gamma}{4\pi\epsilon\beta \sqrt{1 + (\frac{\theta_c \sigma_z}{2\sigma^*})^2}} \quad (2-1)$$

The expected number of produced events of a particular process during the LHC collisions (without taking into account the detector acceptance) is predicted as  $N = \sigma \cdot L$ , where  $\sigma$  is the theoretical production cross section ( $\sigma$ ), and  $L$  is the integrated luminosity ( $L$ ) defined as the integral of the instantaneous luminosity over the LHC operation time. The units of integrated luminosity are the inverse barn ( $\text{b}^{-1}$ ), where  $1 \text{ b}^{-1} = 10^{24} \text{ cm}^{-2}$ .



Table 2-1. Parameters used in the calculation of the instantaneous luminosity at the Large Hadron Collider proton-proton collisions [85].

| LHC design experimental parameters                | Design or nominal values |
|---------------------------------------------------|--------------------------|
| $f$ , revolution frequency                        | 11.245 kHz               |
| $n_b$ , number of proton bunches per beam         | 2808                     |
| $N_b$ , number of proton per bunch                | $1.15 \times 10^{11}$    |
| $\beta$ , optical beta function at the IP         | 55 cm                    |
| $\sigma_z$ , RMS bunch length                     | 7.55 cm                  |
| $\sigma_*$ , transverse RMS beam size             | $16.7 \mu\text{m}$       |
| $\gamma$ , relativistic gamma factor              | 7461                     |
| $\epsilon$ , normalized transverse beam emittance | $3.75 \mu\text{m rad}$   |
| $\theta_c$ , crossing angle at the IP             | $285 \mu\text{rad}$      |

The baseline plan of the LHC at past, present and future operations is presented in Figure 2-2. The LHC particle beams were circulated for the first time in September 2008 but days later a failure of an electrical connection produced led to severe damage to the superconducting magnets, their connections and the vacuum pipe. After investigating the problem and over a year of repairing, the LHC beams were circulated back again in November 2009, and the first run of the LHC (Run-I) started at the beginning of 2010.



Figure 2-2. The Large Hadron Collider schedule starting from Run-1 operations (2011) to the next decades and beyond. In red, it is presented the expected proton-proton collision center-of-mass energy and the instantaneous luminosity.

In the LHC Run-1 (2010-2012), pp collision data was collected at  $\sqrt{s} = 7$  and 8 TeV with bunches crossing every 50 nanoseconds. The collected Run-1 dataset was analyzed for the discovery of the Higgs boson. Then, during the first long shutdown (LS1), 2013-2014, the LHC superconducting magnets were trained for the operation with 6.5 TeV proton beams, thus pp collisions at 13 TeV of center-of-mass energy. Moreover, detector upgrades took place. For instance, the CMS experiment performed the upgrade of its Level-1 trigger (hardware and architecture), in preparation for the Run-2 conditions.

The Run-2 operation of the LHC took place from 2015-2018, with 13 TeV pp collisions. The LHC reached a record instantaneous luminosity of  $2.1 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ , corresponding to an average of 55 PU interactions. In CMS, the measured Run-2 average of PU interactions was around 34 as presented in Figure 2-3 A). The total integrated luminosity delivered by the LHC during the Run-2 operations was around  $164 \text{ fb}^{-1}$ , as presented versus time in Figure 2-3 B).

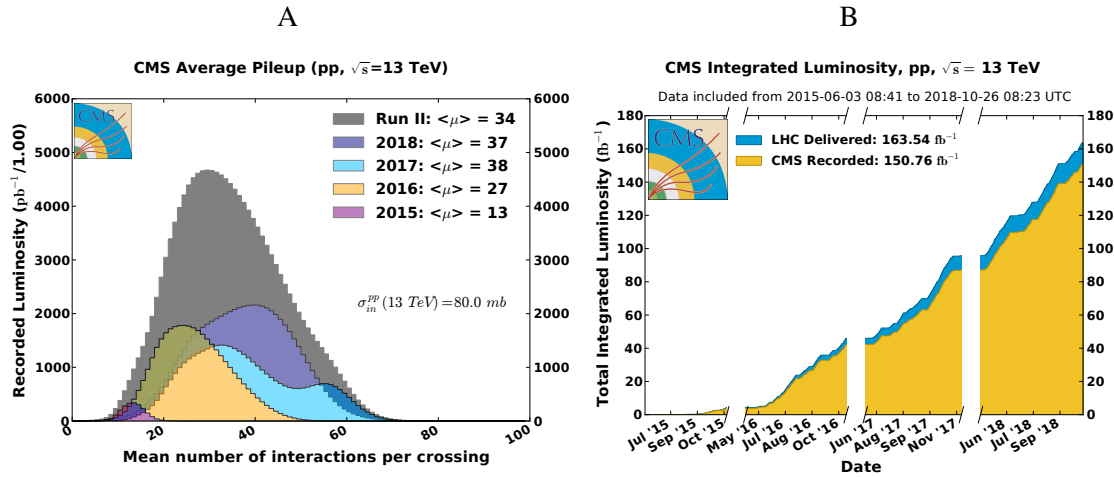


Figure 2-3. CMS operation conditions in proton-proton collisions during the LHC Run-2 data-taking. A) Distribution of pile-up average versus recorded luminosity, B) Total integrated luminosity versus time.

At the time of writing this dissertation, the LHC is going under the second long-shutdown (LS2), 2019–2021, in preparation for the Run-3 operations. Due to the current Covid-19 pandemic, the LHC schedule has been delayed. Run-3 is expected to start at the beginning of 2022, and to last until 2024 or 2025. Then, after the third long shutdown (LS3) of 2–3 years, the LHC will operate at its ultimate high luminosity configuration, the HL-LHC.

The HL-LHC operation period, expected to last for about a decade, is known as the LHC Phase-2. Phase-1 is the current operation until the end of Run-3. The HL-LHC accelerator is expected to produce 14 TeV pp collisions and to reach an instantaneous luminosity of  $5\text{--}7.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ , corresponding to 200–300 average PU interactions. The expected dataset corresponds to 3000-4000  $\text{fb}^{-1}$  of integrated luminosity (roughly 20-27 times more data than the one used in this dissertation). Major consolidation and upgrades of the LHC accelerator are ongoing in preparation for the HL-LHC era. The LHC experiments will undergo an important upgrade program to successfully operate in the challenging collision conditions (more in Section 2.2.5).

## 2.2 The CMS Experiment

### 2.2.1 Description

The Compact Muon Solenoid (CMS) experiment is a general-purpose experimental apparatus located 100 meters underground within a cavern at the LHC Point 5 (P5) in Cessy, France [86]. The CMS detector surrounds the LHC IP with dedicated particle detector layers (or subsystems) to reconstruct the collisions products (electron, muons, photons, charged and neutral hadrons) and study the underlying physics phenomena. This detector has a cylindrical symmetry and is uniquely characterized by a high magnetic field superconducting solenoid. More details on the detector layers are presented by the next subsection.

The conventional coordinate system used by CMS is illustrated in Figure 2-4. The origin is fixed at the interaction point inside the detector. The X-axis points towards the center of the LHC, the Y-axis is perpendicular to the plane defined by the LHC ring, and the Z-axis is parallel to the beam line following the anticlockwise direction. The azimuthal angle  $\phi$  is measured with respect to the X-axis over the X-Y plane. The polar angle  $\theta$  is defined with respect to the Z-axis over the Y-Z plane. The  $\theta$  angle is converted to the variable pseudorapidity  $\eta$  by the relation  $\eta = -\ln[\tan(\frac{\theta}{2})]$ . The angular distance between two points is measured in units of  $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$ . The transverse component of vectorial properties such as the momentum are defined as the projection of the vector in the X-Y plane (e.g. transverse momentum  $p_T$ ).

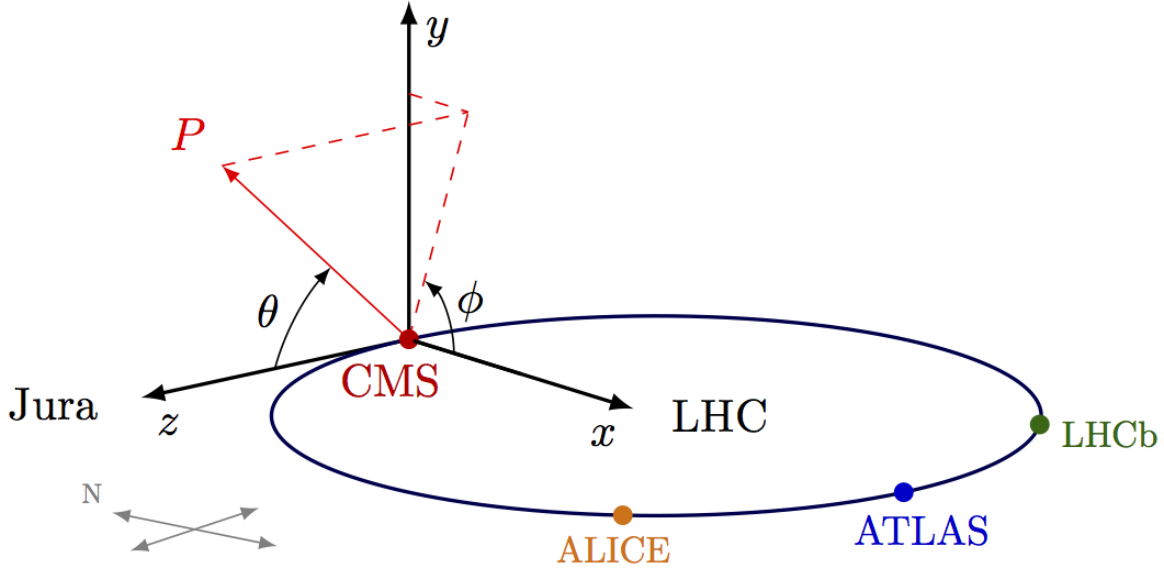


Figure 2-4. The CMS conventional coordinate system.

### 2.2.2 Detector Layers

At the LHC design luminosity, an average of 20 pile-up interactions are expected to occur simultaneously at the interaction point every 25 nanoseconds. Consequently, detectors with large number of channels and good timing resolution are needed to reconstruct and identify the products of the interaction under study with respect to the other interactions in the same bunch crossing. Moreover, the high-particle multiplicity environment ( $\sim 1000$  charged particles per bunch crossing) produce high radiation levels which can cause detector damage and thus, radiation-hard detectors/electronics are needed.

The CMS detector was designed to conduct a successful physics program involving diverse final states under the aforementioned data-taking constraints by satisfying the following detector requirements [86]:

- Muon identification with good reconstructed momentum resolution ( $< 10\%$  at  $P \sim 1$  TeV)
- High momentum resolution and reconstruction efficiency of charged-particles
- Efficient triggering of  $\tau$  leptons and b-quark hadronic jets
- Good resolution and efficient leptons and photons energy and isolation measurements
- Good missing-transverse energy measurement and dijet-mass through a hermetic, granular and angular coverage

## 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 ( $100 \times 150 \mu\text{m}^2$ )  $\sim 1.9 \text{ m}^2 \sim 124\text{M}$  channels  
Microstrips ( $80\text{--}180 \mu\text{m}$ )  $\sim 200 \text{ m}^2 \sim 9.6\text{M}$  channels

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

MUON CHAMBERS  
Barrel: 250 Drift Tube, 480 Resistive Plate Chambers  
Endcaps: 540 Cathode Strip, 576 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  $\text{PbWO}_4$  crystals

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

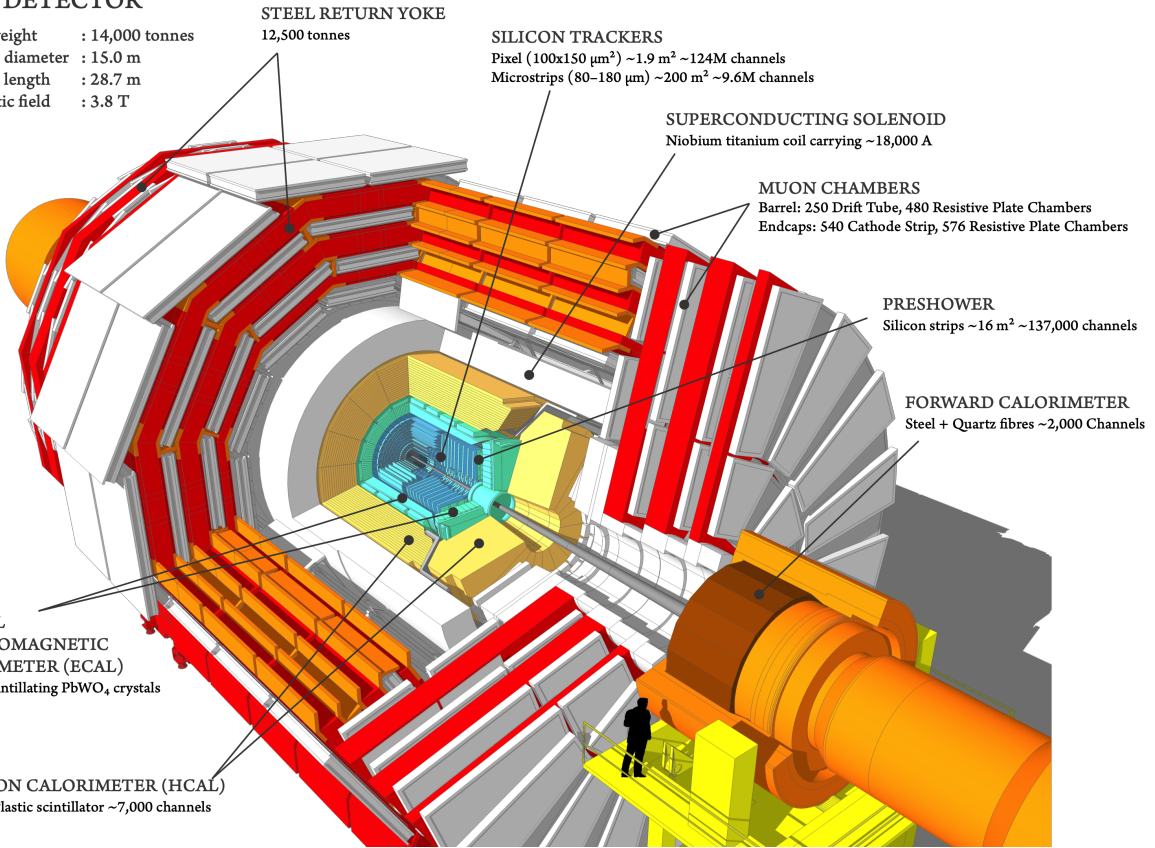


Figure 2-5. Three-dimensional view of the subdetectors of the CMS experiment.

The CMS detector layout and dimensions are presented in Figure 2-5. A tracker, an electromagnetic and a hadron calorimeter are contained within the superconducting solenoid volume. The first detector layer is a full silicon-based inner tracking system composed of pixels and strip trackers, which information is processed to detect the interaction of charged particles and measure their trajectory and momentum. The energy of electromagnetic particles (photon and electrons) is measured from their energy deposits in a hermetic and homogeneous electromagnetic calorimeter layer made of lead tungstate ( $\text{PbWO}_4$ ) crystals. Then, a hadronic calorimeter layer with a wide and hermetic coverage measures the energy of deposits from charged and neutral hadrons (e.g. pions). Then, the solenoid is surrounded by the muon system, which embedded in a steel flux-return yoke [86].

### 2.2.2.1 Inner tracking systems

The purpose of the inner tracking system is to provide an efficient and precise measurement of the trajectories of charged particles ( $|\eta| < 2.5$ ), and to reconstruct the points of interactions of protons, referred to as vertices. This system has a primary importance to reconstruct the points of decay of B hadrons away from the beam line, or secondary vertices, which provide the main input to the jet flavor identification. Furthermore, it brings information for high-level triggers. The detector design has a high granularity, fast response, radiation hardness and as minimum material as possible (to reduce the impact on the momentum resolution from multiple scattering) to operate successfully under high collision rates, high particle flux and radiation levels.

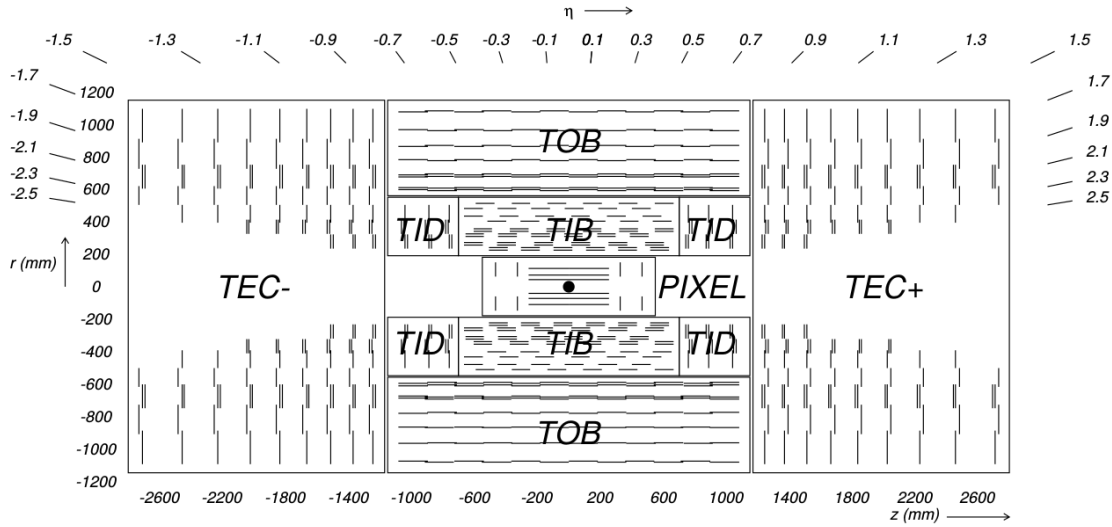


Figure 2-6. Layout of the CMS tracking system in the longitudinal direction [86]. Note: This layout illustrates the Pixel detector before its upgrade.

The layout of tracking system modules is illustrated in Figure 2-6. The Pixel tracking detector is the innermost part of the tracking system and thus the closest to the IP. The previous Pixel detector, used until 2016 data-taking, comprised of pixel detector modules organized in three cylindrical layers (BPIX) and two forward/backward disks (FPIX), covering an area of about 1 m<sup>2</sup> and 66 million pixels with hit resolution between 10-20 μm. In the outermost tracking system is the Tracker detector. It comprises detector strip sensors with a total of 9.3 million strips covering 198 m<sup>2</sup> of area. The strips are organized in three subsystems: Tracker Inner Barrel and

Disks (TIB/TID), Tracker outer barrel (TOB), and Track Endcap (TEC). The Tracker covers a pseudorapidity region up to  $|\eta| = 2.5$  [86].

The new or (upgrade) pixel detector was installed during the 2016/2017 shutdown in preparation for the 2017 data-taking, to maintain and improve the tracking performance under high luminosity conditions (up to  $2 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$  and exceeding an average of 50 pile-up interactions). The new detector consists of four cylindrical layers and three disks [87]. A comparison between the original and new pixel detector organization is illustrated in Figure 2-7.

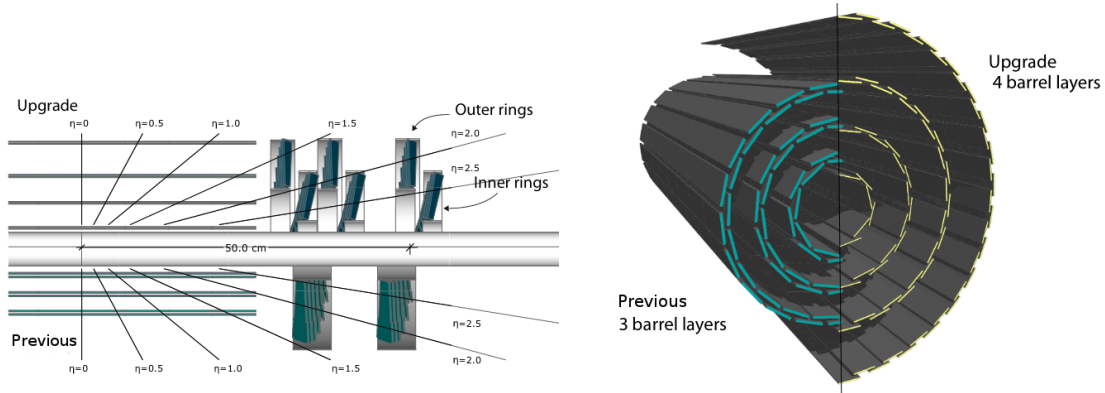


Figure 2-7. The CMS pixel detector [87]. Left: Half longitudinal section of the pixel detector layout previously (bottom) and after the upgrade (top). Right: Pixel barrel layers geometry previously (left) and after the upgrade (right).

### 2.2.2.2 Electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) design was motivated in part by the goal of identifying two photons from the postulated Higgs boson. ECAL is a calorimeter based  $\text{PbWO}_4$  crystals satisfying the requirements to operate under the LHC environment: hermetic, homogeneous, fast, fine granularity and radiation resistant. The radiation length ( $X_0$ ) is the average length to reduce the energy of an incident electron by a factor of  $1/e$ . The thickness of ECAL is larger than  $25 X_0$ . The energy of the electromagnetic showers produced from incident photons/electrons is measured from associated scintillation light as is proportional to the incident particle energy. The scintillation light is collected by avalanche photodiodes (APDs) in the ECAL barrel region (EB,  $|\eta| < 1.48$ ), and vacuum phototriodes (VPTs) in the ECAL endcaps (EE,  $1.48 < |\eta| < 3.0$ ). In addition, a preshower sampling calorimeter (ES) is located in front of the



EE section covering  $1.65 < |\eta| < 2.6$ , as illustrated in Figure 2-8 A), to help to initiate the electromagnetic showers from incoming photons/electrons, and to identify neutral pion decays into two photons.

The ECAL layout is presented in Figure 2-8 B). A total of 61200 (7234) crystals of front-face cross section of around  $22 \times 22 \text{ mm}^2$  ( $28.6 \times 28.6 \text{ mm}^2$ ) and length of 230 (220) mm, are mounted on the EB (EE) region. The EB part is organized in 36 ‘supermodules’, each covering half of the barrel region and divided in four modules. The crystals of the EE region are structured in 2 dees and grouped in units of  $5 \times 5$  crystals denominated as ‘supercrystals’ [86].

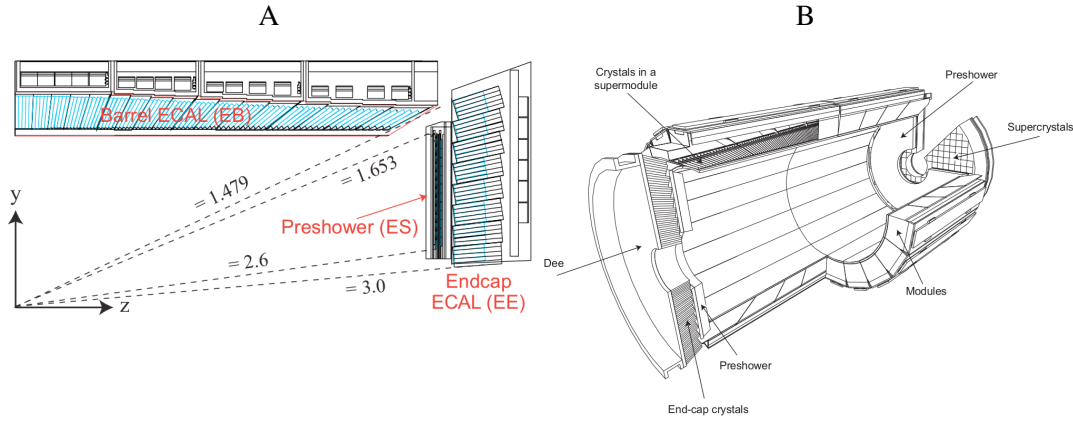


Figure 2-8. The CMS electromagnetic calorimeter (ECAL) [86, 88]. A) Longitudinal section of the geometrical configuration, B) Barrel and endcap configuration layout.

The performance of the ECAL detector was studied through test beam measurements on a barrel supermodule using electron beams of energies between 20 and 250 GeV. The energy resolution ( $\frac{\sigma}{E}$ ) can be parametrized as a function of the energy (E) using eq. 2-2, where S is the stochastic term, N is the noise and C is the constant term. The test beams measured the coefficients to be  $S=2.8\%$ ,  $N=0.12$  and  $C=0.30\%$  [86]. Then, the energy resolution of the was measured using  $Z \rightarrow e^+e^-$  events in the 2010-2011 collision data. The measured resolution of the electrons with  $E_T \approx 45 \text{ GeV}$  was around 2% (2-5%) in  $|\eta| < 0.8$  (elsewhere) [89]. The excellent ECAL calibration and resolution performance were key for the 2012 Higgs boson discovery.

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2 \quad (2-2)$$



### 2.2.2.3 Hadronic calorimeter

The hadronic calorimeter (HCAL) is an absorber/scintillator sampling calorimeter surrounding ECAL. It measures the energy of charged and neutral hadrons, which is crucial for the reconstruction of hadronic jets and the missing transverse energy from neutrinos (or exotic particles). Incident hadrons interact with the absorber via nuclear force and generate hadronic showers in the absorber and scintillator. Then, the associated scintillator light is collected by wavelength-shifting fibres (WLS) and channeled to photodetectors for readout. The interaction length ( $\lambda_I$ ) of a material is the average distance a hadron travels before an inelastic nuclear interaction. Depending on  $\eta$ , the HCAL thickness varies within 10-15  $\lambda_I$ , and thus a good containment of the hadronic shower [86, 88]. HCAL is divided by four subdetectors: Hadron barrel (HB,  $|\eta| < 1.4$ ), Hadron outer (HO), Hadron Endcap (HE,  $1.3 < |\eta| < 3.0$ ) and Hadron forward (HF,  $3.0 < |\eta| < 5.0$ ). The layout of the HCAL detector is presented in Figure 2-9.

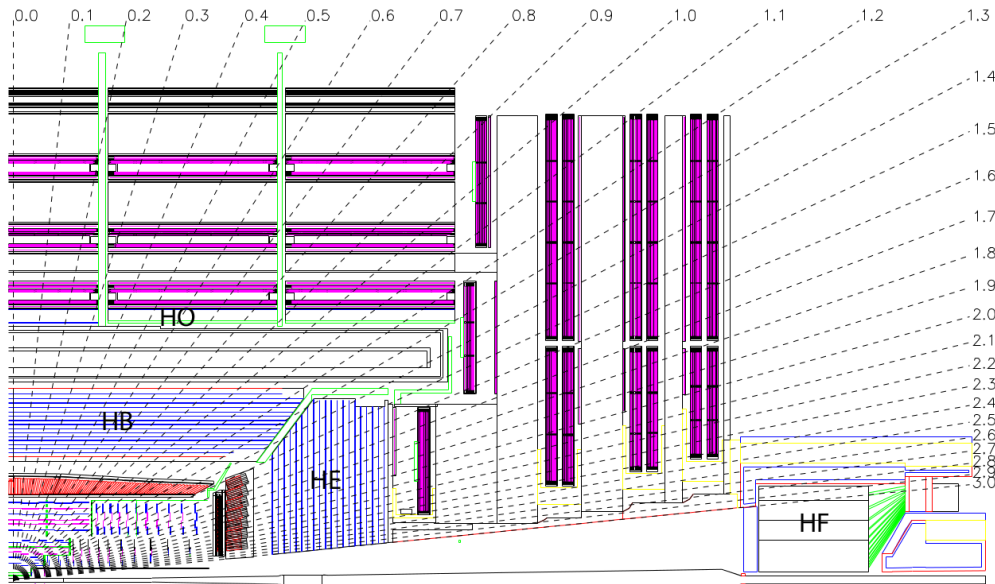


Figure 2-9. Longitudinal layout of the CMS hadronic calorimeter (HCAL) [86].

The HB and HE are brass/plastic scintillator tiles sampling calorimeters located inside the solenoid. In the HB and HE sections, there are a total of 2304 readout ‘towers’ formed from the scintillator tiles with a segmentation of  $(\Delta\eta, \Delta\phi) = (0.087, 0.087)$  and  $(\Delta\eta, \Delta\phi) = (0.17, 0.17)$ , respectively. The HCAL photodetectors are multichannel hybrid photodiodes (HPDs). HO and

HF subdetectors are located outside the solenoid. The HO subdetector provides additional layers of scintillator to sample the energy of hadron showers leaking from the rear of the calorimeters, improving the  $E_T$  resolution. The HF calorimeters, located at around 11 meters from the interaction point, consist of steel absorber and quartz-fiber layers. Signals come from the Cerenkov light emitted from the quartz fibers, which is collected by photomultipliers (PMTs). The HF materials were chosen in order to contain hadronic showers from energetic forward jets in the intense radiation levels in the forward region [86, 88].

#### 2.2.2.4 Superconducting solenoid

The CMS superconducting solenoid (or magnet) plays a key role in bending the trajectories of charged particles to perform a precise measurement of their transverse momentum using the tracking and muon systems. An artistic illustration of the CMS magnet is presented in Figure 2-10. It is design to reach a 4 T magnetic field and 2.6 GJ of stored magnetic energy within a cylindrical region with 6 m of diameter and 12.5 m of length. The magnet ‘cold mass’ consist of a 4-layer winding of a reinforced NbTi conductor within a cryostat operating at 1.2 K [86, 88].

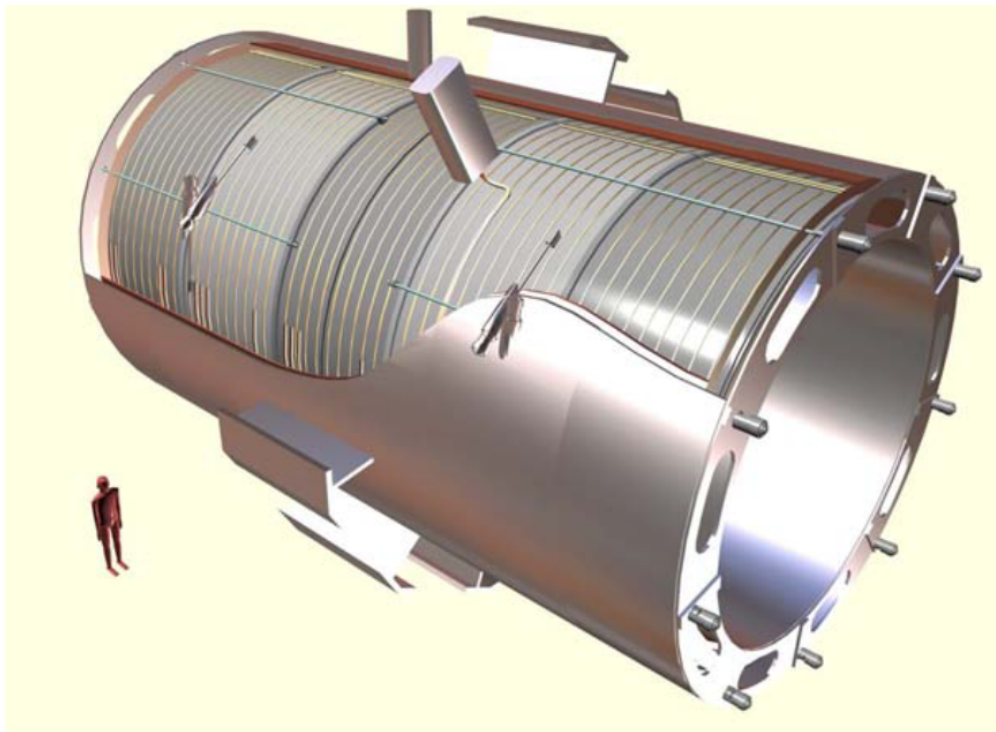


Figure 2-10. Three-dimensional artistic view of the CMS magnet [86].

The magnetic flux is returned through an iron yoke, which comprises 5 wheels (-2,-1,0,1,2) in the barrel region, and two endcaps, as seen in Figure 2-5. The iron yoke holds the detectors of the muon system, and it serves as hadron absorber.

### 2.2.2.5 Muon system

The precise measurement of muons is crucial in the CMS experiment because those particles are produced in many interesting signatures. One example is the Higgs boson decay into a Z boson pair with four muons in the final state ( $H \rightarrow ZZ \rightarrow 4\mu$ ), which was the golden channel in the Higgs boson discovery due to the excellent mass resolution. The muon system is designed for muon identification, precise momentum and charge reconstruction, and triggering over the kinematic range of the LHC. In Figure 2-11 is illustrated the layout of the CMS muon system and its three types gaseous chambers: Drift tubes (DTs) [90], Cathode strip chambers (CSCs) [91] and Resistive Plate Chambers (RPCs) [92]. Depending on  $\eta$ , different detector types are mounted to achieve an efficient detector response in the full region of  $|\eta| < 2.4$  [86, 88].

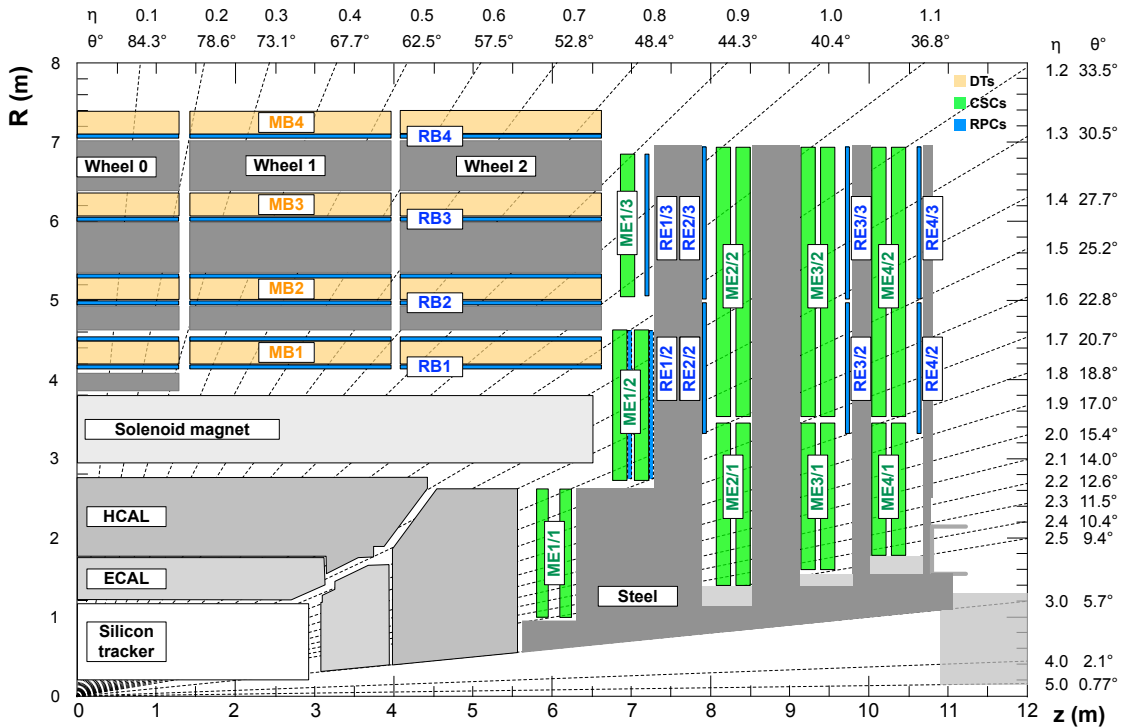


Figure 2-11. One quadrant of the CMS muon system in the longitudinal direction. Drift tubes (DTs), Cathode strip chambers (CSCs) and Resistive Plate Chambers are presented in orange, green and blue color, respectively.

In the barrel region, the neutron-induced background is small, the muon rate is low, and the magnetic field is approximately constant. Consequently, a total of 250 DT chambers are used in the  $|\eta| < 1.2$  region. They are organized in 4 stations (or radial layers) and embedded in the flux-return yoke wheels. In the endcap region, the muon rates and background levels are high, and the magnetic field is not uniform. Therefore, CSCs are used in the  $0.9 < |\eta| < 2.4$  region due to their fast timing, fine segmentation and radiation hardness. A total of 468 CSCs are positioned in the two endcaps. RPCs are characterized by a fast response, good time resolution, but have lower space resolution than CSCs and DTs chambers. A set of RPCs is added in both barrel and endcap regions covering  $|\eta| < 1.6$ , to identify the correct bunch crossing, increase muon hit redundancy to reduce background, and resolve ambiguities from multiple hits in the same chamber [86, 88].

### 2.2.3 Event Reconstruction

A global description of the collision event, where all stable particles are identified, can be achieved by the correlating and combining the detector information associated to their interaction with the different subdetectors (illustrated in Figure 2-12). This type of approach is denominated as Particle-Flow (PF) reconstruction, and first was used by the ALEP experiment at the LEP  $e^+e^-$  collider [93]. The CMS design properties (e.g. fine-granularity of the detectors) enable the use of the PF reconstruction. Since the first implementation of it for physics analysis in 2009 and triggering in 2010, all physics analyses use PF reconstruction.

The basic PF elements are tracks and calorimeter clusters. Then, they are used for PF identification and reconstruction to produce PF objects/candidates associated to muons, electrons, photons, charged and neutral hadrons. The PF candidates are used after for building more complex physics objects such as hadronic jets, tau leptons, and missing transverse momentum. The PF candidates are the objects used for physics analysis. In what follows it is summarized the CMS methods use to reconstruct all event objects in the context of the PF algorithm. More detailed information is found in Ref. [94].

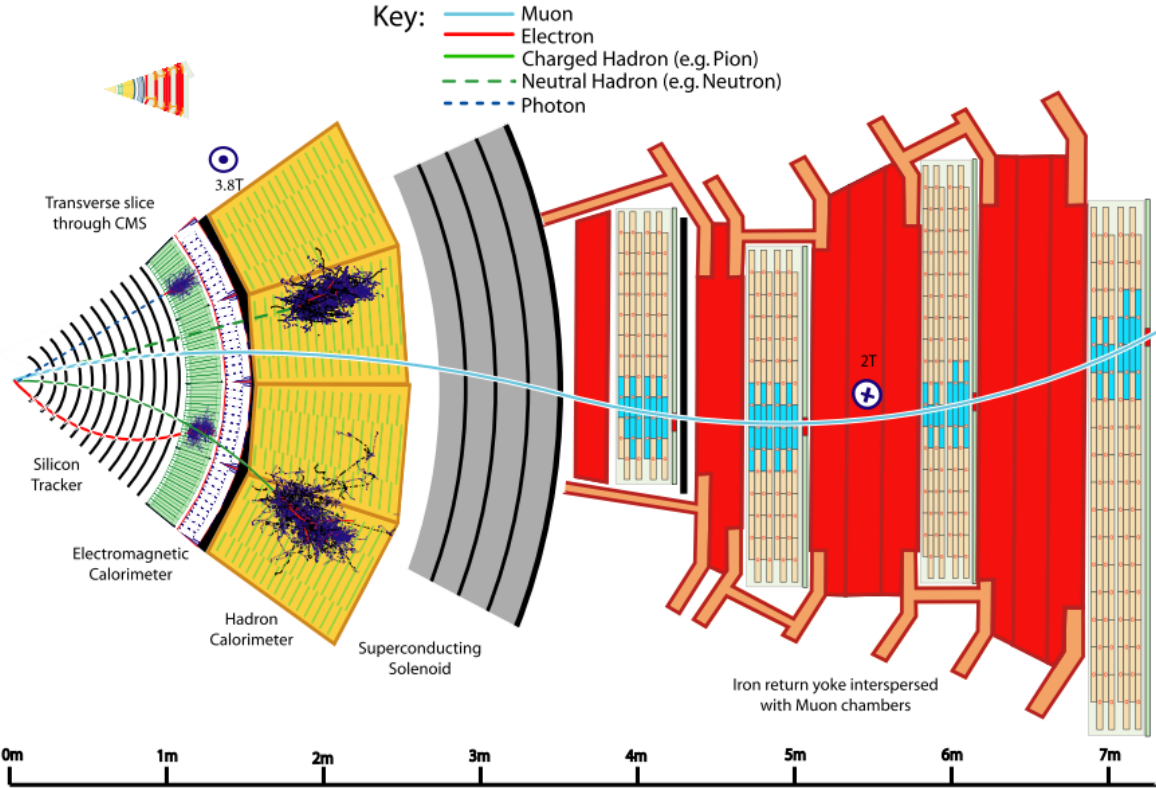


Figure 2-12. A transverse slice of the CMS detector representing the interaction of stable particles (electron, muon, photon, charged and neutral hadrons) with each of the subsystems.

### 2.2.3.1 Tracks and calorimeter clusters

Experimentalists use track-finder algorithms to find a set of detector hits/stubs compatible with a charged particle track. Then, a fit is performed to obtain the track properties (origin, momentum, direction, etc.). The trajectories of charged particles crossing the CMS inner tracking system (inner track) are reconstructed from hits using an iterative tracking method. In this approach, a combinatorial track finder based on Kalman Filtering (KL) is applied in ten successive iterations. Each iteration has a dedicated seeding and type of track target. The hits associated with built tracks are masked in the next iteration. The reconstruction efficiency surpasses the one from single iteration methods and keeps misreconstruction rates at similar levels [94].

The muon system tracking is performed with high efficiency over the different regions of the detector, and independent of the tracker-based or PF reconstruction. There are three possible types of muons:

- Standalone muon: The muon track is reconstructed from the information of the CSC, DT and RPC hits using a pattern recognition algorithm for hit selection and a fit method for getting the track properties
- Tracker muon: It is an inner track with  $p_T > 0.5$  GeV and  $p > 2.5$  GeV, which matches with at least one muon hit when extrapolated to the muon system.
- Global muon: Standalone muon tracks are matched to inner tracks. Then, for each match that is found, all associated hits are used to fit a global muon track.

A dedicated clustering algorithm is used to generate clusters from the energy deposits in the calorimeters. The clustering method is performed independently by ECAL, HCAL and preshower detectors. In essence, the clustering is as follows. First, cluster seeds are found based on cells passing a seed threshold and larger energy with respect to their neighbors. These neighbors can be the four cells sharing a side (HCAL) or eight cells sharing corners and sides (ECAL and preshower). Then, topological clusters are formed by adding to the seed cluster more cells with at least a common corner and energy twice the noise level. Lastly, an algorithm based on a Gaussian-mixture model reconstruct the clusters with position and energy from the topological cluster. This clustering method applied at HF because the electromagnetic or hadronic component of a cell can be separated into HF EM cluster and HF HAD cluster, respectively [94].

### 2.2.3.2 Particle-flow identification and reconstruction

In the PF algorithm, the process starts with an algorithm that links PF elements based on the nearest neighbors obtained with a k-dimensional tree across the  $(\eta, \phi)$  plane. The link conditions depend on the PF elements involved. The output is called PF blocks, consisting of PF elements with a direct or indirect (via common elements) links. Then, the next step is to identify and reconstruct the particles in each PF block [94].

Muons are the first identified particles in the PF block. The muon identification is performed based on the global-muon and tracker-muon properties. The identification minimizes the muon misidentification of charged hadrons. Isolated global muons are identified by taking into account the characteristics of the surrounding tracks and calorimeter deposits within  $\Delta R = 0.3$ . Non-isolated global muons are identified using tight-muon requirements, at least three

muon station segments, and calorimeter deposits compatible with muons, reducing the muon misidentification of high- $p_T$  hadrons. The momentum of the muon is the one from the inner track if  $p_T < 200$  GeV. Otherwise, the  $p_T$  associated to the track fit minimizing the  $\chi^2$  is chosen. Once the PF muon candidates are found, their PF elements are masked against further processing [94].

Electrons are the second objects that are studied in the PF block. When electrons propagate in the inner tracker materials, they often emit bremsstrahlung photons before reaching ECAL. These photons can also produce  $e^+e^-$  pairs, they can emit more photons, and so on. The PF algorithm combines the track and calorimeter information for identification and reconstruction, collecting the energy of soft and energetic radiated photons. Isolated photon identification and reconstruction are performed in the same step [94]

The ECAL-based seeding of electrons uses the information of ECAL clusters ( $E_T > 4$  GeV), to infer the position of the hits associated to well isolated and energetic electrons. The energy is determined by grouping clusters into superclusters within a small  $\eta$  window and extended  $\phi$  window, compatible with the electron magnetic bending. The performance of this method relies on gathering all emitted photons, and thus, this approach is not efficient for non-isolated electrons in jets or soft- $p_T$  electrons. Therefore, a tracker-based seeding for electrons is performed to complement the ECAL-based seeding. First, all inner tracks passing a dedicated preselection ( $p_T > 2$  GeV, number of hits and the fit  $\chi^2$  cuts) are fitted again with a Gaussian-sum filter (GSF). These GSF tracks are more compatible with the electron hypothesis as account for the large energy losses along the trajectory [94].

In the PF block, the electron candidate is seeded from a GSF track linked to an ECAL cluster with no more than three additional track links, whereas a photon candidate is seeded from a ECAL cluster with  $E_T > 10$  GeV and without GSF track link. All ECAL clusters linked to the superclusters or the GSF track are associated to the candidates. Tracks linked to these ECAL clusters are associated to the candidate if the track momentum and the energy of the linked HCAL cluster is compatible with the electron hypothesis. The photon energy is obtained from the total energy of the associated ECAL clusters, which is corrected as a function of energy and  $\eta$ , to



account for any misses during the association process. The photon direction is the one of the supercluster. The electron energy is assigned from corrected ECAL energy and GSF track momentum information, while the electron direction is taken from the GSF track [94].

Non-isolated and isolated PF electron candidates must satisfy a threshold on a multivariate score based on dedicated BDTs trained with up to 14 variables. If photons are isolated from tracks and calorimeter clusters, and the HCAL - ECAL ratio is compatible with the electron hypothesis, then they are classified as PF photons. All tracks and ECAL or preshower clusters associated to the PF electrons and isolated photons are excluded from further processing [94].

The last candidates classified in the PF block are charged hadrons (e.g.  $\pi^\pm$  and protons), neutral hadrons (e.g. neutrons), and non-isolated and soft photons (e.g.  $\pi^0 \rightarrow \gamma\gamma$ ), typically produced in hadronic processes (e.g. jets). Within the tracker acceptance ( $|\eta| < 2.4$ ), all isolated ECAL (HCAL) clusters without a track link are identified as photons (neutral hadrons), while charged hadron candidates are identified from the HCAL clusters with track link. However, beyond the tracker acceptance, there is no distinction between charged and neutral hadrons. In this case, ECAL clusters with a HCAL-link are assumed to be hadron candidates, otherwise, if there is no such link, they are classified as photon candidates. Then, photon and hadron energies are calibrated from their raw calorimeter energies. HF EM and HF HAD clusters are classified as HF photons and HF hadrons, respectively [94].

### 2.2.3.3 Jet reconstruction and flavor identification

The quark or gluon hadronization generates a collimated shower of particles, which experimental signature is known as hadronic jet or jet. The jet is reconstructed using physical objects such as calorimeter towers or clusters (Calo jets), PF candidates (PF jets) or simulated/generated particles (Gen/Ref jets). A jet clustering algorithm determines which objects are included in the jet. A good jet algorithm is infrared and collinear safe, i.e. the set of jets should be unchanged by soft gluon emission and collinear splitting. Typically, the jet algorithms iteratively find two objects in the event which are closest based on a defined metric and combines them. Then, the jet 4-vector is obtained from the sum of all constituent particle 4-vectors.



In the CMS experiment, the most used jet clustering algorithm is the anti- $k_T$  [95, 96]. A radius parameter of  $R = 0.4$  (0.8) is used to reconstruct small-cone jets (large-cone) jets, known as AK4 (AK8) jets. The AK4 PF (Calo) jets are typically used requiring  $p_T > 15$  (20) GeV, otherwise they are considered unreliable. The reconstruction of a di-jet simulated event is illustrated in Figure 2-13, where it is observed the good performance of PF jets with respect to Calo jets. Note that tracks and clusters produced from PU interactions can impact the performance of the jet reconstruction. This is mitigated with the charged-hadron subtraction (CHS) method, which excludes any charged particles pointing towards reconstructed PU vertices from the jet clustering and assigns the jet 4-momenta correction from the impact of neutral PU particles [94].

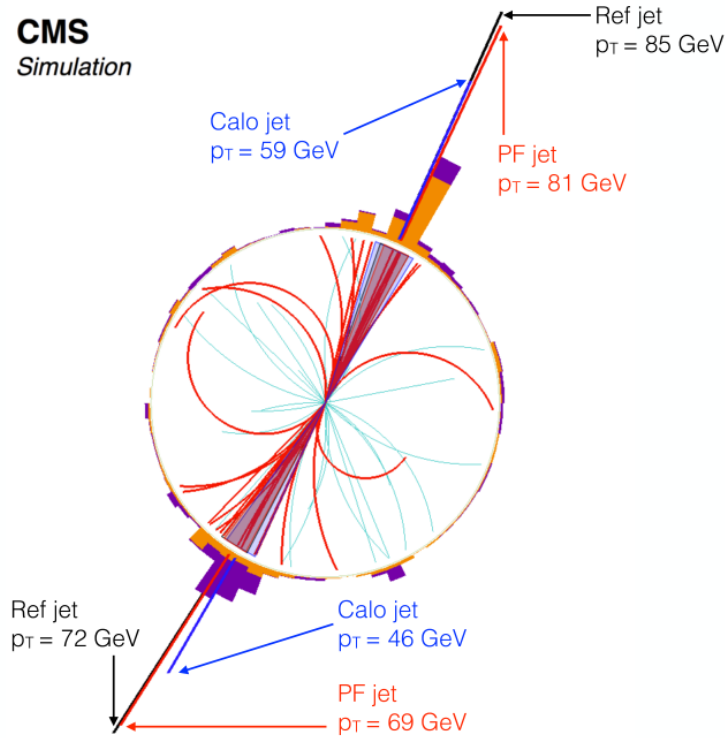


Figure 2-13. Reconstruction of a di-jet in a simulated event without pile-up. The jet reconstruction is performed using calorimeter (Calo) jets, PF jets, and generated (Ref) jets. No jet corrections are applied [94].

The reconstructed jets are calibrated to correct their jet energy scale (JES). Jet energy corrections (JEC) as function of  $p_T$  and  $\eta$  are derived from several dedicated measurements performed using data and simulated events. They address sequentially the impact of pile-up,

uniformity of the detector response, and residual data-simulation differences. Furthermore, the jet energy resolution is measured using di-jet and photon+jet events to derive correction for differences between data and simulation. The energy resolution (JER) of central jets is around 15–20%, 10% and 5% for a jet  $p_T = 30$  GeV, 100 GeV and 1 TeV, respectively [97]. There are jet algorithms developed to reject non-physical jets. Noise jets, originating from calorimeter noise or misreconstruction of electron/muon candidates, are rejected using the PF jet identification criteria, known as PF Jet ID. Moreover, the identification of jets formed from PU particles (PU jets) is performed via the PU ID technique, which permits rejecting those type of jets using a BDT score [98].

Several SM and BSM physics processes have a final state containing heavy-flavor hadronic jets, i.e. jets from b quark (b jets) and c quark hadronization (c jets). Therefore, it is crucial to develop dedicated algorithms for jet flavor assignment or ‘tagging’, to efficiently reconstruct those processes, and reject background associated to light-flavor jets from light-quark and gluon hadronization. Heavy flavor tagging algorithms exploit the physical properties and correlations of the reconstructed jet and its constituents. For instance, one distinctive heavy-flavor jet characteristic is the production of a displaced secondary vertex. As b (c) hadrons (produced during b quark hadronization) have a lifetime of around 1.5 (1) ps, consequently, they have a flight distance a few mm - cm with respect to the primary vertex before decaying into other particles (e.g. charged lepton). The experimental signature is a displaced secondary vertex formed from displaced tracks within the jet. This is illustrated in Figure 2-14.

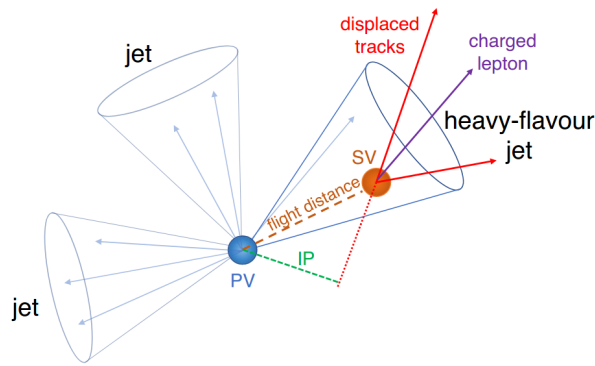


Figure 2-14. Heavy-flavor (b quark or c quark) jet kinematic properties [99].

CMS analysts have used heavy-flavor tagging algorithm in data analysis since Run-1 [100]. For Run-2 data analysis, new and more sophisticated algorithms based on deep learning methods have been developed [99, 101]. The DeepJet algorithm for AK4 jets is the best to date in CMS [102]. It is based on a sophisticated deep neural network with a total of 650 inputs from four categories: PF charged candidates, PF neutral candidates, secondary vertex features, and global event variables. The DeepJet multi-output layer is integrated into a multiclassifier which performs b-tagging, c-tagging, and quark/gluon tagging. The performance of this algorithm surpasses significantly the previous algorithms [76] as illustrated in Figure 2-15. This dissertation uses the DeepJet for b-tagging.

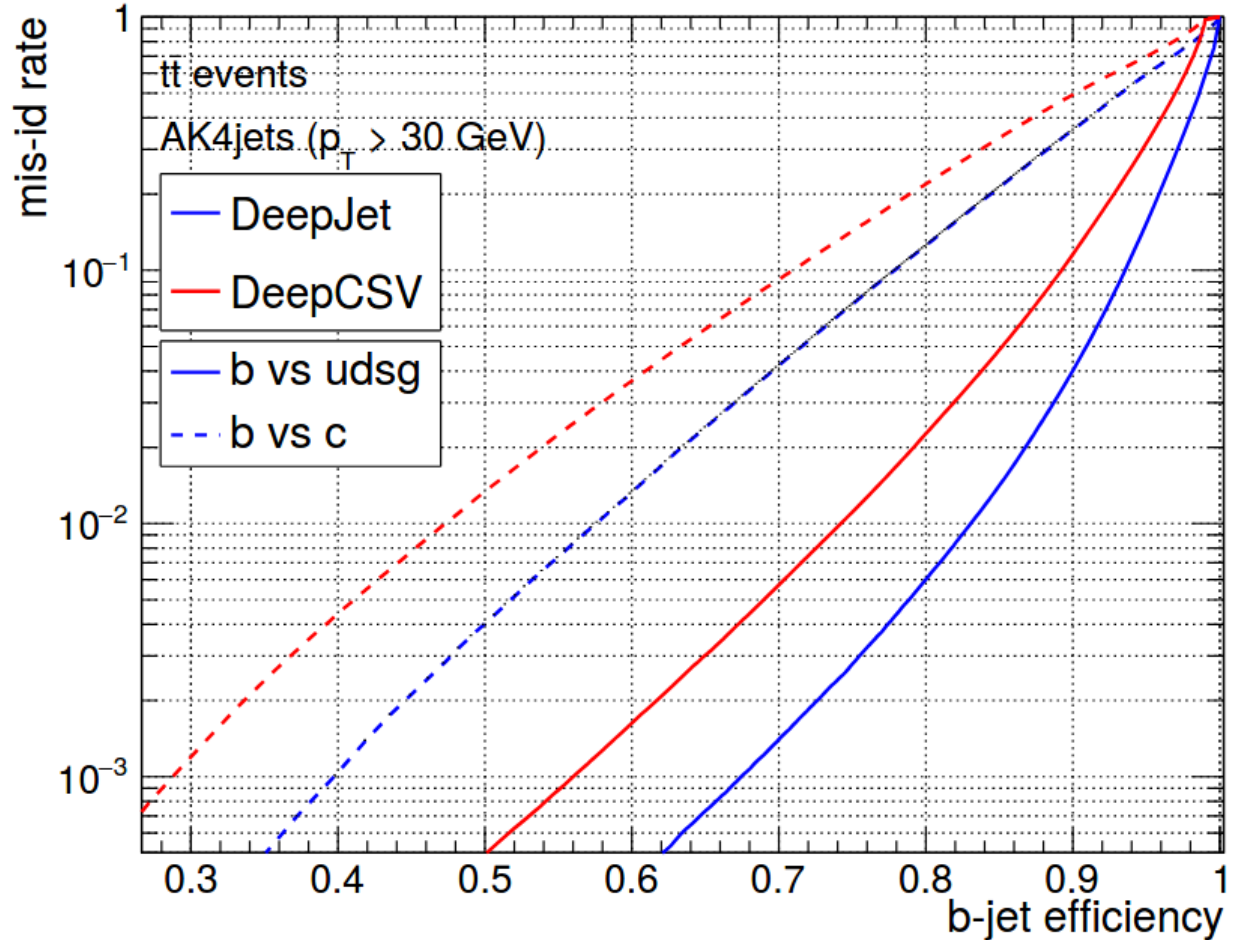


Figure 2-15. Performance of the DeepJet algorithm with respect to the DeepCSV algorithm in simulated hadronic top quark pair ( $t\bar{t}$ ) events [102]. Mis-id rates vs b-tagging efficiency for jets with  $p_T > 30$  GeV.

#### 2.2.3.4 Tau reconstruction

The  $\tau$  leptons are the most massive charged leptons in Nature ( $m_\tau \sim 1.77$  GeV). The  $\tau$ -leptonic decays (e.g.  $\tau^- \rightarrow e^- \bar{\nu}_e \nu_\tau$ ) have a 35.2% branching fraction. Moreover,  $\tau$  leptons are massive enough to have almost all remaining decays involving hadrons (charged hadrons and neutral pions), known as  $\tau$ -hadronic decays ( $\tau_h$ ).

As the two neutrinos escape detection, the  $\tau_l$  decays are identified through the involved electron or muon candidate. On the other hand,  $\tau_h$  decays are identified and reconstructed using the Hadron-Plus-Strip (HPS) algorithm since Run-1 data. This method reconstructs the  $\tau_h$  decay by combining the information from charged hadrons and  $\pi^0$  candidates. The  $\pi^0$  candidate is reconstructed by clustering electrons and photons involved in photon conversions within rectangular  $\eta \times \phi$  regions named ‘strips’. The main  $\tau_h$  backgrounds are quark and gluon jets, electrons, and muons candidates. Isolation requirements and a dedicated MVA-based discriminants are used to separate  $\tau_h$  from jets, electron, and muons. For Run-2 data, the HPS algorithm was modified to have dynamic-size regions to collect more effectively  $\pi^0$  decay products. Studies on  $\tau$  reconstruction using 13 TeV data are presented in Ref. [103]

#### 2.2.3.5 Missing transverse momentum reconstruction

In a collision event, the presence of particles that do not interact with the detector (e.g. neutrinos) is indirectly inferred from an apparent missing transverse momentum ( $\vec{p}_T^{miss}$ ). The precise measurement of  $\vec{p}_T^{miss}$  is crucial for the study of processes involving neutrinos. A method to determine the uncorrected missing transverse momentum ( $\vec{p}_T^{miss,raw}$ ) is defined as the negative sum of the transverse momentum vector of all PF candidates in the event. Some mismeasurement sources are non-linear calorimeter response to hadrons, thresholds on calorimeter energy and track transverse momentum, detector inefficiencies, and pile-up. This is improved by adding the calibrated corrections applied to the PF jets as shown in equation 2-3 [94]. The study of the reconstruction performance of this object at 13 TeV pp collision data is found in Ref.[104].

$$\vec{p}_T^{miss} = \vec{p}_T^{miss,raw} - \sum_{jets} (\vec{p}_{T,jet}^{corr} - \vec{p}_{T,jet}) \quad (2-3)$$

## 2.2.4 Trigger System

The LHC proton beams cross at high collision rates (up to 40 MHz). As the detector data payload from a bunch crossing or ‘event’ is around 1 megabyte (MB), the potential data bandwidth is at the order of 100 Terabytes per second (TB/s). There are not currently the resources to store and process this huge amount of information, and consequently, the event rate has to be reduced by a factor of  $\approx 10^5$ . At the same time, particle physicists know that there is no need to store all events as most of them are low energy multi-jet events and the interesting events are rare (e.g. the Higgs boson production rate is  $\sim 1$  Hz).

The trigger is the online filtering system that reduces the event rate to manageable levels and constitutes the first step of the event selection. It only keeps/accepts events with interesting for the CMS physics analyses, reducing the event rate to around 1 kHz. There are two steps (or levels) in which the rate is reduced: (1) Level-1 (L1), and (2) High Level Trigger (HLT) [86, 88].

The trigger menu is the set of selection criteria used to cover the physics program with the highest selection efficiency under the trigger rate budget. For instance, a selection criterion (sometimes called trigger path) is to require 2 muons of opposite sign and one with  $p_T > 17$  GeV and other  $p_T > 10$  GeV. The L1 and HLT systems have separate menus with typically 300 and 600 trigger paths, respectively.

Events collected by the trigger then send to the Tier-0 at CERN for tape archival, organization and processing. They are split to one or more primary datasets (PDs) based on the HLT path results, and promptly fully reconstructed for offline physics analysis. Examples of PDs used for this analysis or for associated measurements are: SingleMuon, BTagCSV, and JetHT.

### 2.2.4.1 The Level-1 Trigger

The L1 is hardware based subsystem comprising custom designed  $\mu$ TCA boards using field-programmable gate arrays (FPGAs). This allows a flexible and modular design of algorithms for filtering. The L1 does a fast readout of the event information with a coarse granularity using only the information from calorimeters and muon detectors. However, the L1 system capabilities are constrained by the detector readout electronics. It has up to 4  $\mu$ s to make a decision and is

limited to a 100 kHz of output rate. The front-end buffers store the full event waiting for the trigger decision from the L1 global trigger [86, 88].

The current L1 system is not the one used in the LHC Run-1, which was replaced during the year-end technical stop between 2015 and 2016. This upgrade replaced all the hardware, electronic boards, optical links, firmware and software to maintain a high performance for the data-taking condition of the LHC Run-2 and Run-3. The Run-2 performance was excellent and allowed to pursue rich physics program despite the challenges such as increased luminosity, PU, and aging of the subdetectors [105]. The L1 trigger architecture used during the 2016-2018 operations is illustrated in Figure 2-16 and briefly described below:

- **Calorimeter (calo) trigger:** It combines the inputs from ECAL and HCAL into trigger towers with their corresponding calibrated position and energy at the calo layer 1. Then, hadronic jets,  $e/\gamma$  and  $\tau$  candidates are formed at the calo layer 2. In each of the objects pile-up subtraction, isolation and calibrations are applied. Then, global quantities such as  $E_T$  and MET are computed.
- **Muon trigger:** It combines the information from the three muon detectors in regional track finders (or Muon Track Finder Layer), where muon tracks are built from the muon segments with  $p_T, \eta, \phi$  and quality assignment. There are three independent track finders, each covering specific  $\eta$  regions: Barrel Muon Track Finder (BMTF,  $|\eta| < 0.9$ ), Overlap Muon Track Finder (OMTF,  $0.9 < |\eta| < 1.2$ ), and Endcap Muon Track Finder (EMTF,  $1.2 < |\eta| < 2.4$ ). Then, the Global Muon Trigger ( $\mu$ GMT) ranks the track finder muons, looks for duplicates and checks for isolation from calo layer 2 trigger towers. Lastly, the best muons are sent to the Global trigger.
- **Global trigger:** It collects all objects from the calo layer 2 and  $\mu$ GMT. Then, it computes complex properties from those objects and used them in about 500 algorithms called L1 trigger paths. Events satisfying the trigger menu are accepted and sent from the readout buffers to the HLT.

#### 2.2.4.2 The High Level Trigger

The HLT is a software based system which algorithms run on a farm of around 26000 commercial CPUs. The HLT algorithms work with higher precision data relative to L1 ones, because it accesses all the event information to perform complex calculations similar to the ones performed during the offline reconstruction. It further reduces the event rate to around 1 kHz. The data bandwidth output is at the order of gigabytes per second (GB/s). The HLT average latency is about 0.3 seconds [86, 88].

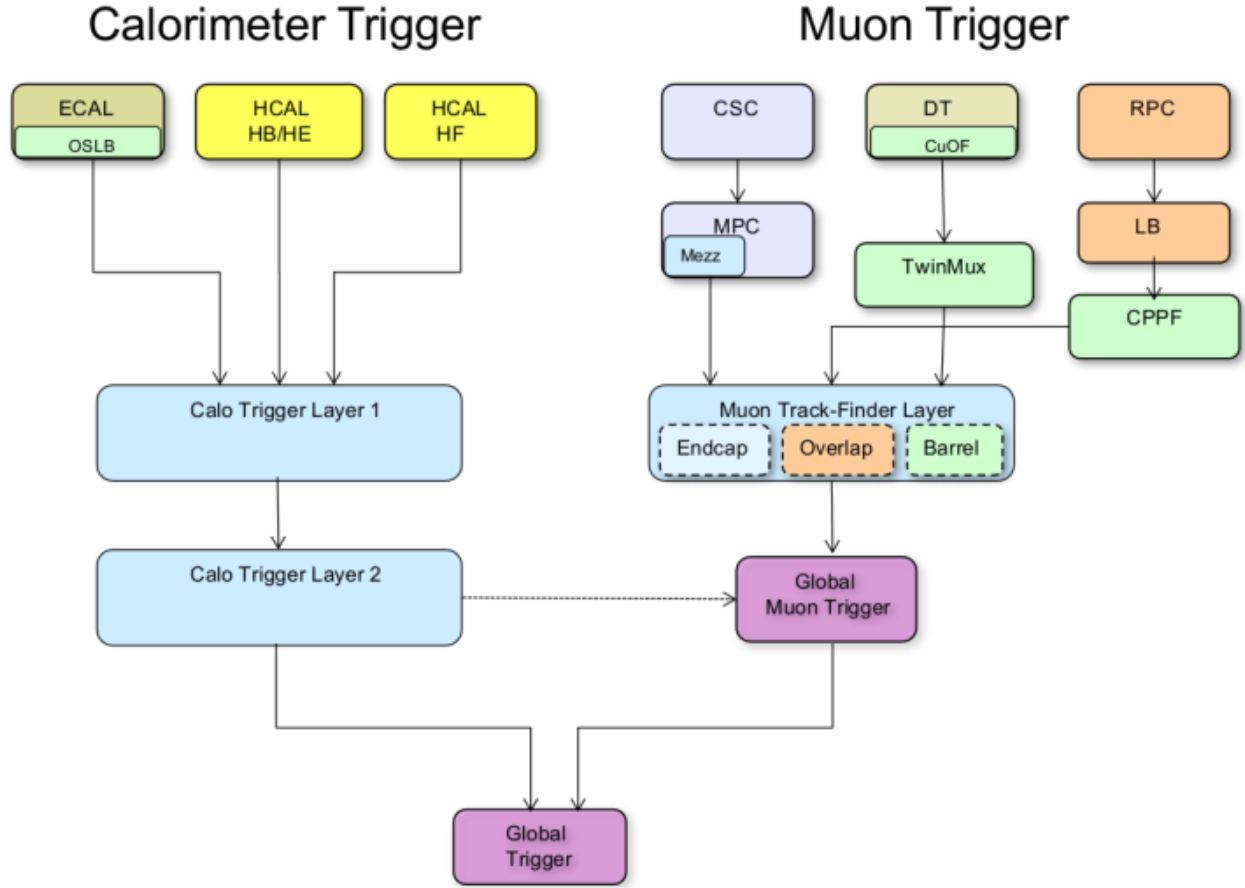


Figure 2-16. The CMS Level-1 trigger architecture during the 2016-2018 operations.

### 2.2.5 Upgrades in Preparation for the HL-LHC

The HL-LHC is the ultimate configuration of the LHC. It is expected to reach an instantaneous luminosity of up to  $7.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ . Consequently, over a decade of operation has the potential to delivery an integrated luminosity of  $4000 \text{ fb}^{-1}$  to each experiment. To maintain a good data-taking performance and extend the physics selectivity, the CMS experiment will upgrade its subdetectors, electronics, trigger and data acquisition systems. The Phase-2 upgrade will allow for the exploration of a broad physics program (e.g. scalar sector characterization) to be pursued.

The upgrade of the CMS subdetectors are planned to take place after the Run-3 operations [106]. The preparation is key to achieve an excellent detector response (high granularity, fast readout, radiation hardness) and event object reconstruction under the high

particle multiplicity and radiation levels at the HL-LHC [107]. These upgrades are summarized in the following points:

- The inner tracking system will be replaced featuring small size pixel sensors and an Outer Tracker consisting of strip and macro pixel sensors. The extended coverage is up to  $|\eta| = 3.8$ . It will allow reconstructing track candidates for the L1 trigger (L1 tracks) with  $|\eta| < 2.4$  [108].
- New readout electronics will be installed in the barrel calorimeters. It will provide higher granularity and timing information [109].
- The endcap calorimeters will be replaced by a new high granularity calorimeter (HGCAL). This new detector will provide shower separation and identification. It will have both timing and finely segmented energy information [110].
- The current detectors of the muon system will be kept, but their electronics will be upgraded to provide more accurate information to the L1 trigger. New muon detectors will be installed to improve the coverage (from  $|\eta|$  up to 2.4 to 2.8) and redundancy in the forward region. The new detectors are additional improved RPC (iRPC) and gas electron multiplier (GEM) chambers [111]. The Phase-2 muon system layout is presented in Figure 2-17.
- A new timing detector (MTD) will be installed in front of the calorimeters to provide timing measurement of charged tracks with a resolution of about 30-45 picoseconds [112].

