

MUCH ADO ABOUT NOTHING: SEARCHES FOR HIGGS BOSON DECAYS  
TO INVISIBLE PARTICLES

by

AMANDA LYNN STEINHEBEL

A DISSERTATION

Presented to the Department of Physics  
and the Division of Graduate Studies of the University of Oregon  
in partial fulfillment of the requirements  
for the degree of  
Doctor of Philosophy

June 2021



## DISSERTATION APPROVAL PAGE

Student: Amanda Lynn Steinhebel

Title: Much Ado About Nothing: Searches Higgs Boson Decays to Invisible Particles

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Physics by:

|                    |                              |
|--------------------|------------------------------|
| Stephanie Majewski | Chairperson                  |
| Jim Brau           | Advisor                      |
| Tim Cohen          | Core Member                  |
| Marina Guenza      | Institutional Representative |
| and                |                              |

|              |                                           |
|--------------|-------------------------------------------|
| Andy Karduna | Interim Vice Provost for Graduate Studies |
|--------------|-------------------------------------------|

Original approval signatures are on file with the University of Oregon Division of Graduate Studies.

Degree awarded June 2021

© 2021 Amanda Lynn Steinhebel  
This work is licensed under a Creative Commons  
**Attribution-NonCommercial-NoDerivs (United States) License.**



## DISSERTATION ABSTRACT

Amanda Lynn Steinhebel

Doctor of Philosophy

Department of Physics

June 2021

Title: Much Ado About Nothing: Searches Higgs Boson Decays to Invisible Particles

The Standard Model (SM) predicts a branching ratio of the Higgs boson decaying to invisible particles of  $\mathcal{O}(0.001)$ , though current measurements have only set weak upper limits on this value. The small SM-allowed rate can be enhanced if the Higgs boson decays into new particles such as dark matter. This dissertation considers three searches for this signature.

Two efforts involve  $139 \text{ fb}^{-1}$  of  $\sqrt{s} = 13 \text{ TeV}$  proton-proton collision data collected at the LHC with the ATLAS detector. Special attention is paid to the Vector Boson Fusion production channel - the most sensitive LHC channel to this measurement. The lack of statistically significant tension with the SM allowed an upper limit on the observed (expected) branching ratio  $\mathcal{B}_{h \rightarrow \text{inv}}$  to be placed at  $\mathcal{B}_{h \rightarrow \text{inv}} < 0.13$  (0.13) at 95% confidence level. This result also informed a statistical combination of three independent ATLAS searches for this signature through different Higgs production mechanisms, refining the upper limit to



$\mathcal{B}_{h \rightarrow \text{inv}} < 0.11$  (0.11) obs (exp) - the strictest limit set to date. As ATLAS collects more data, this value will continue to improve.

The proposed International Linear Collider ‘Higgs factory’ will provide unprecedented precision of this electroweak measurement. Preliminary studies presented in this dissertation use  $1800 \text{ fb}^{-1}$  of  $\sqrt{s} = 250 \text{ GeV}$  polarized electron–positron data modeled with the SiD detector and predict an upper limit of  $\mathcal{B}_{h \rightarrow \text{inv}} \lesssim 0.003$ . This could provide the first evidence of non-SM Higgs boson decays.

This dissertation includes previously published and unpublished coauthored material.

## CURRICULUM VITAE

NAME OF AUTHOR: Amanda Lynn Steinhebel

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, Oregon  
College of Wooster, Wooster, Ohio

DEGREES AWARDED:

Doctor of Philosophy, Physics, 2021, University of Oregon  
Master of Science, Physics, 2018, University of Oregon  
Bachelor of Arts, Physics and Mathematics, 2015, College of Wooster

AREAS OF SPECIAL INTEREST:

High-energy hadron collider physics, high-energy lepton collider physics,  
Higgs boson couplings, dark matter, detector development and design

PROFESSIONAL EXPERIENCE:

University of Oregon, Graduate Research Assistant, ATLAS Experiment and  
SiD Consortium, Feb. 2016 - present

University of Oregon, Graduate Teaching Assistant, Oct. 2015 - Feb. 2016

College of Wooster, Undergraduate Researcher, Aug. 2014 - May 2015

GRANTS, AWARDS AND HONORS:

Lokey Doctoral Science Fellowship, University of Oregon, 2020

First-Year Academic Award, University of Oregon, 2015

PUBLICATIONS:

List of publications with significant contributions:

A. Steinhebel. Search for invisibly decaying vector boson fusion produced Higgs bosons with 139/fb of pp collisions with the ATLAS detector. In *Proceedings of 40th International Conference on High Energy physics — PoS(ICHEP2020)*, volume 390, page 107, 2020.

ATLAS Collaboration. Combination of searches for invisible Higgs boson decays with the ATLAS experiment, Technical Report ATLAS-CONF-2020-052, CERN, Geneva, Oct 2020. URL <http://cds.cern.ch/record/2743055>.

ATLAS Collaboration. Search for invisible Higgs boson decays with vector boson fusion signatures with the ATLAS detector using an integrated luminosity of  $139 \text{ fb}^{-1}$ , Technical Report ATLAS-CONF-2020-008, CERN, Geneva, Oct 2020. URL <http://cds.cern.ch/record/2715447>.

J. Brau and A. Steinhebel. Studies of the Response of the SiD Silicon-Tungsten ECal, arXiv:1703.08605 [physics.ins-det]

Additional publications can be found at <https://inspirehep.net/authors/1629836>

## ACKNOWLEDGEMENTS

I would like to thank my advisor, Jim Brau, for providing opportunities for me to grow as a scientist all around the world. Thank you for supporting my constant learning and new ideas, and meeting them with patience.

Thank you to every University of Oregon faculty member that taught and supported me on this journey, and especially Stephanie Majewski. You embody superwoman and have always been able to make time for me in your impossible schedule. Thank you to Anne McGinley for humoring even my most ignorant of email requests and never failing to make me feel valued. I also want to thank my past professors from the College of Wooster - John Lindner, Cody Leary, Susan Lehman, Robert Wooster, and Nelia Mann - for refusing to let me doubt myself and encouraging me to set my standards high.

Thank you to my fellow University of Oregon graduate students Alice Greenberg and Kara Zappitelli for late night study sessions for the qualifying exam and the initiative to begin and sustain the University of Oregon Women in Physics Group. My entire graduate experience would have been different without you two by my side in Eugene.

Thank you to my European support system - John Myers, Alison Jelden, Jason Barkeloo, Kate Whalen, Aparajita Dattagupta, Jochen Heinrich, Walter Hopkins, Nicholas Luongo, Alexander Bachiu, and especially Elliot Parrish. During my time at CERN, I learned as much about myself and my relationships as I did about physics. I will always be utterly grateful for my time spent there, and for the role you all have played in making me feel welcomed, supported, and loved. I

could not have gotten here without your friendships. I am grateful for all the other friends I have made along the way, too many to list, in physics and beyond.

Thank you to my friends and colleagues who helped me prepare this work  
- Tim Cohen, Kate Whalen, Jochen Heinrich, Ben Carlson, Loan Truong, Jan Strube, and Michael Belden. I owe you all a beer.

Thank you to my family for your support and excitement over a physics graduate education, though you rarely knew what I was talking about or why I would be excited about it. Thank you for believing in me to create my own path.

Finally, thank you to my dearest friends, Ava Myers and Michael Belden. Thank you for laughing, crying, ranting, and traveling with me so many times over all these years. I love you both more than I can express. I would not be who I am, here, without you.

This work is dedicated to every young girl who is afraid of being labeled a ‘nerd’  
for liking science or math, and for every woman studying science and feeling  
isolated or like an imposter. Science is for all of us. We belong.

This dissertation is the product of unionized labor as part of the Graduate  
Teaching Fellows Federation, AFT Local 3544.

# TABLE OF CONTENTS

| Chapter |                                                     | Page |
|---------|-----------------------------------------------------|------|
| I.      | INTRODUCTION . . . . .                              | 1    |
| II.     | STANDARD MODEL HIGGS AND DARK MATTER . . . . .      | 3    |
|         | 2.1. The Standard Model . . . . .                   | 3    |
|         | 2.2. The Higgs Boson . . . . .                      | 6    |
|         | 2.3. Dark Matter . . . . .                          | 13   |
| III.    | LHC AND ATLAS . . . . .                             | 17   |
|         | 3.1. The Large Hadron Collider . . . . .            | 17   |
|         | 3.2. ATLAS . . . . .                                | 24   |
| IV.     | SIMULATION AND RECONSTRUCTION . . . . .             | 46   |
|         | 4.1. Monte Carlo Simulation . . . . .               | 46   |
|         | 4.2. Physics Objects . . . . .                      | 49   |
| V.      | VBF $H \rightarrow$ INVISIBLE . . . . .             | 67   |
|         | 5.1. $h \rightarrow$ invisible Motivation . . . . . | 67   |

| Chapter                                                         | Page |
|-----------------------------------------------------------------|------|
| 5.2. VBF Motivation . . . . .                                   | 68   |
| 5.3. First look at Run 2 . . . . .                              | 69   |
| 5.4. Closer look at Run 2 . . . . .                             | 103  |
| VI. $H \rightarrow$ INVISIBLE COMBINATION . . . . .             | 112  |
| 6.1. Combination Overview and Strategy . . . . .                | 112  |
| 6.2. Input Analyses . . . . .                                   | 113  |
| 6.3. Shared MC Samples and Duplicated Event Numbers . . . . .   | 118  |
| 6.4. Orthogonality . . . . .                                    | 119  |
| 6.5. Impact on Limit . . . . .                                  | 122  |
| 6.6. Fitting and Results . . . . .                              | 124  |
| VII. $H \rightarrow$ INVISIBLE IN FUTURE COLLIDERS . . . . .    | 127  |
| 7.1. Future Endeavors . . . . .                                 | 127  |
| 7.2. High Luminosity-LHC . . . . .                              | 129  |
| 7.3. HL-LHC $h \rightarrow$ invisible Projections . . . . .     | 132  |
| 7.4. The International Linear Collider . . . . .                | 135  |
| 7.5. SiD Detector Concept . . . . .                             | 141  |
| 7.6. $h \rightarrow$ invisible Projections at the ILC . . . . . | 144  |
| VIII. SUMMARY . . . . .                                         | 150  |
| 8.1. Standard Model Higgs Boson and Dark Matter . . . . .       | 150  |



| Chapter                                                      | Page |
|--------------------------------------------------------------|------|
| 8.2. LHC and ATLAS . . . . .                                 | 151  |
| 8.3. Simulation and Reconstruction . . . . .                 | 152  |
| 8.4. VBF $h \rightarrow$ invisible . . . . .                 | 153  |
| 8.5. $h \rightarrow$ invisible Combination . . . . .         | 156  |
| 8.6. $h \rightarrow$ invisible in Future Colliders . . . . . | 156  |
| 8.7. Conclusion . . . . .                                    | 157  |
| IX. CONCLUSION . . . . .                                     | 159  |
| APPENDICES                                                   |      |
| A. REFERENCE LIST OF ABBREVIATIONS . . . . .                 | 160  |
| B. ELECTROMAGNETIC CALORIMETRY . . . . .                     | 162  |
| C. TRIGGER SYSTEM EVOLUTION FROM RUN 1 TO RUN 2 . . . . .    | 165  |
| D. FORWARD JET VERTEX TAGGING . . . . .                      | 167  |
| E. VBF+MET OVERLAP REMOVAL . . . . .                         | 169  |
| F. VBF+MET YIELDS . . . . .                                  | 171  |
| G. HL-LHC SMEARING OF TRUTH-LEVEL OBJECTS FOR VBF            |      |

| Chapter                                                                  | Page |
|--------------------------------------------------------------------------|------|
| $H \rightarrow$ INVISIBLE . . . . .                                      | 173  |
| G.1. Electrons . . . . .                                                 | 173  |
| G.2. Photons . . . . .                                                   | 174  |
| G.3. Muons . . . . .                                                     | 176  |
| G.4. Neutrinos . . . . .                                                 | 177  |
| G.5. Jets . . . . .                                                      | 178  |
| G.6. Missing Transverse Energy . . . . .                                 | 181  |
| G.7. Missing Transverse Energy with Leptons . . . . .                    | 183  |
| G.8. Missing Transverse Energy with Jets . . . . .                       | 183  |
| H. SID ELECTROMAGNETIC CALORIMETER STUDIES AND<br>OPTIMIZATION . . . . . | 185  |
| H.1. Backscatter and Shower Spreading . . . . .                          | 185  |
| H.2. Geometry . . . . .                                                  | 185  |
| H.3. Shower Calibration . . . . .                                        | 187  |
| H.4. Test Beam Studies . . . . .                                         | 190  |
| H.5. KPiX Test Beam . . . . .                                            | 194  |
| REFERENCES CITED . . . . .                                               | 196  |

## LIST OF FIGURES

| Figure |                                                                                                                                                                              | Page |
|--------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| 2.1.   | Visualization of the fundamental particles described by the Standard Model [1] . . . . .                                                                                     | 4    |
| 2.2.   | The Higgs potential from Eq. 2.1 with $\mu^2 < 0$ [2] . . . . .                                                                                                              | 9    |
| 2.3.   | a) Higgs production from $pp$ collisions [3], b) the dominant production mechanism, gluon-gluon fusion, and c) the subdominant production, vector boson fusion [4] . . . . . | 11   |
| 2.4.   | a) Higgs production from $e^+e^-$ collisions and b) the dominant production mechanism at $\sqrt{s} = 250$ GeV, Higgsstrahlung [5] . . . . .                                  | 12   |
| 2.5.   | Higgs boson decay modes in the SM ( $m_h = 125.10 \pm 0.14$ GeV) [3] . . .                                                                                                   | 13   |
| 2.6.   | DM coupling to the SM through a portal (shaded region) [6] . . . . .                                                                                                         | 14   |
| 2.7.   | Spectrum of allowed DM masses and models [7] . . . . .                                                                                                                       | 14   |
| 3.1.   | The accelerator complex at CERN [8] . . . . .                                                                                                                                | 18   |
| 3.2.   | A Run 1 candidate $Z \rightarrow \mu\mu$ event with 25 reconstructed vertices [9] . . .                                                                                      | 23   |
| 3.3.   | Pileup conditions in Run 2 [10] . . . . .                                                                                                                                    | 24   |
| 3.4.   | The ATLAS detector [11] . . . . .                                                                                                                                            | 26   |
| 3.5.   | Rectangular and cylindrical coordinates used with ATLAS [12] . . . . .                                                                                                       | 27   |
| 3.6.   | The ATLAS inner detector [11, 13] . . . . .                                                                                                                                  | 29   |
| 3.7.   | ATLAS calorimeter system with electromagnetic and hadronic calorimeters [11] . . . . .                                                                                       | 32   |
| 3.8.   | LAr barrel module featuring accordion-shaped electrodes and trigger towers [11] . . . . .                                                                                    | 33   |
| 3.9.   | LAr material budget in $\eta$ [11] . . . . .                                                                                                                                 | 34   |

| Figure                                                                                                                                 | Page |
|----------------------------------------------------------------------------------------------------------------------------------------|------|
| 3.10. Tile barrel wedge design and optical readout system [11] . . . . .                                                               | 36   |
| 3.11. Tile material budget in $\eta$ [11] . . . . .                                                                                    | 36   |
| 3.12. Cut-away of the muon system [11] . . . . .                                                                                       | 37   |
| 3.13. Cross section of the MDT chamber arrangement [11] . . . . .                                                                      | 38   |
| 3.14. Muon trigger system [11] . . . . .                                                                                               | 39   |
| 3.15. ATLAS toroid magnet system and tile calorimeter steel [11] . . . . .                                                             | 40   |
| 3.16. The barrel toroid installed underground (human on platform for scale)<br>[11] . . . . .                                          | 41   |
| 3.17. Forward calorimeter layout [14] . . . . .                                                                                        | 42   |
| 3.18. Flow of the trigger and data acquisition system for Run 2 [15] . . . . .                                                         | 44   |
| 3.19. Contribution of different data streams to the output HLT total<br>bandwidth [16] . . . . .                                       | 45   |
| 4.1. Pictorial view of the steps necessary to generate MC events [17] . . . . .                                                        | 47   |
| 4.2. Signatures left in different ATLAS systems for different particles<br>[18] . . . . .                                              | 50   |
| 4.3. Energy deposit clustering with different jet-finding algorithms [19] . . . . .                                                    | 53   |
| 4.4. JVT score for pileup (PU) and hard scatter (HS) jets with $20 < p_T < 30$ GeV [20] . . . . .                                      | 56   |
| 4.5. Tracks associated with a $b$ -jet, including the secondary vertex and<br>impact parameter $d_0$ [21] . . . . .                    | 57   |
| 4.6. Signature from a boosted top quark [22] . . . . .                                                                                 | 59   |
| 4.7. Distributions of a jet substructure variable for large- $R$ jets originating<br>from top quarks and QCD background [23] . . . . . | 59   |
| 4.8. MET trigger efficiencies for Run 2 [24] . . . . .                                                                                 | 66   |
| 5.1. Candidate $h \rightarrow \text{inv.}$ event from Run 2 . . . . .                                                                  | 70   |
| 5.2. Feynman diagrams of the signal process (top) and example $Z$ +jets<br>irreducible backgrounds . . . . .                           | 71   |

| Figure                                                                                                                                                                                                    | Page |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| 5.3. $\eta$ value of the third highest $p_T$ jet in events compared between a) EMTopo and b) PFlow jet collections . . . . .                                                                              | 74   |
| 5.4. Binning scheme for VBF+MET first look. Percentage values reflect the fraction of signal events in each bin, and increasing intensity of grayscale reflects increasing bin sensitivity [25] . . . . . | 78   |
| 5.5. Gluon-initiated jets have larger values of <b>NTracks</b> than quark-initiated jets, as shown in VBF Higgs and $Z$ strong simulated events . . . . .                                                 | 79   |
| 5.6. Using truth information regarding jet-initiating partons emulates a perfect tagger. With perfect full detector tagging, the expected limit ( $\sigma_\mu Z$ ) improves by a relative 5.5%. . . . .   | 80   |
| 5.7. $S/B$ yields and associated expected limit ( $\sigma_\mu Z$ Norm) with different quark/gluon tagging related cuts . . . . .                                                                          | 81   |
| 5.8. The $p_T$ distribution of a) leading and b) subleading jets in the signal region. . . . .                                                                                                            | 82   |
| 5.9. $CL_S$ scan values over the signal strength $\mu$ for 36.1 fb $^{-1}$ of blinded data using an $m_{jj}$ -only binning scheme a) before and b) after additional quark-tagged binning. . . . .         | 82   |
| 5.10. Event distribution in SR-like validation region ( $100 < E_T^{\text{miss}} < 150$ GeV) from VBF trigger . . . . .                                                                                   | 84   |
| 5.11. Distributions of two-jet events broken into categories based upon number of contaminating PU jets. . . . .                                                                                          | 86   |
| 5.12. Distributions of two-jet events after requiring $E_T^{\text{miss}} > 150$ GeV and $p_T(j_0, j_1) > 80, 50$ GeV . . . . .                                                                            | 87   |
| 5.13. Regions of interest in the $m_{jj}/\Delta\phi_{jj}$ plane . . . . .                                                                                                                                 | 87   |
| 5.14. Full truth $\Delta\phi_{jj}$ and $m_{jj}$ distributions of the merged nominal QCD+afterburner filtered file collection after reweighting, merging, and renormalizing the filtered file . . . . .    | 89   |
| 5.15. $E_T^{\text{miss}}$ distribution for VBF trigger-passing 2 jet events with additional $E_T^{\text{jet, no JVT}}$ requirements . . . . .                                                             | 92   |
| 5.16. Visual of $V$ +jets CRs used to estimate contribution to SR . . . . .                                                                                                                               | 93   |

| Figure                                                                                                                                                  | Page |
|---------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| 5.17. Visual of data and MC in ZCR used to calculate the normalization factor $\beta$ . . . . .                                                         | 94   |
| 5.18. Overview of Rebalance and Smear MJ background estimation technique . . . . .                                                                      | 95   |
| 5.19. Visualization of postfit data and backgrounds in all SR and CR bins . . . . .                                                                     | 102  |
| 5.20. Postfit background estimation and data as a function of a) $\Delta\phi_{jj}$ and b) $m_{jj}$ . . . . .                                            | 103  |
| 5.21. Top 20 most impactful NPs in the fit . . . . .                                                                                                    | 104  |
| 5.22. Updated binning scheme for VBF+MET (compare to Figure 5.4) . . . .                                                                                | 106  |
| 5.23. Definition of FJVT CR as compared to the SR . . . . .                                                                                             | 108  |
| 5.24. Distribution in low-MET FJVT CR (leading jet fJVT < 0.2) of a) $m_{jj}$ and b) $\Delta\phi_{jj}$ . . . . .                                        | 108  |
| 5.25. Low-MET FJVT CR normalization factors ( $1\sigma$ statistical uncertainties shown in green band) . . . . .                                        | 110  |
| 6.1. Yields and overlap fraction of all analyses for signal samples . . . . .                                                                           | 121  |
| 6.2. Yields and overlap fraction of all analyses in observed data events for (a) all SR and CR events, (b) for only events in the signal regions. . . . | 122  |
| 6.3. Observed and expected NLL scan of combinations . . . . .                                                                                           | 125  |
| 6.4. Observed and expected 95% CL limits for input analyses and combinations . . . . .                                                                  | 126  |
| 7.1. LHC schedule for future running [26] . . . . .                                                                                                     | 127  |
| 7.2. A simulated $t\bar{t}$ HL-LHC event with two hundred reconstructed vertices [27] . . . . .                                                         | 130  |
| 7.3. Placement of HGTD (green and purple disks) in the upgraded ATLAS detector for HL-LHC [28] . . . . .                                                | 131  |
| 7.4. Smeared jet collection from VBF signal sample compared to truth . . . .                                                                            | 134  |

| Figure                                                                                                                                                                                                               | Page |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| 7.5. Simulated production of $ZH \rightarrow \ell\ell b\bar{b}$ in the LHC ( $\ell = e$ ) with ATLAS in Run 1 (a) and in the ILC ( $\ell = \mu$ ) with ILD (b) [29, 30] . . . . .                                    | 136  |
| 7.6. ILC layout [31] . . . . .                                                                                                                                                                                       | 137  |
| 7.7. Projected precision of Higgs boson couplings of HL-LHC with and without ILC250 input [32] . . . . .                                                                                                             | 138  |
| 7.8. Anticipated timeline for the ILC [31] . . . . .                                                                                                                                                                 | 140  |
| 7.9. The SiD detector [30] . . . . .                                                                                                                                                                                 | 141  |
| 7.10. The SiD ECal (dark grays) and HCal (light gray) configuration, with overlapping trapezoidal modules with thin and thick absorber layers . . . . .                                                              | 142  |
| 7.11. Recoil mass distributions of events passing all kinematic criteria of Table 7.1 . . . . .                                                                                                                      | 147  |
| C.1. Trigger systems for a) Run 1 [33] and b) Run 2 [15] . . . . .                                                                                                                                                   | 165  |
| G.1. Truth and smeared electron $p_T$ distributions from a $W \rightarrow e\nu$ sample with a) MAXHTPTV 140-280 and b) smearing resolution . . . . .                                                                 | 174  |
| G.2. Truth and smeared photon a) $p_T$ distributions, b) ratio of distributions, and c) smearing resolution (c) from a VBF $Z \rightarrow \gamma\gamma$ sample . . . . .                                             | 176  |
| G.3. Truth and smeared muon $p_T$ distributions for a) $p_T < 140$ GeV, b) the full smeared $p_T$ range, and c) smearing resolution from a $W \rightarrow \mu\nu$ sample with MAXHTPTV 140-280 . . . . .             | 178  |
| G.4. $\eta$ distribution of truth muons compared to muons with smeared $p_T > 3$ TeV . . . . .                                                                                                                       | 178  |
| G.5. Jet $\eta$ distribution for truth and smeared jet collections with and without JVT consideration for VBF $h \rightarrow$ inv. signal . . . . .                                                                  | 180  |
| G.6. Smeared and pileup jet $p_T$ distributions for $W \rightarrow e\nu$ MAXHTPTV 0-70 sample with JVT applied up to a) 100 GeV and b) 200 GeV . . . . .                                                             | 180  |
| G.7. Truth and smeared jet $p_T$ distributions for a) VBF $h \rightarrow$ inv. signal, b) $Z \rightarrow \nu\nu$ sample with MAXHTPTV 0-70, and c) smearing resolution for VBF $h \rightarrow$ inv. signal . . . . . | 181  |
| G.8. Approximated soft $E_T^{\text{miss}}$ distribution for VBF $h \rightarrow$ inv. signal . . . . .                                                                                                                | 182  |

| Figure                                                                                                                                                                                                                                        | Page |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| G.9. Truth and smeared $E_T^{\text{miss}}$ distributions for a) VBF $h \rightarrow \text{inv. signal}$ , b) $Z \rightarrow \nu\nu$ sample with MAXHTPTV 0-70, and c) smearing resolution for VBF $h \rightarrow \text{inv. signal}$ . . . . . | 183  |
| G.10. Calculated $E_T^{\text{jet, no JVT}}$ distribution from smeared quantities for VBF $h \rightarrow \text{inv. signal}$ . . . . .                                                                                                         | 184  |
| H.1. a) Absorber depth and b) number of sampling layers of SiD as a function of $\phi$ . The overlap region falls within [4,15] . . . . .                                                                                                     | 186  |
| H.2. Fraction of electromagnetic shower energy deposited beyond the ECal in the HCal . . . . .                                                                                                                                                | 187  |
| H.3. Scaled energy resolution of the ECal barrel with and without leakage correction . . . . .                                                                                                                                                | 189  |
| H.4. Contribution of linear and nonlinear terms of Eq. H.1 to calibrated energy . . . . .                                                                                                                                                     | 190  |
| H.5. Estimated leakage by different techniques for a scan of incident photon energies . . . . .                                                                                                                                               | 190  |
| H.6. Performance of linear (blue) and nonlinear (red) techniques . . . . .                                                                                                                                                                    | 191  |
| H.7. Uncertainty in calibration per energy resolution . . . . .                                                                                                                                                                               | 191  |
| H.8. Data collected with the SiD ECal prototype a) before and b) after cleaning . . . . .                                                                                                                                                     | 192  |
| H.9. Test beam data runs and simulated data match for a) silicon-first and b) tungsten-first configurations . . . . .                                                                                                                         | 193  |
| H.10. Efficiency of incident particle counting algorithm for simulated events with two incident particles . . . . .                                                                                                                           | 194  |
| H.11. a) Counting algorithm applied to the full test beam data set and b) comparison of identified two-particle events between simulation and data . . . . .                                                                                  | 195  |



## LIST OF TABLES

| Table                                                                                                                                                                                                                                                | Page |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| 2.1. SM particles with their quantum numbers and masses [34]. $Y_W$ lists<br>left- (LH) and right-handed (RH) values. . . . .                                                                                                                        | 7    |
| 3.1. LHC data set nomenclature and parameters . . . . .                                                                                                                                                                                              | 24   |
| 3.2. The $ \eta $ expanse of ATLAS subsystems . . . . .                                                                                                                                                                                              | 42   |
| 5.1. ATLAS upper limits (UL) of $h \rightarrow \text{inv.}$ and $\mathcal{B}_{\text{undet.}}$ . . . . .                                                                                                                                              | 69   |
| 5.2. Percent contribution to background events in SR . . . . .                                                                                                                                                                                       | 72   |
| 5.3. Summary of event generators used for simulation . . . . .                                                                                                                                                                                       | 73   |
| 5.4. Signal and control region definitions . . . . .                                                                                                                                                                                                 | 76   |
| 5.5. HLT trigger requirements for the VBF trigger, and additional offline<br>plateau selection requirements for fully-efficient trigger emulation . . . .                                                                                            | 84   |
| 5.6. Percent of events from truth level that fall within the given region after<br>reconstruction . . . . .                                                                                                                                          | 88   |
| 5.7. Yield values from the merged filtered file + nominal MC QCD<br>sample . . . . .                                                                                                                                                                 | 89   |
| 5.8. Expected and observed limits on $\mathcal{B}_{h \rightarrow \text{inv.}}$ calculated at the 95% C.L.<br>with $139 \text{ fb}^{-1}$ . . . . .                                                                                                    | 101  |
| 5.9. Relative impact $\Delta$ of the 95% CL expected upper limit on $\mathcal{B}_{h \rightarrow \text{inv.}}$ if a<br>group of uncertainties is “removed”, i.e. if the corresponding nuisance<br>parameters are fixed to the nominal values. . . . . | 104  |
| 5.10. SR cuts that have been updated (unlisted variables remain identical to<br>Table 5.4) . . . . .                                                                                                                                                 | 105  |
| 5.11. The usage of MJ estimates by SR bins. A “-” indicates that the<br>method is not used. . . . .                                                                                                                                                  | 107  |
| 5.12. FJVT CR normalization factors along with their statistical and<br>systematic uncertainties. The discrepancy in SR1 between the FJVT                                                                                                            |      |

| Table                                                                                                                                                                                                                                                                                                                     | Page |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| and the R&S is handled by taking the average of the two predictions<br>(labeled “comb”.’) with an uncertainty that covers the difference. . . . .                                                                                                                                                                         | 109  |
| 5.13. MJ yield estimations . . . . .                                                                                                                                                                                                                                                                                      | 110  |
| 6.1. NPs correlated between Run 1 and Run 2 analyses . . . . .                                                                                                                                                                                                                                                            | 113  |
| 6.2. Event selection criteria for included $t\bar{t} + E_{\text{T}}^{\text{miss}}$ signal regions . . . . .                                                                                                                                                                                                               | 115  |
| 6.3. Expected and observed ULs at 95% C.L. from the individual $t\bar{t}H$ input<br>analyses and their combination . . . . .                                                                                                                                                                                              | 116  |
| 6.4. Impact of groups of NPs on the best-fit POI value . . . . .                                                                                                                                                                                                                                                          | 124  |
| 6.5. Summary of 95% CL limits on $\mathcal{B}_{h \rightarrow \text{inv}}$ of the individual search regions<br>. . . . .                                                                                                                                                                                                   | 125  |
| 7.1. Cutflow for $h \rightarrow \text{inv.}$ search with leptonic $Z$ decays . . . . .                                                                                                                                                                                                                                    | 147  |
| 7.2. Cutflow for $h \rightarrow \text{inv.}$ search with hadronic $Z$ decays . . . . .                                                                                                                                                                                                                                    | 148  |
| 7.3. Combined final yields of Tables 7.1 and 7.2 for 1800 fb $^{-1}$ of LR/RL<br>polarized beams . . . . .                                                                                                                                                                                                                | 149  |
| E.1. Overlap removal between objects for VBF+MET searches [25] . . . . .                                                                                                                                                                                                                                                  | 169  |
| E.2. Overlap removal of truth smeared objects for VBF+MET HL-LHC<br>projection studies . . . . .                                                                                                                                                                                                                          | 170  |
| F.1. Observed and expected background event yields with associated<br>uncertainties in the SR and CRs after the<br>likelihood fit. A dash (-) indicates that the corresponding yield is < 0.5<br>events. . . . .                                                                                                          | 171  |
| F.2. Yields of data, signal and major backgrounds in the SR for each bin<br>after the likelihood fit. Minor backgrounds from $t\bar{t}$ , $VV$ , $VVV$ ,<br>and VBF $H \rightarrow W^+W^- / \tau^+\tau^-$ are summed up as “Other”. A dash (-)<br>indicates that the corresponding yield is less than 0.5 events. . . . . | 172  |
| G.1. Optimization of forward ( $ \eta  > 3.8$ ) jet $p_{\text{T}}$ requirement in $E_{\text{T}}^{\text{miss}}$<br>definition . . . . .                                                                                                                                                                                    | 182  |

## CHAPTER I

### INTRODUCTION

The Standard Model of particle physics, which describes fundamental particles and their interactions, has held up to rigorous experimental testing. Today, much of this testing is done at the Large Hadron Collider (LHC) - a hadron collider focusing on  $\sqrt{s} = 13$  TeV proton-proton collisions. In 2018, Run 2 came to an end after collecting  $139 \text{ fb}^{-1}$  of proton-proton data. The large resulting data set is used both to refine measurements of the properties of known particles and to search for new particles. The presence of new physics not currently encompassed by the Standard Model is motivated by observations that the theory cannot yet explain. This includes the identity of dark matter. A simple extension of the Standard Model introduces dark matter that couples directly to the Higgs boson, resulting in Higgs bosons decaying to particles that do not directly interact with particle detectors. The presence of new physics can be deduced by measuring this signature,  $h \rightarrow \text{invisible}$ , at a rate different from that predicted by the Standard Model.

This dissertation considers three modern searches for  $h \rightarrow \text{inv.}$  signatures in hadron and lepton colliders. The full LHC Run 2 data set is considered, first in a search for Higgs bosons produced through vector boson fusion - the most sensitive  $h \rightarrow \text{inv.}$  channel at the LHC. Results from this search are included in a statistical combination with two other full Run 2  $h \rightarrow \text{inv.}$  searches and the Run 1  $h \rightarrow \text{inv.}$  combination to refine the  $h \rightarrow \text{inv.}$  measurement. This resulted in the strictest experimentally observed upper limit on the  $h \rightarrow \text{inv.}$  branching ratio to date. Simulated studies of the International Linear Collider, a proposed electron-positron

lepton collider, are also presented. These studies illustrate the precision power of a lepton machine for such electroweak measurements.

Chapter II provides theoretical background for the Standard Model with emphasis on Higgs boson properties, including its production and decays. It considers extensions of the theory that could incorporate dark matter. Chapter III presents an overview of the Large Hadron Collider and the ATLAS detector, used to collect proton–proton collision data. Chapter IV overviews simulated data creation and identification and reconstruction strategies for physics objects from detector signatures. Chapter V details the search for  $h \rightarrow \text{inv.}$  produced through vector boson fusion utilizing the full Run 2 LHC data set, and Chapter VI combines this result and others in a statistical combination to refine the  $h \rightarrow \text{inv.}$  measurement. These two chapters contain material coauthored with the ATLAS collaboration. Chapter VII considers future prospects of this search with upgraded LHC running scenarios and in the lepton collider environment of the proposed International Linear Collider. Chapter VIII is a stand-alone non-technical summary of this thesis, written for the general population. Conclusions are presented in Chapter IX.

## CHAPTER II

### STANDARD MODEL HIGGS AND DARK MATTER

#### 2.1. The Standard Model

Since the discovery of the electron in 1897 [35], a zoo of fundamental particles has been both theoretically predicted and experimentally observed. This century-worth of work informs the Standard Model of particle physics (SM) - a theoretical framework that describes the fundamental particles that have been discovered thus far and their interactions (see Figure 2.1). The most recent experimental addition is the 2012 discovery of the Higgs boson (see Section 2.2). Thus far, the model has been extremely successful at describing particles and their interactions, but it cannot yet account for all observed phenomena. Modern research is aimed both at the continued testing of the accuracy of the theory, as well as exploring potential extended theories that incorporate new physics currently beyond the Standard Model (BSM) that may account for these unpredicted phenomena. The goal is to achieve more complete understanding of the fundamental phenomena that govern how the universe behaves.

Each particle is defined by properties generally called quantum numbers that determine possible interactions. Like the periodic table of chemical elements, SM particles can be categorized based upon these shared characteristics. SM forces are conveyed by the exchange of gauge bosons (blue in Figure 2.1) between fermions (red and green in Figure 2.1)<sup>1</sup>. Also called vector bosons, they all carry spin of 1. The Higgs boson (see Section 2.2, purple in Figure 2.1) also carries integer

---

<sup>1</sup>Gluons also self-interact, conveying the strong force between other gluons.

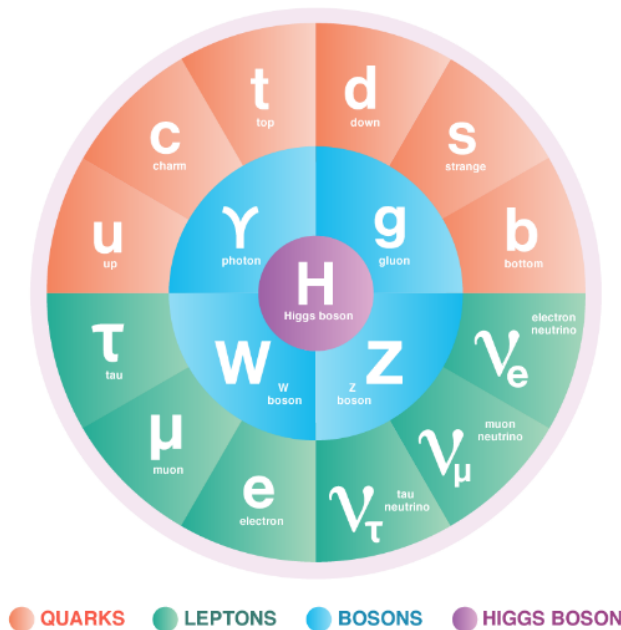


FIGURE 2.1. Visualization of the fundamental particles described by the Standard Model [1]

spin (though the spin is zero) and is responsible for the mass generation of all SM particles.

Electromagnetism is conveyed by the photon and affects particles that carry electromagnetic (EM) charge. This includes charged leptons (green in Figure 2.1) and quarks (red in Figure 2.1), which carry spin of  $1/2$  and are collectively known as fermions. They have electromagnetic charge ( $1/2$  for leptons and  $2/3$  or  $-1/3$  for quarks). Photons themselves do not carry EM charge.

The weak force is communicated by massive  $W$  and  $Z$  bosons and affects particles carrying weak isospin and weak hypercharge. SM fermions have left- or right-handed chirality, which is associated with its weak isospin value ( $T = 1/2$  for left-chiral and  $T = 0$  for right-chiral). Left-handed fermions form doublet structures, pairing up-type ( $u, c, t$ ) and down-type ( $d, s, b$ ) quarks or charged and neutral leptons. They are differentiated by the third component of weak isospin  $T_3$ ,

where  $T_3 = 1/2$  for up-type quarks and neutrinos and  $T_3 = -1/2$  for down-type quarks and charged leptons. Right-handed fermions remain singlets with  $T_3 = 0$ :

$$\begin{pmatrix} u \\ d \end{pmatrix}_L, u_R, d_R \text{ and } \begin{pmatrix} \nu \\ e^- \end{pmatrix}_L, e_R^- \text{ with } T_3 \text{ configuration } \begin{pmatrix} +1/2 \\ -1/2 \end{pmatrix}, 0.$$

The EM charge  $Q$  is related to  $T_3$  and hypercharge  $Y_W$  by  $Q = T_3 + Y_W/2$ . Neither the  $W$  nor  $Z$  boson carries weak hypercharge, but the  $W$  does carry EM charge and therefore has nonzero  $T_3$ . This means that any interaction involving the  $W$  impacts the  $T_3$  value of the involved fermions. Namely, only left-handed fermions can interact with  $W$  bosons. The  $Z$  boson does not carry EM charge or  $T_3$ , so interactions involving the  $Z$  boson can occur with left- or right-handed fermions where only momentum, energy, and spin are impacted.

At high energies, the electromagnetic and weak forces unify into the electroweak force. In order to preserve gauge invariance, the gauge fields must be massless. Similarly, the parity violation of the weak interaction treating left- and right-handed particles differently should imply that fermions are massless as well. Table 2.1 only contains five particles (photons, gluons, and neutrinos) that are predicted to be truly massless (although these predictions are active areas of experimental probing). The Higgs boson breaking the overall electroweak symmetry provides this necessary mechanism (see Section 2.2).

The strong force, communicated via gluons, affects particles with color charge. They themselves, like quarks, carry color charge (red, green, or blue). Though quarks are fundamental they cannot exist alone in free space as the strong force, as described by Quantum Chromodynamics (QCD), becomes non-perturbative at low energies. This leads to “confinement”, and implies that

observable particles have no net color charge. Since quarks alone have nonzero color charge, they must form composite particles with other quarks and the help of gluons to create color neutral states. These composite particles, called hadrons, generally feature two or three valence quarks (which determine the hadron's identity) and a sea of light quarks and gluons. As two particles with color charge move away from each other, the force between them increases, unlike the EM or weak forces. The stored energy becomes high enough to generate a quark-antiquark pair from the vacuum, which creates color neutral hadrons with the original particles through a process called hadronization. In a collider setting, a single quark or gluon will hadronize into a spray of particles in the detector called a jet (see Section 4.2.1). Leptons contain no color charge so isolated, or unconfined, leptons are observed.

For every particle (see Table 2.1), there is also an anti-particle of identical mass but opposite electromagnetic charge. Electrically neutral bosons (photons and  $Z$  bosons) can be their own anti-particles.

## 2.2. The Higgs Boson

The Higgs boson is unique within the scope of the SM - it is the only fundamental scalar boson. Its existence was theorized in 1964 independently by Peter Higgs [36, 37]; François Englert and Robert Brout [38]; and Gerald Guralnik, Carl Hagen, and Tim Kibble [39]. Higgs and Englert were awarded the 2013 Nobel Prize in Physics [40] for this work, after a Higgs-like scalar particle with mass of 125 GeV was discovered at the LHC (see Section 3.1) in 2012 [41, 42].



|                               | Particle   | Spin          | EM<br>Charge   | Color<br>Charge? | $Y_W$<br>(LH, RH)            | Mass                                |
|-------------------------------|------------|---------------|----------------|------------------|------------------------------|-------------------------------------|
| Quarks                        |            |               |                |                  |                              |                                     |
| 1 <sup>st</sup><br>generation | u          | $\frac{1}{2}$ | $+\frac{2}{3}$ | ✓                | $+\frac{1}{3}, +\frac{4}{3}$ | $2.16^{+0.49}_{-0.26}$ MeV          |
|                               | d          | $\frac{1}{2}$ | $-\frac{1}{3}$ | ✓                | $+\frac{1}{3}, -\frac{2}{3}$ | $4.67^{+0.48}_{-0.17}$ MeV          |
| 2 <sup>nd</sup><br>generation | c          | $\frac{1}{2}$ | $+\frac{2}{3}$ | ✓                | $+\frac{1}{3}, +\frac{4}{3}$ | $1.27 \pm 0.02$ GeV                 |
|                               | s          | $\frac{1}{2}$ | $-\frac{1}{3}$ | ✓                | $+\frac{1}{3}, -\frac{2}{3}$ | $93^{+11}_{-5}$ MeV                 |
| 3 <sup>rd</sup><br>generation | t          | $\frac{1}{2}$ | $+\frac{2}{3}$ | ✓                | $+\frac{1}{3}, +\frac{4}{3}$ | $172.76 \pm 0.30$ GeV               |
|                               | b          | $\frac{1}{2}$ | $-\frac{1}{3}$ | ✓                | $+\frac{1}{3}, -\frac{2}{3}$ | $4.18^{+0.03}_{-0.02}$ GeV          |
| Leptons                       |            |               |                |                  |                              |                                     |
| 1 <sup>st</sup><br>generation | e          | $\frac{1}{2}$ | -1             | ×                | -1, 0                        | $0.5109989461 \pm 0.0000000031$ MeV |
|                               | $\nu_e$    | $\frac{1}{2}$ | 0              | ×                | -1, -2                       | < 1.1 eV                            |
| 2 <sup>nd</sup><br>generation | $\mu$      | $\frac{1}{2}$ | -1             | ×                | -1, 0                        | $105.6583745 \pm 0.0000024$ MeV     |
|                               | $\nu_\mu$  | $\frac{1}{2}$ | 0              | ×                | -1, -2                       | < 1.1 eV                            |
| 3 <sup>rd</sup><br>generation | $\tau$     | $\frac{1}{2}$ | -1             | ×                | -1, 0                        | $1776.86 \pm 0.12$ MeV              |
|                               | $\nu_\tau$ | $\frac{1}{2}$ | 0              | ×                | -1, -2                       | < 1.1 eV                            |
| Bosons                        |            |               |                |                  |                              |                                     |
| Vector                        | $\gamma$   | 1             | 0              | ×                | 0                            | < $10^{-18}$ eV                     |
|                               | $g$        | 1             | 0              | ✓                | ×                            | 0                                   |
|                               | $W$        | 1             | $\pm$          | ×                | 0                            | $80.379 \pm 0.012$ GeV              |
|                               | $Z$        | 1             | 0              | ×                | 0                            | $91.1876 \pm 0.0021$ GeV            |
| Scalar                        | H          | 0             | 0              | ×                | +1                           | $125.10 \pm 0.14$ GeV               |

TABLE 2.1. SM particles with their quantum numbers and masses [34].  $Y_W$  lists left- (LH) and right-handed (RH) values.

### 2.2.1. Higgs Mechanism

The Higgs field is a critical element of the SM. It provides a means of generating the mass of fermions and gauge bosons by spontaneously breaking electroweak symmetry. The Higgs mechanism mixes the four massless gauge fields ( $W_1$ ,  $W_2$ ,  $W_3$ , and  $B$ ) into the physical  $\gamma$ ,  $W^\pm$ , and  $Z$  bosons. The  $W$  and  $Z$  bosons acquire a mass while the photon remains massless (as required by the remaining unbroken U(1) symmetry of electromagnetism).

The electroweak component of the Higgs contribution to the SM Lagrangian<sup>2</sup>,

$$\mathcal{L}_H \ni (D^\nu \phi)^\dagger (D_\nu \phi) - V(\phi) = (D^\nu \phi)^\dagger (D_\nu \phi) - \mu^2 |\phi^\dagger \phi| + \lambda |\phi^\dagger \phi|^2, \quad (2.1)$$

introduces a complex scalar Higgs doublet field,

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}. \quad (2.2)$$

The covariant derivative

$$D_\mu = \partial_\mu + \frac{ig'}{2} B_\mu Y + \frac{ig}{2} \tau \cdot \mathbf{W}_\mu$$

involves the massless hypercharge gauge field ( $B_\mu$ ) and three massless weak gauge bosons ( $W_\mu^1, W_\mu^2, W_\mu^3$ ) where  $\tau$  are proportional to Pauli matrices and  $g$  and  $g'$  are weak and hypercharge coupling constants, respectively. If  $\mu^2 < 0$  (see Figure 2.2), there are infinitely many potential minima in a circle around the origin at a radius of

$$|\phi|^2 = -\frac{\mu^2}{2\lambda} \equiv \frac{v^2}{2},$$

where  $v$  is the Higgs vacuum expectation value or VEV.

Without loss of generality,  $\phi^+ = 0$  and  $\phi^0 = v/\sqrt{2}$  are chosen, breaking  $SU(2)$  and  $U(1)_Y$  symmetries. Perturbations can be parameterized [2] with the scalar

---

<sup>2</sup>See [43, 44] for technical details regarding quantum field theories and for a derivation of the SM Lagrangian.

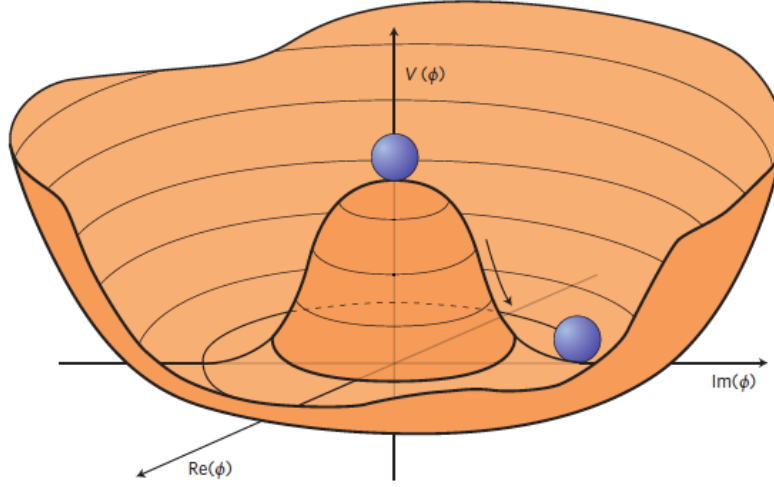


FIGURE 2.2. The Higgs potential from Eq. 2.1 with  $\mu^2 < 0$  [2]

fields  $h$  and  $\eta^a$  like

$$\phi = \exp\left(\frac{i\eta \cdot \tau}{2v}\right) \begin{pmatrix} 0 \\ (v + h)/\sqrt{2} \end{pmatrix}. \quad (2.3)$$

Each scalar field has an associated scalar particle, where  $m_{\eta^a}^2 = 0$  and  $m_h^2 = 2|\mu|^2$ .

These three massless particles are Goldstone bosons. Each is related to a broken symmetry. Gauge transformations allow rotation such that

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h \end{pmatrix}. \quad (2.4)$$

Substituting Eq. 2.4 into Eq. 2.1 reveals mass terms for three of the four originally massless boson fields, where the Goldstone bosons were “eaten” in order to provide mass to the boson fields. Linear combinations of the gauge fields define

mass eigenstates for the four physical bosons,

$$m_W = \frac{gv}{2} , \quad m_Z = \frac{v}{2} \sqrt{g^2 + g'^2}, \quad m_\gamma = 0 ,$$

where the photon remains massless. The added scalar Higgs field is generated a mass as well,  $m_h = 2|\mu|^2$ . Like all Lagrangian parameters,  $\mu$  is not predicted by the theory and must be experimentally determined.

The Lagrangian also reveals couplings. The gauge bosons couple to the Higgs field proportionally to their mass:

$$\mathcal{L} \ni \frac{vg^2}{4} (hW^{\mu-}W_{\mu}^+) + \frac{g^2 + g'^2}{4} (hZ_{\mu}Z^{\mu}) = \frac{2m_W^2}{v} (hW^{\mu-}W_{\mu}^+) + \frac{2m_Z^2}{v} (hZ_{\mu}Z^{\mu}) .$$

The gauge bosons also couple directly to fermions, whose masses are also derived from the Higgs field.

$W^3$  and  $B$  mix to form the orthogonal  $Z$  boson and  $\gamma$  mass eigenstates with the weak mixing angle,

$$\tan \theta_W = \frac{g'}{g} .$$

This also relates the  $W$  and  $Z$  boson masses, where

$$m_W = \frac{m_Z}{\sqrt{1 + g^2/g'^2}} = m_Z \cos \theta_W ,$$

and is directly related to the coupling constant for electromagnetism,  $e = g \sin \theta_W$ .

These parameters have been extracted experimentally :  $v = 246.22$  GeV,

$\sin^2 \theta_W = 0.22337 \pm 0.00010$  , and  $m_h = 125.10 \pm 0.14$  GeV [34].

### 2.2.2. Production and Decay

Higgs couplings to fermions and gauge bosons are important for Higgs production in a collider setting, where accelerated fermions interact. Particles with larger masses are generally the most relevant, as the coupling strength to the Higgs is proportional to particle mass. The Large Hadron Collider (see Section 3.1) collides protons - hadrons composed of three valence quarks and a sea of additional gluons and quark-antiquark pairs. Interactions between either these quarks or gluons can produce Higgs bosons (see Figure 2.3 a)).

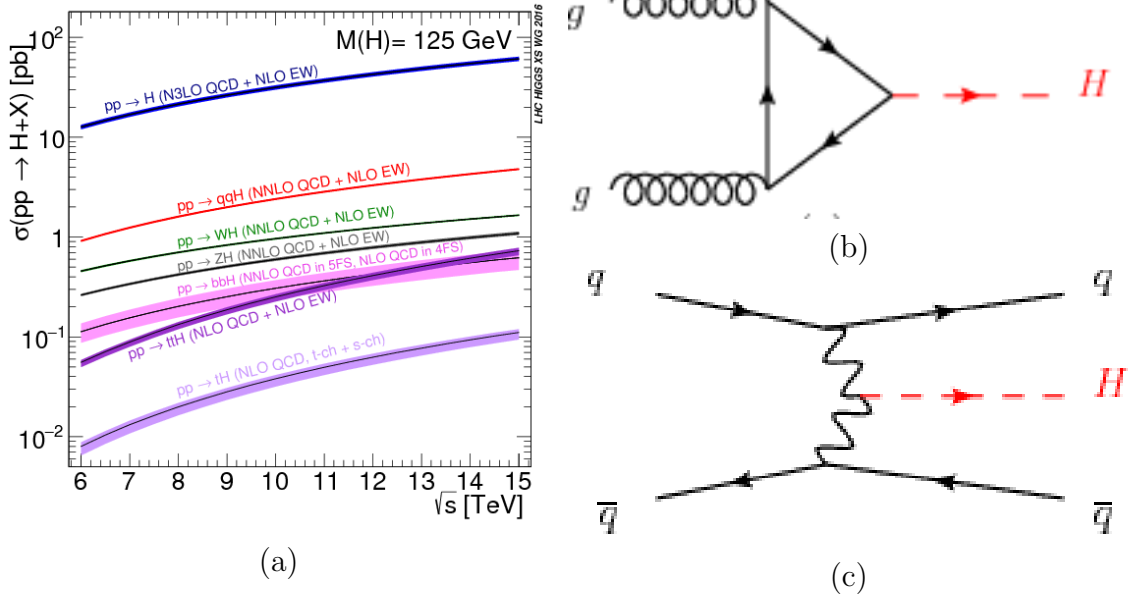


FIGURE 2.3. a) Higgs production from  $pp$  collisions [3], b) the dominant production mechanism, gluon-gluon fusion, and c) the subdominant production, vector boson fusion [4]

The probability with which a process occurs is determined by its cross section ( $\sigma$ ). The Higgs production process with the largest cross section is gluon-gluon fusion ( $pp \rightarrow H$  in Figure 2.3 a), see Figure 2.3 b)), followed by vector

boson fusion ( $pp \rightarrow qqH$  in Figure 2.3, also called VBF or VBS for vector boson scattering), see Figure 2.3 c)).

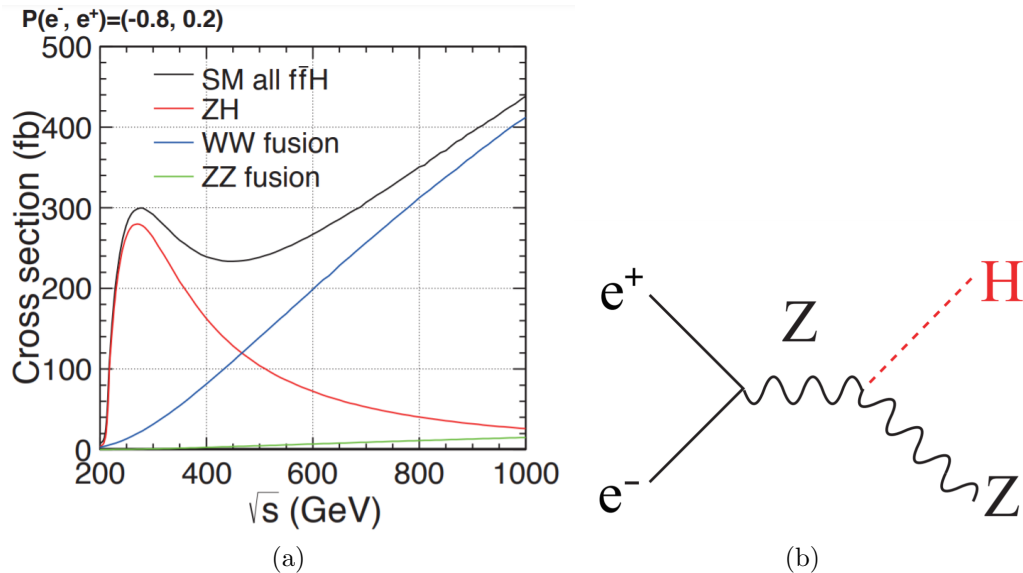


FIGURE 2.4. a) Higgs production from  $e^+e^-$  collisions and b) the dominant production mechanism at  $\sqrt{s} = 250$  GeV, Higgsstrahlung [5]

The dominant Higgs production channel in an  $e^+e^-$  collider with  $\sqrt{s} = 250$  GeV (see Section 7.4) is Higgsstrahlung (see Figure 2.4), where the Higgs boson is produced in association with a  $Z$  boson. This also occurs with  $pp$  collisions but at a lower rate (dark green  $pp \rightarrow ZH$  in Figure 2.3 a)).

Higgs bosons have a theorized total width<sup>3</sup> of  $\sim 4$  MeV with a lifetime of  $\sim 1.6 \times 10^{-22}$  s. This means, when produced in colliders, they decay before they can be directly measured and therefore are indirectly studied through their decay products. The invariant mass of decay products is a spectrum that follows a Breit-Wigner distribution defined by the parent particle mass and full decay width  $\Gamma$ . The individual contribution from specific decay channels account for some fraction of this total width (see Figure 2.5 for a 125 GeV Higgs). This partial

<sup>3</sup>The best measured experimental value is  $\Gamma < 0.013$  GeV [34].

width associated with one decay channel is some fraction of  $\Gamma$ , also called the branching ratio ( $\mathcal{B}$ ). The branching ratio is the fraction of time a particle decays via the particular decay channel.

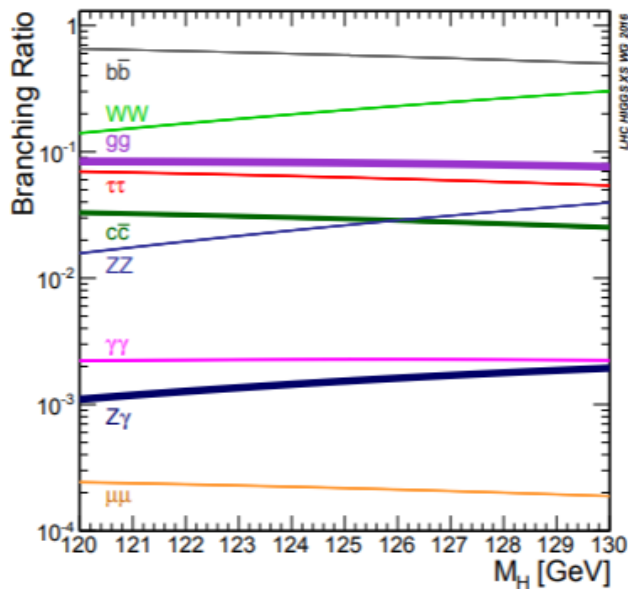


FIGURE 2.5. Higgs boson decay modes in the SM ( $m_h = 125.10 \pm 0.14$  GeV) [3]

### 2.3. Dark Matter

Cosmological and astrophysical measurements are used to deduce the existence of dark matter (DM) - a particle (or sector of particles) that interacts gravitationally and comprises the majority of matter in the universe [45, 46, 47]. There is no explanation in the SM, but extensions of the SM allow for such new particles. An array of experimental techniques is used to search for DM through production at colliders, with astrophysical observations, and in specifically-designed direct detection experiments (see Figure 2.6).

Dark matter models are varied, producing DM candidates with a range of masses and couplings (see Figure 2.7). Weakly-interacting candidates (weakly

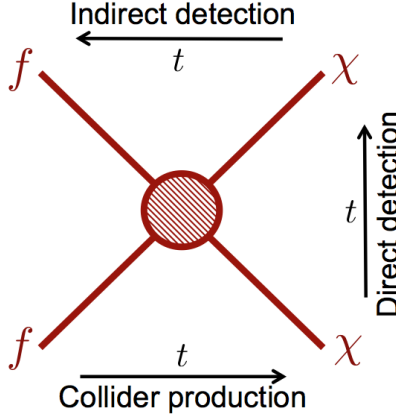


FIGURE 2.6. DM coupling to the SM through a portal (shaded region) [6]

interactive massive particles, or WIMPs) with GeV-TeV-scale masses have been extensively experimentally considered, but additional model are also valid provided that the candidate particle can reproduce the experimentally observed relic density. Most models also involve couplings between DM and SM particles.

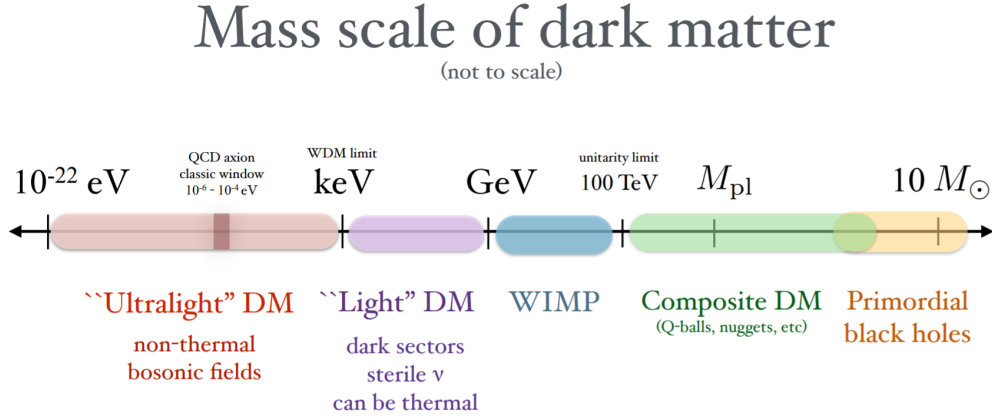


FIGURE 2.7. Spectrum of allowed DM masses and models [7]

A natural avenue for DM additions is through SM-DM portals, or new, low-dimension, singlet operators [7] (the shaded region of Figure 2.6). There are three such portals in the SM with dimension less than four:



- **Kinetic mixing portal** - DM candidates couple to the SM with a new ‘dark photon’ which couples to the SM photon with some mixing angle. This results from the addition of a new U(1) gauge symmetry and dark photon vector field which interacts with the hypercharge field strength tensor.
- **Neutrino portal** - DM candidates couple to the SM by introducing right-handed sterile neutrinos that couple to SM left-handed neutrinos with some mixing angle [48]. Alternatively, BSM decay products of the sterile neutrinos could be DM candidates.
- **Higgs portal** - DM candidates couple to the SM through an added field which interacts only with the Higgs field. This particle itself could be the DM candidate, or could be a mediator between the SM and DM.

The Higgs portal is the simplest theoretical method to introduce a new DM candidate  $\chi$ , which couples only to the SM Higgs [49]. The addition carries only two<sup>4</sup> new parameters - the coupling strength  $\lambda_\chi$  and mass  $m_\chi$ . For scalar  $S$ , fermionic  $F$ , and vector  $V$  fields (where  $\chi = S, F, \text{ or } V$ ), the Lagrangian would be amended by

$$\begin{aligned}\Delta\mathcal{L}_S &= -\frac{1}{2}m_S^2 S^2 - \frac{1}{4}\lambda_S\phi^\dagger\phi S^2, \\ \Delta\mathcal{L}_F &= -\frac{1}{2}m_F\bar{F}F - \frac{1}{4}\frac{\lambda_F}{\Lambda}\phi^\dagger\phi\bar{F}F, \\ \Delta\mathcal{L}_V &= \frac{1}{2}m_V^2 V_\mu V^\mu + \frac{1}{4}\lambda_V\phi^\dagger\phi V_\mu V^\mu.\end{aligned}$$

The new particle can be stable, as DM must be, due to a model-dependent  $Z_2$  parity [50]. The new  $\chi$  particle could also be a mediator between the SM and an

---

<sup>4</sup>A self interaction term of the form  $\lambda\chi^4$  is also introduced with the additional parameter  $\lambda$ , though this term is not relevant to discussion.

additional new particle  $\zeta$  that serves as the DM candidate. This  $\zeta$  is not required to be a scalar.

In a collider setting, this process can be searched for whenever a Higgs boson is produced. Any DM candidate is expected to interact rarely with the SM, most likely escaping the detector without any direct interaction. This carries energy away from the interaction, which can be measured indirectly (see Section 4.2.4). Since they are not directly detectable, these particles are “invisible” to the detector. Therefore, the signal of  $h \rightarrow \chi\chi$  looks like  $h \rightarrow \text{invisible}$ .

The process of  $h \rightarrow \text{inv.}$  is allowed in the SM by  $h \rightarrow ZZ^* \rightarrow \nu\nu\nu\nu$  with a branching ratio of  $\mathcal{B}_{h \rightarrow \text{inv}} \sim 1.13 \times 10^{-3}$ . The addition of  $h \rightarrow \chi\chi$  would increase  $\mathcal{B}_{h \rightarrow \text{inv}}$  above the SM-expected value, so the best way to detect DM through a Higgs portal model is by measuring  $\mathcal{B}_{h \rightarrow \text{inv}}$  [51, 52, 53]. If the DM mass  $M^2 = m_\chi^2 + \lambda_\chi v^2/2 < m_h^2/4$  (the kinematically allowed decay of an on-shell Higgs), then the measured decay width will be amended dependent upon the DM mass and Higgs coupling:

$$\begin{aligned}\Gamma_{h \rightarrow SS}^{\text{inv.}} &= \frac{\lambda_S^2 v^2 \beta_S}{64\pi m_h} , \\ \Gamma_{h \rightarrow FF}^{\text{inv.}} &= \frac{\lambda_F^2 v^2 m_h \beta_F^3}{32\pi \Lambda^2} , \\ \Gamma_{h \rightarrow VV}^{\text{inv.}} &= \frac{\lambda_V^2 v^2 m_h^3 \beta_V}{256\pi M_V^4} \left( 1 - 4 \frac{M_V^2}{m_h^2} + 12 \frac{M_V^4}{m_h^4} \right) ,\end{aligned}$$

where  $\beta_{S,F,V} = \sqrt{1 - 4M_{S,F,V}^2/m_h^2}$ . These additions could impact the expected branching ratio by  $\mathcal{O}(10\%)$  [54].

These searches compliment direct DM detection experiments [55, 56, 57, 58] and indirect astrophysical DM measurements [59]. The findings from all methods of searching in combination help to search a wide swatch of phase space.

## CHAPTER III

### LHC AND ATLAS

#### 3.1. The Large Hadron Collider

The Large Hadron Collider (LHC) [60], the world's largest particle accelerator, is a superconducting accelerator that accelerates beams of protons and heavier ions (such as lead) to speeds approaching the speed of light. It is operated and maintained through CERN, based in Meyrin, Switzerland and came online to begin taking data in 2009. Superconducting magnets are used to direct the particle beams around the subterranean 27 km circular tunnel which spans the Franco-Swiss border. The beams collide at four points around the ring, at which the main LHC experiments (ATLAS [11], CMS [61], LHCb [62], and ALICE [63]) have built detectors in order to study the particles that result from the interactions.

The circular design boasts many performance features. For example, the beams are directed around the circular ring and collided many times before they become unusable. This allows for large amounts of data to be collected. It is also a safe choice for hadron collisions, as the hadrons used (protons) are relatively heavy. Particles subjected to changes of direction as they are accelerated undergo synchrotron radiation and lose energy as a result. The intensity of this phenomenon scales as  $m^{-4}$ , so protons lose little energy when compared to lighter particles such as electrons.

In addition to proton-proton physics, the LHC also supports a heavy-ion physics program in which beams of lead nuclei interact inside the detectors. Consideration of heavy-ion physics is beyond the scope of this work.

### 3.1.1. Accelerator Complex

The LHC requires a specific beam structure and energy before collisions, and a large accelerator complex is required to achieve these goals. Protons are gathered from gaseous hydrogen atoms, which are comprised of one proton and one electron. These atoms pass through an electric field which separates the protons from the electrons allowing only the protons to be selected. The protons are then passed through multiple accelerators gaining more energy at each step (see Figure 3.1).

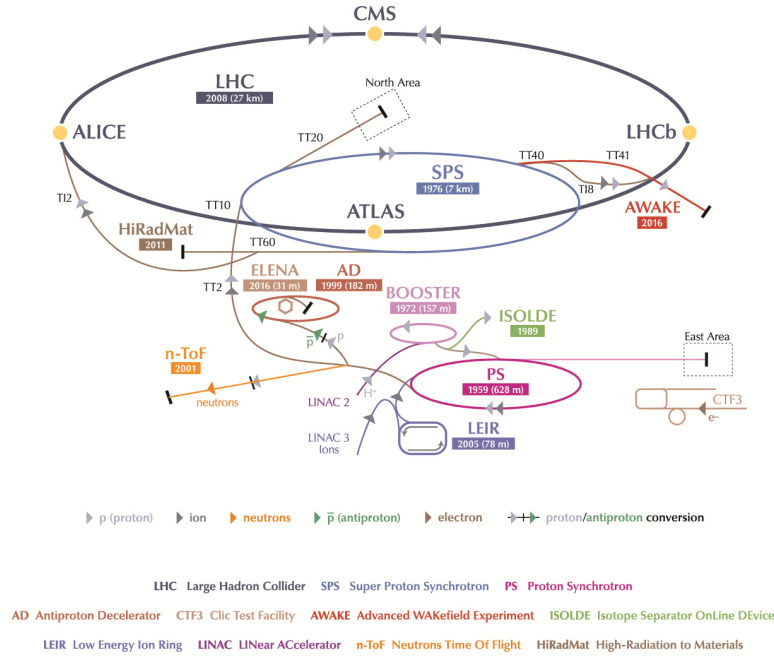


FIGURE 3.1. The accelerator complex at CERN [8]

Protons first encounter the 36 m long linear accelerator (Linac) 2, which uses radiofrequency (RF) cavities and cylindrical conductors. The positively charged protons are repulsed by a positively charged cavity, and attracted to a

negatively charged cavity. The cavities oscillate between positive and negative charges, always pushing the protons toward negatively charged cavities and away from positively charged ones. Cavities are spaced such that any particle reaching the cavity before the rest of the bunch experiences a smaller acceleration than the group, and likewise late-arriving particles experience greater acceleration so that they can rejoin the bunch. This creates bunches of protons with similar momenta that experience the RF together. With each push, the protons accelerate faster until they gain 50 MeV upon exiting Linac 2 [64]. Linac 2 served CERN for 40 years until its successful decommissioning in 2018 following Run 2 [65].