The CMS experiment will keep a trigger with a two-tier strategy at the HL-LHC. The corresponding Phase-2 upgrade of the L1 and HLT systems is currently under design and testing. The first studies of the Phase-2 L1 system capabilities are presented and summarized in the corresponding technical design report (L1TDR) [113], whereas the HLT TDR is under preparation.

The Phase-2 L1 system will have a new and flexible architecture (presented in Figure 2-18) to adapt to the HL-LHC data-taking challenges. The L1 maximum output rate and latency will be increased to 750 kHz and to 12.5  $\mu$ s, respectively. Furthermore, these new capabilities will allow for the introduction of a new correlator layer, which will combine inputs from the various trigger systems (muon detectors, calorimeters, tracking detectors) to perform a particle-flow like reconstruction of the event at the hardware trigger level. The L1 system will continue to use state-of-the-art FPGAs equipped with 28 Gb/s transceivers and high-speed optical links, following the successful performance of these technologies during Run-2 [105].



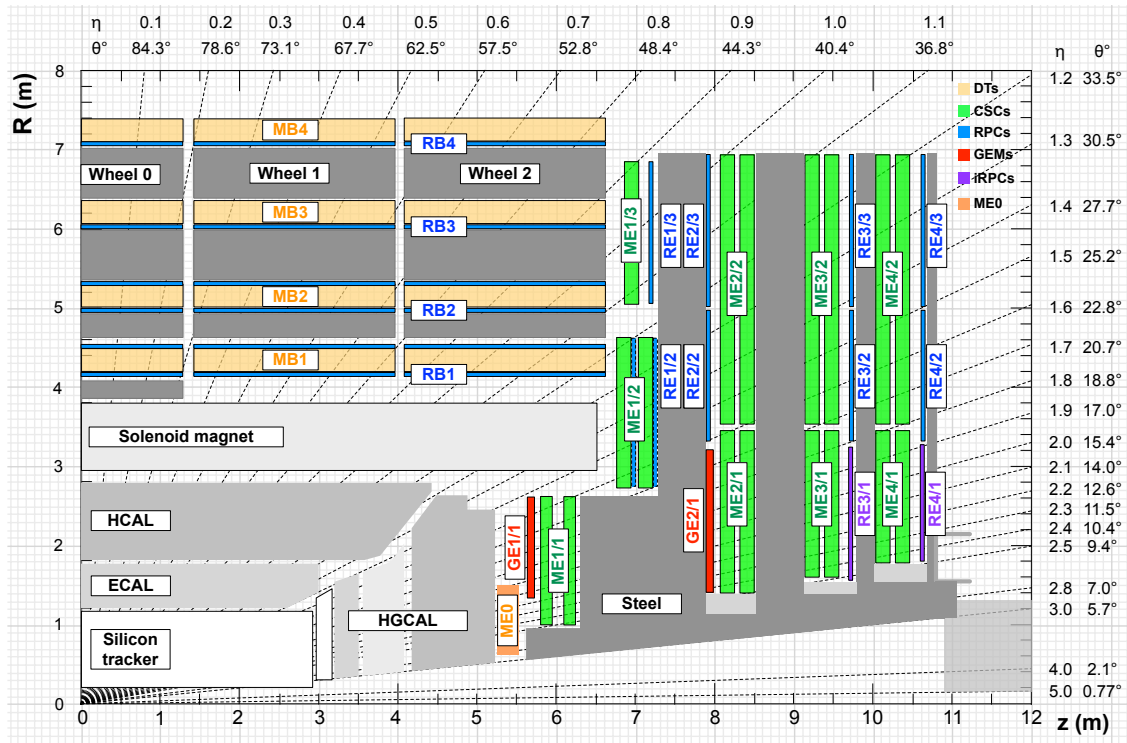


Figure 2-17. The Phase-2 CMS muon system [111].

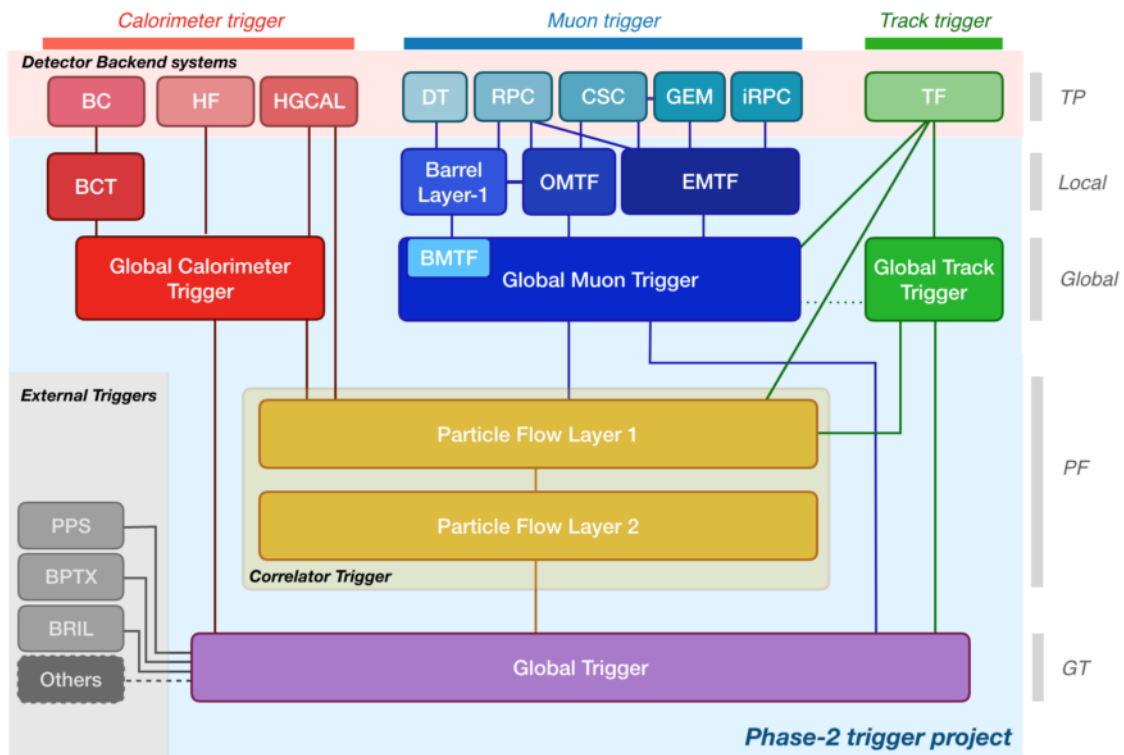


Figure 2-18. The Phase-2 Upgrade of the CMS Level-1 trigger architecture [113].

## CHAPTER 3 ANALYSIS STRATEGY

### 3.1 Challenges and Methods

This dissertation investigates the production of Higgs boson pairs through gluon fusion and vector boson fusion mechanisms at the LHC, focusing on the four b quarks decay channel ( $HH \rightarrow b\bar{b}b\bar{b}$ ). The analysis is performed using the full CMS Run-2 dataset, which is further described in Section 3.2. The signal events must contain two Higgs boson candidates, whose decay into b quark pairs are reconstructed as two jets of 0.4 radius, commonly named as the fully ‘resolved’ Higgs pair topology. It provides the maximal significance for HH production for the SM and the BSM hypotheses with low and medium HH masses (i.e.  $m_{HH} < 1$  TeV). However, the signal identification is very challenging due to the copious production of QCD multi-jet background events. Consequently, several innovative techniques in these areas of object identification, event categorization, signal identification, and background modeling have been included for the maximization of the search sensitivity.

The first set of event preselection requirements is the CMS online trigger system during data-taking, where L1 multijet triggers and HLT triggers select events with four jets and at least three of them identified as b jets, as detailed in Section 3.3. Then, the offline event preselection is applied to find events containing at least four jets identified as b jets (Section 3.4). A dedicated jet pairing method is developed to identify correctly the jet combination that reconstructs each Higgs boson candidate, as described in Subsection 3.4.2. If preselected events contain two additional jets satisfying basic kinematic and VBF topological criteria, these jets are tagged as coming from the hadronization of the outgoing partons in a VBF process or VBF jets.

The data analysis is carried out using the event categories and subcategories presented in Sections 3.5 and 3.7. Events that do not contain VBF jet candidates are classified as ggF events. Events that contain the two VBF jet candidates are further classified as VBF events, or reassigned to the ggF category, using a multivariate Boosted Decision Tree (BDT) score named ggFKiller, that is trained to separate the two main HH production modes. Events in the ggF category are further classified into a low and a high  $m_{HH}$  subcategories, while events in the VBF category are further divided into SM and BSM subcategories based on the output of the ggFKiller score. This

further subcategorization aims at better capturing the changes in the signal kinematics for different hypotheses of anomalous HH couplings production. The event selections are carried out independently but in similar fashion for the 2016, 2017 and 2018 datasets.

One of the biggest challenges of this analysis is the accurate model of the background events, which are expected to be predominantly composed of QCD multi-jet ( $\sim 90\text{-}95\%$ ) and top-quark pair ( $t\bar{t}$ ) ( $\sim 5\text{-}10\%$ ) production. Other minor background processes are single Higgs boson (mostly  $t\bar{t}H(b\bar{b})$  mechanism) and ZZ production. The  $t\bar{t}$ , single H, and ZZ production processes are well-modeled with MC simulations. The simulation of QCD multi-jet is a very complex process that is only simulated at the LO accuracy with the available generators, leading to an inaccurate modeling of final states with multiple heavy-flavor jets. Furthermore, the kinematic phase space enriched with the signal process is very small compared to the inclusive QCD multijet production, resulting in a very low selection efficiency of the simulated QCD events and consequently a low statistics despite millions of generated events. This poses a serious limitation to have an accurate background model. For this reason, an innovative, multidimensional and fully data-driven technique based on a machine learning algorithm is applied to accurately model the background shape and event yields, as presented in Chapter 4.

The kinematic properties, object features and correlations among the various reconstructed objects in the final state (b-jets, VBF-jets, Higgs bosons) are exploited to search for the HH signature in both SM and BSM scenarios. In each ggF subcategory, an MVA discriminant is trained to separate the SM ggF signal from background events, and is used as the observable for signal extraction. In VBF subcategories, the observable for signal extraction is the invariant mass of the reconstructed Higgs boson pair ( $m_{HH}$ ) in the SM subcategory, whereas a counting-experiment is performed in the BSM subcategory. The results for both SM and BSM couplings are reported in Chapter 5. For the statistical analysis, the 2017 and 2018 datasets are merged into a single 2017-2018 data analysis with a common background model and two independent signal contributions and their corresponding uncertainties. The final Run-2 result is obtained by the statistical combination of the results in the 2016 and the 2017-2018 analyses.

### 3.2 Dataset

The results presented in this dissertation are derived using a dataset of 13 TeV pp collisions, collected by the CMS experiment during the LHC operation in 2016, 2017 and 2018. This is known as the ‘full Run-2’ dataset. Each of the datasets is divided in run periods (denoted by a letter) consisting of set of collision runs (run range), as presented in Table 3-1. The primary datasets used in 2016, 2017 and 2018 are BTagCSV, BTagCSV and JetHT, respectively. They contain events passing the HLT trigger paths used in the data analysis, discussed Section 3.3.

The 2016, 2017 and 2018 datasets are certified independently through a data quality evaluation performed by a team of experts in detector and physics objects. The lumisection is a subsection of a run in which the instantaneous luminosity is approximately constant (approximately 23 seconds). Experts select only the lumisections of data which are considered good for physics analysis. The certified total integrated luminosity and its uncertainty per dataset is presented in Table 3-2. The Run-2 dataset has a total integrated luminosity of around  $138 \text{ fb}^{-1}$

Table 3-1. Summary details of the datasets used for analysis corresponding to the 2016 (top block), 2017 (center block), and 2018 (bottom block) proton-proton collisions, of the corresponding run range, and integrated luminosity ( $L_{\text{int}}$ ).

| Dataset | Primary dataset | Run Period | Run range     | $L_{\text{int}} [\text{fb}^{-1}]$ |
|---------|-----------------|------------|---------------|-----------------------------------|
| 2016    | BTagCSV         | B          | 273150–275376 | 5.82                              |
|         | BTagCSV         | C          | 275656–276283 | 2.61                              |
|         | BTagCSV         | D          | 276315–276811 | 4.29                              |
|         | BTagCSV         | E          | 276831–277420 | 4.07                              |
|         | BTagCSV         | F          | 277932–278808 | 3.14                              |
|         | BTagCSV         | G          | 278820–280385 | 7.65                              |
|         | BTagCSV         | H          | 281613–284044 | 8.74                              |
|         | BTagCSV         |            |               |                                   |
| 2017    | BTagCSV         | B          | 297047–299329 | 4.79                              |
|         | BTagCSV         | C          | 299368–302029 | 9.63                              |
|         | BTagCSV         | D          | 302031–302663 | 4.25                              |
|         | BTagCSV         | E          | 303824–304797 | 9.31                              |
|         | BTagCSV         | F          | 305040–306460 | 13.54                             |
|         | BTagCSV         |            |               |                                   |
| 2018    | BTagCSV         | A          | 315257–316995 | 14.03                             |
|         | BTagCSV         | B          | 317080–319310 | 7.07                              |
|         | BTagCSV         | C          | 319337–320065 | 6.90                              |
|         | BTagCSV         | D          | 320500–325175 | 31.75                             |

Table 3-2. The full Run-2 dataset per year and its integrated luminosity and uncertainty. The integrated luminosity ( $L_{\text{int}}$ ) corresponds to the data certified as good for physics and used in the analysis.

| Year  | $L_{\text{int}}$ [ $\text{fb}^{-1}$ ] | Uncertainty |
|-------|---------------------------------------|-------------|
| 2016  | 36.33                                 | 1.2%        |
| 2017  | 41.53                                 | 2.3%        |
| 2018  | 59.74                                 | 2.5%        |
| Total | 137.60                                | 1.6%        |

### 3.3 Trigger

#### 3.3.1 Requirements

The events of interest are characterized by four jets arising from the b-quark hadronization. At the LHC, QCD multijet events are produced approximately  $10^6$  times more often than the signal events. Consequently, the trigger strategy is to start the multijet background rejection at trigger level using HLT trigger paths featuring an online b-tagging algorithm (CSV in 2016, and DeepCSV from 2017 to 2018). Table 3-3 lists the trigger paths chosen and found optimal for this analysis. They were active and unprescaled during the Run 2 data-taking period. Note that the trigger paths consists of a list of filters applied in sequence during data-taking. The list of filters can be found in Appendix A. The analysis HLT paths require three jets tagged by the

Table 3-3. HLT Trigger Paths used in the analysis by year. The trigger efficiency for SM signals produced via gluon and vector boson fusion are also presented.

| Dataset | HLT Trigger Path Name                                                 |
|---------|-----------------------------------------------------------------------|
| 2016    | HLT_QuadJet45_TripleBTagCSV<br>HLT_DoubleJet90_Double30_TripleBTagCSV |
| 2017    | HLT_PFHT300PT30_QuadPFJet_75_60_45_40_TriplePFBTagCSV                 |
| 2018    | HLT_PFHT330PT30_QuadPFJet_75_60_45_40_TriplePFBTagDeepCSV             |

corresponding b-tagging algorithm. For 2016, a logical-OR of two triggers is used, with one path requiring four jets with a  $p_T$  threshold of 45 GeV, and the other requiring two jets above 30 GeV and two jets above 90 GeV. For 2017 and 2018, a single path is used with four  $p_T$ -thresholds of 40, 45, 60, and 75 GeV and a minimal  $H_T$  requirement of 300 (330) GeV for 2017 (2018). The same trigger paths are modeled and applied to the simulated events. The efficiency for SM  $ggF$  HH signal in 2016, 2017 and 2018 is 35%, 21% and 22%, respectively.

### 3.3.2 Efficiency Measurement

The difference in trigger performance between data and MC simulation is studied through the measurement of trigger efficiency in each year. This is carried out in an unbiased sample of events passing a single muon trigger (HLT\_IsoMu24 trigger). The samples used for the efficiency measurement are the SingleMuon for data and top-quark pair production ( $t\bar{t}$ ) MC simulated events (inclusive for 2016, full leptonic samples for 2017 and 2018) for MC. A dedicated set of preselection requirements are applied to find a sample of events enriched with  $t\bar{t}$  production in the electron-muon channel. These events are chosen for the measurement due to the relative high statistics and the presence of two real b-quarks for the b-tagging online efficiency evaluation.

The trigger efficiency in the selected data and simulated events is evaluated by measuring the efficiency of each of the trigger filters associated to the trigger paths over a sample of events that have passed all the previous filters of the path. Some examples of the measured trigger efficiencies of three filters corresponding to the full Run-2 triggers are illustrated in Figure 3-1. As it can be seen from the filter efficiency distributions, the simulation of the trigger is not perfect in simulated events, and thus it has to be corrected in the analysis. Finally, filter efficiency fits associated to the data and  $t\bar{t}$  simulation measurement are obtained by fitting them with dedicated functions. More details on the efficiency measurement and fits is in Appendix A.

Any trigger efficiency differences of the simulation with respect to the data are addressed through event-by-event correction scale factors (SFs) based on the aforementioned measurements. First, the total efficiencies are evaluated event by event as the product of the filter efficiencies obtained evaluating the data and  $t\bar{t}$  MC simulation fits. Then, the event-by-event SF factor is computed as the ratio between the total data efficiency (from data fits) and the total MC efficiency (from  $t\bar{t}$  MC fits). The fit uncertainties are propagated through the SF calculator to provide systematic variations of the SFs, which are used to compute the trigger systematic uncertainties. The trigger SFs derived for signal events are illustrated as a function of  $m_{HH}$  in Figure 3-2.

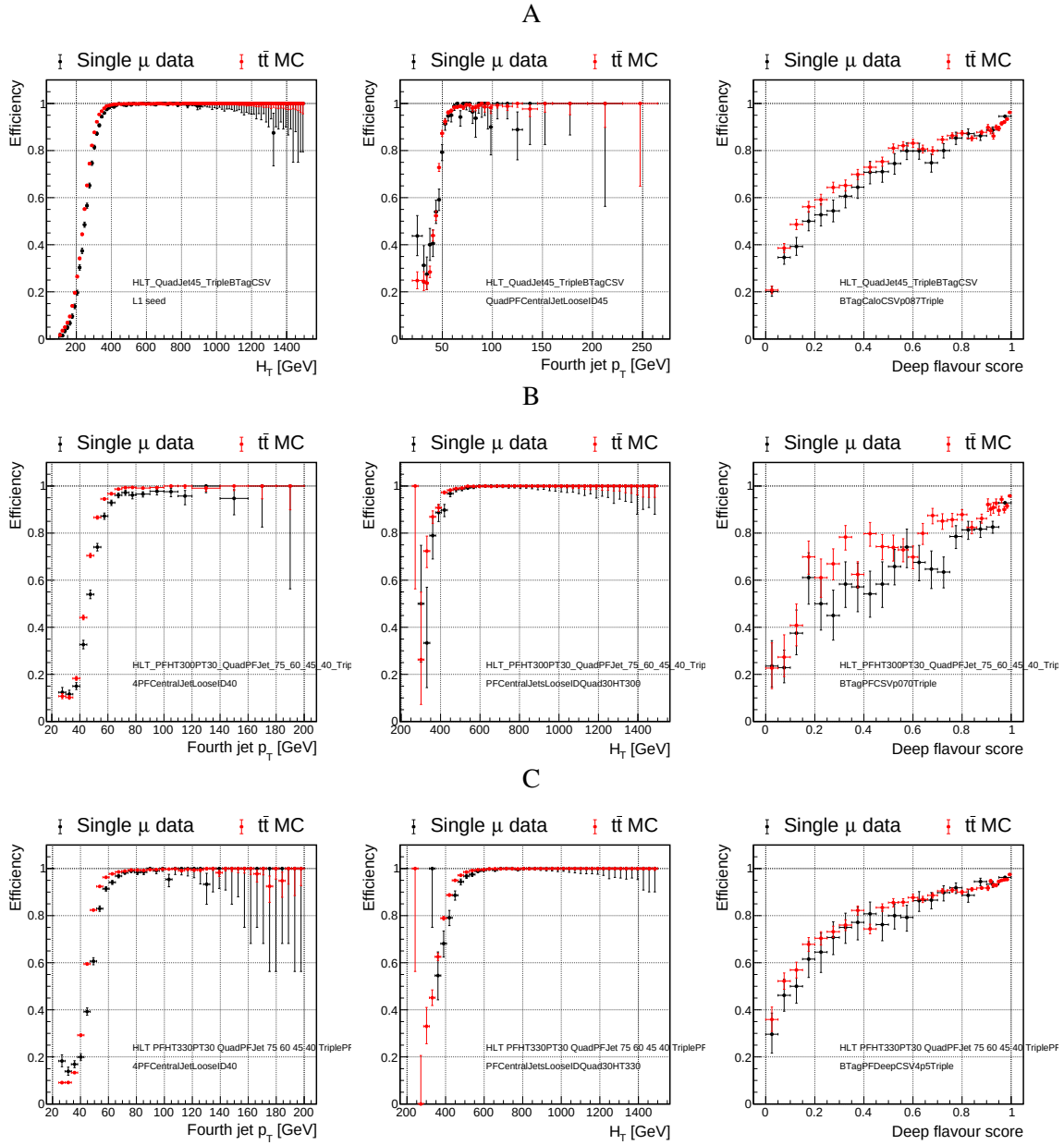


Figure 3-1. Examples of the measured trigger filter efficiency in single muon ( $\mu$ ) data (black) and top-quark pair ( $t\bar{t}$ ) MC simulation (red) events. A) HLT\_QuadJet45\_TripleBTagCSV filters used in 2016, B) HLT\_PFHT300PT30\_QuadPFJet\_75\_60\_45\_40\_TriplePFBTagCSV filters used in 2017, C) HLT\_PFHT330PT30\_QuadPFJet\_75\_60\_45\_40\_TriplePFBTagDeepCSV filters used in 2018.

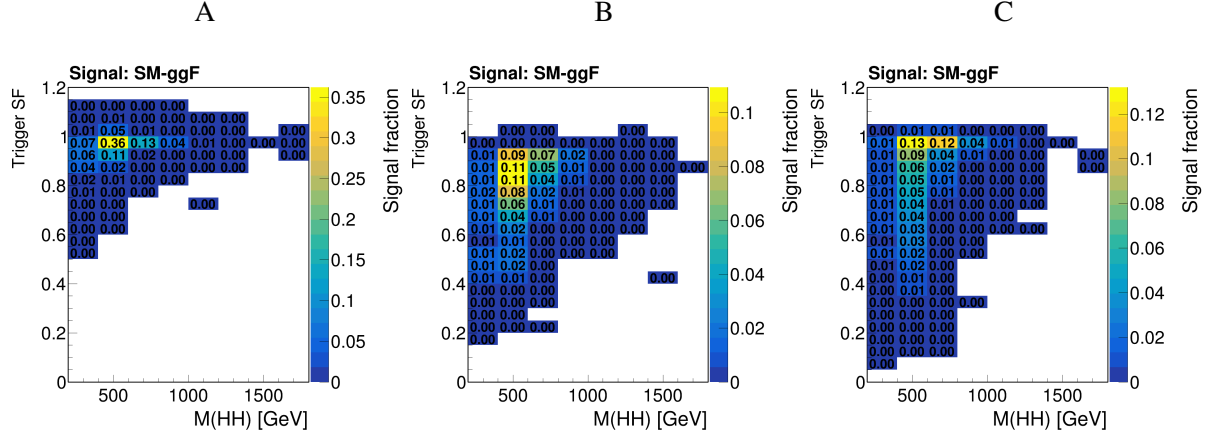


Figure 3-2. Fraction of SM ggF signal events (z-axis) that are attributed a specific value of the trigger scale factor (y-axis) as a function of Higgs boson pair mass (x-axis). A) 2016 trigger, B) 2017 trigger and C) 2018 trigger.

The validation of the trigger MC efficiency measurement is carried out by a closure test performed in  $t\bar{t}$  and signal simulated events passing the event preselection requirements presented in Section 3.4. The test consists of comparing the total MC efficiency obtained from the MC fits, and the one obtained from the simulated trigger bits. A very good agreement was found between the two estimations within the uncertainties in  $t\bar{t}$  and signal closure tests. This validates the trigger MC measurement and also by proxy the usage of the trigger SFs. The closure test for SM ggF signal events is illustrated in Figure 3-3.

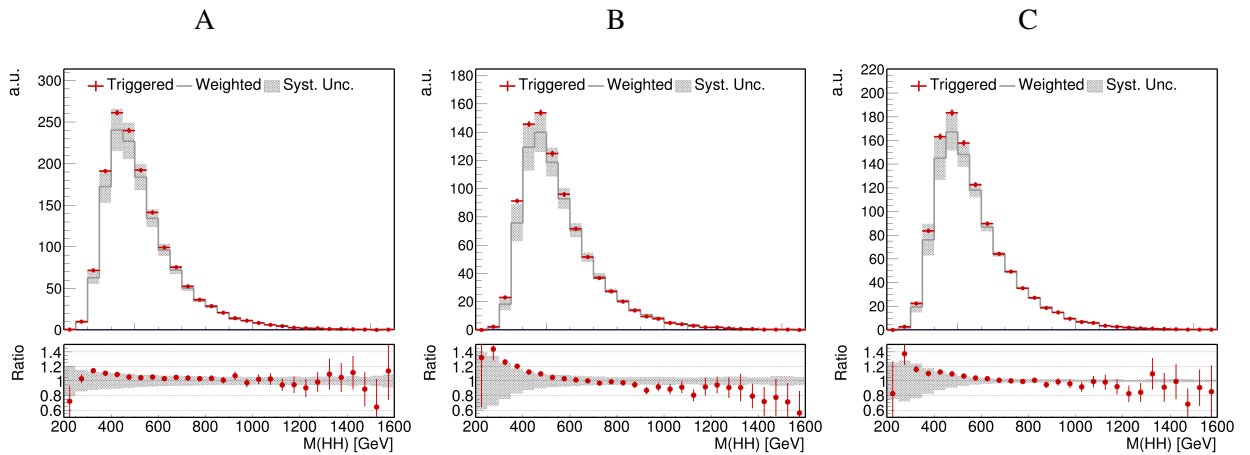


Figure 3-3. Closure test of the trigger simulated efficiency in the SM ggF signal events. A) 2016 trigger, B) 2017 trigger, C) 2018 trigger.



### 3.4 Event Preselection

The ‘offline’ event preselection presented in this section is the following step after applying the online requirements of the triggers listed in Section 3.3. The preselection starts by vetoing events containing isolated muons or electrons satisfying the requirements in Subsection 3.4.1 are excluded from further analysis. Then, jet preselection criteria in Subsections 3.4.2 and 3.4.3 aim at reconstructing the Higgs boson pair (HH) candidate and the VBF-jet pair (jj) candidate (if existent) from the collection of reconstructed jets in the event.

#### 3.4.1 Isolated Lepton Veto

The signal events of interest do not contain isolated leptons (muons and electrons). Therefore, we require the absence of isolated muons and electron in the event. This veto mainly rejects top-quark production events with one or two isolated leptons in the final state. Furthermore, it makes this analysis orthogonal with respect to searches for HH in other final states containing isolated leptons (muons or electrons) in the final state, e.g. the  $HH \rightarrow b\bar{b}W^+W^-$  channel. The criteria used per year to identify isolated muons and electrons is listed in Table 3-4.

Table 3-4. Summary of requirements for isolated electrons and muons.

| Requirement         | Isolated Electron | Isolated Muon |
|---------------------|-------------------|---------------|
| Min. $p_T$ [GeV]    | 15                | 10            |
| Max. $ \eta $       | 2.4               | 2.4           |
| Particle Id         | Loose             | Loose         |
| Max. PFIso          | 0.15              | 0.15          |
| Max. BarrelDxy [cm] | 0.05              | 0.05          |
| Max. BarrelDz [cm]  | 0.10              | 0.10          |
| Max. EndcapDxy [cm] | 0.10              | 0.10          |
| Max. EndcapDz [cm]  | 0.20              | 0.20          |

#### 3.4.2 Higgs Boson Pair Candidate

To identify the four b jet candidate from the Higgs boson pair, the jet collection is filtered event-by-event using the requirements listed in Table 3-5. The high-performing DeepJet b-tagging algorithm is used to identify b-jets using the medium working point, corresponding approximately 75% b-tagging efficiency and 1% mis-id rate. Then, the four jets with the highest DeepJet score are chosen as the four b-jet candidates in the event. The usage of the DeepJet

algorithm improves the sensitivity to SM HH production by about 16-18% compared to the previous algorithms (e.g. DeepCSV). In addition, the four selected b-jet candidates are required to match the trigger objects that ‘fired’ the trigger based on the angular distance  $\Delta R < 0.5$ .

Table 3-5. Summary of jet properties required for b-jet candidates. The pile-up identification (PUID) requirement is only applied to jets with  $p_T < 50\text{GeV}$ .

| Dataset | $p_T$ [GeV] | $ \eta $ | PUID   | JetID | DeepJet |
|---------|-------------|----------|--------|-------|---------|
| 2016    | $> 30$      | $< 2.4$  | Medium | Tight | Medium  |
| 2017    | $> 40$      | $< 2.5$  | Medium | Tight | Medium  |
| 2018    | $> 40$      | $< 2.5$  | Medium | Tight | Medium  |

A B-hadron decay into a final state including a neutrino is around 35% of the time. Then, this neutrino results in a mismeasurement of the b jet energy during the PF reconstruction, and produces a worsening of the resolution of the reconstructed Higgs boson di-jet mass ( $m_{b\bar{b}}$ ), which is a crucial variable for signal and background separation. CMS experts developed a technique that helps to correct the b jet energy and by proxy improves the  $m_{b\bar{b}}$  resolution. This method is based on a multivariate regression based on deep neural networks (DNNs) targeting the generator level b-quark  $p_T$  [80] using real b jets from top-quark pair events. The DNN output provides b jets with a correction factor to the jet  $p_T$  and energy, and an estimation of the jet energy resolution. The correction factor is applied to the four selected b-jet candidates in the event.

One key part of the signal identification is the reconstruction of the two Higgs boson candidates. As there are four b jet candidates (say b1, b2, b3, and b4 without specific ordering), there are three ways to pair them into two di-jets or Higgs candidates:  $A = [H(b1, b2) ; H(b3, b4)]$ ,  $B = [H(b1, b3) ; H(b2, b4)]$  and  $C = [H(b1, b3) ; H(b2, b4)]$ . The pairing method should avoid the accumulation of background events in phase space regions with high signal concentration, e.g. near the Higgs boson mass ( $m_H = 125\text{ GeV}$ ). For instance, a method for selecting the pairings with masses closest to the Higgs boson mass ( $\sim 125\text{ GeV}$ ) is not optimal because increases the background population near the largest population of the signal events (the mass peak).

A pairing method is developed to find the correct pairing that reconstruct the two Higgs boson candidates without any significant background sculpting near the signal-enriched region

centered at the reconstructed Higgs boson masses. The pairing method was studied in preselected 4b events matched at generator level using the following four signal benchmarks: SM ggF, SM VBF, BSM ggF ( $\kappa_\lambda = 5$ ) and BSM VBF ( $\kappa_{2V} = 2$ ).

As the first step of the pairing method, in each of the three pairing possibilities ( $i=A, B, C$ ), we order the two reconstructed Higgs boson candidates, so that  $H_1^i$  is the highest- $p_T$  candidate and  $H_2^i$  the lowest  $p_T$  candidate. Note that each of the pairing possibilities can be represented as a point in the plane defined by the reconstructed mass of  $H_1$  ( $m_{H_1}$ ) and of  $H_2$  ( $m_{H_2}$ ) or  $m_{H_1} - m_{H_2}$  plane, as illustrated in the Figure 3-4 sketch.

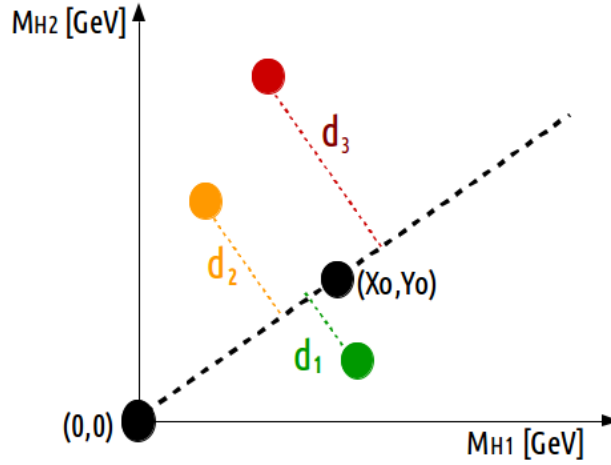


Figure 3-4. Sketch of the pairing method using the two-dimensional plane of the two reconstructed Higgs boson masses ( $m$ ), where  $H_1$  is leading- $p_T$  di-jet and  $H_2$  is subleading- $p_T$  di-jet.

A possible good pairing method is the ‘closest to the diagonal’ or  $d_1$ -pairing method. In this method, the selected pairing is the one that minimizes the distance from its associated point to the diagonal line passing through the points  $(0,0)$  and  $(X_0, Y_0)$  using equation 3-1 (where  $k = X_0/Y_0$ ) as illustrated with the green point and  $d_1$  distance in Figure 3-4. By defining  $k$  equal to the ratio of the reconstructed Higgs boson average peak positions ( $k = 125/120 \sim 1.04$ ), the  $d_1$ -pairing success for the signal benchmarks is the following: 83% for SM VBF, 89% for BSM VBF, 82% SM ggF and 76% for BSM ggF.

$$d = \frac{|m_{H_1} - k m_{H_2}|}{\sqrt{1 + k^2}} \quad (3-1)$$

The correct pairing properties were studied further in the signal benchmarks in order to improve the  $d_1$ -based method. The findings are summarized as follows: (1) When the  $d_1$ -pairing method fails, the correct pairing mostly corresponds to the one with second-closest distance from the associated point to the diagonal line or the  $d_2$ -pairing (orange point in Figure 3-4), (2) the  $d_1$ -pairing method begins to fail when  $|d_1 - d_2| < 30\text{GeV}$ , (3) when the  $d_1$ -pairing method fails, the reconstructed Higgs transverse momentum in the center of mass reference frame ( $p_T^*(H)$ ) associated to the  $d_2$ -pairing tends to be larger than the one from the  $d_1$ -method.

From this aforementioned findings, a better pairing algorithm is developed and described as follows: (1) If  $|d_1 - d_2| > 30\text{ GeV}$ , the pairing is obtained from the  $d_1$ -method (2) otherwise, the pairing is obtained from the method ( $d_1$  or  $d_2$ ) with the largest  $p_T^*(H)$ . Figure 3-5 illustrates the distributions of the reconstructed  $H_2$  and  $H_1$  invariant mass in the SM ggF signal simulation without (blue) and with (red) the application of the b-jet energy regression. The improved pairing success is the following: 91% for SM VBF, 98% for BSM VBF, 96% SM ggF and 82% BSM ggF.

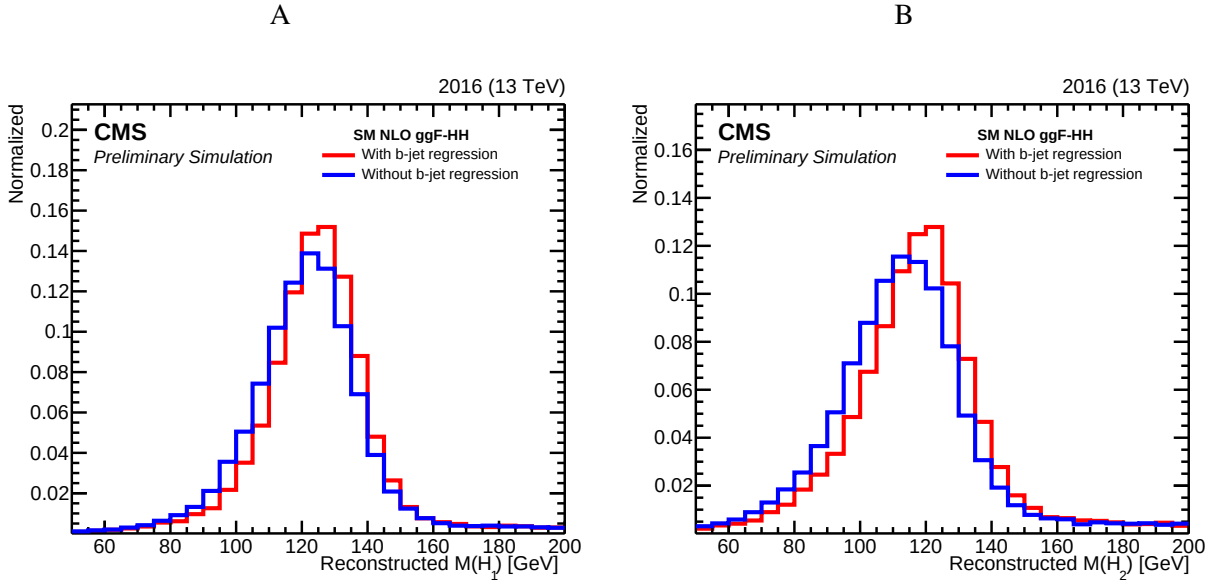


Figure 3-5. Invariant mass of the two reconstructed Higgs bosons in SM ggF HH simulated events of the 2016 data-taking conditions: A) Highest- $p_T$  di-jet ( $H_1$ ), B) Lowest- $p_T$  di-jet ( $H_2$ ). The distributions are illustrated without (blue) and with (red) the application of the b-jet energy regression.

### 3.4.3 Vector Boson Fusion Jet Pair Candidate

To identify whether the HH events are coming from a ggF or a VBF production process, we examine the remaining jets in the collection that have not been selected for the HH reconstruction step to search for the two VBF-jet candidates. The first step is to apply the jet kinematic and quality requirements listed in Table 3-6 to obtain the list of VBF jet candidates.

It is important to mention that during 2017 data collection was observed a large ECAL Endcap (EE) noise in the forward region. This EE noise increased the jet multiplicity with ‘noise-jets’ characterized by a raw  $p_T < 50$  GeV and  $2.6 < |\eta| < 3.1$ . In the 2017 preselection, we require the tightest working point of the PUID algorithm on jets located at that  $\eta$  region.

Table 3-6. Summary of jet kinematic and quality requirements for VBF-jet candidates. The PileupID requirement is only applied to jets with  $p_T < 50$  GeV.

| Dataset | $p_T$ [GeV] | $ \eta $ | PileupID | JetID |
|---------|-------------|----------|----------|-------|
| 2016    | $> 25$      | $< 4.7$  | Medium   | Tight |
| 2017    | $> 25$      | $< 4.7$  | Medium   | Tight |
| 2018    | $> 25$      | $< 4.7$  | Medium   | Tight |

Simulation studies show that preselected VBF signal events with at least two VBF jet candidates (where the four b quarks from the HH decay and the two outgoing partons from the VBF process are all reconstructed as six jets) have the following properties: (1) The VBF outgoing partons ( $q_1$  and  $q_2$ ) have opposite  $\eta$ -sign ( $\eta_{q1} \times \eta_{q2} < 0$ ) in 94-98% of the cases depending on  $\kappa_{2V}$  coupling, and (2) the highest- $p_T$  VBF-jet candidate is one of the real VBF jets in 93-94% of the cases. Based on this information, the VBF-jet candidate with the highest- $p_T$  is chosen as the highest- $p_T$  VBF jet ( $j_1$ ). Then, we check all the opposite  $\eta$ -sign VBF-jet candidates to select the lowest- $p_T$  VBF-jet ( $j_2$ ). If multiple jets with opposite  $\eta$ -sign jet are found (around 30% of the time), then the one with the largest  $p_T$  is chosen to be  $j_2$ . It is important to mention here that we tested different ways to make the  $j_2$ -choice: largest  $p_T$  (nominal), largest VBF pair pseudorapidity separation, largest VBF pair mass. We found that the three options have similar VBF purity as seen in Figure 3-6 for VBF signal with  $\kappa_{2V} = 2$ . However, we expect that nominal choice to have better signal and background separation because it would not sculpt the distributions of  $m_{j_1 j_2}$  and  $\Delta\eta(j_1, j_2)$  at high values (signal-enriched regions).

Preselected events with a fully reconstructed VBF-jet pair are classified to the ‘Pre-VBF’ region and denominated as ‘pre-VBF’ events, otherwise are denominated classified to the ‘Pre-ggF’ region and denominated as ‘pre-ggF’ events. The event yields in data and MC simulated events is presented in Table 3-7.

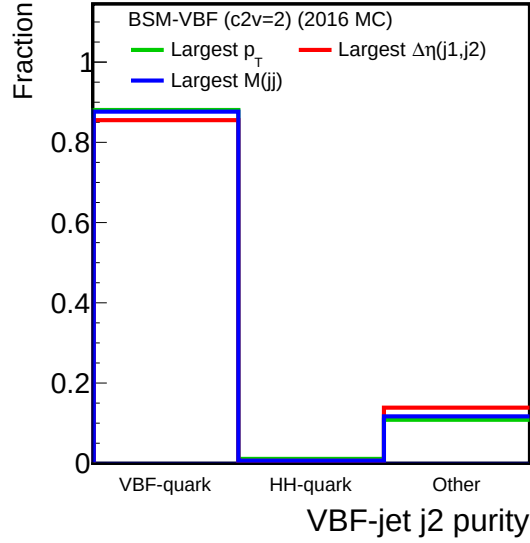


Figure 3-6. Purity of the lowest transverse momentum VBF jet candidate (j2) using three different methods in events with multiple opposite  $\eta$ -sign jets: (green) largest  $p_T$ , (red) largest VBF pair pseudorapidity separation, (blue) largest VBF pair mass. The simulated events correspond to VBF signal with  $\kappa_{2V} = 2$  in the 2016 data-taking conditions.

Table 3-7. Event yields of data and MC simulation after preselection requirements. The processes ggF, VBF and VBF( $\kappa_{2V} = 2$ ) are multiplied by a factor of  $10^2$ .

| Dataset/Presel.           | 2016/ggF | 2016/VBF | 2017/ggF | 2017/VBF | 2018/ggF | 2018/VBF |
|---------------------------|----------|----------|----------|----------|----------|----------|
| ggF                       | 1174.3   | 443.0    | 763.9    | 270.6    | 1338.9   | 527.1    |
| VBF                       | 8.8      | 18.4     | 6.6      | 12.2     | 11.8     | 25.1     |
| VBF ( $\kappa_{2V} = 2$ ) | 607.7    | 976.3    | 444.9    | 580.3    | 910.7    | 1300.8   |
| QCD multijet              | 121147.7 | 52818.2  | 52149.2  | 25017.0  | 86381.0  | 39741.9  |
| $t\bar{t}$                | 4614.0   | 6761.0   | 2611.5   | 3973.1   | 5369.2   | 8648.7   |
| Single Higgs              | 296.4    | 441.7    | 208.1    | 302.1    | 392.4    | 601.4    |
| $ZZ \rightarrow 4b$       | 30.5     | 7.1      | 17.3     | 4.7      | 29.9     | 8.7      |
| Total MC Bkg.             | 126088.6 | 60028.0  | 54986.1  | 29296.9  | 92172.5  | 49000.7  |
| Data                      | 153170   | 69702    | 110500   | 57594    | 164307   | 96274    |

## 3.5 Event Categorization

### 3.5.1 Motivation

The HH search aims at having production mode (ggF and VBF) event categories with the highest signal purity possible to maximize the measurement of both the ggF and VBF signals. As observed in Table 3-7, there is about 26-28% of the SM ggF signal events are classified in the pre-VBF event region. Furthermore, the ratio between the SM ggF and SM VBF signals in Pre-VBF events is approximately 21.0-24.1, and the ratio between SM ggF and VBF ( $\kappa_{2V}=2$ ) signals in pre-VBF events is 0.41-0.47. Consequently, pre-ggF and pre-VBF regions do not have a good production mode purity. As a next step, the strategy is to use a production mode discriminant or score, to reduce ggF signal assigned to the pre-VBF event region and re-assign it back to the ggF signal-enriched region, thus creating higher purity ggF and VBF categories.

### 3.5.2 Production Mode Score: ggFKiller

The production mode score discussed above is named ‘ggFKiller’. It is trained and used in the kinematic phase space of the pre-VBF region. The ggFKiller aims to maximize the following: (1) the efficiency of the assignment of SM ggF signal events to the ggF signal-enriched region, (2) the sensitivity to measure both SM VBF signal and VBF signal with anomalous  $\kappa_{2V}$  couplings.

The chosen multivariate analysis (MVA) algorithm is a Boosted Decision Tree (BDT) implemented in the XGBoost package [114]. The VBF signal with the  $\kappa_{2V} = 2$  and SM ggF signal models are defined as the ‘signal’ events and ‘background’ events in the training, respectively. The specific choice of  $\kappa_{2V}$  is motivated by the largest BSM kinematic enhancement among the available VBF signal simulated samples. The ggFKiller is built using 14 kinematic variables and correlations among the reconstructed Higgs bosons and the VBF-jets. The list of variables is presented in Table 3-8. Three independent trainings are performed depending on the simulation year. For illustration, the signal (VBF with  $\kappa_{2V} = 2$ ) and background (SM ggF) input variable distributions of the 2016 simulated data-taking conditions are presented in Figure 3-7. More information regarding the ggFKiller training and optimization is presented in Appendix B.

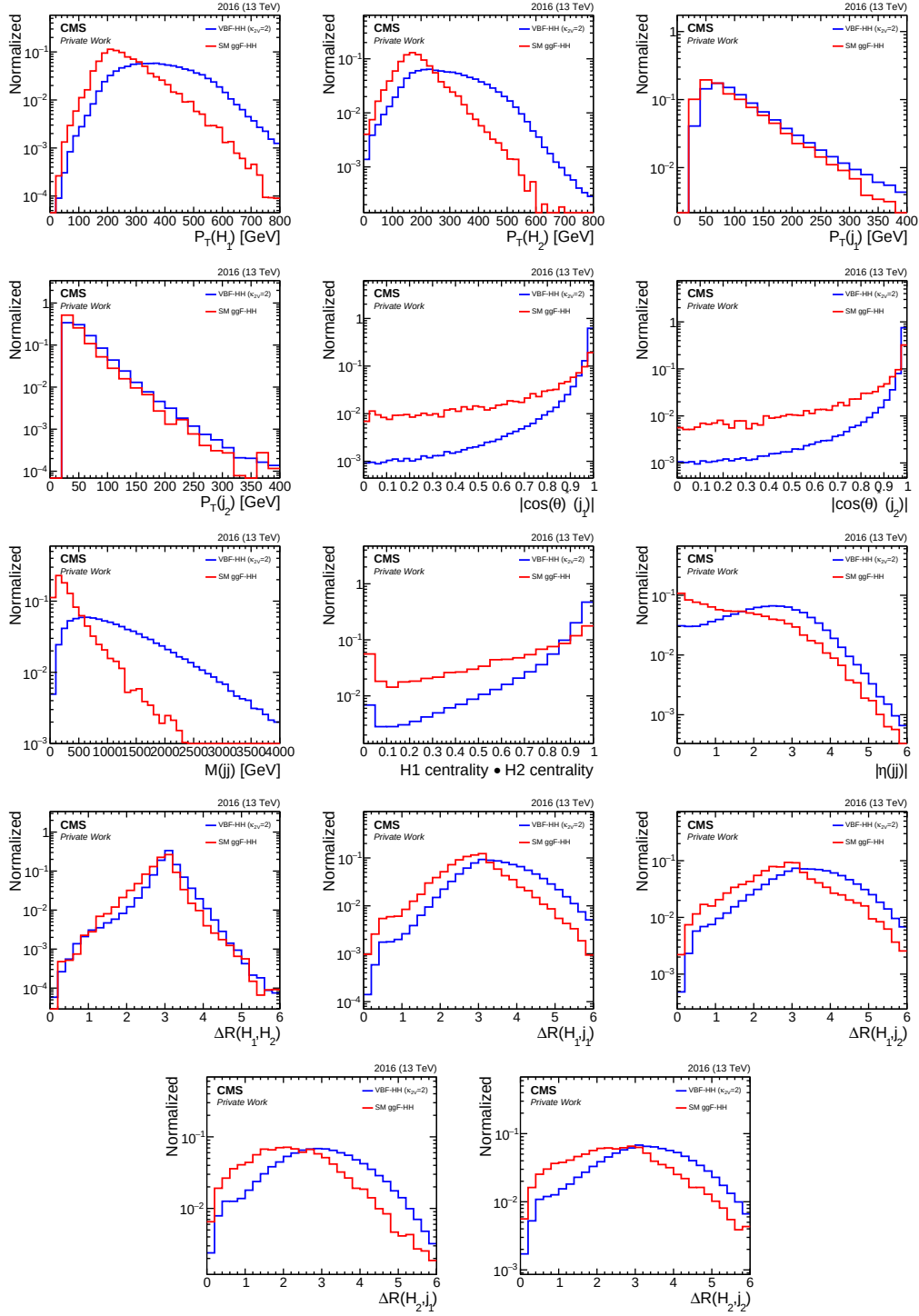


Figure 3-7. Distribution of the ggFKiller input variables in 2016 simulation. The VBF with  $\kappa_{2V} = 2$  (signal) is in solid blue, whereas SM ggF (background) is in solid red.



Table 3-8. Variables used to train the ggFKiller score. The product of Higgs centralities is defined as  $\exp[-(\frac{\eta(H_1)-\eta_{avg}}{\Delta\eta})^2 - (\frac{\eta(H_2)-\eta_{avg}}{\Delta\eta})^2]$ , where  $\eta_{avg} = \frac{\eta(j_1)+\eta(j_2)}{2}$  and  $\Delta\eta = \eta(j_1) - \eta(j_2)$ .

| Variable                                          | Meaning                                                       |
|---------------------------------------------------|---------------------------------------------------------------|
| $p_T(H_1)$ ( $p_T(H_2)$ )                         | Transverse momentum of the $H_1$ ( $H_2$ ) candidate          |
| $p_T(j_1)$ ( $p_T(j_2)$ )                         | Transverse momentum of the $j_1$ ( $j_2$ ) candidate          |
| $ \eta(jj) $                                      | VBF-jet pair pseudorapidity                                   |
| $M(jj)$                                           | VBF-jet pair invariant mass                                   |
| $\Delta R(H_1, H_2)$                              | $\Delta R$ distance between two Higgs bosons                  |
| $\Delta R(H_1, j_1)$                              | $\Delta R$ distance between $H_1$ and $j_1$                   |
| $\Delta R(H_1, j_2)$                              | $\Delta R$ distance between $H_1$ and $j_2$                   |
| $\Delta R(H_2, j_1)$                              | $\Delta R$ distance between $H_2$ and $j_1$                   |
| $\Delta R(H_2, j_2)$                              | $\Delta R$ distance between $H_2$ and $j_2$                   |
| $ \cos(\theta)^*(j_1) $                           | $ \cos(\theta) $ of $j_1$ in the six-jet center of mass frame |
| $ \cos(\theta)^*(j_2) $                           | $ \cos(\theta) $ of $j_2$ in the six-jet center of mass frame |
| $H1\text{-centrality} \cdot H2\text{-centrality}$ | Product of the Higgs boson centralities                       |

The Figure 3-8 illustrates the distributions of the ggFKiller response of three signal simulated events in the pre-VBF event region: for VBF ( $\kappa_{2V} = 2$ ), SM VBF and SM ggF. The observed response shows that VBF signal with anomalous  $\kappa_{2V}$  couplings have a high score, whereas the SM ggF signals a low score. It is important to note that the SM VBF signal tends to have a flat distribution and an enhancement at high ggFKiller values.

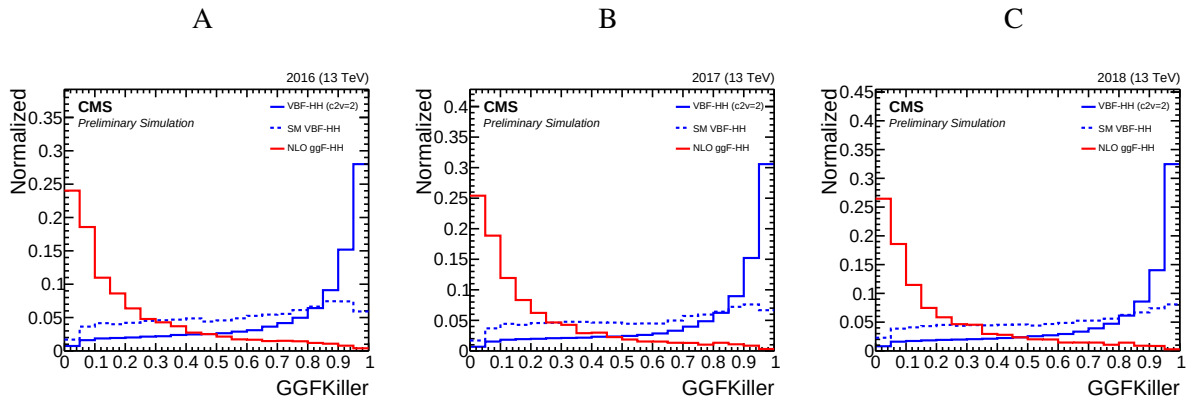


Figure 3-8. The ggFKiller output distributions for VBF ( $c_{2V}=2$ ) (solid blue), SM VBF (dashed blue), SM ggF (solid red) signals in pre-VBF events. The simulated data-taking conditions correspond to A) 2016, B) 2017, C) 2018.

### 3.5.3 Performance of the Production Mode Categorization

The ggFKiller score is used to reduce the ggF signal leakage into pre-VBF events and to create a more pure production mode categories. For this search, a score of ggFKiller=0.5 is found to be the optimal to further categorize events by production mode. The choice of this boundary is discussed in Subsection 3.7.1. The ggF Category combines two groups: Pre-ggF events, and Pre-VBF events with ggFKiller between 0 and 0.5, whereas the VBF Category is composed of Pre-VBF events with ggFKiller between 0.5 and 1.

Event yields of the ggF and VBF categories are reported in Table 3-9. Depending on the year, we observe that the ggF category contains around 96-97.0% of the SM ggF signal. Furthermore, the categorization has improved the relative ratios between SM ggF with respect to the VBF signals in the table: (1) SM ggF / SM VBF is around 4.5-5.3, and (2) SM ggF / VBF ( $\kappa_{2V} = 2$ ) is 0.03-0.07.

Table 3-9. Event yields of data and MC simulation after categorization into ggF and VBF categories. The processes ggF, VBF and VBF( $\kappa_{2V} = 2$ ) are multiplied by a factor of  $10^2$ .

| Dataset/Categ.           | 2016/ggF | 2016/VBF | 2017/ggF | 2017/VBF | 2018/ggF | 2018/VBF |
|--------------------------|----------|----------|----------|----------|----------|----------|
| ggF                      | 1559.4   | 58.0     | 1003.3   | 31.2     | 1798.0   | 67.9     |
| VBF                      | 16.5     | 10.8     | 11.9     | 6.9      | 22.2     | 14.7     |
| VBF( $\kappa_{2V} = 2$ ) | 805.1    | 778.9    | 558.5    | 466.6    | 1152.6   | 1058.8   |
| QCD                      | 165189.0 | 8776.9   | 74259.5  | 2906.6   | 120202.3 | 5920.5   |
| $t\bar{t}$               | 10458.8  | 916.1    | 6121.9   | 462.7    | 12928.2  | 1089.8   |
| Single Higgs             | 684.4    | 53.6     | 483.5    | 26.7     | 936.7    | 57.1     |
| $ZZ \rightarrow 4b$      | 36.7     | 0.9      | 21.4     | 0.6      | 37.6     | 1.0      |
| Total MC Bkg.            | 176368.9 | 9747.5   | 80886.3  | 3396.6   | 134104.8 | 7068.4   |
| Data                     | 211586   | 11286    | 159965   | 8129     | 245374   | 15207    |

The production mode classification has a very good performance. For ggF signals, we observe that the fraction of events classified in the ggF categories as a function of the  $\kappa_\lambda$  coupling is at least  $\sim 95\%$  and has a more or less flat  $\kappa_\lambda$ -dependency, as presented in Figure 3-9 A). Due to the training conditions, the VBF signal efficiency after the ggFKiller cut is the most optimal for  $\kappa_{2V}$  anomalous couplings. For VBF signals, we observe a good efficiency of signal events passing the ggFKiller cut as a function of the  $\kappa_{2V}$  and  $\kappa_V$  couplings, as illustrated Figure 3-9 (B-D).

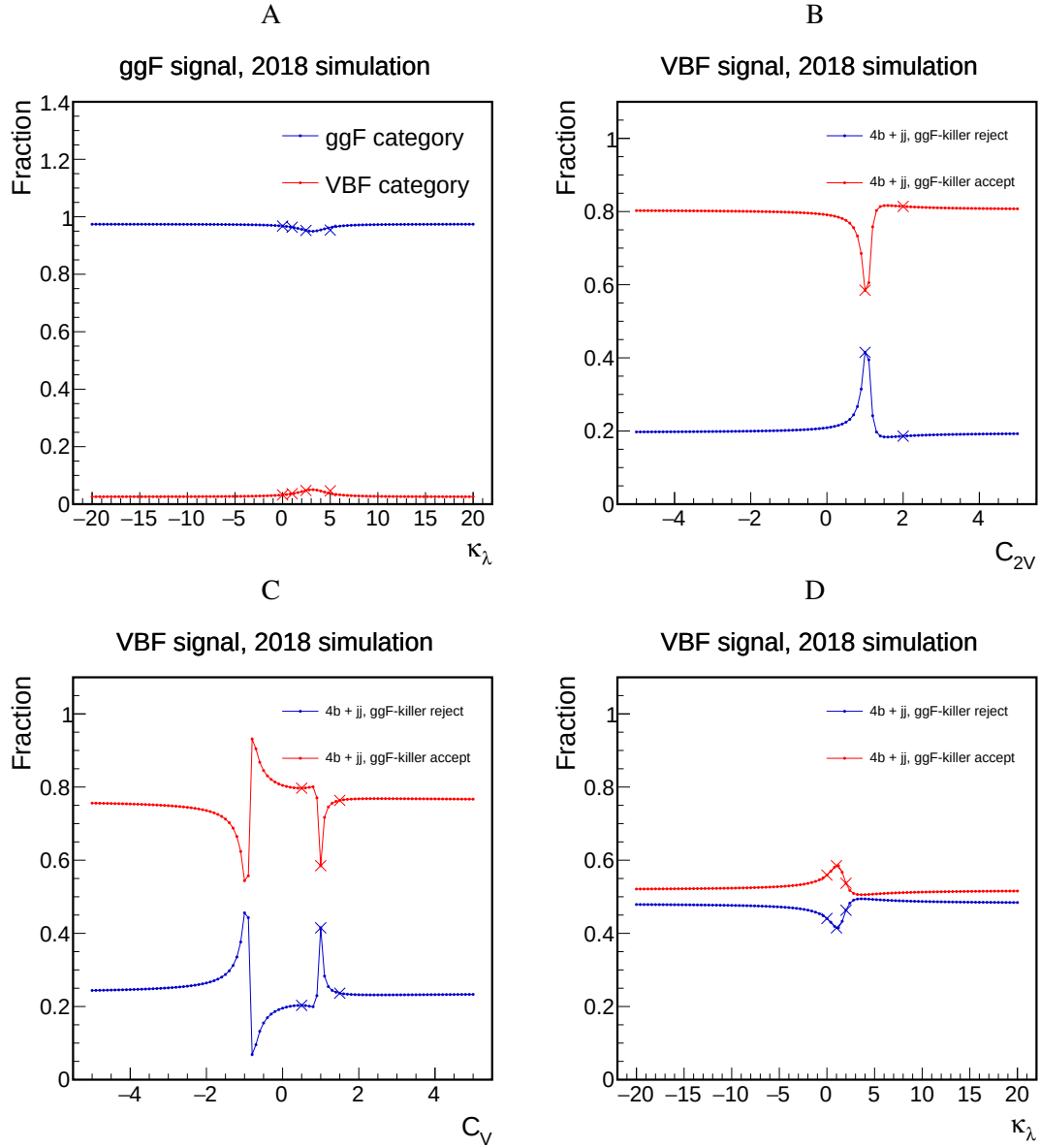


Figure 3-9. Performance of the ggFKiller in the production mode classification as a function of the couplings. A) Fraction of inclusive ggF signal events that are classified in the ggF and VBF categories as a function of  $\kappa_\lambda$ . B/C/D) Fraction of VBF signal events from the pre-VBF region classified in the ggF category and VBF categories as function of the  $\kappa_{2V}/\kappa_V/\kappa_\lambda$  coupling. The crosses indicate the values obtained directly from the simulated samples, serving as cross-check to the values computed with the signal modeling procedure represented by the full circles. The simulated data-taking conditions correspond to 2018.

### 3.6 Event Regions

The analysis strategy demands the definition of regions of the phase space, which events are used to perform dedicated tasks such as background modeling and signal strength measurement. In this work, the set of preselected events are defined depending on the b-tagging multiplicity and the invariant masses of the two reconstructed Higgs boson candidates ( $m_{H_1}$  and  $m_{H_2}$ ). The definition and purpose of these regions are explained in this section.

#### 3.6.1 The ‘4b’ and ‘3b’ Regions

All events passing the trigger and preselection requirements presented in Section 3.4 are denominated as ‘4b’ events. On the other hand, events passing the trigger in which the fourth jet by b-tagging score fails the preselection b-tagging cut (Table 3-5) are denominated as the ‘3b’ events. Event yields for data and MC simulation are presented in Table 3-7 for the ‘3b’ region.

Table 3-10. Event yields of data and MC simulation after categorization requirements in the 3b region. The processes ggF, VBF and VBF( $\kappa_{2V} = 2$ ) are multiplied by a factor of  $10^2$ .

| Dataset/Categ.           | 2016/ggF  | 2016/VBF | 2017/ggF | 2017/VBF | 2018/ggF  | 2018/VBF |
|--------------------------|-----------|----------|----------|----------|-----------|----------|
| ggF                      | 2419.2    | 77.8     | 1188.9   | 32.0     | 2404.3    | 78.7     |
| VBF                      | 48.9      | 16.2     | 29.3     | 7.2      | 62.0      | 18.9     |
| VBF( $\kappa_{2V} = 2$ ) | 1424.8    | 880.0    | 719.7    | 343.0    | 1689.9    | 886.7    |
| QCD                      | 1725011.0 | 68994.7  | 532655.1 | 16344.5  | 1058750.0 | 43059.3  |
| $t\bar{t}$               | 100496.2  | 5819.9   | 50288.3  | 2489.5   | 111774.6  | 6518.1   |
| Single Higgs             | 1657.5    | 106.9    | 973.8    | 45.8     | 2053.0    | 107.3    |
| $ZZ \rightarrow 4b$      | 90.9      | 2.0      | 41.6     | 1.0      | 85.3      | 2.1      |
| Total MC Bkg.            | 1827255.6 | 74923.5  | 583958.8 | 18880.8  | 1172662.9 | 49686.8  |
| Data                     | 2285706   | 86800    | 1263243  | 49416    | 2273811   | 107918   |

Depending on the production mode category, there are 7-10 times more data events in the 3b region than in the 4b region. While this signal contamination in the 3b region cannot be avoided and is due to the limited efficiency of the b tag algorithm, the signal to background ratio in the 3b region is significantly small compared to the 4b region. For instance, for 2016 SM ggF signal, the corresponding signal/background at the ggF 3b region is about 15% the one in the ggF 4b region. Therefore, the ‘3b’ region data events are suitable as input for a data-driven background model and further explained in Chapter 4.

### 3.6.2 Analysis and Validation Regions

The ultimate goal of the dissertation is to extract a possible signal from the collected data. A key step is then to define a signal-enriched region with large signal-to-background ratio for the main data analysis, denominated analysis signal region ( $A_{SR}$ ). At the same time, we have to define a region with low signal-to-background ratio and close to the SR to construct the background model, named analysis control region ( $A_{CR}$ ).

In the two-dimensional phase space of the two reconstructed Higgs boson masses or  $m_{H_1}$ - $m_{H_2}$  plane, the HH signal events populate the region surrounding the peak located around  $m_{H_1}=125$  GeV and  $m_{H_2}=120$  GeV. Note that the difference between  $m_{H_1}$  and  $m_{H_2}$  arises from the  $p_T$  ordering imposed to the Higgs boson candidates. For this search, the optimal choice of the  $A_{SR}$  and  $A_{CR}$  regions is a circle (eq. 3-2) and a ring (eq. 3-3) centered at the coordinates  $(C1,C2)=(125$  GeV,120 GeV) in the  $m_{H_1}$ - $m_{H_2}$  plane. The  $A_{SR}$  and  $A_{CR}$  regions are sketched in the Figure 3-10.

$$\sqrt{(m_{H_1} - C_1)^2 + (m_{H_2} - C_2)^2} < 25 \text{ GeV} \quad (3-2)$$

$$25 \text{ GeV} \leq \sqrt{(m_{H_1} - C_1)^2 + (m_{H_2} - C_2)^2} < 50 \text{ GeV} \quad (3-3)$$

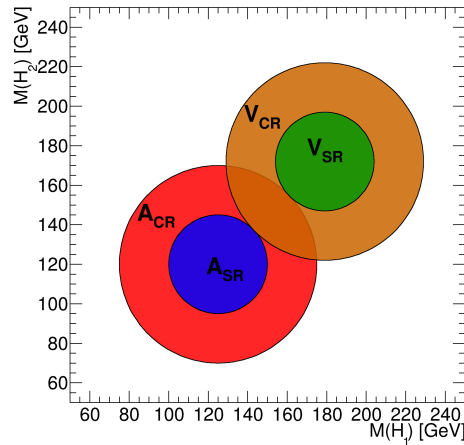


Figure 3-10. Definition of signal and control regions for the analysis and validation regions.

As will be detailed in Chapter 4, events from a phase space orthogonal to  $A_{SR}$  is needed to fully validate the background model. To that end, we define the validation signal region circle ( $V_{SR}$ ) and validation control regions ring ( $V_{CR}$ ) also in the  $m_{H_1}$ - $m_{H_2}$  plane but shifting their center along the pairing diagonal line to  $(C1,C2)=(179 \text{ GeV}, 172 \text{ GeV})$ . The analysis and validation regions used in this dissertation are illustrated in the sketch in Figure 3-10. Furthermore, the distribution of the SM ggF and VBF ( $\kappa_{2V} = 2$ ) signal simulated events across the regions are shown in Figure 3-11.

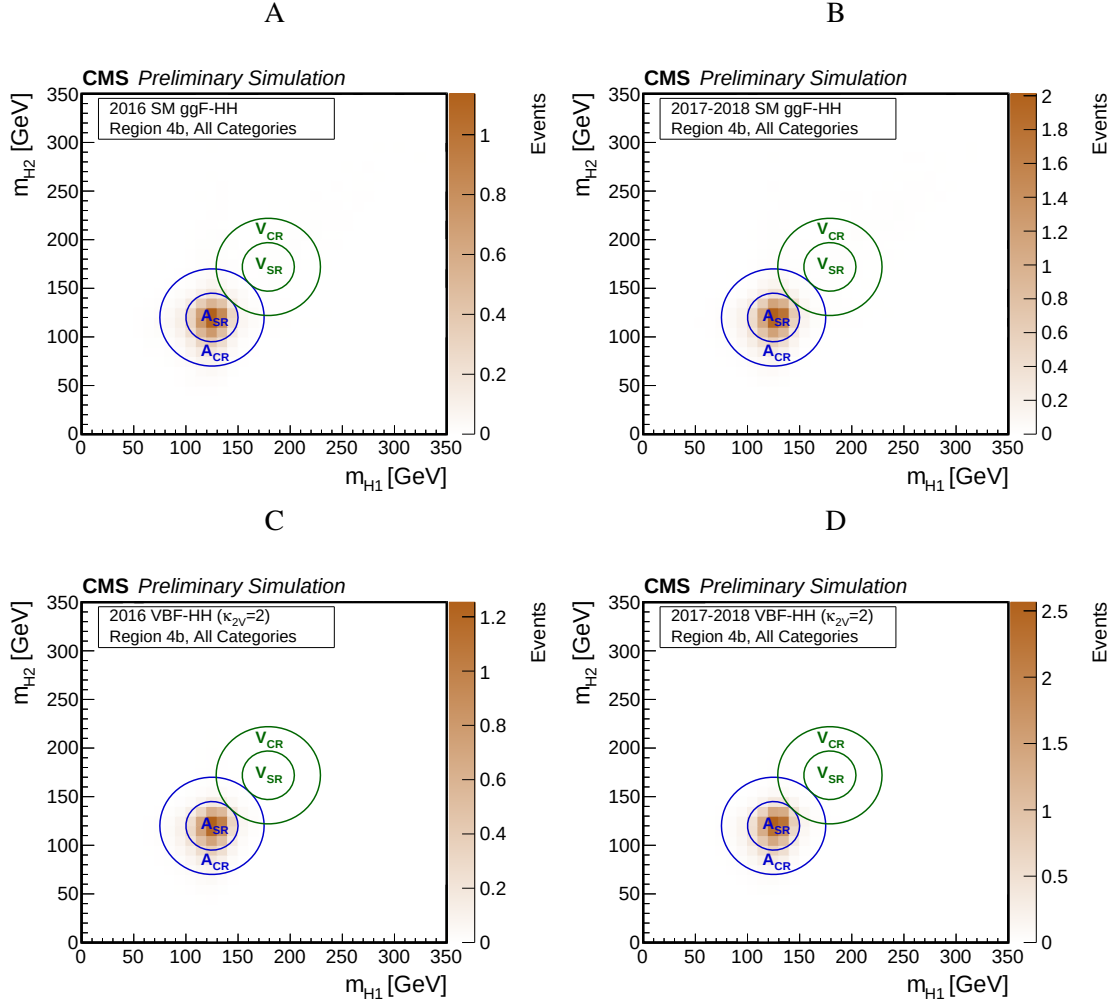


Figure 3-11. Distribution of the preselected signal events in the  $m_{H_1}$ - $m_{H_2}$  plane for A) SM ggF in 2016 conditions, B) SM ggF in 2017-2018 conditions, C) VBF with  $\kappa_{2V} = 2$  in 2016 conditions, D) VBF with  $\kappa_{2V} = 2$  in 2017-2018 conditions. The simulated events are normalized to their theoretical cross section and integrated luminosity. The definition of signal and control region in the analysis and validation region are presented.

A summary of the efficiencies to select ggF and VBF signals as a function of their coupling values is given in Figure 3-12 for the 2018 data-taking conditions. The efficiency lines illustrate the application of event requirements going from the online trigger to the analysis signal region. The 2016 and 2017 efficiencies have a similar behavior, as seen in Appendix C

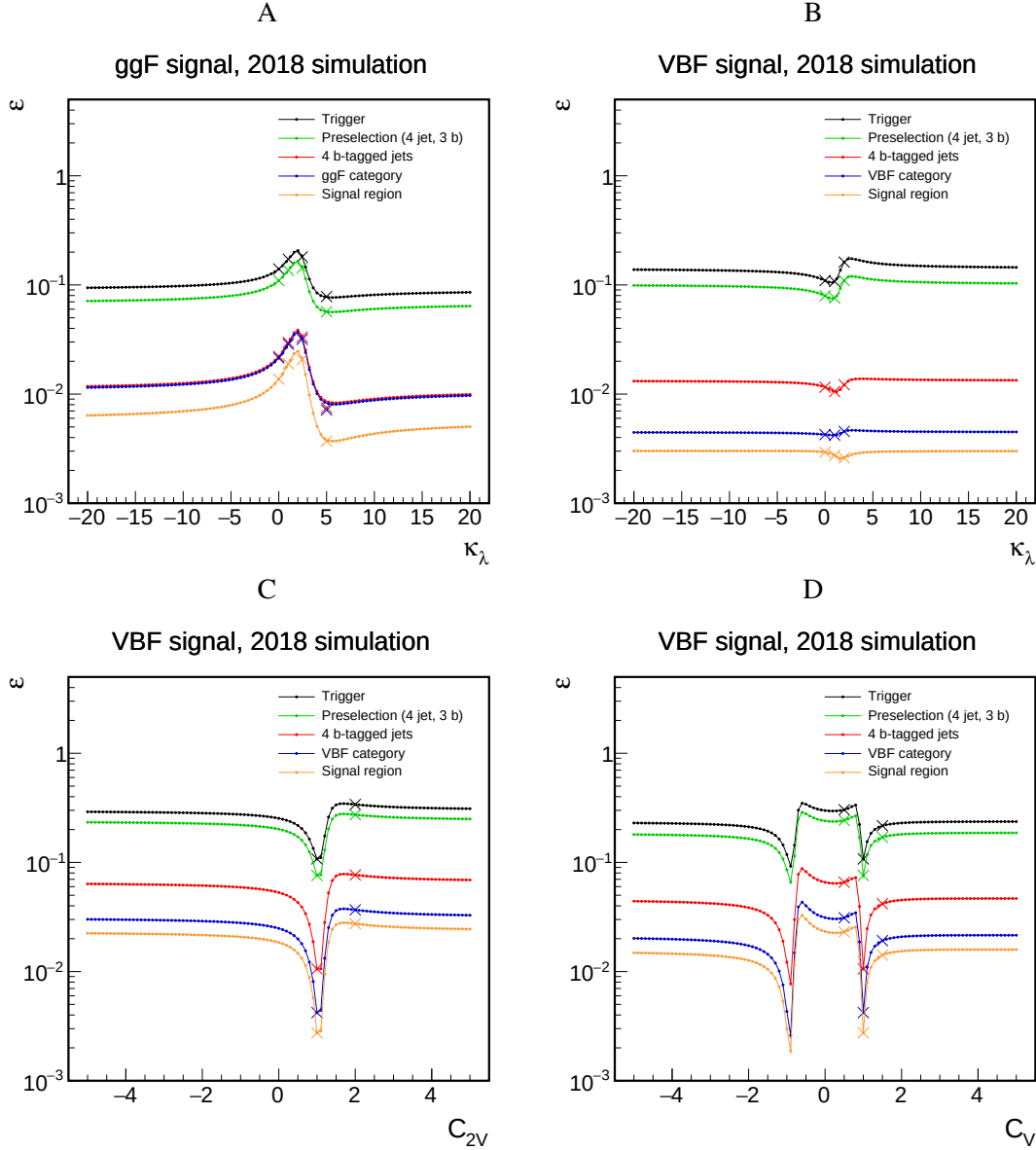


Figure 3-12. Selection efficiency ( $\epsilon$ ) of the 2018 selections versus the ggF and VBF couplings.  $\epsilon$  vs  $\kappa_\lambda$  for the A) ggF and B) VBF signals, C)  $\epsilon$  versus  $\kappa_{2V}$  and D)  $\epsilon$  versus  $\kappa_V$  for the VBF signals. In all figures, the couplings that are not displayed on the x-axis are assumed to be fixed to the SM value. The crosses indicate the values obtained directly from the simulated samples, serving as cross-check to the values computed with the signal modeling procedure (denoted by the full circles).

### 3.7 Event Subcategorization

#### 3.7.1 Subcategory Boundaries

As seen in Chapter 1, the coupling values can change the spectrum of the properties of the HH system. For this purpose, after events are classified into either the ggF or VBF categories, the next step is to split the categories into subcategories. This is to optimize further the analysis by targeting specific ggF and VBF signatures. The ggF and VBF analysis subcategories are defined based on  $m_{HH}$  and ggFKiller score, respectively. The four subcategories are defined in Table 3-11. The associated efficiencies as function of the couplings are illustrated in Figure 3-15.

Table 3-11. Selection requirements used to define the event subcategorization.

| Subcategories                                     | Selection requirements                                    |
|---------------------------------------------------|-----------------------------------------------------------|
| ggF Category 1 or 'Low-mass' ggF                  | ggF category events, $m_{HH} < 450$ GeV                   |
| ggF Category 2 or 'High-mass' ggF                 | ggF category events, $m_{HH} \geq 450$ GeV                |
| VBF Category 1 or 'SM-like' VBF                   | VBF category events, $0.5 \leq \text{ggFKiller} < 0.970$  |
| VBF Category 2 or 'Anomalous- $\kappa_{2V}$ ' VBF | VBF category events, $0.970 \leq \text{ggFKiller} \leq 1$ |

Two ggF subcategories are defined depending on the  $m_{HH}$  kinematics. The boundary for the ggF categories is set at  $m_{HH} = 450$  GeV. This choice is driven by the ggF signal properties and the capability to validate the background model detailed in Section 4.2. For  $m_{HH} > 450$  GeV, ggF signals have similar distributions under the various  $\kappa_\lambda$  hypotheses, as illustrated in Figure 3-13 A), and where we expect their ratio to the SM curve to be flat. Consequently, we can train an optimal BDT output observable for  $\kappa_\lambda$  hypotheses, and there is no gain in increasing further the threshold. Inversely, lowering the threshold much below 450 GeV poses a constraint on the  $m_{H_1}$  and  $m_{H_2}$  values in the low  $m_{HH}$  category (or cat. 1) for background events, biasing them towards lower values. This effect impact how well '3b' and '4b' control region events populate the top-right corner of the  $(m_{H_1}, m_{H_2})$  plane, reducing the statistical power to validate and trained a good validation background model. This effect is illustrated by different boundaries in Figure 3-14.

Two VBF subcategories are defined based on the SM-like and anomalous- $\kappa_{2V}$  signatures. The VBF signal properties depending on the  $\kappa_{2V}$  are captured by the ggFKiller score, as seen in



Figure 3-13 B), and therefore, it is used to set the boundaries. There are two boundaries defined using the VBF couplings sensitivity as figure of merit: ( $\beta_1$ ) The transition between ggF and VBF categories, and ( $\beta_2$ ) the transition from SM-like and BSM-like kinematics. Here sensitivity means the upper limits on the signal strength parameter ( $\mu$ ) using the background model and observables described in Section 4.2 and Subsection 3.7.2, respectively. The procedure is the following:

- The value of  $\beta_2$  is defined by studying the  $\kappa_{2V} = 2$  sensitivity in the VBF cat. 2 (a counting experiment) for thresholds starting from 0.900 in steps of 0.005. The boundary  $\beta_2 = 0.97$  is found optimal for this analysis.
- The value of the  $\beta_1$  is found by testing ggFKiller points: 0.3, 0.4, 0.5, 0.6, 0.7. Most of the expected anomalous  $\kappa_{2V} = 2$  sensitivity is at high ggFKiller scores, thus the choice of the  $\beta_1$  boundary only impacts the SM sensitivity. Consequently, the figure of merit is the SM VBF sensitivity in the SM-like subcategory. The tests showed the sensitivity improves by a few percent when increasing the thresholds. However, there is a significant amount of control region data loss affects our capability to build the data-driven background model. For example, the 2016 SM sensitivity improvements to up to 2% (5%) by requiring a value of 0.6 (0.7) instead of 0.5, but the loss of CR events is around 23% (55%). Therefore, ggFKiller=0.5 is the optimal boundary.

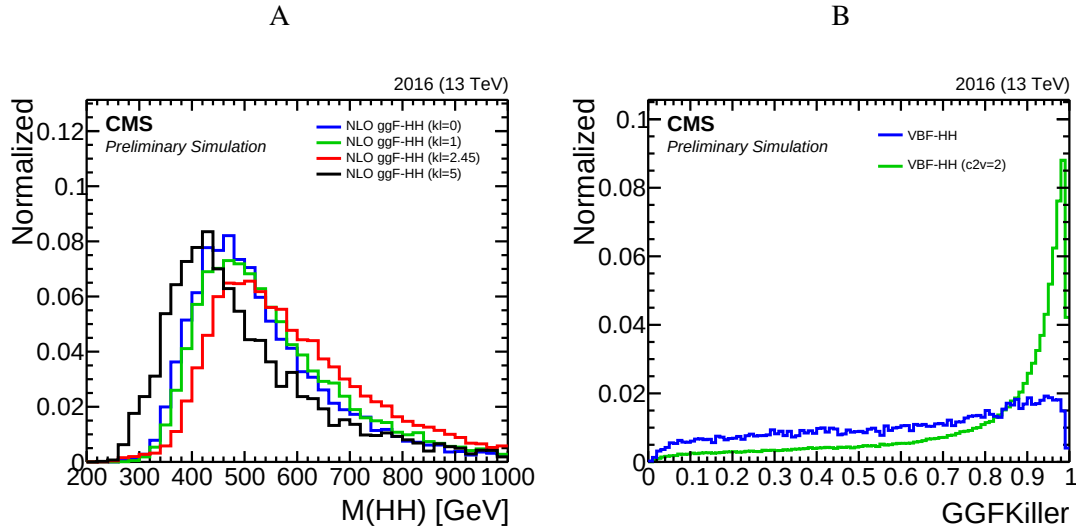


Figure 3-13. Variables used to create ggF and VBF subcategories in the ggF and VBF signal region. A) Invariant mass of the Higgs boson pair for ggF  $\kappa_\lambda = 0, 1, 2.45, 5$  signals, B) ggFKiller score for VBF  $\kappa_{2V} = 1, 2$  signals. The simulated data-taking conditions correspond to 2016.

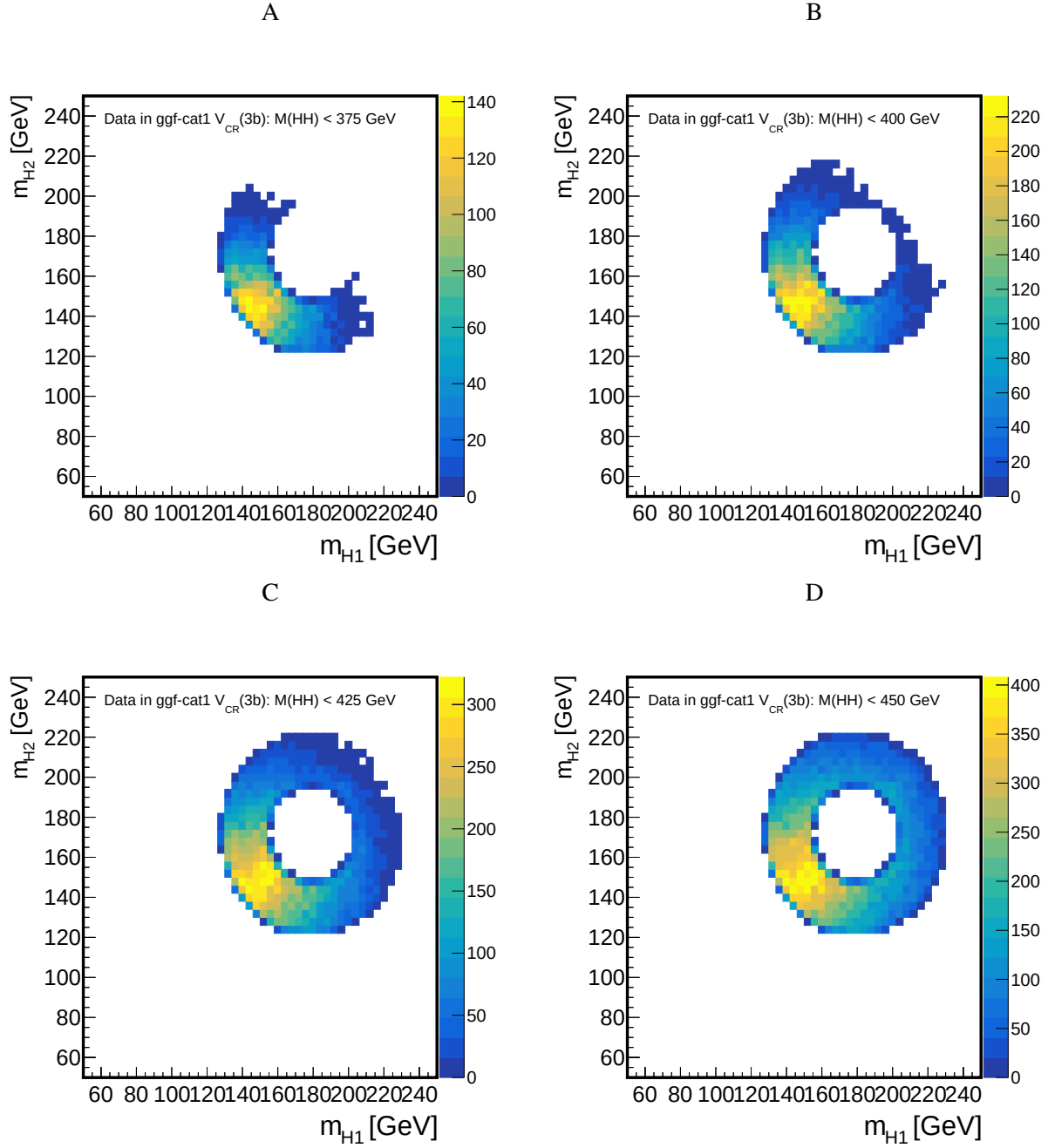


Figure 3-14. Distribution of data events in the ggF category 1 in the  $(m_{H_1}, m_{H_2})$  plane for the  $V_{CR}(3b)$  region using different mass thresholds (in GeV). A) 375, B) 400, C) 425, and D) 450.

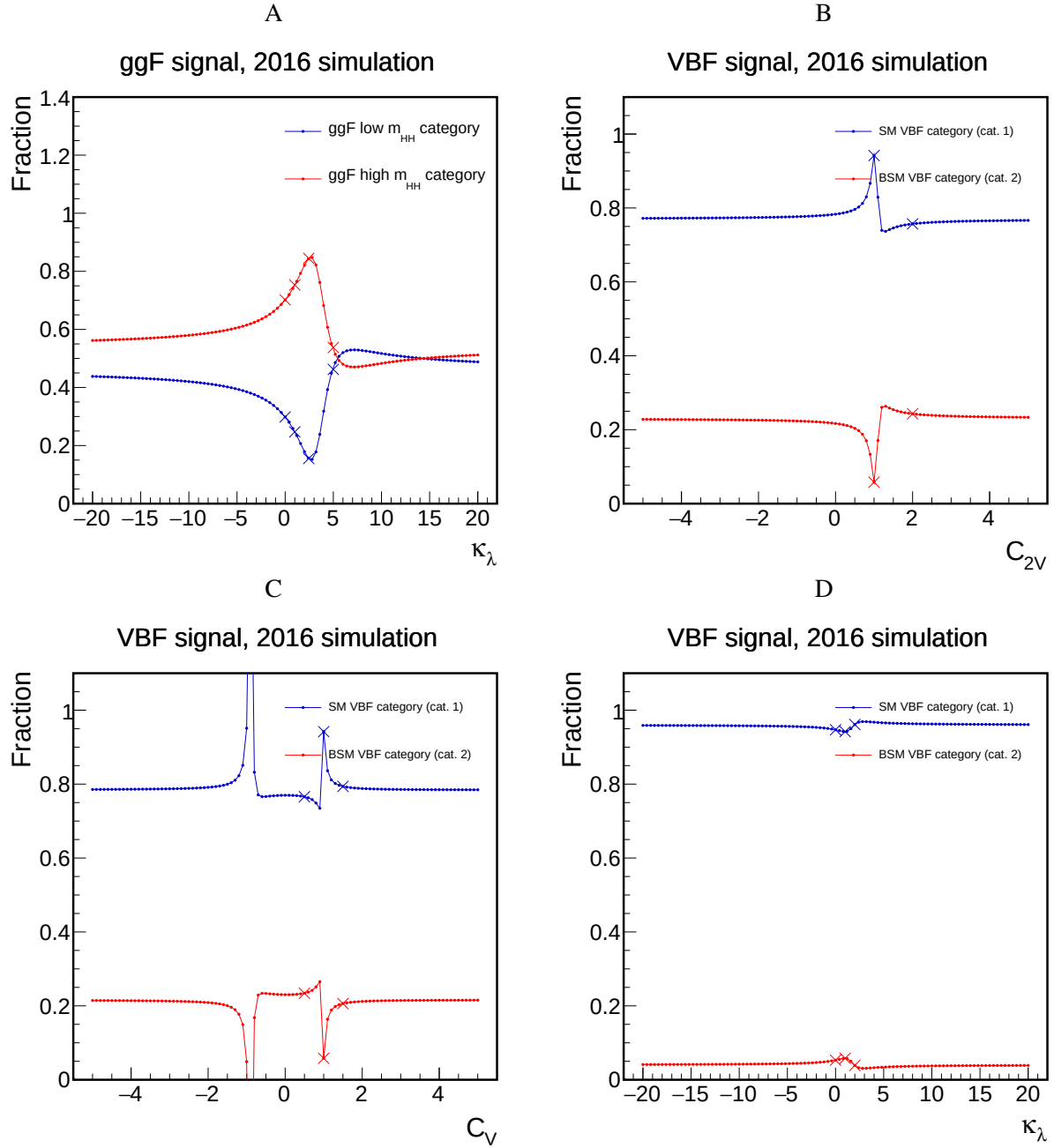


Figure 3-15. Performance of the production mode subcategorization versus the ggF and VBF couplings. A) Fraction of ggF signal events from the ggF Category in the low-mass ggF subcategory and high-mass subcategory as a function of  $\kappa_\lambda$ . D/C/D) Fraction of VBF signal events from the VBF category classified in the SM-like subcategory and anomalous- $\kappa_{2V}$  (BSM) subcategory as function of the  $\kappa_{2V}/\kappa_V/\kappa_\lambda$  coupling. The crosses indicate the values obtained directly from the simulated samples, serving as cross-check to the values computed with the signal modeling procedure (denoted by the full circles). The simulated data-taking conditions correspond to 2016.

### 3.7.2 Signal Extraction Observables

Optimal observables are chosen for each subcategory of the data analysis to obtain the final results presented in Chapter 5. For ggF categories, the observable is the distribution of a multidimensional variable that aims at separating the signal and background events by combining the information of angular correlations, object kinematic properties, and object quality variables. The chosen multivariate analysis (MVA) algorithm is a Boosted Decision Tree (BDT) implemented in the XGBoost package [114]. The training of these BDTs is possible due to rich variable information and the large number of data-driven background events (presented in Chapter 4).

The BDT outputs are trained per category and per datasets. The signal and background samples are the SM ggF model and the data-driven background model, respectively. The BDT is built using the 16 input variables listed in Table 3-12. Furthermore, Figures 3-16 and 3-17 show the distributions of the signal and background for 2017-2018 ggF category 1 and ggF category 2, respectively. The equivalent distributions for 2016 are presented in Appendix G.

Table 3-12. BDT input variables in ggF Category 1 and ggF Category 2.  $H_1$  is the highest- $p_T$  Higgs candidate and  $H_2$  is the lowest- $p_T$  Higgs candidate.

| BDT input variables                                                                |
|------------------------------------------------------------------------------------|
| Mass of the $H_1$ candidate, $M(H_1)$                                              |
| Mass of the $H_2$ candidate, $M(H_2)$                                              |
| Mass of the Higgs pair system, $m_{HH}$                                            |
| Transverse momentum of the $H_1$ candidate, $P_T(H_1)$                             |
| Transverse momentum of the $H_2$ candidate, $P_T(H_2)$                             |
| Pseudorapidity separation between the two Higgs candidates, $\Delta\eta(H_1, H_2)$ |
| $\Delta R$ distance between two b jets of the $H_1$ candidate, $\Delta R(H_1(bb))$ |
| $\Delta R$ distance between two b jets of the $H_2$ candidate, $\Delta R(H_2(bb))$ |
| $ \cos(\theta)^*(H) $ in HH frame                                                  |
| $ \cos(\theta)^*(b) $ in $H_1$ frame                                               |
| Sum of four b jets' regressed $p_T$                                                |
| Transverse momentum of the HH system, $p_T(HH)$                                    |
| Number of tight b-tags in 3 highest b-tags                                         |
| Sum of 3b's resolution scores                                                      |
| Minimal $\Delta R$ distance between two b jets, $\text{Min} \Delta R(bb) $         |
| Maximum pseudorapidity separation between two b jets, $\text{Max} \Delta\eta(bb) $ |

Figure 3-16. Input variables for the BDT output training in the 2017-2018 ggF Category 1. The SM ggF (signal) is in solid red, whereas the data-driven background modeling is in solid blue.

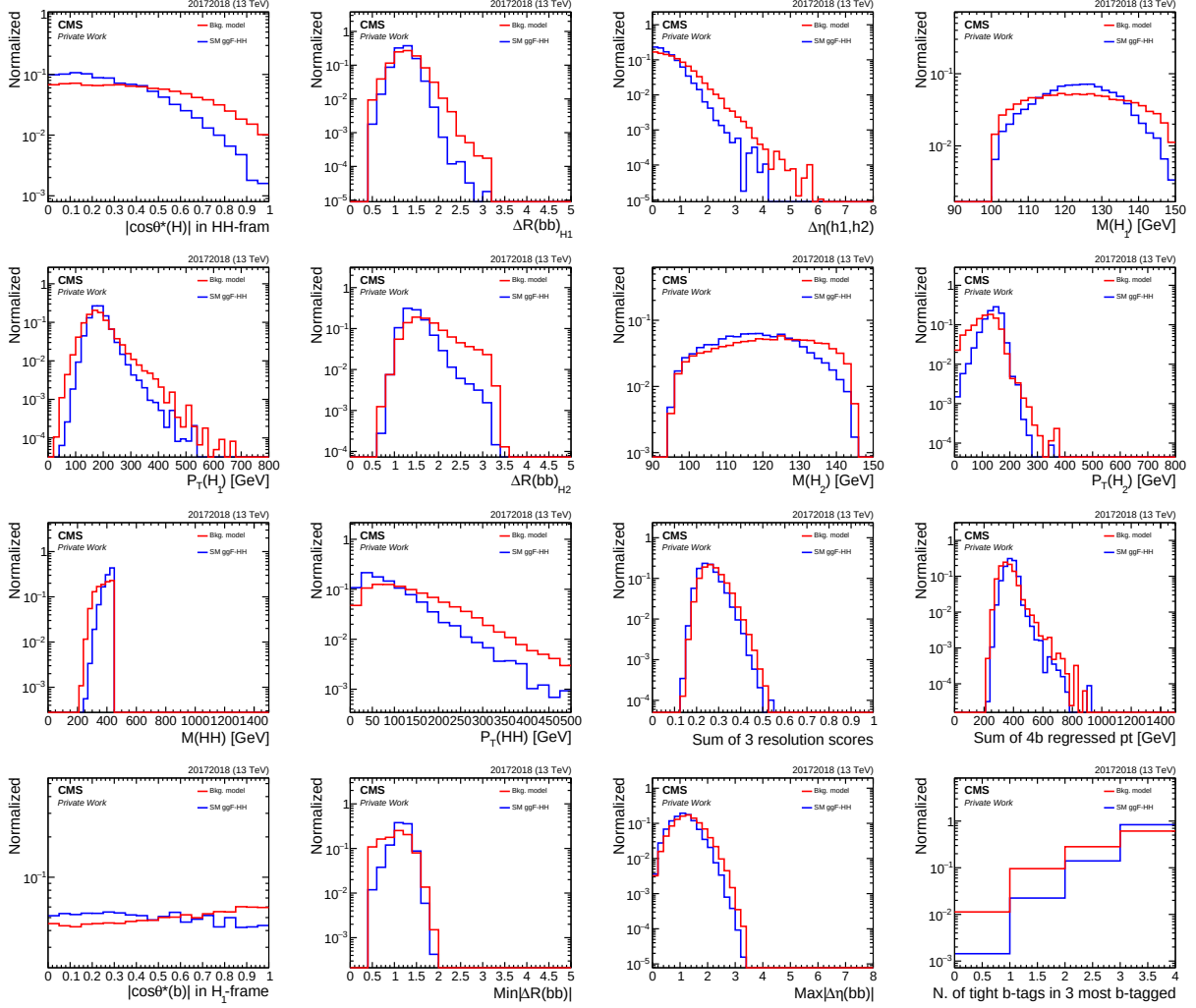
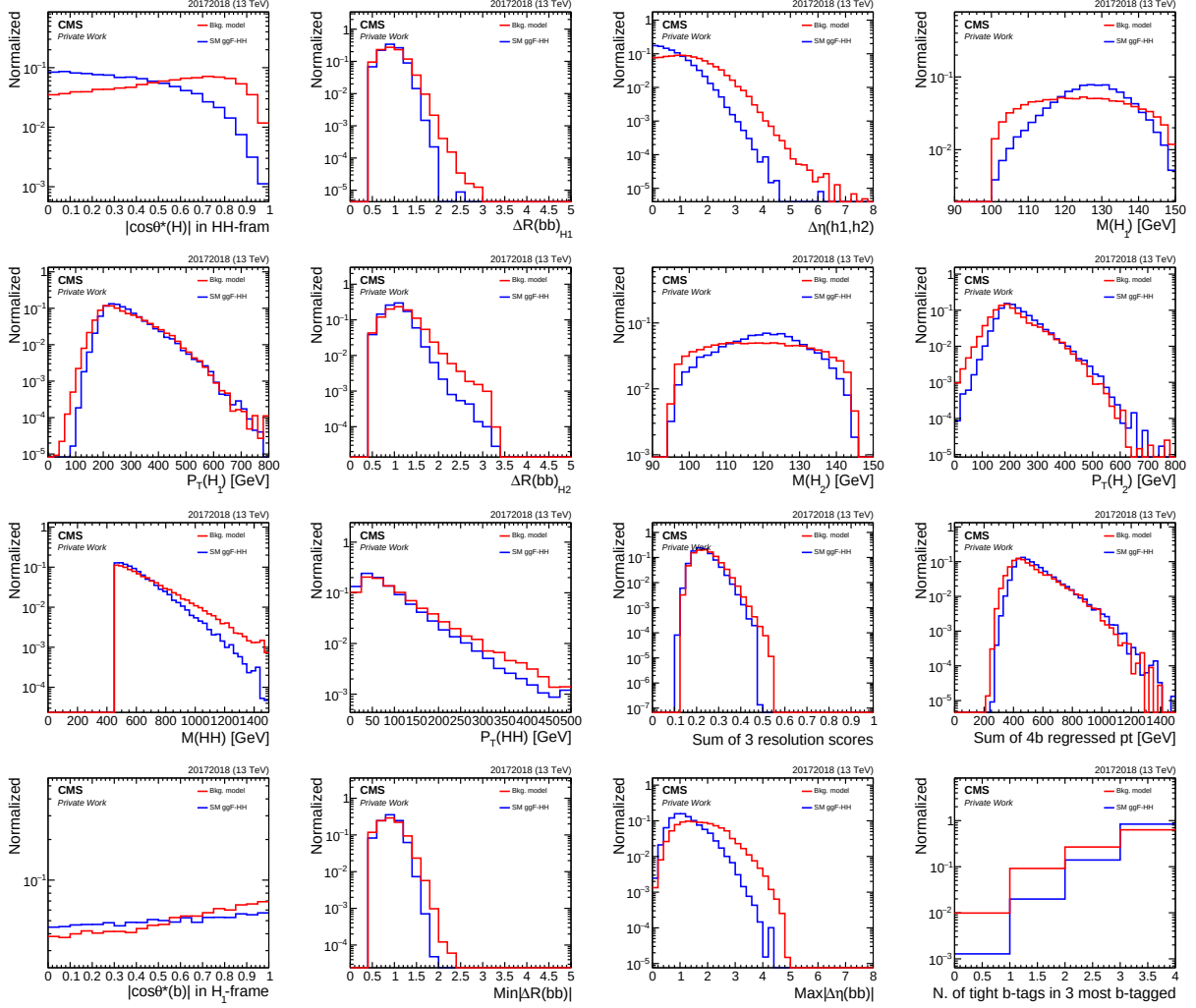


Figure 3-17. Input variables for the BDT output training in the 2017-2018 ggF Category 2. The SM ggF (signal) is in solid red, whereas the data-driven background modeling is in solid blue.



A good practice is to avoid the introduction of a bias on the BDT distribution for background events by reusing the same training/test samples to evaluate the final results. As it will be seen in Chapter 4, the analysis sensitivity is dominated by the background model uncertainties. If the background event samples are reduced due to the BDT training, then the statistical uncertainty on the background model will increase, and there will be direct impact on the final result. For this reason, a special training method is designed to allow for the 100% of the background model to be used for the signal extraction. The procedure is summarized as follows:

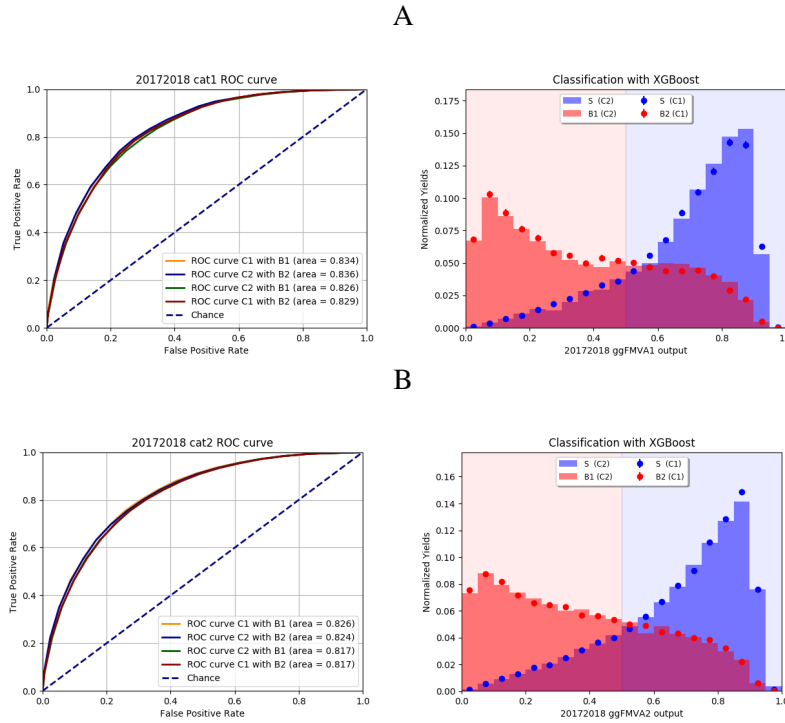
1. Prepare the expected signal (S) and background model (B) samples in the signal region
2. Split the background model into two halves, denoted as B1 and B2 samples.
3. A first classifier (C1) is trained using the signal S and the background B1 sample using the hyperparameter optimization procedure used for the ggFKiller score with nine grids.
4. A second classifier (C2) is trained with the signal (S) and the background model B2 samples using the exact same resulting hyperparameters of the classifier C1.

Detailed information of the training results of the C1 and C2 classifiers is presented in Appendix G. Next, compatibility checks are performed to ensure that the C1 and C2 will give the same performance in any event sample: (1) The ROC-AUC curves associated to the evaluation of the classifier C1 in the S and B1 samples and the evaluation of the classifier C2 in the S and B2 samples are compatible; (2) The ROC-AUC curves associated to the evaluation of the classifier C1 in the S and B2 samples and the evaluation of the classifier C2 in the S and B1 samples are compatible; (3) The distributions have similar shapes when evaluating the S (B1) samples using classifier C2 versus the S (B2) samples using classifier C1.

The compatibility checks show that the two classifiers are compatible and give the same results. This is illustrated for the 2017-2018 dataset in Figure 3-18. Lastly, for each ggF category and dataset, the BDT output used for the final results is built by evaluating the C1 and C2 classifiers using the following recipe:

- Background model events in sample B1: 100% uses classifier C2
- Background model events in sample B2: 100% uses classifier C1
- Simulated events: A 50% uses classifier C1 and the other 50% uses classifier C2
- Observed data events: A 50% uses classifier C1 and the other 50% uses classifier C2

Figure 3-18. Compatibility checks on the classifiers used for the 2017-2018 ggF observables. A) Category 1 and B) Category 2. The left figure shows the ROC-AUC curves performance checks. The right figure illustrates the distributions for signal S and background model B1 (B2) samples using the C2 (C1) classifier.



For VBF categories, data-driven background model statistics is small when compared to the one of the ggF categories. This makes it challenging to prepare a BDT output, and consequently, a different strategy is followed. For VBF category 1, mass variables associated to the Higgs boson pairs are expected to have similar separation performance. The chosen variable is the invariant mass of the Higgs boson pair system ( $m_{HH}$ ). For the VBF category 2, the expected number of background events is at the order of 5-20 events, depending on the dataset. Therefore, the observable is a single bin used to perform a counting experiment.



## CHAPTER 4 MODELING OF PHYSICS PROCESSES

### 4.1 Monte Carlo Simulation

#### 4.1.1 Motivation

An accurate modeling of the physics processes occurring in collisions is needed in order to search for a specific signal process in the experimental data. The physics of pp collisions is modeled numerically based on existing theoretical predictions by solving complex integrals and equations over a multidimensional space of involved particles, kinematic properties, and theoretical parameters. For modeling purposes, their analytical or numerical evaluation is impractical. However, the Monte Carlo (MC) method was formally developed by John Von Neumann and Stanislaw Ulam (1940s), to study the distance that neutrons will likely travel in a material. This method is able to randomly sample predictions of important system variables, assuming they are fully described by probability density functions. Particle physicists have developed MC simulation tools to model physics processes.

MC event generators are used to model the physics phenomena in high-energy collisions. An example of the physics phenomena described by an event generator is seen in Figure 4-1. These generators provide a description of the different regimes of the collision event: hard parton scattering, parton showering, hadron formation or hadronization, and unstable hadron decays. These tools generate events with a theoretical precision that can include higher orders in the perturbative calculations beyond “tree-level” Feynman diagrams (e.g. NLO). Some well-established tools are the matrix element and parton scattering generator MadGraph [115], and PYTHIA [116] which evolves the parton-level event to the final state.

To search for events of interest in the experimental data, we need to know how the detector layers will respond to them and reconstructs them. This information is obtained via MC detector simulation using programs such as GEANT4 [117] or FLUKA [118]. The general procedure is to implement a detailed detector model (geometry, materials, alignment, etc.). Then, the program models the propagation of the event generated particles through the detector layers and their interaction with the materials. The detector interactions are translated into realistic signals (digitization) as well, and then the same event reconstruction used in real data is applied.

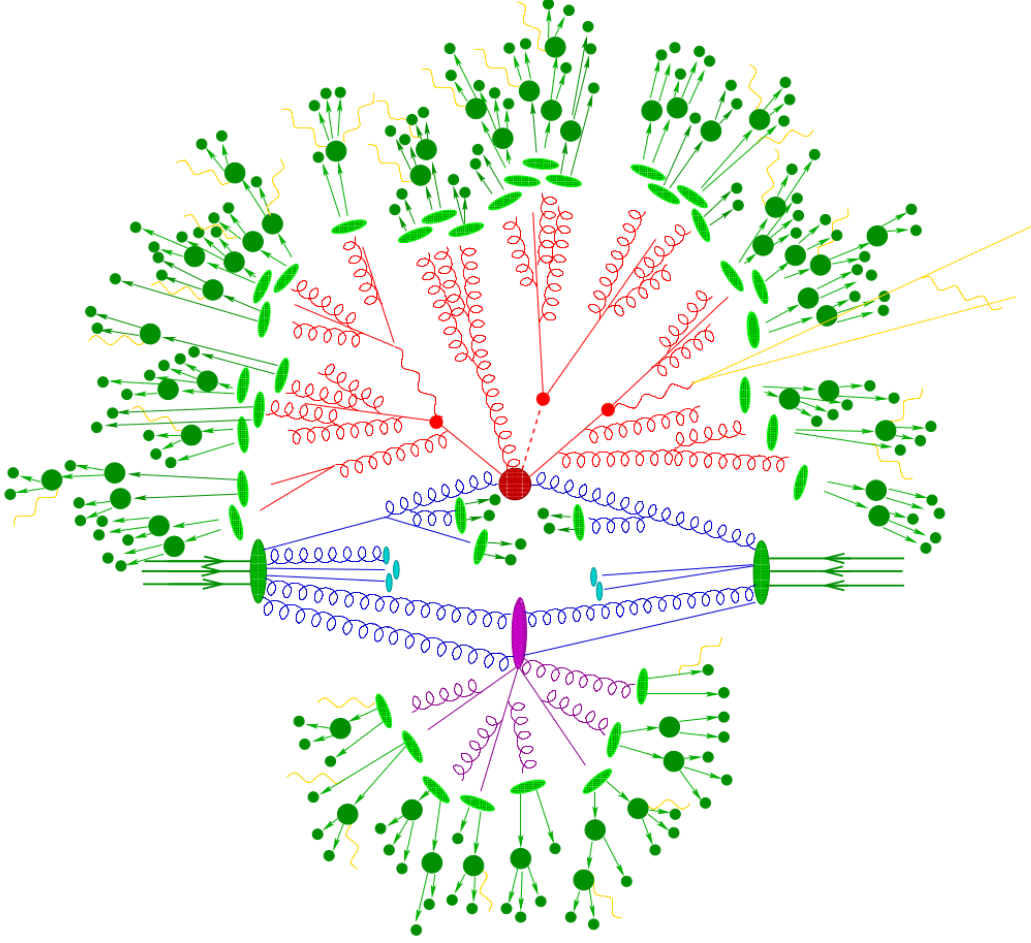


Figure 4-1. Sketch of processes involved in a proton-proton collision [119]. The hard scattering process (big red circle) corresponds to the production of a Higgs boson in association with two top-quarks or  $t\bar{t}H$  (three small red circles). The secondary interaction or underlying event (big purple ellipse) is also shown. Before hadron formation (green light ellipses), parton showering and radiation takes place (red). Hadron decays (dark green circles) and photon emission (yellow) are shown as well.

#### 4.1.2 Simulated Samples

The production of signal and background processes in pp collisions are simulated using Monte Carlo (MC) simulations. The full simulation is carried out independently per year or ‘campaign’, considering the detector and data-taking conditions (i.e. PU profile and luminosity).

For ggF anomalous coupling modeling, signals with  $\kappa_\lambda = 0, 1, 2, 2.45, 5$  are simulated at NLO precision in QCD using the POWHEG generator [120, 121, 122] with the NNPDF3 parton distribution function set at the NNLO precision. For VBF anomalous coupling modeling, signals

with seven different combination of  $(\kappa_V, \kappa_{2V}, \kappa_\lambda)$  couplings are generated at LO precision using the MadGraph5\_aMC@NLO generator. In addition, ggF signals are simulated at LO accuracy in the MadGraph5\_aMC@NLO generator, and used for BSM EFT benchmark modeling (Table 1-5).

Background events are simulated to study the background event properties, but they are not used in the nominal background model. The dominant background events, QCD multijet production, are generated using MadGraph5\_aMC@NLO at LO accuracy and by regions of the scalar sum of the transverse momentum (or HT) of scattering partons at matrix element. The  $t\bar{t}$  production is generated at NLO accuracy using the POWHEG generator and normalized to the NNLO theoretical cross section [123]. Other minor backgrounds (Single Higgs production,  $ZZ \rightarrow 4b$ ) are also simulated. The list of samples and their normalization cross sections are presented in Tables D-1, D-2 and D-3 in the Appendix D.

All generated events are interfaced with PYTHIA8 [116] for parton showering, hadronization, and decays. Additional pp collisions are superimposed on the main event to simulate the PU particle environment. Then, the CMS detector response simulation is carried out using GEANT4 [117]. Lastly, the CMS event reconstruction is performed as if the simulated events were data.

Several corrections are applied to the simulated events to better model the recorded data. Object-level corrections correspond to jet energy scale and resolution corrections, and to b-tag efficiency scale factor. An additional event-level correction is applied to the simulated events to better model the pileup distribution measured in the data. An event weight is applied and corresponds to the ratio of the normalized distributions of pileup interactions in simulation and in data. The pileup weights are normalized to a total sum of one for every sample processed, ensuring that the total normalization of each simulated process is not altered. A weight to correct for differences in the trigger efficiency between data and simulation is applied (see Section 3.3).

#### 4.1.3 HH Signal Model

A generalized signal HH model is used to compute the total cross section and kinematic distributions as function of arbitrary values of ggF and VBF couplings. Moreover, a dedicated

ggF EFT signal model is implemented. The HH model initially developed in the context of the analysis presented here is used in all the other non-resonant HH analyses prepared or under preparation by the CMS Collaboration.

For the ggF signal, a linear combination of three NLO ggF signal samples is used to model the total and differential cross sections (i.e. kinematic variables) for an arbitrary  $\kappa_\lambda$  and  $\kappa_t$  coupling values. The linear combination coefficients are computed using polynomial functions of the  $\kappa_\lambda$  and  $\kappa_t$  couplings derived from the known dependence on the production cross section. The concepts behind the method are described in Section E.1 of the Appendix E. The signal yields are scaled to match the cross section at NNLO accuracy taking into account the  $\kappa_\lambda$ -dependency, as illustrated in Figure 4-2 A). The optimal combination of ggF samples optimal is the one corresponding to  $\kappa_\lambda=1, 2.45$  and 5 samples.

The VBF signal modeling follows a similar approach to the one in the ggF signal modeling. However, the linear combination is carried out using six VBF simulated signals. The parametrization of the coefficients is further described in Section E.2 of the Appendix E. Using the expected LO cross section dependence on the coupling values and the SM N<sup>3</sup>LO prediction, the total signal yields are scaled to match the N<sup>3</sup>LO accuracy. The optimal choice of signal samples is presented in Table 4-1. The resulting VBF HH production cross section dependency on the three individual couplings is illustrated in Figure 4-2 B-D).

Table 4-1. Couplings variations used for the ggF and VBF signal models. The ggF (VBF) listed values of cross section times the branching fraction,  $\sigma(pp \rightarrow HH) \cdot B(HH \rightarrow bbbb)$ , are taken from the generator and scaled to the NNLOFTapprox (N<sup>3</sup>LO) precision by a factor  $k = \sigma_{\text{NNLOFTapprox}}^{\text{SM}} / \sigma_{\text{generator}}^{\text{SM}} = 1.16$  ( $k = \sigma_{\text{N}^3\text{LO}}^{\text{SM}} / \sigma_{\text{generator}}^{\text{SM}} = 1.04$ ).

| Model                                           | Coupling values | $\sigma(pp \rightarrow HH) \cdot B(HH \rightarrow bbbb)[\text{pb}]$ |
|-------------------------------------------------|-----------------|---------------------------------------------------------------------|
| ggF ( $\kappa_\lambda, \kappa_t$ )              | 1,1             | 0.010517                                                            |
| ggF ( $\kappa_\lambda, \kappa_t$ )              | 2.45,1          | 0.004455                                                            |
| ggF ( $\kappa_\lambda, \kappa_t$ )              | 5,1             | 0.031072                                                            |
| VBF ( $\kappa_\lambda, \kappa_V, \kappa_{2V}$ ) | 1,1,1           | 0.000585                                                            |
| VBF ( $\kappa_\lambda, \kappa_V, \kappa_{2V}$ ) | 1,1,0           | 0.009169                                                            |
| VBF ( $\kappa_\lambda, \kappa_V, \kappa_{2V}$ ) | 1,1,2           | 0.004823                                                            |
| VBF ( $\kappa_\lambda, \kappa_V, \kappa_{2V}$ ) | 2,1,1           | 0.000482                                                            |
| VBF ( $\kappa_\lambda, \kappa_V, \kappa_{2V}$ ) | 0,1,1           | 0.001558                                                            |
| VBF ( $\kappa_\lambda, \kappa_V, \kappa_{2V}$ ) | 1.1.5,1         | 0.022412                                                            |

An EFT ggF signal model is used to simulate the 12 benchmarks presented in Table 1-5 at NLO precision. First, each benchmark signal follows the recipe of the ggF LO modeling detailed in Section E.1 in Appendix E. Then, a reweighting procedure is applied to these events, in order to correct the  $m_{HH}$  spectrum from LO to NLO accuracy in QCD following Ref. [46].

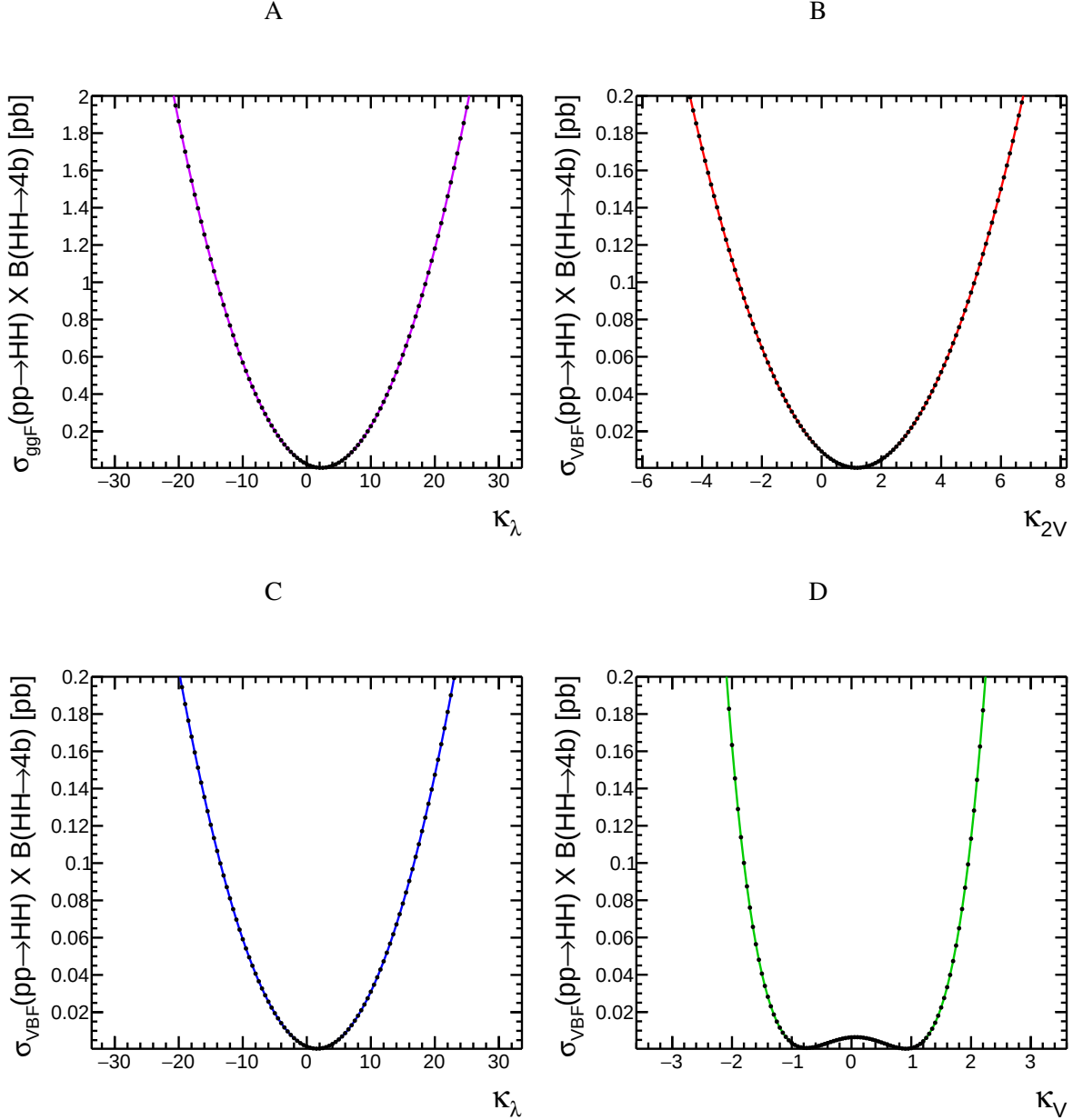


Figure 4-2. Higgs boson pair production cross section times the branching fraction into four bottom quarks as a function of the couplings using the signal modeling. A) ggF mode vs  $\kappa_\lambda$ , B) VBF mode vs  $\kappa_{2V}$ , C) VBF mode vs  $\kappa_\lambda$ , D) VBF mode vs  $\kappa_V$ .

## 4.2 Data-driven Background Model

### 4.2.1 Motivation

In LHC searches for  $HH \rightarrow bbbb$ , the dominant QCD multijet background events are not properly modeled by MC simulation. Consequently, a background model is typically derived using real collision data events. These kinds of methods are known as a ‘data-driven’ methods.

In the published CMS 2016 data-only version of this search [59], the background model strategy used a data-driven method. An artificial ensemble of background events (QCD,  $t\bar{t}$  and others) are produced by mixing event hemispheres from real ‘4b’ data events using a method called the ‘Hemisphere-mixing’. This technique has a good modeling when tested in QCD MC events. However, when applied to the data (containing other minor backgrounds), a systematic bias on the background shape is observed. The background shape and normalization uncertainties can affect the analysis sensitivity by 30 and 9%, respectively.

The published ATLAS 2015-2016 data-only analysis uses a background model that combines a QCD multijet data-driven estimation and  $t\bar{t}$  simulation [55]. The ATLAS data-driven multijet model part takes events from a ‘2b’ region to model ‘4b’ region events. There are kinematic differences between the two datasets, which are corrected with event-by-event correction weights or reweighting procedure. The weights are derived using an iterative method involving several 1D histograms. This method performs well within the uncertainties. However, being an iterative method based on 1D reweightings, it is hard to use it to model a multivariate discriminant that we can use to boost the signal separation from the background.

The background model presented in this dissertation is developed for the Run-2 result, replacing the one in the previous CMS 2016-only result. It uses a machine learning algorithm proposed to solve the problem of simultaneously reweighting of multiple distributions called the ‘Boosted Decision Tree reweighting’ or BDT reweighting from here [124]. It is the first time that this method is applied in a CMS analysis. It has been applied in ATLAS [125, 126] and LHCb searches [127] in the past. The machine learning algorithm is described in Subsection 4.2.2, and its implementation in the data analysis is presented in Subsection 4.2.3.

### 4.2.2 Boosted Decision Tree Reweighting

In high energy physics is very common to study a ‘base’ dataset (e.g. MC simulated events) with discrepancies and imperfections among multiple variables with respect to a ‘target’ dataset (e.g. collected data). A reweighting procedure is commonly introduced to improve the variables’ agreement. Mathematically, a reweighting problem between two datasets is equivalent to find the density ratio function  $\rho(x) = f_{\text{target}}(x)/f_{\text{base}}(x)$ , where  $x$  corresponds to an  $N$ -dimensional vector of the variables involved in the reweighting.

One approach to determine  $\rho(x)$  is the ‘Reweighting with bins’ method. In this method, the multivariable phase space is split into  $k$  bins, and then each bin of base events are multiplied by a factor  $\alpha_k = \frac{w_{k,\text{target}}}{w_{k,\text{base}}}$ , where  $w_{k,\text{base}}$  and  $w_{k,\text{target}}$  correspond to the total weight of events in the bin- $k$  in the base and target dataset. In practice, the ratio of histograms representing the densities of events in this space. However, although the technique is very intuitive and simple, it has the following limitations:

- Only a few variables can be reweighted simultaneously with good performance
- Choosing what variables to re-weight is not simple as one variable often brings disagreement to others
- It can give an unstable reweighting rule for bins of the input density histograms that are scarcely populated because of the large statistical fluctuations

A machine learning algorithm to solve this exact problem is proposed in Ref. [124]. In this algorithm, a decision tree is used to split the phase space into two regions (or leafs of the tree) with the largest event weight difference, i.e. find the regions that require more reweighting. To that end, the metric in equation 4-1 is maximized, where  $w_{\text{leaf},\text{base}}$  and  $w_{\text{leaf},\text{target}}$  correspond to the total weight of events in the leaf of the base and target dataset.

$$\chi^2 = \sum_{\text{leaf}} \frac{(w_{\text{leaf},\text{base}} - w_{\text{leaf},\text{target}})^2}{w_{\text{leaf},\text{base}} + w_{\text{leaf},\text{target}}} \quad (4-1)$$

The BDT reweighter consists of a forest of  $N$  trees, each trained one by one with these steps:

1. Build the decision tree that maximizes the  $\chi^2$  function. It determines the cut that maximizes the  $\chi^2$  function for each variable. Then, the sample is sliced into two using the variable

with maximal  $\chi^2$ . For each resulting slice, the same procedure is repeated until defined criteria (maximum depth, minimum events in the leaf, etc.) is satisfied.

2. Compute the weight in a given leaf as  $\mu_{\text{leaf}} = \log \left( \frac{w_{\text{leaf,base}}}{w_{\text{leaf,target}}} \right)$
3. Before training a new tree, the current tree weights are applied to the base sample as  $w_{\text{tree}} = \exp(\lambda \mu_{\text{leaf}})$ , where  $\lambda$  is the learning rate.

In the last step, the event-by-event weight is determined as the product of the contributions from the individual trees (i.e. boosting),  $w = \prod_{i=1}^N w_i$ . By following this procedure, the BDT reweighter captures the differences between the base and the target datasets, taking into consideration multiple variables and their correlations. This method is expected to have better performance than algorithms such as bin-reweighting, re-used general-purpose classification scores based on boosted decision trees (BDTs) or artificial neural networks (ANNs).

### 4.2.3 Implementation

#### 4.2.3.1 Overview

The derivation of the data-driven background model begins selecting real data events from a high statistics region orthogonal to the ‘4b’ real data. As seen in Subsection 3.6.1, the ‘3b’ region has around 7-10 times more data events than the ‘4b’ region. In consequence, events from this region are chosen as the ‘base’ dataset to model the background in the ‘target’ 4b region. If CMS had a ‘2b’ trigger during the data-taking, we could exploit data from an additional 2b region for the background modeling.

The 3b and 4b data events are expected to differ in event yields and kinematic properties. The latter is due to differences in b-tagging efficiency and jet activity, among other effects. Furthermore, a multivariate model is needed to accurately and simultaneously model the input variables of the MVAs used in the ggF categories (described in Subsection 3.7.2). Consequently, a multidimensional transfer model is built to transfer the event yield and event distributions from the 3b region to the 4b region. To this end, the analysis signal ( $A_{\text{SR}}$ ) region and analysis control ( $A_{\text{CR}}$ ) region defined in Subsection 3.6.2 are note. Note that  $A_{\text{SR}}(4b)$  is the target region for the



search because it has the best signal-to-background ratio. The  $A_{SR}(4b)$  background model is derived using information from the  $A_{CR}(3b)$ ,  $A_{CR}(4b)$  and  $A_{SR}(3b)$  regions.

The multidimensional transfer model is prepared using  $A_{CR}$  data events and then is applied to the  $A_{SR}(3b)$  events to derive the  $A_{SR}(4b)$  background estimate. It addresses the background shape and events yields separately. The background normalization is imposed in the background model template. It is computed correcting the  $A_{SR}(3b)$  normalization, taking into consideration the yield differences seen in  $A_{CR}(3b)$  and  $A_{CR}(4b)$  data. For this purpose, a 3b-to-4b normalization transfer factor ( $\alpha$ ) is derived. It can be a constant or a function of a sensitive variable. The shape component corrects any residual differences between 3b and 4b data on important analysis variables. To this end, a BDT reweighter is trained to learn the differences between  $A_{CR}(3b)$  and  $A_{CR}(4b)$  data events. It provides a reweighting function to reshape the variables from  $A_{SR}(3b)$  events to better model  $A_{SR}(4b)$  ones.

The background model method is fully validated in a separate region, where both the validation signal ( $V_{SR}$ ) and control ( $V_{CR}$ ) regions are redefined to be incompatible with the SM  $m_H$  mass hypothesis. The requirements of the validation region are presented in Subsection 3.6.2. The validation tests are detailed in Subsection 4.2.4. Moreover, extra tests in Subsection 4.2.5 check the impact of SM HH signal and single Higgs production contamination in the modeling.

Each event subcategory has its own  $A_{SR}(4b)$  background estimate. Note that as the VBF category 2 signal extraction is defined as a single counting-experiment (see more in Chapter 5), no shape correction is needed. To avoid any bias on the background model, the  $A_{SR}(4b)$  region data are ‘blinded’ until the method is finalized, reviewed, and the data unblinding is approved by CMS.

#### 4.2.3.2 Background normalization

The normalization transfer factor  $\alpha$  can be calculated by taking the ratio of number of events in the  $A_{CR}(4b)$  and  $A_{CR}(3b)$  regions. Then, it is applied to the  $A_{SR}(3b)$  normalization to predict the  $A_{SR}(4b)$  background normalization using the following equation:

$$N(A_{SR}^{4b})_{\alpha} = \frac{N(A_{CR}^{4b})}{N(A_{CR}^{3b})} \cdot N(A_{SR}^{3b}) = \alpha \cdot N(A_{SR}^{3b}) \quad (4-2)$$

The previous method assumes that  $\alpha$  is a constant across the  $(m_{H_1}, m_{H_2})$  plane. Within the  $(m_{H_1}, m_{H_2})$  plane, events are naturally described by two quantities defined by the projections of each point of the plane along the axes parallel and perpendicular to the diagonal used for the reconstruction of the two  $m_{H_1}$  candidates ( $m_{H_1} = 1.04 \times m_{H_2}$ ). We refer to these two quantities as the “parallel” and “perpendicular” masses, and denote them with  $m_{\text{paral}}$  and  $m_{\text{perp}}$ , respectively. Events have a similar kinematics as function of  $m_{\text{perp}}$ , since an increase in the mass of  $m_{H_1}$  is compensated by a reduction of the mass of  $m_{H_2}$ . Inversely, smaller and larger values of  $m_{\text{paral}}$  correlate with softer and harder event kinematics, impacting the value of the  $\alpha$  transfer factor. Since the number of events in the 3b region is not flat as a function of  $m_{\text{paral}}$ , this can bias the determination of the background normalization. The effect magnitude is illustrated in Fig. 4-3.

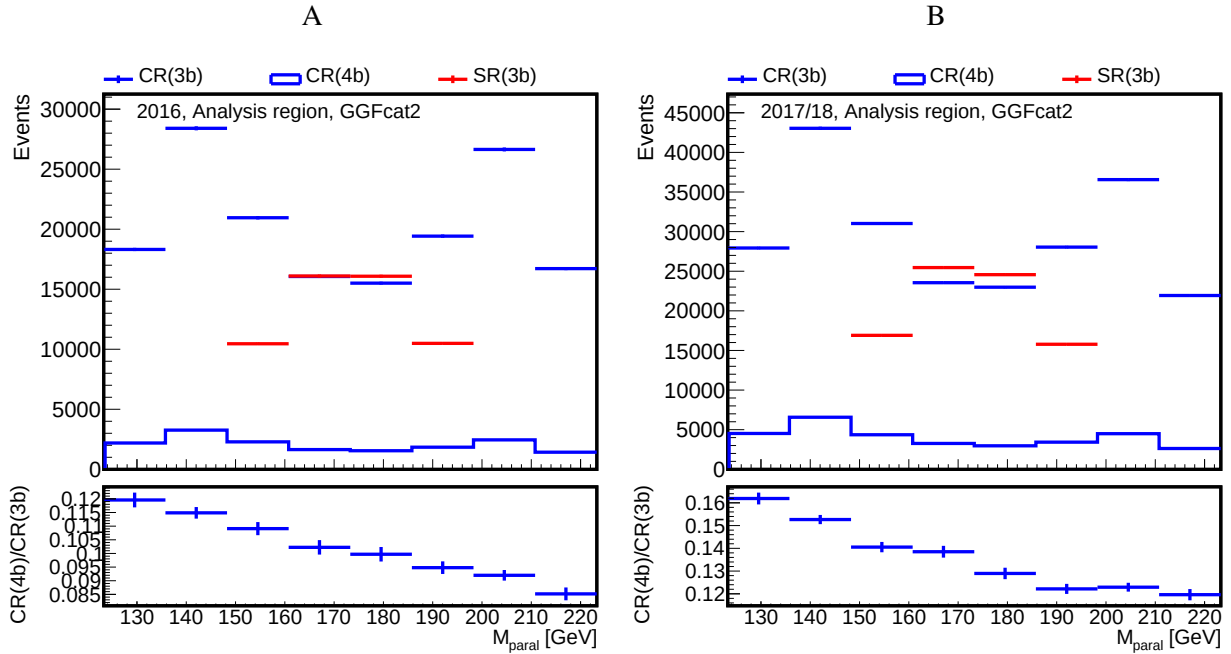


Figure 4-3. Dependence of the normalization transfer factor on the  $m_{\parallel}$  variable, illustrated for the ggF category 2 in A) 2016 data and B) 2017-2018 data.

For ggF subcategories, the adopted strategy is to first divide the CR data in eight bins of  $m_{\parallel}$ , and compute the value of  $\alpha$  in each of them. Then, this dependency is used on  $A_{\text{SR}}(3b)$  events to derive an alternative estimation of the SR(4b) normalization following equation 4-3. For VBF subcategories, however, the number of events is too small and possible trends are covered by the

statistical fluctuations of data and by their associated uncertainties. Consequently, for this case, the strategy is to apply the constant transfer factor following equation 4-2.

$$N(A_{SR}^{4b})_{\alpha,\parallel} = \sum_{ibin} \left( \frac{N(A_{CR}^{4b})}{N(A_{CR}^{3b})} \right)_{ibin} \cdot \left( N(A_{SR}^{3b}) \right)_{ibin} = \sum_{ibin} (\alpha)_{ibin} \cdot \left( N(A_{SR}^{3b}) \right)_{ibin} \quad (4-3)$$

In Tables 4-2 and 4-3 are shown the values of the SR(4b) background normalization predictions  $N(A_{SR}^{4b})_{pred}$  in all categories for the 2016 and 2017-2018 analyses, respectively. Note that there are two background normalization uncertainty sources of statistical nature: the one associated to the SR(3b) data statistical power, and the one associated to the CR statistical power.

Table 4-2. Summary of the quantities involved in the estimation of the background normalization in the 2016 data analysis. The prediction used in the background model is denoted as  $N(A_{SR}^{4b})_{pred}$ .

| Subcategory                         | ggF Cat1          | ggF Cat2         | VBF Cat1        | VBF Cat2        |
|-------------------------------------|-------------------|------------------|-----------------|-----------------|
| $N(A_{CR}^{3b})$                    | 94464±307.3       | 162022±402.5     | 7823±88.4       | 103±10.1        |
| $N(A_{SR}^{3b})$                    | 34594±186.0       | 53129±230.5      | 2991±54.7       | 38± 6.2         |
| $N(A_{CR}^{4b})$                    | 9860± 99.3        | 16643±129.0      | 1151±33.9       | 11± 3.3         |
| $\alpha$                            | 0.10437±0.00110   | 0.10272±0.00084  | 0.14713±0.0046  | 0.10680±0.03388 |
| $N(A_{SR}^{4b})_{\alpha}$           | 3610.9 ±24.8±38.1 | 5457.4±30.5±44.6 | 440.1 ±8.9±13.8 | 4.1±0.7±1.3     |
| $N(A_{SR}^{4b})_{\alpha,\parallel}$ | 3394.8 ±23.3±59.5 | 5384.5±30.1±69.8 | -               | -               |
| $N(A_{SR}^{4b})_{pred}$             | 3394.8 ±23.3±59.5 | 5384.5±30.1±69.8 | 440.1 ±8.9±13.8 | 4.1±0.7±1.3     |

Table 4-3. Summary of the quantities involved in the estimation of the background normalization in the 2017-2018 data analysis. The normalization used in the background model is denoted as  $N(A_{SR}^{4b})_{pred}$ .

| Subcategory                         | ggF Cat1         | ggF Cat2           | ggF Cat1        | ggF Cat2        |
|-------------------------------------|------------------|--------------------|-----------------|-----------------|
| $N(A_{CR}^{3b})$                    | 101819±319.1     | 235074±484.8       | 13533±116.3     | 233±15.3        |
| $N(A_{SR}^{3b})$                    | 38802±197.0      | 82715±287.6        | 5162± 71.8      | 119±10.9        |
| $N(A_{CR}^{4b})$                    | 15787±125.6      | 32222±179.5        | 2540± 50.4      | 37±6.1          |
| $\alpha$                            | 0.15505±0.00133  | 0.13707±0.00081    | 0.18768±0.00406 | 0.15880±0.02810 |
| $N(A_{SR}^{4b})_{\alpha}$           | 6016.2±37.5±51.6 | 11337.9±49.6±67.0  | 968.9±14.7±21.0 | 18.9±1.7±3.3    |
| $N(A_{SR}^{4b})_{\alpha,\parallel}$ | 5748.6±35.8±81.4 | 10998.2±48.2±104.2 | -               | -               |
| $N(A_{SR}^{4b})_{pred}$             | 5748.6±35.8±81.4 | 10998.2±48.2±104.2 | 968.9±14.7±21.0 | 18.9±1.7±3.3    |

#### 4.2.3.3 Background shape

The background normalization procedure presented in Subsection 4.2.3.2 addresses the normalization of the SR(4b) background estimate. However, as explained in Subsection 4.2.3.1, there are residual differences between the ‘3b’ and ‘4b’ data that need to be corrected in the

SR(3b) distributions to have an accurate description of the important variables of the analysis. They are corrected through a re-shaping or reweighting procedure using event-by-event weights obtained from a BDT reweighter.

The implementation of the BDT-reweighting method available in the *hep\_ml* toolkit. The BDT reweighter is a multidimensional re-weighting rule, which training configuration is very similar to the one of any general-purpose BDT score. Therefore, the definition of the datasets, input variables, and hyperparameters is needed. The  $A_{CR}(3b)$  and  $A_{CR}(4b)$  data events are defined as the ‘model’ and ‘target’ datasets, respectively.

In ggF categories 1 and 2, the training variables are the ones used in the BDT output used for signal extraction in collision data which is further explained in Chapter 5. In the VBF category 1, the training variables are kinematic properties of the reconstructed Higgs bosons, the VBF-jet system and the ggFKiller. In addition, all categories include the regressed  $p_T$  of the four b jets as training variables. The complete list of BDT-reweighter input variables are summarized in the Table 4-4 and Table 4-5 for ggF categories and VBF category 1, respectively. In order to obtain unbiased weights, the training is performed using a 2-fold cross-validation procedure also implemented in the python *hep\_ml* toolkit. The re-shaping event-by-event weight is the average output of the two trained BDT-reweighters.

A grid of points is tested for the hyperparameter optimization of each individual BDT-reweighter. The point contains a variation of the following hyperparameters: number of trees, maximum depth, minimum leaf samples, and learning rate. For each point of the grid, a BDT reweighter is trained, and then the re-weighting is tested by a two-step improvement check.