The beam is then sent to the Proton Synchrotron Booster (PSB), which uses RF to accelerate the protons up to 1.4 GeV before injection into the Proton Synchrotron (PS). This circular accelerator was first commissioned in 1959 and served as CERN's flagship synchrotron. The PSB's injection rate helps to set the bunch structure of the resulting LHC physics beam. Rather than filling the PS with the maximum number of bunches possible, bunches are injected from the PSB in two batches 1.2 seconds apart creating a structured beam of 72 bunches of more than 100 billion ( $10^{11}$ ) protons each with 25 ns bunch spacing [66]. The PS accelerates this beam to 25 GeV and injects 3 or 4 batches to the Super Proton Synchrotron (SPS). Like the PS, the SPS was originally built as a stand-alone accelerator that has been re-purposed as an LHC injector. The SPS accelerates the proton beam to 450 GeV before it is split and injected into the LHC ring itself - one in each direction. At this point, the beam consists of 2808 bunches. The LHC uses superconducting magnets to further accelerate the beam up to its collision energy of 6.5 TeV. In total, it takes around three minutes to fill the LHC ring.

### 3.1.2. Magnets

The superconducting magnets used to steer and focus the LHC beams are the most powerful in the world. They were designed with the best industrially-scalable technology at the time and create a uniform magnetic field of 8.3 T. A cooling system is necessary to keep the magnets at cryogenic temperatures below 2K in order to maintain superconductivity.

1232 dipole magnets are used for steering the beams around the circular LHC tunnel. Quadrupole magnets are used to focus the beam before it enters a detector so that particle interactions are maximized. The beam is squeezed from 0.2 mm to 16 nm across inside the detectors [67].

### 3.1.3. Luminosity and Pileup

The bunch structure of the LHC beam was designed such to optimize the amount of proton–proton interactions. The rate at which data are collected is called the luminosity, and when integrated over time it quantifies the amount of data collected. A large luminosity enriches the LHC’s physics program, as many rare and exotic processes have very small cross sections. Although increased collision energy increases the probability of these processes to occur, a large data set further increases the chances of the observation and measurement of these rare processes.

If the LHC used beams with bunches containing one single proton each, the bunch luminosity would be dependent on the probability of a collision,  $\mathcal{L}_b = 1/\sigma$ . The bunch structure of the LHC beams optimizes this value by increasing the number of protons in each bunch which may interact. If the number of protons

in each bunch is  $N$  (a value around  $10^{11}$ ) then the bunch luminosity increases to

$$\mathcal{L}_b = \frac{N_1 N_2}{\sigma} . \quad (3.1)$$

The circular machine also increases total luminosity, as beams traverse around the ring and collide multiple times. Therefore, Eq. 3.1 can be rewritten as

$$\mathcal{L}_b = \frac{\mu f}{\sigma} , \quad (3.2)$$

where  $f$  is the LHC revolution frequency of 11246 Hz and  $\mu$  is the average number of inelastic interactions per bunch crossing.

The total instantaneous luminosity is found by summing the luminosity of  $N_b$  bunches, which can be represented by scaling the average bunch luminosity by  $N_b$ , or

$$\mathcal{L} = N_b \langle \mathcal{L}_b \rangle = N_b \frac{\langle \mu \rangle f}{\sigma} , \quad (3.3)$$

where quantities in angled brackets represent the mean value [68]. Experiments use luminosity detectors and algorithms to measure this instantaneous luminosity during data-taking. During the most recent period of data-taking, the LHC delivered a peak instantaneous luminosity of  $\mathcal{L} = 2.10 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ . This is more than a factor of two excess over the design luminosity of  $\mathcal{L} = 10^{34} \text{ cm}^{-2}\text{s}^{-1}$  [69]. The integrated luminosity then considers the cumulative amount of data collected over an LHC run period.

The average number of interactions per bunch crossing,  $\langle \mu \rangle$  from Eq. 3.3, is large and referred to as pileup (PU). This is a complicating factor when trying to identify physics events to study. The most common LHC process is inelastic scattering of protons from colliding bunches. These interactions produce soft,

or low-momentum, QCD objects seen in the detectors. The most interesting physics questions can be answered by considering hard scatter interactions where the quark and gluon constituents of the protons interact, but these processes are far more rare. The majority of interactions in each bunch crossing are soft, uninteresting, inelastic scattering that constitute this in-time pileup and pollute the signal of the hard scatter event. This is illustrated in Figure 3.2 with a bunch crossing interaction containing 25 reconstructed vertices, or 25 independent interactions. The signatures left from particles from these separate events are color-coded to represent their originating vertex. This illustrates the amount of backgrounds occurring in the midst of rarer physics processes, although the figure shows lower pileup than the average LHC run.

Out-of-time pileup can occur when interactions from one bunch crossing are measured with the following bunch crossing due to the time resolution of some subsystems exceeding the time interval between bunch crossings. Pileup can be mitigated by applying minimum momentum requirements on physics objects to exclude the soft pileup objects.

The LHC has been collecting physics data since the machine was turned on in 2009. These initial beams were injected with 450 GeV and interacted with center of mass energy  $\sqrt{s} = 900$  GeV. The first large campaign of data-taking, Run 1, began in 2010 and lasted through 2012. The LHC magnets were now used to accelerate the beams to a center of mass collision energy of  $\sqrt{s} = 7$  TeV in 2011 and  $\sqrt{s} = 8$  TeV in 2012 with bunches of 50 ns spacing. The machine and many detectors then underwent a period of upgrades to begin Run 2, the most recent data-taking period which lasted from 2015-2018. In this time, the LHC operated at  $\sqrt{s} = 13$  TeV with 25 ns bunch spacing. Data was collected in these runs (see



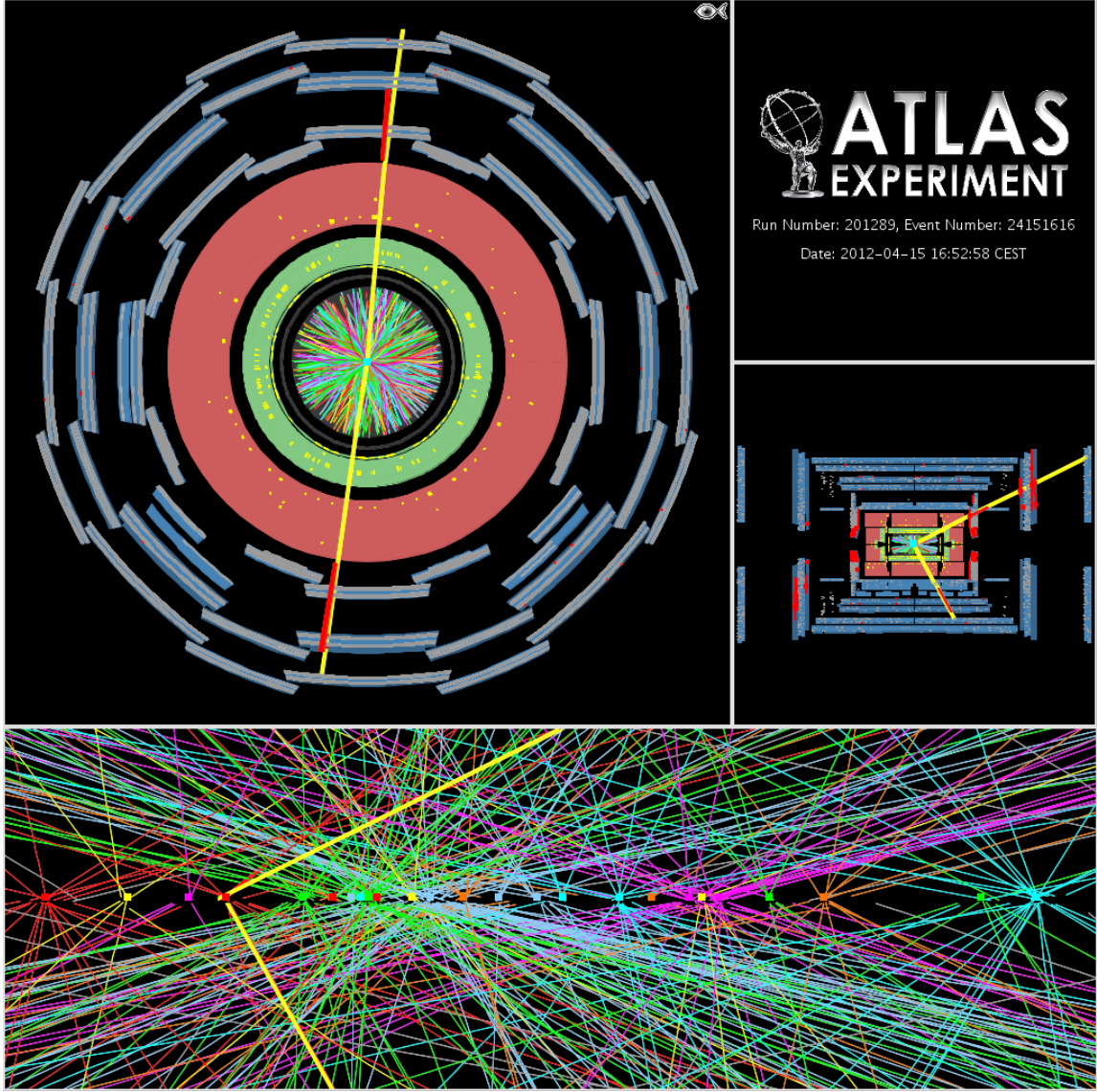


FIGURE 3.2. A Run 1 candidate  $Z \rightarrow \mu\mu$  event with 25 reconstructed vertices [9]

Table 3.1) is expressed in units of inverse barns ( $1 \text{ b} = 10^{-28} \text{ m}^2$ ). The cross section for physics processes is reported in units of barns, so the number of events of a given physics process can be found by multiplying its cross section by the total integrated luminosity of a data set. It also shows the average pileup ( $\langle \mu \rangle$ ) in each run. The four years of Run 2 each had very different pileup environments (see Figure 3.3) leading to a large overall value. The increased pileup in Run 2 often

requires that only detector objects associated with a defined hard scatter vertex be considered. Additional discriminatory variables such as the Jet Vertex Tagger (see Section 4.2.1.2) aim to flag and exclude pileup contributions to physics analyses.

| Run | Year      | Collision Energy [TeV] | Data collected [ $\text{fb}^{-1}$ ] | $\langle\mu\rangle$ |
|-----|-----------|------------------------|-------------------------------------|---------------------|
| 1   | 2011      | 7                      | 4.7                                 | 9.1                 |
| 1   | 2012      | 8                      | 20.3                                | 20.7                |
| 2   | 2015-2018 | 13                     | 139                                 | 33.7                |

TABLE 3.1. LHC data set nomenclature and parameters

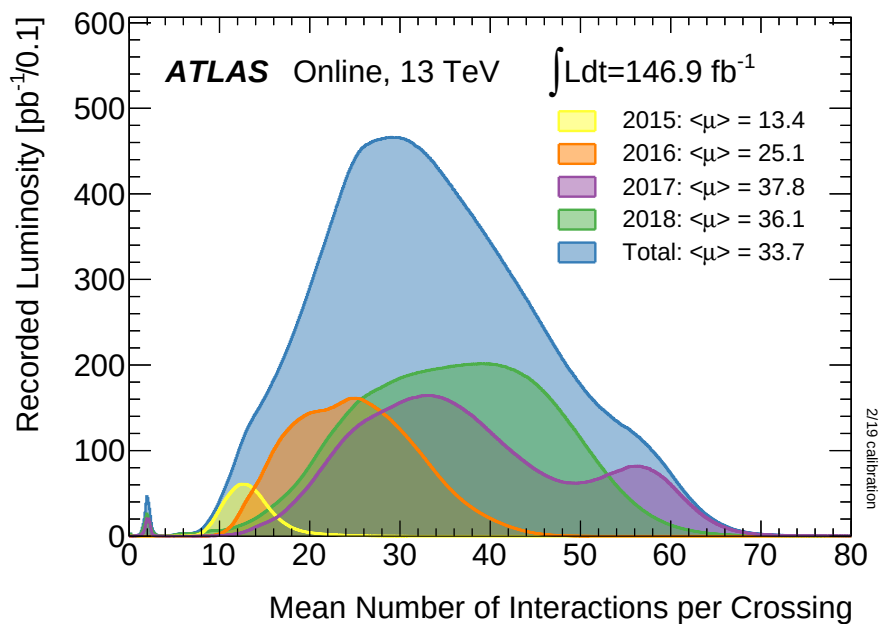


FIGURE 3.3. Pileup conditions in Run 2 [10]

### 3.2. ATLAS

The ATLAS detector [11] is a general-purpose particle detector with nearly  $4\pi$  solid angle coverage around interaction Point 1, designed to be sensitive to a large range of physics signatures. At 44 m in length and weighting nearly 7000 tonnes, it is the larger of two such detectors at the LHC - CMS sits across the ring

at Point 6 and serves comparable physics goals. The detector design features large toroidal magnets, from which the detector and collaboration take their names (**A**Toroidal **L**H**C** **A**pparatu**S**).

Subdetectors are arranged concentrically around the interaction point and used to measure different types of incident particles (see Figure 3.4). These detectors are sensitive to different characteristics of the final state particles that result from the interacting proton beams. The signatures of each subdetector system are used to identify particles and reconstruct the event that occurred at the interaction point. The symmetric cylindrical portion around the interaction point is called the central barrel region, and contains (radially outward from the interaction point) the inner detector (Section 3.2.2), electromagnetic and hadronic calorimeters (Section 3.2.3), and a muon system (Section 3.2.4). These detectors are bathed in magnetic fields provided by a system of magnets (Section 3.2.5). The area very close to the beamline requires specialized detectors (Section 3.2.6). The data collected by all these subdetector systems are recorded and coordinated via a trigger system (Section 3.2.7).

The ATLAS experiment requires an international collaboration of scientists and engineers for maintenance and operation of the detector and to analyze and interpret the data. Each detector subsystem, including the trigger, is individually monitored during data-taking by trained shifters in the ATLAS Control Room. Monitoring to ensure proper operation and high data quality is done around the clock while the beam is running. Problems noted in the Control Room are relayed to on-call experts for the subsystems in order to identify and solve any runtime problems in real time. I monitored the performance of the trigger system for more than 135 hours at the end of Run 2 in the ATLAS Control Room.

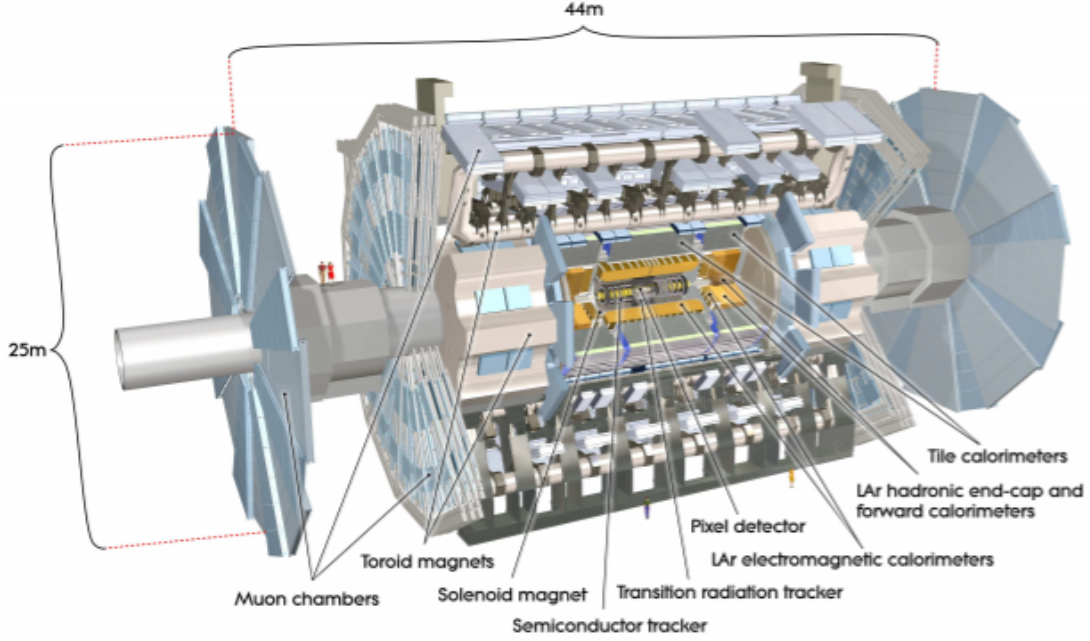


FIGURE 3.4. The ATLAS detector [11]

### 3.2.1. Coordinates and Variables

ATLAS uses a right-handed coordinate system with the origin set at the interaction point (IP). The beams travel along the longitudinal  $z$  direction, leaving the  $xy$  plane transverse to the beamline with the positive  $x$  axis point toward the center of the LHC ring and the positive  $y$  axis pointing upwards toward the surface (see Figure 3.5 a)).

Since the detector itself is designed as a nearly symmetric cylinder, cylindrical coordinates are more commonly used in analyses (and for the remainder of this work). An azimuthal angle  $\phi$  and polar angle  $\theta$  are defined such that  $\phi$  measures around the beamline and  $\theta$  measures the distance from the  $z$  direction (see Figure 3.5 b)).

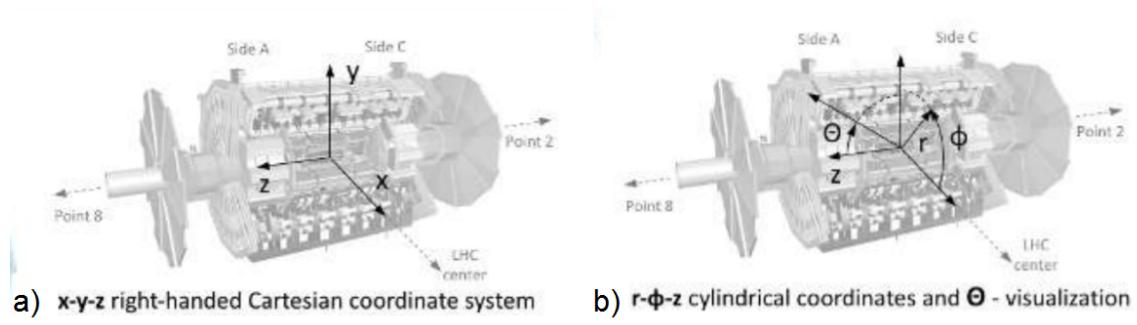


FIGURE 3.5. Rectangular and cylindrical coordinates used with ATLAS [12]

Rapidity ( $y$ ) is constructed using measured characteristics of particles like

$$y \equiv \frac{1}{2} \ln \left( \frac{E + p_z}{E - p_z} \right) \approx -\ln \tan \left( \frac{\theta}{2} \right) \equiv \eta , \quad (3.4)$$

where  $E$  is the particle's energy and  $p_z$  is its longitudinal momentum in the  $z$  direction<sup>1</sup>. In the limit of  $p \gg m$  where  $m$  is the particles mass, rapidity can be approximated as pseudorapidity ( $\eta$ ) [70]. The high energies probed by the LHC largely ensure that this approximation can be made. Pseudorapidity is preferred over rapidity as no measurement of energy or momentum is required. Pseudorapidity is a function of  $\theta$  but its use is preferred to  $\theta$  alone as massless particles are produced uniformly in  $\eta$  but not  $\theta$ , and differences in pseudorapidity are invariant under longitudinal Lorentz boosts. Pseudorapidity (see Figure 3.5 b)) is measured from the  $y$  axis toward the  $z$  axis where  $\eta = 0$  points along the positive  $y$  axis. Larger values of  $\eta$  are regions closer to the beam pipe, generally referred to as forward (as opposed to the central  $\eta$  region).

Angular distance between physics objects is measured as

$$\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} . \quad (3.5)$$

---

<sup>1</sup> $p_z = p \cos \theta$

### 3.2.2. Inner Detector

The first subdetector system a central particle produced at the IP encounters is the inner detector (ID), also referred to as the tracker. The ID is contained within a 2 T axial magnetic field (Section 3.2.5) that causes the path of charged particles to bend as they pass through. As the name suggests, the ID records the tracks of these passing charged particles. The intention of this system is to precisely measure the curvature of these charged tracks without substantially reducing the energy of particles to provide precision information of the particles location and momentum. This is also crucial for vertex identification (used, for example, in pileup rejection - Section 3.1.3).

As the ID is the detector component closest to the beamline, it experiences a high dose of radiation. Nearly 1000 particles interact with the ID upon every bunch crossing [11], so the precision of the system is paramount. To achieve this, the ID itself contains multiple subsystems [71] (see Figure 3.6). These are separated into central barrel regions and endcap elements, covering in total  $|\eta| < 2.5$ .

The central silicon pixel detector sits 50.5 mm away from the IP and uses three layers of highly granular silicon to achieve high precision in a dense track environment. Each pixel is  $50 \times 400 \mu\text{m}^2$ , so the full system contains more than 80 million channels that are individually read out. When a charged particle passes through the silicon, an electron-hole pair is created. The electron drifts due to an applied high-voltage to the read out electronics, and the locations of these “hits” are recorded and connected together to form a coherent track along the particles trajectory. The outer two central layers only extend to  $|\eta| < 1.7$ , so four endcap disks extend the coverage in  $1.7 < |\eta| < 2.5$ .

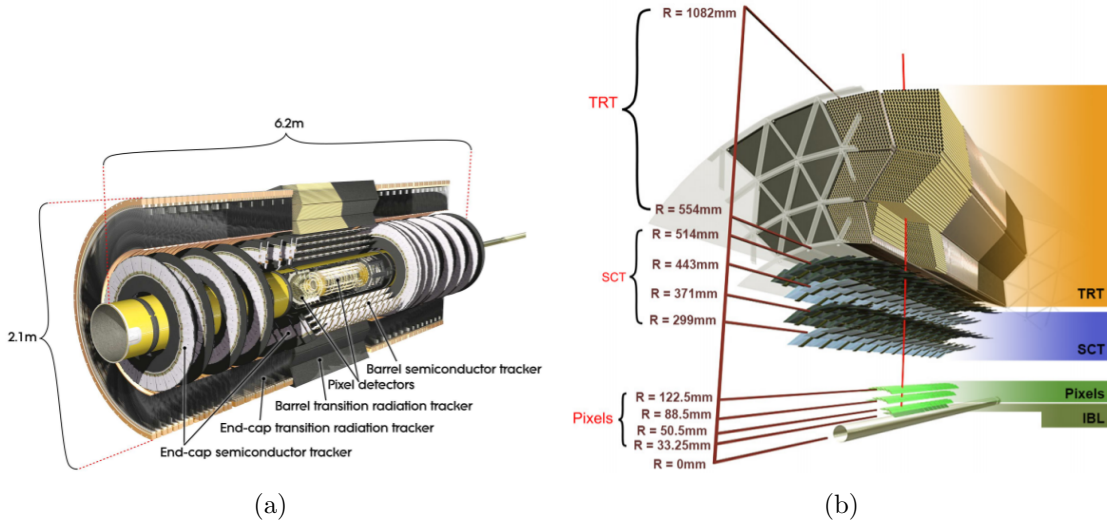


FIGURE 3.6. The ATLAS inner detector [11, 13]

During the scheduled year-end shutdown in 2016, the pixel detector was upgraded with a fourth central pixelated silicon layer even closer to the beam pipe. The Insertable B-Layer (IBL) [72] sits 33.25 mm from the beam pipe and adds 12 million additional pixels of  $50 \times 250 \mu\text{m}^2$  up to  $|\eta| < 2.58$ . Information from this additional fourth layer increases the ability of the pixel detector to accurately reconstruct both primary (from the bunch crossing interaction) and secondary (from particle decays, Section 4.2.1.3) vertices. The high rate of radiation also causes damage to the detector requiring that elements be replaced once the damage causes inefficiencies in measurement. The presence of the IBL shields the innermost pixel layer from such damage and ensures that resolution close to the beam pipe is retained.

The tracks are extended in the Semiconductor Tracker (SCT). The granularity of the SCT is slightly relaxed compared to the pixel detector as silicon strips are used rather than pixels. Each SCT module contains two layers of strips positioned orthogonal to each other such that one layer measures  $(R - \phi)$  position



and the other measures ( $z$ ). With this configuration, high location precision is retained with 6.3 million central readout channels and tracks can be distinguished as long as they are separated by at least  $200\text{ }\mu\text{m}$ . The barrel has four layers and extends to  $|\eta| < 1.4$ . The forward region from  $1.4 < |\eta| < 2.5$  is covered with nine endcap wheels.

Beyond the high-precision silicon layers is the Transition Radiation Tracker (TRT), which utilizes 300,000 straw tubes [73]. High-momentum particles do not bend dramatically as a result of the 2 T magnetic field. Their resolution benefits from the addition of more hits to allow for larger deflection. The barrel TRT drift tubes run along the direction of the beam pipe up to  $|\eta| < 0.7$ . Radial straws complete forward coverage, and both regions use 36 layers of straws, providing the opportunity for 36 additional hits to define a charged particle track. Each gas-filled 4 mm diameter Kapton tube has a gold-plated tungsten wire running through creating an anode. When a charged particle passes through the gas (an argon-xenon mixture), gas molecules are ionized and the freed electrons drift toward the wire anode to be amplified and read out. The argon-xenon gas mixture occupies the space between straws as well, which assists the identification of electrons that produce photons in the space between straws. This so-called transition radiation occurs when a charged particle passes across the interface of two media with different dielectric constants - here, the surrounding gas and straw tube boundary. Xenon is specifically chosen for this process, which allows for discrimination between transition radiation photons from electrons and charged pions as pion signals are relatively localized but transition radiation photons from electrons smear out the electron signature in the TRT [74] (see Figure 4.2).



Position resolution in the drift tubes is not as high as that in silicon ( $130\text{ }\mu\text{m}$  precision in the TRT), but this is compensated by the large number of layers of the TRT. Overall, momentum resolution is roughly equal in the silicon and straw tube systems.

### 3.2.3. Calorimetry

Beyond the ID lies a two-fold system of calorimeters designed to measure the energy of electrons, photons, and hadrons [75]. Where the ID is intended to measure the curvature of the trajectory of a charged particle with minimal impact, calorimeters aim to measure the energy of particles by inducing interactions between the detector and the incident particle. These interactions create daughter particles which themselves interact further with the calorimeter creating showers of particles of sequentially lower energy that spread through the detector. In this way, the incident particle itself is lost but its original energy can be measured through the consideration of its resulting shower.

The necessity to induce showering and also measure energy deposits of daughter particles leads to designs that utilize two different materials. Passive absorber layers are typically atomically dense material, which increases the probability of interactions with the incident particle thus encouraging showering. The showers are measured and read out with active readout layers. Calorimeters with this design are known as sampling calorimeters. As only the energy deposited in active layers can be read out, corrections must be made to account for the stochastic nature of this process. The way in which showers develop for different particles is known and can be used to identify the source of a shower. In order to fully capture the energy of an incident particle, the calorimeter must be deep

enough to contain the shower with little loss of daughter particles escaping out the back before detection - a component called leakage.

Electromagnetic calorimeters rely upon electromagnetic interactions of incident particles to induce showering, and are used mainly to measure the energy of incident photons and electrons. Pair production and brehmsstrahlung processes produce the measured electromagnetic showers (see Appendix B for more general details about shower development and measurement). Hadronic calorimeters measure incident charged and neutral hadrons that interact with the calorimeter through electromagnetic and strong processes. The ATLAS calorimeter system (see Figure 3.7) has an electromagnetic sampling calorimeter utilizing liquid argon [14] and a hadronic sampling calorimeter with scintillating tiles [76] which provide central coverage up to  $|\eta| < 3.2$  and full coverage to  $|\eta| < 4.9$  (see Section 3.2.6). The central barrel region of the calorimetry system contains the full expanse of the ID.

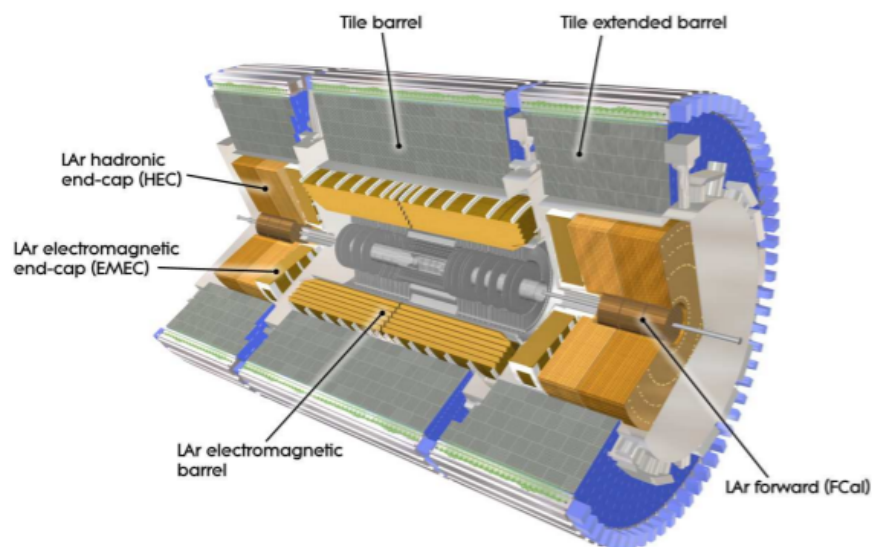


FIGURE 3.7. ATLAS calorimeter system with electromagnetic and hadronic calorimeters [11]

The liquid argon calorimeter (LAr) system uses active layers of highly atomically dense liquid argon and absorbing lead plates with readout electronics. Argon is also intrinsically radiation resistant and known to provide stable responses over time. Within the pseudorapidity range of  $|\eta| < 3.2$ , LAr is comprised of a barrel up to  $|\eta| < 1.475$  and two electromagnetic endcaps with thicknesses of 22 and 24  $X_0$ , where radiation length ( $X_0$ ) is approximately the distance an electron can travel before losing half of its energy to photon emission (see Appendix B). In both the barrel and endcap systems, the lead absorber is bent in an accordion style to ensure equal coverage around the entire  $\phi$  range with no azimuthal cracks (see Figure 3.8). In the barrel alone, 1024 absorber layers are used.

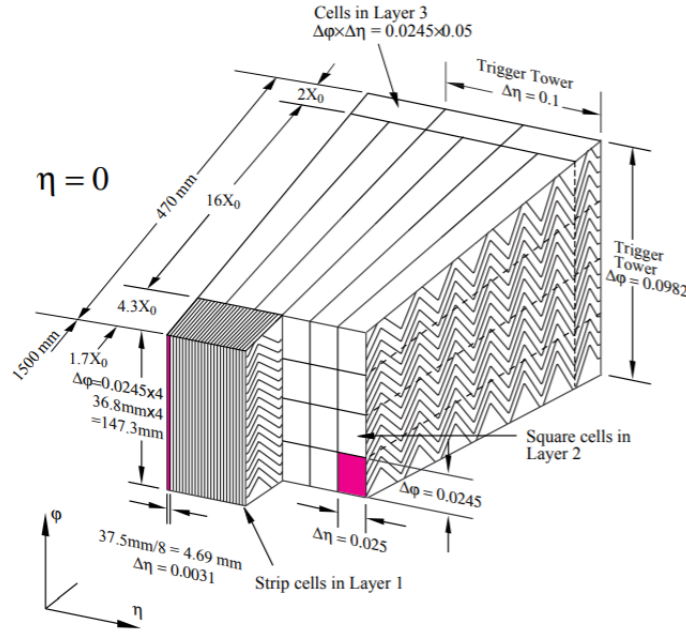


FIGURE 3.8. LAr barrel module featuring accordion-shaped electrodes and trigger towers [11]

The calorimeter is divided into segments in  $\Delta\eta \times \Delta\phi$ , with the most granular sections falling in the most central region that is also covered by the

ID (the barrel and endcaps). Within this range there are more active layers for particles to interact with, leading to more precision. The first of three layers is segmented the most finely in  $\eta$  with  $0.003 \times 0.1$  cells. This allows the  $\eta$  position of electromagnetic showers to be measured with high accuracy when hits are matched to the subsequent showers in the second and third layers with cells of  $0.025 \times 0.025$  and  $0.050 \times 0.025$ , respectively (see Figure 3.8). The largest fraction of energy of the shower is collected in the  $16 X_0$  second layer.

The two endcaps include two wheels that cover  $1.375 < |\eta| < 2.5$  and an inner wheel extending coverage to  $|\eta| < 3.2$ . A presampler single LAr layer extends to  $|\eta| < 1.8$  and corrects for the energy lost before the particle interacts with the bulk of the detector from interactions in the ID or with the solenoid. These different components provide full coverage up to  $|\eta| < 3.2$ , but result in an asymmetric distribution of overall material (see Figure 3.9). The transition region between the barrel and endcaps cause the decrease in material around  $|\eta| = 1.4$ . To recover energy deposits in this “crack” area of  $1.37 < |\eta| < 1.52$ , a scintillator layer is included.

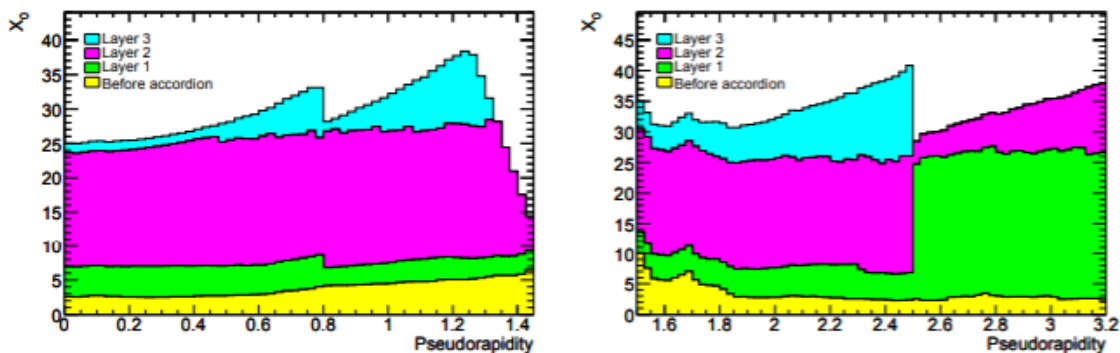


FIGURE 3.9. LAr material budget in  $\eta$  [11]

The energy resolution [77] for the electromagnetic objects measured by LAr is parameterized as

$$\frac{\sigma}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c ,$$

where  $a$ ,  $b$ , and  $c$  are  $\eta$ -dependent parameters,  $E$  is the incident particle energy, and  $\oplus$  is a quadratic sum. LAr was designed for a resolution of

$$\frac{\sigma}{E} \sim \frac{10\%}{\sqrt{E/\text{GeV}}} \oplus 0.7\% .$$

The analog signal is read out with electrodes, amplified, and digitized.

Another sampling calorimeter is used to measure the energy of hadrons in the central region with absorbing steel layers and active scintillator tiles (TileCal) that run parallel to the beam line [76]. As this region is shielded from the bulk of the radiation from the interaction, radiation-resistant argon was not necessary. This allows for the more practical scintillator construction. A barrel and extended barrel segments provide coverage up to  $|\eta| < 1.7$  with 64 azimuthal modules (see Figure 3.10). These wedges create projective cracks through the calorimeter, which LAr avoids with the accordion design. Signal from both sides of the scintillating tile is read out with wavelength-shifting fibers into separate photomultiplier tubes which are then grouped in  $\eta$ . Three layers account in total to a depth of  $9.7\lambda$ , where the interaction length  $\lambda$  is the mean free path of a hadron (see Figure 3.11).

Hadronic showers are measured up to  $|\eta| < 3.2$  with the addition of hadronic endcaps (HEC) that use liquid argon, similarly to LAr. The HECs contain two wheels on each side, and fall directly behind the LAr endcap segments (see Figure 3.7) extending coverage to  $|\eta| < 3.2$ . Each wheel has 32 identical azimuthal wedges and two readout layers containing liquid argon and absorbing copper

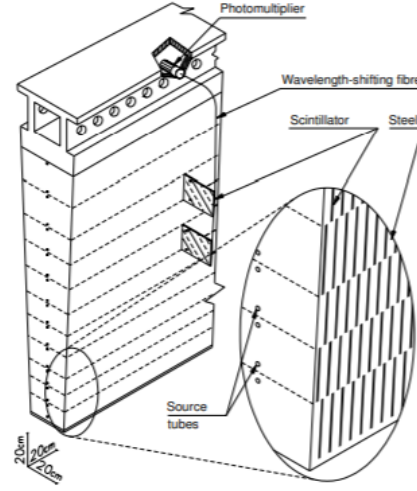


FIGURE 3.10. Tile barrel wedge design and optical readout system [11]

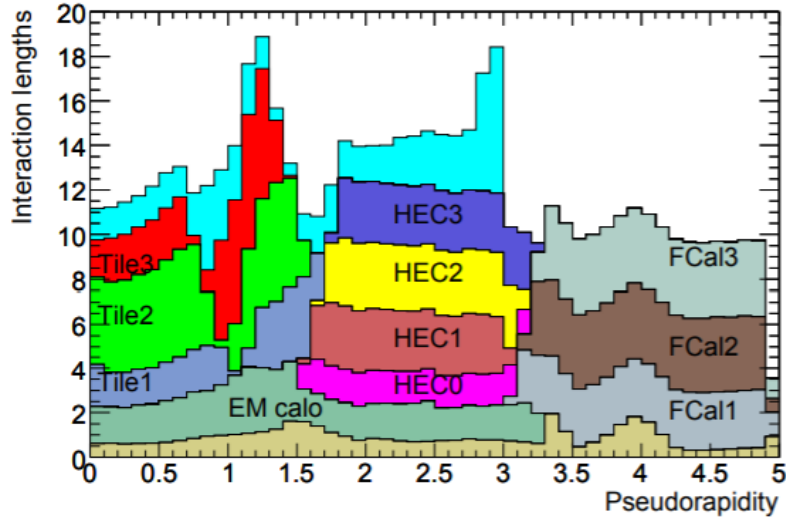
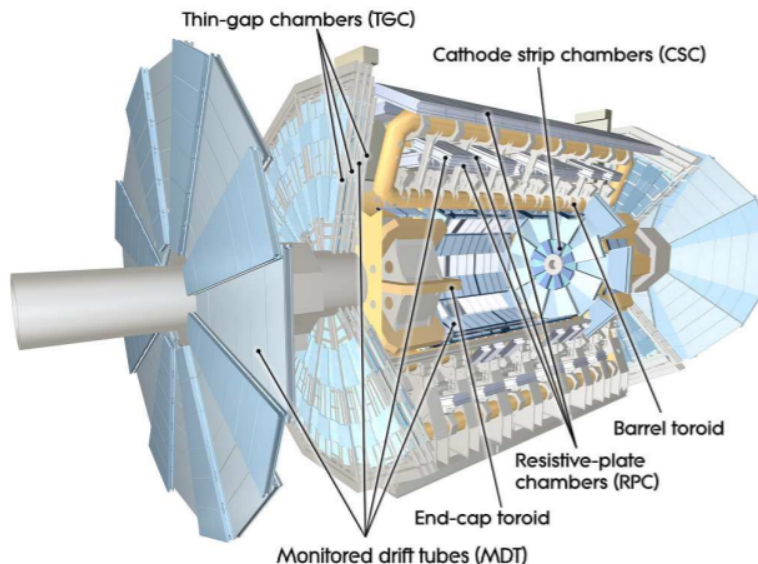


FIGURE 3.11. Tile material budget in  $\eta$  [11]

plates. These copper plates are parallel to each other and roughly perpendicular to incident particles, rather than accordion-style like the electromagnetic calorimeter.

### 3.2.4. Muon System

The outermost subdetector system is the muon spectrometer (MS) which measures the momentum of muons [78]. The task of the spectrometer is to collect tracks from passing muons. Like the tracking system of the ID, an additional magnetic field is provided by toroidal magnets (Section 3.2.5) to bend the paths of the muons so that their momenta can be deduced from curvature of the tracks left in the spectrometer. This is complimented by a trigger system (see Figure 3.12) that helps identify the proper signals to read out.



**Figure 1.4:** Cut-away view of the ATLAS muon system.

FIGURE 3.12. Cut-away of the muon system [11]

The ATLAS muon spectrometer uses monitored drift tubes (MDTs) in the barrel region up to  $|\eta| < 2.7$  for precision measurements of muon tracks, with a shortened inner layer covering only up to  $|\eta| < 2.0$ . The more forward region from  $2.0 < |\eta| < 2.7$  experiences a higher flux of background, so cathode strip chambers (CSCs) are used for precision tracking. Muons act as minimum ionizing

particles (MIPs) and lose roughly the minimum energy as allowed in the material [34]. Both tracking systems pick up signals from these MIPs to build the tracks with a momentum resolution up to  $\Delta p_T/p_T = 10\%$ . The position of each chamber must be understood to the micrometer level, so the system contains precision-mounted alignment sensors to ensure positioning is retained - hence “monitored” in the name.

The drift tubes in the MDT are 30 mm diameter aluminum tubes filled with a mixture of argon gas that surrounds a central wire that run parallel to the beamline. Passing muons ionize the dense gas and free electrons drift toward the central wire for detection within 500 ns. The tubes are arranged into chambers which form three concentric rings around the barrel with the outermost falling 10 m from the IP (see Figure 3.13).

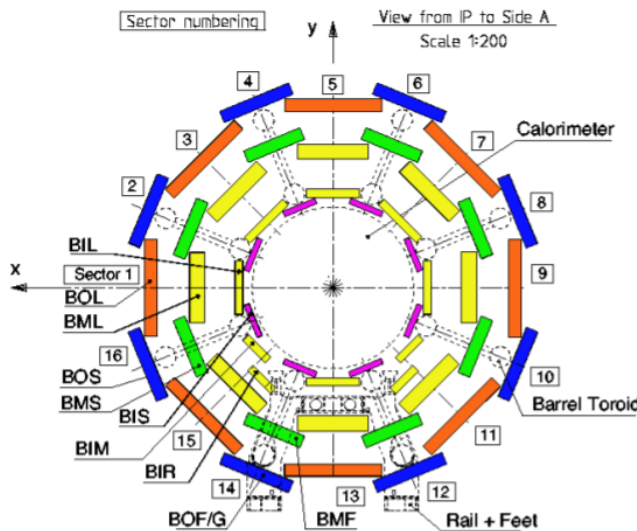


FIGURE 3.13. Cross section of the MDT chamber arrangement [11]

The more forward CSCs use chambers with multiple wires and a cathode strip readout. This provides better time resolution and rate capability than the



central MDTs which is necessary in the forward region of increased particle flux. Three layers are used here as well.

Both systems require triggering systems to determine which signals should be read out. The long response time of the MDTs is incompatible with the 25 ns bunch spacing of the LHC, so resistive plate chambers (RPCs) provide quick signals in the barrel region informing the system whether or not to read out the MDTs. In the endcap region, thin gap chambers (TGCs) provide triggering information (see Figure 3.14). These trigger systems help define  $p_T$  thresholds in the muon system and measure the coordinate of a muon's momentum that is orthogonal to that which the tracking chambers measure. Their quick response time also provides bunch crossing identification.

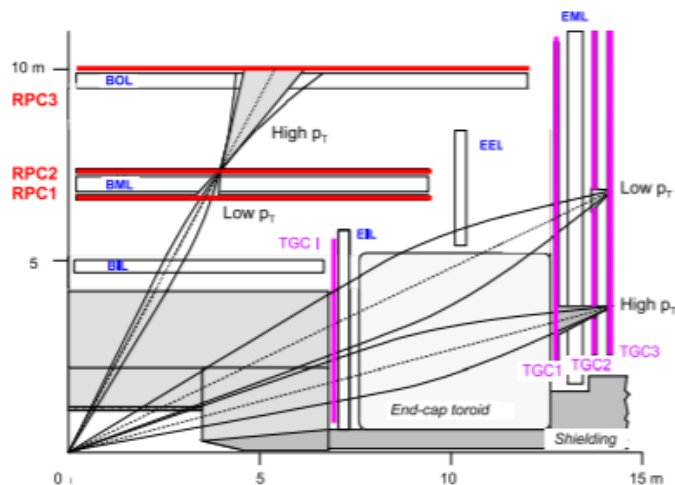


FIGURE 3.14. Muon trigger system [11]

The RPCs are three sets of parallel plates with opposite charge in the range  $|\eta| < 1.05$ . A passing muon ionizes the gas between the plates, and the free electrons cause avalanches that are read out by external metal strips. With relatively low background particle flux, the RPCs only need moderate spatial resolution unlike the endcap TGCs which span  $1.05 < |\eta| < 2.4$  and require high

spatial and timing resolution. They use a multi-wire approach similar to the CSC precision trackers.

### 3.2.5. Magnets

ATLAS employs four superconducting magnets (see Figure 3.4).

A 5.3 m long solenoid magnet outside the inner detector provides the 2 T magnetic field that engulfs the ID. The solenoid length does not extend to the full length of the ID, so the field strength decreases to about 1 T at the edge of the solenoid length and 0.4 T at the forward edge of the ID [71]. The thin coil is designed such that there is not a large material density in front of the calorimeters.

The iconic toroidal magnet sits between the calorimeter and muon systems (see Figure 3.15). A toroid system covers each endcap and the barrel region providing 4.1 T and 3.9 T respectively. This field bends the path of muons before interaction with the muon spectrometer. The toroid systems each have eight coils encompassed by stainless-steel vacuum chambers that keep the superconductors at cryogenic temperatures (see Figure 3.16). They are designed such that the overall magnetic field is roughly perpendicular to all incident particles.

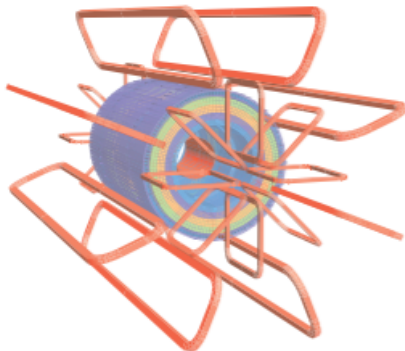


FIGURE 3.15. ATLAS toroid magnet system and tile calorimeter steel [11]

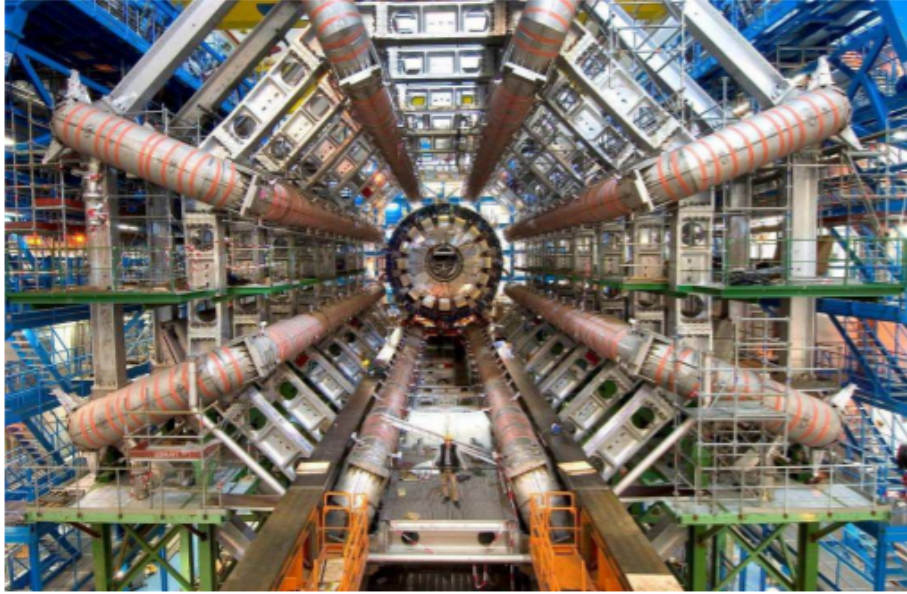


FIGURE 3.16. The barrel toroid installed underground (human on platform for scale) [11]

Both solenoids and toroids create interior magnetic fields of relative uniformity that can be designed such that they are perpendicular to incident particles. A toroid was chosen for the largest magnet subsystem in order to minimize potential particle interactions with additional material [79]. A solenoid of similar size and power would necessarily require a higher material budget leading to poorer muon performance and a higher construction cost.

### 3.2.6. Forward Calorimetry

The main ATLAS subsystems of Sections 3.2.2, 3.2.3, and 3.2.4 detect most physics objects up to  $|\eta| < 3.2$  (see Table 3.2).

Beyond this range, calorimetric coverage up to  $|\eta| < 4.5$  is provided by the liquid argon forward calorimeter (FCal) (see Figure 3.7). The FCal uses liquid argon and is segmented in three layers with the overall depth of  $9.5\lambda$ . The first layer closest to the IP, optimized for electromagnetic showers, is made of copper

| Subsystem | Barrel range [max $ \eta $ ] | Endcap range [max $ \eta $ ] |
|-----------|------------------------------|------------------------------|
| ID        | 1.7                          | 1.7 - 2.5                    |
| LAr       | 1.475                        | 1.375 - 3.2                  |
| TileCal   | 1.7                          | 1.7 - 3.2                    |
| MS        | 2.7                          | 2.0 - 2.7                    |

TABLE 3.2. The  $|\eta|$  expanse of ATLAS subsystems

while other two layers use tungsten for the measurement of hadronic interactions. These metals form tubes with concentric rods that run longitudinally along the calorimeter with liquid argon in the gaps acting as the sensitive medium (see Figure 3.17).

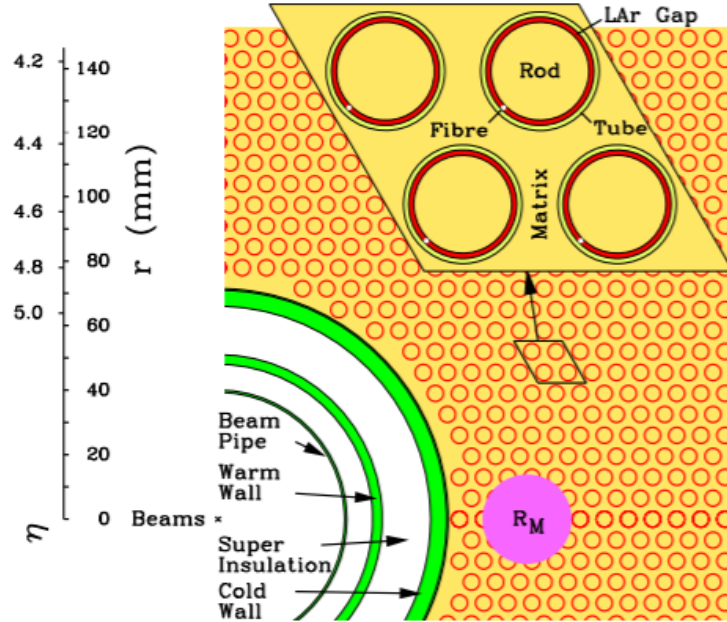


FIGURE 3.17. Forward calorimeter layout [14]

The FCal is contained within the endcap structures of the other calorimeters with its closest face 4.7m from the IP. This has the advantage of providing nearly full  $|\eta|$  coverage in the calorimeter system, but its close location to the IP results

in large radiation fluxes. It was designed to be a dense, compact calorimeter in order to avoid energy spillage into the surrounding endcaps.

### 3.2.7. Trigger and Data Acquisition

Nearly one billion collisions occur every second, at a rate of 40 MHz. Each raw event is roughly 1.6MB, so to retain all this information nearly 64 TB/s would need to be read out and stored [80]. The amount of data that can be stored is limited by the bandwidth of detector readout systems and computer storage space. This drives the need for a trigger and data acquisition (TDAQ) system, which identifies events with interesting physics phenomena in real-time during data-taking.

The ATLAS trigger system for Run 2 uses two independent levels - the hardware-based Level 1 (L1) trigger, and the software-based High Level Trigger (HLT) [81] (see Appendix C for a comparison of the trigger between Run 1 and Run 2). An event passing selection criteria from the L1 system is subject to further selection by the HLT systems before it can be read out into permanent data storage (see Figure 3.18). A trigger menu details the types of desired physics objects and groups selected events based upon which L1 and HLT criteria they pass.

The hardware-based L1 reduces the LHC provided 40 MHz event rate to roughly 100 kHz by considering signals in the calorimeter [82] and muon systems. Regions of Interest (RoI) are defined in  $\eta$  and  $\phi$  as areas with significant energy deposition<sup>2</sup> likely resulting from a physics object (electron, photon, muon, jet,  $\tau$ ) being created.

---

<sup>2</sup>Muon  $p_T$  is used instead of energy - see Section 4.2.3.

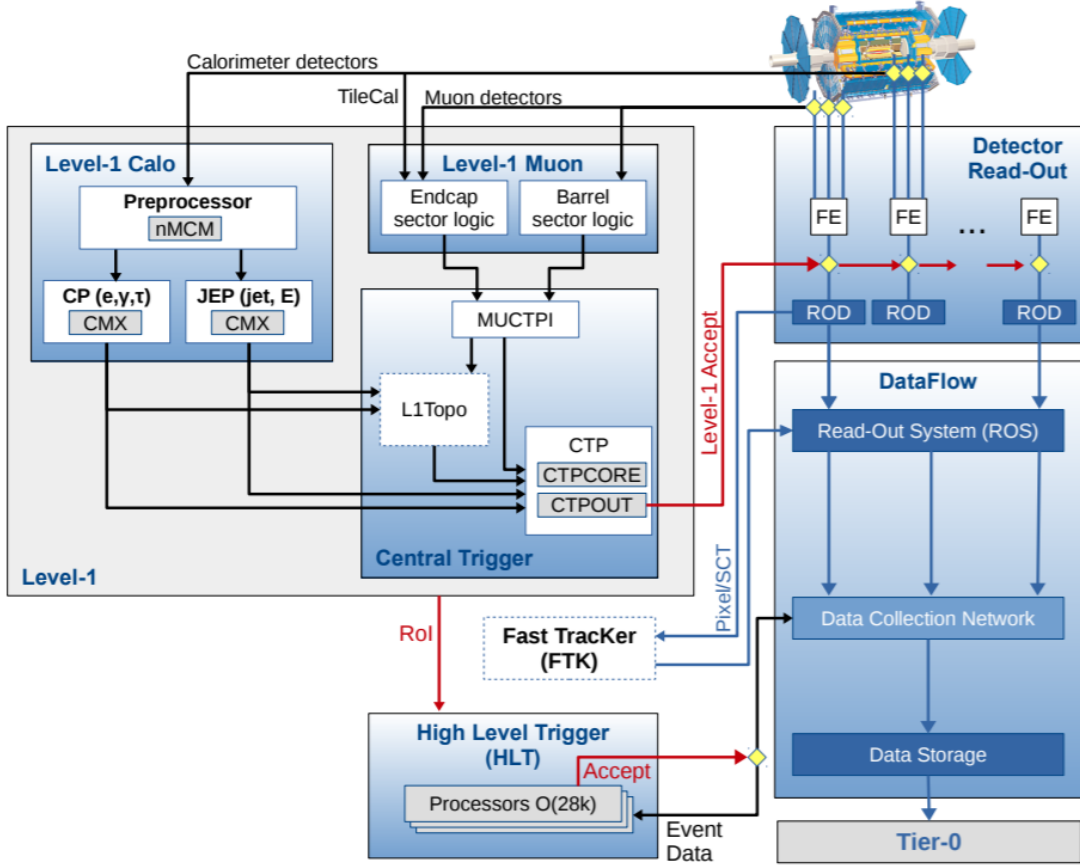


FIGURE 3.18. Flow of the trigger and data acquisition system for Run 2 [15]

More general detector-wide kinematic considerations are made with the Central Trigger Processor, which considers object multiplicities. For example, events with two forward jets (see Section 5.3.3) or missing transverse energy (see Section 4.2.4) can be selected. The central trigger processor decides whether the event is accepted (passes the criteria) based upon timing information from the detector subsystems. This checks whether an event that passes an L1 trigger occurred in coincidence with a collision at the IP. This decision is made within  $2.5 \mu\text{s}$  while data is being collected.

If accepted, the RoIs are then passed to the software-based HLT. In software, the RoI are used to create pseudo-reconstructed physics objects within 300 ms

by running over a processor farm with roughly 40,000 processor cores. The menu defines different object or event criteria called chains [83]. For example, the menu may define a trigger chain requiring at least one jet as defined at L1 with more than 100 GeV. The event is saved if it contains objects that pass the associated HLT criteria (see Section 4.2.1.4). Partial or full event reconstruction can be done based upon requirements of the readout stream. Events may be written to different streams for the purposes of physics analysis, calibration, or monitoring. The physics stream dominates the output saved to long-term storage (see Figure 3.19) as the full event is always saved (which is not the case for the other streams). The HLT reduces the event rate from 100 kHz to 1 kHz which must be shared between all the 2500 trigger chains.

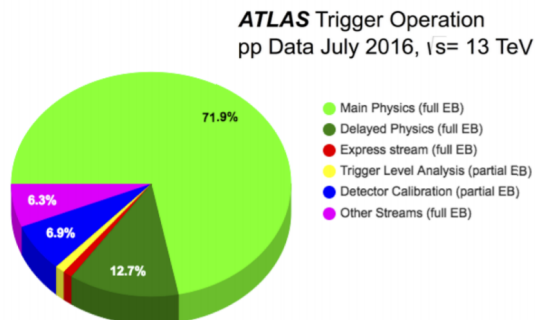


FIGURE 3.19. Contribution of different data streams to the output HLT total bandwidth [16]

In order to regulate the rate that some triggers read out, prescale factors are used. A factor of  $N$  means that only every  $N$ -th passing event is saved, keeping the trigger rate controlled regardless of factors such as increased luminosity or pileup. In addition to the partial event readout, prescaling is another technique used to reduce the bandwidth of calibration and monitoring streams while allotting the most space possible to read out the physics stream.

## CHAPTER IV

### SIMULATION AND RECONSTRUCTION

#### 4.1. Monte Carlo Simulation

In order to draw conclusions from ATLAS data, simulated models of physics events and the ATLAS detector are used. This allows analysts to understand the data output and any potential behaviors that could indicate a deviation from the SM theoretical prediction. Measurements are stochastic by nature, so individual collisions cannot be exactly modeled and therefore exact theoretical predictions cannot be made. Instead, Monte Carlo (MC) simulations of physics processes and detector response are used to model predicted particle behavior. MC simulations rely upon the repeated random sampling of possible physical processes to evolve the underlying event. The particle energy and environment can play a factor in this generation. An underlying physics event is first defined before it hadronizes and interacts with a simulation of the ATLAS detector (see Figure 4.1).

##### 4.1.1. Generation and Simulation

The first step of simulated event generation involves identifying initial conditions of interacting particles. Protons are composite particles; the interactions of note occur when their fundamental contents (quarks and gluons, collectively known as partons) interact with fundamental constituents from the other beam. The proton in its entirety is accelerated to 6.5 TeV, but the individual parton velocity cannot be exactly known. At the high energies of the LHC, deep inelastic scattering is possible and the sea quarks can play a nontrivial role in interactions.



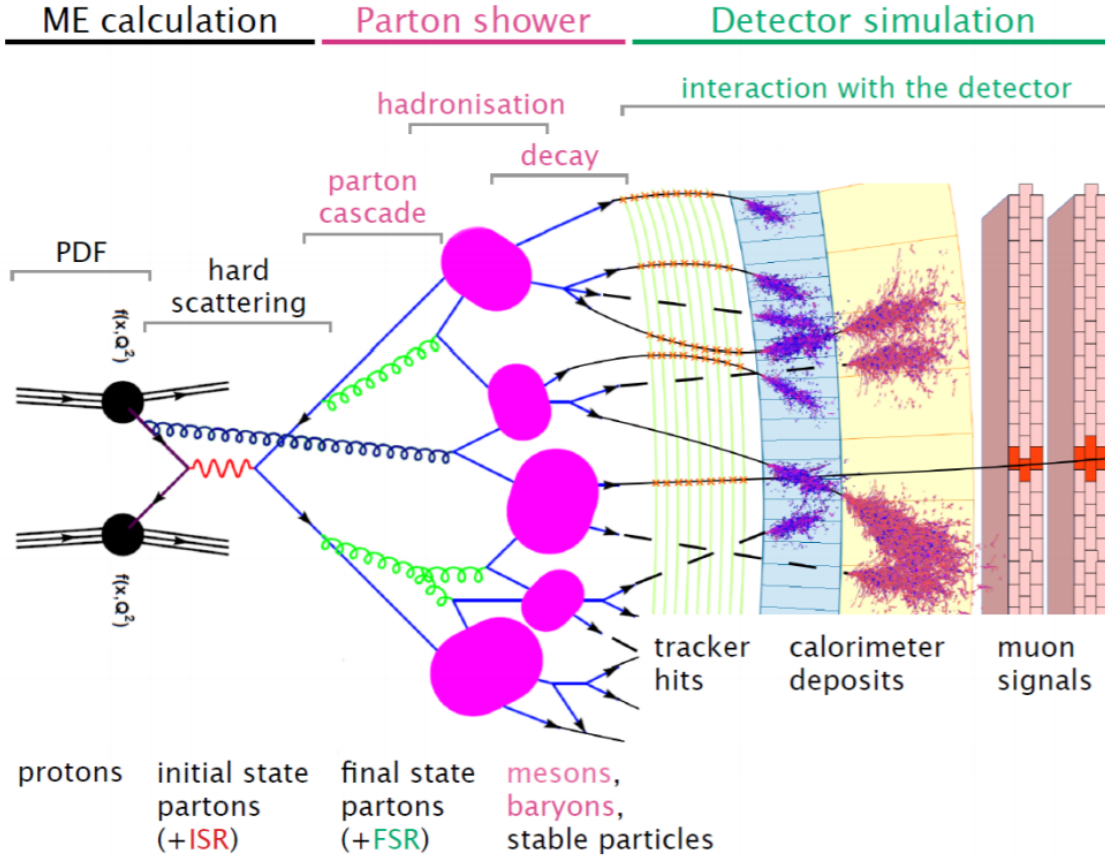


FIGURE 4.1. Pictorial view of the steps necessary to generate MC events [17]

The probability that any parton, valence or sea, carries some momentum fraction of the beam  $x$  is modeled by Parton Distribution Functions (PDFs) [84]. These functions both inform experiment and are refined themselves with further experimental input.

A specific hard scatter (HS) event is defined with incident partons and a desired final state. This underlying physics process can be represented by a Feynman diagram. The event's cross section and a PDF set are combined to give the probability of this specific interaction occurring. This probability is called the matrix element (ME) and is a theoretical calculation. The perturbative nature of the theory requires that the ME calculation be made at a specified order

of perturbation theory, with higher orders being associated with higher levels of complexity both theoretically and computationally but also higher levels of precision. Higher order terms are denoted as next-to-leading order or NLO, NNLO, etc.

Once the ME is generated, the output partons evolve in time through a parton shower (PS) calculation. The PS accounts for hadronization and other QCD processes, evolving the hard scatter event into the final state particles that would interact with the detector. For example, color-changed particles can spontaneously emit gluons as QCD radiation before or after the hard scatter process, creating initial- (ISR) or final-state radiation (FSR). The PS generator accounts for this radiation and its subsequent showering.

PDF sets are common for most samples as an underlying physical aspect of the calculation. Specific tunings of parameters used in ME [85] are used to tune the ME calculation to reflect its environment (for example, to match conditions of Run 2 and the ATLAS detector). Different software packages (generators) are used to generate and shower the HS events. They differ in the perturbative order of the calculation and occasionally in the strategy of parton showering.

#### **4.1.2. Detector Simulation**

Thus far, MC generation emulates physics processes in a generic sense. The final step is to model how these processes would appear within the ATLAS detector itself. This is done with a detailed model of the ATLAS detector with GEANT4 [86] software, which simulates interaction and signal readout in the detector. The final product creates collections of collision events that emulate the signals collected during data-taking. The same processing software is then used to

reconstruct physics objects from both MC and data samples. For MC events that contain a truth record, this reconstruction can be used to understand detector and reconstruction performance.

## 4.2. Physics Objects

Particles from simulated MC events and data interact with the ATLAS subdetector systems (see Section 3.2) in different ways. The signals from multiple different subsystems are considered together in order to identify and reconstruct each physics object (electron, muon, etc).

The ID and muon systems form tracks by connecting charge deposits from passing charged particles. The curvature of these tracks indicates particle momentum and electric charge. Energy deposits from the calorimeters are clustered and calibrated to reconstruct the initial energy and position of particles. Tracks and calorimeter clusters are considered together in particle identification and reconstruction (see Figure 4.2).

In order to reduce the superfluous consideration of pileup objects, the precise track information gathered from the ID is used to define vertices or locations of beam interactions. The vertex that maximizes  $p_T$

$$\sum_t p_{T,t}^2 ,$$

where  $t$  considers all tracks with  $p_T > 400$  MeV in the event, is identified as the primary vertex (PV). This is considered to be the vertex containing physics worthy of study, and all other vertices<sup>1</sup> are considered pileup. Due to uncertainties

---

<sup>1</sup>Some vertices displaced from the PV are still considered for  $b$ -jet tagging (see Section 4.2.1.3).

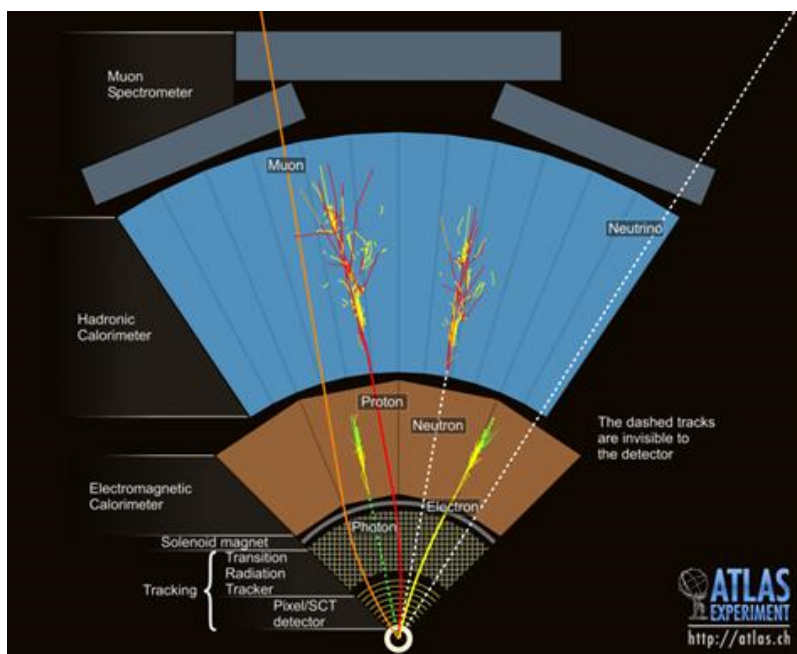


FIGURE 4.2. Signatures left in different ATLAS systems for different particles [18]

in tracking, vertex identification comes with large associated uncertainties. It is also complicated by events with forward physics objects that do not fall within the tracker volume.

#### 4.2.1. Jets

Single quarks and gluons are the most common end result of a proton-proton interaction, due to the low cross section of soft QCD processes. Through hadronization (see Section 2.1), collections of hadrons are produced which then interact with the detector. These hadrons leave a collimated spray in the detector, showering through QCD processes of gluon emission and pair production much like electromagnetic showers from electrons and photons. These signatures in the detector are called jets, and are associated with one initiating parton. The direction of the initiating parton is usually evident from a core of higher-energy

deposits, with the shower developing around this core and growing in size as it passes through the detector.

Jets can be valuable contributors to the  $h \rightarrow \text{inv.}$  signal, as they are frequently associated with Higgs boson production mechanisms and can be used to identify  $h \rightarrow \text{inv.}$  events.

#### 4.2.1.1. Reconstruction Techniques

Jets are identified in ATLAS with jet-finding algorithms. The simplest jet-finding algorithm may identify the high-energy core and build a cone with a predefined radius around the core, collecting all energy deposits within the cone and adding them to create the reconstructed jet energy. Reliable jet-finding algorithms must additionally be collinear and infrared safe. This protects the number of identified jets from changing due to the splitting or merging of high transverse momentum particles or the presence of soft gluon emission between jets, respectively [87].

Jets are expected to leave signal in the ID and calorimeters (see Figure 4.2). Charged hadrons, containing roughly 60% of the total energy of the jet, leave tracks in the ID and hadrons will interact with the calorimeters. Photons contained within the jet carry roughly 30% of the jet's energy, and shower in the LAr electromagnetic calorimeter. Since jets are detector signatures, the selected jet-finding algorithm can use information from one or all of these sources. Significant deposits in calorimeter cells that are topologically connected are grouped together into topo-clusters associated with particle showers [88], which can define calorimeter-only jets (EMTopo jets). Topo-clustering reduces noise that can be associated with the simple summing of all energy deposits in a cone.

Corrections to the measured energy are then made using tracker information. Including tracker information at the stage of jet-finding aims to identify individual particles within the jet (PFlow jets). This technique improves the accuracy of the measurement of charged hadrons and enhances the transverse momentum resolution of low- $p_T$  jets [89]. The inspiration for this ‘particle flow’ (PFlow) technique hails from the lepton collider LEP which first successfully used the idea of ‘energy flow’ [90]. Its success is enhanced by very highly granular tracker and calorimeter systems which allow for individual particles to be identified (see Section 7.4).

The choice of algorithm affects the resulting collection of identified jets. For PFlow jets, this algorithm considers topo-clusters and tracks associated with the PV. These tracks are then matched to topo-clusters beginning with the highest  $p_T$  tracks. The particle energy is computed considering both tracker and calorimeter information. It is common for a particle to deposit energy in multiple topo-clusters, so the algorithm considers whether multiple topo-clusters must be further clustered together to recover all the energy from the particle. This process identifies a particle flow object that can then be used for jet reconstruction.