In the first check, just like in Ref. [124], the Kolmogov-Smirnov (K-S) distance (or called statistic) between  $A_{CR}(3b)$  and  $A_{CR}(4b)$  datasets are computed for each variable and compared before and after the BDT-reweighting is applied. Then, we check that the K-S distances improve in all variables, i.e. the K-S distance is smaller after reweighting. In the second check, we quantify how good the reweighted/model events match the target events. This task can be challenging to quantify simultaneously in multiple variables. However, to this end, we train a

Table 4-4. BDT reweighter input variables in ggF Category 1 and ggF Category 2.  $H_1$  is the highest- $p_T$  Higgs candidate and  $H_2$  is the lowest- $p_T$  Higgs candidate.

| BDT Reweighter Input variables                                                     |
|------------------------------------------------------------------------------------|
| Regressed $p_T$ of the leading- $p_T$ b jet of the $H_1$ candidate                 |
| Regressed $p_T$ of the trailing- $p_T$ b jet of the $H_1$ candidate                |
| Regressed $p_T$ of the leading- $p_T$ b jet of the $H_2$ candidate                 |
| Regressed $p_T$ of the trailing- $p_T$ b jet of the $H_2$ candidate                |
| Mass of the $H_1$ candidate, $M(H_1)$                                              |
| Mass of the $H_2$ candidate, $M(H_2)$                                              |
| Mass of the Higgs pair system, $m_{HH}$                                            |
| Transverse momentum of the $H_1$ candidate, $P_T(H_1)$                             |
| Transverse momentum of the $H_2$ candidate, $P_T(H_2)$                             |
| Pseudorapidity separation between the two Higgs candidates, $\Delta\eta(H_1, H_2)$ |
| $\Delta R$ distance between two b jets of the $H_1$ candidate, $\Delta R(H_1(bb))$ |
| $\Delta R$ distance between two b jets of the $H_2$ candidate, $\Delta R(H_2(bb))$ |
| $ \cos(\theta)^*(H) $ in HH frame                                                  |
| $ \cos(\theta)^*(b) $ in $H_1$ frame                                               |
| Sum of four b jets' regressed $p_T$                                                |
| Transverse momentum of the HH system, $p_T(HH)$                                    |
| Number of tight b-tags in the 3 highest b-tagged jets                              |
| The sum of resolution scores of the 3 highest b-tagged jets                        |
| Minimal $\Delta R$ distance between two b jets, $\text{Min} \Delta R(bb) $         |
| Maximum pseudorapidity separation between two b jets, $\text{Max} \Delta\eta(bb) $ |

Table 4-5. BDT reweighter input variables in VBF Category 1.  $H_1$  is the highest- $p_T$  Higgs candidate and  $H_2$  is the lowest- $p_T$  Higgs candidate.

| BDT Reweighter Input variables                                                      |
|-------------------------------------------------------------------------------------|
| Regressed $p_T$ of the leading- $p_T$ b jet of the $H_1$ candidate                  |
| Regressed $p_T$ of the trailing- $p_T$ b jet of the $H_1$ candidate                 |
| Regressed $p_T$ of the leading- $p_T$ b jet of the $H_2$ candidate                  |
| Regressed $p_T$ of the trailing- $p_T$ b jet of the $H_2$ candidate                 |
| Mass of the $H_1$ candidate, $M(H_1)$                                               |
| Mass of the $H_2$ candidate, $M(H_2)$                                               |
| Mass of the Higgs pair system, $m_{HH}$                                             |
| Transverse momentum of the $H_1$ candidate, $P_T(H_1)$                              |
| Transverse momentum of the $H_2$ candidate, $P_T(H_2)$                              |
| Pseudorapidity separation between the two Higgs candidates, $\Delta\eta(H_1, H_2)$  |
| Azimuthal angle separation between the two Higgs candidates, $\Delta\phi(H_1, H_2)$ |
| Mass of the VBF-jet pair system, $M(jj)$                                            |
| Pseudorapidity separation between the two VBF jets, $\Delta\eta(j1, j2)$            |
| ggFKiller score                                                                     |

general-purpose BDT discriminant score by a two-fold cross validation using scikit-learn [128]. The discriminant aims at predicting if a given event is from the model dataset or the target dataset.

If the discriminant can not tell the difference between the model and target dataset, i.e. performs similar to a random guess or  $\text{ROC-AUC} \sim 0.5$ , then the hyperparameters are good for the BDT-reweighting procedure. If multiple grid points are found to satisfy the aforementioned checks, then the chosen point is the one with ROC-AUC closest to 0.5. As an illustration, the classifier check of the 2017-2018 ggF category 2 is in Figure 4-4. The optimized hyperparameters for all categories and years are presented in Tables F-1 and F-2 in Appendix F.

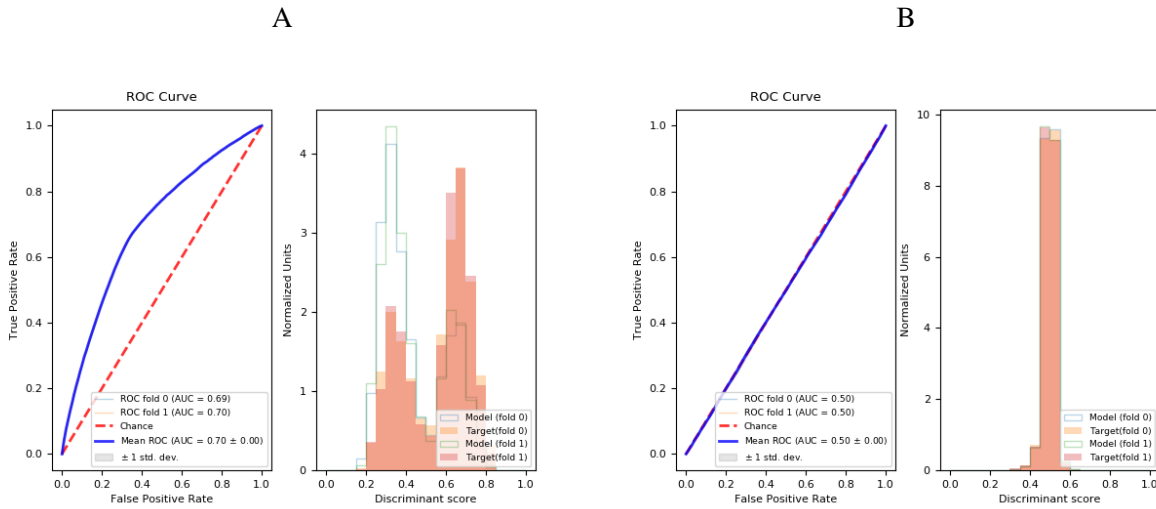


Figure 4-4. Example of the discriminant check step before A) and after B) BDT reweighter weights are applied in the model in control region data. The model corresponds to the one used for the 2017-2018 ggF category 2.

Once the BDT-reweighters are trained, their associated weights are applied to re-shape the  $A_{\text{CR}}(3b)$  distributions and compared to the data  $A_{\text{CR}}(4b)$  as a closure test. An example of the closure achieved by the BDT reweighting in the control regions can be seen in Figure 4-5. A good closure is observed between data and the background model is observed in all input variables. The distributions are presented in the following figures in Appendix F: Figures F-1 to F-4 for the ggF categories 1 and 2, and the Figures F-5 to F-6 for the VBF category 1.

The derived transfer factors and reweighting rules are applied in the  $A_{SR}(3b)$  data to estimate the  $A_{SR}(4b)$  background in the observables used for the signal extraction. It is important to emphasize that the methods presented here are optimized and settled without looking at the  $A_{SR}(4b)$  data, i.e. a so-called blind analysis is performed. The methods followed a procedure of review within the CMS Collaboration before using the  $A_{SR}(4b)$  data to derive the results. Before that, the method is validated in the validation region studies (see Subsection 4.2.4).

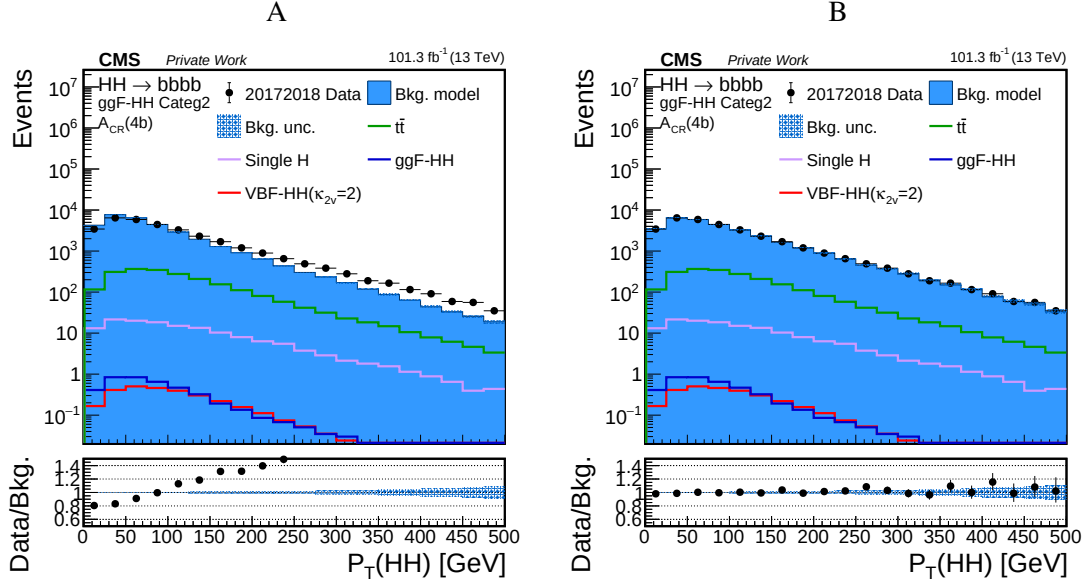


Figure 4-5. Example of data versus background modeling agreement without A) and with B) BDT reweighting in the analysis control regions data of the 20172018 ggF category 2.

#### 4.2.4 Validation Studies

The background model technique is fully validated by studying it in a region orthogonal to the  $A_{SR}$  region and with a comparable statistical power. In essence, a transfer model is derived using the validation control regions defined in Subsection 3.6.2:  $V_{CR}(3b)$  and  $V_{CR}(4b)$ . Then, the transfer model is applied to the  $V_{SR}(3b)$  data to model the background in the  $V_{SR}(4b)$  data. First, the validation produce consists in checking the closure of the background normalization and reweighting with the observed data in the  $V_{SR}(4b)$  region. Moreover, the closure of all observables used for the signal extraction is tested.

#### 4.2.4.1 Normalization validation test

Tables 4-6 and 4-7 present the observed data and expected background normalization values for all categories in the 2016 and 2017-2018 data. The sources of the  $N(V_{SR}^{4b})_{pred}$  uncertainty are: the statistical uncertainty from the SR(3b) data and the normalization transfer factor  $\alpha$  from the control regions, respectively. One can see that the transfer factor model is able to predict correctly the normalization within the uncertainties. There are some categories with a residual difference at the order of up to two standard deviations, which are considered as a systematic uncertainty during the statistical analysis in the  $A_{SR}(4b)$  region.

Table 4-6. Summary of the transfer factor and normalization estimation in the 2016 validation study. The normalization used in the background model is denoted as  $N(V_{SR}^{4b})_{pred}$ .

| Subcategory                  | ggF Cat1                   | ggF Cat2                   | VBF Cat1                 | VBF Cat2              |
|------------------------------|----------------------------|----------------------------|--------------------------|-----------------------|
| $N(V_{CR}^{3b})$             | $58203 \pm 241.3$          | $190979 \pm 437.0$         | $6146 \pm 78.4$          | $90 \pm 9.5$          |
| $N(V_{SR}^{3b})$             | $30703 \pm 175.2$          | $70566 \pm 265.6$          | $2257 \pm 47.5$          | $30 \pm 5.5$          |
| $N(V_{CR}^{4b})$             | $6108 \pm 78.2$            | $16775 \pm 129.5$          | $857 \pm 29.3$           | $9 \pm 3.0$           |
| $N(V_{SR}^{4b})$             | $3349 \pm 57.9$            | $6062 \pm 77.8$            | $317 \pm 17.8$           | $1 \pm 1.0$           |
| $\alpha$                     | $0.10494 \pm 0.00141$      | $0.08784 \pm 0.0007$       | $0.13944 \pm 0.00508$    | $0.1 \pm 0.03496$     |
| $N(V_{SR}^{4b})_{\alpha}$    | $3222.1 \pm 23.6 \pm 43.3$ | $6198.3 \pm 29.7 \pm 49.4$ | $314.7 \pm 6.8 \pm 11.5$ | $3.0 \pm 0.5 \pm 1.0$ |
| $N(V_{SR}^{4b})_{\alpha,  }$ | $3207.6 \pm 23.5 \pm 73.1$ | $6088.8 \pm 29.2 \pm 77.4$ | -                        | -                     |
| $N(V_{SR}^{4b})_{pred}$      | $3207.6 \pm 23.5 \pm 73.1$ | $6088.8 \pm 29.2 \pm 77.4$ | $314.7 \pm 6.8 \pm 11.5$ | $3.0 \pm 0.5 \pm 1.0$ |

Table 4-7. Summary of the transfer factor and normalization estimation in the 2017-2018 validation study. The normalization used in the background model is denoted as  $N(A_{SR}^{4b})_{pred}$ .

| Subcategory                  | ggF Cat1                   | ggF Cat2                    | VBF Cat1                  | VBF Cat2              |
|------------------------------|----------------------------|-----------------------------|---------------------------|-----------------------|
| $N(V_{CR}^{3b})$             | $72379 \pm 269.0$          | $260855 \pm 510.7$          | $11194 \pm 105.8$         | $192 \pm 13.9$        |
| $N(V_{SR}^{3b})$             | $38387 \pm 195.9$          | $88771 \pm 297.9$           | $3980 \pm 63.1$           | $51 \pm 7.1$          |
| $N(V_{CR}^{4b})$             | $10967 \pm 104.7$          | $29842 \pm 172.7$           | $1864 \pm 43.2$           | $23 \pm 4.8$          |
| $N(V_{SR}^{4b})$             | $5981 \pm 77.3$            | $10254 \pm 101.3$           | $670 \pm 25.9$            | $4 \pm 2.0$           |
| $\alpha$                     | $0.15152 \pm 0.00155$      | $0.11440 \pm 0.00070$       | $0.16652 \pm 0.00417$     | $0.11979 \pm 0.02643$ |
| $N(V_{SR}^{4b})_{\alpha}$    | $5816.4 \pm 36.4 \pm 59.5$ | $10155.5 \pm 41.4 \pm 62.1$ | $662.7 \pm 11.1 \pm 16.6$ | $6.1 \pm 0.9 \pm 1.3$ |
| $N(V_{SR}^{4b})_{\alpha,  }$ | $5683.3 \pm 35.6 \pm 95.5$ | $10042.3 \pm 40.9 \pm 98.7$ | -                         | -                     |
| $N(V_{SR}^{4b})_{pred}$      | $5683.3 \pm 35.6 \pm 95.5$ | $10042.3 \pm 40.9 \pm 98.7$ | $662.7 \pm 11.1 \pm 16.6$ | $6.1 \pm 0.9 \pm 1.3$ |

#### 4.2.4.2 BDT reweighting validation test

This test aims at checking that the BDT reweighting method derived in control regions performs well when is applied to the signal region. To this end, BDT-reweighters are trained using the same procedure as in the analysis region, but instead using the  $V_{CR}$  information. An example



of the discriminant check of the validation training is shown in Figure 4-6 for the 2017-2018 ggF category 2. The optimized hyperparameters are presented in Tables F-3 and F-4 in Appendix F.

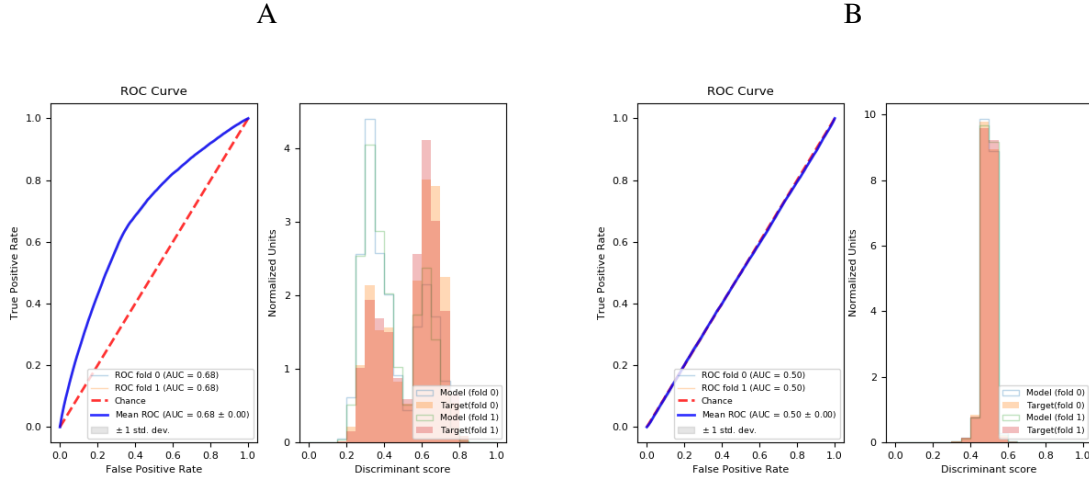


Figure 4-6. Example of the discriminant check step before A) and after B) BDT reweighter weights are applied in the model in validation control region data. The model corresponds to the one used for the 2017-2018 ggF category 2.

After the BDT-reweighters are trained, the weights are applied to the  $V_{CR}(3b)$  events to model  $V_{CR}(4b)$  ones. A very good agreement is observed between the data and the model in all variables involved in the  $V_{CR}$  reweighting. The reweighting variables distributions of data and the background model are illustrated for ggF categories in Figures F-7-F-10 and for VBF category in Figures F-11-F-12. The improvement achieved from the BDT-reweighting in the  $V_{CR}(4b)$  region is illustrated for one of the reweighting variables in Figure 4-7.

The BDT reweighting method is then checked by studying the application of the method in the  $V_{SR}$  region. A good agreement between the observed data and background model in all variables involved in the re-weighting is observed in the  $V_{SR}(4b)$  region. Figures F-13 to F-16 present the performance of our background estimate in the ggF  $V_{SR}$  regions, whereas Figures F-17 to F-18 present the performance of our background estimate in the VBF  $V_{SR}$  regions. The improvement with respect to the case without reweighting is illustrated in Figure 4-8. These results show that the data-driven technique is capable of accurately modeling the background in the signal regions of the analysis.

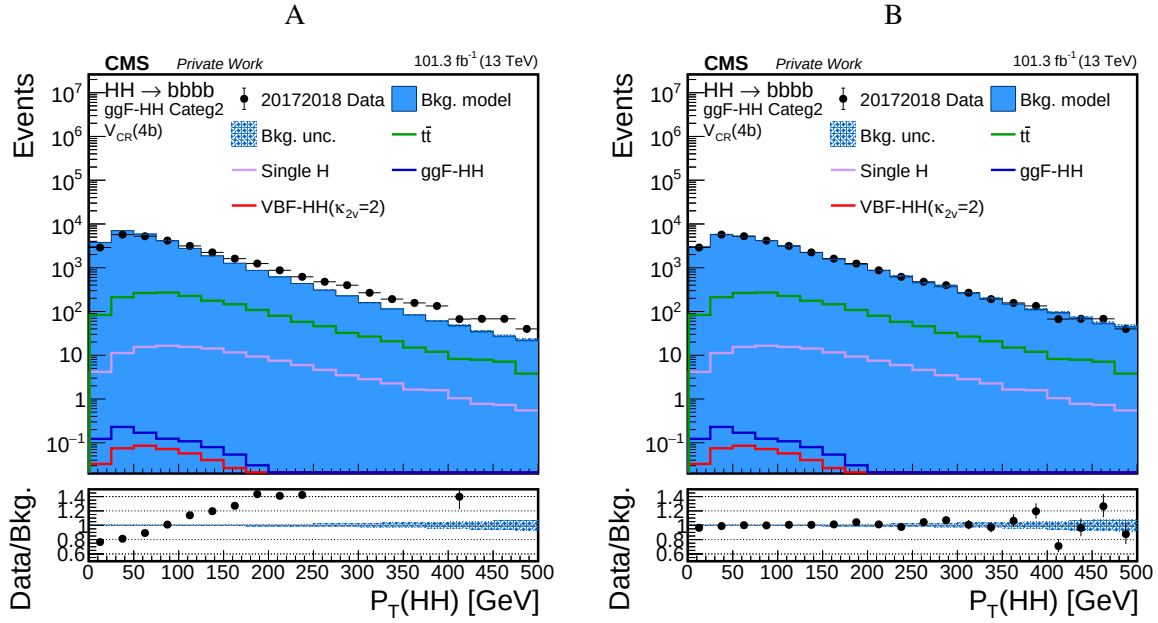


Figure 4-7. Example of data versus background modeling agreement A) without and B) with BDT reweighting in the validation control region of the 20172018 ggF category 2.

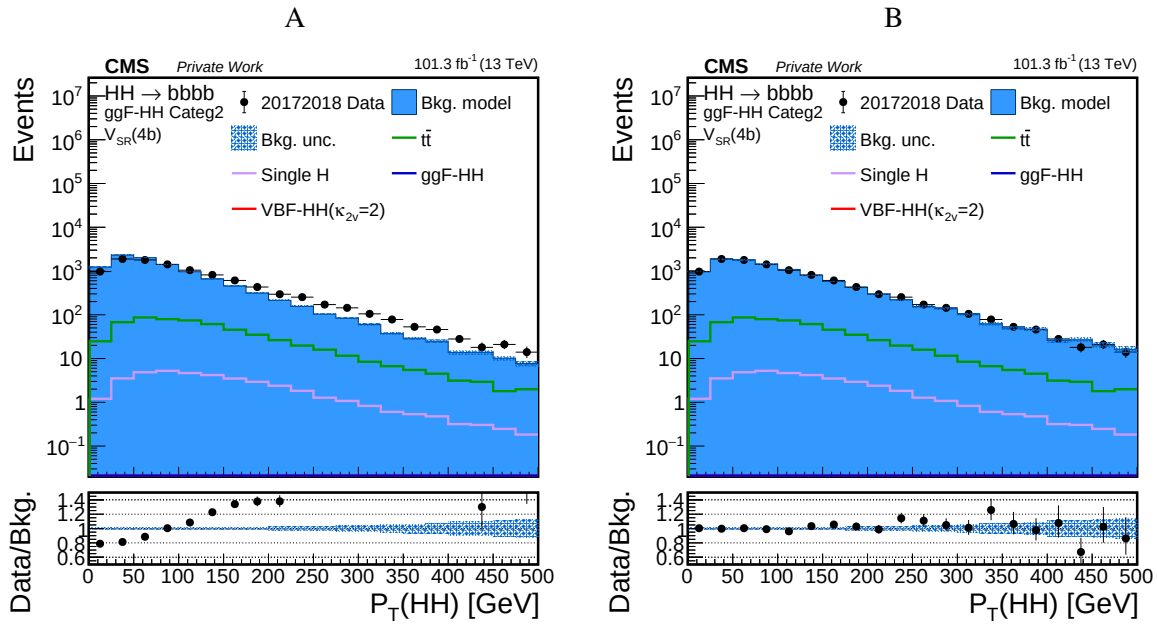


Figure 4-8. Example of data versus background modeling agreement without A) and with B) BDT reweighting in the validation signal region of the 2017-2018 ggF category 2.

#### 4.2.4.3 Statistical closure test in the validation region

As presented in Chapter 5, the signal extraction is performed in eight observables associated to four subcategories in two datasets. One crucial test is to confirm that the background estimate models well the data in all the observables. However, this is not possible to do this before  $A_{SR}(4b)$  unblinding approval. Consequently, this test is performed in the  $V_{SR}(4b)$  region.

The closure test consists on studying the observables by performing a background-only fit to the observed data by floating the background nuisance parameters and their uncertainties (more in 5). For each observable, the post-fit background model is compared to the observed data in order to check that there are not any striking trends pointing to a mismodeling. The resulting  $V_{SR}(4b)$  distributions are illustrated in Figure 4-9 and Figure 4-9 for 2016 and 2017-2018 data, respectively. A very good agreement is observed across all observables.

In addition, a goodness of fit test based on the saturated method [129] is performed in each observable to quantify the compatibility between the background model and observed data. A high compatibility ( $p\text{-value} > 5\%$ ) is observed across all observables of the  $V_{SR}(4b)$  region, as presented in Table 4-8. A similar test is carried out during the unblinding of the  $A_{SR}(4b)$  region observables, and further described in Section 5.1.2.

Table 4-8. Data-background model compatibility p-value obtained from the goodness of fit test applied to the observables in the  $V_{SR}(4b)$  regions.

| Dataset   | Subcategory    | Observable           | Goodness of fit test p-value |
|-----------|----------------|----------------------|------------------------------|
| 2016      | ggF category 1 | BDT output (GGFMVA1) | 42%                          |
| 2016      | ggF category 2 | BDT output (GGFMVA2) | 57%                          |
| 2016      | VBF category 1 | $m_{HH}$             | 55%                          |
| 2016      | VBF category 2 | Counting experiment  | 22%                          |
| 2017-2018 | ggF category 1 | BDT output (GGFMVA1) | 45%                          |
| 2017-2018 | ggF category 2 | BDT output (GGFMVA2) | 12%                          |
| 2017-2018 | VBF category 1 | $m_{HH}$             | 83%                          |
| 2017-2018 | VBF category 2 | Counting experiment  | 42%                          |
| Run-2     | All            | All                  | 53%                          |

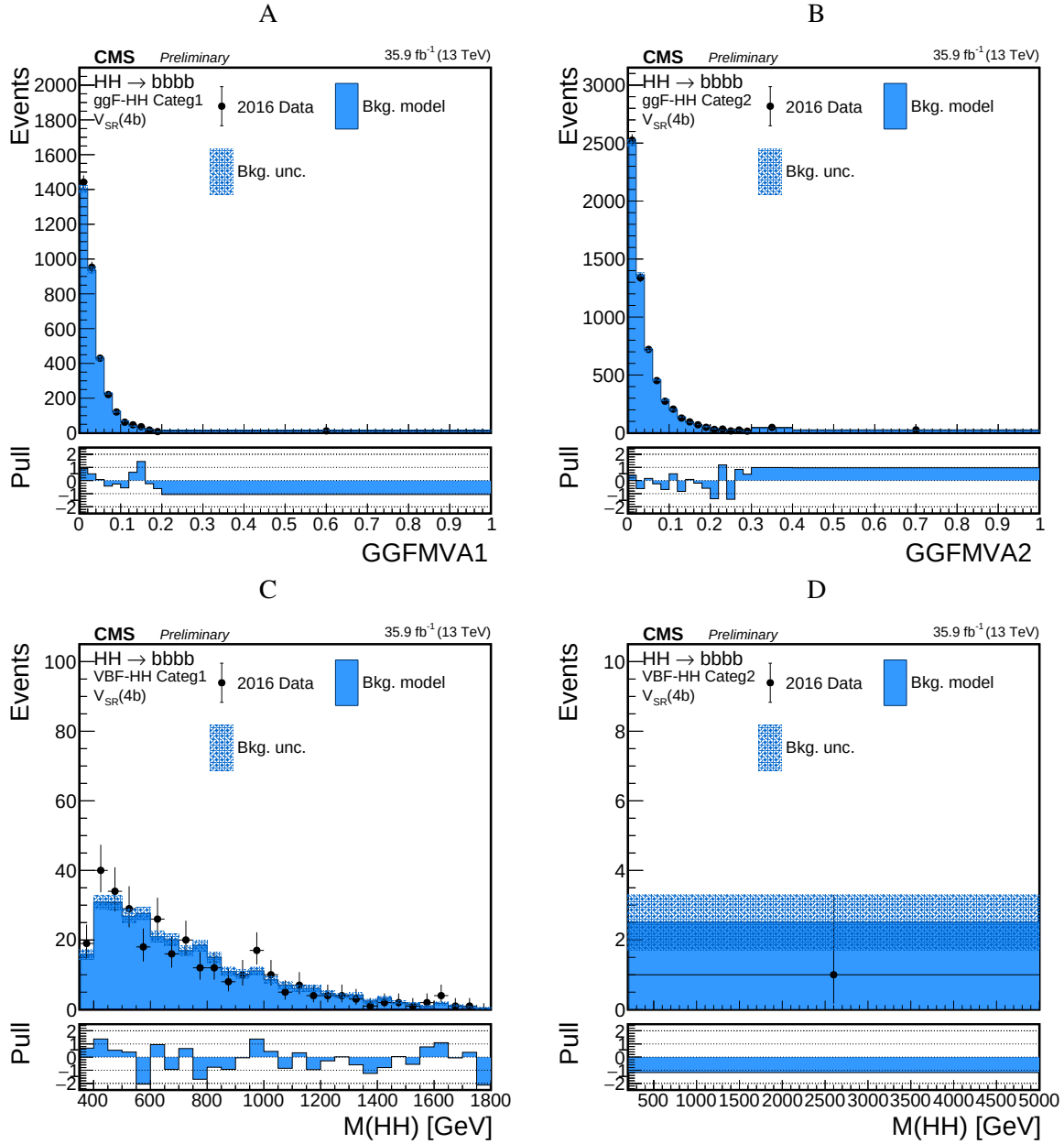


Figure 4-9. Distributions of the observables in all validation signal regions for 2016 data. A) ggF category 1, B) ggF category 2, C) VBF category 1, D) VBF Category 2. The presented background distributions are the one obtained after fitting the background prediction to the data floating the background uncertainties.

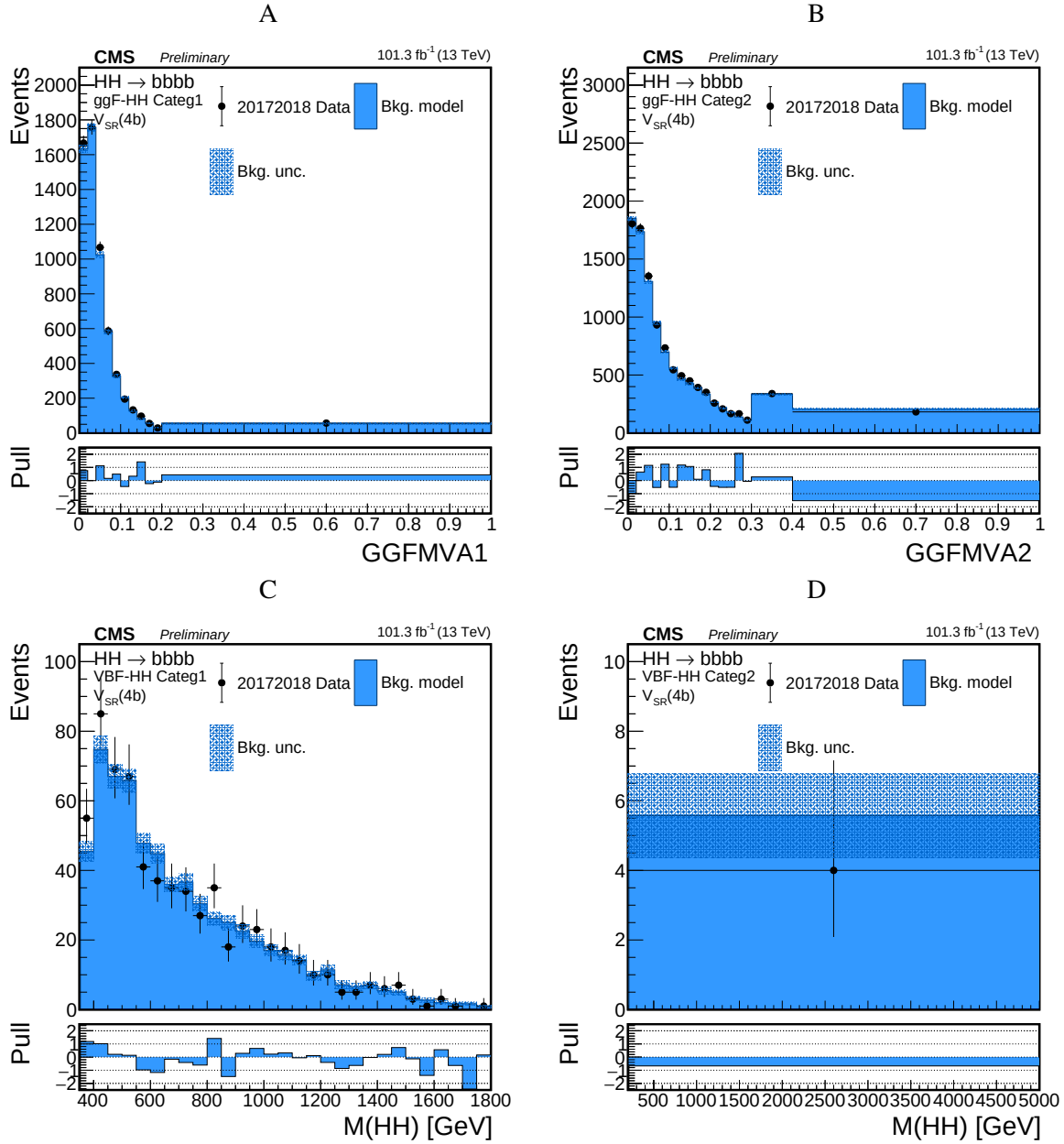


Figure 4-10. Distributions of the observables in all validation signal regions for 2017-2018 data. A) ggF category 1, B) ggF category 2, C) VBF category 1, D) VBF Category 2. The presented background distributions are the one obtained after fitting the background prediction to the data floating the background uncertainties.

## 4.2.5 Additional Checks

### 4.2.5.1 Self-bias test

The presence of signal in the 3b regions may impact the modeling of the background. Signal events failing the four b tagging requirement can be assigned to the part of the 3b region used to train the BDT reweighting method and to derive the background model. Given that these events are identified as background, but have the kinematic properties of the HH process, they can mimic the signal signature in the  $A_{SR}(4b)$ , introducing a bias on the background estimation, and thus reduce the sensitivity of the analysis.

A signal events classified in the  $A_{CR}(3b)$  typically result from an erroneous jet pairing, or from wrongly selection of an additional jet in the event that does not come from the  $HH \rightarrow b\bar{b}b\bar{b}$  process. These events consequently have different properties with respect to the ‘good’  $A_{SR}(4b)$  events because they have similar distributions to the background events in the  $A_{SR}(3b)$ , and thus have a negligible impact on the training of the background BDT reweighting. However, as events from the  $A_{SR}(3b)$  region are used to model the background template, the contamination of HH signal events in that region may introduce a spurious ‘signal-induced’ background component that can mask the actual signal presence in the  $SR(4b)$ . In order to understand the  $A_{SR}(3b)$  signal impact, a ‘self-bias’ test is performed. The test is run with toy experiments using, in each studied subcategory of the analysis separately, the following inputs:

- $b_{true}$  : The expected distribution of background events in the  $A_{SR}(4b)$ . It represents the ‘true’ background source
- $s$  : The signal distribution in the  $A_{SR}(4b)$  region
- $b_{sig}$  : The distribution of signal events in the  $A_{SR}(3b)$  corrected using the exact same transfer model used in the background modeling. It represents the signal events erroneously treated as background.

The self-bias test evaluates the bias induced on the best fitted signal strength ( $\hat{\mu}$ ) when a signal of strength  $\mu_{inj}$  is injected. In essence, this is done in three steps: (1) generate 2000 toys according to  $\mu_{inj} \cdot s + b_{true}$ , (2) fit each toy with  $\mu \cdot s + \mu_{inj} \cdot b_{sig} + b_{true}$ , (3) compare the average and pull of the best fit  $\hat{\mu}$  as function of  $\mu_{inj}$ . As a cross-check, the toys are fit with  $\mu \cdot s + b_{true}$ .

The test is carried out for values of  $\mu_{\text{inj}}$  up to approximately a factor of 4 times larger than the expected 95% CL upper limit on the SM ggF and VBF signal strengths obtained in two ggF categories and two VBF categories, respectively. A simple statistical model is used with a log-normal uncertainty of 5% on  $s$ , and a log-normal uncertainty of 5% on  $b_{\text{true}}$  and  $b_{\text{sig}}$ . The uncertainties have a negligible impact on the sensitivity, and the test shows the impact of the self-bias with essentially statistical uncertainties only. The bias impact on the full results (that includes systematic uncertainties) is consequently expected to be smaller.

The outcome of the test is illustrated for the 2016 analysis in Figures 4-11 and 4-12. They report the average value of  $\hat{\mu}$  in toys ( $\langle \hat{\mu} \rangle$ ), and the  $\langle \hat{\mu} \rangle$  pull with respect to  $\mu_{\text{inj}}$  normalized to the 1 standard deviation of the  $\hat{\mu}$  distribution ( $\sigma_{\hat{\mu}}$ ). A mild effect of the presence of  $b_{\text{sig}}$  can be observed, but the best-fit signal strength is well compatible within the expected one standard deviation of  $\hat{\mu}$  spread, to the one obtained without the  $b_{\text{sig}}$  term in the fit. Therefore, at the current level of sensitivity, the contamination of the signal in the background estimation regions has a negligible impact.

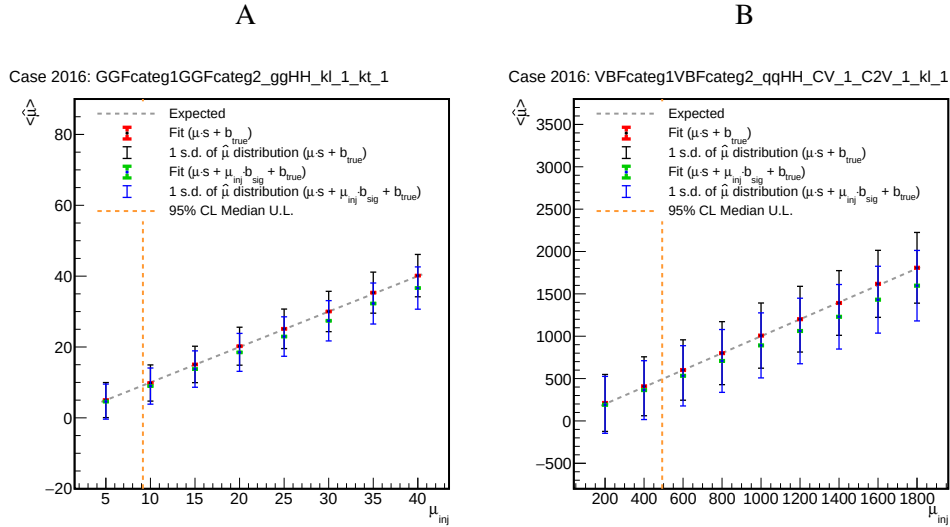


Figure 4-11. Example of the self-bias impact on fitted signal strength. A) ggF signal in the 2016 ggF categories and B) VBF signal in the 2016 VBF categories. The green (red) points and blue (black) error-bars show the average and one standard deviation of the best fit signal strength  $\hat{\mu}$  distribution when the fit is performed under the model with (without) the self-bias term  $b_{\text{sig}}$ . The vertical line denotes the 95% CL upper limit in that category.

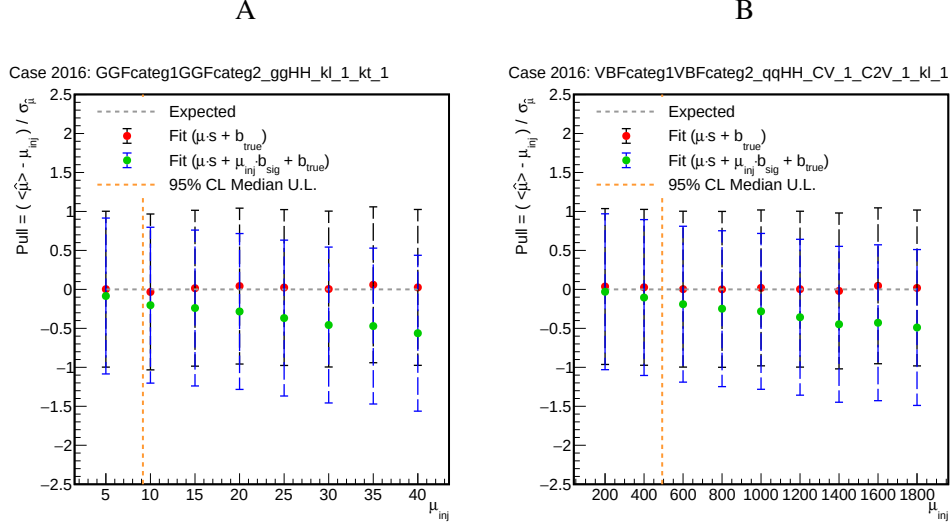


Figure 4-12. Example of the self-bias impact on the fitted signal strength pull. A) ggF signal in the 2016 ggF categories and B) VBF signal in the 2016 VBF categories. The points show the pull of the average of the best fit signal strength  $\hat{\mu}$  distribution with respect to the injected signal strength  $\mu_{\text{inj}}$ , normalized to one standard deviation of the  $\hat{\mu}$  distribution. The fit is performed under the model with (without) the self-bias term  $b_{\text{sig}}$  and shown in green (red). The vertical line denotes the 95% CL upper limit in that category.

#### 4.2.5.2 Single Higgs impact

The dominant single Higgs contribution comes from  $t\bar{t}H(b\bar{b})$  events (65%-95% of the total single H depending on the category and year). The contribution of single Higgs production to the background events is assumed to be covered by the background model. This hypothesis is studied using  $t\bar{t}H(b\bar{b})$  events in the  $A_{\text{SR}}(4b)$  region observables of the ggF categories in 2016 and 2017+2018.

The study evaluates the expected distribution of  $t\bar{t}H(b\bar{b})$  events, and how much of it is covered by our background model. The expected coverage is evaluated by transforming  $t\bar{t}H(b\bar{b})$  in the  $A_{\text{SR}}(3b)$  region with the background model method, here denominated as the ‘ $t\bar{t}H(b\bar{b})$  model’. The results are presented in Figure 4-13. The signal region distributions correspond to the background model statistical uncertainty (in blue), expected  $t\bar{t}H(b\bar{b})$  events (in green), expected ‘ $t\bar{t}H(b\bar{b})$  model’ (in orange), and the expected SM-HH signal scaled by expected upper limit (in red).



The outcome shows that the background estimate models partially the expected  $t\bar{t}H(b\bar{b})$  events. This effect is expected since we do not expect the method to learn to model  $t\bar{t}H(b\bar{b})$  in the CR because most of the  $t\bar{t}H$  background will be in the SR given the resonant  $H(b\bar{b})$  signature. However, this is not a concern because this contribution is overshadowed by the statistical uncertainty of the background estimate.

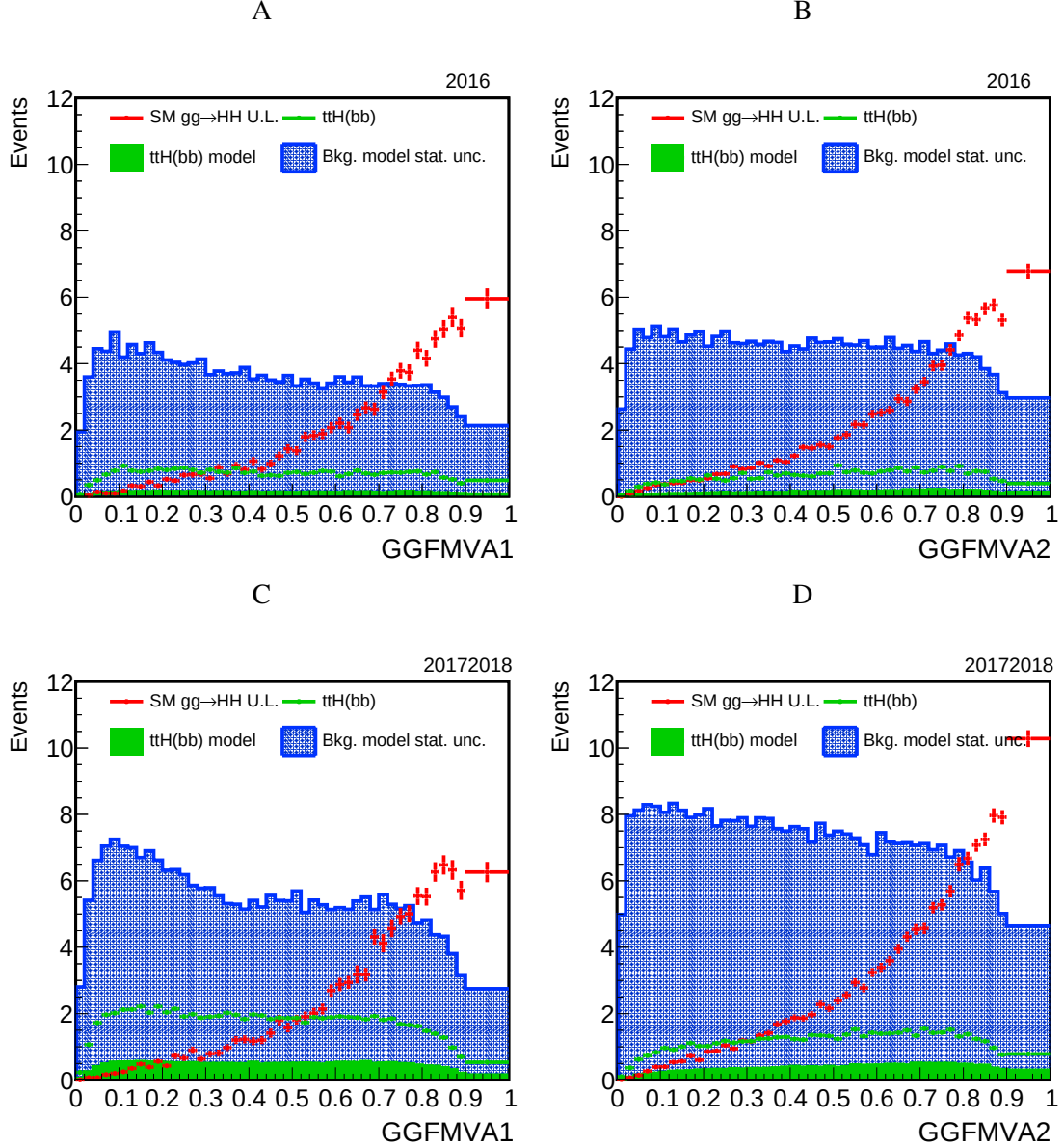


Figure 4-13. Impact of the Higgs boson production with associate top-quark pairs in the four observables of the ggF  $A_{SR}(4b)$  regions. A) ggF category 1 in 2016, B) ggF category 2 in 2016, C) ggF category 1 in 2017-2018, D) ggF category 2 in 2017-2018.

## 4.3 Systematic Uncertainties

### 4.3.1 Signal Uncertainties

The modeling of the signal yields and shape can be affected by sources of uncertainty associated to experimental measurements, event generation, and theoretical predictions. These systematic uncertainties are evaluated in all signals (6 ggF, 3 VBF, 12 EFT nodes) for the 2016, 2017 and 2018 simulation campaigns. The following sources are considered:

- **Luminosity:** As the signal yields are normalized according to the measured integrated luminosity, a normalization uncertainty is assigned to account for the uncertainty from the CMS luminosity measurements. The amounts are 1.2% for 2016 [130], 2.3% for 2017 [131], and 2.5% for 2018 [132].
- **Pileup reweighting:** A normalization uncertainty is derived with an alternative reweighting of the number of pileup interaction. This alternative reweighting is made by shifting the minimum bias cross section by 4.6%.
- **b-tagging algorithm:** Difference between data and simulation on the performance of the deepJet algorithm are addressed using scale factors provided by the CMS BTAG POG. Then, their impact on the signal modeling is evaluated by varying these factors within their uncertainties and propagating the effect in the analysis.
- **Trigger:** The efficiency of the simulated triggers is corrected to match the one measured in data using a correction scale factor using the procedure described in Section 3.3. The uncertainty on this procedure is derived by varying the factor within its uncertainties and propagating them on the full data analysis.
- **Jet energy scale (JES):** There are a total of 21 uncertainty sources in the jet energy scale, which are at the order of few percent and depend on  $p_T$  and  $\eta$ . The analysis uses the ‘reduced’ set recommended by the CMS JETMET POG, which roughly groups them by detector regions into a total of 11 sources. The impact of each uncertainty on the signal yield is calculated independently by varying the jet energy scale within the recommended uncertainties and rerunning the full analysis selection.
- **Jet energy resolution (JER):** The uncertainty on the jet energy resolution is taken into account by varying the jet resolution within the recommended uncertainties, and rerunning the full analysis selection. This uncertainty is considered as a signal shape uncertainty.
- **Level-1 Prefiring:** A gradual shift in the timing of the ECAL L1 trigger inputs at the forward region was observed during the 2016 and 2017 operations. This led to decrease of efficiency for events with a significant amount of ECAL energy in the  $2 < |\eta| < 3$  region. This effect is not simulated by default. and thus, correction factors and uncertainties were provided by trigger experts to account for it. Then, a signal normalization uncertainty associated to these corrections is derived by varying them within their uncertainties and propagating them to the final result.

- Event generator: The impact of the normalization and renormalization scales in the event generation are addressed by first propagating their corresponding variations (implemented as event-by-event weights) in the analysis. Then, the variations are combined by taking the envelope of them, and use the uncertainty.
- PDF: The uncertainty associated to the choice of PDF is addressed using variations of the PDF weights.
- Hadronization: Variations of the parton shower weights are used to assign an uncertainty on the modeling of the parton shower. For this purpose, the uncertainty is derived by combining four variations in quadrature.
- ISR recoil scheme The uncertainties in the modeling of additional radiation in VBF topologies is estimated by comparing samples generated with the `dipoleRecoil` PYTHIA option turned on and off. Nominal samples are generated without `dipoleRecoil` since this option was not yet available for the 2016 MC campaign. The difference in the yields of the VBF signals simulated with and without this option is considered as a systematic uncertainty, and the effect observed is a change in the sharing of the signal in the ggF and VBF categories corresponding to an uncertainty of 10-15% depending on the year and category considered. For the 2016 simulation, where no `dipoleRecoil` exists, the largest uncertainty between 2017 and 2018 is used.
- Theoretical HH cross section: These uncertainties correspond to  $+2.2\% / -5.0\%$  (scale),  $\pm 3\%$  (PDF), and  $\pm 2.6\%$  ( $m_t$ ) for ggF production, and to  $+0.03\% / -0.04\%$  (scale) and  $\pm 2.1\%$  (PDF +  $\alpha_s$ ) for VBF production. They are included only when quoting the limit with respect to the SM (i.e. signal strength  $\mu$ ) and for the likelihood scans, and not for the cross section upper limit. For anomalous  $\kappa_\lambda$  values, a  $\kappa_\lambda$ -dependent uncertainty is used following the LHC HH group recommendations. Note: a new recommendation for the ggF uncertainty was published by the LHC HH group after these analysis results were released as a PAS [1], thus these are not accounted for in this dissertation.
- Final state branching fraction: The uncertainty is  $\pm 2.5\%$ , as obtained by propagating the theoretical uncertainty in the branching ratio of the  $H \rightarrow b\bar{b}$  decay, assuming  $\mathcal{B}(H \rightarrow b\bar{b}) = 0.5824 \pm 0.65\% (\text{theory})^{+0.72\%}_{-0.74\%} (m_b)^{+0.78\%}_{-0.80\%} (\alpha_s)$  for  $m_H = 125$  GeV [133]. It is included only when quoting the cross section limit with respect to the SM (i.e. signal strength  $\mu$ ) and for the likelihood scans, but not for the cross section upper limit.

### 4.3.2 Background Model Uncertainties

For the data-driven background model, a dedicated set of systematic uncertainties are considered to account for different effects that impact the predicted normalization and shape of the background templates. These uncertainties are summarized in Tables 4-9 and 4-10. They are treated uncorrelated across categories and datasets. The following sources are considered:

- Bin-by-bin: These uncertainties account for the limited number of events available in the  $A_{SR}(3b)$  region. They are the bin-by-bin statistical uncertainty of the  $A_{SR}(3b)$  region propagated into the  $A_{SR}(4b)$  region. The Barlow-Beeston-lite approach is used [134].
- Transfer factor: The  $A_{CR}(3b)$  and  $A_{CR}(4b)$  have limited number of events. This impacts the measurement of the transfer factor used to compute the  $A_{SR}(4b)$  predicted normalization. A normalization uncertainty is assigned to account for this effect in all categories, as presented in Tables 4-2 and 4-3.
- Validation normalization: A normalization uncertainty to account for the residual differences between the observed data and background normalization in some categories (up to 2 standard deviations) in the validation studies, presented in Table 4-6 and Table 4-7. The uncertainty is calculated by adding the extra contribution to the total normalization uncertainty that reduces the discrepancy with data to one standard deviation level.
- Validation statistical precision: From background validation studies, we observed a very good compatibility between the data and the model across all categories and years in the  $V_{SR}(4b)$  region. To validate this procedure, the  $V_{SR}(4b)$  region should have a comparable statistical precision with respect to the  $A_{SR}(4b)$  region. The statistical precision is defined as  $p = 1/\sqrt{N_{pred}}$ , where  $N_{pred}$  is the predicted background normalization. An extra normalization uncertainty is assigned to the categories with  $p(A_{SR}^{4b}) < p(V_{SR}^{4b})$  using the following formula:  $\sigma_{val} = \sqrt{p(V_{SR}^{4b})^2 - p(A_{SR}^{4b})^2}$ .
- Shape: For the two ggF categories, the reweighting of the observables has a small dependency on the definition of the  $A_{CR}$  region (where the BDT reweighting method is trained). A shape uncertainty is derived to account for this effect by deriving alternative templates in restricted portions of the default  $A_{CR}$  region. The portions used for the alternative trainings are illustrated in Figure 4-15. Since the kinematic properties of the events are highly correlated with  $m_{paral}$ , the comparison of the trainings between regions A and B allows for estimating the quality of a training that interpolates from the CR into the SR using events with similar (B) or softer/harder (A) kinematic properties compared to the target SR. The alternative templates are built for the two GGF categories. For the VBF category 1, the aforementioned method is not viable, because there are not enough events in the regions A and B to ensure a good training of comparable quality as the nominal one. Consequently, the method used is to compare the shape of the observables between data and background model in the  $V_{SR}(4b)$  region. Then, a linear function is fitted to the data/model ratio. The resulting fits are compatible with a horizontal line, as illustrated in Fig. 4-14. Next, alternative templates are estimated by reweighting the bins of the nominal template with the linear curve with the slope varied by one standard deviation up or down. These templates are used as the shape uncertainty of the observables.

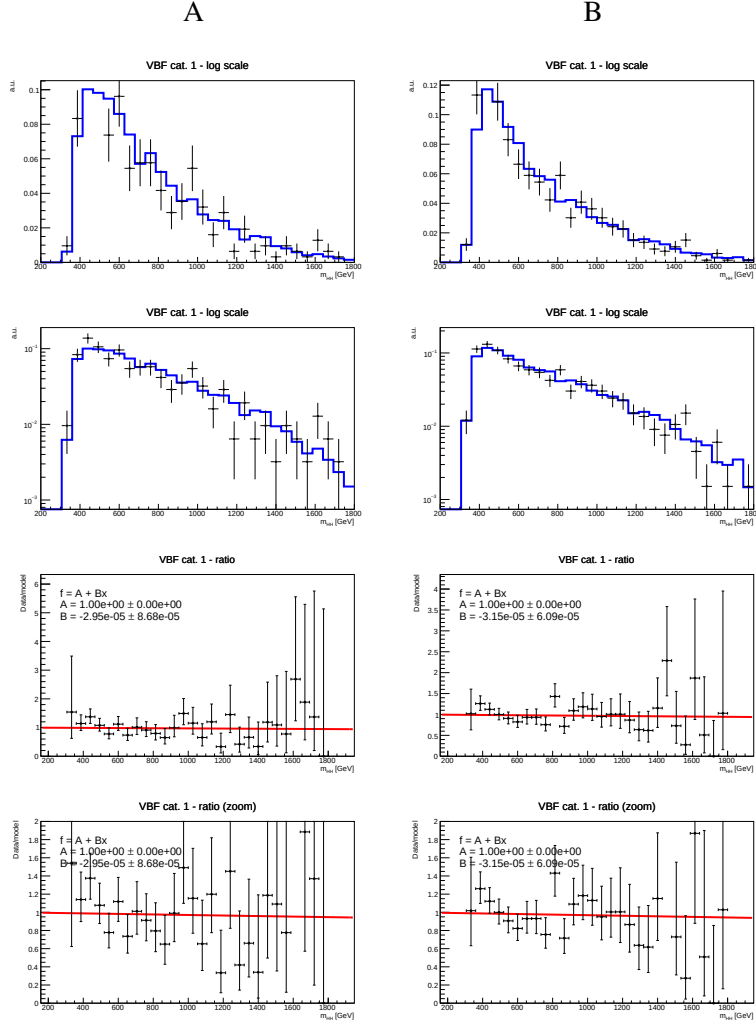


Figure 4-14. Linear fits used to derive a shape uncertainty for VBF category 1. A) 2016 fits and B) 2017-2018 fits.

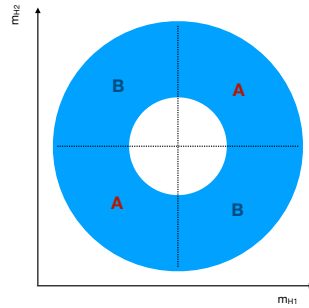


Figure 4-15. Regions used to determine the background shape uncertainty in ggF categories.

Table 4-9. Summary of the background model uncertainties in the 2016 analysis.

| Uncertainty sources          | ggF Cat1 | ggF Cat 2 | VBF Cat1 | VBF Cat 2 |
|------------------------------|----------|-----------|----------|-----------|
| Transfer factor (%)          | 1.8      | 1.3       | 3.2      | 31.7      |
| Validation normalization (%) | 3.2      | -         | -        | -         |
| Validation precision (%)     | 0.4      | -         | 3.0      | 29.5      |
| Bin-by-bin                   | ✓        | ✓         | ✓        | ✓         |
| Shape                        | ✓        | ✓         | ✓        | -         |

Table 4-10. Summary of the background model uncertainties in the 2017-2018 analysis.

| Uncertainty sources          | ggF Cat1 | ggF Cat 2 | VBF Cat1 | VBF Cat 2 |
|------------------------------|----------|-----------|----------|-----------|
| Transfer factor (%)          | 1.4      | 0.9       | 2.2      | 17.7      |
| Validation normalization (%) | 4.7      | 1.5       | -        | -         |
| Validation precision (%)     | -        | 0.3       | 2.2      | 33.3      |
| Bin-by-bin                   | ✓        | ✓         | ✓        | ✓         |
| Shape                        | ✓        | ✓         | ✓        | -         |

### 4.3.3 Expected Impact of Nuisance Parameters

The signal and background uncertainties of the analysis are defined as nuisance parameters  $\theta$  in the statistical model further described in Section 5.1.1. The predicted or ‘pre-fit’ nuisance parameter have a default value ( $\theta_0$ ) and  $\pm 1$  standard deviation uncertainties ( $\Delta_\theta$ ). The presence of a signal can be determined by measuring the best-fit to the signal strength ( $\hat{\mu}$ ). The signal strength  $\mu$  is the parameter used to scale the signal cross section with respect to the values used for the signal normalization. The measurement of  $\hat{\mu}$  is carried out by a maximum likelihood fit of the signal plus background ( $\mu s + b$ ) model to the observed data, in which the nuisance parameters are floated. The corresponding fitted or ‘post-fit’ value and uncertainty of the nuisance parameters are denoted as  $\hat{\theta}$  and  $\Delta$ , respectively.

To understand the importance or ‘impact’ of a systematic uncertainty in the statistical analysis, the procedure is to determine how the corresponding nuisance parameter  $\theta$  affects the measurement of  $\hat{\mu}$ . Then, one evaluates the shift  $\Delta\hat{\mu}$  induced when the nuisance parameter is set at its  $\pm 1$  post-fit values. Large  $\Delta\hat{\mu}$  values imply a strong dependency of the signal strength on that systematic uncertainty. Moreover, one can compare the pre-fit and post-fit nominal value by computing the pull defined as  $(\hat{\theta} - \theta_0)/\Delta_0$ . The pull error bars are defined as the ratio between  $\Delta$  and  $\Delta_0$ . If they are smaller than 1, the nuisance parameter uncertainties are constrained in the fit.

Before the looking at the data, the blinded procedure is to compute the expected nuisance parameter impacts. They are evaluated using the ‘Asimov’ dataset, i.e. the prediction at the median expected from the  $\mu s + b$  model. Moreover, the nuisance parameters and signal strength are set to their pre-fit values. As the HH signal is not observed yet at the LHC,  $\mu$  is set at 7.3, which is the expected 95 % CL upper limit on  $\mu$  using all Run-2 categories. This blinded limit is obtained using the method described in Subsection 5.1.1 based on the a-priori background distribution instead of the observed data. The nuisance parameters with the largest expected impact are presented in Figure 4-16. The behavior of the systematic uncertainties is as expected. They show that the background model uncertainties are the dominant uncertainties in this search. In particular, the background uncertainties associated to the ggF category 2 have the most impact. The largest impact of a signal uncertainty comes from the trigger measurement.

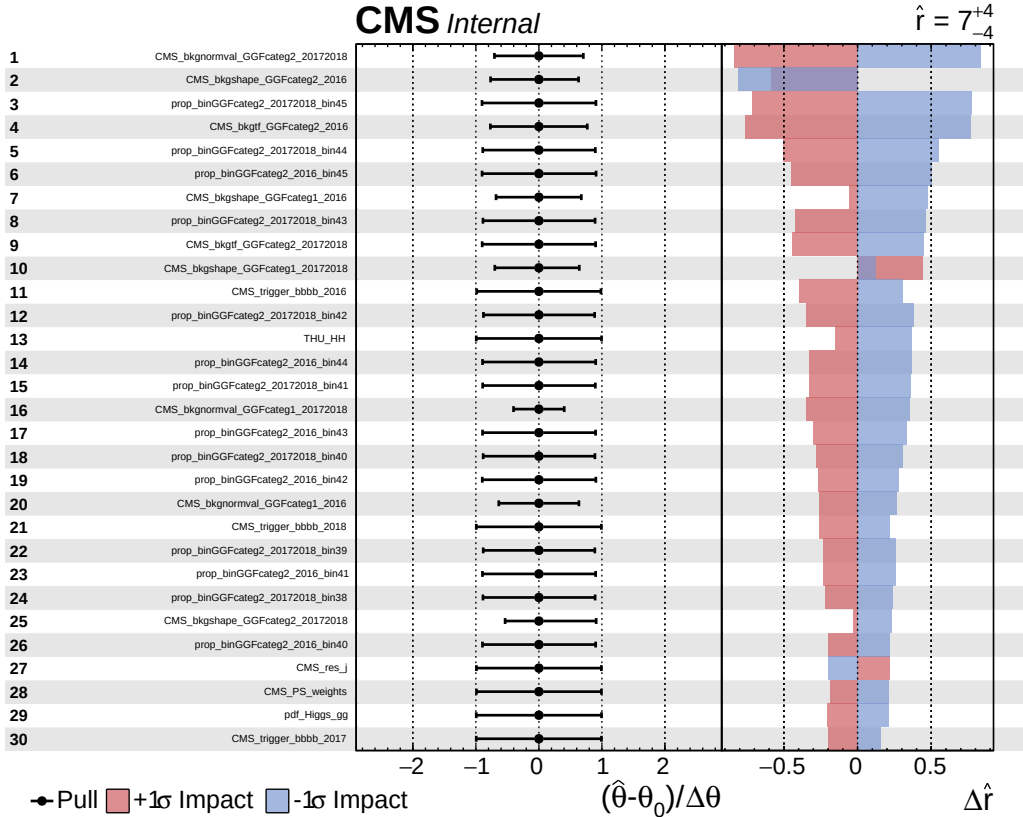


Figure 4-16. Expected impact of the highest systematic uncertainties on the measurement of the signal strength ( $\hat{r}$ ) using all Run-2 categories.

## CHAPTER 5 RESULTS

### 5.1 Statistical Analysis

#### 5.1.1 Statistical Model

The CMS and ATLAS Collaborations follow a procedure based on frequentist statistics in order to characterize the exclusion of a signal or an observation of an excess in the observed data. The general aspects of this formalism are introduced in this subsection, but more details can be found in Refs. [135, 136]. In the CMS Collaboration, these statistical methods are implemented by experts in a general-purpose framework denominated as Higgs Combine Tool. This tool is currently used by the majority of searches.

In the following formalism, the data (experimental observation or pseudo-data) yields is denoted as  $n$ . The signal and background expected yields are (denoted  $s$  and  $b$ , respectively) are defined as functions of their uncertainty nuisance parameters  $\theta$  (where  $\tilde{\theta}$  is the default value):  $s = s(\theta)$  and  $b = b(\theta)$ . The systematic error pdfs  $\rho(\theta|\tilde{\theta})$  are interpreted as the posterior constructed from auxiliary (real or imaginary) measurements  $\tilde{\theta}$ . Then, from Bayes' Theorem,  $\rho(\theta|\tilde{\theta}) \sim p(\tilde{\theta}|\theta) \cdot \pi_{\theta}(\theta)$ , where  $p(\tilde{\theta}|\theta)$  is measurements pdf and  $\pi_{\theta}(\theta)$  is the priors pdf. Then, in a frequentist statistical test, the pdf  $p(\tilde{\theta}|\theta)$  is used as a constraint term in the likelihood and to sample distributions of a test statistic. For instance, an uncertainty describing a positive and negative variation with respect to the default nuisance parameter (e.g. background model normalization) is typically defined as a log-normal uncertainty:

$$p(\tilde{\theta}|\theta) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(\tilde{\theta} - \theta)^2}{2}\right) \quad (5-1)$$

In a frequentist statistical test, the null hypothesis ( $H_0$ ) is tested against the alternative hypothesis ( $H_1$ ). Within the particle physics context, the absence of a new signal requires the test of the signal plus background case ( $H_{\mu_{s+b}}$ ) against the background-only ( $H_b$ ) case. On the other hand, the observation of a new signal requires the test of the background-only case against the signal plus background case. The agreement of the observed data with a given hypothesis is quantified by the probability or “p-value” of finding data with similar or greater incompatibility with the hypothesis. The hypothesis is rejected if a p-value is found to be under a threshold.



In this dissertation, the values of the observables are determined and put in histograms of  $i$  bins to perform a binned statistical analysis to obtain the results. Then, by defining  $P(\text{data}|\mu s(\theta) + b(\theta))$  as the product of Poisson probabilities per histogram bin (equation 5-2), the likelihood function is constructed using equation 5-3.

$$P(\text{data}|\mu s + b) = \prod_i \frac{(\mu s_i + b_i)^{n_i}}{n_i!} e^{-(\mu s_i + b_i)} \quad (5-2)$$

$$\mathcal{L}(\text{data}|\mu s + b) = P(\text{data}|\mu s(\theta) + b(\theta)) \cdot p(\tilde{\theta}|\theta) \quad (5-3)$$

It is common to express the absence of signal as the exclusion limit on the parameter  $\mu$ . The observed exclusion limits are calculated using a test statistic  $\tilde{q}_\mu$  based on the profiled likelihood ratio, presented in equation 5-4. The numerator is a conditional likelihood function, where  $\hat{\theta}_\mu$  is the conditional maximum-likelihood estimator of  $\theta$ , i.e. the value of  $\theta$  maximizing the likelihood for a given  $\mu$ . The denominator is the maximized likelihood function, where  $\tilde{\theta}$  and  $\tilde{\mu}$  are the maximum-likelihood estimators.

$$\tilde{q}_\mu = -2 \ln \frac{\mathcal{L}(\text{data}|\mu, \hat{\theta}_\mu)}{\mathcal{L}(\text{data}|\hat{\mu}, \hat{\theta})}, \text{ with } 0 \leq \tilde{\mu} \leq \mu \quad (5-4)$$

The next step is to find the observed value of the test statistic  $\tilde{q}_\mu$  for a given  $\mu$  value. Then, we find the values of the nuisances parameters maximizing the likelihood for the  $H_b$  and  $H_{\mu s+b}$  hypotheses, denoted as  $\hat{\theta}_0^{obs}$  and  $\hat{\theta}_\mu^{obs}$ , respectively. To evaluate the probabilities, we need to know the  $H_b$  and  $H_{\mu s+b}$  associated pdfs:  $f(\tilde{q}_\mu|0, \hat{\theta}_0^{obs})$  and  $f(\tilde{q}_\mu|\mu, \hat{\theta}_\mu^{obs})$ , respectively. These pdfs are typically derived by generating toys or by using formulae derived the ‘asymptotic-limit’ approximation of the profile likelihood ratio. Then, exclusion limits are obtained using the modified frequentist  $\text{CL}_s$  method [137]. For a given  $\mu$  hypothesis, the  $\text{CL}_s$  parameter is defined by equation 5-5. The  $H_{\mu s+b}$  hypothesis is excluded at  $(1-\alpha)$  confidence level (CL) when  $\text{CL}_s \leq \alpha$ . Therefore, the 95% CL upper level limit on  $\mu$  ( $\mu^{95\%CL}$ ) is the value satisfying that  $\text{CL}_s \leq 0.05$ .

$$\text{CL}_s(\mu) = \frac{\text{CL}_{s+b}}{\text{CL}_b} = \frac{p_{\mu s+b}}{1 - p_b} = \frac{P(\tilde{q}_\mu \geq \hat{\theta}_\mu^{obs} | H_{\mu s+b})}{P(\tilde{q}_\mu \geq \hat{\theta}_\mu^{obs} | H_b)} = \frac{\int_{\tilde{q}_\mu^{obs}}^{\infty} f(\tilde{q}_\mu|\mu, \hat{\theta}_\mu^{obs}) d\tilde{q}_\mu}{\int_{\tilde{q}_0^{obs}}^{\infty} f(\tilde{q}_\mu|0, \hat{\theta}_0^{obs}) d\tilde{q}_\mu} \quad (5-5)$$

The expected exclusion limits are computed by generating pseudo-data from background-only toy experiments. Then, one follows the same procedure as presented above to obtain the  $\mu^{95\%CL}$  value per toy. From the toys information, a cumulative probability function is built in order to find the quantile crossing corresponding to the median and standard deviation (s.d.) bands: 2.5% (-2 s.d), 16% (-1 s.d), 50% (median), 68% (+1 s.d.), 95% (+2 s.d.). Note that if the procedure is CPU intensive, ‘asymptotic-limit’ formulae are used.