A standard choice, used in this discussion of  $h \rightarrow \text{inv.}$ , is the anti- $k_t$  clustering algorithm [19] with a radius parameter<sup>2</sup> of  $R = 0.4$ . A distance parameter between each particle flow object ( $i, j$ ) (or each individual topo-cluster, in the case of EMTopo jets) and between each object and the beam ( $B$ ) is computed as

$$d_{ij} = \min(k_{t,i}^{-2}, k_{t,j}^{-2}) \frac{\Delta_{ij}^2}{R^2}, \quad d_{iB} = k_{t,i}^{-2},$$

---

<sup>2</sup>This is not a radial distance like Eq. 3.5, but rather a radius parameter of order 1 that is analogous to a cone size for cone-based jet finding algorithms [91].

where  $k_t$  is the transverse momentum,  $\Delta_{ij}$  is the distance<sup>3</sup> between objects  $i$  and  $j$ , and  $R = 0.4$ . The smallest  $d_{ij}$  is first identified and the  $i$  and  $j$  particle flow objects are clustered together if  $d_{ij} > d_{iB}$ . If  $d_{ij} < d_{iB}$ , then object  $j$  is eliminated from consideration. This new cluster is considered as the new object  $i$  in this iterative process, which continues clustering particle flow objects together until there are no additional objects to consider. The radius parameter  $R$  acts as a cut-off for energy clustering, but does not define a size of the jet. Objects with  $\Delta > R$  are still considered. This process results in a circular jet (see Figure 4.3), which aids in ease of calibration and modeling.

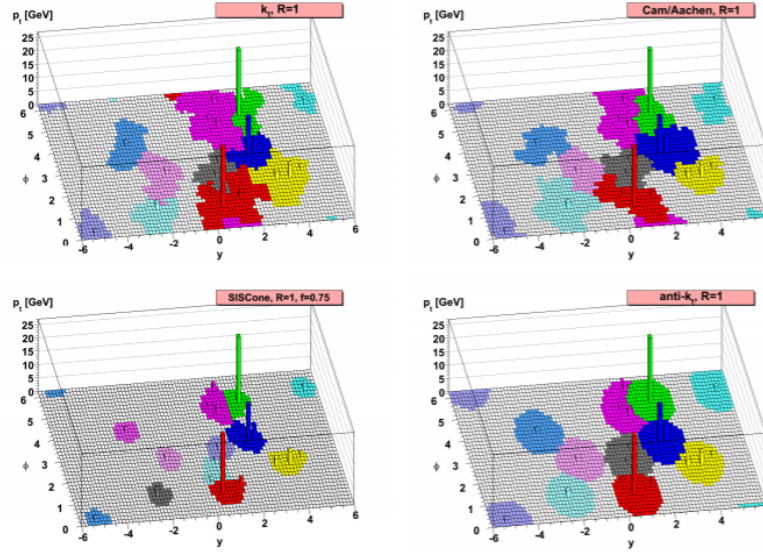


FIGURE 4.3. Energy deposit clustering with different jet-finding algorithms [19]

The jets are corrected for the impact of homogenous pileup [92] and other experimental effects such as detector inhomogeneities [93].

---

<sup>3</sup> $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$

#### 4.2.1.2. Jet Vertex Tagging

As described in Section 3.1.3, the most frequent interaction in the LHC is proton–proton inelastic scattering which leaves soft QCD radiation in the detector. The presence of these additional jets can complicate the construction of event-wide quantities as they are not associated with the hard scatter event. In ATLAS, a variable called the jet vertex tagger (JVT) is used to flag jets as originating from pileup events, and removing them from physics consideration [20]. JVT uses a multivariate combination of two independent variables - the corrected jet vertex fraction (corrJVF), and a ratio of track  $p_T$  to jet  $p_T$  ( $R_{pT}$ ) - built largely from tracker information to assign each jet a score related to the probability that it is pileup in a way that keeps the hard scatter jet efficiency independent of the degree of pileup.

Due to pileup interactions many physics analyses first define a primary vertex ( $PV_0$ ) as the hard scatter vertex, where  $N$  vertices may be reconstructed but  $PV_0$  is the one with the largest  $\Sigma p_T^2$  of tracks. The jet vertex fraction (JVF) can be calculated with any number of vertices as

$$\text{JVF} = \frac{\Sigma_k p_T^{\text{trk}_k}(PV_0)}{\Sigma_m p_T^{\text{trk}_m}(PV_0) + \Sigma_{n \geq m} p_T^{\text{trk}_n}(PV_n)} = \frac{\Sigma_k p_T^{\text{trk}_k}(PV_0)}{\Sigma_m p_T^{\text{trk}_m}(PV_0) + p_T^{\text{PU}}} , \quad (4.1)$$

where the sum over  $k$  represents tracks associated with the jet in question and  $p_T^{\text{PU}}$  is the scalar sum of all tracks associated with pileup. JVF measures the fraction of track momenta associated with the primary vertex and was used as an early pileup discrimination variable, but implementing a cut on JVF introduces a dependence on the number of reconstructed vertices to the jet efficiency due to the term in the denominator. To avoid this dependence, JVF is corrected by the total number of



pileup tracks in the event that do not originate from  $PV_0$ ,

$$\text{corrJVF} = \frac{\sum_k p_T^{\text{trk}_k}(PV_0)}{\sum_m p_T^{\text{trk}_m}(PV_0) + \frac{p_T^{\text{PU}}}{k \cdot n_{\text{trk}}^{\text{PU}}}} , \quad (4.2)$$

where  $k = 0.01$ .

Another variable that is stable across the number of reconstructed vertices is  $R_{\text{pT}}$ , or the ratio of a jets track momenta to the total measured transverse momentum,

$$R_{\text{pT}} = \frac{\sum_k p_T^{\text{trk}_k}(PV_0)}{p_T^{\text{jet}}} . \quad (4.3)$$

Pileup jets are not associated with the primary vertex, so  $R_{\text{pT}}$  peaks at 0 whereas hard scatter jets have a wider and positive  $R_{\text{pT}}$  distribution as it measures the charged  $p_T$  fraction of jet constituents. Since only the primary vertex is considered, there is no dependence on the total number of reconstructed vertices.

The corrJVF and  $R_{\text{pT}}$  values for each jet are then considered. Both values are expected to be small for pileup jets and large for hard scatter jets. JVT considers the corrJVF/ $R_{\text{pT}}$  plane for a dictionary of known hard scatter and pileup jets and finds the ratio of hard scatter to pileup jets in a region surrounding the (corrJVF,  $R_{\text{pT}}$ ) of the jet in question. This provides a probability for the jet being hard scatter-like (see Figure 4.4).

As JVT is reliant upon predefined dictionaries of jets, it is only applicable to jets within certain  $p_T$  ranges ( $20 < p_T < 60$  GeV). Individual analyses are then free to make requirements on the JVT score for jets to qualify for participation in the analysis. Official collaboration-wide recommendations are also set by a central group, which provides working points associated with JVT cut values. For

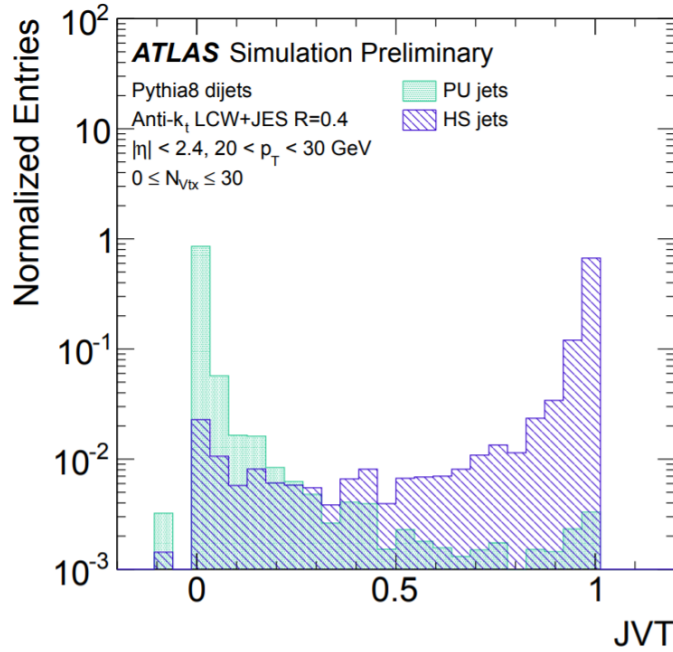


FIGURE 4.4. JVT score for pileup (PU) and hard scatter (HS) jets with  $20 < p_T < 30$  GeV [20]

example, the tightest working point for PFlow jets requires a JVT cut of 0.5 and has an average efficiency of 96%.

JVT is a tracker-based quantity, relying on the accurate measurement and reconstruction of tracks. Therefore, its use is limited to the barrel region of ATLAS with tracker coverage ( $|\eta| < 2.4$ ). The lack of forward tracker information complicates the association of forward jets to vertices. Pileup jets occur isotropically, so a different method is needed to control pileup in the forward region. The forward jet vertex tagger (fJVT) identifies different types of forward pileup jets and uses their forward calorimeter signatures for discrimination from hard scatter jets [94]. More details about the fJVT algorithm can be found in Appendix D.

As with JVT, different centrally-defined working points for fJVT are defined. Up to 51% of forward pileup jets can be rejected using the fJVT discriminant while retaining an 85% hard scatter efficiency.

#### 4.2.1.3. *b*-jet Tagging

Third generation *b* quarks are long-lived relative to hadronization due to their large masses. This allows them to travel short distances - for example, a *B*-meson can travel more than  $450 \mu\text{m}$  before decaying [21]. The location of this decay is distinct from the original particle interaction at the primary vertex, and called a secondary vertex. Tracks from the *B*-hadron decay (associated with the secondary vertex) are used to define the transverse impact parameter  $d_0$  - a measure of distance of closest approach of the secondary track to the primary vertex in the  $r - \phi$  plane [95] (see Figure 4.5). A longitudinal impact parameter ( $z_0 \sin \theta$ ) is similarly defined in the longitudinal plane.

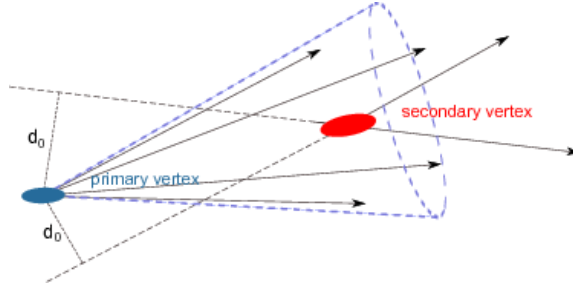


FIGURE 4.5. Tracks associated with a *b*-jet, including the secondary vertex and impact parameter  $d_0$  [21]

The identification of jets initiated by heavy-flavor quarks is done with the multivariate algorithm MV2c10 that tags the likelihood of heavy-flavor initiation. The multivariate algorithm combines the outputs of three independent *b*-tagging

algorithms that themselves consider the impact parameter, secondary vertex reconstruction, and secondary vertex placement for each jet.

Operating points are defined that provide various  $b$ -jet identification efficiency values, related to cut requirements on the multivariate output. These efficiencies vary as a function of jet  $p_T$ , even within one operating point, with degrading efficiency at lower jet  $p_T$ . The standard operating point provides, on average, 77% efficiency in correctly tagging  $b$ -jets with  $p_T > 25$  GeV. This also has a  $c$ -jet rejection rate of 6 and light-flavor rejection rate of 134.

Heavy-flavor jets are not expected in the  $h \rightarrow \text{inv.}$  signal, but some Higgs boson production modes can involve top or bottom quarks such as  $t\bar{t}H$  (see Section VI).

#### 4.2.1.4. Jet Substructure Trigger

Jet objects can be used for triggering, and some triggers are designed such that they reject background events such as those from pileup and the QCD multijet background. One strategy for identifying objects of interest such as vector and Higgs bosons and top quarks offline (at the analysis stage) involves considering large-radius parameter jets of  $R = 1.0$  and searching for substructure within the jet [96]. If the object is sufficiently boosted (see Figure 4.6), its decay products (colored  $R = 0.4$  jets in Figure 4.6) will be collimated enough that all their energy is contained within a single jet (green  $R = 1.0$  jet in Figure 4.6). Having multiple decay products contributing to the jet creates substructure within the large-radius parameter jet, which is lacking for jets that originate from QCD-like multijet interactions (see Figure 4.7).

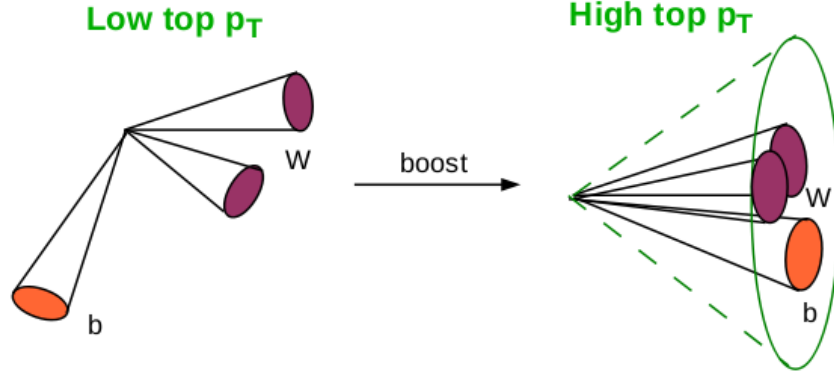


FIGURE 4.6. Signature from a boosted top quark [22]

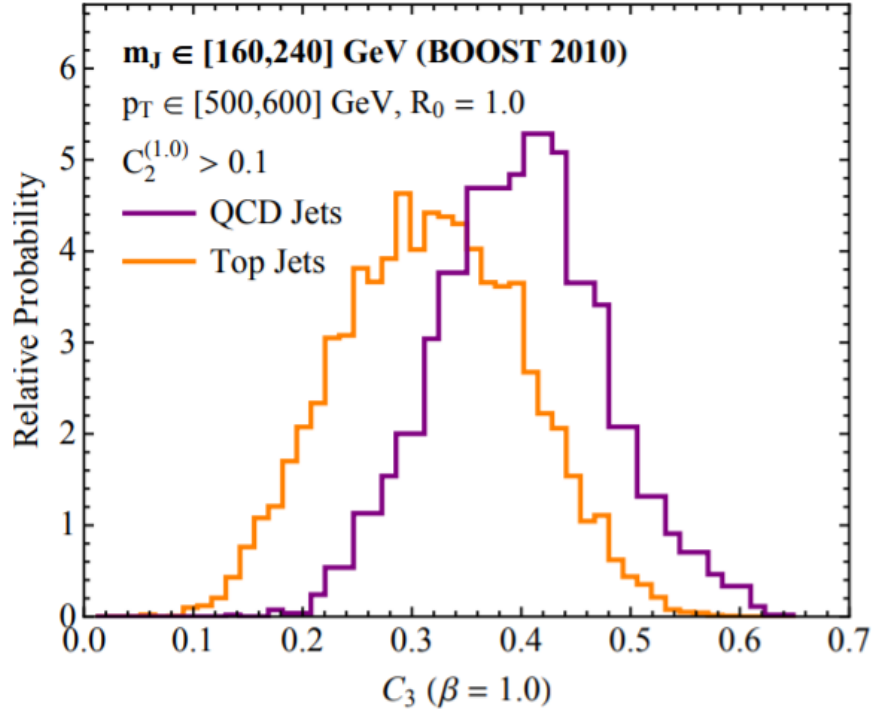


FIGURE 4.7. Distributions of a jet substructure variable for large- $R$  jets originating from top quarks and QCD background [23]

I was the first to begin implementation of a trigger that could consider variables that quantify the substructure of jets within the HLT itself [97], which would create data sets enriched with boosted objects for study. Two powerful

variables [98] were identified for initial implementation - the  $p_T$ -weighted jet width,

$$W = \frac{\sum_i \Delta R^i p_T^i}{\sum_i p_T^i} ,$$

and KtDeltaR (the distance between the last two merged jet constituents when a large- $R=1.0$  anti-kt jet is reclustered with the kt algorithm until two constituents remain).

The code was designed such that parameters from trigger chain names would evoke the consideration of one or both of these variables, and require a pass for the event to satisfy the chain. This leads to a data set enriched with physics objects containing substructure and fewer contaminating QCD events. Details of implementation can be found in [99]. This work was then migrated into the Run 3 menu with the ability to build upon the foundation I set for running in Run 3.

#### 4.2.2. Electrons and Photons

Compared to jets, electrons and photons are relatively straightforward detector signatures. Both form electromagnetic showers in the LAr calorimeter and deposit the majority of their energy in this system (see Appendix B for differences between electron and photon electromagnetic showers). One main difference in their signatures (see Figure 4.2) is the lack of a charged track in the ID from neutral photons.

Electrons are identified by matching a charged track from the ID with a cluster of energy deposits in LAr [100]. Electrons can occur from physics processes of interest at the IP (so-called prompt electrons) or from background processes such as converted photons (photons which pair produce into an electron-positron

pair) or heavy flavor decays. ATLAS uses a multivariate algorithm that considers electromagnetic shower shape and track-to-cluster matching to identify prompt electron candidates with different working points (WPs). Each WP is associated with some level of background rejection and signal selection efficiency. The *Loose* WP allows the most background contamination, but runs the lowest risk of rejecting signal electron candidates due to reconstruction errors or inefficiencies. Electrons can also be considered with a *Medium* or *Tight* WP which each increase the level of background rejection.

Isolation parameters are defined to consider the physical distance between an electron candidate and other physics objects in the event [101]. This can be track- or calorimeter-based, and provides additional background discrimination. Prompt electrons are more likely to be isolated than electrons from background processes.

The efficiency of electron reconstruction is measured using data samples rich in two-electron events, such as  $Z \rightarrow ee$ . The efficiency measured in data is used to correct MC reconstruction efficiency. ATLAS electron reconstruction and identification efficiencies are roughly constant across  $\eta$  and degrade from sub-percent levels to a few percent as electron energy decreases below 30 GeV. Efficiency is highest for the loosest WP.

Electron energy is calibrated such that it is uniform throughout the detector. The electron calibration strategy was optimized for Run 1 [102] and further refined for Run 2 [103].

Photon identification is challenged by the presence of non-prompt photons, largely from hadron decays in jets and hadrons that mimic photon energy depositions showering early in the electromagnetic calorimeter [104]. Shower shape and isolation variables are used for photon identification as well. These are

considered separately for converted and unconverted photons, which leave different electromagnetic signatures. The photon identification efficiency in data ranges from 0.5%-5% depending on photon  $p_T$  and  $\eta$ .

Electrons and photons are not expected in the  $h \rightarrow \text{inv.}$  signal that this dissertation considers. Events containing electrons or photons are vetoed from consideration as signal, but can be further considered in other useful regions (see Section V).

### 4.2.3. Muons

Muon identification relies upon the tracking capabilities of ATLAS for identification (see Figure 4.2), both in the inner detector (ID) and the muon spectrometer (MS). Tracks in each subsystem are independently created, and candidate muon events combine the signals from both subdetectors to form one full muon track [105].

Four different types of muons are defined, based upon the signatures they leave. Combined muons contain isolated tracks in both the ID and MS. These are identified first in the MS and reconstructed from the outside-in, connecting the MS track to a candidate ID track. Hits that define tracks are chosen with a global  $\chi^2$  refit. Low- $p_T$  muons may not leave a full track through the entire MS, so segment-tagged muons are also defined where ID tracks are connected to MS signatures with deposits in only one layer.

Muons act as minimum ionizing particles in the calorimeters, so calorimeter-tagged muons can be identified if an ID track can be matched with a cluster of MIP deposits through the calorimeters. This muon type is prone to inefficiencies, but can be used to recover muon candidates in regions with reduced MS



instrumentation ( $|\eta| < 0.1$ ). Forward muons that may not interact with the ID can still be identified from MS tracks. These extrapolated muons require a MS track and compatibility with originating from the IP.

Like electrons and photons, similar identification WPs and isolation requirements are identified for muons. Muon momenta are calibrated from combined muons, using the transverse momenta measured independently in the ID and MS. The accuracy of the corrected, overall transverse momentum is a function of detector  $\eta$  and  $\phi$ , and requires the input of MC studies of detector performance to muon signatures.

Muons are not expected in the  $h \rightarrow \text{inv.}$  signal that this dissertation considers. Events containing muons are vetoed from consideration as signal, but can be further considered in other useful regions (see Section V).

#### **4.2.4. Missing Transverse Energy**

Particles such as neutrinos or theorized dark matter interact minimally with other SM particles, and therefore also with the ATLAS detector. There is not a signature of depositions that can be identified with them, like for jets, photons, and leptons. Instead, their presence in an interaction is deduced by measuring the energy that they carry away from an interaction.

Even in collider environments, momentum must be conserved in interactions. Due to PDFs, individual interacting parton velocities in the  $z$  (or beam) direction cannot be known, but the momentum transverse to this direction is known to be small. The transverse plane measured by the detector must therefore also conserve momentum. Once all physics objects in an event are identified, any imbalance in transverse momentum represents that which is carried away by a particle

that did not interact with the detector. This missing transverse momentum is calculated with the negative vector sum of all physics objects and qualifying tracks in the event. This vector provides a direction (in the transverse plane) of these unmeasured particles carrying an energy indicated by the magnitude of the vector called the missing transverse energy (MET, or  $E_T^{\text{miss}}$ ). An additional element from tracks pointing to the PV but not necessarily associated with a physics object is included in the MET calculation. This contribution is called the soft term [106].

Pileup must be removed from the MET calculation, as it originates from a different momentum-conserving vertex. This is managed by requiring that jets involved in the calculation pass JVT requirements. The mismeasurement of physics objects, such as the incorrect assignment of JVT for jets or misidentified and miscalibrated leptons, can introduce “fake MET”.

MET is a crucial element of the  $h \rightarrow \text{inv.}$  signal, where the invisible particles are included in the search requirements through a requirement of MET.

#### 4.2.4.1. Missing Jet Energy

An additional MET-like variable called  $E_T^{\text{jet, no JVT}}$  is defined. Rather than considering momentum conservation of the PV, this variable considers all jets from any vertex with any JVT score.  $E_T^{\text{jet, no JVT}}$  is computed as the negative vector sum of all jets with  $p_T > 20$  GeV regardless of their JVT score. This allows pileup jets to contribute to the calculation.

This variable is used in the  $h \rightarrow \text{inv.}$  search in the VBF channel in an effort to understand fake MET introduced by improper JVT tagging. If a hard scatter jet is inappropriately removed due to JVT, a restriction on  $E_T^{\text{jet, no JVT}}$  would

remove that event from signal consideration. It supplements the fJVT requirement of the characteristic forward jets (see Section 5.3.3).

#### 4.2.4.2. MET Triggers

MET is an object that can be used to seed triggers. This requires that deposits within the full detector are taken into consideration, for fast L1 and HLT MET calculations. There are different MET calculation algorithms that can be used within the trigger. The calculation is complicated by the presence of pileup and general mismeasurements, both of which contribute fake MET to an event [107]. At L1, MET is calculated as the negative sum of measured calorimeter energy with coarse spatial granularity. The more advanced HLT uses calibrated calorimeter clusters to roughly reconstruct jets, and calculates MET with each jet with  $p_T > 7$  GeV.

MET thresholds at the HLT level are required for these triggers. Higher pileup conditions introduce higher trigger rates, so these thresholds were changed over the course of Run 2 to control the trigger rate. The lowest unscaled triggers had MET thresholds of 70 GeV in 2015 and 110 GeV in 2016-2018. Fake MET most dramatically impacts trigger performance at low values of MET (see Figure 4.8), leading to low efficiency. These triggers become fully efficient (“plateau”) with offline MET  $\gtrsim 180 - 200$  GeV.

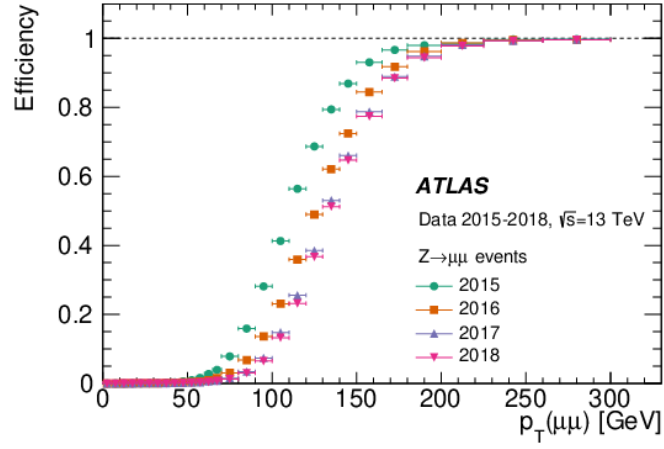


FIGURE 4.8. MET trigger efficiencies for Run 2 [24]

## CHAPTER V

### VBF $h \rightarrow$ invisible

#### 5.1. $h \rightarrow$ invisible **Motivation**

Much can be gained from studying the Higgs boson and obtaining precise measurements of its properties. There is a SM-predicted process ( $H \rightarrow ZZ^* \rightarrow \nu\nu\nu\nu$ ) in which a Higgs boson ultimately decays to particles that are invisible to the ATLAS detector - neutrinos. The branching ratio for this process is small,  $\mathcal{B}_{h \rightarrow \text{inv}} \sim 0.001$  [108], and difficult to measure. This process in particular carries potential implications for Higgs portal dark matter (see Section 2.3), where the addition of a Higgs decay to DM ( $H \rightarrow \chi\chi$ ) would leave an identical detector signature. This addition could impact the expected branching ratio measurement by  $\mathcal{O}(10\%)$  [54]. Any deviation of the measured branching ratio from the SM-only expectation could hint toward a DM contribution.

Discovery is claimed when a measurement deviates by more than  $5\sigma$  from the SM prediction. In the absense of such statistical significance, an upper limit is set instead. Through this work, upper limits (UL) are reported through calculation via the  $\text{CL}_S$  prescription at 95% confidence level (C.L.) [109] unless otherwise stated. Typically, two limits are reported - the observed (obs) value found with data, and the expected (exp) value which assumes only SM processes and is calculated without measured data<sup>1</sup>.

The branching ratio of Higgs to BSM decays can be inferred by measuring all visible Higgs decay channels and considering what fraction of the total width

---

<sup>1</sup>Instead, an Asimov data set from simulated events meant to be representative of overall data but without statistical fluctuations is used [110].

is unaccounted for. BSM Higgs decays could contribute to the  $h \rightarrow \text{inv.}$  measurement, or be present in BSM final states that are not currently directly searched for. At the end of Run 1, this was considered by combining ATLAS and CMS data [111]. The overall branching ratio to such undetermined final states was  $\mathcal{B}_{\text{undet.}} < 0.34$ . This leaves room for BSM to contribute to Higgs measurements.

The  $h \rightarrow \text{inv.}$  signature can be directly searched for by looking for events with MET (see Section 4.2.4) from the invisible decay. This is done most precisely by defining one Higgs production channel to study (see Figure 2.3). Prior to the full analysis of Run 2, upper limits of  $\mathcal{B}_{h \rightarrow \text{inv.}}$  were  $\sim 25\%$  (see Table 5.1).

The current strictest experimentally-set upper limit on  $\mathcal{B}_{h \rightarrow \text{inv.}}$  using the full Run 2 data set is 0.11 [112] (see Section VI) - two orders of magnitude above the expected SM rate. This result, from a statistical combination of multiple search channels, is driven by the VBF channel [25] (see Section 5.2) which measured  $\mathcal{B}_{h \rightarrow \text{inv.}} < 0.13$ . This result also informed a full Run 2 study of combined Higgs coupling measurements, refining the branching ratio of Higgs to BSM states to  $\mathcal{B}_{\text{undet.}} < 0.19$  [113].

## 5.2. VBF Motivation

In the LHC, the Higgs boson production mode (see Figure 2.3) with highest cross section is gluon-gluon fusion (ggF), in which two initial state gluons interact through a loop and produce a Higgs boson. Though this process produces the most Higgs bosons, it is a difficult  $h \rightarrow \text{inv.}$  signal to search for. Additional ISR or FSR is required in order to tag events for study.  $h \rightarrow \text{invisible}$  is considered as an interpretation of the jet+MET (or ‘monojet’) search which measures  $\mathcal{B}_{h \rightarrow \text{inv.}} < 0.34$  (0.39) exp (obs) [122].

| Data set                      | Measurement      | Source | Observed UL | Expected UL |
|-------------------------------|------------------|--------|-------------|-------------|
| Run 1                         | $V(\text{had})H$ | [114]  | 0.78        | 0.86        |
|                               | $Z(\text{lep})H$ | [115]  | 0.75        | 0.62        |
|                               | VBF              | [116]  | 0.28        | 0.31        |
|                               | Combination      | [117]  | 0.25        | 0.27        |
| Run 2, 36 fb <sup>-1</sup>    | $V(\text{had})H$ | [118]  | 0.83        | 0.58        |
|                               | $Z(\text{lep})H$ | [119]  | 0.67        | 0.39        |
|                               | VBF              | [120]  | 0.37        | 0.28        |
|                               | Combination      | [121]  | 0.26        | 0.17        |
| Run 2                         | VBF              | [25]   | 0.13        | 0.13        |
|                               | $t\bar{t}H$      | [112]  | 0.40        | 0.36        |
|                               | Combination      | [112]  | 0.11        | 0.11        |
| $\mathcal{B}_{\text{undet.}}$ | Run 1            | [111]  | 0.34        | -           |
| $\mathcal{B}_{\text{undet.}}$ | Run 2            | [113]  | 0.19        | -           |

TABLE 5.1. ATLAS upper limits (UL) of  $h \rightarrow \text{inv.}$  and  $\mathcal{B}_{\text{undet.}}$

The next most frequent production mode is vector boson fusion (VBF), which requires two incident quarks to radiate vector bosons ( $Z$  or  $W$ ) which interact and produce a Higgs boson (see Figure 2.3 c)). This leaves an experimental signature of two kinematically-distinct jets from the original incident quarks and the Higgs signature. For Run 2, the VBF cross section of  $3.78 \pm 0.08$  pb is more than an order of magnitude lower than ggF, however the channel is more sensitive due to the experimental signature. The distinctive VBF jets (see Figure 5.1) provide the ability to effectively tag signal events while rejecting backgrounds.

### 5.3. First look at Run 2

The first consideration of the full Run 2 data set resulted in a CONF note presented at Moriond 2020 [25]. This chapter contains material coauthored with the ATLAS collaboration. I conducted studies to optimize object definition and event selection and refined legacy code to calculate and illustrate the result, among

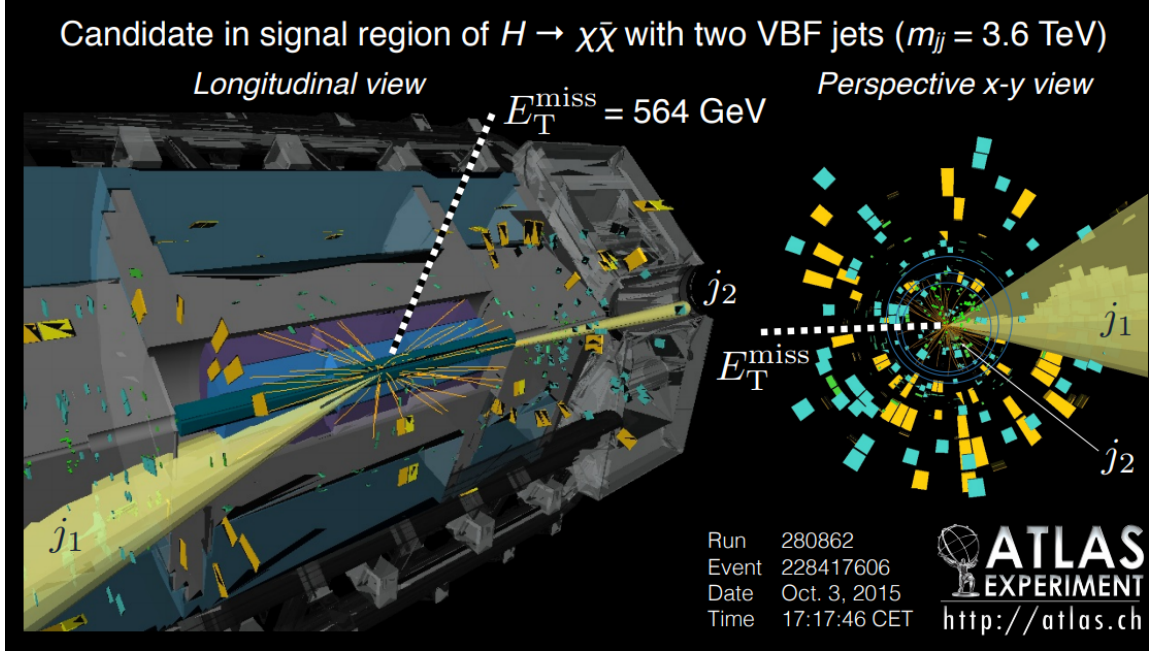


FIGURE 5.1. Candidate  $h \rightarrow \text{inv.}$  event from Run 2

smaller contributions. Other members of the ATLAS collaboration contributed to the analysis through further optimization studies, background modeling technique development, and systematic uncertainty derivation and quantification. In combination, this all informed the final result.

### 5.3.1. Overview and Analysis Strategy

The desired signature requires no leptons, two well-separated jets with a large dijet mass, and significant MET (referred to as a VBF+MET search, indicative of the signature). A signal region (SR) is defined through kinematic cuts that is enriched with this signal. However, a number of background processes share this signature and contaminate the SR. The leading background is  $V$ +jets events - VBF  $Z \rightarrow \nu\nu$  or VBF  $W \rightarrow \ell\nu$  where the lepton is lost<sup>2</sup> (see Figure 5.2).

<sup>2</sup>Leptons are largely lost if they fail object definitions, especially  $p_T$  or  $\eta$  requirements.



There are multiple production mechanisms for these  $V$ +jets backgrounds that can generally be categorized as electroweak (EWK) or strong based upon the number of vertices with EWK couplings (four or two, respectively). Strong and EWK cross sections scale with boson  $p_T$  (PTV). The EWK contribution has the same kinematics as the  $h \rightarrow \text{inv.}$  signature, and the high collision energy of Run 2 increased the cross section of the strong contribution causing it to have a dominant effect.

These backgrounds are measured with control regions (CRs) enriched in these processes with visible leptons (VBF  $Z \rightarrow \ell\ell$  or VBF  $W \rightarrow \ell\nu$ ). Two CRs are thus defined - a one-lepton region enriched with  $W$ +jets events (WCR) and a two-lepton region enriched with  $Z$ +jets events (ZCR). The contribution in CRs as measured with MC is extrapolated through lepton universality to estimate the invisible background contribution to the SR in data (see Section 5.3.4.1).