In equation 5-6 is presented the test statistic  $q_0$  used for discovery of a new signal. It is important to note that by this definition, a background deficit with respect to the data does not contribute to rejecting the  $H_b$  hypothesis.

$$q_0 = -2 \ln \frac{\mathcal{L}(\text{data}|0, \hat{\theta}_0)}{\mathcal{L}(\text{data}|\hat{\mu}, \hat{\theta})}, \text{ with } \hat{\mu} \geq 0 \quad (5-6)$$

To quantify the level in which the  $H_b$  hypothesis is rejected, the p-value ( $p_0$ ) corresponding to the observed test statistic  $q_0^{obs}$  is evaluated using equation 5-7. The  $f(q_0|0, q_0^{obs})$  pdf is the distribution of the  $q_0$  assuming the background-only hypothesis, and can be estimated with toys or by asymptotic approximation formulae. The p-value is typically converted into statistical significance with a one-sided Gaussian tail by the relation in equation 5-8.

$$p_0 = P(q_0 \geq \hat{\theta}_0^{obs} | H_b) = \int_{q_0^{obs}}^{\infty} f(q_0|0, \hat{\theta}_0^{obs}) dq_0 \quad (5-7)$$

$$p = \int_Z^{\infty} \frac{1}{\sqrt{2\pi}} \exp(-x^2/2) dx \quad (5-8)$$

In the particle physics community, a discovery of new phenomena is equivalent to the observation of an excess of data with respect to the background with  $5\sigma$  significance ( $p = 2.87 \times 10^{-7}$ ). Searches for a new particle with an unknown mass ( $m_X$ ) typically perform scans over a range of mass, where each mass hypothesis has its own ‘local’ p-value. However, a local excess must take into account the probability to observe such an excess anywhere in the mass range or the “looking-elsewhere ” effect. A ‘global’ p-value is derived to address it using dedicated methods [138].

### 5.1.2 Validation of the Statistical Model

This analysis received the green light for the unblinding of the data on March 2021 after presenting the analysis strategy described above. Before proceeding with the unblinding of the observed distribution and obtaining the final results presented later in Sections 5.2 and 5.3, statistical tests are performed to validate the statistical model with observed data.

Figure 5-1 presents the nuisance parameters with the largest impacts when fitting the observed data. The results demonstrate a good behavior of the post-fit values and uncertainties relative to the pre-fit ones. Furthermore, as seen before in Subsection 4.3.3, the nuisance parameters with the largest impact correspond to the background model uncertainties in the ggF category 2 and predominantly in the 2017-2018 dataset. The largest impact signal uncertainty comes from the 2016 trigger efficiency measurement.

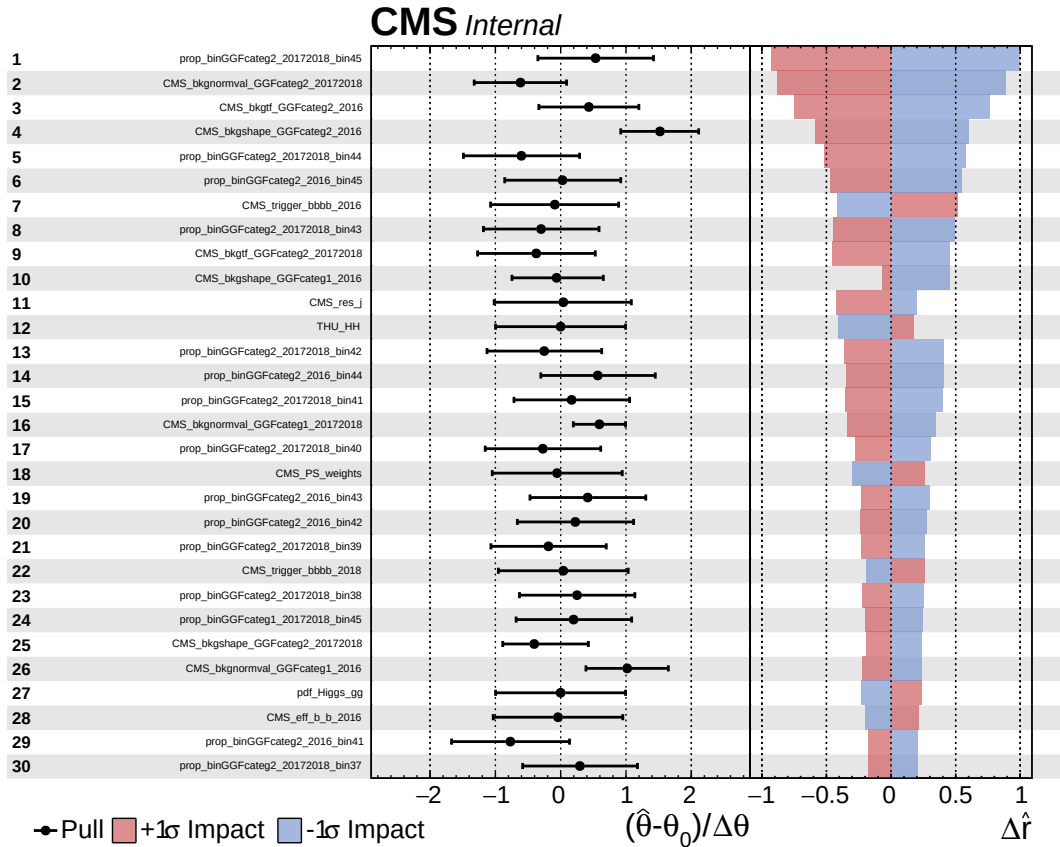


Figure 5-1. Observed impact of the highest systematic uncertainties on the measurement of the signal strength ( $\hat{r}$ ) using all analysis categories.

To quantify how compatible the background model is to the observed data, a goodness of fit test based on the saturated model method [129] is evaluated individually in each category, and simultaneously in all Run-2 categories. This method similar to the  $\chi^2$ , but has the advantage to be evaluated for arbitrary binned channels and parameters. A test statistic is computed first in the observed data. Then, it is computed for 1000 generated toys with  $\mu$  left floating such that the estimation is independent of the presence or absence of signal. The test statistic toys distribution is compared to the one in observed in data to compute the compatibility p-value, as illustrated in Figure 5.7. The resulting p-values are presented Table 5-1. They demonstrate the excellent performance of the background model in each individual category and in all categories.

Table 5-1. Data-background model compatibility p-value obtained from the goodness of fit test applied to the observables in the  $A_{SR}(4b)$  region.

| Dataset   | Subcategory    | Observable          | Goodness of fit test p-value |
|-----------|----------------|---------------------|------------------------------|
| 2016      | ggF category 1 | BDT output          | 76%                          |
| 2016      | ggF category 2 | BDT output          | 57%                          |
| 2016      | VBF category 1 | $m_{HH}$            | 72%                          |
| 2016      | VBF category 2 | Counting experiment | 99%                          |
| 2017-2018 | ggF category 1 | BDT output          | 23%                          |
| 2017-2018 | ggF category 2 | BDT output          | 62%                          |
| 2017-2018 | VBF category 1 | $m_{HH}$            | 45%                          |
| 2017-2018 | VBF category 2 | Counting experiment | 41%                          |
| Run-2     | All            | All                 | 72%                          |

### 5.1.3 Observed Distributions

The distributions of the signal region observables in all categories (ggF cat.1, ggF cat.2, VBF cat.1 and VBF cat. 2) are presented in Figures 5-2 and 5-3 for the 2016 and 2017-2018 datasets, respectively. The shown background model is the ‘post-fit’ background and uncertainties obtained by maximum likelihood fit of the background-only model to the observed data. The background model describes well the data within the uncertainties in all categories. Figure 5-4 combines all distributions into a single one, made with bins of similar expected signal-to-background ratio, where SM ggF signal is assumed. No significant excess of data events is observed in the highest bins of the distribution by looking at the data/background ratio panel. An underfluctuation of the data relative to the background model is observed in a few very

sensitive bins, which correspond to the 2017-2018 ggF category 2. The underfluctuation in this category was studied in detail during the review, and it was concluded that it is of statistical nature. Therefore, the observed distributions are consistent with the background-only hypothesis.

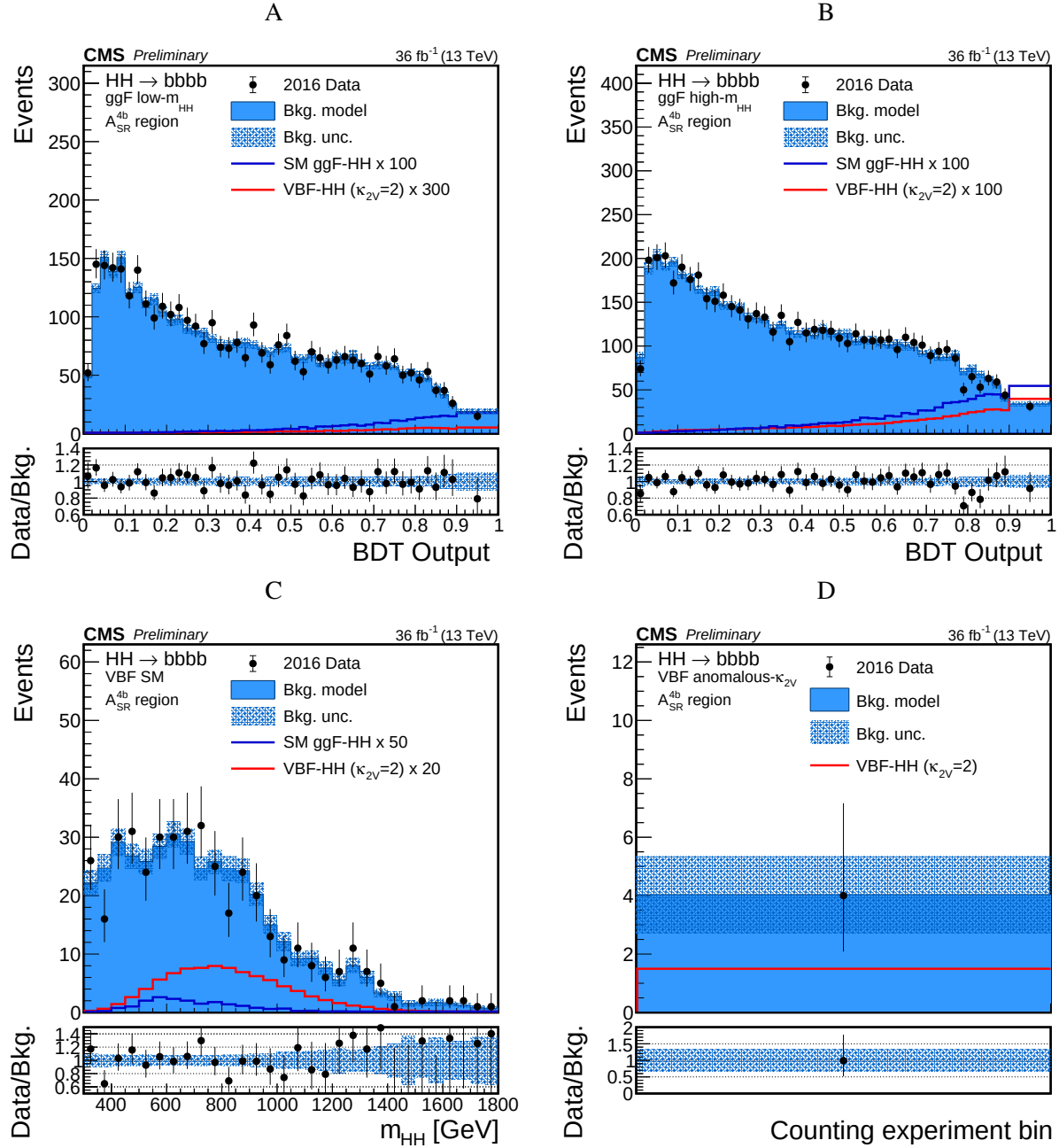


Figure 5-2. Observed distribution in the signal regions for 2016 data. A) ggF category 1, B) ggF category 2, C) VBF category 1, D) VBF Category 2. The presented background distributions are the one obtained after fitting the background prediction to the data floating the background uncertainties.

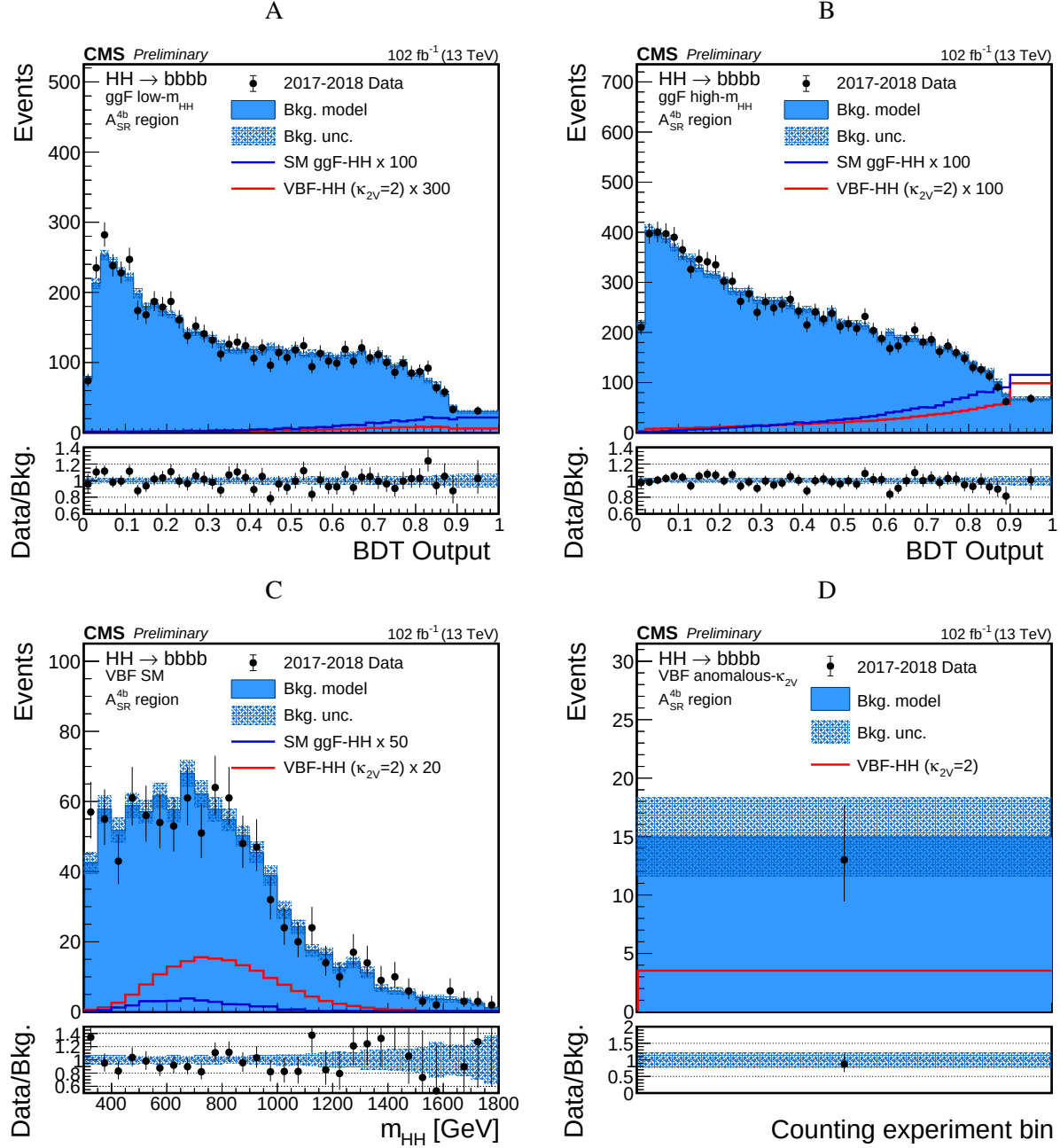


Figure 5-3. Distributions of the observables in the signal regions for 2017-2018 data. A) ggF category 1, B) ggF category 2, C) VBF category 1, D) VBF Category 2. The presented background distributions are the one obtained after fitting the background prediction to the data floating the background uncertainties.

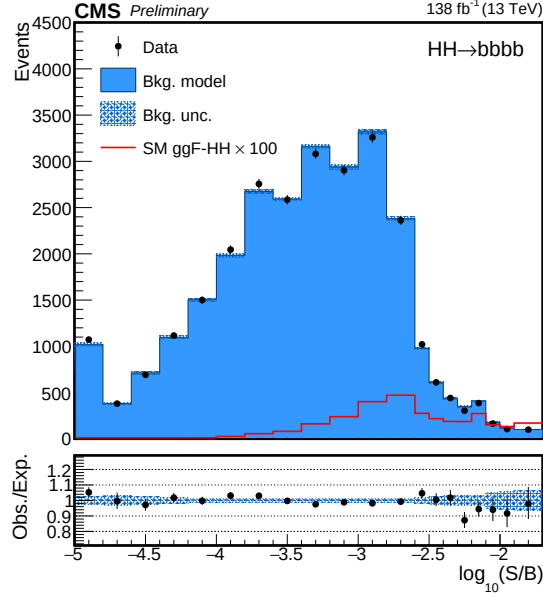


Figure 5-4. Distribution of the expected signal-to-background ratio for events selected in all categories and datasets.

## 5.2 Upper Limits

Under the hypothesis of absence of signal in the observed data with respect to the background, one can set 95 % CL upper limits on the signal strength parameter  $\mu$ , which is then translated into a limit on the HH production cross section. The limits are computed using the asymptotic CLs method using the statistical model described in Subsection 5.1.1. The final results are obtained combining the limits from all categories and years. In the figures below, green and yellow bands indicate the regions containing 68 and 95% of the distribution of limits expected under the background-only hypothesis, respectively.

The observed (expected) upper limits on the SM signal strength parameter ( $\mu = \sigma_{\text{HH}}/\sigma_{\text{HH}}^{\text{SM}}$ ) is set 3.7 (7.3). The breakdown of the contributions to this constraint by year and category are presented in Figure 5-5 and Figure 5-6, respectively. The results show that the largest observed and expected constraint is obtained in the 2017-2018 ggF category 2.

The results are also interpreted as a limit on the SM production cross section. The observed (expected) upper limit on the HH production cross section is set at 118 (239) fb. At the time of writing this dissertation, this is the stringent constraint on this process at the LHC. In addition, the

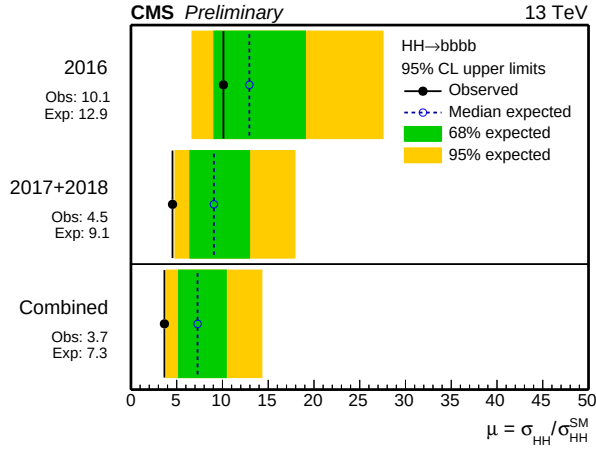


Figure 5-5. Observed and expected 95% CL limits on the standard model Higgs boson pair (HH) production signal strength ( $\mu$ ) breakdown by year.

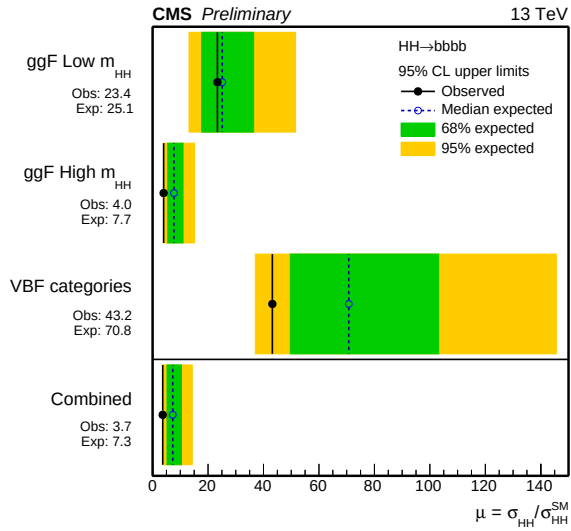


Figure 5-6. Observed and expected 95% CL limits on the standard model Higgs boson pair (HH) production signal strength ( $\mu$ ) breakdown by category.

observed (expected) limit on the VBF production cross section is set at 390 (711) fb, where the SM ggF signal is fixed as background.

Cross section limits are also obtained in the context of the anomalous SM couplings and non-SM couplings. Couplings that are not indicated on the horizontal axis are set to the SM prediction. Upper limits on the HH production cross section are derived as a function of the  $\kappa_\lambda$  coupling modifier as illustrated in Figure 5-7 A). These limits are compared to the theoretical



cross section (pink line) to obtain the observed (expected) allowed interval of  $-2.3 < \kappa_\lambda < 9.4$  ( $-5.0 < \kappa_\lambda < 12.0$ ) at 95% CL. Moreover, upper limits on the VBF HH production cross section are also set as a function of the  $\kappa_{2V}$  coupling modifier, presented in Figure 5-7 B). The observed (expected) allowed interval of  $-0.1 < \kappa_{2V} < 2.2$  ( $-0.4 < \kappa_{2V} < 2.5$ ) at 95% CL. In addition, Figure 5-8 shows the observed (expected) upper limits are also set on the production cross section of the twelve EFT benchmarks presented section 1.3.2

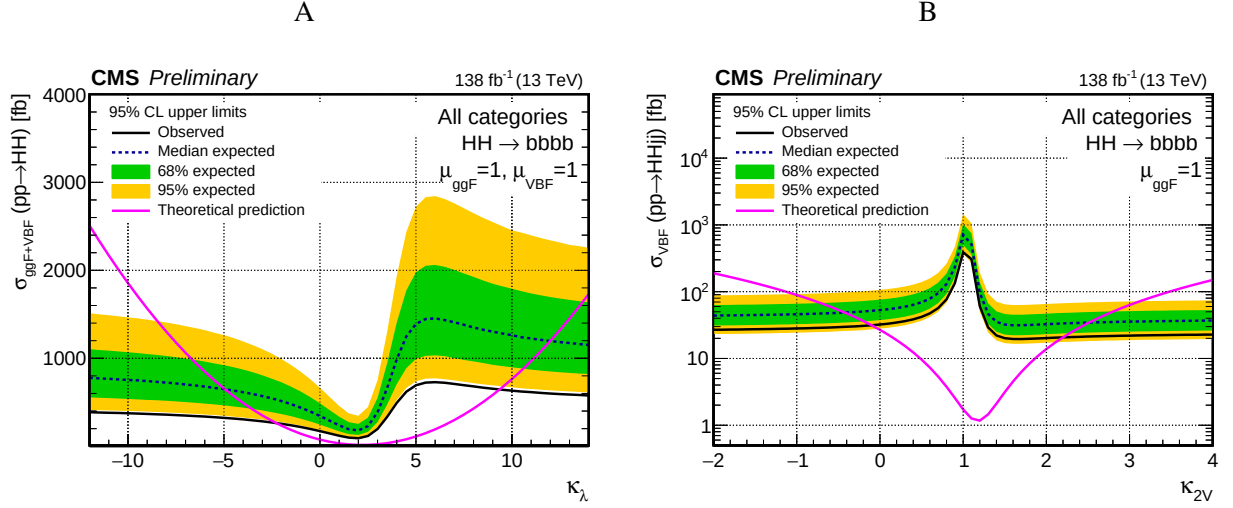


Figure 5-7. Observed and expected 95% CL upper limits on Higgs boson pair cross section as a function of the coupling modifiers. A)  $\kappa_\lambda$  and B)  $\kappa_{2V}$ .

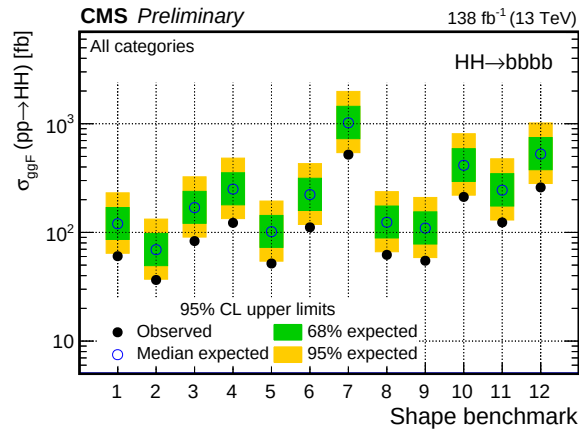


Figure 5-8. Observed and expected 95% CL upper limits on the gluon fusion Higgs boson pair production cross section for the 12 benchmarks derived from the effective field theory parametrization.

### 5.3 Likelihood Scans

Under the assumption that HH production exists according to the SM expectation, one can perform a measurement of the coupling modifiers through one dimensional or two-dimensional negative log-likelihood scans. The likelihood shape and minimum originates from an interplay between the cross section and the signal acceptance in the analysis categories.

Figure 5-9 shows the one-dimensional negative log-likelihood scans as function of the  $\kappa_\lambda$  (A) and  $\kappa_{2V}$  (B) coupling modifiers, where the coupling modifiers that are not shown in the figures are set to their SM values. The scans are evaluated with an Asimov dataset assuming the SM expected signal hypothesis (in light blue), and the observed data (black). In both cases, the observed likelihood is compatible with the SM hypothesis at 95% CL.

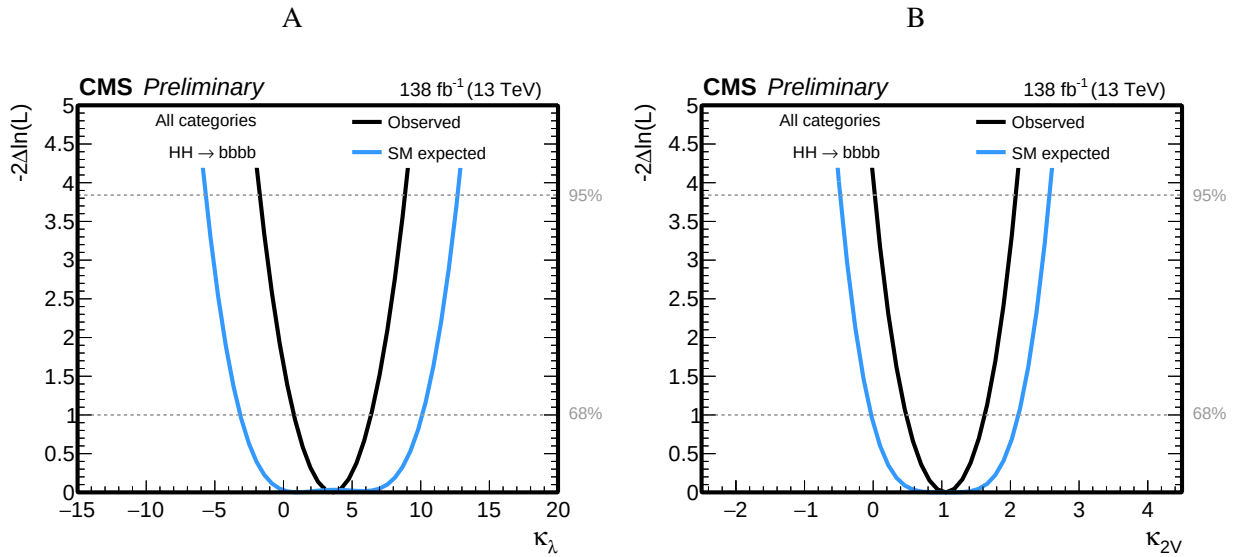


Figure 5-9. Observed and expected negative log-likelihood scan as a function of the couplings. A)  $\kappa_\lambda$  and B)  $\kappa_{2V}$  couplings.

Two-dimensional negative log-likelihood scans in the  $(\kappa_\lambda, \kappa_{2V})$  and  $(\kappa_{2V}, \kappa_V)$  planes are performed and illustrated in Figures 5-10 and 5-10, respectively. The minimum or best-fit (black star) and the 68% and 95% CL contours are presented for both the Asimov dataset assuming the SM hypothesis and the observed data. The coupling modifiers not shown in the figures are set to the SM values. Both results in observed data are compatible with the SM expectation (circle) at 95% CL.

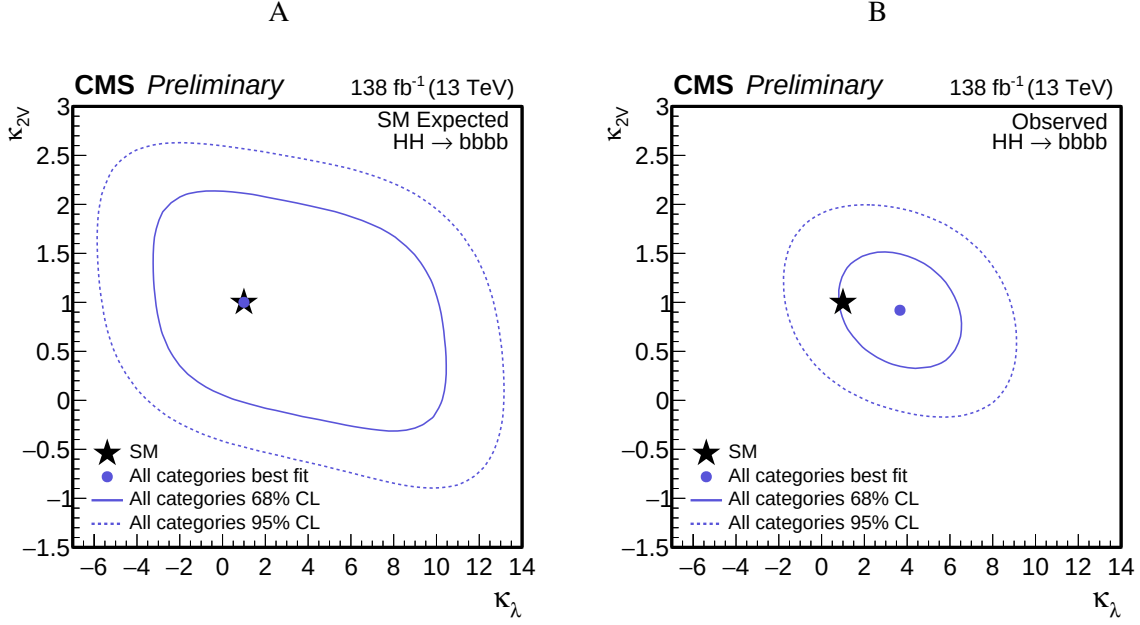


Figure 5-10. Negative log-likelihood contours at 68% and 95% CL as a function of the  $(\kappa_\lambda, \kappa_{2V})$  couplings. A) SM hypothesis with Asimov dataset, and B) Observed data. The  $\kappa_t$  and  $\kappa_V$  couplings are set at the SM prediction.

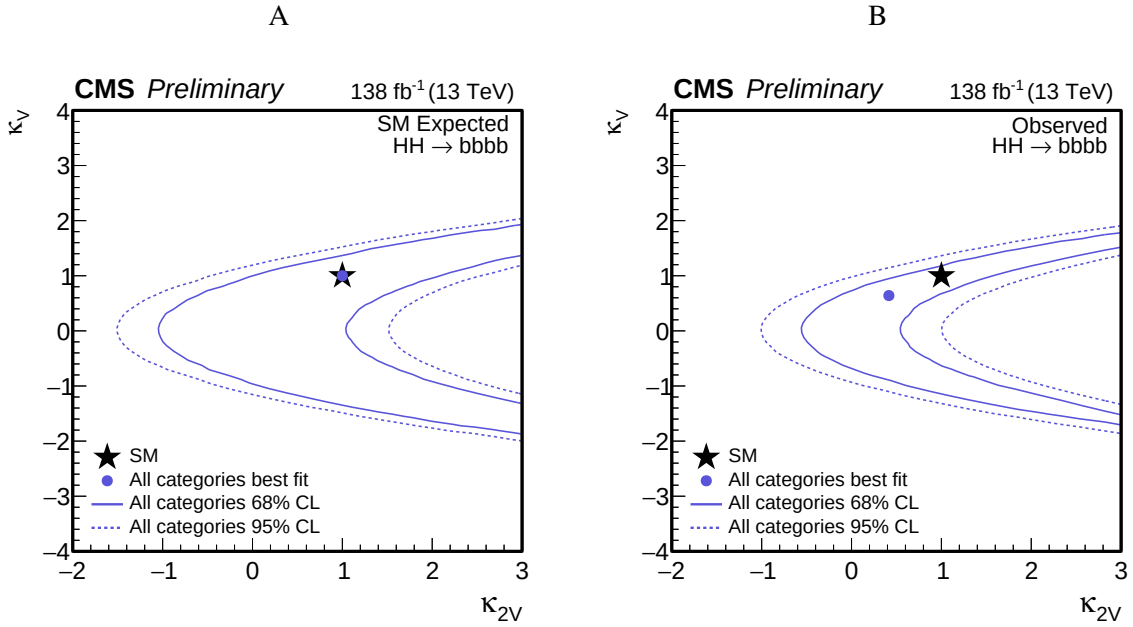


Figure 5-11. Negative log-likelihood contours at 68% and 95% CL as a function of the  $(\kappa_{2V}, \kappa_V)$  couplings. A) SM hypothesis with Asimov dataset, and B) Observed data. The  $\kappa_\lambda$  and  $\kappa_t$  couplings are set at the SM prediction.

## CHAPTER 6 CONCLUSIONS

The Higgs energy potential is connected to the physics of the very early universe, the mechanism electroweak symmetry breaking, and the stability of the Universe. The measurement of Higgs boson self-coupling  $\lambda_{HHH}$  is crucial for the reconstruction of this potential. If we observe any deviations of  $\lambda_{HHH}$  relative to the SM, this could have profound consequences in our understanding of nature. At the LHC, non-resonant HH production gives direct access to this coupling. However, this process has a small SM theoretical cross section ( $\sigma_{SM}^{HH}$ ), making it a very rare process to see at the current data luminosities. The study of the HH signature can shed light on the nature of the scalar sector. By being sensitive to BSM physics, we use the HH signature as a probe to study it and potentially discover it already with the current datasets.

During the LHC Run-2 period, the CMS experiment collected an unprecedented amount of pp collision data at  $\sqrt{s} = 13$  TeV totaling  $138 \text{ fb}^{-1}$ . This dissertation presents an original Run-2 search for non-resonant HH production in the  $b\bar{b}b\bar{b}$  channel, exploring both the gluon fusion and vector boson fusion production mechanisms. It implements innovative techniques in the areas of object identification, event categorization, signal identification, and background modeling. In particular, it deploys, for the first time in CMS, a full data-driven background modeling using an ML algorithm known as BDT-reweighting. The presented result improves the expected analysis sensitivity by a factor of 5 with respect to previous 2016-only result, where a factor of 2 comes from the increase of integrated luminosity and the rest comes from the better and new analysis methods. The observed data does not show an excess of events with respect to the background-only hypothesis, and 95% CL upper limits on the SM and BSM production cross sections are set. These results are complemented with measurements of the couplings associated to the HH process using one-dimensional and two-dimensional likelihood scans.

The work of this dissertation sets an observed upper limit on the SM HH production cross section at  $3.6 \times \sigma_{SM}^{HH}$ . At the time of writing this thesis, it is the best single-channel limit at the LHC. It sets tight constraints on the production cross section of anomalous Higgs couplings and benchmark signals for new physics in an EFT scenario. It excludes the values of the Higgs boson self-coupling modifier outside the range  $-2.3 < \kappa_\lambda < 9.4$  at 95% CL, and the values of the

coupling modifier of the di-vector-boson-di-Higgs-boson interaction outside the range  $-0.1 < \kappa_{2V} < 2.2$  at 95% CL. Furthermore, the measurements of the couplings show the observation compatible with the SM hypothesis at 95% CL.

The results obtained in this dissertation will be a crucial piece of the Run-2 combination of all non-resonant searches planned at the end of 2021 / early 2022. Just by statistically combining the results of the three most sensitive channels ( $b\bar{b}b\bar{b}$ ,  $b\bar{b}\gamma\gamma$  and  $b\bar{b}\tau^+\tau^-$ ), the CMS experiment is expected to set a constraint on the SM cross section at  $2-3 \times \sigma_{SM}^{HH}$ , exceeding preliminary expectations from the extrapolation of the 2016 HH combination ( $\sim 13 \times \sigma_{SM}^{HH}$ ) to the Run-2 integrated luminosity ( $\sim 7 \times \sigma_{SM}^{HH}$ ). Therefore, there is almost a factor of 3 sensitivity gain (on top of the luminosity), which has been achieved through new and sophisticated analysis methods in every channel.

The LHC Run-3 is expected to start in early 2022. During this 3-4 year running period, the LHC will produce pp collisions at  $\sqrt{s} = 13$  or 14 TeV, and will collect a dataset which will at least double the integrated luminosity of the one collected in Run-2. It will be a great opportunity to establish a solid HH strategy with multiple decay channels and production modes. It will be essential to develop new strategies to maximize the sensitivity of the searches for SM and BSM physics with this unprecedented dataset. Based on the experience of this work, it will be important to address the current limitations of the  $b\bar{b}b\bar{b}$  channel: acceptance (the trigger, b jet identification), systematic uncertainties (accuracy of the background model and control region statistical precision), and multijet background rejection.

The Run 3 era will be the playground for new ideas and not just more refined techniques. There is an ongoing effort with the goal of developing a very efficient  $HH \rightarrow b\bar{b}b\bar{b}$  Run-3 trigger. The new trigger requires only two jets to be identified as b jets instead of the three required in the Run-2 triggers, without increasing the jet  $p_T$  requirements to control the trigger rate. The ‘2b’ requirement will allow for additional ‘b-tag’ control region data to be collected, to derive and validate the future data-driven background model. The new trigger is also being developed using the novel jet flavor tagging algorithm ParticleNet [139], which uses a state-of-the-art machine

learning architecture based on point clouds. This tagger has demonstrated to surpass the performance of the standard Run-2 algorithms. The same tagger could be used for the offline analysis. Furthermore, the analysis background model based on BDT-reweighting (or other similar methods) is limited systematically by the statistical power of the control and validation region data. The lack of statistical power can be addressed by using the information of the extra ‘2b’ region. An alternative way for exploration is to use data-driven techniques that could generate a synthetic dataset from real data at high fidelity. These techniques can be ML generative models such as generative adversarial networks (GANs), which are currently studied to speed up the huge production of MC simulation needs for the HL-LHC. Lastly, a large dataset of data-driven background model events will allow for more sophisticated discriminants (e.g. DNNs) to be trained and used for background rejection and signal extraction.

The HL-LHC operation is expected to start in 2027 and continue for at least a decade, collecting 14 TeV pp collision data at higher luminosities. Consequently, this will be the ideal opportunity to measure SM HH production given the large expected dataset ( $\sim 3000\text{-}4000 \text{ fb}^{-1}$ ). According to the last HL-LHC projection at  $3000 \text{ fb}^{-1}$ , the expected significance for this process is 3.0 and 2.6 standard deviations for ATLAS and CMS, respectively. The ATLAS and CMS combined result yields an expected significance of 4.0 standard deviations, thus the evidence for HH production will be at reach. The actual results will be surely better than these projections with the progress and innovation of future experimental methods. Based on the improvements already seen in Run-2, a larger dataset will allow for more sensitive results to be obtained. However, the experimental challenges at high PU will need to be overcome.

The work of this dissertation shows that the  $b\bar{b}b\bar{b}$  channel can be one of the most sensitive HH channels at the LHC, despite being considered very challenging a few years ago. It opens the door to promising prospects for a solid and long-term program starting from Run-2 combinations to the HL-LHC era analyses, with the eventual possibility to deepen our understanding of the fundamental laws of nature. My wish is for this work to serve as a stepping stone and a guide for the future searches in this channel.

## APPENDIX A TRIGGER EFFICIENCY MEASUREMENT

In this appendix is presented more information, omitted in section 3.3, regarding the measurement of the trigger efficiency in data and simulation.

An event selection to get an enriched sample of  $t\bar{t}$  events with a dileptonic final state ( $t\bar{t} \rightarrow W^+bW^-b \rightarrow e^\pm \nu_e \mu^\mp \nu_\mu$ ) is applied for the measurement. First, events pass the HLT\_IsoMu24 trigger path. They must contain exactly muon with  $p_T > 30$  GeV satisfying the medium identification working point and a PF relative isolation  $< 0.1$ . Events are vetoed if they contain additional muons of relative isolation  $< 0.3$  with the same  $p_T$  and identification criteria. Events are also required to contain an electron with  $p_T > 10$  GeV, relative isolation  $< 0.1$ . The electron and muon are required to have opposite electric charge. Events are also required to have at least 4 jets with  $p_T > 25$  GeV,  $|\eta| < 2.4$ , the tight working point of the jet ID algorithm, deepJet score  $> 0$ . Selected jets with  $p_T < 50$  GeV are requested to pass the medium working point of the PU ID algorithm. The four jets with the highest deepJet score are selected. The trigger objects passing the required filters are counted only if they are matched within  $\Delta R = 0.5$  with the four jets. This last requirement is not applied to filters applied on the HT since it is a property of the event and not of the single jets.

The filters associated to the HLT trigger paths used in three years are shown in tables A-1, A-2, A-3 and A-4. The efficiency for each filter as measured in data and in simulation is shown in Figures A-1, A-2, and A-3 for 2016, in Figures A-4 for 2017, and in Figures A-5 for 2018. The efficiency measurement has the important subtleties to mention:

- The nominal procedure cannot be applied directly to filters with b-tagging requirements because there are only two jets which are real b-jets in  $t\bar{t}$  events. Consequently, the efficiency of an online b-tag is evaluated by selecting the highest b-tagged jet (the most probable one to be a real b-jet) of the event and checking that it matches a trigger object passing the online b-tag. Then the triple b-tagging filter efficiency is evaluated as the probability that three out of the four selected jets pass online b-tag.
- The 2016 trigger requirements are the logical ‘OR’ of two paths. Hence, the efficiency of the OR is evaluated using equation A-1, where  $\varepsilon(\text{Double90Quad30} \& \text{Quad45}) = \varepsilon(\text{Double90Quad30}) \times \varepsilon'(\text{Quad45})$ , with  $\varepsilon'(\text{Quad45})$  defined as the efficiency of the ‘Quad45’ trigger over a sample of events passing the ‘Double90Quad30’ trigger.

$$\varepsilon(\text{OR}) = \varepsilon(\text{Double90Quad30}) + \varepsilon(\text{Quad45}) - \varepsilon(\text{Double90Quad30}\&\text{Quad45}) \quad (\text{A-1})$$

Table A-1. Filters associated to the HLT\_DoubleJet90\_Double30\_TripleBTagCSV trigger path in the 2016 dataset. The filters are sorted in the order of their application. The number of objects required and the variable as a function of which the efficiency is measured are shown.

| Filter name                                       | # of required objects | Variable         |
|---------------------------------------------------|-----------------------|------------------|
| hltL1sTripleJetVBFiorHTTIorDoubleJetCIorSingleJet | 1                     | $\sum_{jet} p_T$ |
| hltQuadCentralJet30                               | 4                     | $p_T^4$          |
| hltDoubleCentralJet90                             | 2                     | $p_T^2$          |
| hltBTagCaloCSVp087Triple                          | 3                     | see text         |
| hltQuadPFCentralJetLooseID30                      | 4                     | $p_T^4$          |
| hltDoublePFCentralJetLooseID90                    | 2                     | $p_T^2$          |

Table A-2. Filters of the HLT\_QuadJet45\_TripleBTagCSV trigger path in the 2016 dataset. The filters are sorted following the order of their application. The number of objects required and the variable as a function of which the efficiency is measured are shown.

| Filter name                         | # of objects | Variable         |
|-------------------------------------|--------------|------------------|
| hltL1sQuadJetCIorTripleJetVBFiorHTT | 1            | $\sum_{jet} p_T$ |
| hltQuadCentralJet45                 | 4            | $p_T^4$          |
| hltBTagCaloCSVp087Triple            | 3            | see text         |
| hltQuadPFCentralJetLooseID45        | 4            | $p_T^4$          |

Table A-3. Filters of the HLT\_PFHT300PT30\_QuadPFJet\_75\_60\_45\_40\_TriplePFBTagCSV trigger path in 2017 dataset. The filters are sorted following the order of their application. For every filter, the number of objects required and the variable as a function of which the efficiency is measures is reported.

| Filter name                                               | # of objects | Variable         |
|-----------------------------------------------------------|--------------|------------------|
| hltL1sQuadJetC60IorHTT380IorHTT280QuadJetIorHTT300QuadJet | 1            | $\sum_{jet} p_T$ |
| hltQuadCentralJet30                                       | 4            | $p_T^4$          |
| hltCaloQuadJet30HT300                                     | 1            | $\sum_{jet} p_T$ |
| hltBTagCaloCSVp05Double                                   | 2            | see text         |
| hltPFCentralJetLooseIDQuad30                              | 4            | $p_T^4$          |
| hlt1PFCentralJetLooseID75                                 | 1            | $p_T^1$          |
| hlt2PFCentralJetLooseID60                                 | 2            | $p_T^2$          |
| hlt3PFCentralJetLooseID45                                 | 3            | $p_T^3$          |
| hlt4PFCentralJetLooseID40                                 | 4            | $p_T^4$          |
| hltPFCentralJetsLooseIDQuad30HT300                        | 1            | $\sum_{jet} p_T$ |
| hltBTagPFCSVp070Triple                                    | 3            | see text         |



Table A-4. Filters of the HLT\_PFHT330PT30\_QuadPFJet\_75\_60\_45\_40\_TriplePFBTagDeepCSV trigger path in the 2018 dataset. The filters are sorted following the order of their application. For every filter, the number of objects required and the variable as a function of which the efficiency is measures are reported.

| Filter name                                             | # of objects | Variable         |
|---------------------------------------------------------|--------------|------------------|
| hltL1sQuadJetC50to60IorHTT280to500IorHTT250to340QuadJet | 1            | $\sum_{jet} p_T$ |
| hltQuadCentralJet30                                     | 4            | $p_T^4$          |
| hltCaloQuadJet30HT320                                   | 1            | $\sum_{jet} p_T$ |
| hltBTagCaloDeepCSVp17Double                             | 2            | see text         |
| hltPFCentralJetLooseIDQuad30                            | 4            | $p_T^4$          |
| hlt1PFCentralJetLooseID75                               | 1            | $p_T^1$          |
| hlt2PFCentralJetLooseID60                               | 2            | $p_T^2$          |
| hlt3PFCentralJetLooseID45                               | 3            | $p_T^3$          |
| hlt4PFCentralJetLooseID40                               | 4            | $p_T^4$          |
| hltPFCentralJetsLooseIDQuad30HT330                      | 1            | $\sum_{jet} p_T$ |
| hltBTagPFDeepCSVp24Triple or hltBTagPFDeepCSV4p5Triple  | 3            | see text         |

The filter efficiency are obtained from a fit to the data and simulation distributions. A crystal ball cumulative distribution function (CRY\_S\_CDF) is used as fitting function for filters dependent on  $p_T$  and  $H_T$ . For some filters the CRY\_S\_CDF function does not reproduce the full turn on curve, thus for this case, the CRY\_S\_CDF function is multiplied by an error (ERF) function with the same inflection point coordinate but with the width as a free parameter. The b-tagging filter efficiencies are fitted with a five-degree polynomial. The physical meaning of the CRY\_S\_CDF x ERF parameters is indicated in table A-5. The fits performed in the data and MC measurement are presented from Figures A-6 to A-11 in (2016), Figures A-12 to A-15 in 2017, and Figures A-16 to A-19 in 2018.

Table A-5. Physical meaning of the parameters of the CRY\_S\_CDF x ERF fitting function.

| Parameter | Physical meaning                |
|-----------|---------------------------------|
| Par0      | threshold                       |
| Par1      | width                           |
| Par2      | efficiency low asymptote        |
| Par3      | efficiency high asymptote       |
| Par4      | crystal ball n parameter        |
| Par5      | crystal ball $\alpha$ parameter |
| Par6      | ERF width                       |

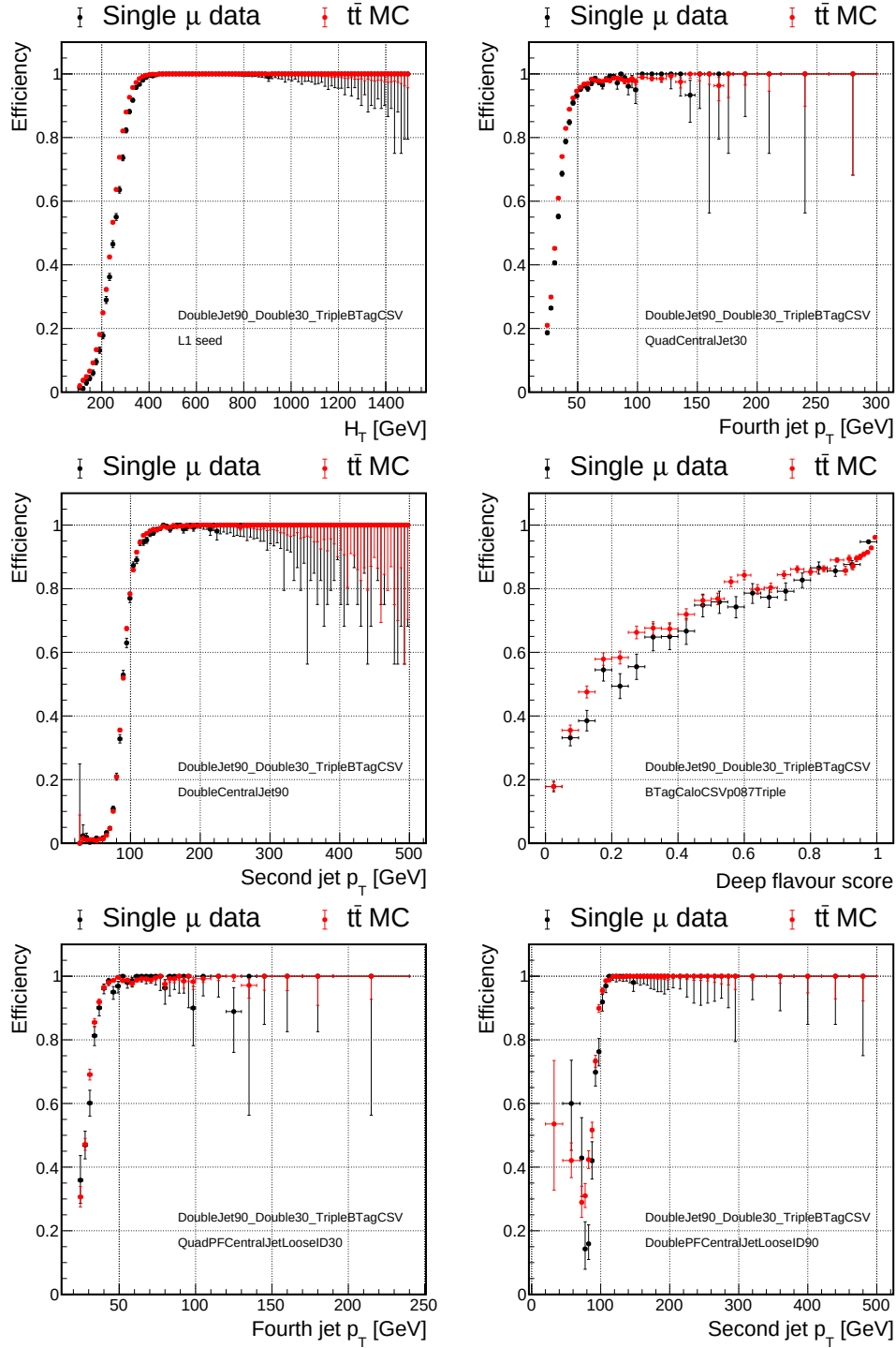


Figure A-1. Efficiency measured in single muon data (black) and  $t\bar{t}$  MC (red) for filters of HLT\_DoubleJet90\_Double30\_TripleBTagCSV, corresponding to the 2016 dataset.

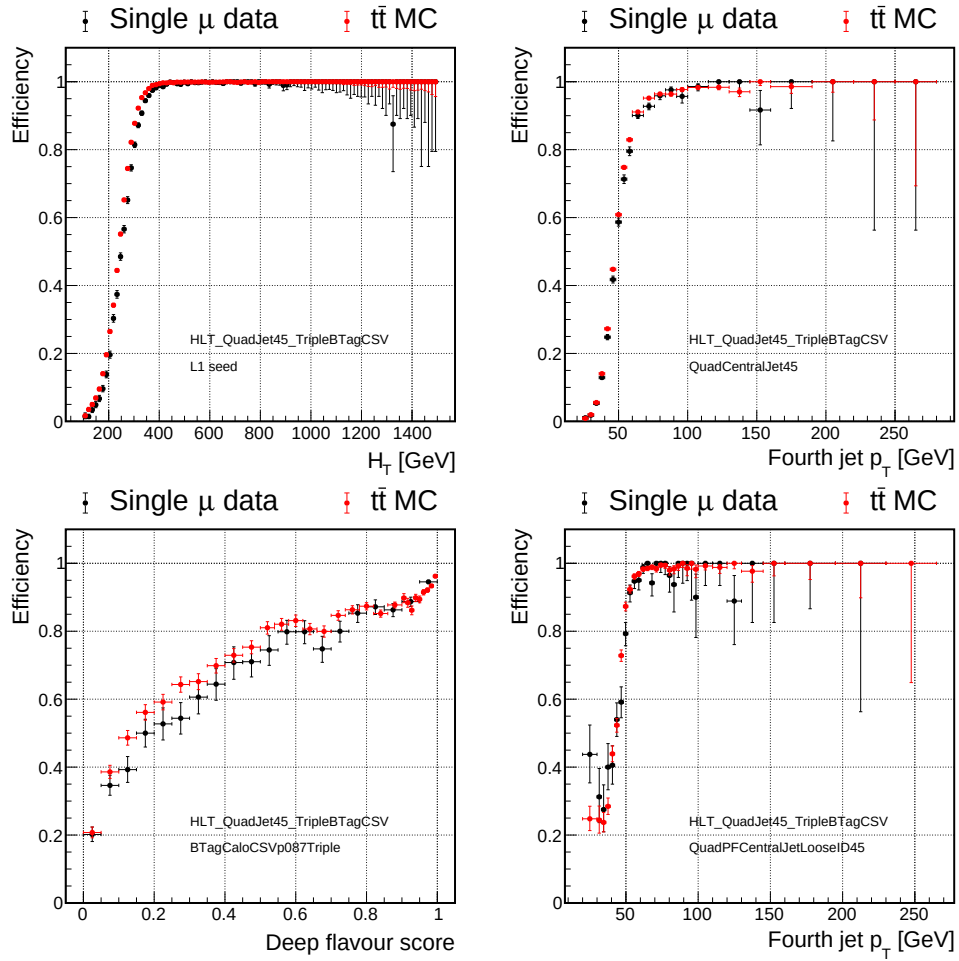


Figure A-2. Efficiency measured in single muon data (black) and  $t\bar{t}$  MC (red) for filters of HLT\_QuadJet45\_TripleBTagCSV, corresponding to the 2016 dataset.

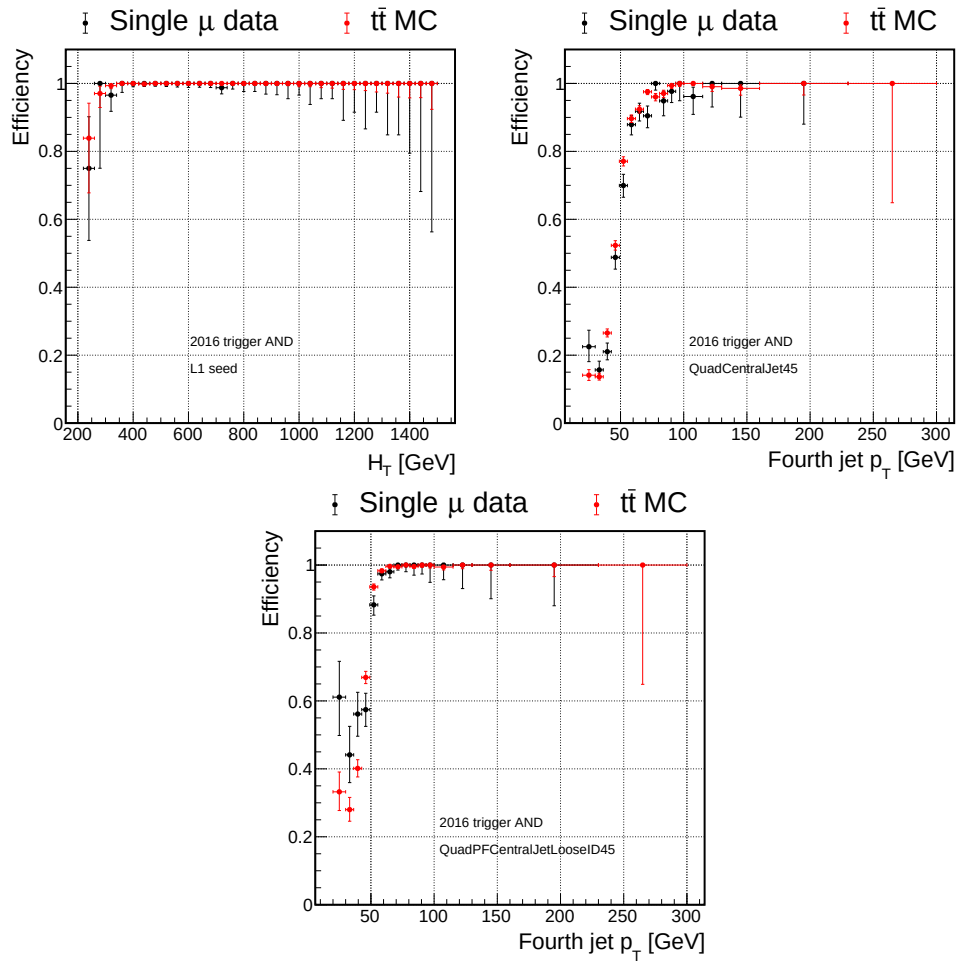


Figure A-3. Efficiency measured in single muon data and  $t\bar{t}$  MC for filters of HLT\_QuadJet45\_TripleBTagCSV evaluated over a sample of events passing HLT\_DoubleJet90\_Double30\_TripleBTagCSV, corresponding to the 2016 dataset.

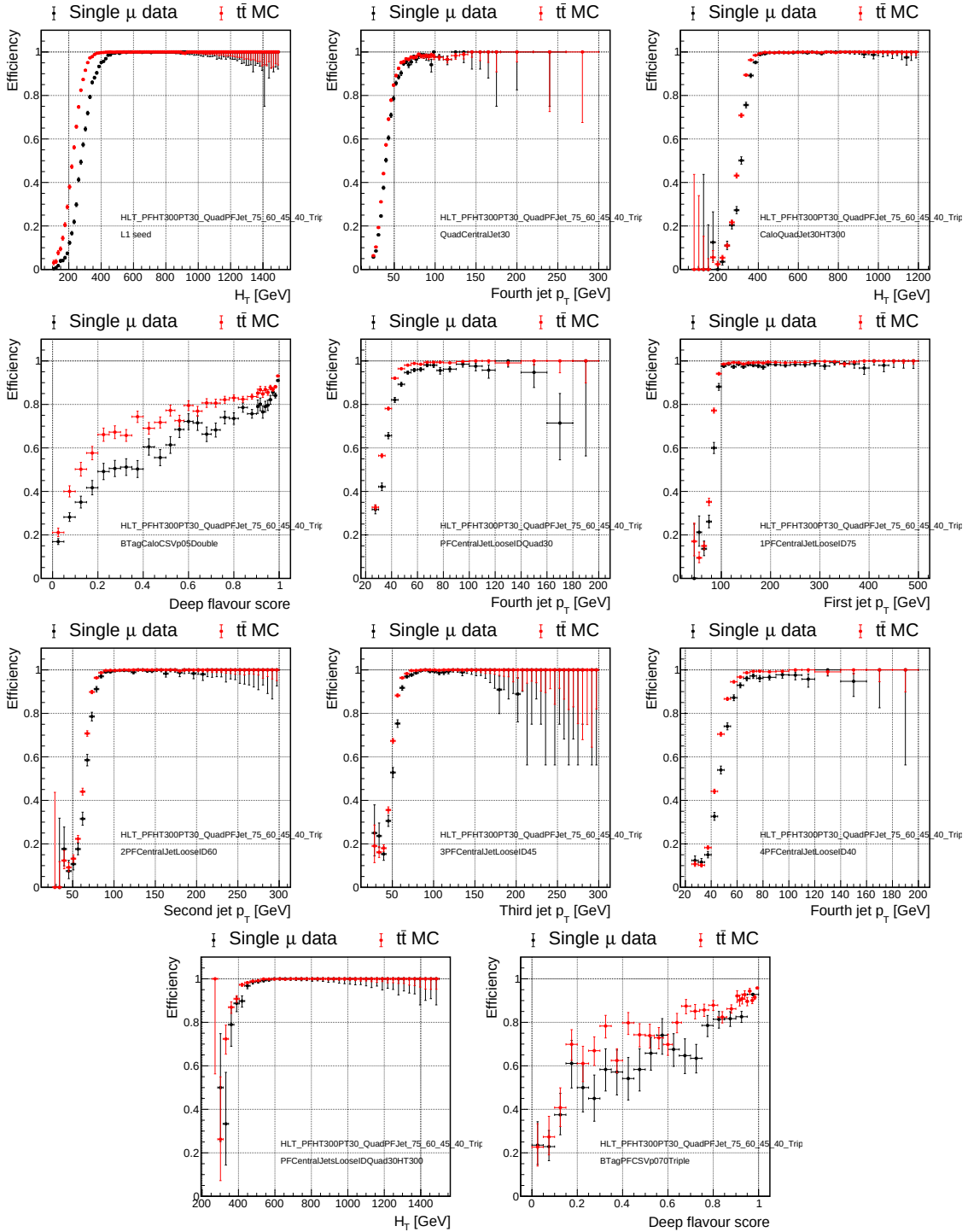


Figure A-4. Efficiency measured in single muon data (black) and  $t\bar{t}$  MC (red) for filters of HLT\_PFHT300PT30\_QuadPFJet\_75\_60\_45\_40\_TriplePFBTagCSV, corresponding to the 2017 dataset.

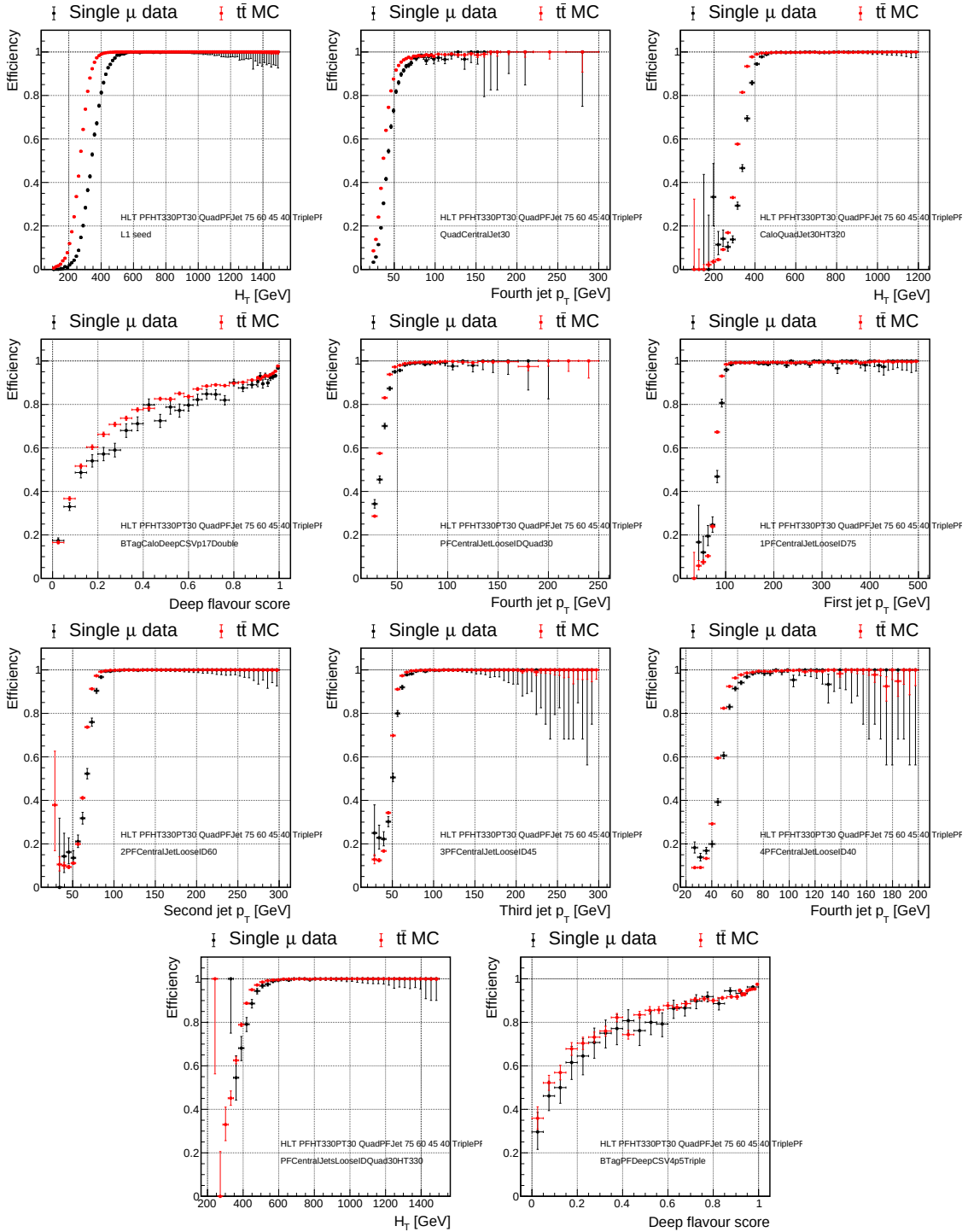


Figure A-5. Efficiency measured in single muon data (black) and  $t\bar{t}$  MC (red) for filters of HLT\_PFHT330PT30\_QuadPFJet\_75\_60\_45\_40\_TriplePFBTagDeepCSV, corresponding to the 2018 dataset.

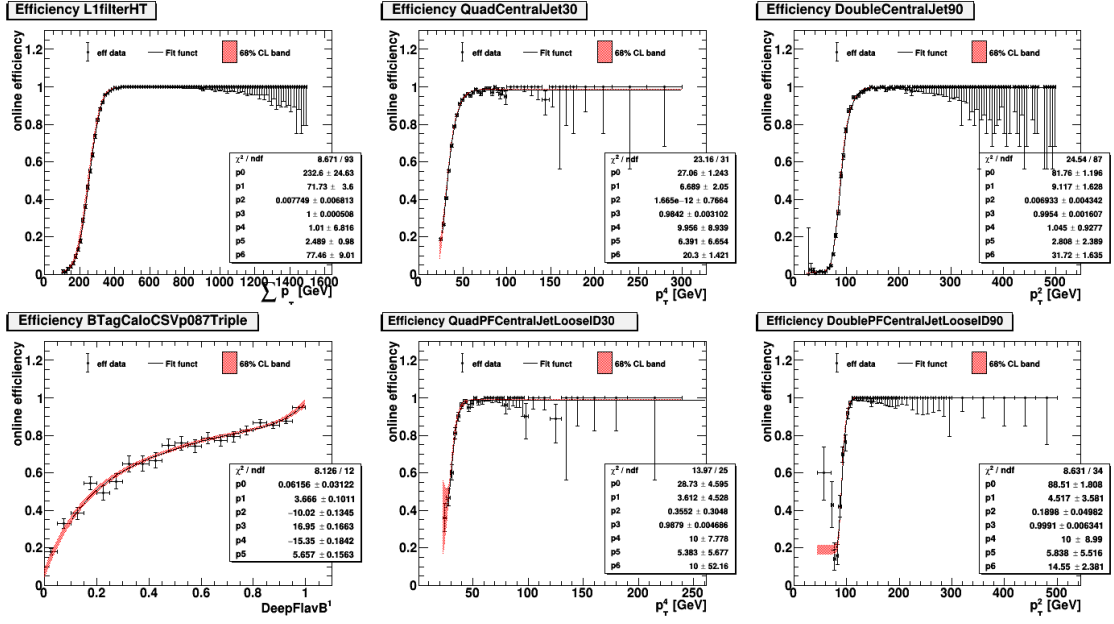


Figure A-6. Efficiency fits in SingleMuon filters of HLT\_DoubleJet90\_Double30\_TripleBTagCSV in 2016. The red band represents the 1 sigma error on the fit obtained from the fit parameter error propagation.

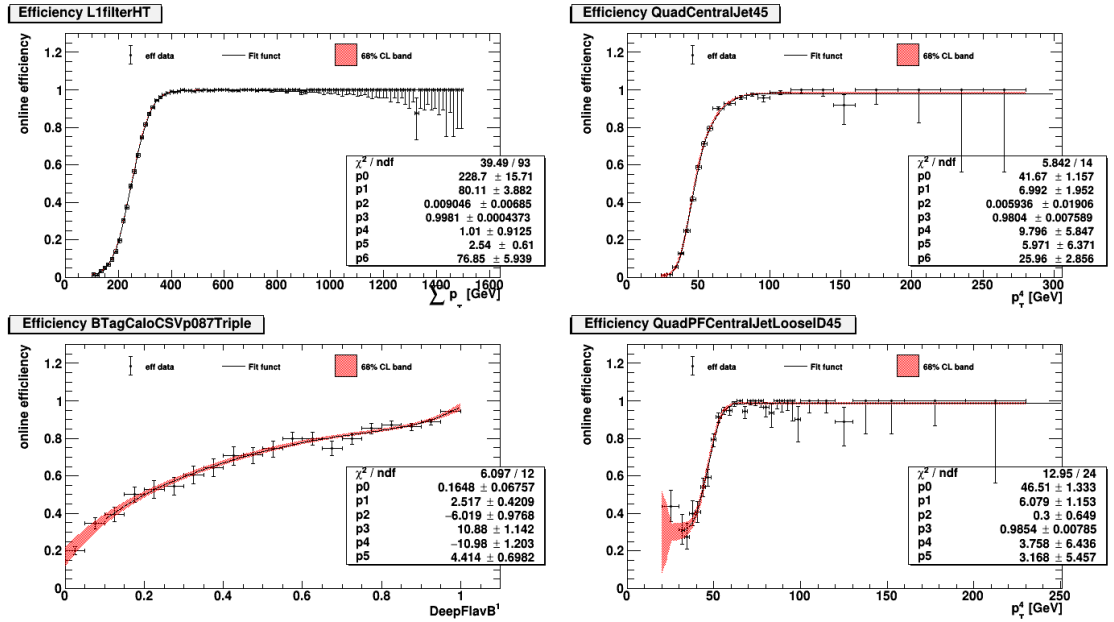


Figure A-7. Efficiency fits in SingleMuon for filters of HLT\_QuadJet45\_TripleBTagCSV in 2016. The red band represents the 1 sigma error on the fit obtained from the fit parameter error propagation.