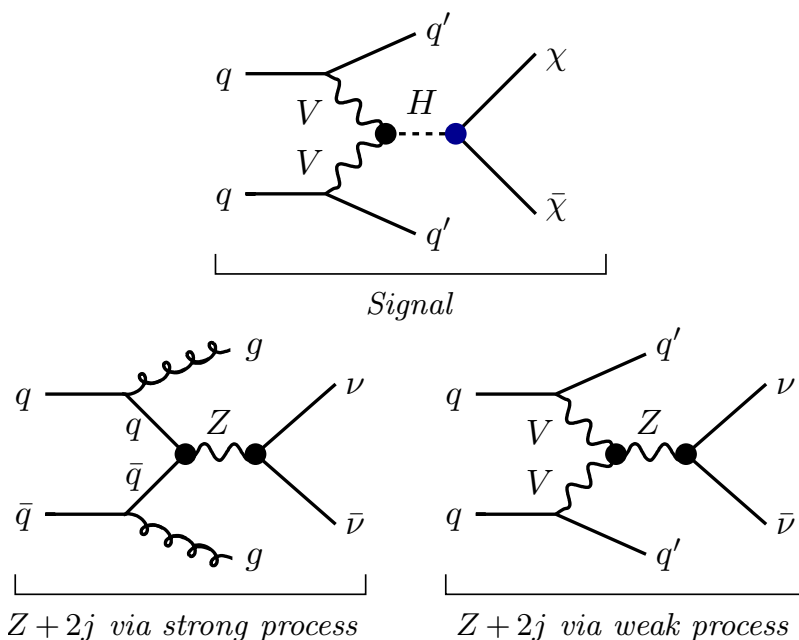


FIGURE 5.2. Feynman diagrams of the signal process (top) and example  $Z$ +jets irreducible backgrounds

Soft QCD interactions resulting in jets are another background referred to as the QCD multijet (MJ) background. This accounts for roughly 5% of background events. Additional small contributions, together accounting for 1% of all backgrounds (see Table 5.2), include  $t\bar{t}$  and multi-boson ( $VV$ ,  $VVV$ ) processes.

| Process    | Contribution<br>[% of total bkg] |
|------------|----------------------------------|
| $V$ +jets  | 94.2                             |
| $Z$ strong | 44.0                             |
| $Z$ EWK    | 17.1                             |
| $W$ strong | 24.2                             |
| $W$ EWK    | 8.9                              |
| Multijet   | 4.8                              |
| Others     | 1.0                              |

TABLE 5.2. Percent contribution to background events in SR

All these background sources must be understood with the help of MC samples. Understanding the background contribution within the data then allows for a search of signal-like events. The SM hypothesis can then be tested to judge whether BSM influence is observed.

#### 5.3.1.1. Data and Monte Carlo Samples

This search utilizes  $138.9 \text{ fb}^{-1}$  of  $\sqrt{s}=13 \text{ TeV}$  p-p collision data, collected from 2015-2018 with MET [123], single-, and double- lepton triggers. In all cases, the lowest unprescaled trigger available at the time was used. This results in data sets with differing MET thresholds. Data runs are only considered if they are recorded during stable beam conditions with all ATLAS subsystems fully operational. 12-28% of the efficiency of identifying muons is recovered by considering both single muon triggers and a MET trigger for the WCR.

MC samples simulating all signal and backgrounds processes are utilized in the search (see Table 5.3) with special attention paid to  $V$ +jets.  $V$ +jets strong and EWK processes are modeled separately. Strong  $V$ +jets samples are simulated at NLO (LO) in  $\alpha_s$  for up to two (four) final state partons. EWK  $V$ +jets samples are further broken into diboson and VBF diagrams and simulated separately with interference assumed to be negligible. VBF channel MEs are calculated at NLO in  $\alpha_s$ , while diboson channels are calculated to LO and reweighted to NLO cross sections. Interference between the strong and EWK modes introduce an  $\alpha_{\text{EW}}^3$  term which is modeled independently.

| Process                                             | Generator                                                | ME Order                                       | PDF               | Parton Shower      | Tune   |
|-----------------------------------------------------|----------------------------------------------------------|------------------------------------------------|-------------------|--------------------|--------|
| Strong $V$ +jets                                    | SHERPA v2.2.1,<br>SHERPA v2.2.7<br>( $m_{jj}$ -filtered) | NLO (up to<br>2-jets),<br>LO (up to<br>4-jets) | NNPDF3.0nnlo      | SHERPA<br>MEPS@NLO | SHERPA |
| Electroweak $V$ +jets                               | SHERPA v2.2.1<br>reweighted by<br>HERWIG 7 in $m_{jj}$   | LO (up to<br>4-jets), NLO<br>(reweighting)     | NNPDF3.0nnlo      | SHERPA<br>MEPS@LO  | SHERPA |
| $VV$ +jets<br>(including $gg \rightarrow VV$ +jets) | SHERPA v2.2.1 or<br>SHERPA v2.2.2                        | NLO (up to<br>1-jet), LO<br>(up to 3-jets)     | NNPDF3.0nnlo      | SHERPA<br>MEPS@NLO | SHERPA |
| Electroweak $VV$ +jets                              | SHERPA v2.2.1 or<br>SHERPA v2.2.2                        | LO                                             | NNPDF3.0nnlo      | SHERPA<br>MEPS@LO  | SHERPA |
| $V$ +jets $\alpha_{\text{EW}}^3$<br>interference    | MADGRAPH5_aMC@NLO                                        | LO                                             | PDF4LHC15         | PYTHIA8            |        |
| $t\bar{t}$                                          | POWHEGBOX v2                                             | NLO                                            | NNPDF3.0nnlo      | PYTHIA8            | A14    |
| QCD multijet                                        | PYTHIA 8.230                                             | LO                                             | NNPDF2.31o        | PYTHIA8            | A14    |
| ggF Higgs                                           | POWHEG NNLOPS                                            | NNLO                                           | PDF4LHC15<br>NNLO | PYTHIA8            | AZNLO  |
| VBF Higgs                                           | POWHEG                                                   | NLO                                            | PDF4LHC15         | PYTHIA8            | AZNLO  |
| $VH$ Higgs                                          | POWHEGBOX v2                                             | NLO                                            | PDF4LHC15         | PYTHIA8            | AZNLO  |

TABLE 5.3. Summary of event generators used for simulation

### 5.3.1.2. Object and Event Selection

Jets are reconstructed with the anti- $k_t$  algorithm using radius parameter  $R = 0.4$ . Prior to this work, EMTopo jets were used. A switch to PFlow jets was encouraged for ease of eventual combination with other analysis channels, but at the time fJVT had not been fully calibrated for PFlow jets yet. I further noticed modeling irregularities (an excess of low- $p_T$  jets in forward regions) in the EMTopo jet collection. I found that this was likely due to a change in calorimeter calibration settings in Run 355258, as the effect is pronounced in runs that chronologically follow this run. Since PFlow jets use tracker information as well as calorimeter information, the effect is smaller in the PFlow collection (see Figure 5.3).

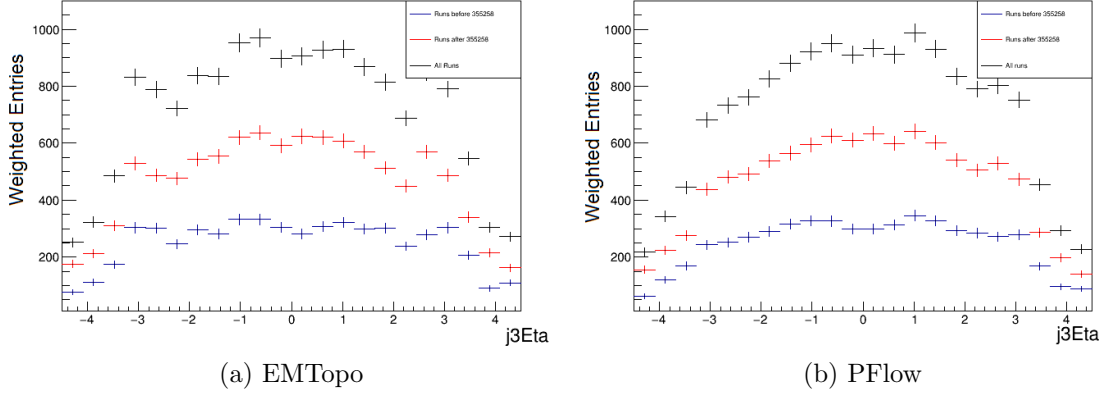


FIGURE 5.3.  $\eta$  value of the third highest  $p_T$  jet in events compared between a) EMTopo and b) PFlow jet collections

The PFlow jet collection therefore is used with fJVT values copied from EMTopo jets. For fJVT assignment, PFlow jets must have  $p_T > 50$  GeV and fall within  $\Delta R < 0.3$  of the associated EMTopo jet. This matching is more than 99% efficient in analysis regions.

Jets are considered for analysis provided that  $p_T > 25$  GeV,  $|\eta| < 4.5$ , and the Medium JVT WP<sup>3</sup> is satisfied. The two leading jets must satisfy Loose or Very Tight fJVT WPs<sup>4</sup> depending on the  $E_T^{\text{miss}}$  contained in the event.

Jets containing  $b$ -hadrons are identified with the multivariate discriminant MV2c10 algorithm. A 77% WP is used.

Events with isolated photons with  $p_T > 20$  GeV within  $|\eta| < 2.37$  are vetoed. Leptons are used both to populate CRs and to veto SR events. Two different reconstruction criteria are used for these two requirements. First, loose ‘baseline’ leptons are defined requiring  $p_T > 4.5(4)$  GeV within  $|\eta| < 2.47(2.7)$  for electrons (muons) with no isolation requirement. If a lepton candidate as such is identified, the event is rejected from signal selection. Events populating the lepton CRs have stricter lepton requirements. A ‘signal’ lepton considered for a CR must be associated with the PV, have  $p_T > 30$  GeV, fulfill tighter identification, and pass a loose isolation requirement.

In order to suppress the contribution of fake electron reconstruction (where a jet from a MJ event is reconstructed as an electron) to the WCR, a requirement is made on MET significance  $\mathcal{S}_{\text{MET}}$ , where

$$\mathcal{S}_{\text{MET}} = \frac{E_T^{\text{miss}}}{\sqrt{p_T(j_0) + p_T(j_1) + p_T(e)}} .$$

If a jet was misidentified as an electron, the resulting  $E_T^{\text{miss}}$  would be inappropriately large. A minimum threshold on this requirement thus rejects such

---

<sup>3</sup>If  $p_T < 60$  GeV and  $|\eta| < 2.4$ , require  $\text{JVT} > 0.2$ . This carries a HS selection efficiency of 98%.

<sup>4</sup>Loose fJVT requires  $\text{fJVT} < 0.5$ . Very Tight fJVT requires  $\text{fJVT} < 0.2$ , a WP developed for this analysis to reject events with  $E_T^{\text{miss}} \leq 180$  GeV.

events in the CR. Events failing the  $\mathcal{S}_{\text{MET}}$  requirement are saved and considered separately in a “fake- $e$ ” CR enriched in events with false electrons.

To avoid considering the same energy deposit in the reconstruction of multiple physics objects, a selection of criteria is defined to remove such overlapping objects (see Appendix E).

Kinematic requirements are made on these objects (see Table 5.4) to define the SR and CRs. Since the CRs are meant to utilize visible leptons to extrapolate their effect with undetected leptons,  $E_{\text{T}}^{\text{miss}}$  in the CRs is calculated treating the visible leptons as invisible particles<sup>5</sup>.

| Cut                                    | SR                       | WCR                                               | ZCR                |
|----------------------------------------|--------------------------|---------------------------------------------------|--------------------|
| $N_\ell$                               | 0                        | 1                                                 | 2                  |
| $p_{\text{T}}(\ell_0)$                 | -                        | $> 30 \text{ GeV}$                                |                    |
| $N_\gamma$                             | 0                        |                                                   |                    |
| $N_{\text{jet}}$                       | 2,3,4                    |                                                   |                    |
| $N_{b \text{ jet}}$                    | $< 2$                    |                                                   |                    |
| $E_{\text{T}}^{\text{miss}}$           | $> 200 \text{ GeV}$      |                                                   |                    |
| $E_{\text{T}}^{\text{miss}}$ soft term | $< 20 \text{ GeV}$       |                                                   |                    |
| $E_{\text{T}}^{\text{jet, no JVT}}$    | $> 180 \text{ GeV}$      |                                                   |                    |
| $ m_{\ell\ell} - m_Z $                 | -                        | -                                                 | $< 25 \text{ GeV}$ |
| $E_{\text{T}}^{\text{miss}}$ (nolep)   | -                        | -                                                 | $< 70 \text{ GeV}$ |
| $\mathcal{S}_{\text{MET}}$             | -                        | $> 4 \sqrt{\text{GeV}} \text{ (} e \text{ only)}$ | -                  |
| VBF jets:                              |                          |                                                   |                    |
| $p_{\text{T}}(j_0, j_1)$               | $> (80, 50) \text{ GeV}$ |                                                   |                    |
| $\Delta\phi_{jj}$                      | $< 2$                    |                                                   |                    |
| $\Delta\eta_{jj}$                      | $> 3.8$                  |                                                   |                    |
| $\eta_0 \cdot \eta_1$                  | $< 0$                    |                                                   |                    |
| $m_{jj}$                               | $> 0.8 \text{ TeV}$      |                                                   |                    |
| If $N_{\text{jet}} > 2$ :              |                          |                                                   |                    |
| $p_{\text{T}}(j_2, j_3)$               | $> 25 \text{ GeV}$       |                                                   |                    |
| Centrality $C_i$                       | $< 0.6$                  |                                                   |                    |
| $m_{rel}^i$                            | $< 0.05$                 |                                                   |                    |

TABLE 5.4. Signal and control region definitions

---

<sup>5</sup>Leptons are included in the  $E_{\text{T}}^{\text{miss}}$  calculation, as usual, for  $E_{\text{T}}^{\text{miss}}$  (nolep)

The SR is designed to isolate VBF  $h \rightarrow \text{inv.}$  events, but additional Higgs events from ggF production also populate the SR. These events are considered to be signal, and account for 12.5% of signal events.

Though exactly two jets are expected, ISR or FSR can contribute additional jets to the event. This is how a ggF event may fall within the SR. A third jet veto is effective at reducing background rates from MJ and strong  $V+\text{jets}$  processes, but some signal is also sacrificed. Unlike previous searches for this signature, this work allows for up to two ISR and FSR jets on the condition that these extra jets meet additional criteria.

Little radiation is expected between the leading VBF jets, so any additional jets must fall close to a VBF jet. This is quantified through centrality,

$$C_i = \exp \left( -\frac{4}{(\eta_0 - \eta_1)^2} \left( \eta_i - \frac{\eta_0 + \eta_1}{2} \right)^2 \right) ,$$

where the  $\eta$  of the third or fourth jet ( $i$ ) is compared to the leading two jets (0 and 1).  $C_i = 1$  when jet  $i$  is central between the leading jets, and  $C_i = 1/e$  if jet  $i$  is aligned with a leading jet. Thus an upper limit on  $C_i$  is enforced at 0.6.

Contributions from FSR are more likely for this signature. This is selected for by comparing the relative dijet mass of the third or fourth jet  $i$  to the dijet mass of the leading jets,

$$m_{rel}^i = \frac{\min\{m_{0,i}m_{1,i}\}}{m_{0,1}} .$$

The events that population the SR and CRs are further binned orthogonally based upon kinematic variables (see Figure 5.4). Events with three or four jets are all included in one bin, the “multijet bin”. Roughly 18% of  $h \rightarrow \text{inv.}$  signal falls in this region. The remaining events necessarily contain  $N_{\text{jet}} == 2$  which is broken

into five bins of  $m_{jj}$  and two bins of  $\Delta\phi_{jj}$ . The most sensitive bins fall at low  $\Delta\phi_{jj}$  and high  $m_{jj}$ , as indicated by the intensity of grayscale in Figure 5.4).

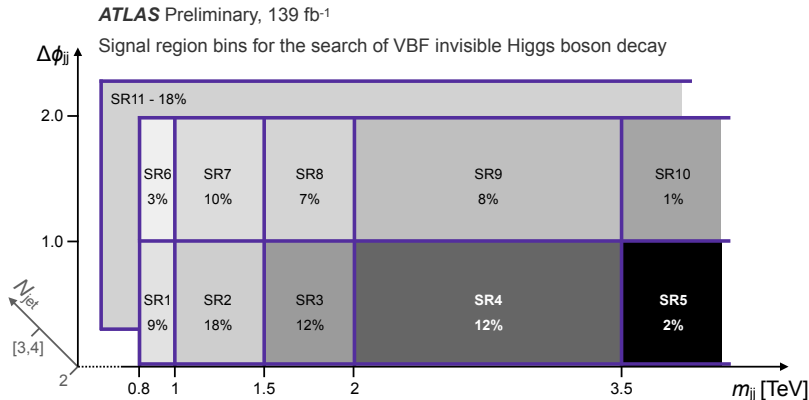


FIGURE 5.4. Binning scheme for VBF+MET first look. Percentage values reflect the fraction of signal events in each bin, and increasing intensity of grayscale reflects increasing bin sensitivity [25]

### 5.3.2. Quark/Gluon Tagging

The VBF signature characteristic of this search involves two forward quark-initiated jets. The strong  $V$ +jets background, however, may include gluon-initiated jets (see Figure 5.2). I conducted a study to determine whether including additional requirements on jet initiating parton could further reject background in the SR.

Though difficult to distinguish in practice, different variables have been identified and studied that can provide some level of discrimination between quark- and gluon-initiated jets [124, 125]. The most successful of these, the number of tracks contained within a jet, is used as the basis of an ATLAS CP BoostedJetTagging tool, JetQGTagger [126]. Due to differing Casimir factors ( $C_F = 3$  for gluons,  $C_A = 4/3$  for quarks), gluons are more likely to radiate soft particles and create overall wider jets in the tracker and calorimeter. Due to this,



gluon-initiated jets are expected to have a larger number of tracks than quark-initiated jets. This effect is seen for simulated VBF Higgs and  $Z$  strong events (see Figure 5.5). The discrimination is stronger as jet  $p_T$  increases.

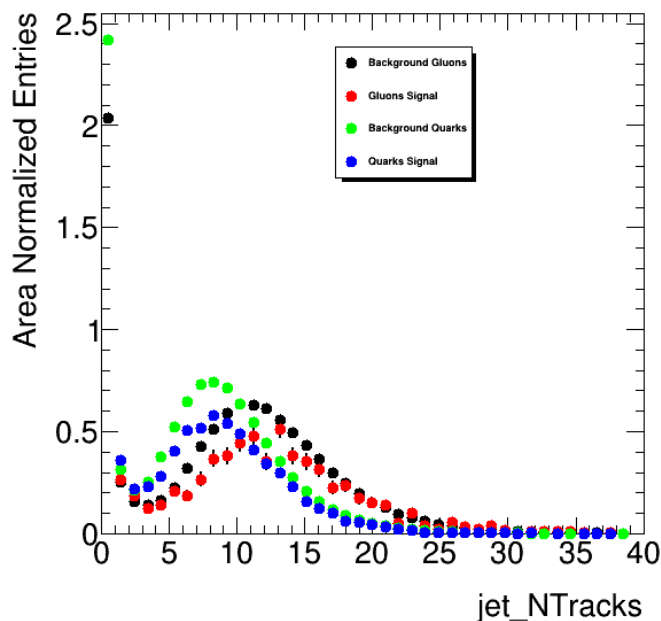


FIGURE 5.5. Gluon-initiated jets have larger values of NTracks than quark-initiated jets, as shown in VBF Higgs and  $Z$  strong simulated events

Truth level studies (see Figure 5.6) investigated potential gain from a “perfect tagger” by considering simulated parton ID. An “expected limit” is defined as

$$\sigma_\mu Z = 2\sqrt{N_B + (N_Z * N_{ZCR})^2 + (N_W * 0.015)^2}/N_S ,$$

where  $N$  is the number of signal ( $S$ ) or background ( $B$ ) events in the bin. This is normalized by the number of events in the ZCR ( $N_{ZCR}$ ) and the number of  $Z$  and  $W$  events in the bin of interest. With an additional SR requirement of that both VBF jets be quark-initiated, a relative 5.5% improvement of  $\sigma_\mu Z$  is seen. However, the discriminating variable used in the tagger requires information from the tracker

which is limited to only the central barrel region of ATLAS. With perfect tagging in this region and no quark-gluon tagging information assumed in the forward region, the relative improvement decreases to 0.6% over the nominal  $\sigma_\mu Z$ .

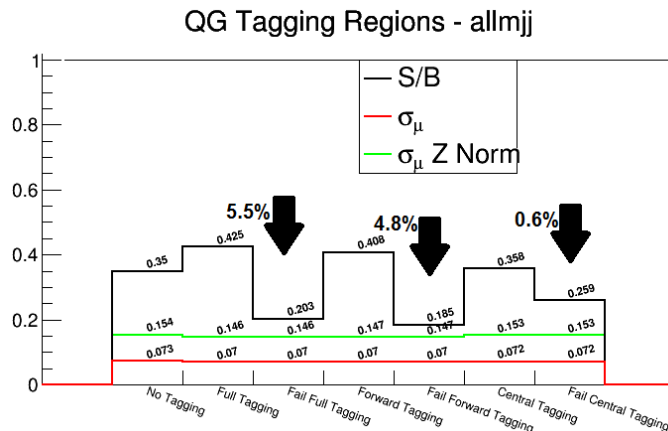


FIGURE 5.6. Using truth information regarding jet-initiating partons emulates a perfect tagger. With perfect full detector tagging, the expected limit ( $\sigma_\mu Z$ ) improves by a relative 5.5%.

The centralized tagging tool assigns “quark-like” or “gluon-like” tags to jets based upon predefined working points and utilizing a threshold cut on the number of tracks variable. The optimum threshold cut value was found to be 5 for our samples, as this value minimized  $\sigma_\mu Z$  (see Fig. 5.7). In order to retain statistics, “pass” and “fail” regions were considered as two separate bins rather than requiring a hard cut. The threshold requirement is made on both leading and subleading jets (the VBF tagged jets), as this increased performance.

The expected improvement is small due to the kinematics of the expected signal. This quark/gluon discrimination technique is limited in scope for this search for the following reasons:

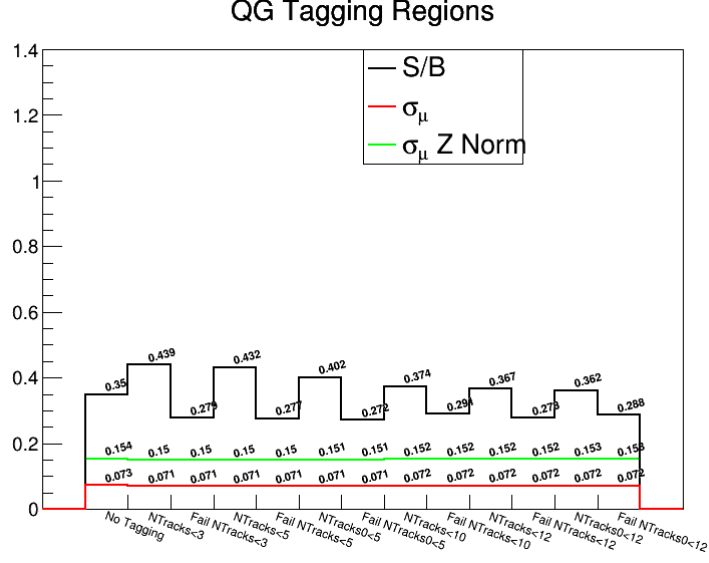


FIGURE 5.7.  $S/B$  yields and associated expected limit ( $\sigma_\mu$  Z Norm) with different quark/gluon tagging related cuts

1. **NTracks** is a tracker-based variable, and many VBF jets are forward.

Forward jets therefore cannot be tagged in this way and pass the requirement by default (**NTracks**==0 by definition).

2. VBF jets have a large dijet transverse mass, but low individual momenta (see Fig. 5.8). The JetQGTagger has minimum total systematic uncertainties from 200-800 GeV, and performance decreases as jet  $p_T$  decreases [126].

With the  $36.1 \text{ fb}^{-1}$   $m_{jj}$ -only signal region binning scheme (with bin boundaries of [1.0, 1.5, 2.0, +] TeV) and  $36.1 \text{ fb}^{-1}$  blinded data set, the true expected nominal limit from HistFitter was 0.165. To consider quark/gluon tagging, each of these three bins was further split similarly to the pass/fail scheme of Fig. 5.7. The limit from this six-bin fit was 0.166, showing that no significant improvement is gained from the inclusion of quark/gluon tagging information (see Figure 5.9).

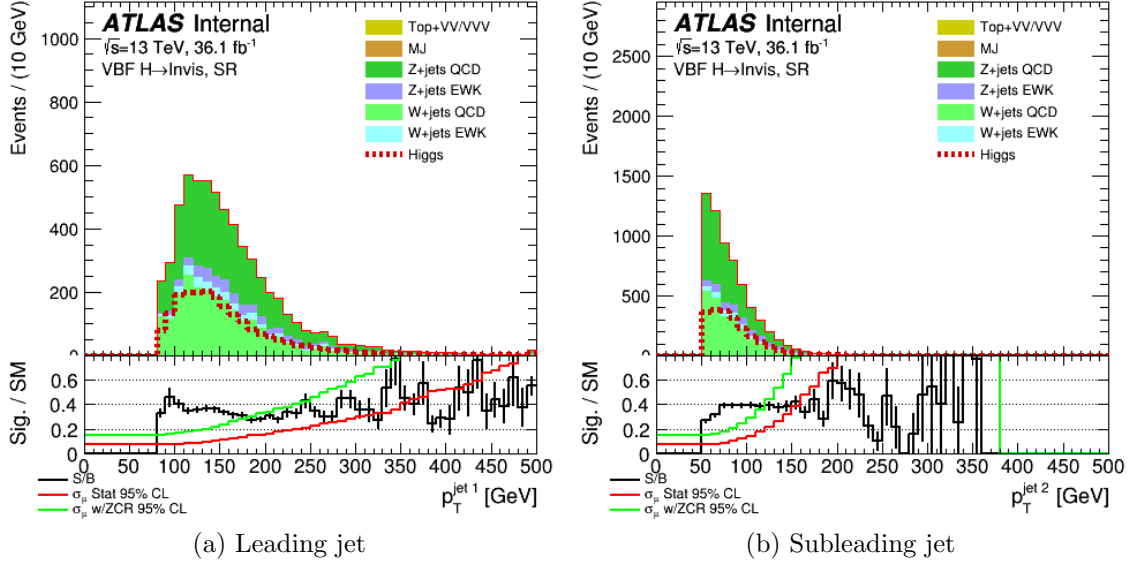


FIGURE 5.8. The  $p_T$  distribution of a) leading and b) subleading jets in the signal region.

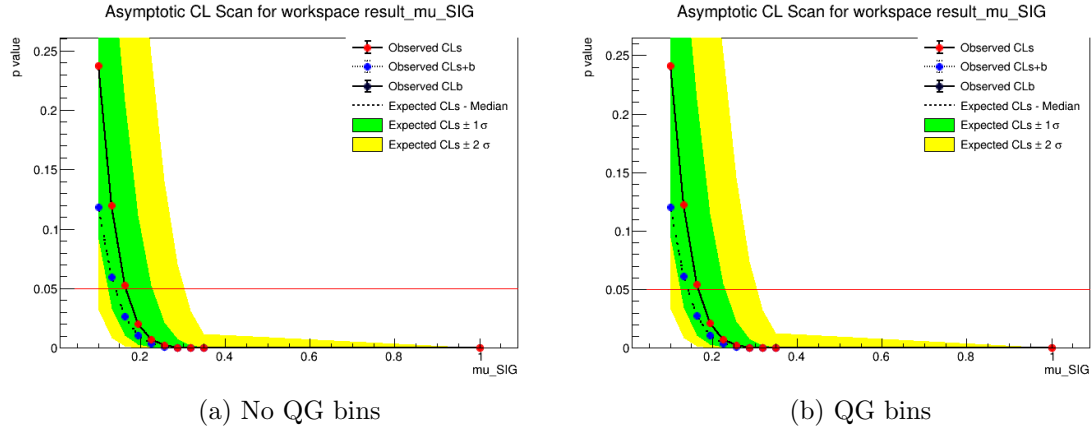


FIGURE 5.9.  $CL_S$  scan values over the signal strength  $\mu$  for  $36.1 \text{ fb}^{-1}$  of blinded data using an  $m_{jj}$ -only binning scheme a) before and b) after additional quark-tagged binning.

### 5.3.3. VBF Trigger and MJ Studies

A new trigger designed to target the VBF signature was introduced into the Run 2 menu in 2018. I used the output of this trigger and its simulated

performance to consider the MJ background of the VBF+MET search. The MJ background is estimated through data-driven techniques, as it mainly enters the SR due to mismeasurement or pileup features, so its behavior is not well understood in the SR. I created a new MJ sample, enriched with events in the VBF phase space, and used this sample to validate the  $E_T^{\text{jet, no JVT}}$  cut which is intended to preferentially eliminate MJ.

### 5.3.3.1. VBF Trigger

The MET trigger is only fully efficient beyond  $\sim 180$  GeV. In order to recover some lower-MET events, VBF-targeting triggers were implemented online in the 2018 menu and collected  $\sim 42 \text{ fb}^{-1}$  of data. These triggers,

- HLT\_j70\_j50\_0eta490\_inv1000j50\_dphi24\_xe90\_pufit\_xe50\_L1MJJ-500-NFF and
- HLT\_j70\_j50\_0eta490\_inv1100j70\_dphi20\_deta40\_L1MJJ-500-NFF ,

are seeded by L1 events with two jets with an online calculate dijet mass (MJJ) exceeding 500 GeV, where at least one jet was central. This central requirement reduces the amount of contaminating multijet events. Since the trigger focuses on the VBF jets, it is of more general use than only VBF+MET.

At the HLT, additional online<sup>6</sup> jet selections are made (see Table 5.5). The second listed trigger was considered for inclusion in this analysis. As it supplements signal events from the standard MET triggers but at lower MET values, a validation region with loosened  $100 < E_T^{\text{miss}} < 150$  GeV and  $E_T^{\text{jet, no JVT}}$  requirements was considered. The signal to background ratio was low in

---

<sup>6</sup>Online HLT measurements are made with different (faster) algorithms than the more robust offline reconstruction. Therefore, online and offline measurements are not expected to agree.

this region, leading to a relative 1% increase in search sensitivity before inclusion of the MJ background. This is partially due to the  $\Delta\phi_{jj}$  distribution peaking near high values of  $\Delta\phi_{jj}$  indicating many background-like multijet events are still captured in the trigger (see Figure 5.10). This serves as a limitation for the use of the trigger in the search, and it was abandoned.

|                    | HLT Selection  | Offline Plateau selection |
|--------------------|----------------|---------------------------|
| $p_T(j_0, j_1)$    | 70, 50 GeV     | 90, 80 GeV                |
| $ \eta (j_0, j_1)$ | $< (4.9, 4.9)$ | $< (3.2, 4.9)$            |
| $m_{jj}$           | 1.1 TeV        | 1.3 TeV                   |
| $\Delta\phi_{jj}$  | $< 2$          | $< 2$                     |
| $\Delta\eta_{jj}$  | $> 4$          | $> 4$                     |

TABLE 5.5. HLT trigger requirements for the VBF trigger, and additional offline plateau selection requirements for fully-efficient trigger emulation

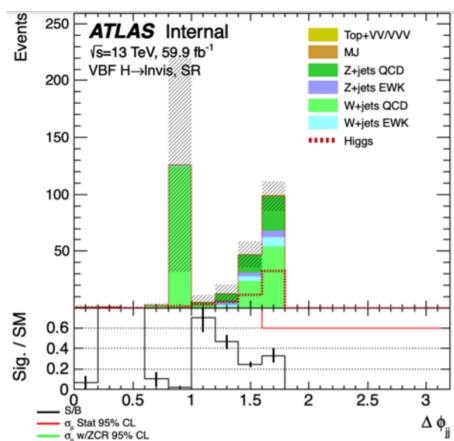


FIGURE 5.10. Event distribution in SR-like validation region ( $100 < 150$  GeV) from VBF trigger

$< E_T^{\text{miss}}$

### 5.3.3.2. Multijet MC Sample

The extreme  $m_{jj}/\Delta\phi_{jj}$  phase space targeted by this search is effective at removing multijet - especially from MC. In previous studies, no MC multijet events propagate to the SR. This leads to complications in understanding how this

background behaves in this region. Its presence in VBF-triggered events provides an opportunity to consider its impact. In order to overcome this limitation in our understanding, I created a Monte Carlo QCD sample with the intention of understanding this background and investigating the power of our SR cuts (such as the cut on  $E_T^{\text{jet, no JVT}}$ , as described in Appendix 4.2.4.1) in reducing the background.

To understand the nominal QCD behavior, a collection of EMTopo dijet events with low- $p_T$  leading jets was considered. These events are known to contain pileup events that cause the reconstructed jet  $p_T$  spectrum to be not smooth (after weighting) so a pileup filter was implemented (as suggested by J. Lacey<sup>7</sup>) for smoothing. MC events were only considered if they passed the quality cut,

$$0.6 < \frac{p_T^{\text{reco}}(j_0) + p_T^{\text{reco}}(j_1)}{2 * p_T^{\text{truth}}(j_0)} < 1.4 , \quad (5.1)$$

given at least two jets without consideration of JVT. Events with two reconstructed jets passing the medium JVT working point and the MC quality requirement of Eqn. 5.1 are considered.

From this collection, the two leading  $p_T$  jets (tagged as VBF jets) were flagged as originating from the hard scatter or from pileup based upon a truth-matching technique. In order to be considered HS, the JVT-passing reconstructed jet must fall within  $\Delta R < 0.15$  of a true jet. The two-reconstructed-jet event collection is then broken into three categories: 2PU, where both jets are flagged as PU, 2HS where neither jet is flagged as PU, and 1HS1PU with one flagged PU jet.

---

<sup>7</sup>[https://twiki.cern.ch/twiki/pub/AtlasProtected/JetStudies2012/JLacey\\_MC12\\_JZXW\\_weights\\_and\\_CleaningCuts.pdf](https://twiki.cern.ch/twiki/pub/AtlasProtected/JetStudies2012/JLacey_MC12_JZXW_weights_and_CleaningCuts.pdf)

2HS jet events tend to have large  $\Delta\phi_{jj}$  values. Events with PU jets contaminate the low- $\Delta\phi_{jj}$  and high- $m_{jj}$  signal region the most (see Figure 5.11).