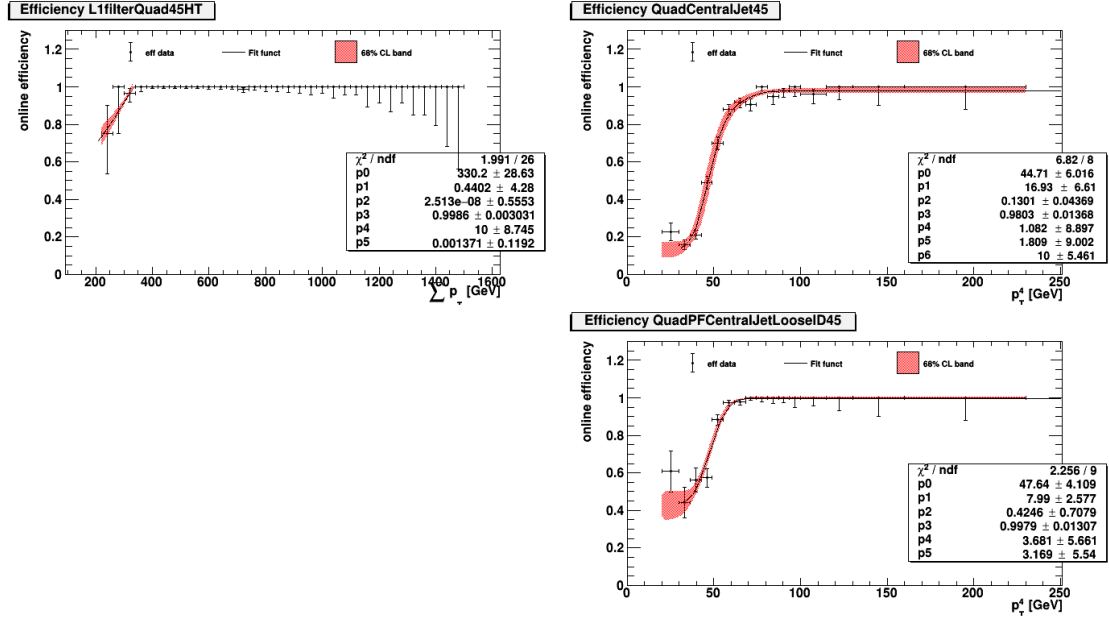


Figure A-8. Efficiency fits in SingleMuon for filters of HLT\_QuadJet45\_TripleBTagCSV evaluated over a sample of events passing HLT\_DoubleJet90\_Double30\_TripleBTagCSV in 2016. The red band represents the 1 sigma error on the fit obtained from the fit parameter error propagation.

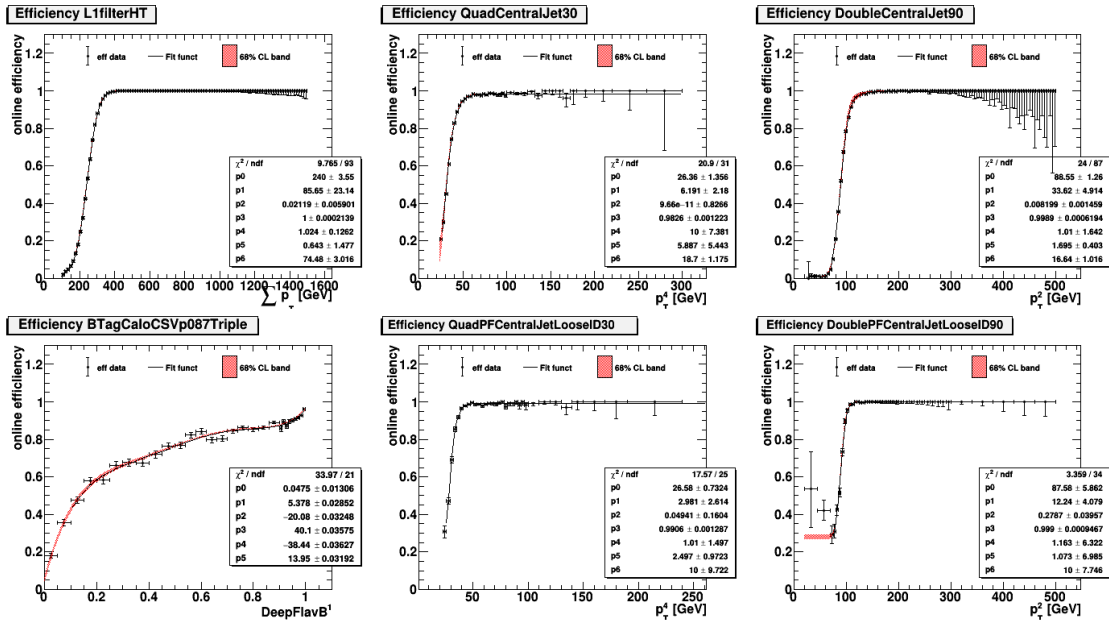


Figure A-9. Efficiency fits in  $t\bar{t}$  for filters of HLT\_DoubleJet90\_Double30\_TripleBTagCSV in 2016. The red band represents the 1 sigma error on the fit obtained from the fit parameter error propagation.



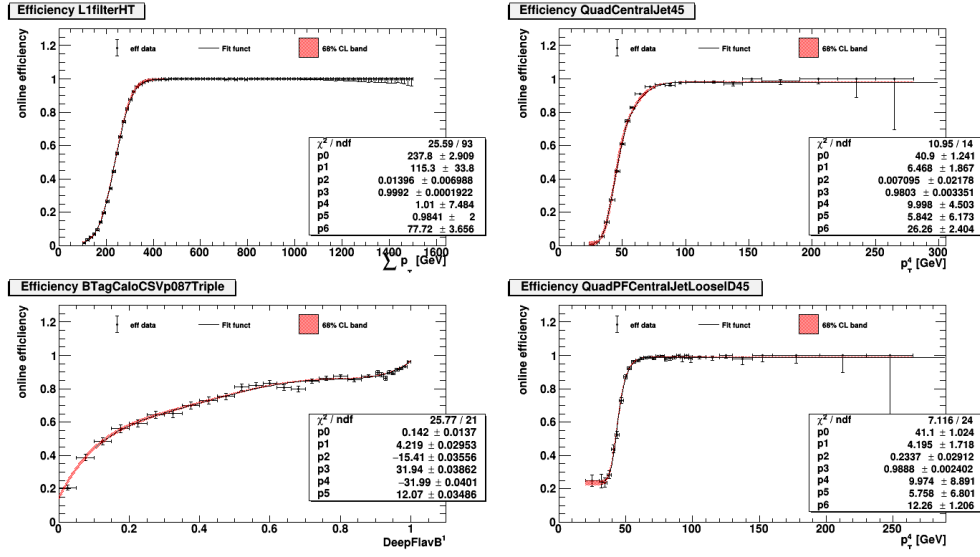


Figure A-10. Efficiency fits in  $t\bar{t}$  for filters of HLT\_QuadJet45\_TripleBTagCSV in 2016. The red band represents the 1 sigma error on the fit obtained from the fit parameter error propagation.

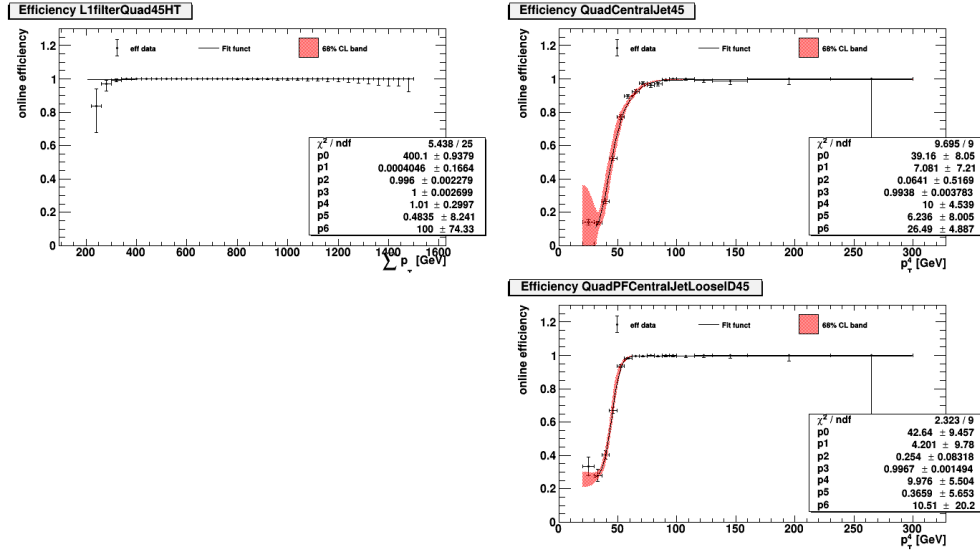


Figure A-11. Efficiency fits in  $t\bar{t}$  for filters of HLT\_QuadJet45\_TripleBTagCSV evaluated over a sample of events passing HLT\_DoubleJet90\_Double30\_TripleBTagCSV in 2016. The red band represents the 1 sigma error on the fit obtained from the fit parameter error propagation. Due to the limited statistic of the first two bins, the fit is not able to model the turn-on curve, however it was verified that this does not cause a problem in the model thanks to the closure test in 2016 signal.

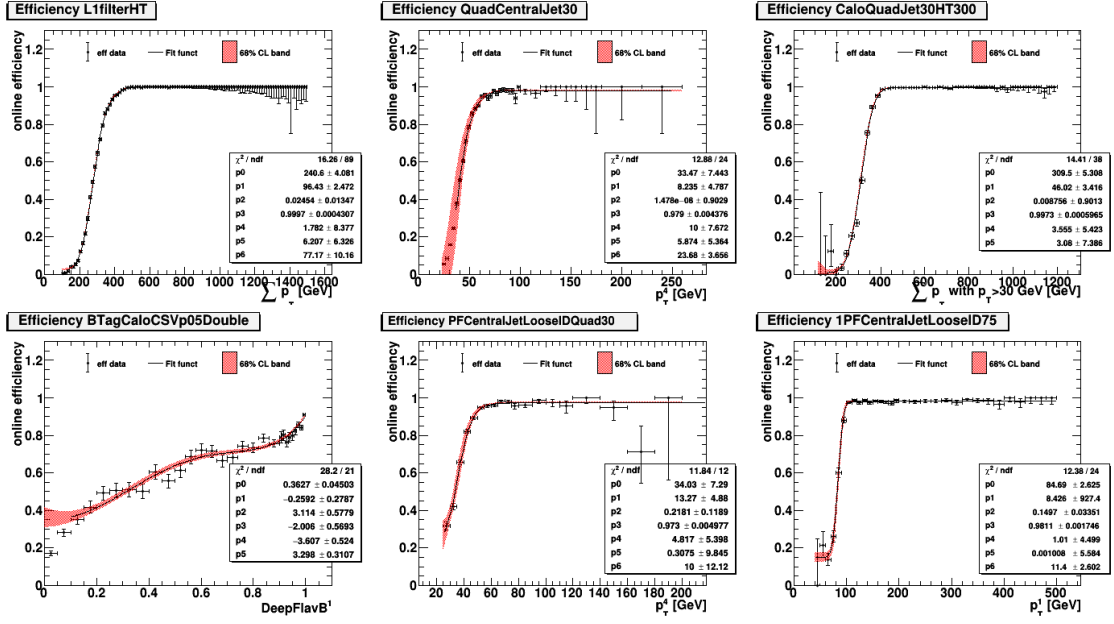


Figure A-12. Efficiency fits in SingleMuon for filters of the trigger path in 2017 (1/2). The red band represents the 1 sigma error on the fit obtained from the fit parameter error propagation.

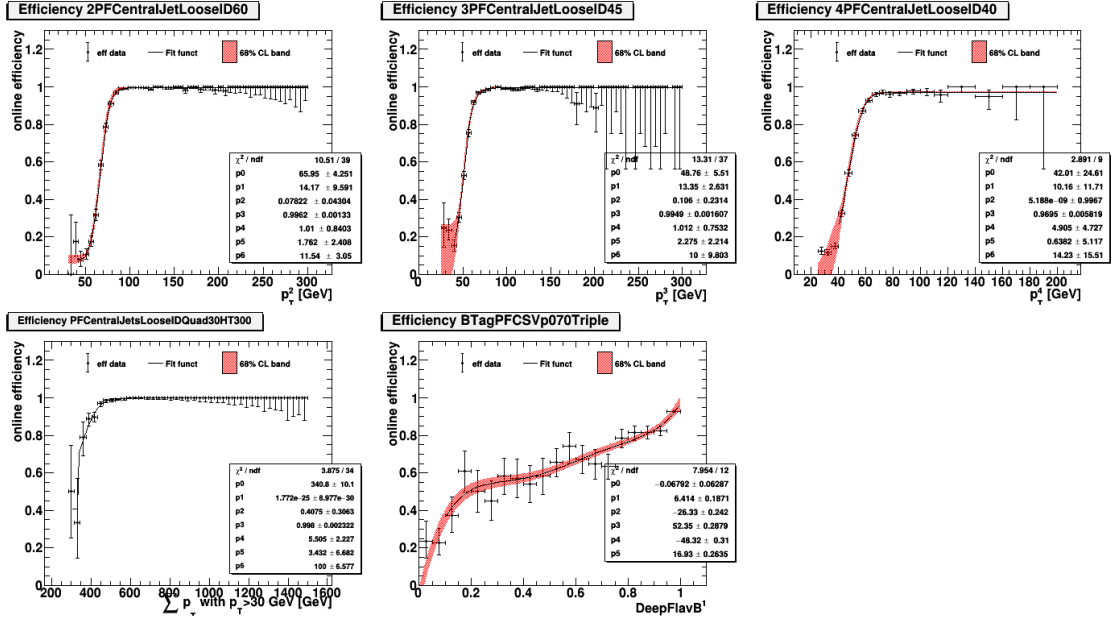


Figure A-13. Efficiency fits in SingleMuon for filters of the trigger path in 2017 (2/2). The red band represents the 1 sigma error on the fit obtained from the fit parameter error propagation.

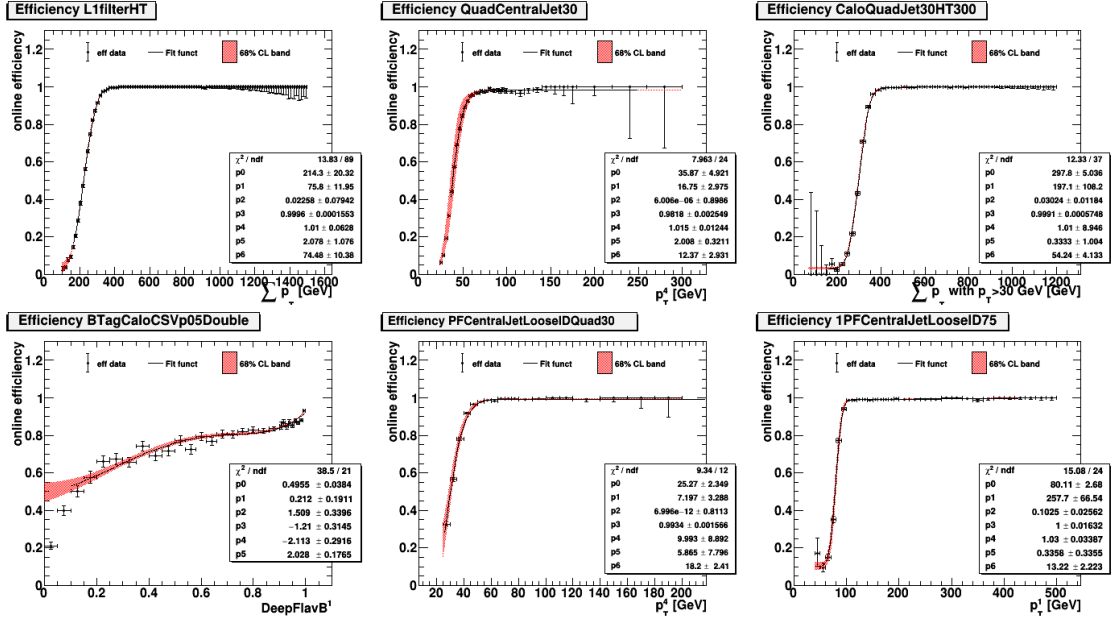


Figure A-14. Efficiency fits in  $t\bar{t}$  for filters of the trigger path in 2017 (1/2). The red band represents the 1 sigma error on the fit obtained from the fit parameter error propagation.

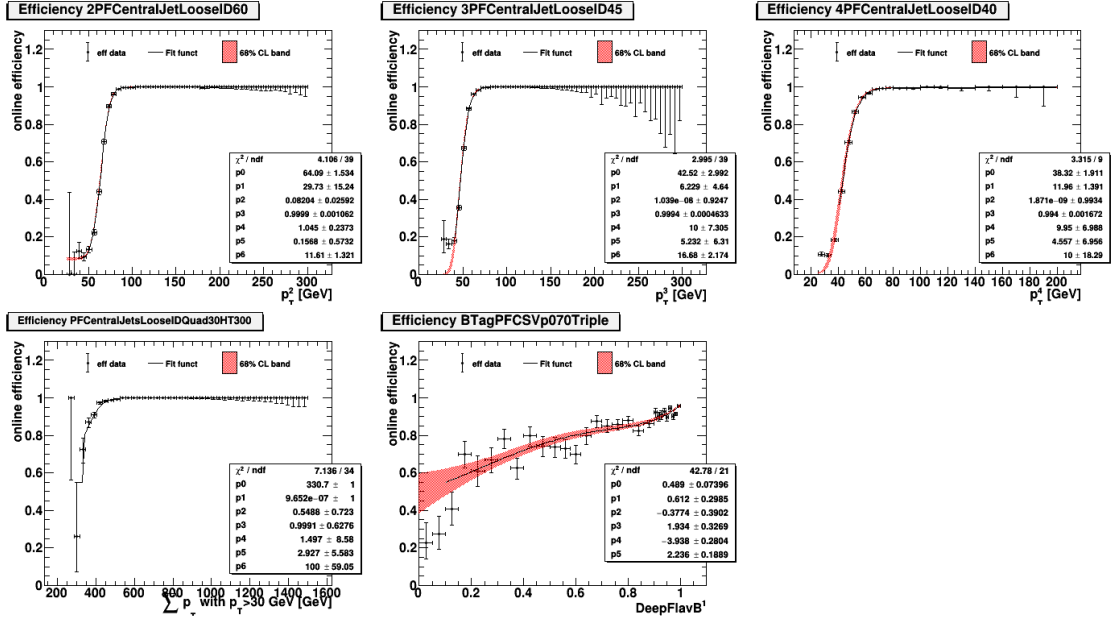


Figure A-15. Efficiency fits in  $t\bar{t}$  for filters of the trigger path in 2017 (2/2). The red band represents the 1 sigma error on the fit obtained from the fit parameter error propagation.

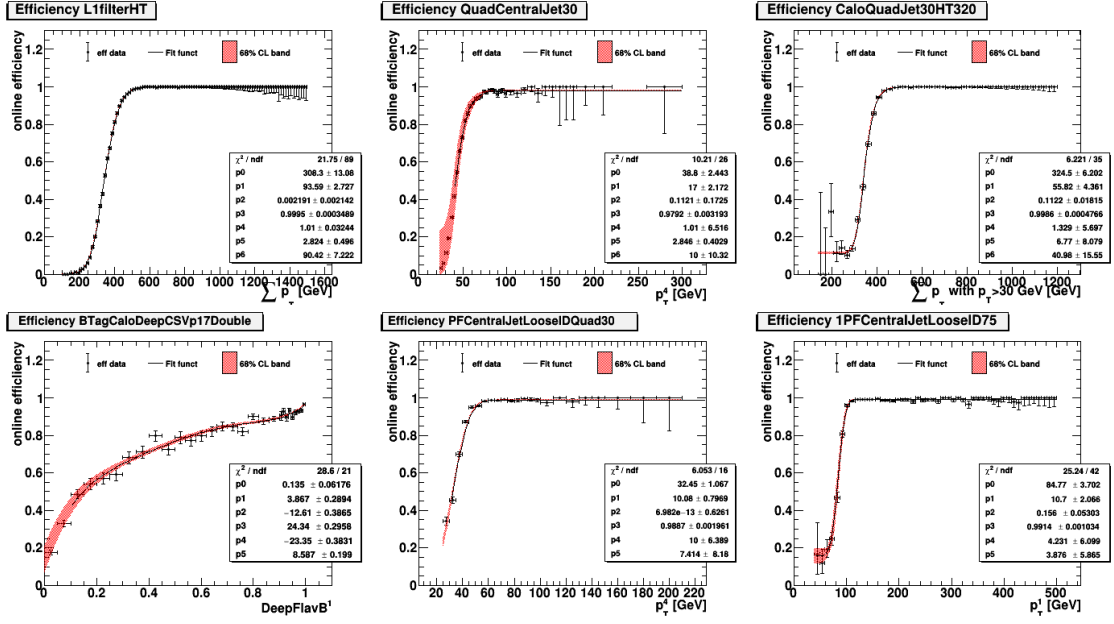


Figure A-16. Efficiency fits in SingleMuon for filters of the trigger path in 2018 (1/2). The red band represents the 1 sigma error on the fit obtained from the fit parameter error propagation.

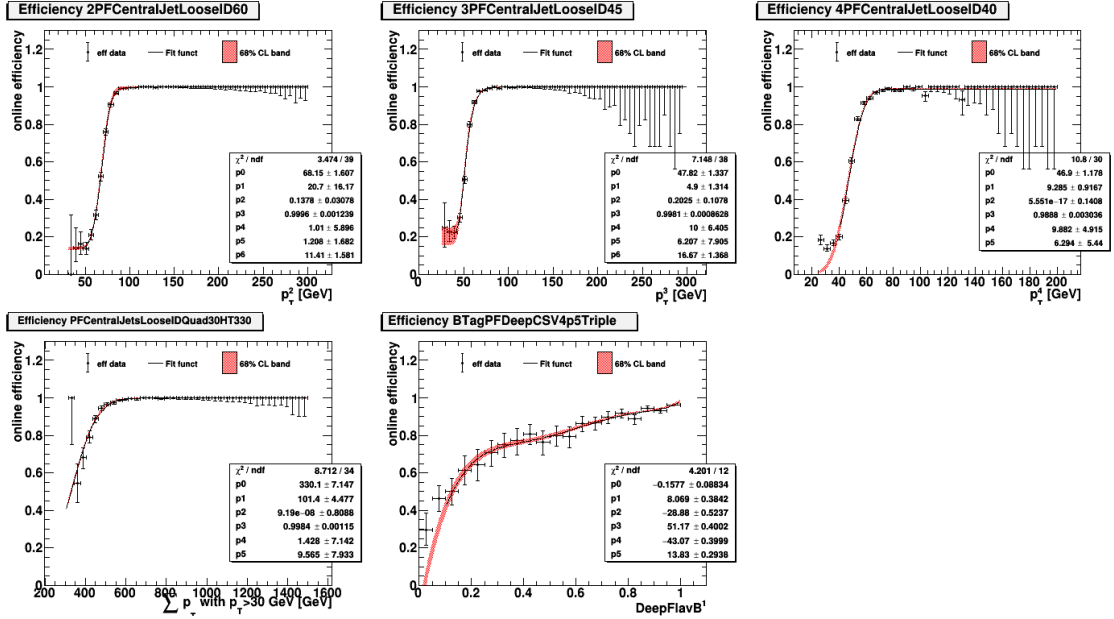


Figure A-17. Efficiency fits in SingleMuon for filters of the trigger path in 2018 (2/2). The red band represents the 1 sigma error on the fit obtained from the fit parameter error propagation.

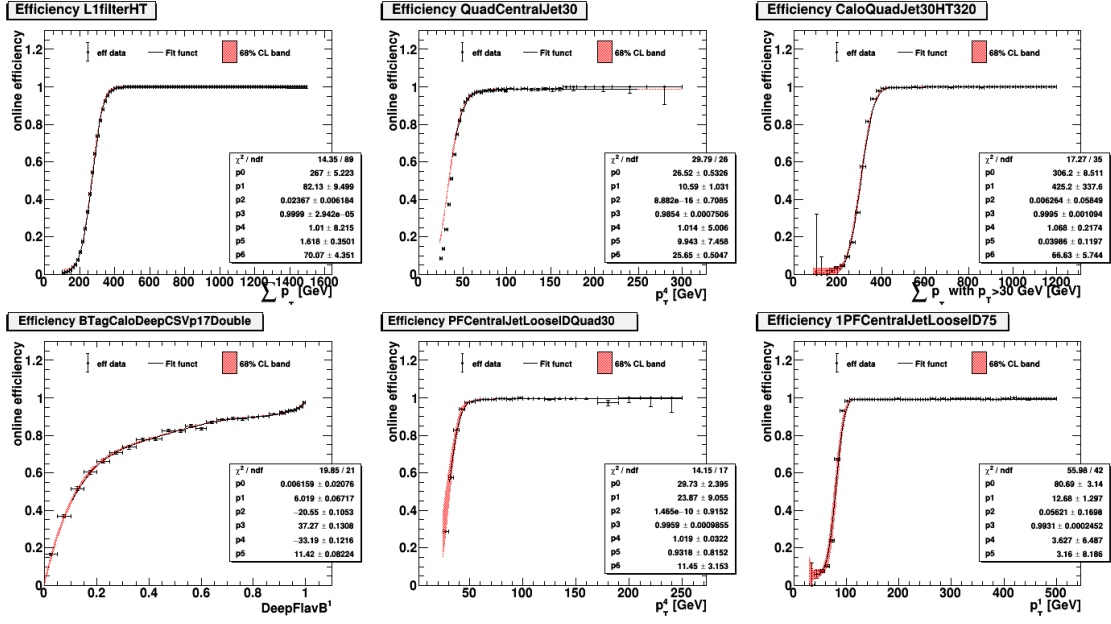


Figure A-18. Efficiency fits in  $t\bar{t}$  for filters of the trigger path in 2018 (1/2). The red band represents the 1 sigma error on the fit obtained from the fit parameter error propagation.

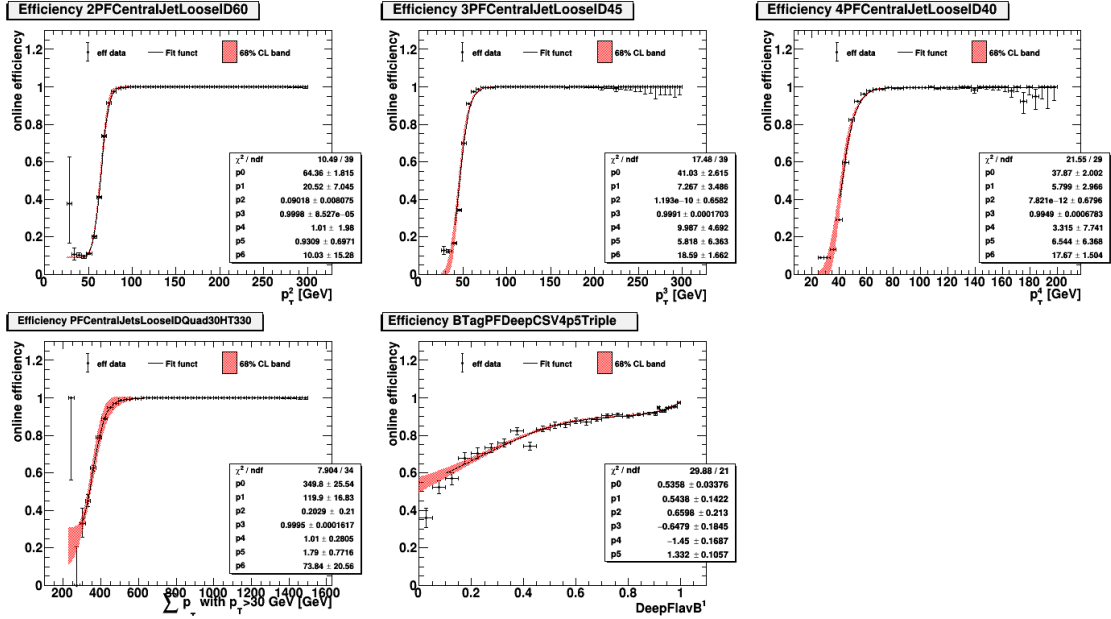


Figure A-19. Efficiency fits in  $t\bar{t}$  for filters of the trigger path in 2018 (2/2). The red band represents the 1 sigma error on the fit obtained from the fit parameter error propagation.

## APPENDIX B GGFKILLER SCORE

This appendix describes more details on the ggFKiller training. The simulation events were separated into 75% training and 25% test. The ggFKiller chosen hyperparameters for the 2016, 2017 and 2018 dataset are presented in Table B-1. For completeness, the distributions of the input variables in signal and background for 2017 and 2018 are presented in Figure B-1, and Figure B-2, respectively. The optimization procedure is summarized as follows:

- Eight grids with 4 points of BDT-hyperparameters each are defined. Each point is a different combination of number of estimators, maximum depth, maximum learning rate
- The best of point of each grid is found using a 2-fold cross-validation training performed using the GridSearchCV method implemented in scikit-learn [128]. In essence, the procedure is the following: 1) The training sample is divided in 2 folds, 2) Iteratively one fold is used for training and the remaining one for the validation, 3) The mean of the 2 ROC-AUCs in the validation sample for each point of the grid is calculated. The best point of the grid is the one which maximizes this mean.
- The classifier associated to the best point of each grid is refitted in the complete training sample, and then applied to the test sample. The best set of BDT-hyperparameters is the one maximizing the ROC-AUC in the test sample and also passing the Kolmogorov-Smirnov overtraining-check (just like is done in TMVA [140]) in the training and test samples with at least 5% probability.

Table B-1. Best hyperparameters of the ggFKiller training.

| Training Parameters | 2016            | 2017            | 2018            |
|---------------------|-----------------|-----------------|-----------------|
| BoostType           | Gradient Boost  | Gradient Boost  | Gradient Boost  |
| Objective           | binary:logistic | binary:logistic | binary:logistic |
| Number of trees     | 400             | 200             | 600             |
| Max. depth          | 2               | 4               | 2               |
| Learning rate       | 0.1             | 0.1             | 0.1             |

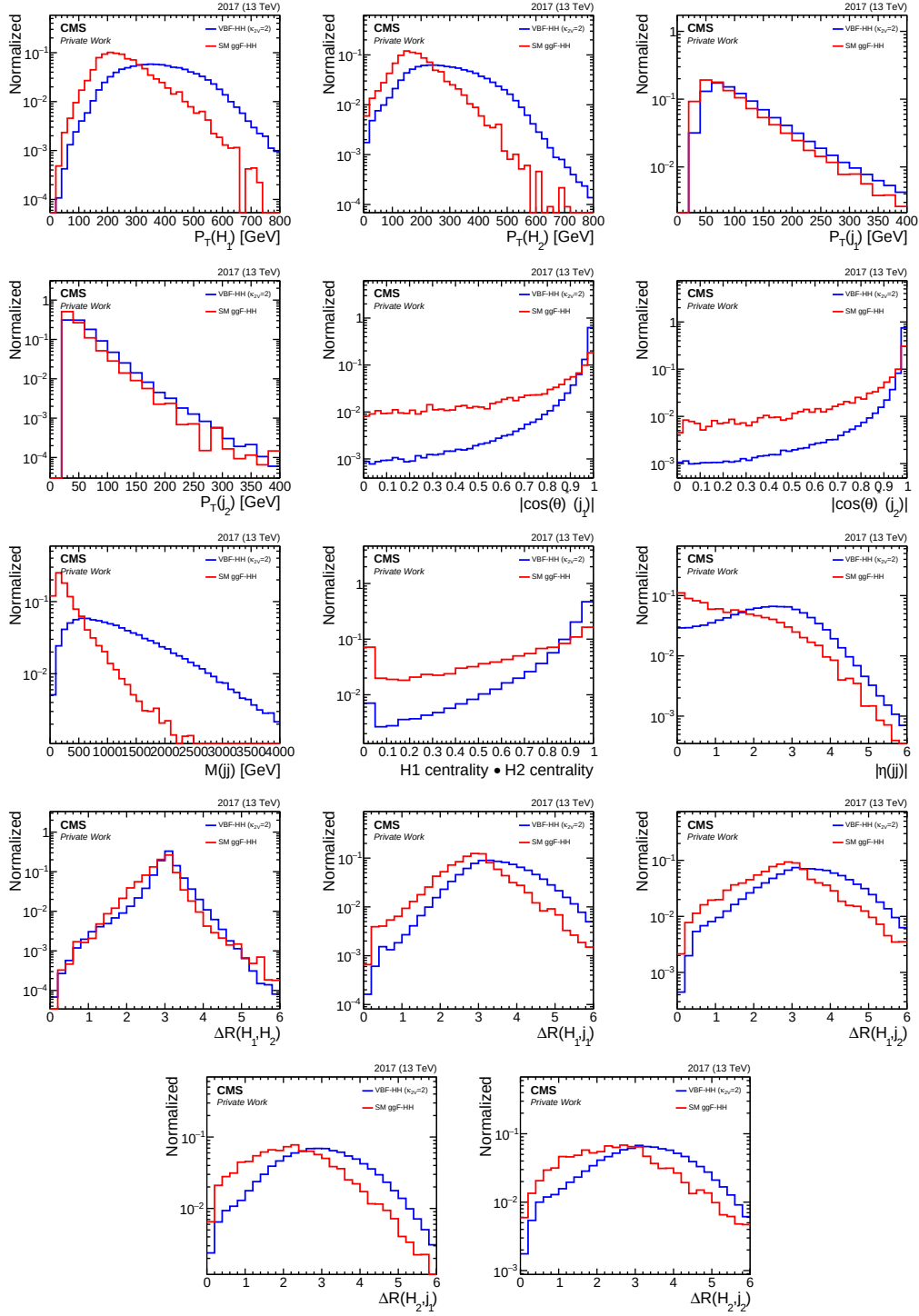


Figure B-1. Distribution of the ggF Killer input variables in 2017 simulation. The VBF with  $\kappa_{2V} = 2$  (signal) is in solid blue, whereas SM ggF (background) is in solid red.

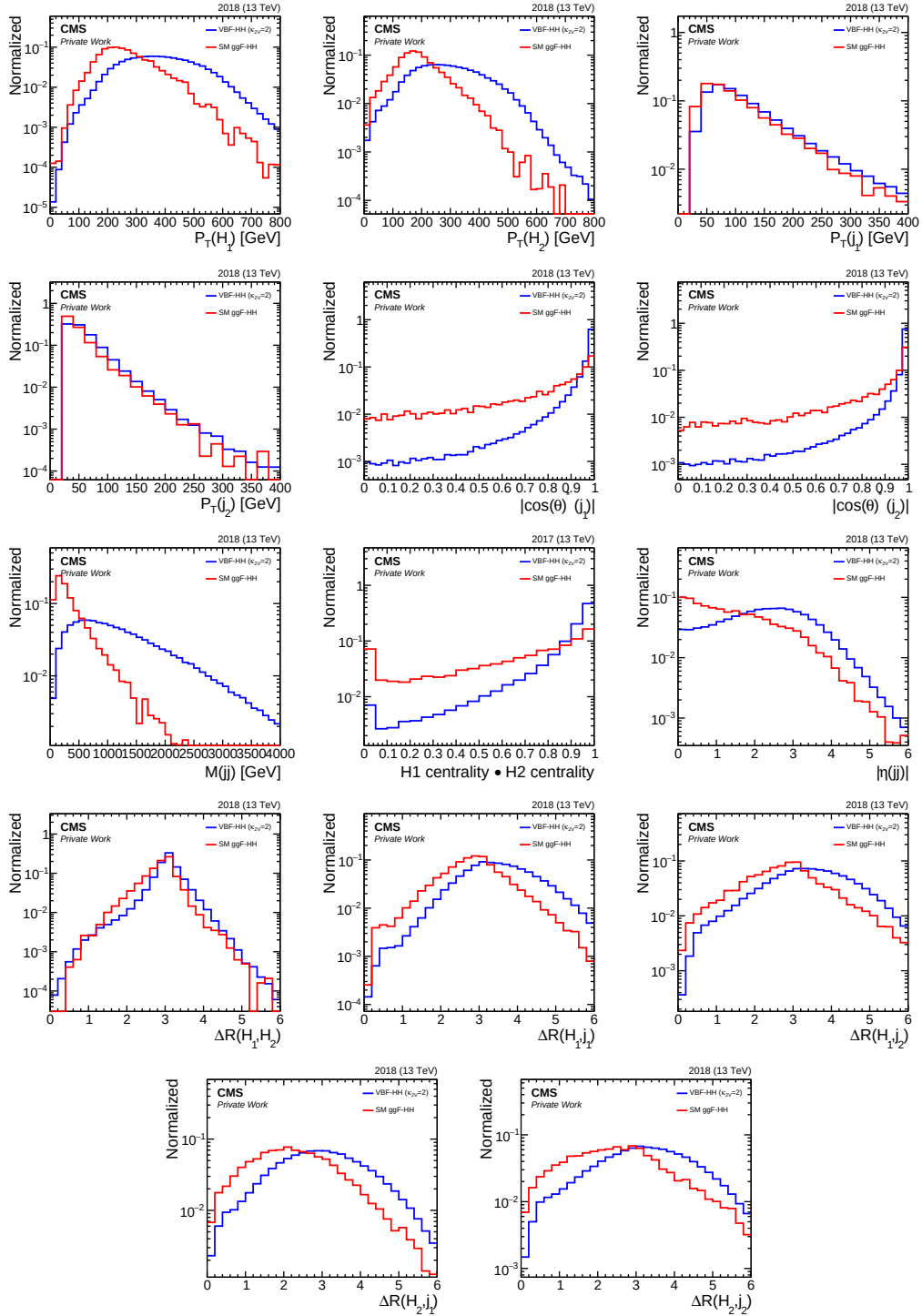


Figure B-2. Distribution of the ggF Killer input variables in 2018 simulation. The VBF with  $\kappa_{2V} = 2$  (signal) is in solid blue, whereas SM ggF (background) is in solid red.



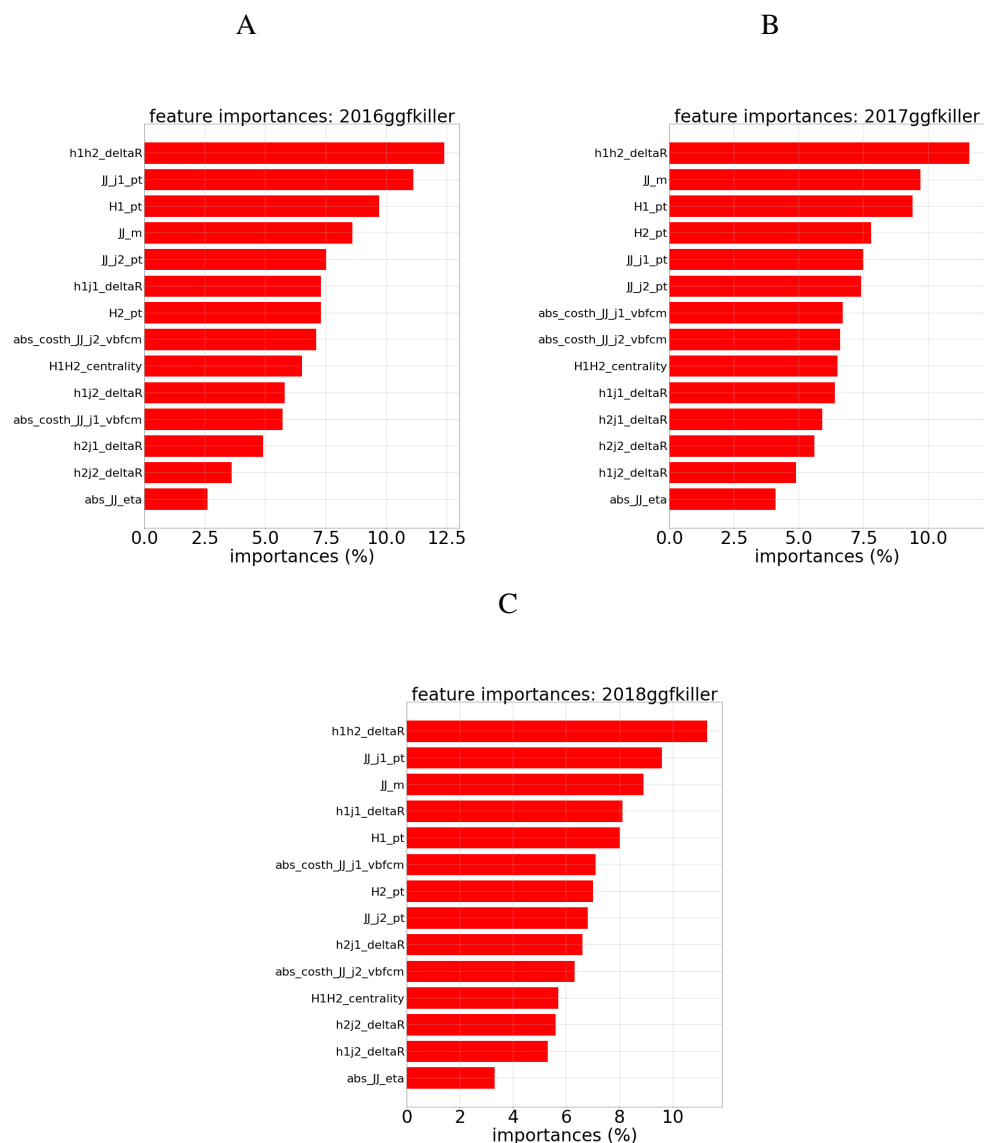
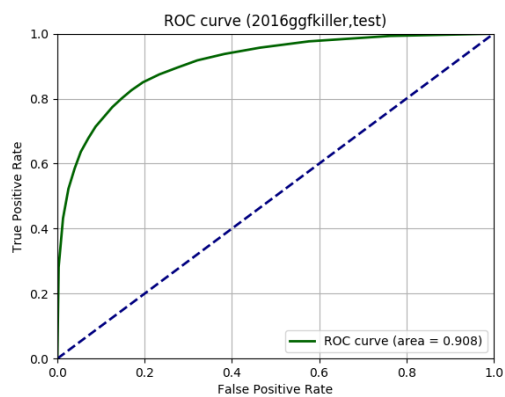
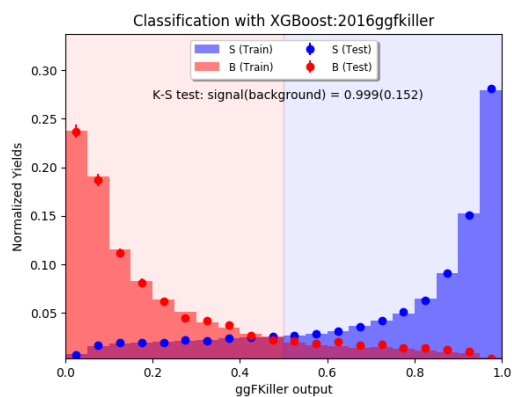


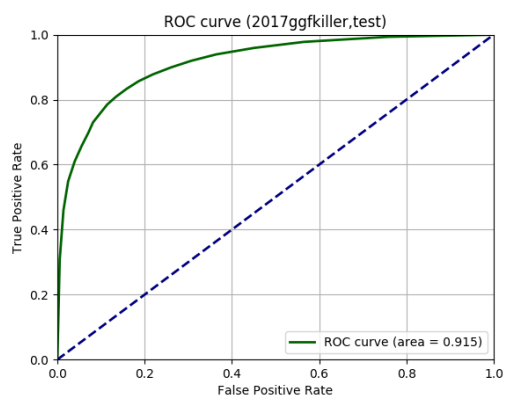
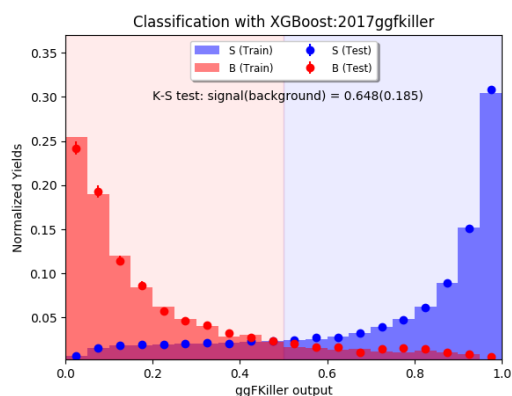
Figure B-3. The relative importance of the ggFKiller training variables. Each row represents a particular simulation year: A) 2016, B) 2017, C) 2018.

The performance of the ggFKiller score in the train and test samples are presented depending on the year on Figure B-4 A) B) and E), whereas the ROC curves in the test sample are presented depending on the year in Figure B-4 B) D) and E). No signs of over-training are seen. The integral of the area under the curve (AUC) for the (train) test samples is (0.914) 0.908 in 2016, (0.923) 0.915 in 2017, and (0.923) 0.912 in 2018. For completeness, the relative variable importance is presented in Figure B-3.

A



B



C

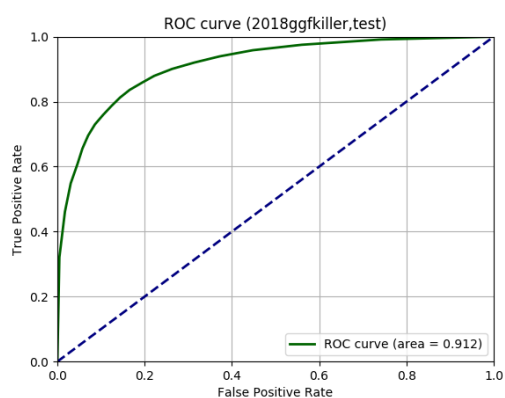
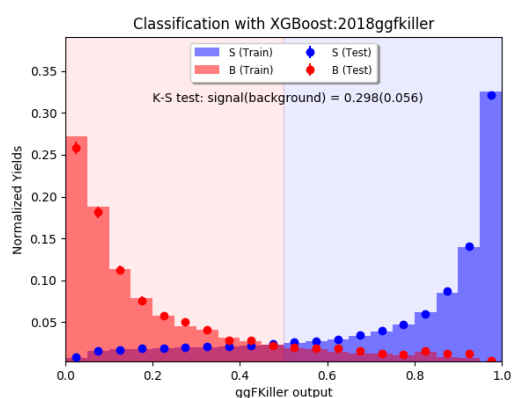


Figure B-4. Over-training check of the ggFKiller score. The left figure is the distribution of the score in train and test samples. The right figure is the test sample ROC curve performance. A) 2016, B) 2017 and C) 2018.

# APPENDIX C SIGNAL EFFICIENCIES VS HIGGS BOSON COUPLINGS

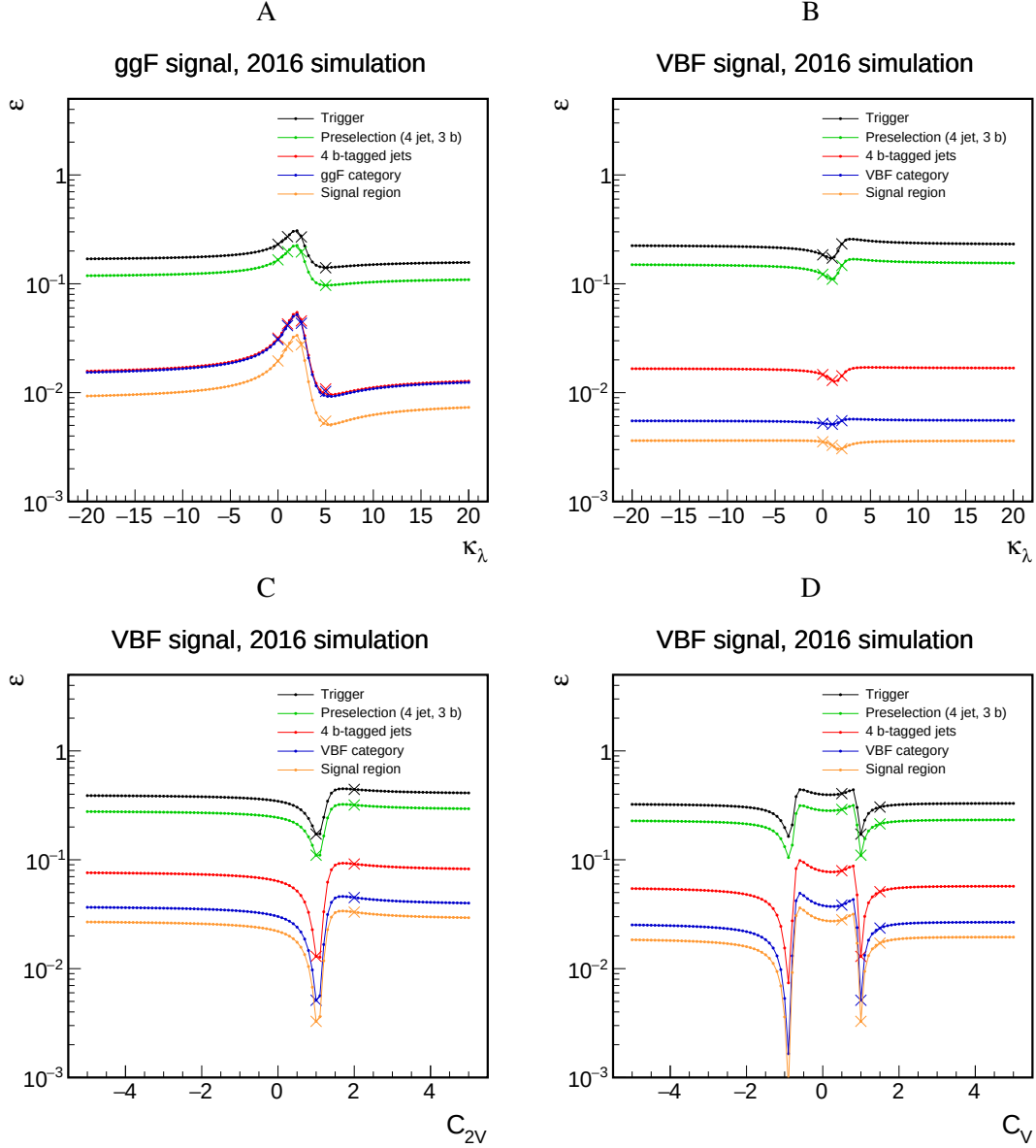


Figure C-1. Selection efficiency ( $\epsilon$ ) of the 2016 selections versus the ggF and VBF couplings.  $\epsilon$  vs  $\kappa_\lambda$  for the A) ggF and B) VBF signals, C)  $\epsilon$  versus  $\kappa_{2V}$  and D)  $\epsilon$  versus  $\kappa_V$  for the VBF signals. In all figures, the couplings that are not displayed on the x-axis are assumed to be fixed to the SM value. The crosses indicate the values obtained directly from the simulated samples, serving as cross-check to the values computed with the signal modeling procedure (denoted by the full circles).

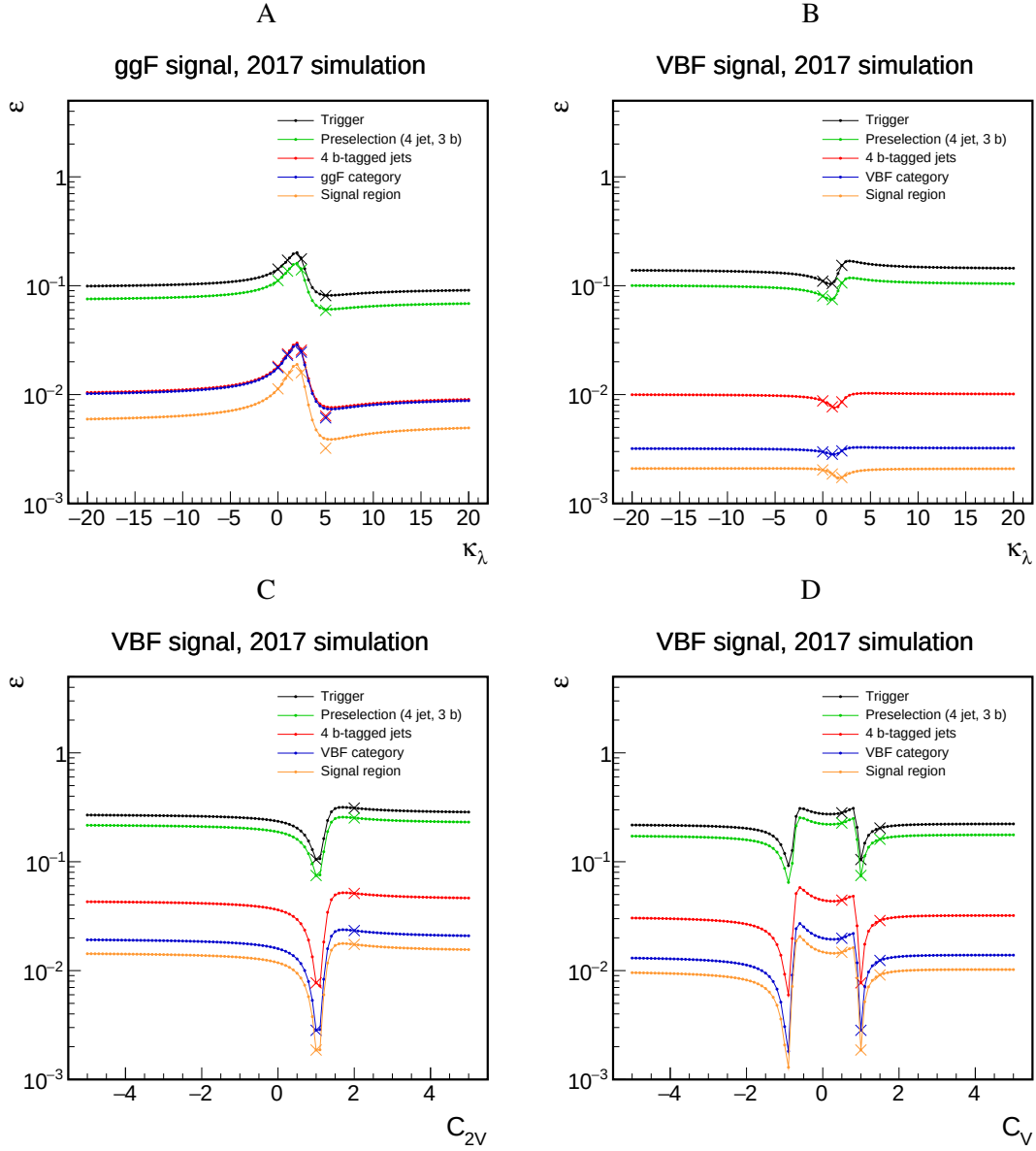


Figure C-2. Selection efficiency ( $\epsilon$ ) of the 2017 selections versus the ggF and VBF couplings.  $\epsilon$  vs  $\kappa_\lambda$  for the A) ggF and B) VBF signals, C)  $\epsilon$  versus  $\kappa_{2V}$  and D)  $\epsilon$  versus  $\kappa_V$  for the VBF signals. In all figures, the couplings that are not displayed on the x-axis are assumed to be fixed to the SM value. The crosses indicate the values obtained directly from the simulated samples, serving as cross-check to the values computed with the signal modelling procedure (denoted by the full circles).

# APPENDIX D

## LIST OF SIMULATION SAMPLES

Simulated samples are used to model the signal processes as well as minor backgrounds. A full list of the MC samples is reported in Tables D-1, D-2, and D-3 for the data taking conditions of 2016, 2017, and 2018, respectively. The normalization cross section also shown.

Table D-1. Summary of the MC samples used for the 2016 conditions. All the samples are in the CMS NANOASIM data tier. The cross section includes the branching fraction.

| Process                                                                                             | Event generators  | Cross section [pb] |
|-----------------------------------------------------------------------------------------------------|-------------------|--------------------|
| QCD, $200 < H_T < 300$ GeV                                                                          | MadGraph, Pythia8 | 1710000            |
| QCD, $300 < H_T < 500$ GeV                                                                          | MadGraph, Pythia8 | 347500             |
| QCD, $500 < H_T < 700$ GeV                                                                          | MadGraph, Pythia8 | 32060              |
| QCD, $700 < H_T < 1000$ GeV                                                                         | MadGraph, Pythia8 | 6829               |
| QCD, $1000 < H_T < 1500$ GeV                                                                        | MadGraph, Pythia8 | 1207               |
| QCD, $1500 < H_T < 2000$ GeV                                                                        | MadGraph, Pythia8 | 120                |
| QCD, $H_T > 2000$ GeV                                                                               | MadGraph, Pythia8 | 25.25              |
| $t\bar{t}$                                                                                          | PowHeg, Pythia8   | 831.76             |
| Single $H \rightarrow b\bar{b}$ (ggF)                                                               | PowHeg, Pythia8   | 28.45              |
| Single $H \rightarrow b\bar{b}$ (VBF)                                                               | PowHeg, Pythia8   | 2.202              |
| Single $H \rightarrow b\bar{b}$ ( $W^+H$ )                                                          | PowHeg, Pythia8   | 0.3299             |
| Single $H \rightarrow b\bar{b}$ ( $W^-H$ )                                                          | PowHeg, Pythia8   | 0.209              |
| Single $H \rightarrow b\bar{b}$ (ZH)                                                                | PowHeg, Pythia8   | 0.359              |
| Single $H \rightarrow b\bar{b}$ ( $t\bar{t}H$ )                                                     | PowHeg, Pythia8   | 0.2953             |
| $ZZ \rightarrow b\bar{b}b\bar{b}$                                                                   | MadGraph, Pythia8 | 0.3682             |
| $gg \rightarrow HH \rightarrow b\bar{b}b\bar{b}$ (NLO), $\kappa_\lambda = 1$                        | PowHeg, Pythia8   | 0.010517           |
| $gg \rightarrow HH \rightarrow b\bar{b}b\bar{b}$ (NLO), $\kappa_\lambda = 0$                        | PowHeg, Pythia8   | 0.023618           |
| $gg \rightarrow HH \rightarrow b\bar{b}b\bar{b}$ (NLO), $\kappa_\lambda = 2.45$                     | PowHeg, Pythia8   | 0.004455           |
| $gg \rightarrow HH \rightarrow b\bar{b}b\bar{b}$ (NLO), $\kappa_\lambda = 5$                        | PowHeg, Pythia8   | 0.031072           |
| VBF $HH \rightarrow b\bar{b}b\bar{b}$ , $\kappa_V = 1$ , $\kappa_{2V} = 0$ , $\kappa_\lambda = 1$   | MadGraph, Pythia8 | 0.009169           |
| VBF $HH \rightarrow b\bar{b}b\bar{b}$ , $\kappa_V = 1$ , $\kappa_{2V} = 1$ , $\kappa_\lambda = 1$   | MadGraph, Pythia8 | 0.000585           |
| VBF $HH \rightarrow b\bar{b}b\bar{b}$ , $\kappa_V = 1$ , $\kappa_{2V} = 2$ , $\kappa_\lambda = 1$   | MadGraph, Pythia8 | 0.004823           |
| VBF $HH \rightarrow b\bar{b}b\bar{b}$ , $\kappa_V = 1$ , $\kappa_{2V} = 1$ , $\kappa_\lambda = 2$   | MadGraph, Pythia8 | 0.000482           |
| VBF $HH \rightarrow b\bar{b}b\bar{b}$ , $\kappa_V = 1$ , $\kappa_{2V} = 1$ , $\kappa_\lambda = 0$   | MadGraph, Pythia8 | 0.001558           |
| VBF $HH \rightarrow b\bar{b}b\bar{b}$ , $\kappa_V = 0.5$ , $\kappa_{2V} = 1$ , $\kappa_\lambda = 1$ | MadGraph, Pythia8 | 0.003656           |
| VBF $HH \rightarrow b\bar{b}b\bar{b}$ , $\kappa_V = 1.5$ , $\kappa_{2V} = 1$ , $\kappa_\lambda = 1$ | MadGraph, Pythia8 | 0.022412           |
| EFT $gg \rightarrow HH \rightarrow b\bar{b}b\bar{b}$ (LO)                                           | MadGraph, Pythia8 | 0.010517           |

Table D-2. Summary of the MC samples used for the 2017 conditions. The LO VBF HH samples (\*\*) were generated with Pythia8 dipole recoil mode on (default mode is off). All the samples are in the CMS NANOASIM data tier. The cross section includes the branching fraction.

| Process                                                                                                 | Event generators  | Cross section [pb] |
|---------------------------------------------------------------------------------------------------------|-------------------|--------------------|
| QCD, $200 < H_T < 300$ GeV                                                                              | MadGraph, Pythia8 | 1547000            |
| QCD, $300 < H_T < 500$ GeV                                                                              | MadGraph, Pythia8 | 322600             |
| QCD, $500 < H_T < 700$ GeV                                                                              | MadGraph, Pythia8 | 29980              |
| QCD, $700 < H_T < 1000$ GeV                                                                             | MadGraph, Pythia8 | 6334               |
| QCD, $1000 < H_T < 1500$ GeV                                                                            | MadGraph, Pythia8 | 1088               |
| QCD, $1500 < H_T < 2000$ GeV                                                                            | MadGraph, Pythia8 | 99.11              |
| QCD, $H_T > 2000$ GeV                                                                                   | MadGraph, Pythia8 | 20.23              |
| $t\bar{t}$ fully hadronic                                                                               | PowHeg, Pythia8   | 377.96             |
| $t\bar{t}$ semileptonic                                                                                 | PowHeg, Pythia8   | 365.34             |
| $t\bar{t}$ fully leptonic                                                                               | PowHeg, Pythia8   | 88.29              |
| Single $H \rightarrow b\bar{b}$ (ggF)                                                                   | PowHeg, Pythia8   | 28.45              |
| Single $H \rightarrow b\bar{b}$ (VBF)                                                                   | PowHeg, Pythia8   | 2.202              |
| Single $H \rightarrow b\bar{b}$ ( $W^+H$ )                                                              | PowHeg, Pythia8   | 0.3299             |
| Single $H \rightarrow b\bar{b}$ ( $W^-H$ )                                                              | PowHeg, Pythia8   | 0.209              |
| Single $H \rightarrow b\bar{b}$ (ZH)                                                                    | PowHeg, Pythia8   | 0.359              |
| Single $H \rightarrow b\bar{b}$ ( $t\bar{t}H$ )                                                         | PowHeg, Pythia8   | 0.2953             |
| $ZZ \rightarrow 4b$                                                                                     | PowHeg, Pythia8   | 0.3682             |
| $gg \rightarrow HH \rightarrow b\bar{b}b\bar{b}$ (NLO), $\kappa_\lambda = 1$                            | PowHeg, Pythia8   | 0.010517           |
| $gg \rightarrow HH \rightarrow b\bar{b}b\bar{b}$ (NLO), $\kappa_\lambda = 0$                            | PowHeg, Pythia8   | 0.023618           |
| $gg \rightarrow HH \rightarrow b\bar{b}b\bar{b}$ (NLO), $\kappa_\lambda = 2.45$                         | PowHeg, Pythia8   | 0.004455           |
| $gg \rightarrow HH \rightarrow b\bar{b}b\bar{b}$ (NLO), $\kappa_\lambda = 5$                            | PowHeg, Pythia8   | 0.031072           |
| VBF $HH \rightarrow b\bar{b}b\bar{b}$ , $\kappa_V = 1$ , $\kappa_{2V} = 1$ , $\kappa_\lambda = 1$       | MadGraph, Pythia8 | 0.009169           |
| VBF $HH \rightarrow b\bar{b}b\bar{b}$ , $\kappa_V = 1$ , $\kappa_{2V} = 1$ , $\kappa_\lambda = 1$       | MadGraph, Pythia8 | 0.000585           |
| VBF $HH \rightarrow b\bar{b}b\bar{b}$ , $\kappa_V = 1$ , $\kappa_{2V} = 2$ , $\kappa_\lambda = 1$       | MadGraph, Pythia8 | 0.004823           |
| VBF $HH \rightarrow b\bar{b}b\bar{b}$ , $\kappa_V = 1$ , $\kappa_{2V} = 1$ , $\kappa_\lambda = 2$       | MadGraph, Pythia8 | 0.000482           |
| VBF $HH \rightarrow b\bar{b}b\bar{b}$ , $\kappa_V = 1$ , $\kappa_{2V} = 1$ , $\kappa_\lambda = 0$       | MadGraph, Pythia8 | 0.001558           |
| VBF $HH \rightarrow b\bar{b}b\bar{b}$ , $\kappa_V = 0.5$ , $\kappa_{2V} = 1$ , $\kappa_\lambda = 1$     | MadGraph, Pythia8 | 0.000585           |
| VBF $HH \rightarrow b\bar{b}b\bar{b}$ , $\kappa_V = 1.5$ , $\kappa_{2V} = 1$ , $\kappa_\lambda = 1$     | MadGraph, Pythia8 | 0.004823           |
| VBF $HH \rightarrow b\bar{b}b\bar{b}$ , (**), $\kappa_V = 1$ , $\kappa_{2V} = 1$ , $\kappa_\lambda = 1$ | MadGraph, Pythia8 | 0.000585           |
| VBF $HH \rightarrow b\bar{b}b\bar{b}$ , (**), $\kappa_V = 1$ , $\kappa_{2V} = 2$ , $\kappa_\lambda = 1$ | MadGraph, Pythia8 | 0.004688           |
| EFT $gg \rightarrow HH \rightarrow b\bar{b}b\bar{b}$                                                    | MadGraph, Pythia8 | 0.010517           |

Table D-3. Summary of the MC samples used for the 2018 conditions. The LO VBF HH samples (\*\*) were generated with Pythia8 dipole recoil mode on (default mode is off). All the samples are in the NANOAODSIM data tier. The cross section includes the branching fraction.

| Process                                                                                   | Event generators  | Cross section [pb] |
|-------------------------------------------------------------------------------------------|-------------------|--------------------|
| QCD, $200 < H_T < 300$ GeV                                                                | MadGraph, Pythia8 | 1547000            |
| QCD, $300 < H_T < 500$ GeV                                                                | MadGraph, Pythia8 | 322600             |
| QCD, $500 < H_T < 700$ GeV                                                                | MadGraph, Pythia8 | 29980              |
| QCD, $700 < H_T < 1000$ GeV                                                               | MadGraph, Pythia8 | 6334               |
| QCD, $1000 < H_T < 1500$ GeV                                                              | MadGraph, Pythia8 | 1088               |
| QCD, $1500 < H_T < 2000$ GeV                                                              | MadGraph, Pythia8 | 99.11              |
| QCD, $H_T > 2000$ GeV                                                                     | MadGraph, Pythia8 | 20.23              |
| $t\bar{t}$ fully hadronic                                                                 | PowHeg, Pythia8   | 377.96             |
| $t\bar{t}$ semileptonic                                                                   | PowHeg, Pythia8   | 365.34             |
| $t\bar{t}$ fully leptonic                                                                 | PowHeg, Pythia8   | 88.29              |
| Single $H \rightarrow b\bar{b}$ (ggF)                                                     | PowHeg, Pythia8   | 28.45              |
| Single $H \rightarrow b\bar{b}$ (VBF)                                                     | PowHeg, Pythia8   | 2.202              |
| Single $H \rightarrow b\bar{b}$ ( $W^+H$ )                                                | PowHeg, Pythia8   | 0.3299             |
| Single $H \rightarrow b\bar{b}$ ( $W^-H$ )                                                | PowHeg, Pythia8   | 0.209              |
| Single $H \rightarrow b\bar{b}$ (ZH)                                                      | PowHeg, Pythia8   | 0.359              |
| Single $H \rightarrow b\bar{b}$ ( $t\bar{t}H$ )                                           | PowHeg, Pythia8   | 0.2953             |
| $ZZ \rightarrow 4b$                                                                       | PowHeg, Pythia8   | 0.3682             |
| $gg \rightarrow HH \rightarrow bbbb$ (NLO), $\kappa_\lambda = 1$                          | PowHeg, Pythia8   | 0.010517           |
| $gg \rightarrow HH \rightarrow bbbb$ (NLO), $\kappa_\lambda = 0$                          | PowHeg, Pythia8   | 0.023618           |
| $gg \rightarrow HH \rightarrow bbbb$ (NLO), $\kappa_\lambda = 2.45$                       | PowHeg, Pythia8   | 0.004455           |
| $gg \rightarrow HH \rightarrow bbbb$ (NLO), $\kappa_\lambda = 5$                          | PowHeg, Pythia8   | 0.031072           |
| VBF $HH \rightarrow bbbb$ , $\kappa_V = 1$ , $\kappa_{2V} = 0$ , $\kappa_\lambda = 1$     | MadGraph, Pythia8 | 0.009169           |
| VBF $HH \rightarrow bbbb$ , $\kappa_V = 1$ , $\kappa_{2V} = 1$ , $\kappa_\lambda = 1$     | MadGraph, Pythia8 | 0.000585           |
| VBF $HH \rightarrow bbbb$ , $\kappa_V = 1$ , $\kappa_{2V} = 2$ , $\kappa_\lambda = 1$     | MadGraph, Pythia8 | 0.004823           |
| VBF $HH \rightarrow bbbb$ , $\kappa_V = 1$ , $\kappa_{2V} = 1$ , $\kappa_\lambda = 2$     | MadGraph, Pythia8 | 0.000482           |
| VBF $HH \rightarrow bbbb$ , $\kappa_V = 1$ , $\kappa_{2V} = 1$ , $\kappa_\lambda = 0$     | MadGraph, Pythia8 | 0.001558           |
| VBF $HH \rightarrow bbbb$ , $\kappa_V = 0.5$ , $\kappa_{2V} = 1$ , $\kappa_\lambda = 1$   | MadGraph, Pythia8 | 0.003656           |
| VBF $HH \rightarrow bbbb$ , $\kappa_V = 1.5$ , $\kappa_{2V} = 1$ , $\kappa_\lambda = 1$   | MadGraph, Pythia8 | 0.022412           |
| VBF $HH \rightarrow bbbb$ (**), $\kappa_V = 1$ , $\kappa_{2V} = 1$ , $\kappa_\lambda = 1$ | MadGraph, Pythia8 | 0.000585           |
| VBF $HH \rightarrow bbbb$ (**), $\kappa_V = 1$ , $\kappa_{2V} = 2$ , $\kappa_\lambda = 1$ | MadGraph, Pythia8 | 0.004823           |
| EFT $gg \rightarrow HH \rightarrow bbbb$ (LO)                                             | MadGraph, Pythia8 | 0.010517           |

## APPENDIX E SIGNAL MODEL

### E.1 Gluon Fusion Production

Two different techniques are described to model the ggF HH signal. The first is the LO-reweighting model, which is the same used in all the 2016 nonresonant public results, including the bbbb analysis [59], and the 2016 HH combination [50]. This event-based reweighting method has the advantage of using the full statistical power of the simulated samples, as it does attribute a positive weight to all simulated HH events. However, it is only applicable to LO signal samples because it cannot model events generated with the presence of parton emission at matrix element level and thus higher degrees of freedom (e.g. the NLO ggF signal simulated events). Consequently, another modeling is developed to model the signal using NLO samples.

**LO modeling:** The ggF production at LO is a  $2 \rightarrow 2$  scattering process. At the hard scattering level, the two Higgs bosons are produced back-to-back with the same transverse momentum. The azimuthal angle of the HH system is assumed to be isotropic. Furthermore, any parton distribution function effects in the simulation can be factored out as the Lorentz boost of the HH system is along the beam line. Consequently, the HH system is fully described by two variables: the invariant mass of the HH system ( $m_{HH}$ ) and the absolute value of the cosine of the polar angle of one Higgs boson with respect to the beam axis in the HH rest frame ( $|\cos \theta^*|$ ).

All available LO EFT HH simulated samples (SM and BSM) are combined and the resulting event distribution in the  $m_{HH}$  and  $|\cos \theta^*|$  variables is represented as a two-dimensional histogram with 55 bins and 4 bins and normalized to unity, respectively. A similar histogram is created for the SM sample only. The content of each histogram bin  $k$  is denoted as  $f_{\text{comb}}^k$  and  $f_{\text{SM}}^k$ , respectively. The event weight for an arbitrary coupling combination is defined as a function of  $m_{HH}$  and  $|\cos \theta^*|$  using the differential HH cross section ratio between the BSM target to the SM at each bin  $k$ .

From an EFT expansion assuming a generic HH production at LO, the ratio of the total cross section to the SM expectation ( $R_{HH}$ ) depends on the values of the five EFT anomalous couplings ( $\kappa_\lambda$ ,  $\kappa_t$ ,  $c_g$ ,  $c_{2g}$ ,  $c_2$ ). The ratio value for the bin  $k$  ( $R_{HH}^k$ ) is a polynomial function of the couplings with coefficients  $A_i^k$  ( $i=1,\dots,15$ ), as seen in equation E-1. From generated events, the



value of  $R_{HH}^k$  is computed for different coupling combinations and interpolated as a function of them to get the associated  $A_i^j$  coefficients. This procedure is further described in Ref. [141].

$$R_{HH}^k = \frac{\sigma_{HH}^k}{\sigma_{HH}^{k,SM}} \stackrel{LO}{=} A_1^k \kappa_t^4 + A_2^k c_2^2 + \left( A_3^k \kappa_t^2 + A_4^k c_g^2 \right) \kappa_\lambda^2 + A_5^k c_2^2 + \left( A_6^k c_2 + A_7^k \kappa_\lambda \kappa_t \right) \kappa_t^2 \\ + \left( A_8^k \kappa_t \kappa_\lambda + A_9^k c_g \kappa_\lambda \right) c_2 + A_{10}^k c_2 c_{2g} + \left( A_{11}^k c_g \kappa_\lambda + A_{12}^k c_{2g} \right) \kappa_t^2 \\ + \left( A_{13}^k \kappa_\lambda c_g + A_{14}^k c_{2g} \right) \kappa_t \kappa_\lambda + A_{15}^k c_g c_{2g} \kappa_\lambda \quad (E-1)$$

The event weight ( $\omega$ ) used to model a generic BSM combination of Higgs boson couplings is fully determined from this parametrization and two-dimensional histograms. A value  $\Omega$  is defined in equation E-2, where it has been indicated explicitly the dependence on the five couplings of the ratio of the total ( $R_{HH}$ ) and differential ( $R_{HH}^j$ ) cross section to the SM prediction, completely determined from the  $A_i^j$  coefficients. Then, the event weight  $\omega$  is defined from  $\Omega$  by normalizing it to the sum over all the  $n$  simulated MC signal events considered (i.e.  $\omega = \frac{\Omega}{\sum_n \Omega}$ ). Therefore, the application of  $\omega$  only modifies the differential event distribution but not the total normalization, and correctly accounts for changes in the acceptance after the selections.

$$\Omega(\kappa_\lambda, \kappa_t, c_2, c_g, c_{2g}; k) \equiv \frac{1}{f_{\text{comb}}^k} \cdot \frac{\sigma_{HH}^k(\kappa_\lambda, \kappa_t, c_2, c_g, c_{2g})}{\sigma_{HH}(\kappa_\lambda, \kappa_t, c_2, c_g, c_{2g})} \\ = \frac{f_{\text{SM}}^k}{f_{\text{comb}}^k} \cdot \frac{R_{HH}^k(\kappa_\lambda, \kappa_t, c_2, c_g, c_{2g})}{R_{HH}(\kappa_\lambda, \kappa_t, c_2, c_g, c_{2g})} \quad (E-2)$$

**NLO modeling:** The LO reweighting method cannot be applied for the case of NLO simulation, as the presence of an additional parton emitted at matrix level requires many more parameters to fully and correctly model the phase space. Consequently, a procedure based on the sum of different samples is used and restricted to the modeling of the  $\kappa_\lambda$  and  $\kappa_t$  couplings only.

The amplitude of the  $gg \rightarrow HH$  process can be generically expressed as  $\mathcal{A} = \kappa_\lambda \kappa_t T + \kappa_t^2 B$ . At the LO of the QCD expansion,  $T$  and  $B$  can be directly associated to the triangle and box diagrams of Fig. 1-6 (a). This relation is also valid at any order of the QCD perturbative expansion because  $T$  and  $B$  will be the sum of all the diagrams with the same order in  $\kappa_\lambda$  and  $\kappa_t$ . Therefore, the ggF cross section can be expressed by the sum of three terms in equation E-3,

where  $t = |T|^2$ ,  $b = |B|^2$ , and  $i = |TB^* + B^*T|$ . The same relation is also valid for any differential cross section ( $d\sigma/dx$ ).

$$\sigma(\kappa_\lambda, \kappa_t) \propto |\mathcal{A}|^2 = \kappa_\lambda^2 \kappa_t^2 t + \kappa_t^4 b + \kappa_\lambda \kappa_t^3 i \quad (\text{E-3})$$

Note that equation E-3 can be written as equation E-4, where  $\mathbf{c}(\kappa_\lambda, \kappa_t) = (\kappa_\lambda^2 \kappa_t^2, \kappa_t^4, \kappa_\lambda \kappa_t^3)$  is the vector of the coupling functions and  $\mathbf{v} = (t, b, i)$  is the vector of the values of the three terms. Therefore, for an arbitrary set of  $(\kappa_\lambda, \kappa_t)$  values, the HH signal can be described as the linear sum of three components. Moreover, the actual shape dependence is just on the ratio  $\kappa_\lambda/\kappa_t$ , as  $\mathbf{c}$  can be redefined as  $\mathbf{c}(\kappa_\lambda, \kappa_t) \rightarrow \kappa_t^4 \mathbf{c}(\kappa_\lambda/\kappa_t) = \kappa_t^4 ((\kappa_\lambda/\kappa_t)^2, 1, \kappa_\lambda/\kappa_t)$ . For simplicity, the  $\kappa_t = 1$  case is considered in the following derivation.

$$\sigma(\kappa_\lambda, \kappa_t) = \mathbf{c}(\kappa_\lambda, \kappa_t) \cdot \mathbf{v} \quad (\text{E-4})$$

The NLO MC POWHEG generator cannot individually generate events corresponding to the three terms. Furthermore, the direct sum of  $t$ ,  $b$ , and  $i$  is not optimal to model  $\kappa_\lambda$  values around 2.5 because of the large cancellations due to the interference, which would require producing samples with large statistics. Equation (E-3) also can be expressed in terms of a linear combination of three produced samples ( $s_1, s_2, s_3$ ) with a specific choice of  $\kappa_\lambda$ . Applying equation (E-3) to all the three samples, and denoting the 3 elements of the coupling vector for sample  $s_k$  as  $(c_k^1, c_k^2, c_k^3)$ , the one obtains the equation E-5. The three  $\sigma_i$  values are known from the POWHEG generator.

$$\begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{pmatrix} = \begin{pmatrix} c_1^1 & c_1^2 & c_1^3 \\ c_2^1 & c_2^2 & c_2^3 \\ c_3^1 & c_3^2 & c_3^3 \end{pmatrix} \begin{pmatrix} t \\ b \\ i \end{pmatrix} \quad (\text{E-5})$$

A simplified form of equation E-5 is  $\boldsymbol{\sigma} = \mathbf{C}\mathbf{v}$ . Then, inverting the matrix  $\mathbf{C}$ , one obtains the relation  $\mathbf{v} = \mathbf{C}^{-1}\boldsymbol{\sigma}$ . By substituting it in equation (E-4), one gets the parametrization defined in equation E-6. This relation also holds for differential  $d\sigma/dx$  distributions.

$$\sigma(\kappa_\lambda) = \mathbf{c}^T(\kappa_\lambda) \mathbf{C}^{-1} \boldsymbol{\sigma} \quad (\text{E-6})$$

Using this method, a generic value of  $\kappa_\lambda$  can be modeled as the sum of three NLO generated samples, scaling each sample  $s_i$  by functions of  $\kappa_\lambda$  that are derived by equation E-6. The choice of the three  $\kappa_\lambda$  values can be optimized based on the statistical precision, as some multiplicative coefficients may be negative depending on the coupling choice. For selecting the optimal coupling choice, redundancy and cross-check of the method, four samples were generated ( $\kappa_\lambda = 0, 1, 2.45, 5$ ). The optimal coupling choice is  $\kappa_\lambda = 1, 2.45, 5$ .

## E.2 Vector Boson Fusion Production

A VBF event-based reweighting method is not feasible due to the high number of parameters needed to describe the HHjj system. Then, the modeling of the VBF signals follows the same idea as the ggF NLO component sum method presented above. In analogy to equation E-3, the cross section can be described as equation E-7 based on the Figure 1-6 (b) diagrams, where  $a$ ,  $b$  and  $c$  are the amplitude squared terms and  $i_{ab}$ ,  $i_{ac}$ , and  $i_{bc}$  are the interference terms.

$$\sigma(\kappa_\lambda, \kappa_V, \kappa_{2V}) = \begin{pmatrix} \kappa_V^2 \kappa_\lambda^2 \\ \kappa_V^4 \\ \kappa_{2V}^2 \\ \kappa_V^3 \kappa_\lambda \\ \kappa_V \kappa_\lambda \kappa_{2V} \\ \kappa_V^2 \kappa_{2V} \end{pmatrix} \cdot \begin{pmatrix} a \\ b \\ c \\ i_{ab} \\ i_{ac} \\ i_{bc} \end{pmatrix} \quad (\text{E-7})$$

In the VBF component sum model, six samples  $s_i$  are generated. Each sample has a cross section  $\sigma_i$  taken from the MadGraph LO generator. Following the same formal procedure as with the ggF model, the final modeling of the VBF cross section parametrization is derived as equation E-8, where  $\mathbf{C}^{-1}$  is a  $6 \times 6$  matrix. To optimize this model based on the statistical precision and validation, seven samples were generated with the following  $(\kappa_\lambda, \kappa_V, \kappa_{2V})$  couplings: (1,1,1), (1,1,0), (1,1,2), (2,1,1), (0,1,1), (1,0.5,1) and (1.1.5,1). The best choice of six samples drops the sample  $(\kappa_\lambda, \kappa_V, \kappa_{2V}) = (1.0.5, 1)$ .

$$\sigma(\kappa_\lambda, \kappa_V, \kappa_{2V}) = \mathbf{c}^T(\kappa_\lambda, \kappa_V, \kappa_{2V}) \mathbf{C}^{-1} \boldsymbol{\sigma} \quad (\text{E-8})$$

## APPENDIX F BACKGROUND MODELING

### F.1 BDT Reweighting Hyperparameters

Table F-1. Best BDT-reweigher hyperparameters in 2016 analysis control region.

| Hyperparameters   | ggF Cat. 1 | ggF Cat. 2 | VBF Cat. 1 |
|-------------------|------------|------------|------------|
| Number of trees   | 50         | 50         | 125        |
| Max. depth        | 2          | 2          | 3          |
| Learning rate     | 0.5        | 0.5        | 0.1        |
| Min. samples leaf | 5          | 20         | 20         |
| Subsample         | 0.6        | 0.6        | 0.6        |

Table F-2. Best BDT-reweigher hyperparameters in 2017-2018 analysis control region.

| Hyperparameters   | ggF Cat. 1 | ggF Cat. 2 | VBF Cat. 1 |
|-------------------|------------|------------|------------|
| Number of trees   | 50         | 150        | 50         |
| Max. depth        | 2          | 3          | 2          |
| Learning rate     | 0.5        | 0.1        | 0.5        |
| Min. samples leaf | 10         | 5          | 20         |
| Subsample         | 0.6        | 0.6        | 0.6        |

Table F-3. Best BDT-reweigher hyperparameters in 2016 validation test.

| Hyperparameters   | ggF Cat. 1 | ggF Cat. 2 | VBF Cat. 1 |
|-------------------|------------|------------|------------|
| Number of trees   | 150        | 75         | 50         |
| Max. depth        | 4          | 3          | 2          |
| Learning rate     | 0.1        | 0.5        | 0.05       |
| Min. samples leaf | 10         | 5          | 5          |
| Subsample         | 0.6        | 0.6        | 0.6        |

Table F-4. Best BDT-reweigher hyperparameters in 2017-2018 validation test.

| Hyperparameters   | ggF Cat. 1 | ggF Cat. 2 | VBF Cat. 1 |
|-------------------|------------|------------|------------|
| Number of trees   | 75         | 125        | 125        |
| Max. depth        | 2          | 3          | 2          |
| Learning rate     | 0.5        | 0.1        | 0.1        |
| Min. samples leaf | 10         | 1          | 10         |
| Subsample         | 0.6        | 0.6        | 0.6        |

## **F.2 Analysis Control Region Results**

This section presents the data and background model distributions with the BDT-reweighting applied in the analysis control region. Figures [F-1](#), [F-2](#) and [F-5](#) show the results in 2016 ggF category 1, 2016 ggF category 2, and 2016 VBF category 1, respectively. Figures [F-3](#), [F-4](#) and [F-6](#) show the results in 2016 ggF category 1, 2016 ggF category 2, and 2016 VBF category 1, respectively.

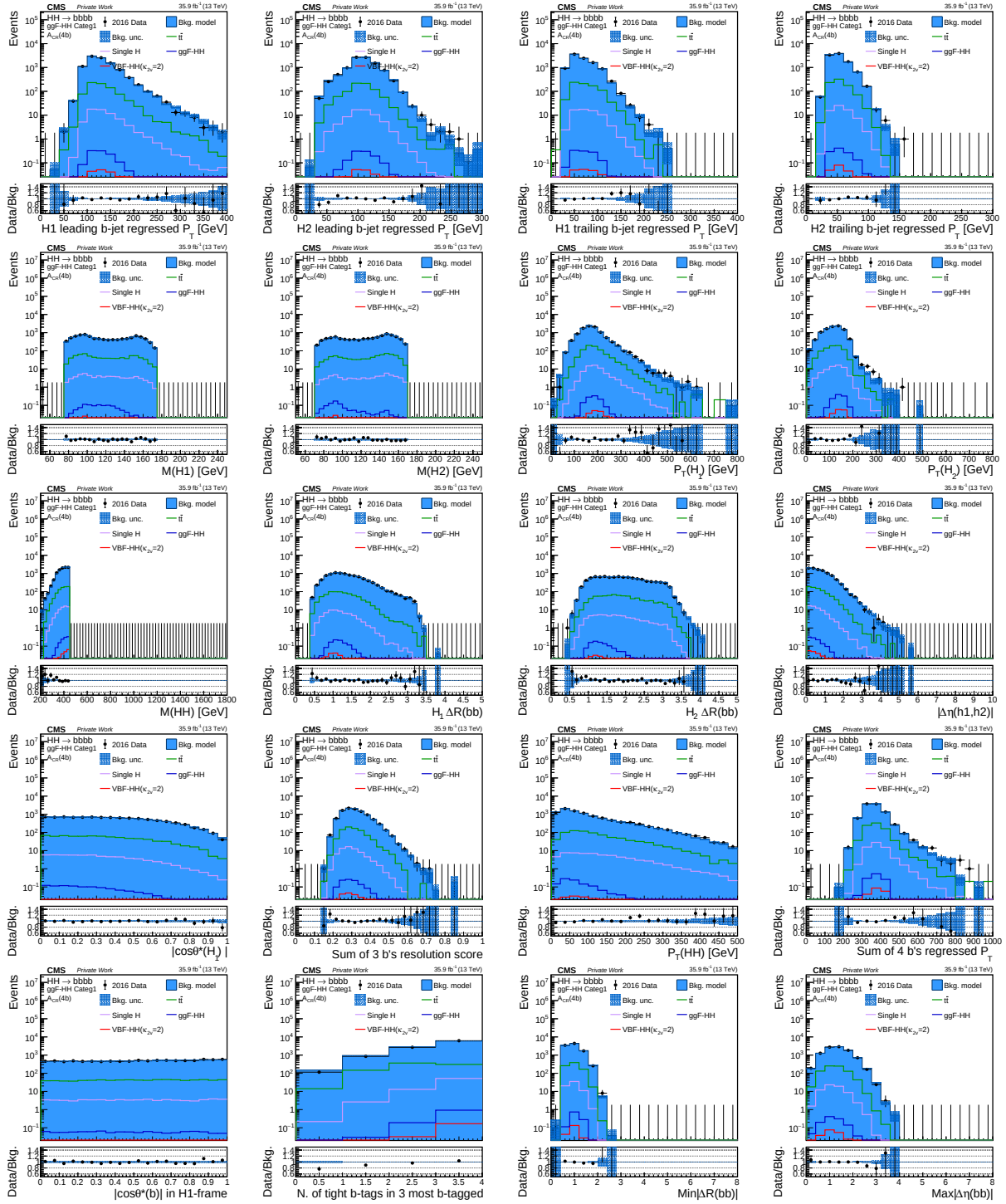


Figure F-1. Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 2016 ggF category 1.

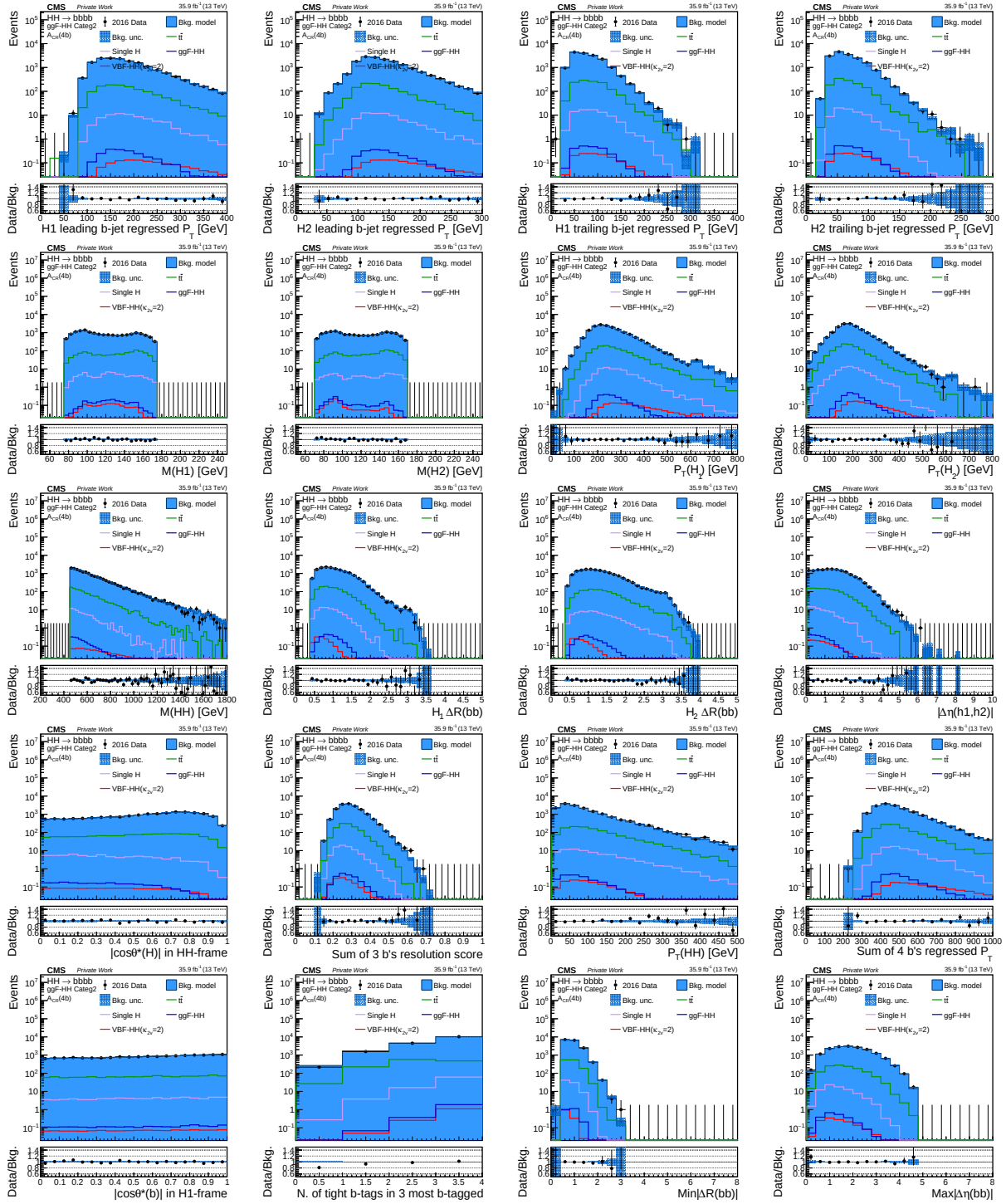


Figure F-2. Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 2016 ggF category 2.

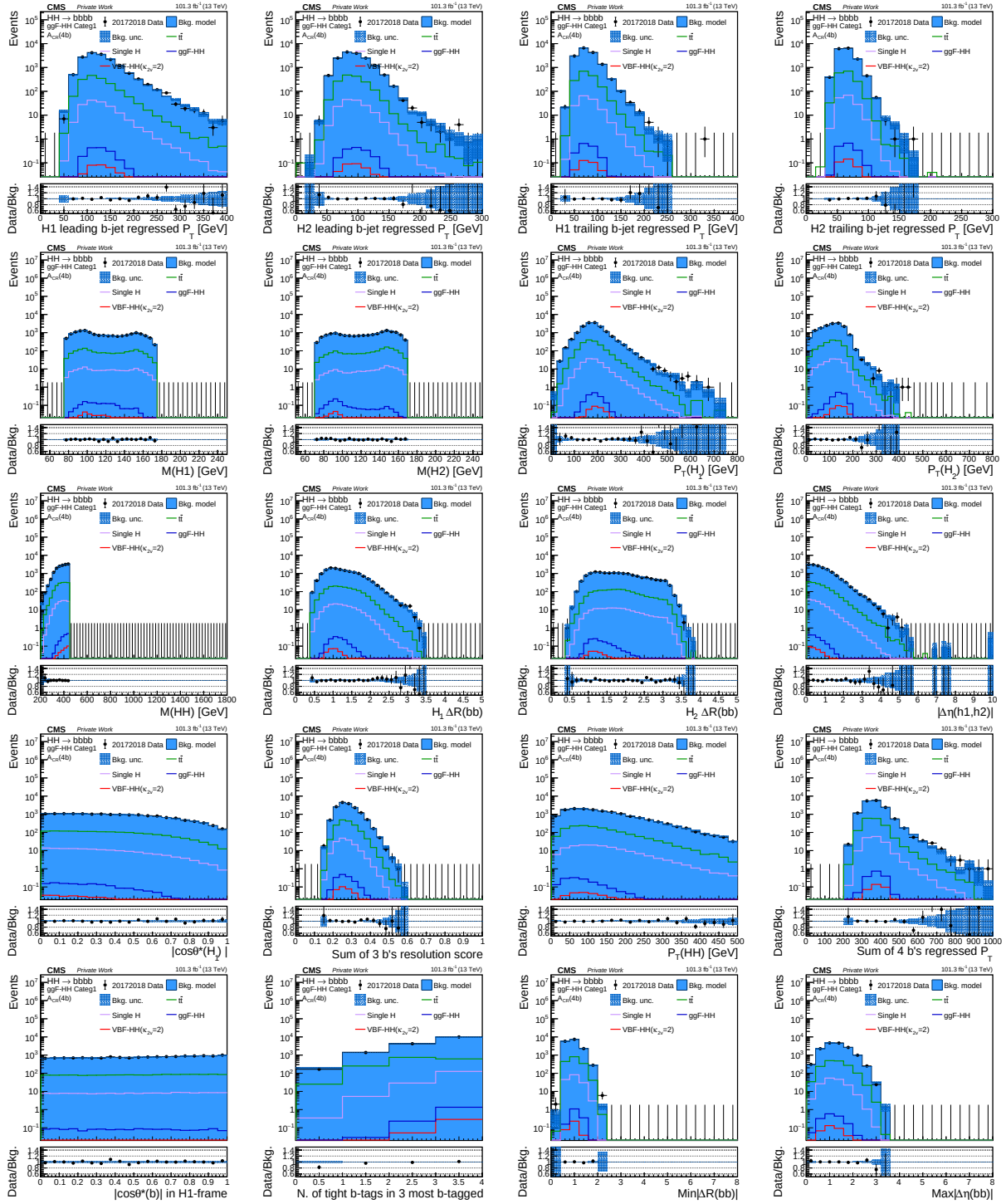


Figure F-3. Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 2017/2018 ggF category 1.



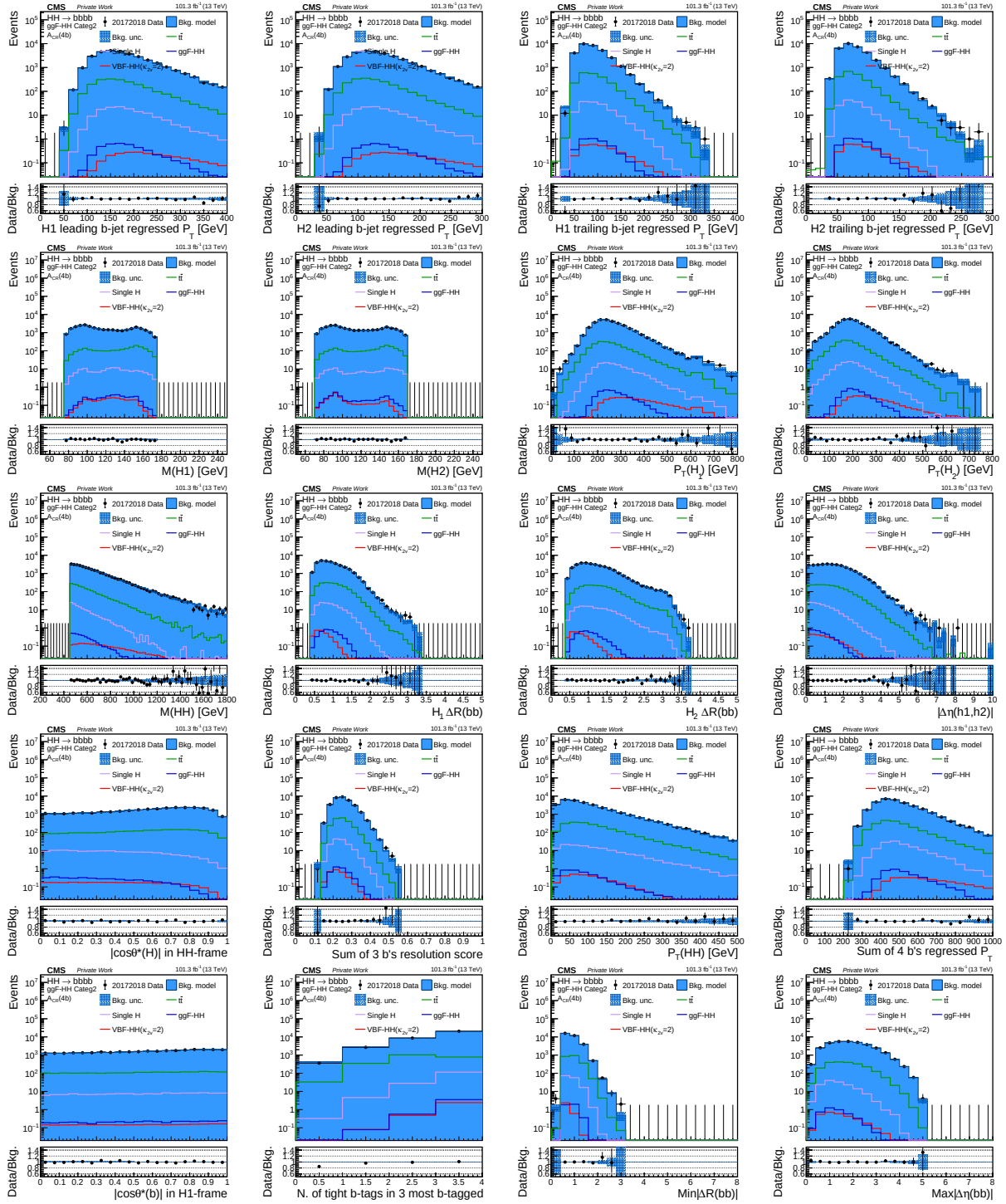


Figure F-4. Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 20172018 ggF category 2.

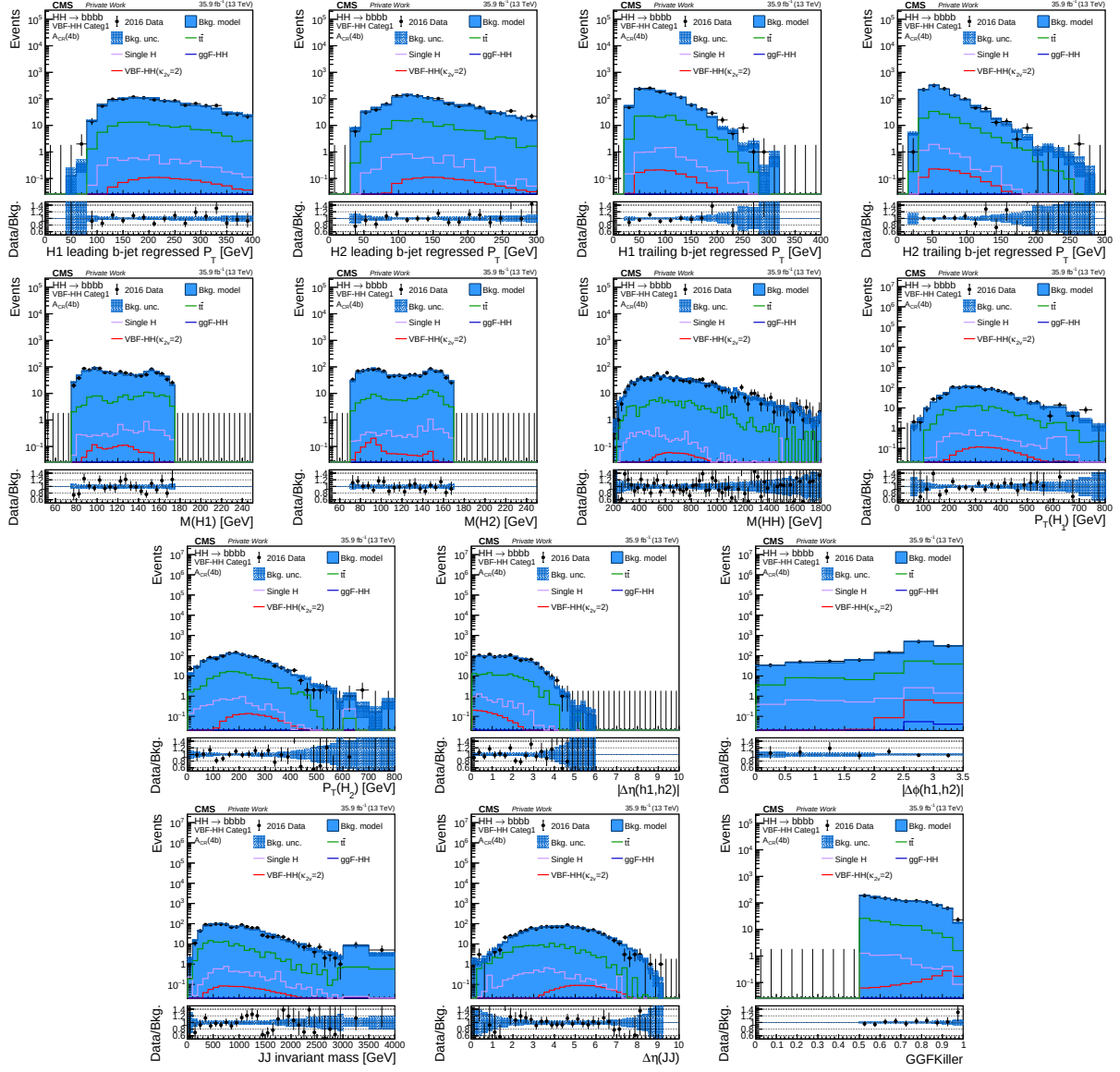


Figure F-5. Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 2016 VBF category 1.

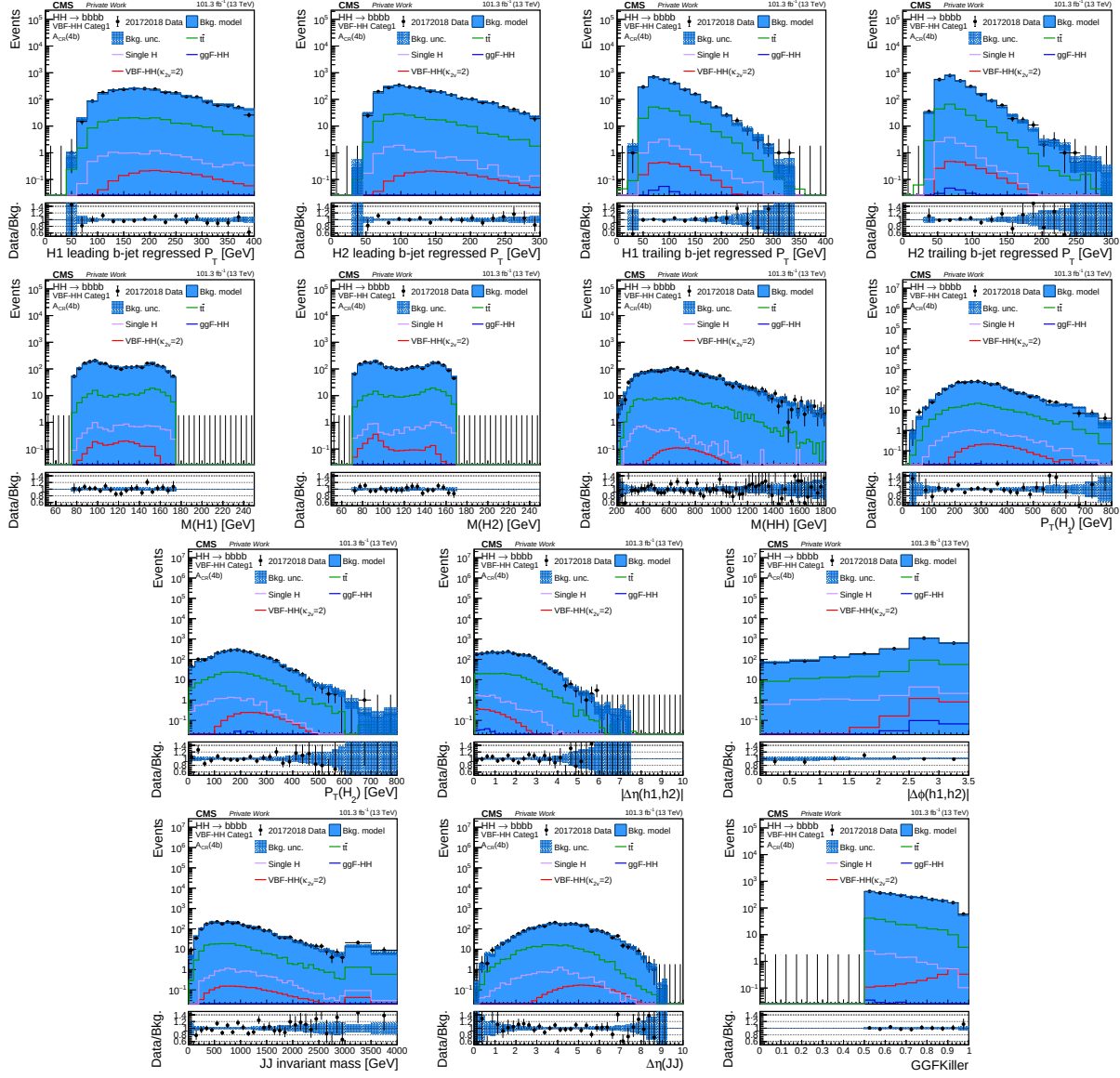


Figure F-6. Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 20172018 VBF category 1.

### F.3 Validation Control Region Results

This section presents the data and background model distributions with the BDT-reweighting applied in the validation control region. Figures F-7, F-8 and F-11 show the results in 2016 ggF category 1, 2016 ggF category 2, and 2016 VBF category 1, respectively. Figures F-9, F-10 and F-12 show the results in 2016 ggF category 1, 2016 ggF category 2, and 2016 VBF category 1, respectively.

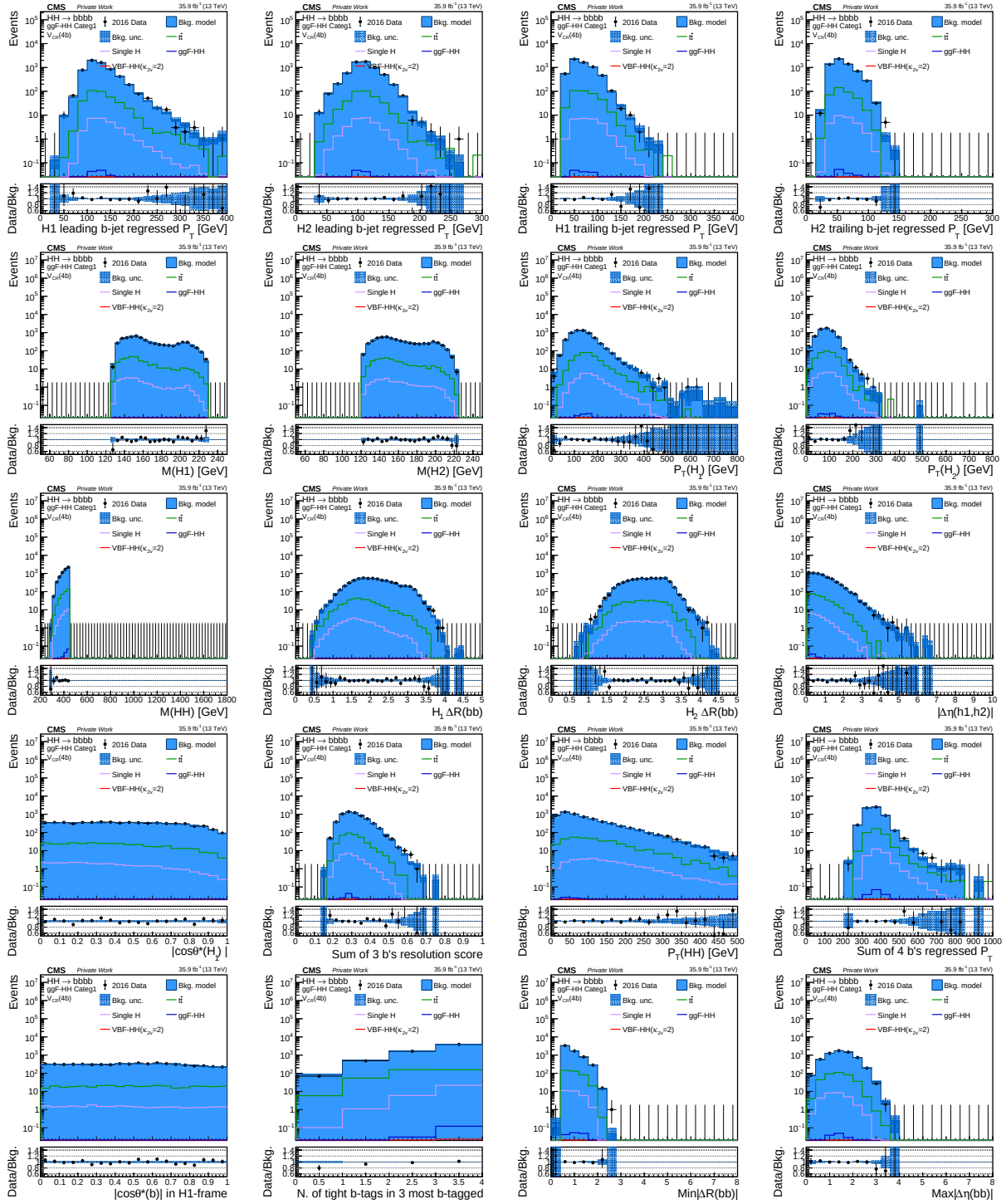


Figure F-7. Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 2016 ggF category 1 in the validation study.

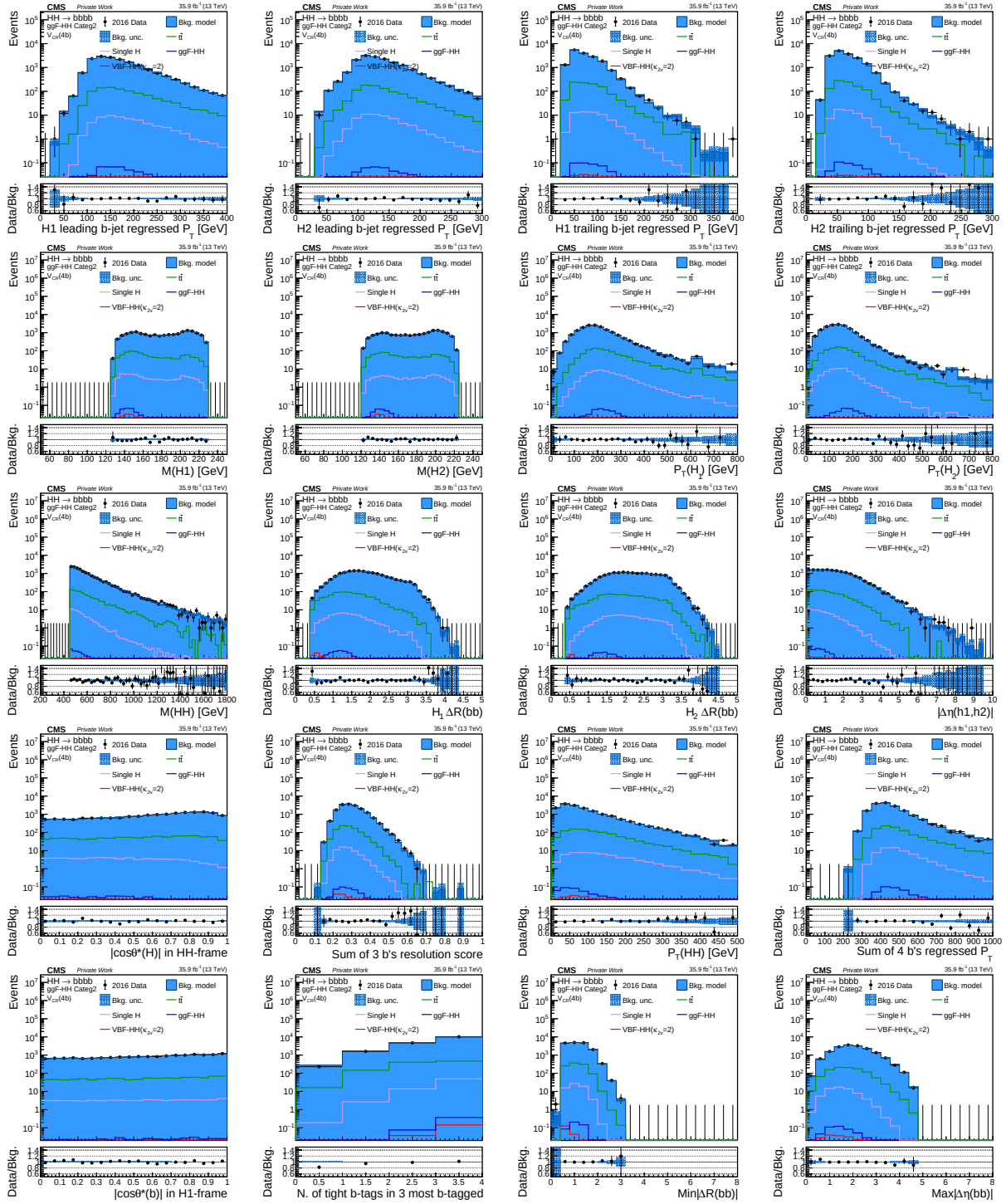


Figure F-8. Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 2016 ggF category 2 in the validation study.

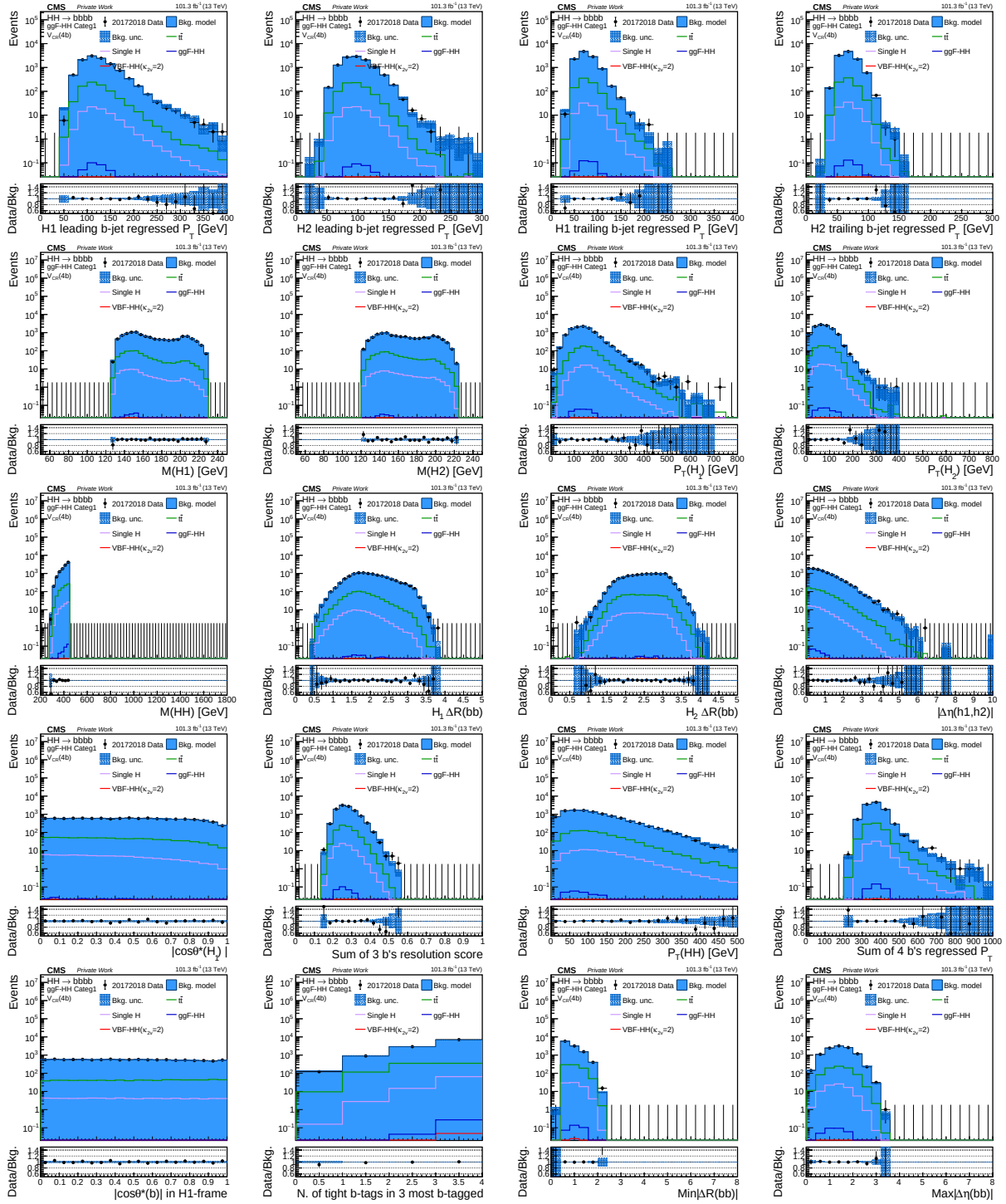


Figure F-9. Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 2017/2018 ggF category 1 in the validation study.

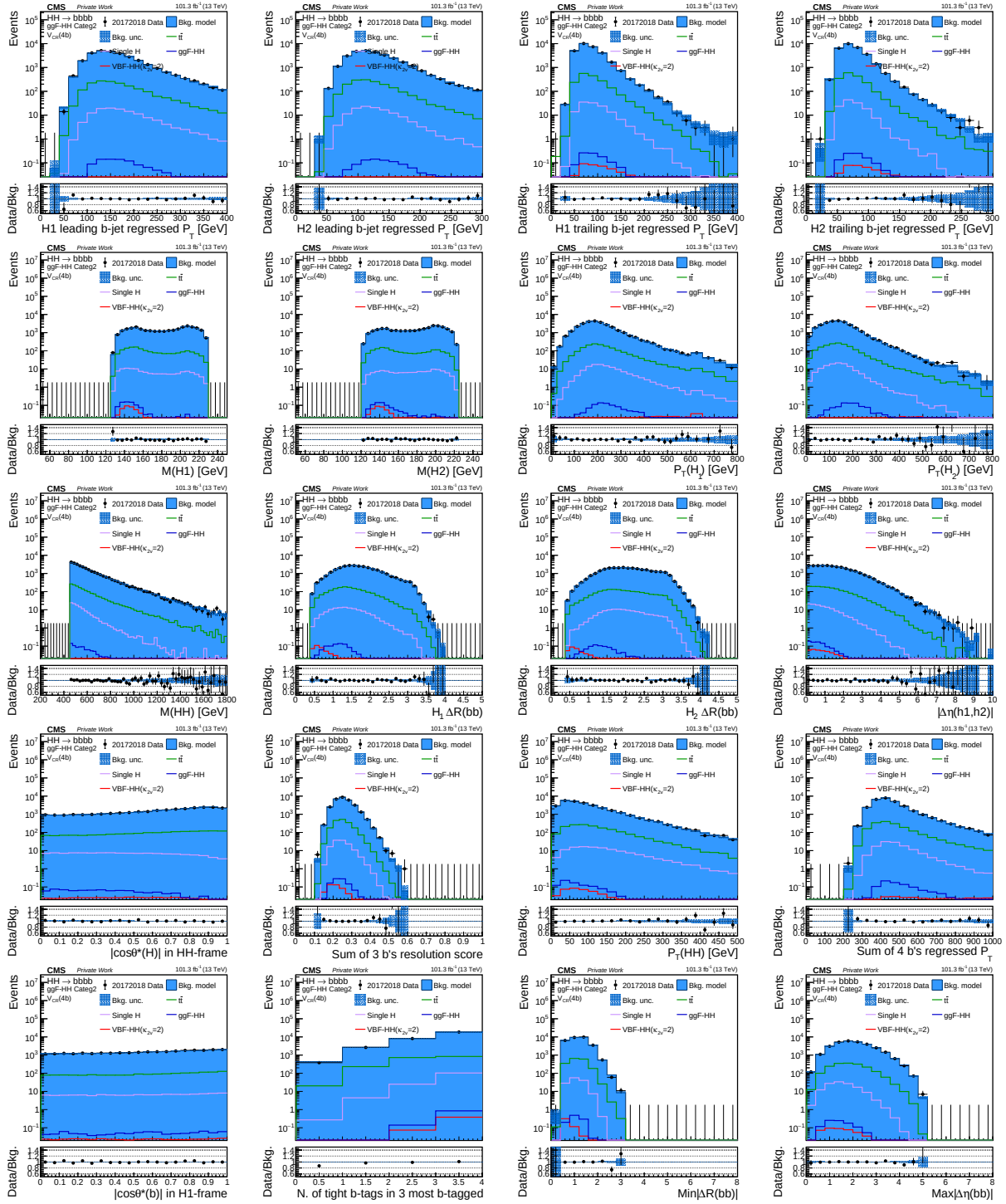


Figure F-10. Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 2017/2018 ggF category 2 in the validation study.



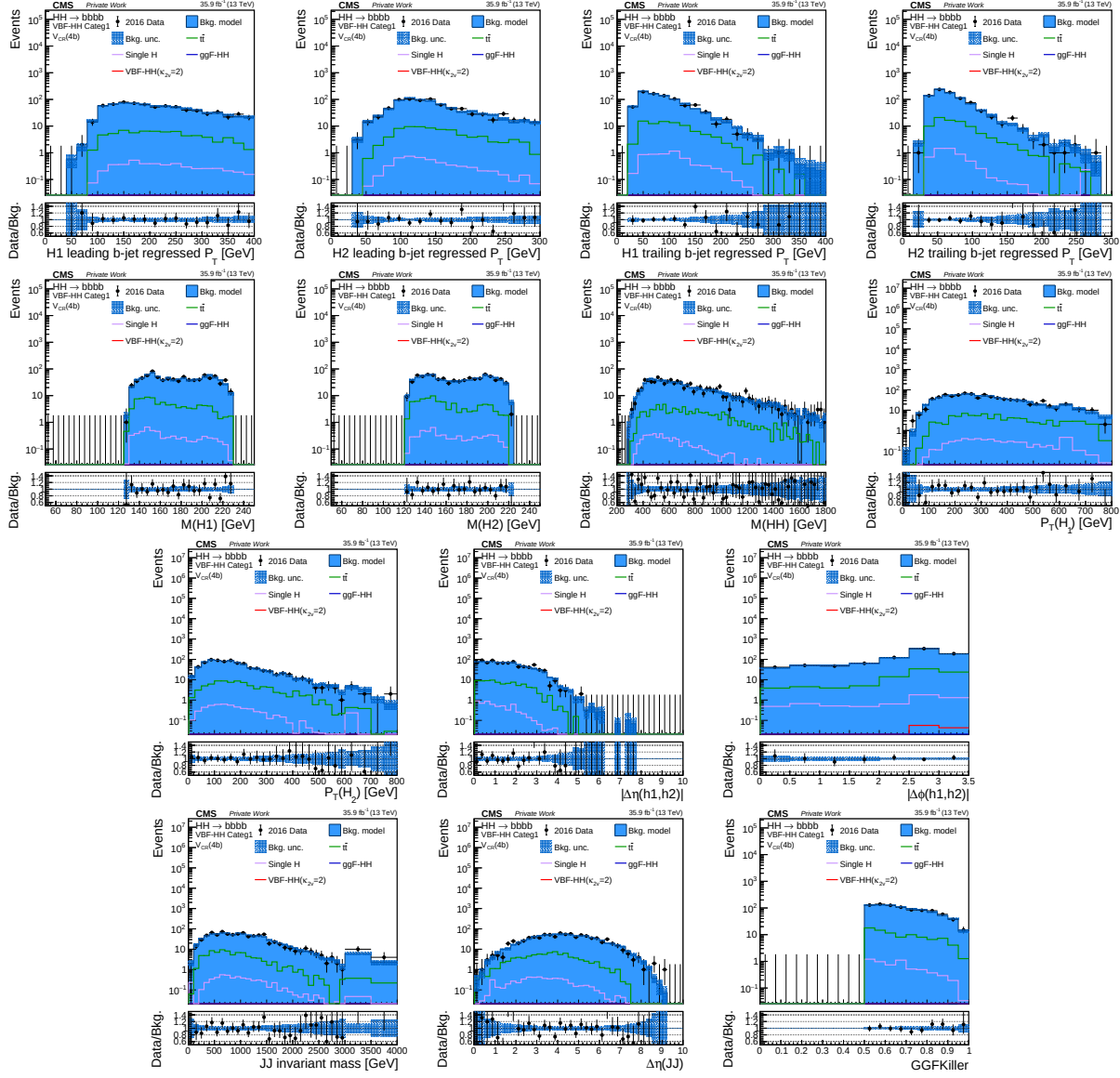


Figure F-11. Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 2016 VBF category 1 in the validation study.

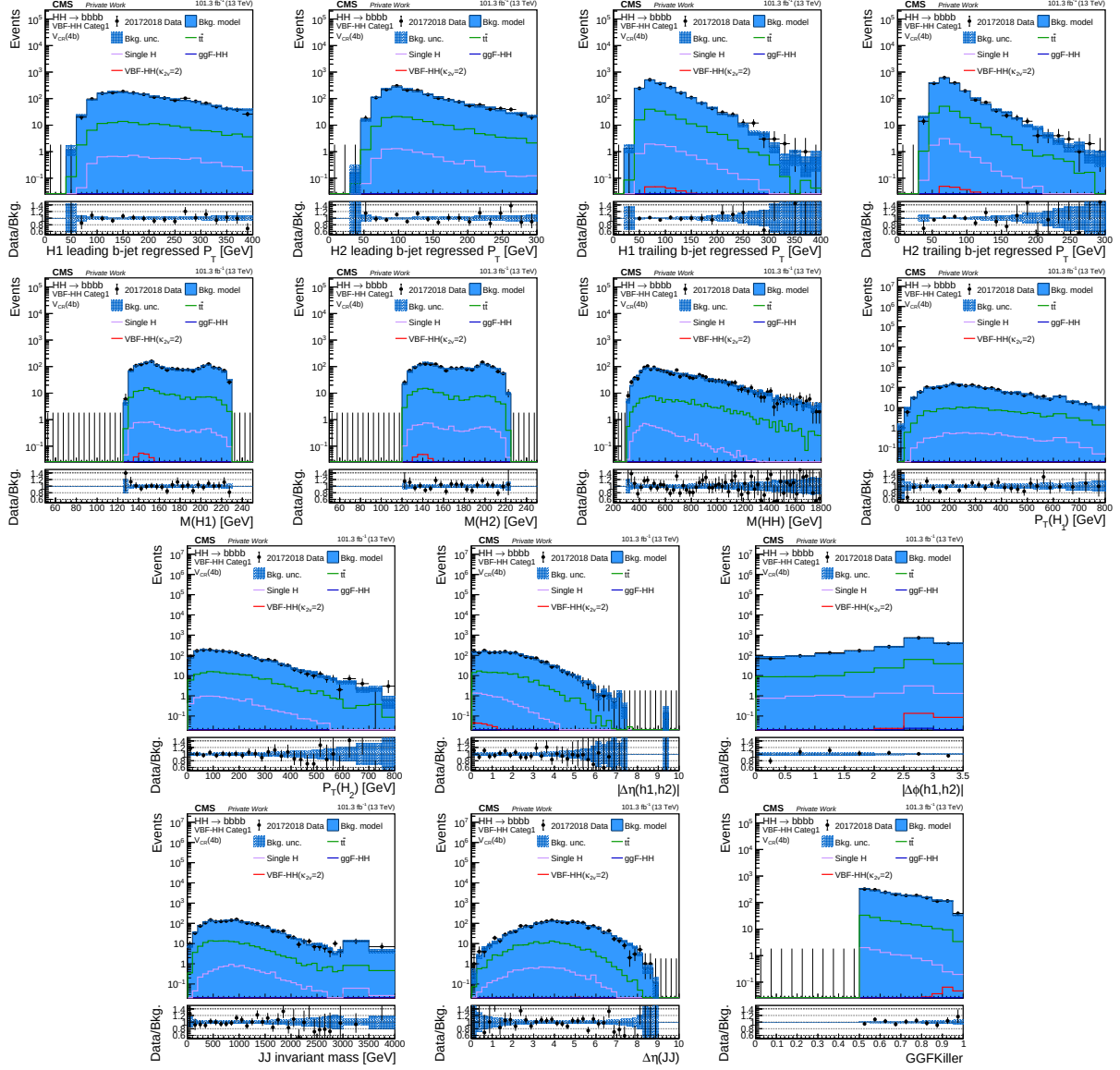


Figure F-12. Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 2017/2018 VBF category 1 in the validation study.

#### **F.4 Validation Signal Region Results**

This section presents the data and background model distributions with the BDT-reweighting applied in the validation signal region. Figures [F-13](#), [F-14](#) and [F-17](#) show the results in 2016 ggF category 1, 2016 ggF category 2, and 2016 VBF category 1, respectively. Figures [F-15](#), [F-16](#) and [F-18](#) show the results in 2016 ggF category 1, 2016 ggF category 2, and 2016 VBF category 1, respectively.

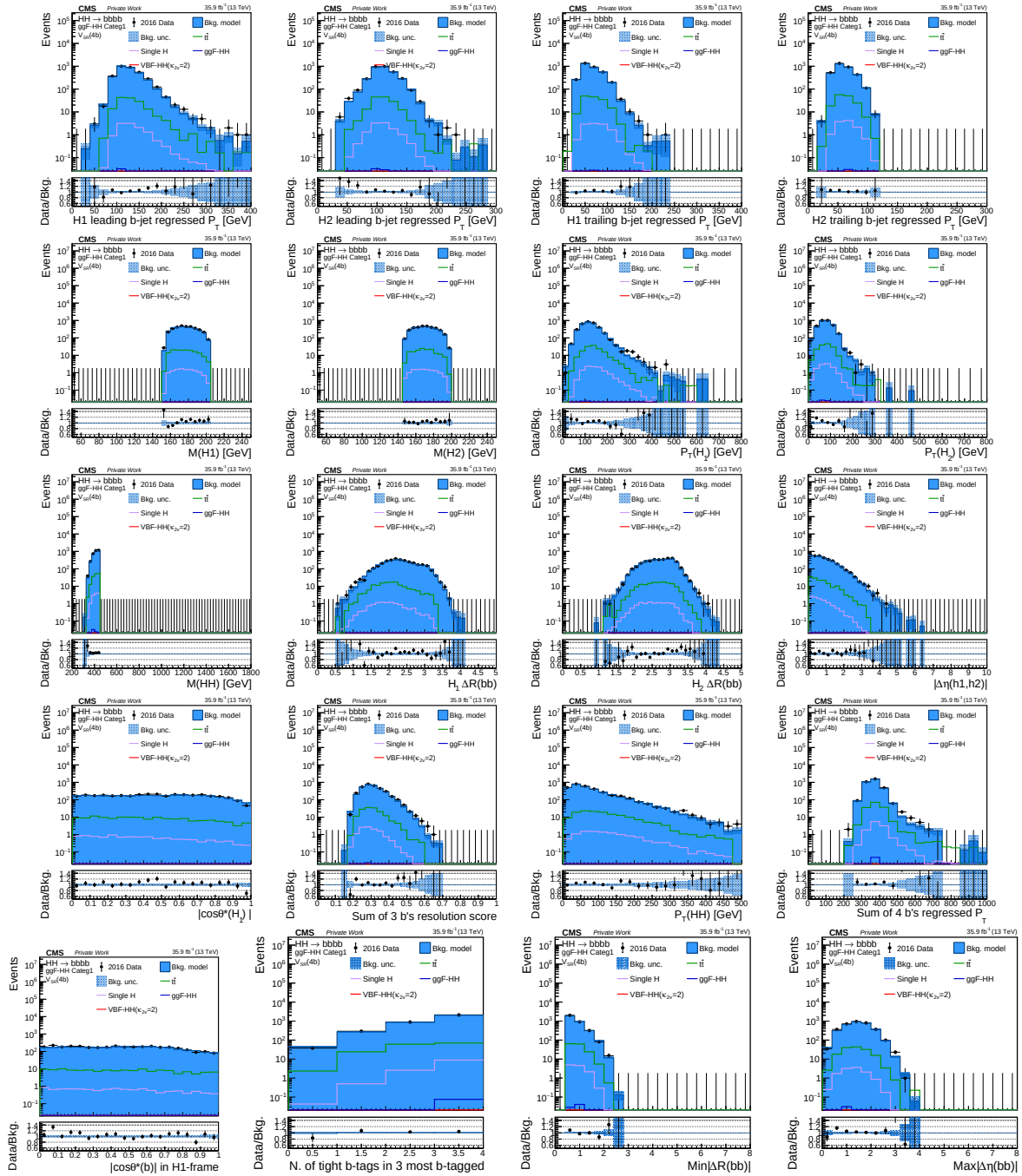


Figure F-13. Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 2016 ggF category 1 in the validation signal region study.

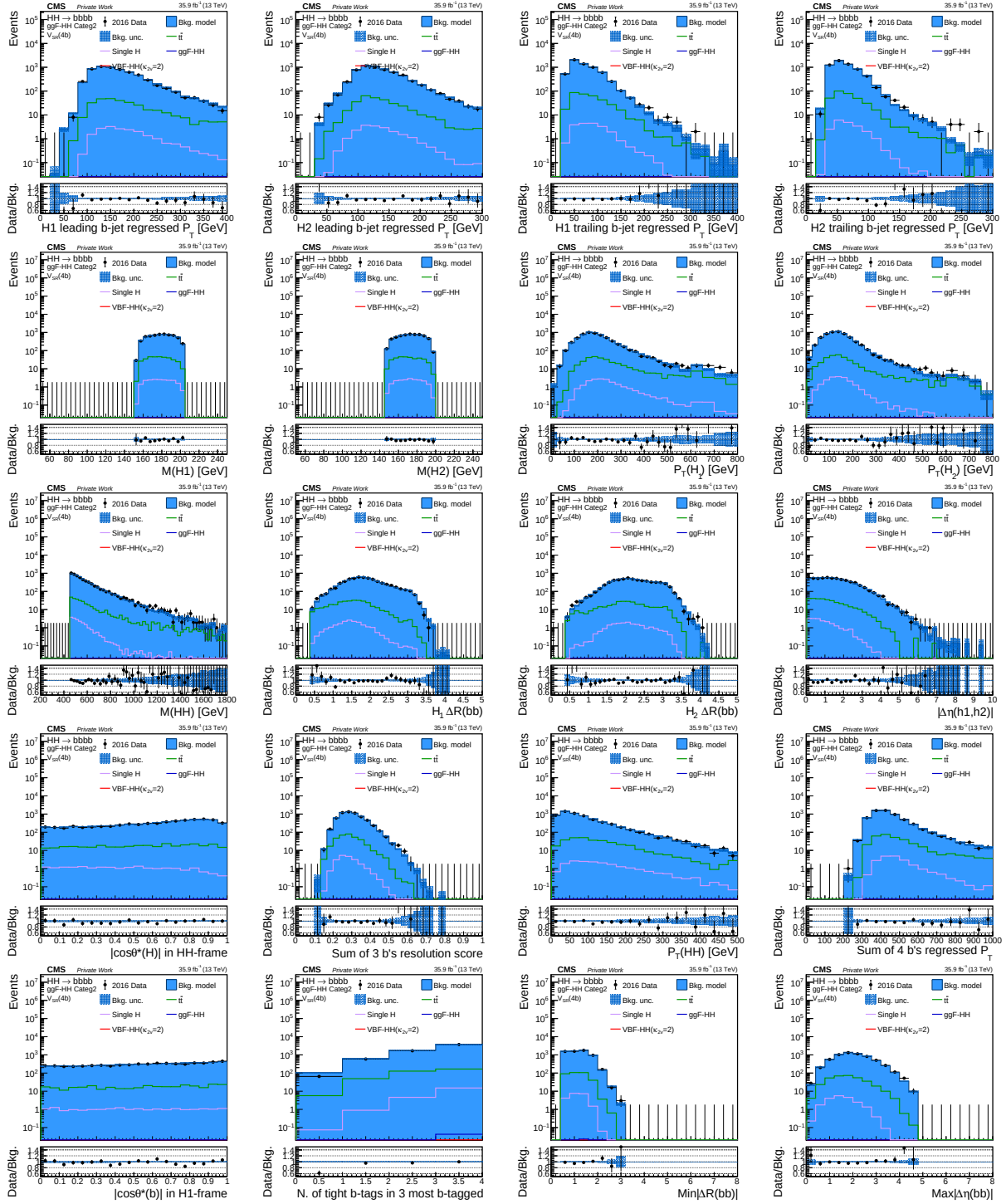


Figure F-14. Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 2016 ggF category 2 in the validation signal region study.