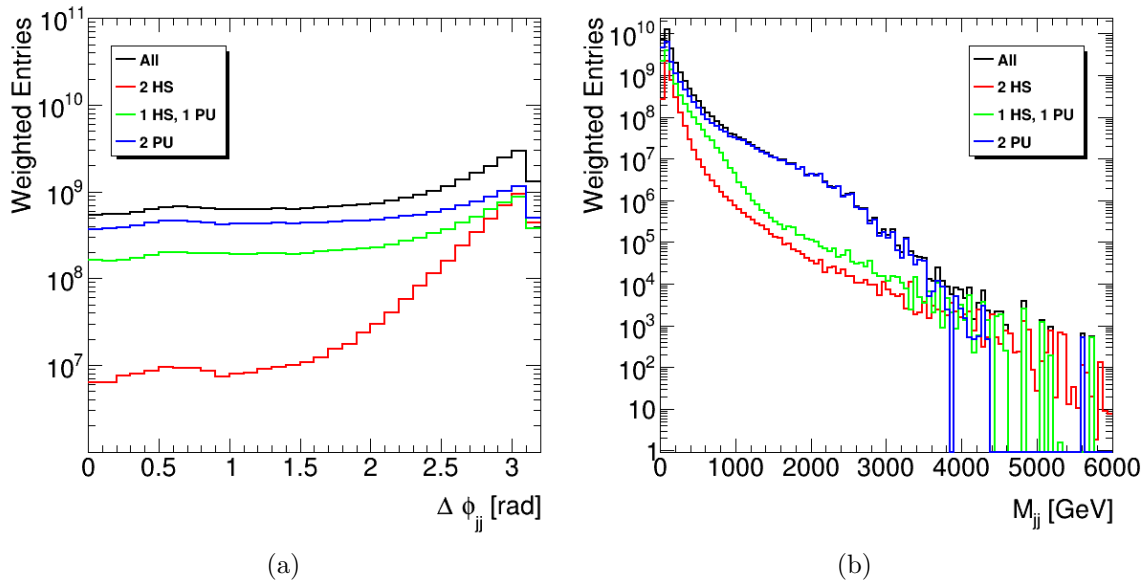


FIGURE 5.11. Distributions of two-jet events broken into categories based upon number of contaminating PU jets.

Implementing a  $E_T^{\text{miss}}$  cut of  $E_T^{\text{miss}} > 150$  GeV and minimum jet  $p_T$  requirements where  $p_T(j_0, j_1) > 80, 50$  GeV more closely resembles the VBF+MET SR. The resulting MJ distribution contains no remaining events in the SR ( $\Delta\phi_{jj} < 2$  and  $m_{jj} > 800$  GeV) (see Figure 5.12). This motivates the need for a designated QCD sample with filtered events that populate the signal region.

Reconstruction of dijet variables can cause modeling discrepancies when compared to truth quantities. Five regions were defined (see Figure 5.13) in the  $m_{jj}/\Delta\phi_{jj}$  plane. A pseudo-SR, Region 0, falls at large  $m_{jj}$  and low  $\Delta\phi_{jj}$ . Validation Regions 1, 2, and 3 border this range down to  $m_{jj}=0.8$  TeV, while Region 4 serves as an underflow bin for low- $m_{jj}$ .



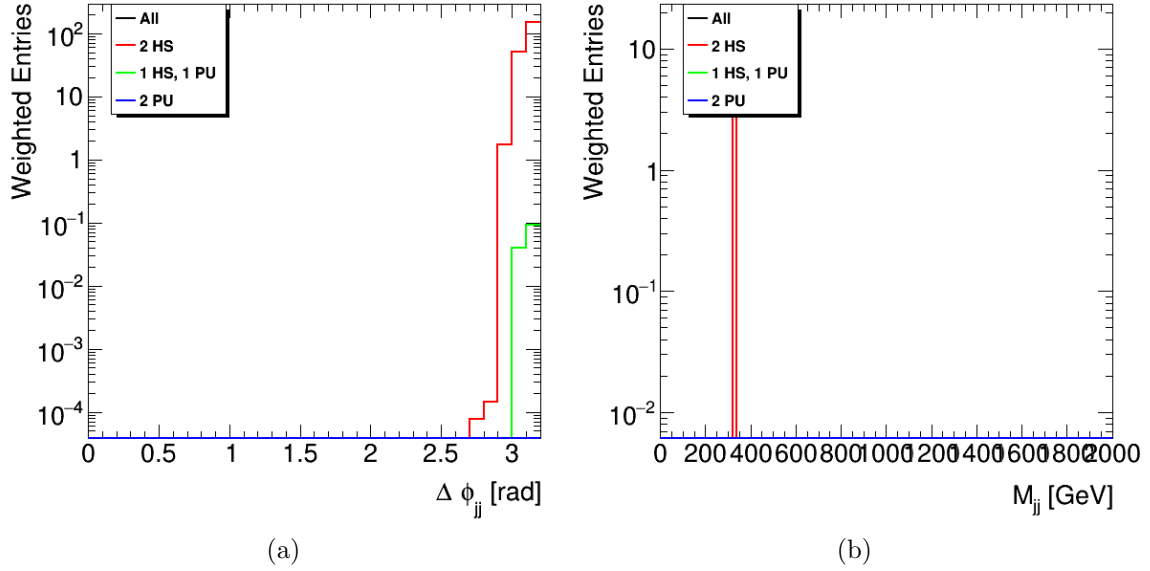


FIGURE 5.12. Distributions of two-jet events after requiring  $E_T^{\text{miss}} > 150$  GeV and  $p_T(j_0, j_1) > 80, 50$  GeV

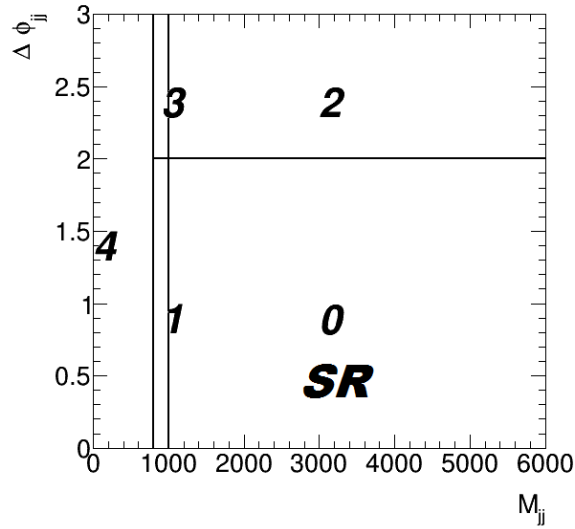


FIGURE 5.13. Regions of interest in the  $m_{jj}/\Delta\phi_{jj}$  plane

Some events migrate within the  $m_{jj}/\Delta\phi_{jj}$  plane from truth-level to after reconstruction for all events with exactly two jets (see Table 5.6). This rate of migration was found to be tolerable, and a truth-level afterburner filter was

implemented in order to isolate and preferentially reconstruct these events that fall within Region 0, requiring truth-level  $m_{jj} > 1$  TeV and  $\Delta\phi_{jj} < 2.0$ .

| [%]     | Reco 0 | Reco 1 | Reco 2 | Reco 3 | Reco 4 |
|---------|--------|--------|--------|--------|--------|
| Truth 0 | 52.99  | 5.98   | 7.57   | 0.40   | 33.07  |
| Truth 1 | 10.11  | 33.7   | 3.37   | 1.69   | 51.12  |
| Truth 2 | 3e-3   | 2e-3   | 99.20  | 0.66   | 0.14   |
| Truth 3 | 0.03   | 0.01   | 9.20   | 76.46  | 14.30  |
| Truth 4 | 0.04   | 0.03   | 0.04   | 0.22   | 99.67  |

TABLE 5.6. Percent of events from truth level that fall within the given region after reconstruction

This afterburner filtered file was merged with a complete nominal QCD collection. To retain orthogonality, only nominal events that fail the truth-level filter ( $m_{jj} > 1$  TeV and  $\Delta\phi_{jj} < 2.0$ ) are included. Weights ( $W$ ) are recalculated for each nominal slice as

$$W = \frac{\sigma\epsilon Lw}{\sum_i^N w_i} ,$$

where  $\sigma$  is the process cross section,  $\epsilon$  is the efficiency,  $L$  is the luminosity,  $w$  is the MC event weight of the event, and the normalization factor  $\sum_i^N w_i$  is the sum of all MC event weights from all  $N$  appropriately passing (if filtered file) or failing (if nominal) events. Differences in generators between the afterburner filtered file and the nominal QCD collection were corrected by renormalizing the truth-level  $m_{jj}$  distribution of the filtered file with respect to the nominal QCD sample (see Figure 5.14).

The full distributions after reconstruction are not smooth after the filter cuts, but are smooth (as expected) after the MC quality cut of Eqn. 5.1. The nominal collection yields  $0.975 \pm 0.388$  signal region events, while the addition of the filtered file increases the yield to  $4.02 \pm 3.02$  events (see Table 5.7). Though events now populate the SR, the statistics in this region are still too low for thorough

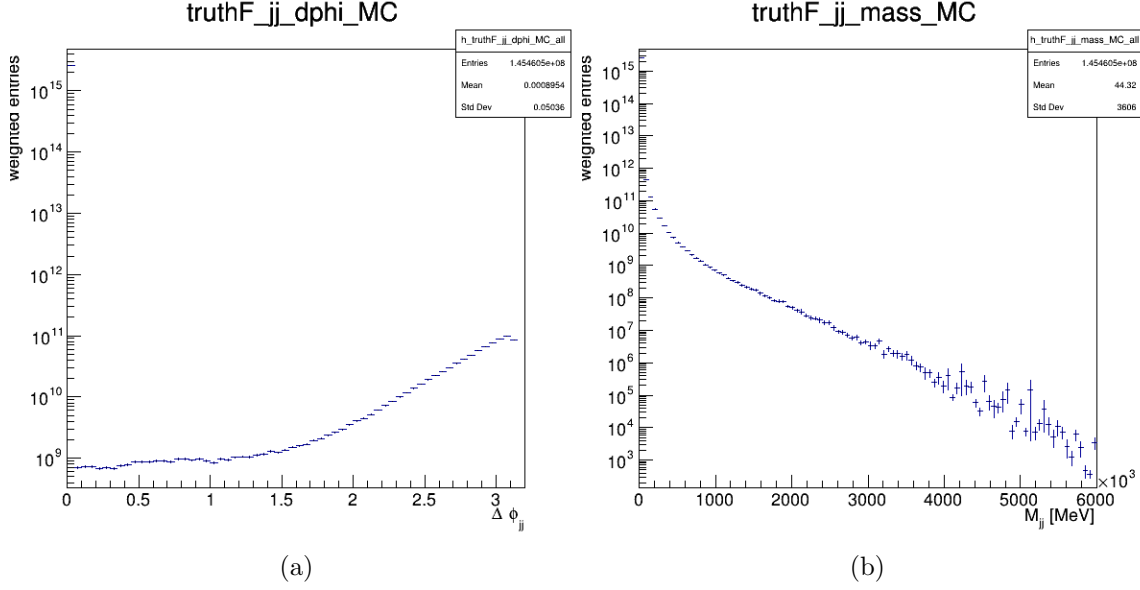


FIGURE 5.14. Full truth  $\Delta\phi_{jj}$  and  $m_{jj}$  distributions of the merged nominal QCD+afterburner filtered file collection after reweighting, merging, and renormalizing the filtered file

investigation. Instead, the sample is used for checks on the performance of the VBF trigger and impact of cuts on the QCD background.

| Cut                                | Integral $\pm$ Error    |
|------------------------------------|-------------------------|
| -                                  | 2.545e15 $\pm$ 2.812e12 |
| MC Filter                          | 1.394e12 $\pm$ 1.506e9  |
| Trigger                            | 2.284e8 $\pm$ 2.186e7   |
| Lead/sublead $p_T$                 | 2.921e7 $\pm$ 7.276e6   |
| njet==2                            | 7.624e6 $\pm$ 3.92e6    |
| $ \Delta\eta_{jj} $                | 7.470e6 $\pm$ 3.924e6   |
| $\Delta\phi(E_T^{\text{miss}}, j)$ | 3.191e6 $\pm$ 2.861e6   |
| $\eta_0 \times \eta_1$             | 4.808e4 $\pm$ 1.031e4   |
| $m_{jj}$                           | 449.6 $\pm$ 51.27       |
| $\Delta\phi_{jj}$                  | 17.10 $\pm$ 10.30       |
| $E_T^{\text{jet, no JVT}}$         | 16.89 $\pm$ 10.30       |
| soft MET term                      | 16.28 $\pm$ 10.30       |
| $E_T^{\text{miss}} > 150$ GeV      | 4.02 $\pm$ 3.02         |
| Full SR - nominal only             | 0.975 $\pm$ 0.388       |

TABLE 5.7. Yield values from the merged filtered file + nominal MC QCD sample

### 5.3.3.3. $E_T^{\text{jet, no JVT}}$ Validation

The MJ background is largely controlled by the  $E_T^{\text{miss}}$  and  $\Delta\phi_{jj}$  requirements but some MJ can inappropriately populate the SR. A true MJ event should contain no  $E_T^{\text{miss}}$ , but jet mismeasurement can introduce fake MET. These mismeasurements include the inappropriate removal of jets from the MET calculation based upon JVT and other jet energy resolution effects. Experimental causes including electronic noise and the punch-through of highly energetic jets also contribute. The high pileup environment of the LHC increases the possibility of these mismeasurements and mis-tagging conspiring to create signal-like topologies. For example, pileup jets from vertices other than the PV can be tagged as a VBF jet leading to coincidental signal-like topologies. This is especially complicated with the minimum jet  $p_T$  requirement imposed by the analysis (jet  $p_T > 25$  GeV), where lower- $p_T$  PU jets that could balance the VBF-tagged PU jet would be lost.

An additional requirement on  $E_T^{\text{jet, no JVT}}$  (see Sec. 4.2.4.1) was included in order to reduce this MJ contamination. Since  $E_T^{\text{jet, no JVT}}$  does not consider JVT score, the effects of inappropriate tagging do not contribute. Events will fail the  $E_T^{\text{jet, no JVT}}$  requirement if a HS jet was inappropriately removed from the  $E_T^{\text{miss}}$  calculation due to JVT or if PU jets are inappropriately included in the  $E_T^{\text{miss}}$  calculation because JVT is only applied for jets with  $p_T > 60$  GeV. This compliments the fJVT requirement of the VBF jets, which checks whether the VBF-tagged jets are consistent with the PV.

This section presents a “proof-of-principle” exercise to investigate the ability of  $E_T^{\text{jet, no JVT}}$  to target and reduce the multijet background.

The efficacy of this  $E_T^{\text{jet, no JVT}}$  cut to target multijet events was investigated with the simulated multijet sample of Section 5.3.3.2<sup>8</sup>. To compare with data, MJ and  $V$ +jets<sup>9</sup> simulation is compared to data from the VBF trigger HLT\_j70\_j50\_0eta490\_invM1100j70\_dphi20\_deta40\_L1MJJ-500-NFF, which has no explicit  $E_T^{\text{miss}}$  requirement. Cuts are made (see Table 5.5) to consider only data and MC events on the trigger plateau. Due to differences in generation of the MJ and  $V$ +jets samples, direct comparisons between data and MC should not be taken seriously. Rather, the trends of data minus  $V$ +jets in response to additional requirements on jet multiplicity and  $E_T^{\text{jet, no JVT}}$  are considered.

Requiring exactly two jets further reduces the number of viable QCD multijet events by 99.8%, all with relatively small  $E_T^{\text{miss}}$  values (see Figure 5.15 a)). From this collection, a series of cuts on  $E_T^{\text{jet, no JVT}}$  are made (see Figure 5.15). With a 50 GeV  $E_T^{\text{jet, no JVT}}$  cut, QCD is reduced by 6.4%. A stricter requirement of 75 GeV reduces 63.1% of the MJ background. The final 120 GeV cut eliminates 90.8% of the QCD background while reducing the  $V$ +jets background by 6.5% (caveat that these samples are filtered on an OR of  $E_T^{\text{miss}}$  definitions with  $E_T^{\text{miss}} > 100$  GeV). Both the multijet QCD background and the excess in data minus  $V$ +jets background is greatly reduced. The  $E_T^{\text{jet, no JVT}}$  cut is removing events in which the other  $E_T^{\text{miss}}$  definitions have fake- $E_T^{\text{miss}}$ .

Therefore, implementing a  $E_T^{\text{jet, no JVT}}$  cut helps to control the contribution of the QCD multijet background and compliments the  $E_T^{\text{miss}}$  requirement that is already applied by reducing MJ with fake  $E_T^{\text{miss}}$  due to PU.

---

<sup>8</sup>This sample contains MC generated to reproduce 2015/2016 run conditions, and was not reweighted for 2018 PU conditions.

<sup>9</sup>The  $V$ +jets samples are skimmed with  $E_T^{\text{miss}} > 100$  GeV.

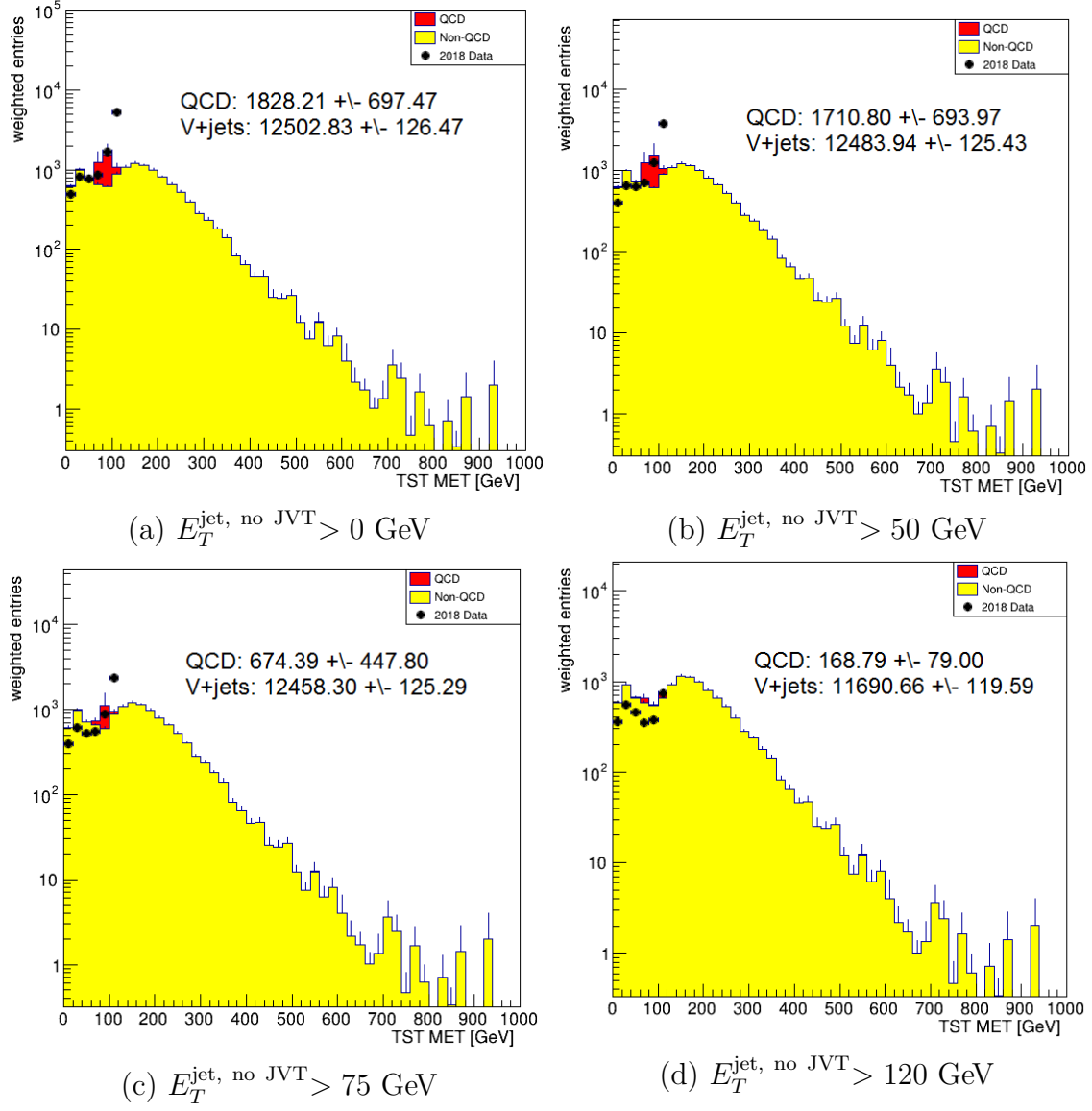


FIGURE 5.15.  $E_T^{\text{miss}}$  distribution for VBF trigger-passing 2 jet events with additional  $E_T^{\text{jet, no JVT}}$  requirements

### 5.3.4. Fitting and Limit Setting

With a signal region defined, the contribution of background events that contaminate the SR must be understood. This is done differently for the different background sources. These measurements include sources of uncertainty that must be propagated to the final result. A fitting procedure is used to understand

and quantify the leading  $V$ +jets backgrounds. In the absence of a statistically significant signal detection, an upper limit is set on  $\mathcal{B}_{h \rightarrow \text{inv}}$ . Fitting and limit-setting is conducted with HistFitter and Root-based statistics software.

### 5.3.4.1. $V$ +jets Background Estimation

The dominant  $V$ +jets background is estimated through a transfer factor method (see Figure 5.16). Data and MC yields from the CRs are first compared (see Figure 5.17) to understand modeling differences and the impact of theoretical uncertainties (related to the MC calculations). The expected background yield

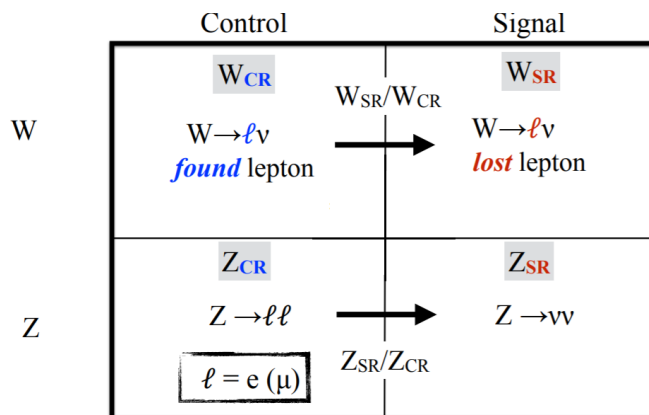


FIGURE 5.16. Visual of  $V$ +jets CRs used to estimate contribution to SR

in the SR from data,  $B_{V,i}^{\text{SR}}$  where  $i$  is a bin index and  $V = W$  or  $Z$ , is found by essentially scaling the MC yield of background events in the SR ( $B_{V,i}^{\text{SR,MC}}$ ) by a normalization factor  $\beta_{V,i}$ . One normalization  $\beta$  is calculated for each bin for the ZCR and WCR individually, leading to 22 total factors. The normalization  $\beta_{V,i}$  is found from the fit but can be thought of like

$$\beta_{V,i} = \frac{N_i^{\text{VCR}} - B_{\text{non-}V,i}^{\text{VCR}}}{B_{V,i}^{\text{VCR,MC}}} ,$$

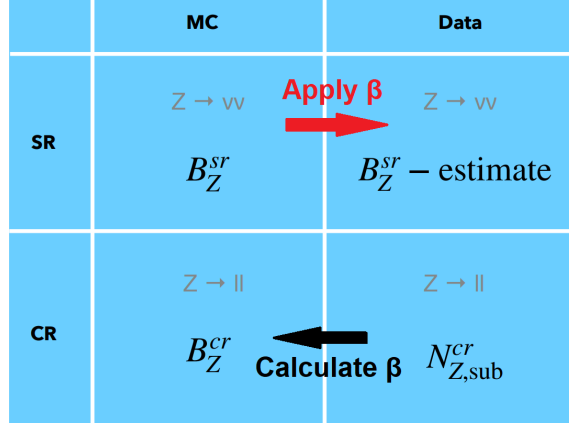


FIGURE 5.17. Visual of data and MC in ZCR used to calculate the normalization factor  $\beta$

where  $N_i^{V\text{CR}}$  is the data yield in the designated CR bin,  $B_{V,i}^{V\text{CR},\text{MC}}$  is the  $V$ +jets MC yield of the designated CR bin, and  $B_{\text{non-}V,i}^{V\text{CR}}$  is the MC yield of all backgrounds other than the  $V$ +jets process in question in the designated CR bin. In this way, each WCR and ZCR bin is treated independently. The estimated  $W$  and  $Z$ +jets yields in each SR bin are

$$B_{V,i}^{\text{SR}} = \beta_{V,i} \cdot B_{V,i}^{\text{SR},\text{MC}}.$$

#### 5.3.4.2. Other Background Estimation

The fake- $e$  background (of events with baseline electrons that fail the  $\mathcal{S}_{\text{MET}}$  cut) is estimated through comparison with data. A ratio ( $R_S$ ) of data events passing(failing) the  $\mathcal{S}_{\text{MET}}$  cut after the subtraction of backgrounds other than MJ is computed. This ratio is used to scale the fake- $e$  SR contribution.

The MJ background arises in the SR mainly due to jet mismeasurement or experimental effects and is estimated through a data-driven technique called Rebalance and Smear (R&S) (see Figure 5.18). Events with no leptons obtained



from a prescaled single-jet trigger are first considered. Jets with  $p_T > 25$  GeV are classified as HS or PU depending upon (f)JVT. The HS and PU collections from each event are separately “rebalanced” where their momenta is adjusted within their uncertainties such  $E_T^{\text{miss}}$  is minimized. The momenta of these rebalanced jets are then smeared within the expected jet response. The smeared HS and PU collections are then recombined and (f)JVT is recalculated (as momenta has changed). These collections are then normalized by a loose multijet control region. The MET triggers used for the search introduce additional inefficiencies not present in the single-jet triggered sample, so corrections for this are also included. The performance of this overall method is validated in multiple validation regions, and associated systematic uncertainties are calculated.

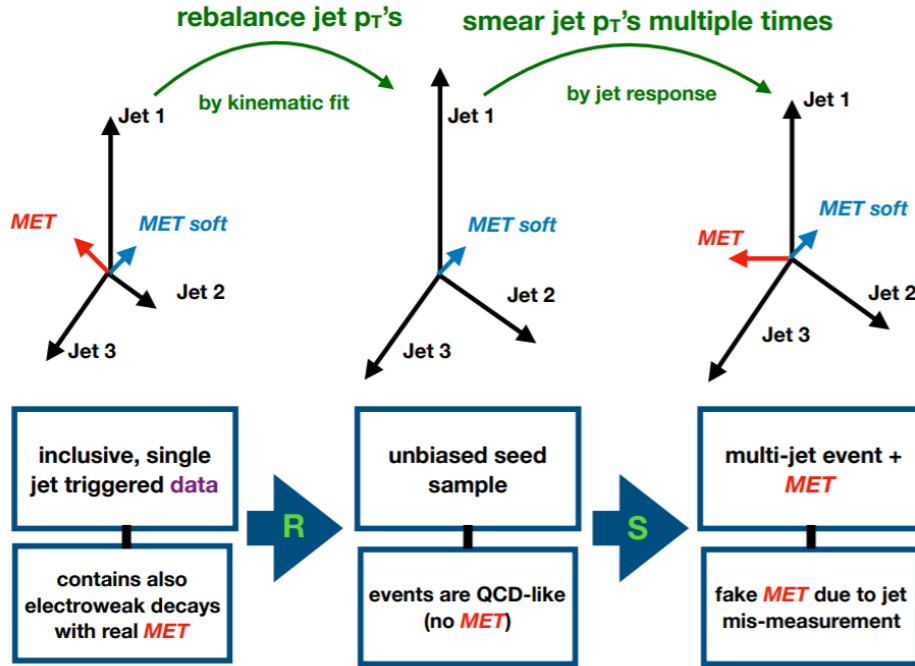


FIGURE 5.18. Overview of Rebalance and Smear MJ background estimation technique

Small backgrounds including  $t\bar{t}$  and multi-boson are estimated directly from MC estimates.

### 5.3.4.3. Systematic Uncertainties

Sources of uncertainty are associated with statistics, theoretical calculations, and measurement uncertainties.

An uncertainty of 1.7% is ascribed on the measured luminosity value, obtained by the LUCID-2 detector [127, 128]. There are potentially large uncertainties associated with  $E_{\text{T}}^{\text{miss}}$ , as any effect from any physics object used in the  $E_{\text{T}}^{\text{miss}}$  calculation must be propagated. Uncertainties for the  $E_{\text{T}}^{\text{miss}}$  soft term are derived independently. Overall, the impact of  $E_{\text{T}}^{\text{miss}}$  systematic uncertainty is considered through centralized ATLAS recommendations. Jet Energy Resolution (JER) and Jet Energy Scale (JES) are categories of uncertainties related to the detection, calibration, and reconstruction of jet objects [129]. Additional sources of uncertainty are also associated with JVT efficiency and pileup modeling. Leptons carry experimental sources of uncertainty as well, including errors associated with particle identification and reconstruction, isolation, trigger efficiency, and detector energy resolution.

Theoretical uncertainties are propagated through the calculation of MEs for MC event generation. These are due to inherent uncertainties in PDFs and parton showering, as well as the finite-order calculations that can be done. Each MC sample is affected differently, depending upon the process generated and the generator used. PDF uncertainties associated with  $V$ +jets ranges from 1-2%.

Four scales are chosen for  $V$ +jets background generation - renormalization, factorization, resummation, and CKKW matching [130]. Uncertainties from

renormalization and factorization scales are found varying each scale up and down by a factor of two from the nominal using on-the-fly varied event weights in SHERPA. This eliminates the need to generate additional samples, and reduces the statistical uncertainty on the estimate. Seven variations are considered - the central value, the two scales independently varied up/down, and both scales coherently varied up/down - and an envelope is used to calculate the overall uncertainty. This is done independently for strong and EWK  $V$ +jets processes and varies as a function of  $m_{jj}$  with strong (EWK) 7-point variations growing from  $^{+27\%}_{-18\%} (^{+11\%}_{-9\%})$  at low  $m_{jj}$  to  $^{+43\%}_{-26\%} (^{+29\%}_{-20\%})$  at high  $m_{jj}$ . The the resummation scale is also varied by a factor of two, while the CKKW scale is varied down to 15 GeV and up to 30 GeV from the nominal 20 GeV. These are applied with event-by-event reweighting factors and range from 4-8%.

No correlation is assumed between  $Z$  and  $W$  or strong and EWK processes, but each  $V$ +jets uncertainty is treated as fully correlated in the fit. Therefore, an uncertainty that increases the  $Z$ +jets strong background in the SR will also increase the corresponding background in the ZCR. In this way, there is a strong cancellation of these uncertainties with the uncertainty on the ratio of variation of each background in the SR to the corresponding CR falling within 1-3% and driven by statistics.

Theoretical uncertainties associated with signal samples arise from similar sources. Production cross section and uncertainties are taken from the LHC Higgs Yellow Report [131]. ggF signal includes renormalization and factorization scale variations, as well as PDF and parton shower uncertainties. Bin migration also contributes an uncertainty of 45%(41%) in the 2-jet (3- and 4-jet) bin.

The VBF signal sample is reweighted with a  $p_T$ -dependent NLO EWK correction of about 2%. Renormalization and factorization uncertainties and correlations were computed by the LHC Higgs working group through independent variation, yielding uncertainties of 1-3%. PDF and PS uncertainties from 1-4% are also included.

#### 5.3.4.4. Fit Model

Expected background yields are obtained by a simultaneous maximum profiled likelihood fit of the background components to the data in all CRs and SRs. The signal strength  $\mu$  is the only global free parameter of interest (POI) in the fit. Each bin has three associated free parameters ( $\beta_Z$ ,  $\beta_W$ , and  $n_{\text{fake}}$  associated with the fake- $e$  background) which are extracted from the fit along with the global signal strength  $\mu$ , which is interpreted as  $\mathcal{B}_{h \rightarrow \text{inv}}$ .

A HistFitter workspace is used along with ROOT wrappers RooFit [132] and RooStats [133]. A general likelihood function follows the form

$$\mathcal{L}(\mu, \theta; \text{data}) = \prod_{c=1}^{N_{\text{cat}}^a} \mathcal{L}_c^a(\mu, \theta; \text{data}) \prod_{k=1}^{N_{\text{cons}}^a} \mathcal{G}(\theta_{i_k^a}; \tilde{\theta}_{i_k^a}^a), \quad (5.2)$$

where  $\mu$  and  $\theta$  are vectors of the POIs and NPs respectively and  $a$  indicates an analysis channel (in this case  $a = 1$  and only considers the VBF channel).  $N_{\text{cat}}^a$  is the number of categories (bins) in each analysis  $a$  and the index  $c$  spans these categories. Gaussian probability functions  $\mathcal{G}$  are used to model NPs, where  $\tilde{\theta}_{i_k^a}^a$  is the global observable corresponding to  $\theta_{i_k^a}$  and  $k$  runs over all constrained

(systematic) uncertainties. The fit maximizes the profiled likelihood ratio,

$$\Lambda(\mu; \theta) = \frac{\mathcal{L}(\mu, \hat{\hat{\theta}}(\mu))}{\mathcal{L}(\hat{\mu}, \hat{\theta})}, \quad (5.3)$$

which is used as the test statistic. In the numerator, NPs are set to their profiled values which maximize the likelihood function for fixed values of  $\mu$ . In the denominator, both the POI and the NPs are set to whatever values jointly maximize the likelihood.

The full likelihood function,

$$\begin{aligned} \mathcal{L}(\mu, \vec{\beta}_Z, \vec{\beta}_W, \vec{n}_{\text{fake}}, \vec{\theta}) = & \prod_i \mathcal{P} \left( N_i^{\text{SR}} \mid \beta_{Z,i} \cdot B_{Z,i}^{\text{SR,MC}} + \beta_{W,i} \cdot B_{W,i}^{\text{SR,MC}} + B_{\text{MJ},i}^{\text{SR}} + B_{\text{other},i}^{\text{SR,MC}} + \mu \cdot S_i^{\text{SR,MC}} \right) \\ & \prod_i \mathcal{P} \left( N_i^{\text{ZCR}} \mid \beta_{Z,i} \cdot B_{Z,i}^{\text{ZCR,MC}} + B_{\text{non-Z},i}^{\text{ZCR}} \right) \\ & \prod_i \mathcal{P} \left( N_i^{W\mu\nu\text{CR}} \mid \beta_{W,i} \cdot B_{W,i}^{W\mu\nu\text{CR,MC}} + B_{\text{non-W},i}^{W\mu\nu\text{CR,MC}} \right) \\ & \prod_i \mathcal{P} \left( N_i^{We\nu\text{CR}} \mid \beta_{W,i} \cdot B_{W,i}^{We\nu\text{CR,MC}} + B_{\text{non-W},i}^{We\nu\text{CR,MC}} + R_{S,i} \cdot n_{\text{fake},i} \right) \\ & \prod_i \mathcal{P} \left( N_i^{\text{fake-}e\text{ CR}} \mid \beta_{W,i} \cdot B_{W,i}^{\text{fake-}e\text{ CR,MC}} + B_{\text{non-W},i}^{\text{fake-}e\text{ CR,MC}} + n_{\text{fake},i} \right) \\ & \prod_j \mathcal{G}(0 \mid \theta_j), \end{aligned}$$

involves Poisson ( $\mathcal{P}$ ) and Gaussian ( $\mathcal{G}$ ) probability density functions with parameters of background ( $B$ ) and signal ( $S$ ) yields.  $R_S$  is the fake- $e$  ratio (see Section 5.3.4.2).