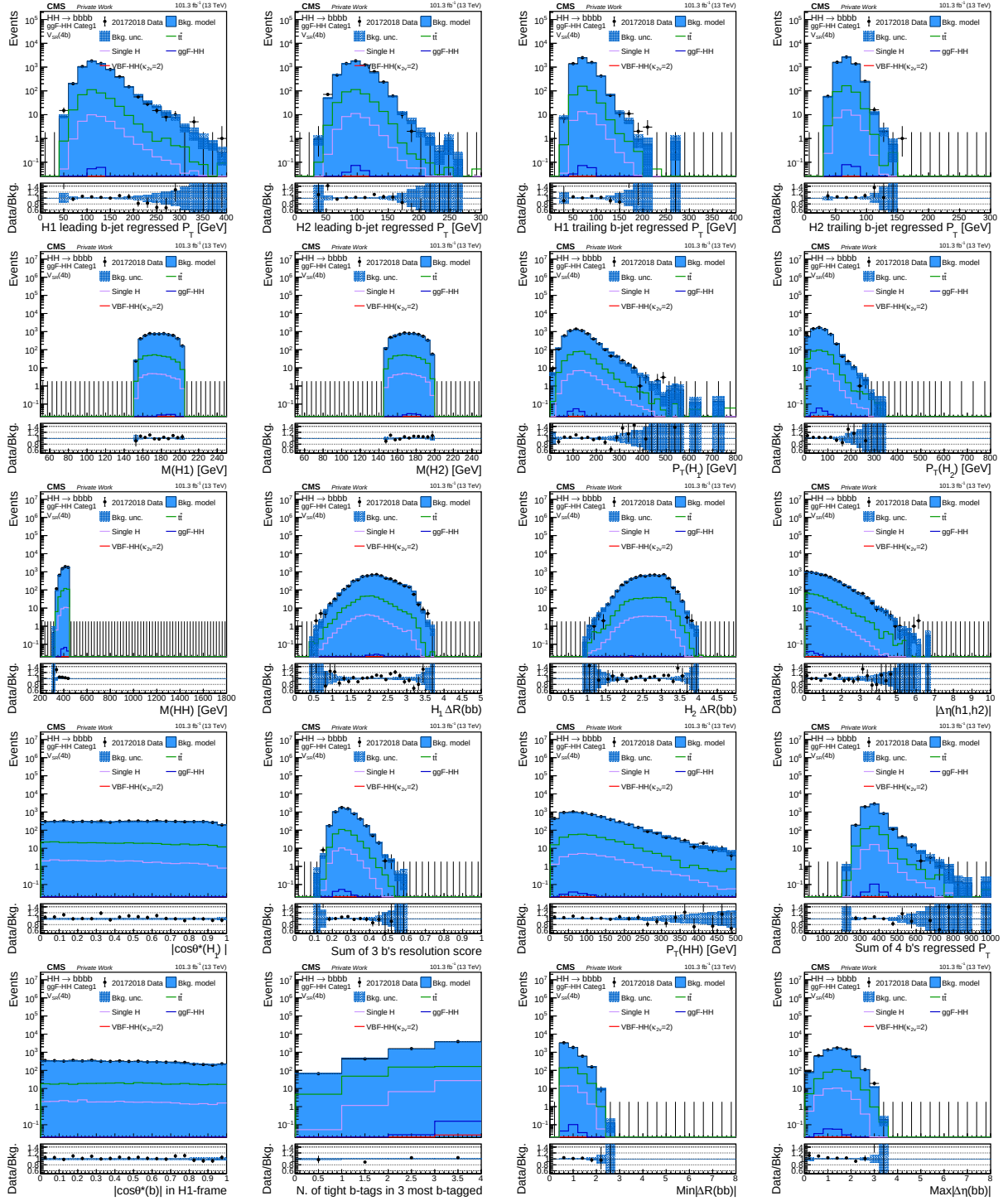


Figure F-15. Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 2017/2018 ggF category 1 in the validation signal region study.

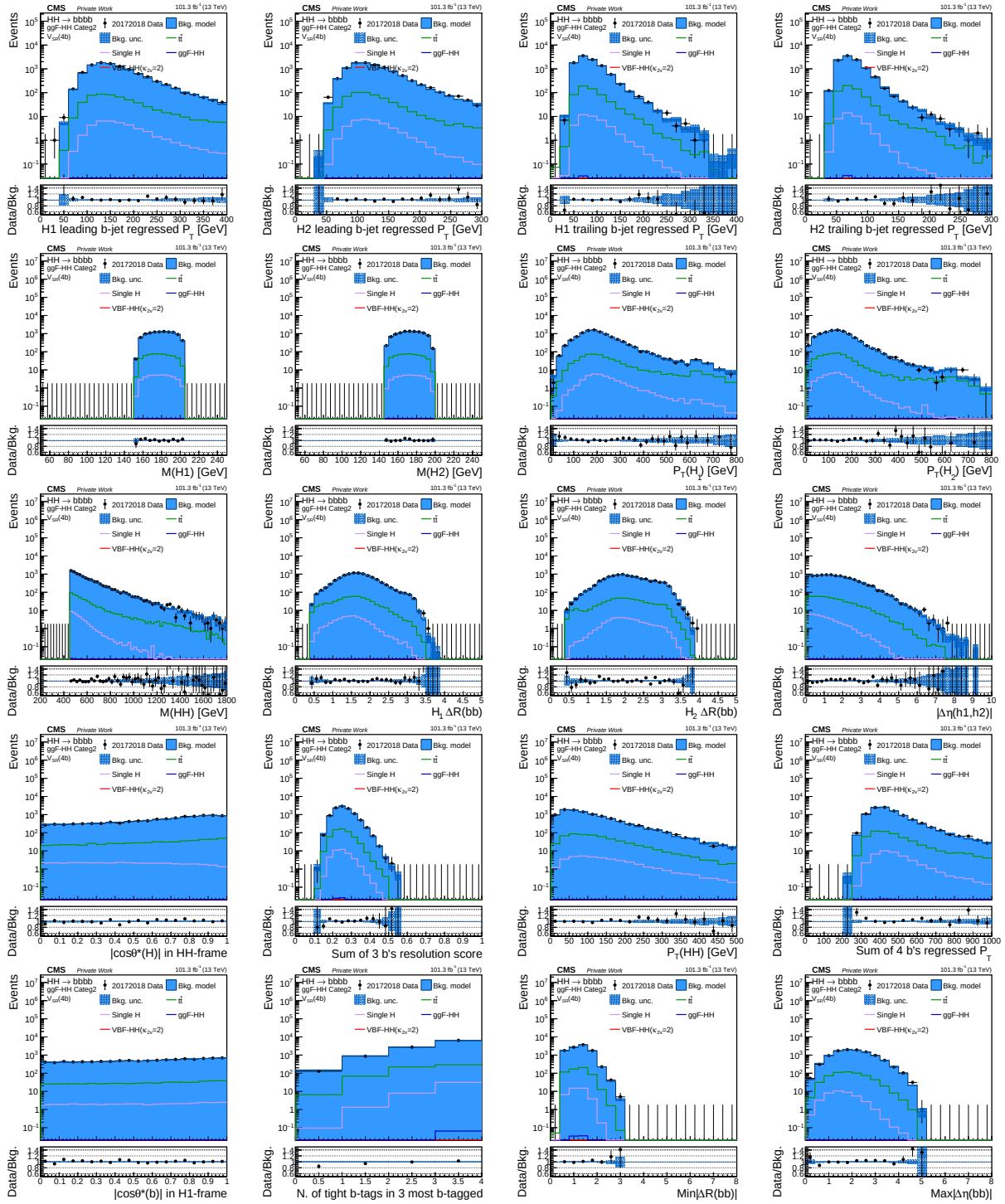


Figure F-16. Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 2017/2018 ggF category 2 in the validation signal region study.

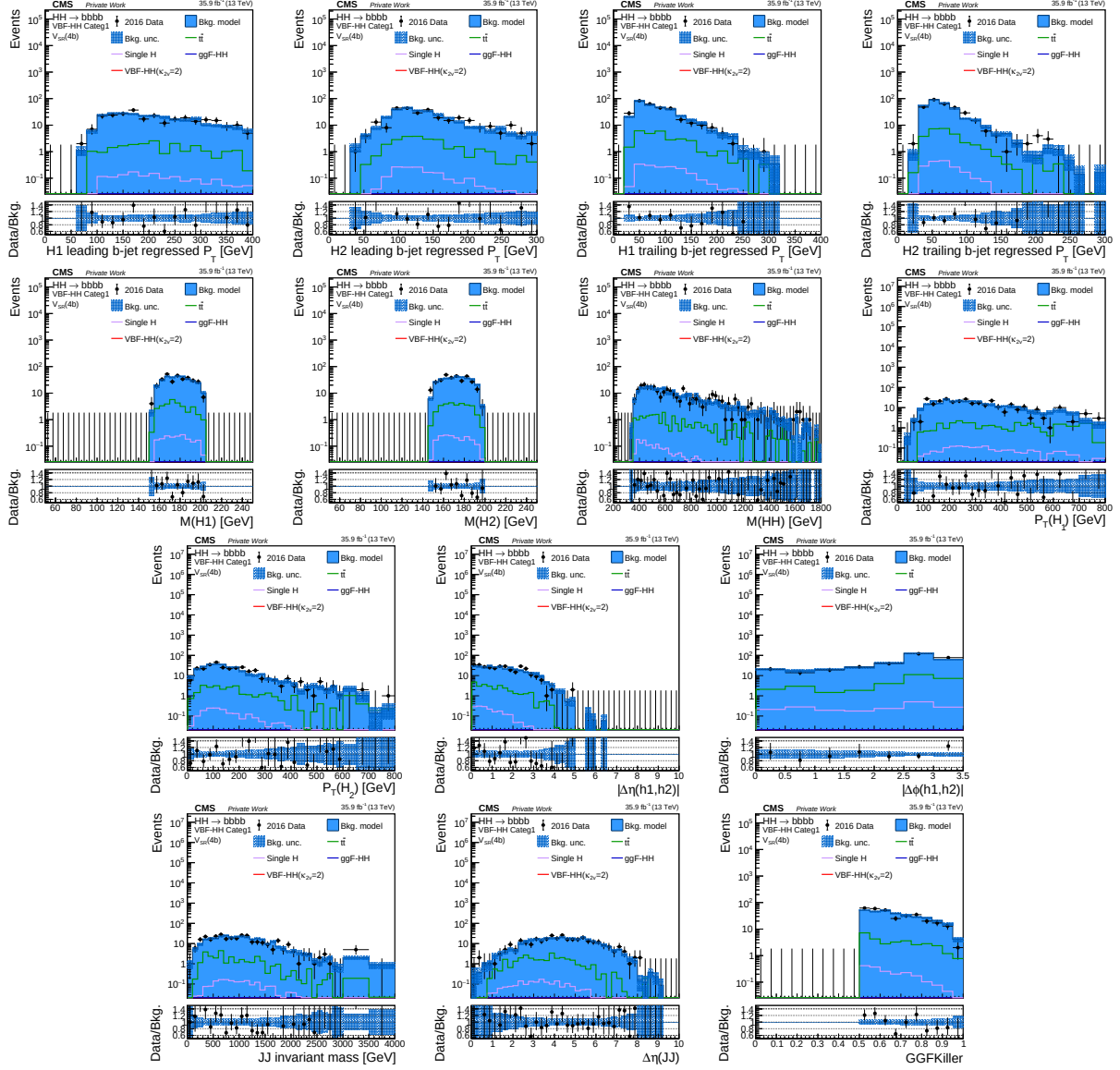


Figure F-17. Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 2016 VBF category 1 in the validation signal region study.



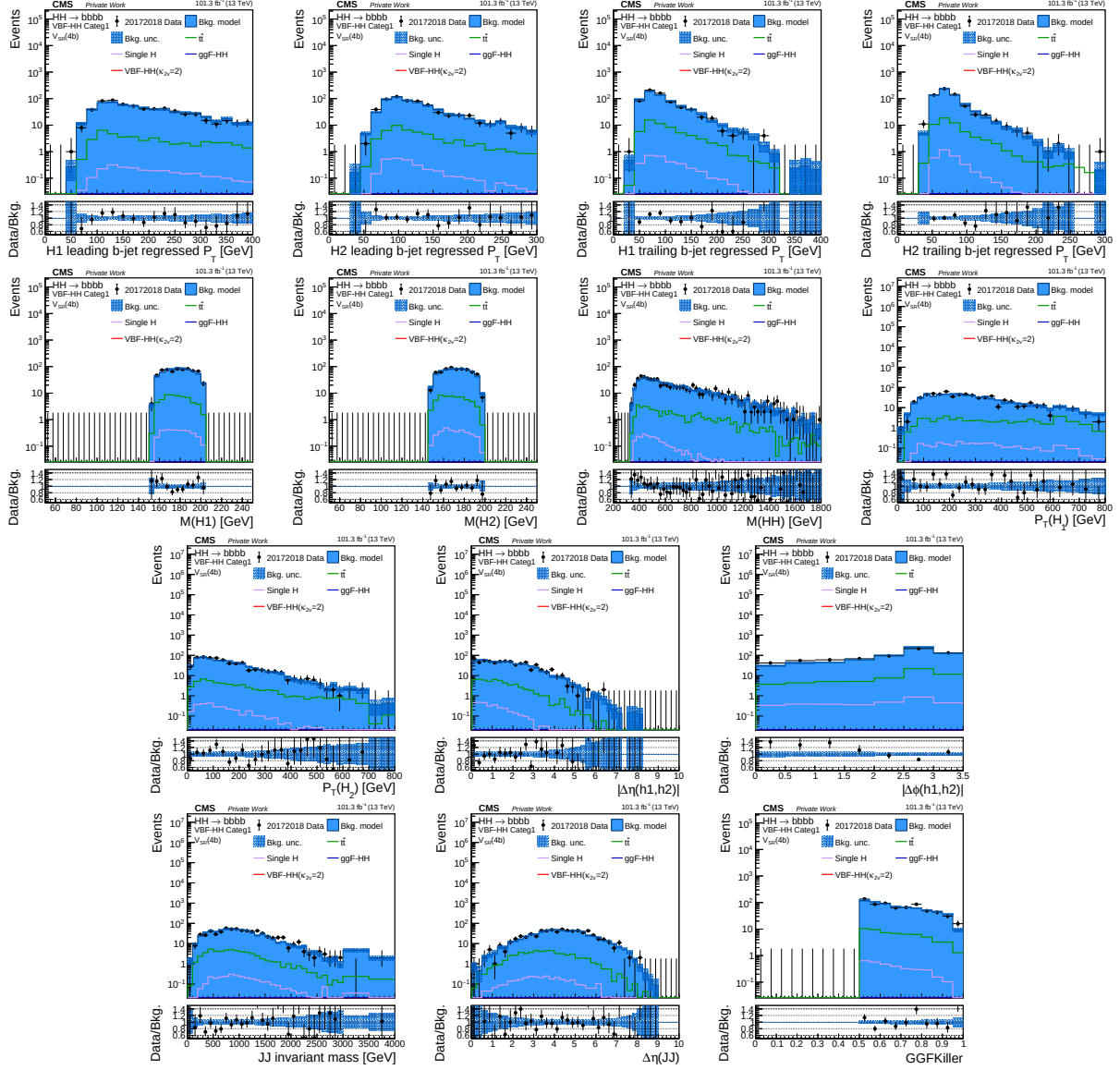


Figure F-18. Data versus background modeling agreement of the variables of the BDT-reweighter training (with BDT-reweighting) in 20172018 VBF category 1 in the validation signal region study.

## APPENDIX G SIGNAL OBSERVABLES

This appendix offers more information regarding the trained classifiers (C1 and C2) used as the ggF observables for the signal extraction. For completion, the distributions of the signal and background samples in 2016 are shown in Figures [G-1](#) and [G-2](#).

The resulting hyperparameters of the classifier C1 are presented in Table [G-1](#). The signal and background distributions of the C1 classifier output in the train and test samples are shown in Figure [G-3](#). The C1 ROC-AUC (train/test) values are presented in Table [G-2](#), and the ranking of variable importance is presented in Figure [G-5](#).

The signal and background distributions of the C2 classifier in the train and test samples are shown in Figure [G-4](#). The C2 ROC-AUC (train/test) values are presented in Table [G-2](#) and are compatible with the C1 classifier training.

Figure G-1. Input variables for the BDT output training in the 2016 ggF Category 1. The SM ggF (signal) is in solid red, whereas the data-driven background modeling is in solid blue.

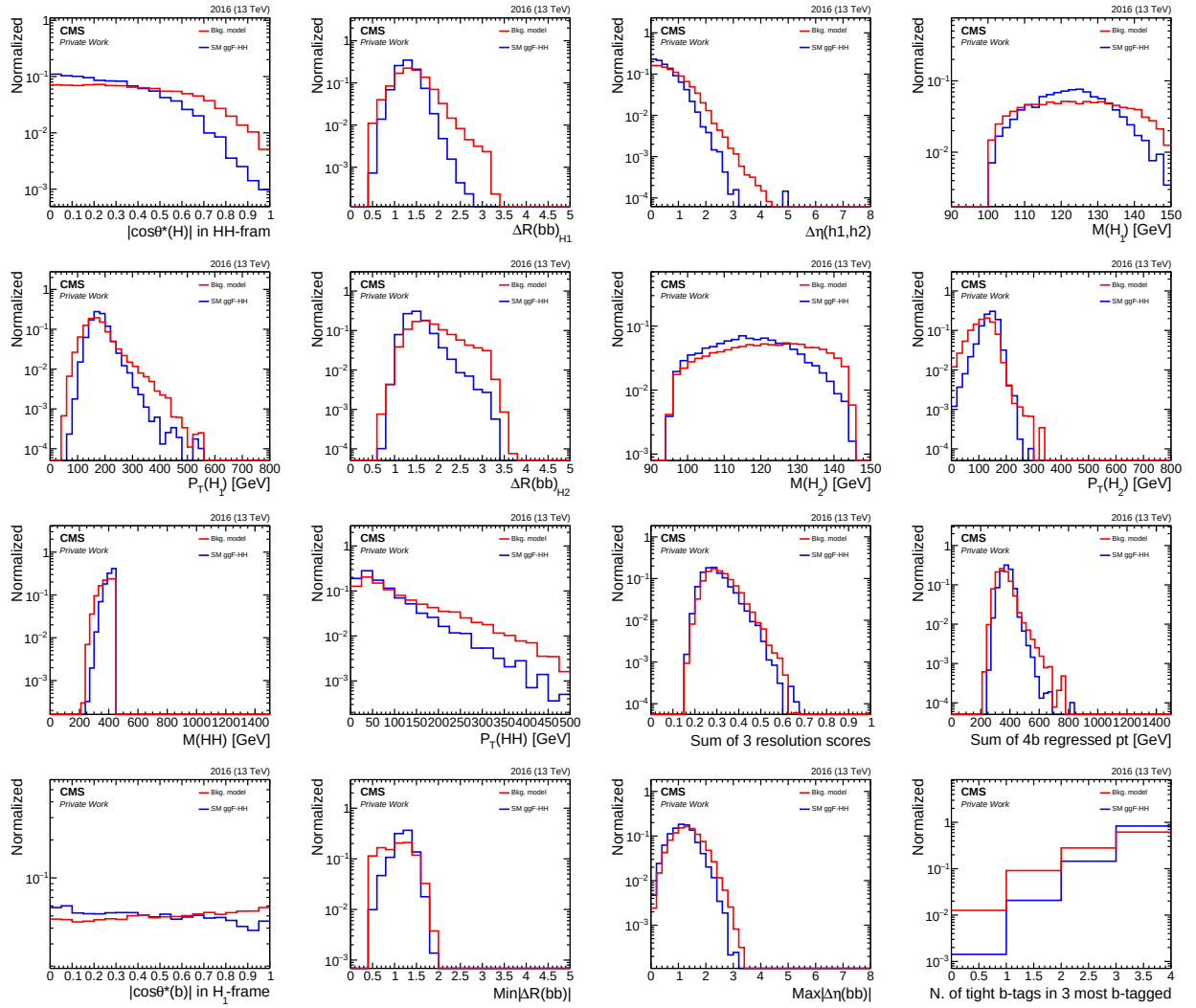


Figure G-2. Input variables for the BDT output training in the 2016 ggF Category 2. The SM ggF (signal) is in solid red, whereas the data-driven background modeling is in solid blue.

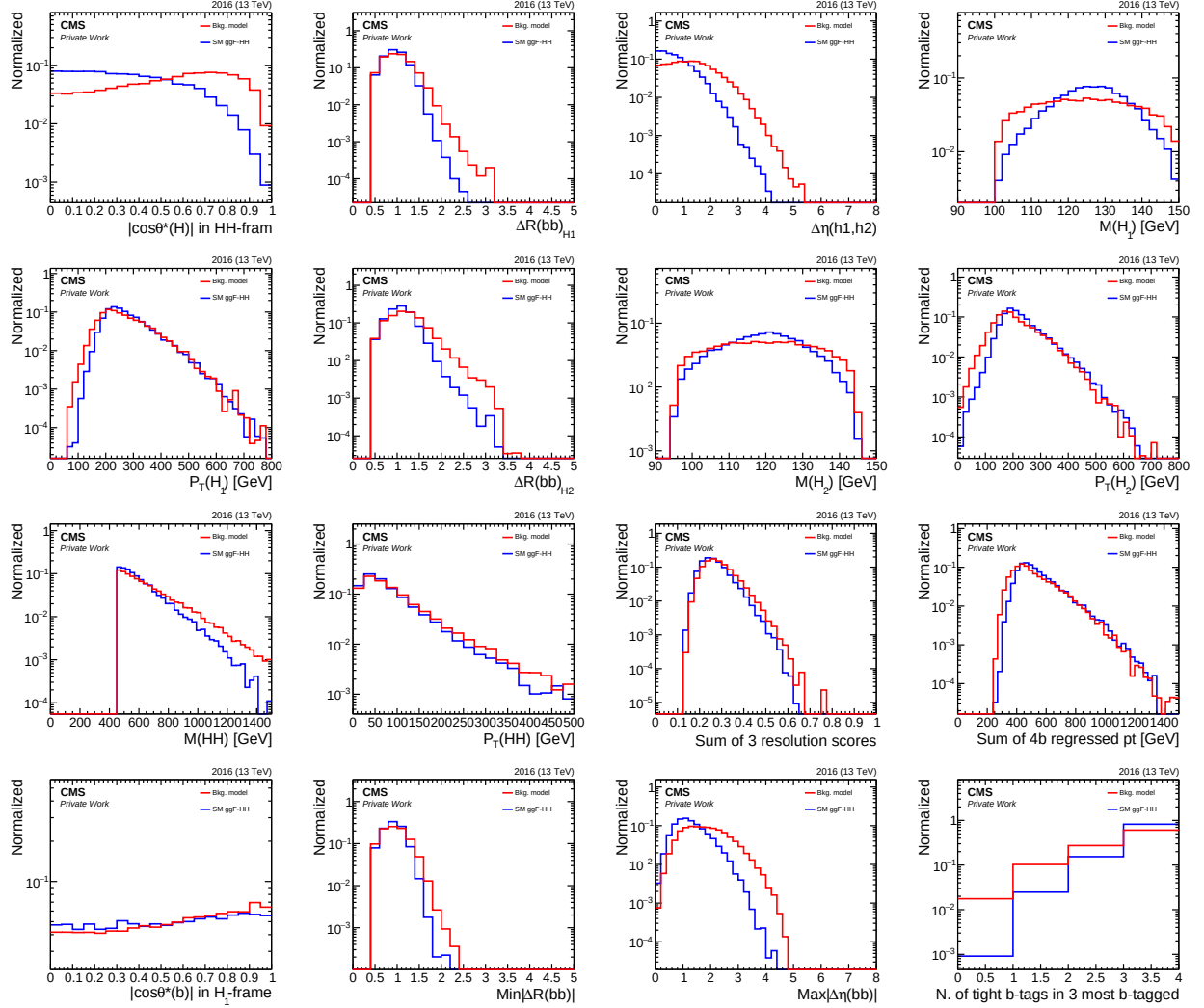


Table G-1. Best hyperparameters of the BDT output training in ggF categories..

| Hyperparameters | 2016 Cat. 1     | 20172018 Cat. 1 | 2016 Cat. 2     | 20172018 Cat. 2 |
|-----------------|-----------------|-----------------|-----------------|-----------------|
| BoostType       | Gradient Boost  | Gradient Boost  | Gradient Boost  | Gradient Boost  |
| Objective       | binary:logistic | binary:logistic | binary:logistic | binary:logistic |
| Number of trees | 400             | 300             | 300             | 200             |
| Max. depth      | 2               | 2               | 2               | 3               |
| Learning rate   | 0.1             | 0.1             | 0.1             | 0.1             |

Figure G-3. Training performance of the classifier C1 used for building the ggF observables. A) 2016 Category 1, B) 2016 Category 2, C) 2017-2018 Category 1, D) 2017-2018 Category 2. The left figure shows the trained classifier C1 distribution in train/test samples. The right figure shows the ROC curve of the test samples.

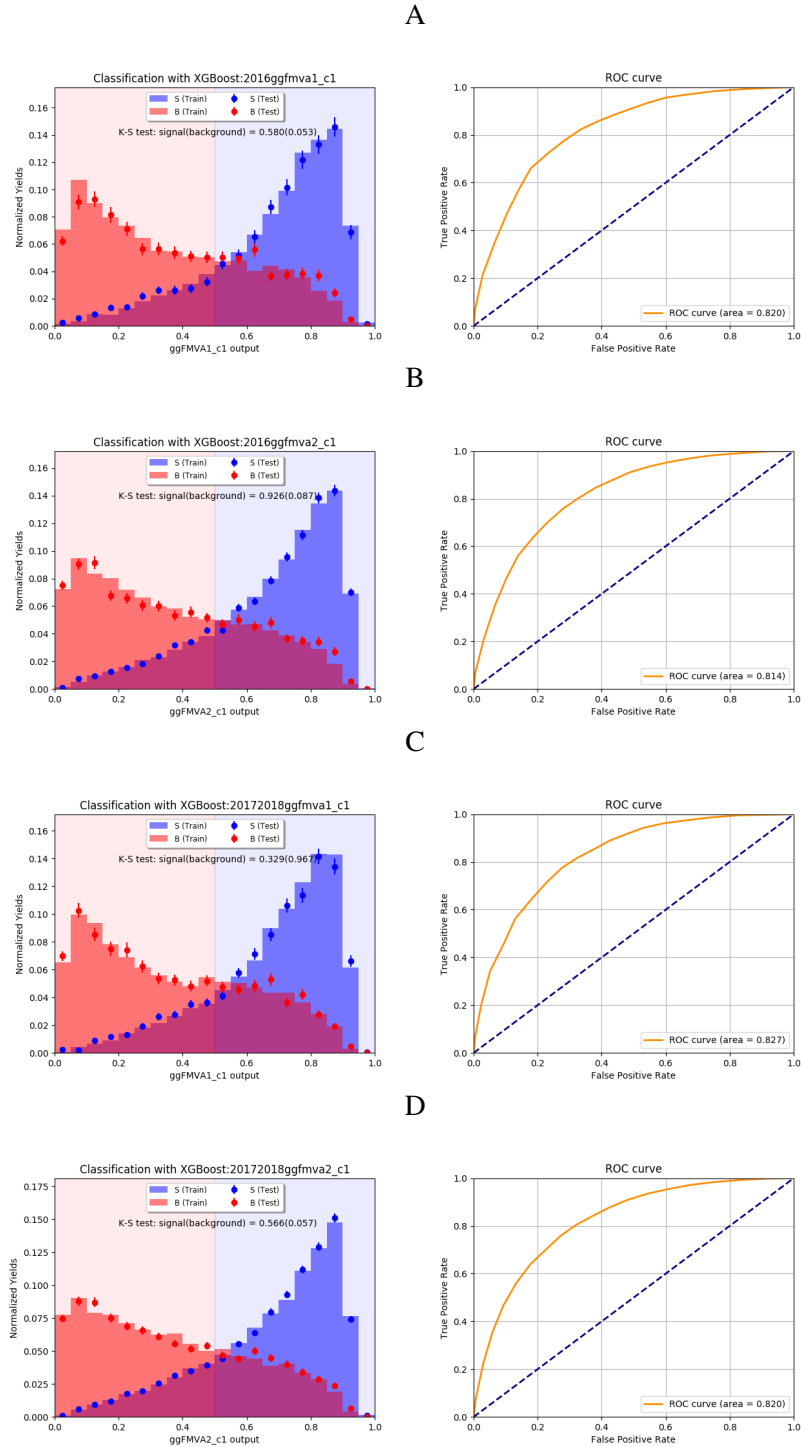
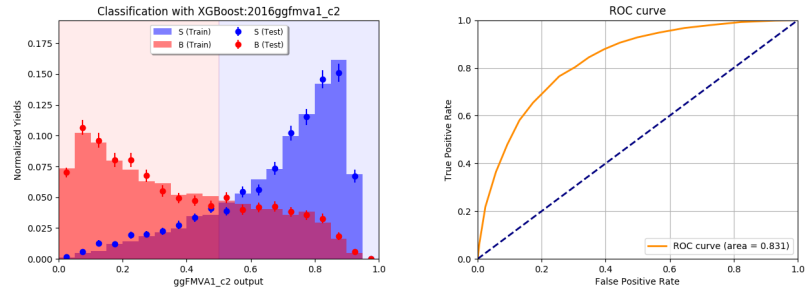
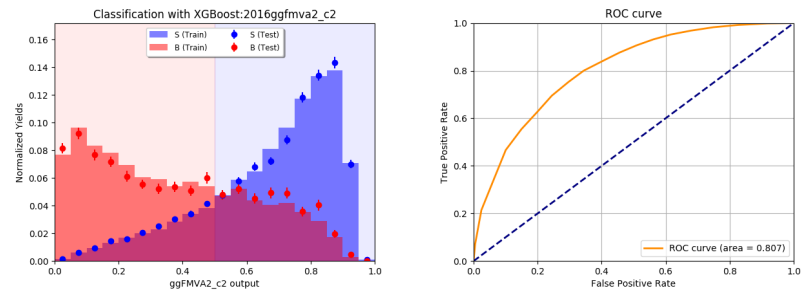


Figure G-4. Training performance of the classifier C2 used for building the ggF observables. A) 2016 Category 1, B) 2016 Category 2, C) 2017-2018 Category 1, D) 2017-2018 Category 2. The left figure shows the trained classifier C1 distribution in train/test samples. The right figure shows the ROC curve of the test samples.

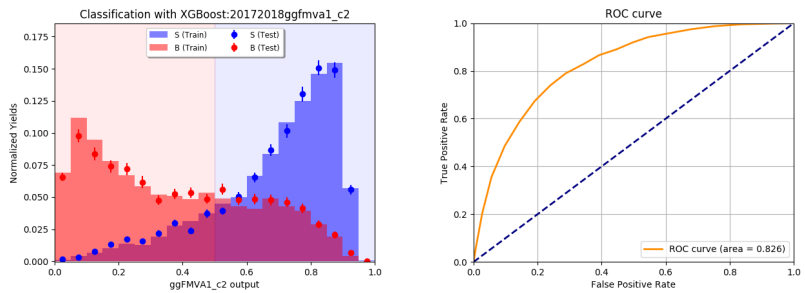
A



B



C



D

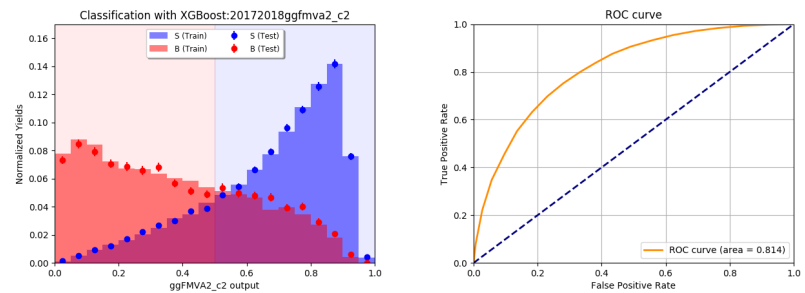


Table G-2. Performance of the BDT output training in Train and Test samples for the classifiers C1 and C2.

| MVA (Sample)     | 2016 (C1) | 2016 (C2) | 2017-2018 (C1) | 2017-2018 (C2) |
|------------------|-----------|-----------|----------------|----------------|
| ggF cat1 (Train) | 0.844     | 0.843     | 0.836          | 0.840          |
| ggF cat1 (Test)  | 0.820     | 0.830     | 0.827          | 0.826          |
| ggF cat2 (Train) | 0.829     | 0.828     | 0.827          | 0.827          |
| ggF cat2 (Test)  | 0.814     | 0.807     | 0.820          | 0.814          |

Figure G-5. The relative importance of the training variables for classifier C1. A) 2016 ggF category 1, B) 2016 ggF category 2, C) 2017-2018 ggF category 1, D) 2017-2018 ggF category 2.

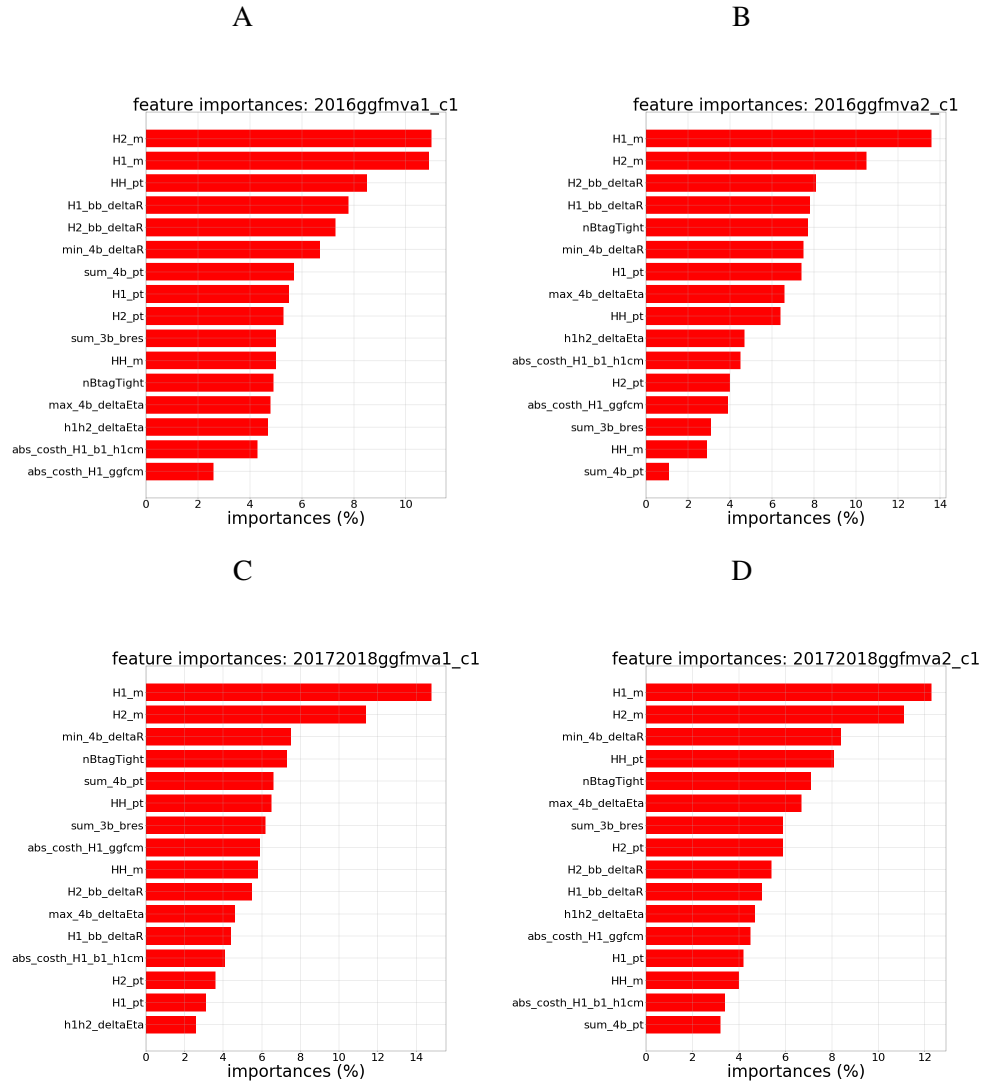
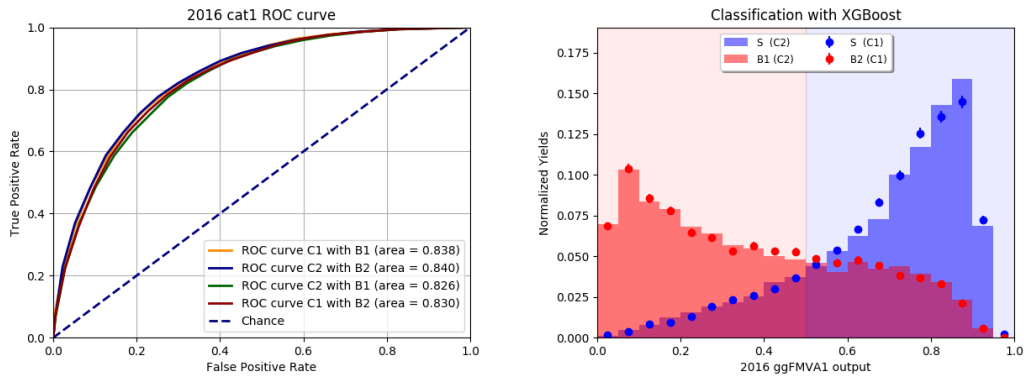
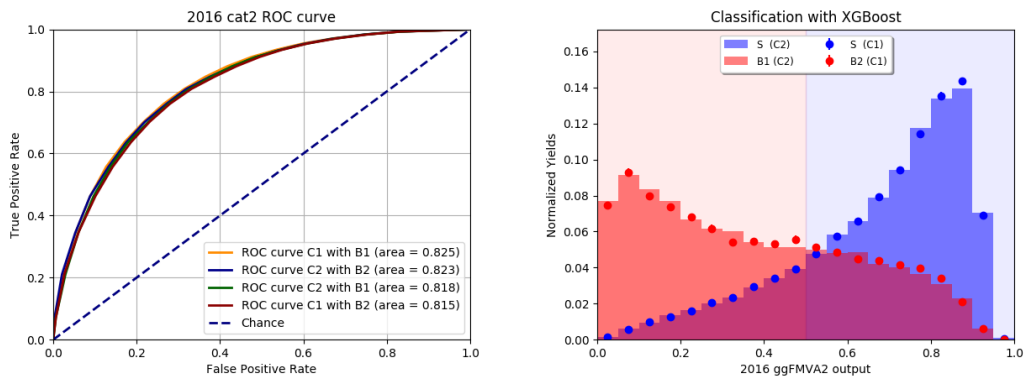


Figure G-6. Compatibility checks on the classifiers used for the 2016 ggF observables. A) Category 1 and B) Category 2. The left figure shows the ROC-AUC curves performance checks. The right figure illustrates the distributions for signal S and background model B1 (B2) samples using the C2 (C1) classifier.

A



B





## REFERENCES

- [1] CMS collaboration, *Search for Higgs boson pair production in the four  $b$  quark final state*, Tech. Rep. [CMS-PAS-HIG-20-005](#) (2021).
- [2] A. Purcell, “Go on a particle quest at the first CERN webfest.” <https://cds.cern.ch/record/1473657> (2012).
- [3] S.G. Karshenboim, *Precision physics of simple atoms: QED tests, nuclear structure and fundamental constants*, *Phys. Rept.* **422** (2005) 1 [[arXiv:0509010](#)].
- [4] M. Peskin and D. Schroeder, *Introduction to Quantum Field Theory*, Addison-Wesley, Reading, U.S.A., 1st ed. (1995), [10.1201/9780429503559](#).
- [5] G. Costa and G. Fogli, *Symmetries and group theory in particle physics. An introduction to space-time and internal symmetries*, vol. 823, Springer, Heidelberg, Germany (2012), [10.1007/978-3-642-15482-9](#).
- [6] P.A. Zyla et al., *Review of Particle Physics*, *PTEP* **2020** (2020) 083C01.
- [7] S. Weinberg, *A Model of Leptons*, *Phys. Rev. Lett.* **19** (1967) 1264.
- [8] A. Salam, *Weak and Electromagnetic Interactions*, *Conf. Proc. C* **680519** (1968) 367.
- [9] S.L. Glashow, *Partial Symmetries of Weak Interactions*, *Nucl. Phys.* **22** (1961) 579.
- [10] P.W. Higgs, *Broken symmetries, massless particles and gauge fields*, *Phys. Lett.* **12** (1964) 132.
- [11] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, *Phys. Rev. Lett.* **13** (1964) 321.
- [12] G.S. Guralnik, C.R. Hagen and T.W.B. Kibble, *Global Conservation Laws and Massless Particles*, *Phys. Rev. Lett.* **13** (1964) 585.
- [13] ATLAS collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett. B* **716** (2012) 1 [[arXiv:1207.7214](#)].
- [14] CMS collaboration, *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*, *Phys. Lett. B* **716** (2012) 30 [[arXiv:1207.7235](#)].
- [15] ATLAS collaboration, *Evidence for the spin-0 nature of the Higgs boson using ATLAS data*, *Phys. Lett. B* **726** (2013) 120 [[arXiv:1307.1432](#)].
- [16] CMS collaboration, *Study of the Mass and Spin-Parity of the Higgs Boson Candidate Via Its Decays to Z Boson Pairs*, *Phys. Rev. Lett.* **110** (2013) 081803 [[arXiv:1212.6639](#)].
- [17] 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** (2015) 191803 [[arXiv:1503.07589](#)].

- [18] ATLAS collaboration, *Combined measurements of Higgs boson production and decay using up to  $80\text{ fb}^{-1}$  of proton-proton collision data at  $\sqrt{s} = 13\text{ TeV}$  collected with the ATLAS experiment*, *Phys. Rev. D* **101** (2020) 012002 [[arXiv:1909.02845](#)].
- [19] ATLAS collaboration, *Measurement of the Higgs boson mass in the  $H \rightarrow ZZ^* \rightarrow 4\ell$  and  $H \rightarrow \gamma\gamma$  channels with  $\sqrt{s} = 13\text{ TeV}$   $pp$  collisions using the ATLAS detector*, *Phys. Lett. B* **784** (2018) 345 [[arXiv:1806.00242](#)].
- [20] CMS collaboration, *Combined Higgs boson production and decay measurements with up to  $137\text{ fb}^{-1}$  of proton-proton collision data at  $\sqrt{s} = 13\text{ TeV}$* , Tech. Rep. [CMS-PAS-HIG-19-005](#) (2020).
- [21] CMS collaboration, *A measurement of the Higgs boson mass in the diphoton decay channel*, *Phys. Lett. B* **805** (2020) 135425 [[arXiv:2002.06398](#)].
- [22] ATLAS collaboration, *Observation of  $H \rightarrow b\bar{b}$  decays and  $VH$  production with the ATLAS detector*, *Phys. Lett. B* **786** (2018) 59 [[arXiv:1808.08238](#)].
- [23] CMS collaboration, *Observation of Higgs boson decay to bottom quarks*, *Phys. Rev. Lett.* **121** (2018) 121801 [[arXiv:1808.08242](#)].
- [24] ATLAS collaboration, *Cross-section measurements of the Higgs boson decaying into a pair of  $\tau$ -leptons in proton-proton collisions at  $\sqrt{s} = 13\text{ TeV}$  with the ATLAS detector*, *Phys. Rev. D* **99** (2019) 072001 [[arXiv:1811.08856](#)].
- [25] CMS collaboration, *Observation of the Higgs boson decay to a pair of  $\tau$  leptons with the CMS detector*, *Phys. Lett. B* **779** (2018) 283 [[arXiv:1708.00373](#)].
- [26] ATLAS collaboration, *Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector*, *Phys. Lett. B* **784** (2018) 173 [[arXiv:1806.00425](#)].
- [27] CMS collaboration, *Observation of  $t\bar{t}H$  production*, *Phys. Rev. Lett.* **120** (2018) 231801 [[arXiv:1804.02610](#)].
- [28] CMS collaboration, *Evidence for Higgs boson decay to a pair of muons*, *JHEP* **01** (2021) 148 [[arXiv:2009.04363](#)].
- [29] ATLAS collaboration, *Evidence for Higgs boson decays to a low-mass dilepton system and a photon in  $pp$  collisions at  $\sqrt{s} = 13\text{ TeV}$  with the ATLAS detector*, *Phys. Lett. B* **819** (2021) 136412 [[arXiv:2103.10322](#)].
- [30] ATLAS collaboration, *Direct constraint on the Higgs-charm coupling from a search for Higgs boson decays to charm quarks with the ATLAS detector*, Tech. Rep. [ATLAS-CONF-2021-021](#) (2021).
- [31] CMS collaboration, *A search for the standard model Higgs boson decaying to charm quarks*, *JHEP* **03** (2020) 131 [[arXiv:1912.01662](#)].

- [32] G. Degrand et al., *Probing the Higgs self coupling via single Higgs production at the LHC*, *JHEP* **12** (2016) 080 [[arXiv:1607.04251](#)].
- [33] F. Maltoni et al., *Trilinear Higgs coupling determination via single-Higgs differential measurements at the LHC*, *Eur. Phys. J. C* **77** (2017) 887 [[arXiv:1709.08649](#)].
- [34] J. Alison et al., *Higgs boson potential at colliders: Status and perspectives*, *Rev. Phys.* **5** (2020) 100045 [[arXiv:1910.00012](#)].
- [35] J. Elias-Miro et al., *Higgs mass implications on the stability of the electroweak vacuum*, *Phys. Lett. B* **709** (2012) 222 [[arXiv:1112.3022](#)].
- [36] T. Robens and T. Stefaniak, *Status of the Higgs Singlet Extension of the Standard Model after LHC Run 1*, *Eur. Phys. J. C* **75** (2015) 104 [[arXiv:1501.02234](#)].
- [37] S. Dawson and I.M. Lewis, *NLO corrections to double Higgs boson production in the Higgs singlet model*, *Phys. Rev. D* **92** (2015) 094023 [[arXiv:1508.05397](#)].
- [38] U. Ellwanger et al., *The Next-to-Minimal Supersymmetric Standard Model*, *Phys. Rept.* **496** (2010) 1 [[arXiv:0910.1785](#)].
- [39] A. Djouadi et al., *The post-Higgs MSSM scenario: Habemus MSSM?*, *Eur. Phys. J. C* **73** (2013) 2650 [[arXiv:1307.5205](#)].
- [40] L. Randall and R. Sundrum, *A large mass hierarchy from a small extra dimension*, *Phys. Rev. Lett.* **83** (1999) 3370 [[arXiv:9905221](#)].
- [41] A. Carvalho, *Gravity particles from Warped Extra Dimensions, predictions for LHC*, [arXiv:1404.0102](#).
- [42] U. Ellwanger and M. Rodriguez-Vazquez, *Simultaneous search for extra light and heavy Higgs bosons via cascade decays*, *JHEP* **11** (2017) 008 [[arXiv:1707.08522](#)].
- [43] T. Robens et al., *Two-real-scalar-singlet extension of the SM: LHC phenomenology and benchmark scenarios*, *Eur. Phys. J. C* **80** (2020) 151 [[arXiv:1908.08554](#)].
- [44] R. Frederix et al., *Higgs pair production at the LHC with NLO and parton-shower effects*, *Phys. Lett. B* **732** (2014) 142 [[arXiv:1401.7340](#)].
- [45] F. Goertz et al., *Higgs boson pair production in the  $D=6$  extension of the SM*, *JHEP* **04** (2015) 167 [[arXiv:1410.3471](#)].
- [46] G. Buchalla et al., *Higgs boson pair production in non-linear Effective Field Theory with full  $m_t$ -dependence at NLO QCD*, *JHEP* **09** (2018) 057 [[arXiv:1806.05162](#)].
- [47] A. Carvalho et al., *Higgs Pair Production: Choosing Benchmarks With Cluster Analysis*, *JHEP* **04** (2016) 126 [[arXiv:1507.02245](#)].
- [48] J. de Blas et al., *Higgs Boson Studies at Future Particle Colliders*, *JHEP* **01** (2020) 139 [[arXiv:1905.03764](#)].

- [49] ATLAS collaboration, *Searches for Higgs boson pair production in the  $hh \rightarrow bb\tau\tau, \gamma\gamma WW^*, \gamma\gamma bb, bbbb$  channels with the ATLAS detector*, *Phys. Rev. D* **92** (2015) 092004 [[arXiv:1509.04670](#)].
- [50] CMS collaboration, *Search for Higgs boson pair production in the  $bb\tau\tau$  final state in proton-proton collisions at  $\sqrt{s} = 8$  TeV*, *Phys. Rev. D* **96** (2017) 072004 [[arXiv:1707.00350](#)].
- [51] ATLAS collaboration, *Combination of searches for Higgs boson pairs in  $pp$  collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, *Phys. Lett. B* **800** (2020) 135103 [[arXiv:1906.02025](#)].
- [52] CMS collaboration, *Combination of searches for Higgs boson pair production in proton-proton collisions at  $\sqrt{s} = 13$  TeV*, *Phys. Rev. Lett.* **122** (2019) 121803 [[arXiv:1811.09689](#)].
- [53] ATLAS collaboration, *Search for Higgs boson pair production in the  $\gamma\gamma b\bar{b}$  final state with 13 TeV  $pp$  collision data collected by the ATLAS experiment*, *JHEP* **11** (2018) 040 [[arXiv:1807.04873](#)].
- [54] ATLAS collaboration, *Search for resonant and non-resonant Higgs boson pair production in the  $b\bar{b}\tau^+\tau^-$  decay channel in  $pp$  collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, *Phys. Rev. Lett.* **121** (2018) 191801 [[arXiv:1808.00336](#)].
- [55] ATLAS collaboration, *Search for pair production of Higgs bosons in the  $b\bar{b}b\bar{b}$  final state using proton-proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, *JHEP* **01** (2019) 030 [[arXiv:1804.06174](#)].
- [56] ATLAS collaboration, *Search for Higgs boson pair production in the  $WW^{(*)}WW^{(*)}$  decay channel using ATLAS data recorded at  $\sqrt{s} = 13$  TeV*, *JHEP* **05** (2019) 124 [[arXiv:1811.11028](#)].
- [57] ATLAS collaboration, *Search for Higgs boson pair production in the  $\gamma\gamma WW^*$  channel using  $pp$  collision data recorded at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, *Eur. Phys. J. C* **78** (2018) 1007 [[arXiv:1807.08567](#)].
- [58] CMS collaboration, *Search for production of Higgs boson pairs in the four  $b$  quark final state using large-area jets in proton-proton collisions at  $\sqrt{s} = 13$  TeV*, *JHEP* **01** (2019) 040 [[arXiv:1808.01473](#)].
- [59] CMS collaboration, *Search for nonresonant Higgs boson pair production in the  $b\bar{b}b\bar{b}$  final state at  $\sqrt{s} = 13$  TeV*, *JHEP* **04** (2019) 112 [[arXiv:1810.11854](#)].
- [60] CMS collaboration, *Search for resonant pair production of Higgs bosons decaying to bottom quark-antiquark pairs in proton-proton collisions at 13 TeV*, *JHEP* **08** (2018) 152 [[arXiv:1806.03548](#)].
- [61] CMS collaboration, *Search for Higgs boson pair production in the  $\gamma\gamma b\bar{b}$  final state in  $pp$  collisions at  $\sqrt{s} = 13$  TeV*, *Phys. Lett. B* **788** (2019) 7 [[arXiv:1806.00408](#)].

- [62] CMS collaboration, *Search for Higgs boson pair production in events with two bottom quarks and two tau leptons in proton–proton collisions at  $\sqrt{s}=13\text{TeV}$* , *Phys. Lett. B* **778** (2018) 101 [[arXiv:1707.02909](#)].
- [63] CMS collaboration, *Search for heavy resonances decaying into two Higgs bosons or into a Higgs boson and a W or Z boson in proton-proton collisions at 13 TeV*, *JHEP* **01** (2019) 051 [[arXiv:1808.01365](#)].
- [64] CMS collaboration, *Search for resonant and nonresonant Higgs boson pair production in the  $b\bar{b}\ell\nu\ell\nu$  final state in proton-proton collisions at  $\sqrt{s}=13\text{ TeV}$* , *JHEP* **01** (2018) 054 [[arXiv:1708.04188](#)].
- [65] ATLAS collaboration, *Constraints on the Higgs boson self-coupling from the combination of single-Higgs and double-Higgs production analyses performed with the ATLAS experiment*, Tech. Rep. [ATLAS-CONF-2019-049](#) (2019).
- [66] ATLAS collaboration, *Search for Higgs boson pair production in the two bottom quarks plus two photons final state in pp collisions at  $\sqrt{s}=13\text{ TeV}$  with the ATLAS detector*, Tech. Rep. [ATLAS-CONF-2021-016](#) (2021).
- [67] ATLAS collaboration, *Search for resonant and non-resonant Higgs boson pair production in the  $b\bar{b}\tau^+\tau^-$  decay channel using 13 TeV pp collision data from the ATLAS detector*, Tech. Rep. [ATLAS-CONF-2021-030](#) (2021).
- [68] ATLAS collaboration, *Search for non-resonant Higgs boson pair production in the  $b\bar{b}\ell\nu\ell\nu$  final state with the ATLAS detector in pp collisions at  $\sqrt{s}=13\text{ TeV}$* , *Phys. Lett. B* **801** (2020) 135145 [[arXiv:1908.06765](#)].
- [69] CMS collaboration, *Search for nonresonant Higgs boson pair production in final states with two bottom quarks and two photons in proton-proton collisions at  $\sqrt{s}=13\text{ TeV}$* , *JHEP* **03** (2021) 257 [[arXiv:2011.12373](#)].
- [70] CMS collaboration, *Search for nonresonant Higgs boson pair production in the 4 leptons plus 2 b jets final state in proton-proton collisions at  $\sqrt{s}=13\text{ TeV}$* , Tech. Rep. [CMS-PAS-HIG-20-004](#) (2020).
- [71] ATLAS collaboration, *Search for the  $HH\rightarrow b\bar{b}b\bar{b}$  process via vector-boson fusion production using proton-proton collisions at  $\sqrt{s}=13\text{ TeV}$  with the ATLAS detector*, *JHEP* **07** (2020) 108 [[arXiv:2001.05178](#)].
- [72] CMS collaboration, *Search for Higgs boson pair production via vector boson fusion with highly Lorentz-boosted Higgs bosons in the four b quark final state at  $\sqrt{s}=13\text{ TeV}$* , Tech. Rep. [CMS-PAS-B2G-21-001](#) (2021).
- [73] CMS collaboration, *Search for resonant Higgs boson pair production in four b quark final state using large-area jets in proton-proton collisions at  $\sqrt{s}=13\text{ TeV}$* , Tech. Rep. [CMS-PAS-B2G-20-004](#) (2021).



- [74] CMS collaboration, *Search for heavy resonances decaying to a pair of boosted Higgs bosons in final states with leptons and a bottom quark-antiquark pair at  $\sqrt{s} = 13$  TeV*, Tech. Rep. [CMS-PAS-B2G-20-007](#) (2021).
- [75] ATLAS collaboration, *Combination of searches for non-resonant and resonant Higgs boson pair production in the  $b\bar{b}\gamma\gamma$ ,  $b\bar{b}\tau^+\tau^-$  and  $b\bar{b}b\bar{b}$  decay channels using  $pp$  collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, Tech. Rep. [ATLAS-CONF-2021-052](#) (2021).
- [76] CMS collaboration, *Identification of heavy-flavour jets with the CMS detector in  $pp$  collisions at 13 TeV*, *JINST* **13** (2018) P05011 [[arXiv:1712.07158](#)].
- [77] CMS collaboration, *Identification of heavy, energetic, hadronically decaying particles using machine-learning techniques*, *JINST* **15** (2020) P06005 [[arXiv:2004.08262](#)].
- [78] ATLAS collaboration, *ATLAS  $b$ -jet identification performance and efficiency measurement with  $t\bar{t}$  events in  $pp$  collisions at  $\sqrt{s} = 13$  TeV*, *Eur. Phys. J. C* **79** (2019) 970 [[arXiv:1907.05120](#)].
- [79] ATLAS collaboration, *Identification of boosted Higgs bosons decaying into  $b$ -quark pairs with the ATLAS detector at 13 TeV*, *Eur. Phys. J. C* **79** (2019) 836 [[arXiv:1906.11005](#)].
- [80] CMS collaboration, *A Deep Neural Network for Simultaneous Estimation of  $b$  Jet Energy and Resolution*, *Comput. Softw. Big Sci.* **4** (2020) 10 [[arXiv:1912.06046](#)].
- [81] ATLAS collaboration, *Measurement prospects of the pair production and self-coupling of the Higgs boson with the ATLAS experiment at the HL-LHC*, Tech. Rep. [ATL-PHYS-PUB-2018-053](#) (2018).
- [82] CMS collaboration, *Prospects for  $HH$  measurements at the HL-LHC*, Tech. Rep. [CMS-PAS-FTR-18-019](#) (2018).
- [83] M. Cepeda et al., *Report from Working Group 2: Higgs Physics at the HL-LHC and HE-LHC*, *CERN Yellow Rep. Monogr.* **7** (2019) 221 [[arXiv:1902.00134](#)].
- [84] E. Mobs, “The CERN accelerator complex. Complexe des accélérateurs du CERN.” <https://cds.cern.ch/record/2197559> (2016).
- [85] O. Brüning et al., *LHC Design Report*, CERN Yellow Reports: Monographs, CERN, Geneva, CH (2004), [10.5170/CERN-2004-003-V-1](#).
- [86] CMS collaboration, *The CMS Experiment at the CERN LHC*, *JINST* **3** (2008) S08004.
- [87] CMS collaboration, *CMS Technical Design Report for the Pixel Detector Upgrade*, Tech. Rep. [CERN-LHCC-2012-016](#), [CMS-TDR-11](#) (2012).
- [88] CMS collaboration, *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software*, Tech. Rep. [CERN-LHCC-2006-001](#), [CMS-TDR-8-1](#) (2006).
- [89] CMS collaboration, *Energy Calibration and Resolution of the CMS Electromagnetic Calorimeter in  $pp$  Collisions at  $\sqrt{s} = 7$  TeV*, *JINST* **8** (2013) P09009 [[arXiv:1306.2016](#)].

- [90] CMS collaboration, *Performance of the CMS Drift Tube Chambers with Cosmic Rays*, *JINST* **5** (2010) T03015 [[arXiv:0911.4855](#)].
- [91] CMS collaboration, *Performance of the CMS Cathode Strip Chambers with Cosmic Rays*, *JINST* **5** (2010) T03018 [[arXiv:0911.4992](#)].
- [92] CMS collaboration, *Performance Study of the CMS Barrel Resistive Plate Chambers with Cosmic Rays*, *JINST* **5** (2010) T03017 [[arXiv:0911.4045](#)].
- [93] ALEPH collaboration, *Performance of the ALEPH detector at LEP*, *Nucl. Instrum. Meth. A* **360** (1995) 481.
- [94] CMS collaboration, *Particle-flow reconstruction and global event description with the CMS detector*, *JINST* **12** (2017) P10003 [[arXiv:1706.04965](#)].
- [95] M. Cacciari et al., *The anti- $k_t$  jet clustering algorithm*, *JHEP* **04** (2008) 063 [[arXiv:0802.1189](#)].
- [96] M. Cacciari et al., *FastJet User Manual*, *Eur. Phys. J. C* **72** (2012) 1896 [[arXiv:1111.6097](#)].
- [97] CMS collaboration, *Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV*, *JINST* **12** (2017) P02014 [[arXiv:1607.03663](#)].
- [98] CMS collaboration, *Jet algorithms performance in 13 TeV data*, Tech. Rep. [CMS-PAS-JME-16-003](#) (2017).
- [99] CMS collaboration, *Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV*, *JINST* **13** (2018) P05011 [[arXiv:1712.07158](#)].
- [100] CMS collaboration, *Identification of b-quark jets with the CMS Experiment*, *JINST* **8** (2013) P04013 [[arXiv:1211.4462](#)].
- [101] CMS collaboration, *Identification of heavy, energetic, hadronically decaying particles using machine-learning techniques*, *JINST* **15** (2020) P06005 [[arXiv:2004.08262](#)].
- [102] E. Bols et al., *Jet Flavour Classification Using DeepJet*, *JINST* **15** (2020) P12012 [[arXiv:2008.10519](#)].
- [103] CMS collaboration, *Performance of reconstruction and identification of  $\tau$  leptons decaying to hadrons and  $\nu_\tau$  in pp collisions at  $\sqrt{s} = 13$  TeV*, *JINST* **13** (2018) P10005 [[arXiv:1809.02816](#)].
- [104] CMS collaboration, *Performance of missing transverse momentum reconstruction in proton-proton collisions at  $\sqrt{s} = 13$  TeV using the CMS detector*, *JINST* **14** (2019) P07004 [[arXiv:1903.06078](#)].
- [105] CMS collaboration, *Performance of the CMS Level-1 trigger in proton-proton collisions at  $\sqrt{s} = 13$  TeV*, *JINST* **15** (2020) P10017 [[arXiv:2006.10165](#)].

- [106] D. Contardo et al., *Technical Proposal for the Phase-II Upgrade of the CMS Detector*, Tech. Rep. [CERN-LHCC-2015-010](#), [LHCC-P-008](#), [CMS-TDR-15-02](#) (2015).
- [107] O. Aberle et al., *High-Luminosity Large Hadron Collider (HL-LHC): Technical design report*, CERN Yellow Reports: Monographs, Geneva, CH (2020), [10.23731/CYRM-2020-0010](#).
- [108] CMS collaboration, *The Phase-2 Upgrade of the CMS Tracker*, Tech. Rep. [CERN-LHCC-2017-009](#), [CMS-TDR-014](#) (2017).
- [109] CMS collaboration, *The Phase-2 Upgrade of the CMS Barrel Calorimeters*, Tech. Rep. [CERN-LHCC-2017-011](#), [CMS-TDR-015](#) (2017).
- [110] CMS collaboration, *The Phase-2 Upgrade of the CMS Endcap Calorimeter*, Tech. Rep. [CERN-LHCC-2017-023](#), [CMS-TDR-019](#) (2017).
- [111] CMS collaboration, *The Phase-2 Upgrade of the CMS Muon Detectors*, Tech. Rep. [CERN-LHCC-2017-012](#), [CMS-TDR-016](#) (2017).
- [112] CMS collaboration, *Technical proposal for a MIP timing detector in the CMS experiment Phase 2 upgrade*, Tech. Rep. [CERN-LHCC-2017-027](#), [LHCC-P-009](#) (2017).
- [113] CMS collaboration, *The Phase-2 Upgrade of the CMS Level-1 Trigger*, Tech. Rep. [CERN-LHCC-2020-004](#), [CMS-TDR-021](#) (2020).
- [114] T. Chen and C. Guestrin, *Xgboost: A scalable tree boosting system*, in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD 785, ACM, New York, NY, U.S.A. (2016) [[arXiv:1603.02754](#)].
- [115] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **07** (2014) 079 [[arXiv:1405.0301](#)].
- [116] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159 [[arXiv:1410.3012](#)].
- [117] GEANT4 collaboration, *GEANT4 – a simulation toolkit*, *Nucl. Instrum. Meth. A* **506** (2003) 250.
- [118] A. Ferrari et al., *FLUKA: A multi-particle transport code (program version 2005)*, CERN Yellow Reports: Monographs, Geneva, CH (2005), [10.5170/CERN-2005-010](#).
- [119] T. Gleisberg et al., *Event generation with SHERPA 1.1*, *JHEP* **02** (2009) 007 [[arXiv:0811.4622](#)].
- [120] S. Alioli et al., *NLO vector-boson production matched with shower in POWHEG*, *JHEP* **07** (2008) 060 [[arXiv:0805.4802](#)].
- [121] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040 [[arXiv:0409146](#)].



- [122] S. Frixione et al., *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, *JHEP* **11** (2007) 070 [[arXiv:0709.2092](#)].
- [123] M. Czakon and A. Mitov, *Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders*, *Comput. Phys. Commun.* **185** (2014) 2930 [[arXiv:1112.5675](#)].
- [124] A. Rogozhnikov, *Reweighting with Boosted Decision Trees*, *J. Phys. Conf. Ser.* **762** (2016) 012036 [[arXiv:1608.05806](#)].
- [125] ATLAS collaboration, *Search for pair production of higgsinos in final states with at least three b-tagged jets in  $\sqrt{s} = 13$  TeV pp collisions using the ATLAS detector*, *Phys. Rev. D* **98** (2018) 092002 [[arXiv:1806.04030](#)].
- [126] ATLAS collaboration, *Search for resonances decaying into a weak vector boson and a Higgs boson in the fully hadronic final state produced in proton–proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, *Phys. Rev. D* **102** (2020) 112008 [[arXiv:2007.05293](#)].
- [127] LHCb collaboration, *Search for decays of neutral beauty mesons into four muons*, *JHEP* **03** (2017) 001 [[arXiv:1611.07704](#)].
- [128] F. Pedregosa et al., *Scikit-learn: Machine learning in python*, *Journal of Machine Learning Research* **12** (2011) 2825 [[arXiv:1201.0490](#)].
- [129] R. D. Cousins, “Generalization of chisquare goodness-of-Fit test for binned data using saturated models, with application to histograms.” [http://www.physics.ucla.edu/~cousins/stats/cousins\\_saturated.pdf](http://www.physics.ucla.edu/~cousins/stats/cousins_saturated.pdf) (2013).
- [130] CMS collaboration, *Precision luminosity measurement in proton-proton collisions at  $\sqrt{s} = 13$  TeV in 2015 and 2016 at CMS*, *Eur. Phys. J. C* **81** (2021) 800 [[arXiv:2104.01927](#)].
- [131] CMS collaboration, *CMS luminosity measurement for the 2017 data-taking period at  $\sqrt{s} = 13$  TeV*, Tech. Rep. [CMS-PAS-LUM-17-004](#) (2018).
- [132] CMS collaboration, *CMS luminosity measurement for the 2018 data-taking period at  $\sqrt{s} = 13$  TeV*, Tech. Rep. [CMS-PAS-LUM-18-002](#) (2019).
- [133] LHC Higgs Cross Section Working Group, “SM Higgs Branching Ratios and Total Decay Widths.” <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageBR> (2016).
- [134] R.J. Barlow and C. Beeston, *Fitting using finite Monte Carlo samples*, *Comput. Phys. Commun.* **77** (1993) 219.
- [135] ATLAS, CMS, LHC HIGGS COMBINATION GROUP collaboration, *Procedure for the LHC Higgs boson search combination in Summer 2011*, Tech. Rep. [CMS-NOTE-2011-005](#), [ATL-PHYS-PUB-2011-11](#) (2011).

- [136] G. Cowan et al., *Asymptotic formulae for likelihood-based tests of new physics*, *Eur. Phys. J. C* **71** (2011) 1554 [[arXiv:1007.1727](#)].
- [137] A.L. Read, *Presentation of search results: The CL(s) technique*, *J. Phys. G* **28** (2002) 2693.
- [138] E. Gross and O. Vitells, *Trial factors for the look elsewhere effect in high energy physics*, *Eur. Phys. J. C* **70** (2010) 525 [[arXiv:1005.1891](#)].
- [139] H. Qu and L. Gouskos, *ParticleNet: Jet Tagging via Particle Clouds*, *Phys. Rev. D* **101** (2020) 056019 [[arXiv:1902.08570](#)].
- [140] A. Hocker et al., *TMVA - Toolkit for Multivariate Data Analysis*, [arXiv:physics/0703039](#).
- [141] A. Carvalho et al., *On the reinterpretation of non-resonant searches for Higgs boson pairs*, *JHEP* **02** (2021) 049 [[arXiv:1710.08261](#)].

## BIOGRAPHICAL SKETCH

Daniel Fernando Guerrero Ibarra was born in Quito, Ecuador on November 29th, 1991. He lived most of his childhood and teenage years in a small town at the heart of the Ecuadorian coast called Quevedo. After finishing high school, he moved back to Quito to pursue a career in science at Escuela Politécnica Nacional, where he enrolled in the physics program. As an undergraduate student, he had the unique opportunity to engage early on in experimental particle physics research in the CMS experiment with the supervision and mentorship of Prof. Edgar Carrera and Prof. Francisco Yumiceva. During this time, he pursued his research in top quark physics and studies on hadron calorimeters at the CMS experiment by visiting the CERN Laboratory in Geneva, Switzerland, and the LHC Physics Center (LPC) at Fermilab in Chicago, IL. After obtaining his bachelor's degree, he joined the Physics Department at the University of Florida to pursue a doctorate focused on experimental high energy physics at the CMS experiment under the supervision of Prof. Jacobo Konigsberg. After finishing his coursework, he moved to CERN for two years where he was involved in the operations of the L1 trigger, and investigated the areas of the trigger upgrades and searches for Higgs boson pair production. During his final year, he completed his data analysis and trigger research through the prestigious Fermilab LPC Graduate Scholar award. He received a Doctorate of Philosophy in Physics at the University of Florida in the fall of 2021.