Uncertainties are introduced to the fit as nuisance parameters (NPs,  $\vec{\theta}$ ) which also depend on  $B$  and  $S$ . There are more than 200 NPs in this search. NPs affect the yields through an exponential response function  $(1 + \varepsilon_{ij})^{\theta_j}$  where  $\varepsilon_{ij}$  is the fractional uncertainty amplitude of bin  $i$  from the uncertainty source  $j$ . This treatment prevents the fitted yields from turning negative in cases of large uncertainties. Each experimental uncertainty and PDF uncertainties are considered to be fully correlated across all signal and control regions. Perturbative theory uncertainties are considered in twelve bins - a set of eleven, uncorrelated bins associated with each SR bin and one additional fully correlated component.

Nominally, all uncertainty values are zero and constrained within Gaussian probability distributions with widths corresponding to the size of their uncertainty. The degree of deviation from zero is calculated with asymptotic methods utilizing an Asimov data set<sup>10</sup> [134]. The Asimov data set and known properties of the test statistic can be used to derive the variance of the parameter of interest by solving for the covariance matrix.

The Asimov data set is used to calculate an expected upper limit on  $\mu$ . Limits on  $\mu$ , both the expected using the Asimov data set and the observed using experimentally collected data, are set at 95% C.L. through the  $CL_S$  frequentist approach [135], relating to an observed upper limit on  $\mathcal{B}_{h \rightarrow \text{inv}}$ .

With this approach, signal and background events are distinguished and yields within each category are considered. In this case, “background” is considered to be the null hypothesis assuming only the SM and “signal” are BSM signals beyond the scope of the SM (here,  $h \rightarrow \chi\chi$ ) offering an alternative hypothesis. This alternative hypothesis, practically, cannot be probed without influence of the SM so in practice this hypothesis truly considers both signal and background (S+B). PDFs are used to derive functions representing the probability density of the null hypothesis ( $CL_B$ ) and alternative hypothesis ( $CL_{S+B}$ ). These are functions of  $-2\ln(Q)$  where  $Q = \mathcal{L}(S + B)/\mathcal{L}(B)$  is a ratio of likelihood functions. A lack of signal detection in repeated experiments leads to a uniform distribution of  $CL_B$ . The signal-only probability distribution function can then be found from the ratio

$$CL_S = \frac{CL_{S+B}}{CL_B}$$

---

<sup>10</sup>An Asimov data set is effectively a perfect data set free of fluctuations, where any estimator of a parameter yields the true parameter value.

and though of as an approximate confidence in a signal-only hypothesis.

Confidence levels are defined from where the  $CL_S$  curve crosses the desired value of probability density (for example, the 95% C.L. is set where  $CL_S$  crosses 5%).

These values provide a conservative estimate.

Different fits are computed before data in the signal region is considered (while the analysis is “blinded”) to ensure stability of the fit, reliability of the background estimation, and to avoid bias. These control region- and background-only fits provide validation that the modeling is sound. A final fit utilizing all data and all MC in all regions informs the observed upper limit value on  $\mu$  and  $\mathcal{B}_{h \rightarrow \text{inv}}$ .

### 5.3.5. Results

The  $S$  and  $B$  yields after the fit are shown in Appendix F, with good agreement of expected background rates to the observed data (see Figure 5.19). In the SR, an overall signal efficiency ( $S/B$ ) of 0.05 is seen. This agreement is also true of key variables  $\Delta\phi_{jj}$  and  $m_{jj}$  (see Figure 5.20). No signal excess was observed, and an upper limit on  $\mathcal{B}_{h \rightarrow \text{inv}}$  was set as 0.13 (0.13) observed (expected) (see Table 5.8).

| Observed | Expected | +1 $\sigma$ | −1 $\sigma$ | +2 $\sigma$ | −2 $\sigma$ |
|----------|----------|-------------|-------------|-------------|-------------|
| 0.132    | 0.132    | 0.183       | 0.095       | 0.248       | 0.071       |

TABLE 5.8. Expected and observed limits on  $\mathcal{B}_{h \rightarrow \text{inv}}$  calculated at the 95% C.L. with 139 fb<sup>−1</sup>.

The 20 NPs that individually most impact the limit (see Figure 5.21) include normalization factors (KZ\*), elements of JER, lepton factors, and MJ sources. The impact on the final limit is illustrated with the top axis and each blue box. The circular marker indicates the pull of the NP from the fitting procedure, against the

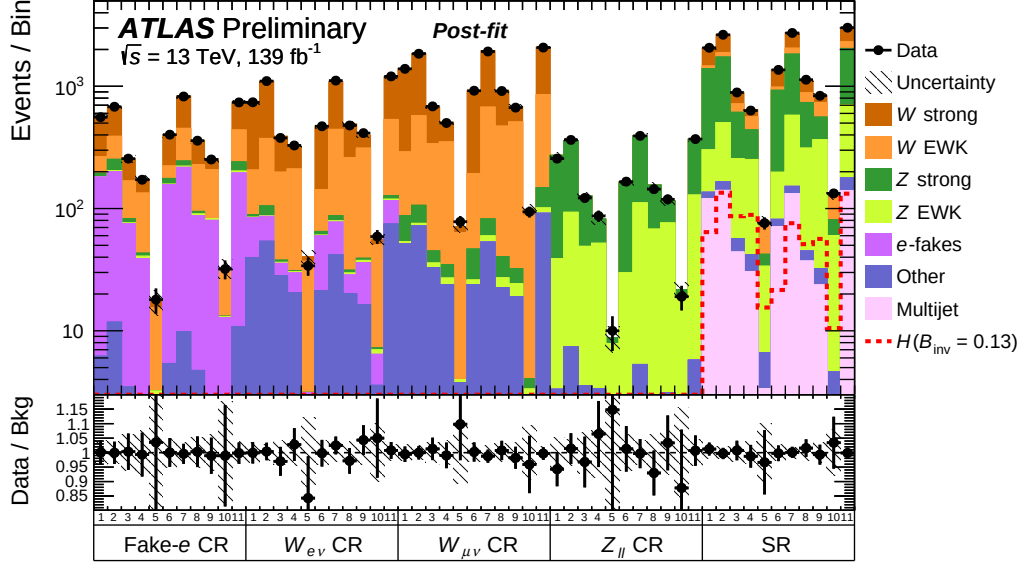
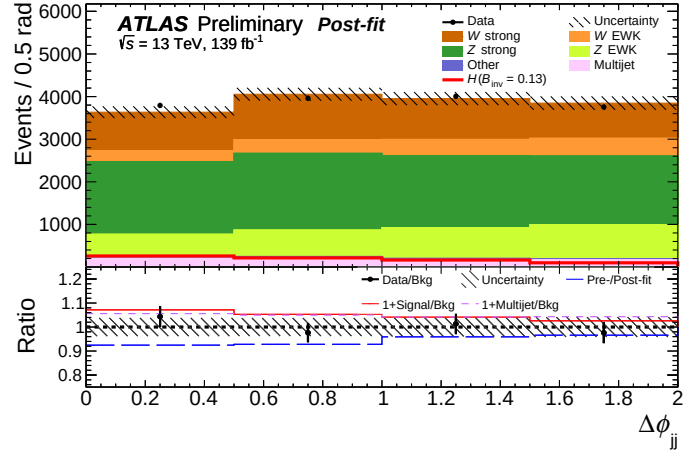


FIGURE 5.19. Visualization of postfit data and backgrounds in all SR and CR bins

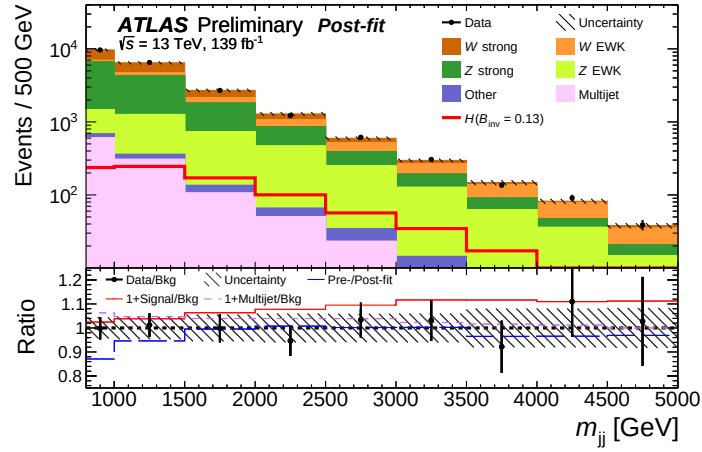
bottom axis. Normalization factors are expected to be pulled away from unity, and all other NPs from zero. A strong pull can indicate an unstable fit or problematic NP. This was noticed during the optimization procedure due to NPs (especially those associated with JER) with large statistical fluctuations in different bins. It was mitigated by smoothing the distribution across the spectrum.

The impact of groups of systematic uncertainties is found by holding all NPs in one group set to the nominal value while allowing all others to float in a new fit. The difference in the resulting signal strength  $\mu$  indicates the degree of impact that each NP group has on the measurement (see Table 5.9). Fixing all NPs to the nominal value results in a limit of 7.3%, illustrating the influence of NPs in the fit. The statistical uncertainty on the yields of simulated events and data in the SR and CRs have an impact of  $\sim 8\%$  and  $\sim 17\%$ , respectively. These are the dominant sources of uncertainty in this search. Leading experimental uncertainties include JER and lepton uncertainties. The derived uncertainties from the data-





(a)



(b)

FIGURE 5.20. Postfit background estimation and data as a function of a)  $\Delta\phi_{jj}$  and b)  $m_{jj}$

driven R&S MJ estimation approach at  $\Delta\mathcal{B}_{h\rightarrow\text{inv}} \sim 7\%$  are the third highest impacting group.

#### 5.4. Closer look at Run 2

A closer consideration of the full Run 2 data set builds heavily off the previous work (see Section 5.3), and is currently blinded. I contributed to a refined

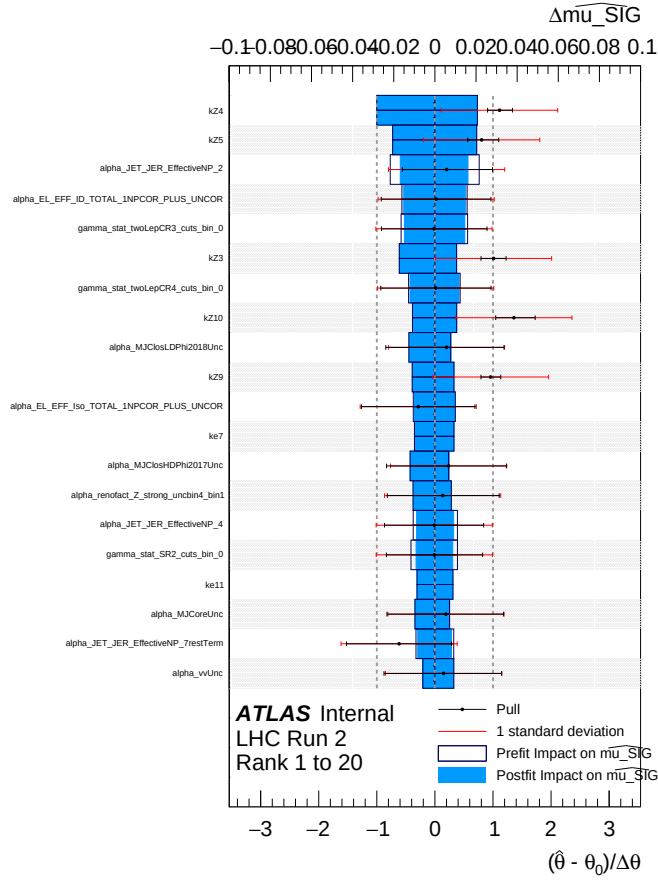


FIGURE 5.21. Top 20 most impactful NPs in the fit

| Source                | $\Delta$ [%] |
|-----------------------|--------------|
| Jet energy scale      | 1.8          |
| Jet energy resolution | 5.5          |
| Lepton                | 4.6          |
| Other                 | 1.9          |
| Multijet              | 7.0          |
| V+jets theory         | 1.6          |
| Signal theory         | 1.0          |
| MC stats.             | 7.9          |
| Data stats.           | 17.3         |

TABLE 5.9. Relative impact  $\Delta$  of the 95% CL expected upper limit on  $\mathcal{B}_{h \rightarrow \text{inv}}$  if a group of uncertainties is “removed”, i.e. if the corresponding nuisance parameters are fixed to the nominal values.

MJ background estimation technique. A paper presenting updated results is being planned for publication this year.

### 5.4.1. General Improvements

This search loosened kinematic cuts (see Table 5.10) to accept additional signal. Notably, the  $E_T^{\text{miss}}$  requirement was loosened to 160 GeV<sup>11</sup>. A new variable,

$$m_T = \sqrt{2p_T^\mu E_T^{\text{miss}} \left[ 1 - \cos(\phi_\mu - \phi_{E_T^{\text{miss}}}) \right]},$$

is introduced to estimate the rate of muons incorrectly reconstructed as jets, leading to  $V$ +jets events that are miscategorized into the signal region. This is done identically to the fake- $e$  background estimation procedure using  $m_T$  instead of  $\mathcal{S}_{\text{MET}}$ .

| Cut                        | Old Region          | Updated Region                               |
|----------------------------|---------------------|----------------------------------------------|
| $E_T^{\text{miss}}$        | $> 200 \text{ GeV}$ | $> 160 \text{ GeV}$                          |
| $E_T^{\text{jet, no JVT}}$ | $> 180 \text{ GeV}$ | $> 140 \text{ GeV}$                          |
| $m_T$                      | -                   | $> 20 \text{ GeV} (\mu \text{ in WCR only})$ |

TABLE 5.10. SR cuts that have been updated (unlisted variables remain identical to Table 5.4)

The binning scheme was also updated (see Figure 5.22). The multijet bin was broken into three  $m_{jj}$  bins, and the  $E_T^{\text{miss}}$  threshold for these and the ten  $N_{\text{jet}} == 2$  bins was increased to  $E_T^{\text{miss}} > 200 \text{ GeV}$ . This allowed for the addition of three  $m_{jj}$ -binned “low-MET” bins for events with  $160 < E_T^{\text{miss}} < 200 \text{ GeV}$  for a total of 16 bins. This allows for more parameters (especially  $V$ +jets normalization factors) in the fit.

---

<sup>11</sup>The Loose fJVT WP is used for leading jets of events with  $E_T^{\text{miss}} > 200 \text{ GeV}$ . Very Tight fJVT is used for events with  $E_T^{\text{miss}} < 200 \text{ GeV}$ .

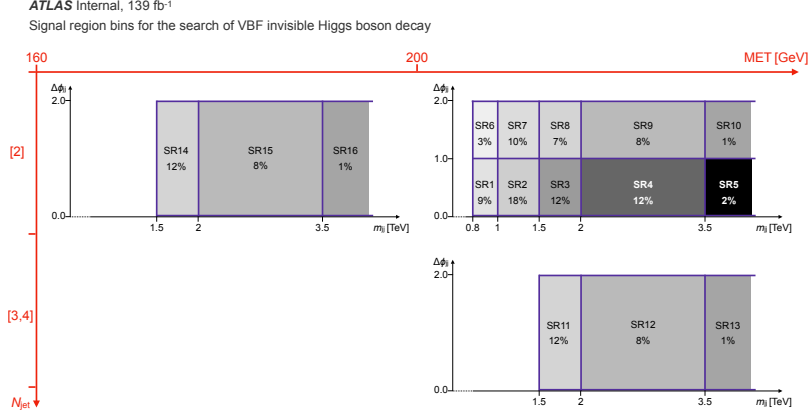


FIGURE 5.22. Updated binning scheme for VBF+MET (compare to Figure 5.4)

Monte Carlo statistics of the  $V$ +jets background provided a large source of uncertainty in the search for VBF  $h \rightarrow \text{inv.}$  utilizing a partial Run 2 data set of 36 fb<sup>-1</sup> [120]. Many problems with this uncertainty were mitigated in the previous work (see Table 5.9), however improvements were still be made. Extensions of 125 million  $V$ +jets events in the extreme VBF phase space are now considered, further reducing Monte Carlo modeling uncertainties.

#### 5.4.2. Improved MJ Estimate

The QCD MJ background estimation technique was also greatly changed from the previous work (see Section 5.3.4.2). The high-pileup environment of Run 2 eventually lead to large uncertainties associated with the R&S strategy (see Table 5.9), which was revised to emulate the fake-lepton control regions. In the new FJVT CR method, the fVT value of the leading jet is used as a discriminating variable and normalization factors based on the ratio of events that pass and fail the fJVT requirement lead to an estimate of the impact of MJ in each bin. I derived these FJVT normalization factors and provided method validation, especially for the newly introduced low-MET bins requiring the Very Tight fJVT

WP. This method benefits from much reduced uncertainties when compared to R&S, but does sacrifice the event-by-event information that R&S was able to retain (such as kinematic shape distributions in addition to simple yield values). Its performance also falters in the multijet bins. Now, a combined strategy is used utilizing both the FJVT CR and R&S methods (see Table 5.11).

| SR bins | FJVT CR                          | Reblance and Smear      |
|---------|----------------------------------|-------------------------|
| 1-5     | 1 CR per bin                     | -                       |
| 6-10    | 1 inclusive CR for normalisation | $m_{jj}$ shape          |
| 11-13   |                                  | Shape and normalisation |
| 14-16   | 1 CR per bin                     | -                       |

TABLE 5.11. The usage of MJ estimates by SR bins. A “-” indicates that the method is not used.

Forward JVT score of the leading jet is chosen as a discriminating variable for the FJVT CR. Events failing the fJVT requirement lead to a CR enriched in MJ events with pileup jets and roughly 95(92.5)% pure in MJ for high(low)-MET bins (see Figures 5.23 and 5.24). Similarly to the fake-lepton background techniques, a normalization factor ( $f_{\text{MJ}}$ ) is calculated by comparing the data-background disagreement ratio between events that pass and fail the leading jet fJVT requirement,

$$f_{\text{MJ}}^i = \frac{N_{\text{data}}^{\text{pass FJVT},i} - N_{\text{non-MJ}}^{\text{pass FJVT},i}}{N_{\text{data}}^{\text{rev. FJVT},i} - N_{\text{non-MJ}}^{\text{rev. FJVT},i}} .$$

In order to increase available statistics, we consider events with  $2 \leq N_{\text{jet}} \leq 4$  for the FJVT CR but extrapolate to the  $N_{\text{jet}} == 2$  region ( $N^{\text{pass FJVT}}$ ). This is another reason that the FJVT CR method cannot be used in the multijet bins, as these events are already used. The low-MET region also does not bin in  $\Delta\phi_{jj}$ . Normalization factors (see Table 5.12) are calculated separately for events with

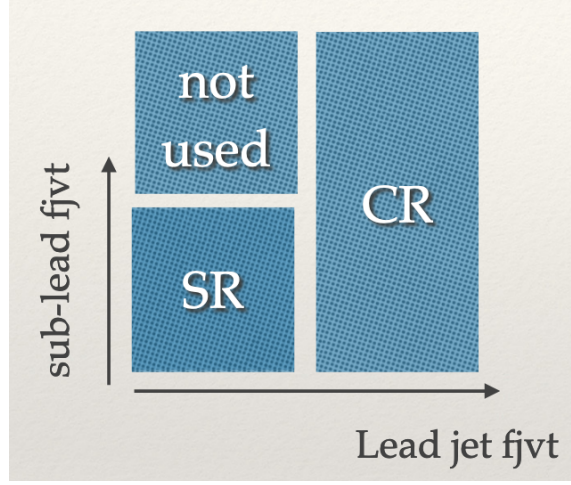


FIGURE 5.23. Definition of FJVT CR as compared to the SR

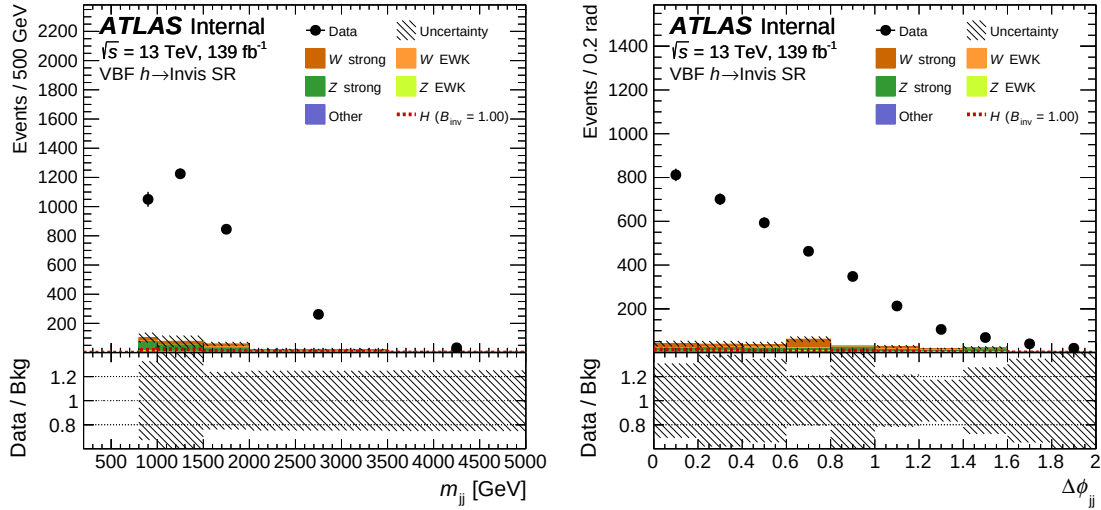


FIGURE 5.24. Distribution in low-MET FJVT CR (leading jet fjvt < 0.2) of a)  $m_{jj}$  and b)  $\Delta\phi_{jj}$

$E_T^{\text{miss}} > 200$  GeV and  $E_T^{\text{miss}} < 200$  GeV, as different fJVT WPs are used. Factors tend to be larger for high-MET bins due to the looser fJVT requirement.

All derived normalization factors are stable against relevant kinematic variables ( $E_T^{\text{miss}}$ , jet  $p_T$ , jet  $\eta$ , and pileup conditions) (see Figure 5.25). For the low-MET bins, samples containing events with  $100 \leq E_T^{\text{miss}} \leq 160$  GeV are used (after normalization with 1- and 2- lepton CRs). The higher pileup and V+jets

| SR Bin | $m_{jj}$    | $\Delta\phi_{jj}$           | $E_T^{\text{miss}}$ | $f_{\text{MJ}}^i$ | Stat. unc.        | Systematic unc.              |
|--------|-------------|-----------------------------|---------------------|-------------------|-------------------|------------------------------|
| SR1    | 0.8-1.0 TeV | $ \Delta\phi_{jj}  < 1$     | $> 200$ GeV         | 2.29 (5.7 comb.)  | $\pm 0.23$ (10%)  | $\pm 0.47$ (21%) (60% comb.) |
| SR2    | 1.0-1.5 TeV | $ \Delta\phi_{jj}  < 1$     | $> 200$ GeV         | 2.19              | $\pm 0.39$ (18%)  | $\pm 0.44$ (20%)             |
| SR3    | 1.5-2.0 TeV | $ \Delta\phi_{jj}  < 1$     | $> 200$ GeV         | 1.63              | $\pm 0.27$ (17%)  | $\pm 0.32$ (20%)             |
| SR4    | 2.0-3.5 TeV | $ \Delta\phi_{jj}  < 1$     | $> 200$ GeV         | 1.52              | $\pm 0.15$ (10%)  | $\pm 0.30$ (20%)             |
| SR5    | $> 3.5$ TeV | $ \Delta\phi_{jj}  < 1$     | $> 200$ GeV         | 1.82              | $\pm 0.49$ (27%)  | $\pm 0.36$ (20%)             |
| SR6    | 0.8-1.0 TeV | $1 <  \Delta\phi_{jj}  < 2$ | $> 200$ GeV         | 1.75              | $\pm 0.21$ (12%)  | $\pm 0.39$ (22%)             |
| SR7    | 1.0-1.5 TeV | $1 <  \Delta\phi_{jj}  < 2$ | $> 200$ GeV         | 1.76              | $\pm 0.25$ (14%)  | $\pm 0.35$ (20%)             |
| SR8    | 1.5-2.0 TeV | $1 <  \Delta\phi_{jj}  < 2$ | $> 200$ GeV         | 2.10              | $\pm 0.54$ (26%)  | $\pm 0.47$ (22%)             |
| SR9    | 2.0-3.5 TeV | $1 <  \Delta\phi_{jj}  < 2$ | $> 200$ GeV         | 2.35              | $\pm 0.78$ (33%)  | $\pm 0.28$ (20%)             |
| SR10   | $> 3.5$ TeV | $1 <  \Delta\phi_{jj}  < 2$ | $> 200$ GeV         | 3.38              | $\pm 1.42$ (42%)  | $\pm 0.68$ (20%)             |
| SR6-10 | $< 0.8$ TeV | $1 <  \Delta\phi_{jj}  < 2$ | $> 200$ GeV         | 1.80              | $\pm 0.20$ (11%)  | $\pm 0.36$ (20%)             |
| SR14   | 1.5-2.0 TeV | $ \Delta\phi_{jj}  < 2$     | 160-200 GeV         | 0.37              | $\pm 0.037$ (10%) | $\pm 0.10$ (27%)             |
| SR15   | 2.0-3.5 TeV | $ \Delta\phi_{jj}  < 2$     | 160-200 GeV         | 0.32              | $\pm 0.032$ (10%) | $\pm 0.064$ (20%)            |
| SR16   | $> 3.5$ TeV | $ \Delta\phi_{jj}  < 2$     | 160-200 GeV         | 0.36              | $\pm 0.065$ (18%) | $\pm 0.083$ (23%)            |

TABLE 5.12. FJVT CR normalization factors along with their statistical and systematic uncertainties. The discrepancy in SR1 between the FJVT and the R&S is handled by taking the average of the two predictions (labeled “comb”.’) with an uncertainty that covers the difference.

contributions to these samples lead to a slight leading jet  $p_T$  dependence (see Figure 5.25 c)). This is possibly enhanced by the jet collection used, where PFlow jets utilize copied fJVT values from the EMTopo collection. This jet mapping is less efficient for low- $p_T$  jets. This variable is not used in the method, so the small effect is deemed negligible.

The MJ yield ( $N_{\text{MJ}}$ ) in any SR bin  $i$  as

$$N_{\text{MJ}}^i = f_{\text{MJ}}^i \times (N_{\text{data}}^{\text{rev},i} - N_{\text{non-MJ}}^{\text{rev},i}) ,$$

where  $N^{\text{rev},\cdot}$  is the number of events in the FJVT CR with reversed leading fJVT ( $\text{fJVT} > 0.5(0.2)$  if  $\text{MET} > (<) 200$  GeV). The predicted yields agree with those estimated from R&S (see Table 5.13), with lower uncertainties.

The R&S strategy itself was slightly updated from previous use. Now, templates used to smear HS and PU jet collections are derived from MC rather than from data. This reduces mis-identification of a jet as HS or PU due to (f)JVT tagging, but also requires that pileup be modeled by MC. To compensate

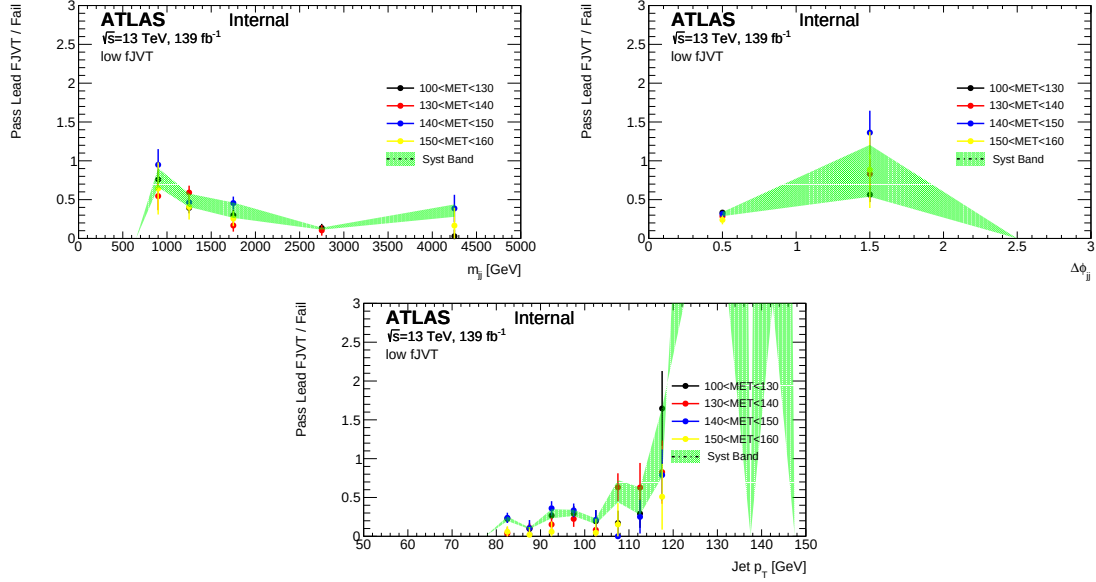


FIGURE 5.25. Low-MET FJVT CR normalization factors ( $1\sigma$  statistical uncertainties shown in green band)

|          | Year      | FJVT CR MJ $\pm$ syst $\pm$ stat | R&S                    | CONF Note yields $\pm$ syst $\pm$ stat |
|----------|-----------|----------------------------------|------------------------|----------------------------------------|
| High-MET | 2015-2016 | 95 $\pm$ 19 $\pm$ 15             | 77 $\pm$ 31            | 85 $\pm$ 85                            |
|          | 2017      | 207 $\pm$ 42 $\pm$ 25            | 325 $\pm$ 135          | 231 $\pm$ 115                          |
|          | 2018      | 187 $\pm$ 38 $\pm$ 22            | 277 $\pm$ 94           | 289 $\pm$ 116                          |
|          | 2015-2018 | 489 $\pm$ 99 $\pm$ 37            | 679 $\pm$ 260          | 605 $\pm$ 316                          |
| low-MET  | 2015-2016 | 117 $\pm$ 24 $\pm$ 7             | 43 $\pm$ 17 $\pm$ 3    | 87 $\pm$ 87 $\pm$ 16                   |
|          | 2017      | 154 $\pm$ 32 $\pm$ 8             | 97 $\pm$ 46 $\pm$ 19   | 111 $\pm$ 44 $\pm$ 20                  |
|          | 2018      | 132 $\pm$ 27 $\pm$ 8             | 170 $\pm$ 57 $\pm$ 77  | 188 $\pm$ 75 $\pm$ 42                  |
|          | 2015-2018 | 403 $\pm$ 83 $\pm$ 13            | 310 $\pm$ 120 $\pm$ 80 | 386 $\pm$ 147 $\pm$ 49                 |

TABLE 5.13. MJ yield estimations

for differences between data and simulation, an additional systematic shape uncertainty is introduced.

### 5.4.3. Results

The improvements outlined in Section 5.4.1 lower the expected upper limit on  $\mathcal{B}_{h \rightarrow \text{inv}}$ . At this time, the analysis remains blinded but Asimov expected limits with full systematics of  $\mathcal{B}_{h \rightarrow \text{inv}} < 0.1132$  is a 15% relative improvement over the



expected upper limit of 0.132 publicized in the first look at the full Run 2 data set (see Section 5.3).

The analysis currently remains blinded but ready to advance upon the refinement of theoretical uncertainties. The intended timeline hopes to begin ATLAS analysis approval in 2021 and publication in a peer reviewed journal. To date, there has been no public result from CMS regarding a full Run 2 data set analysis of VBF  $h \rightarrow \text{inv.}$

## CHAPTER VI

### $h \rightarrow$ invisible COMBINATION

#### 6.1. Combination Overview and Strategy

In order to most fully take advantage of the multitude of work done to study  $h \rightarrow$  inv. and further exploit the large data set of Run 2, a statistical combination of different  $h \rightarrow$  inv. searches (from different Higgs production channels) is performed. A preliminary result using three Run 2 inputs (including the VBF result of Section V) and the Run 1  $h \rightarrow$  inv. combination was publicized in a 2020 conference note [112]. This chapter contains material coauthored with the ATLAS collaboration. I coordinated the study of orthogonality between input analyses and quantified the impact of groups of uncertainties on the expected UL, among smaller contributions. Other members of the ATLAS collaboration contributed to the analysis through the definition and implementation of the correlation scheme and fitting procedure. In combination, this all informed the resulting observed UL of  $\mathcal{B}_{h \rightarrow \text{inv}} < 0.11$  (0.11) obs (exp).

This work will be built upon in a peer reviewed article incorporating more Run 2 analyses. This is anticipated in 2022.

RooWorkspaces from each input analysis is used along with ROOT wrappers RooFit and RooStats to create a combined workspace used to set the combined limit. The likelihood of each input analysis is given by Eq. 5.2. The profiled likelihood ratio of Eq. 5.3 is used as a test statistic. Like the VBF case, NPs are modeled by Gaussian or log-normal probability density functions and an Asimov data set is used to calculate expected results.

The input Run 2 analyses are considered first, creating first a combined  $ttH$  result and then combining with VBF for a Run 2 combination. This Run 2 combination is then combined with Run 1 for a final result. A correlation scheme for NPs (see Table 6.1) is defined between Run 1 and Run 2, as run conditions and algorithms have evolved. No correlations are assumed between most instrumental uncertainties, unless similar methodology is used in both runs to compute them. For example,  $b$ -jet JES uncertainty was estimated with MC simulation and therefore considered correlated. Background and signal modeling uncertainties (for example, parton showering) are considered uncorrelated to reflect the evolution of MC simulation tools.

| NP Name                           |
|-----------------------------------|
| JET_EtaIntercalibration_Modelling |
| JET_Pileup_OffsetMu               |
| JET_Pileup_OffsetNPV              |
| JET_Pileup_PtTerm                 |
| JET_Pileup_RhoTopology            |
| JET_Flavor_Response               |
| JET_Flavor_Composition            |
| JET_BJES_Response                 |

TABLE 6.1. NPs correlated between Run 1 and Run 2 analyses

## 6.2. Input Analyses

Four analyses were considered for this combination (see Table 5.1):

- VBF production (see Section V)
- $ttH$  production with no leptons ( $ttH$  2L)
- $ttH$  production with two leptons ( $ttH$  0L)

– Run 1 combination .

Only Run 2 analyses with public results by the time of HIGGS2020 could be used.

### 6.2.1. VBF Channel

This work is explored in detail in Section V. Due to high sensitivity, this channel drives the combination. Using the full Run 2 data set, the UL  $\mathcal{B}_{h \rightarrow \text{inv}} < 0.13$  (0.13) obs (exp) was set at 95% C.L. [25].

### 6.2.2. $t\bar{t}H$ Channel

The  $t\bar{t}H$  channel itself ( $t\bar{t} + E_{\text{T}}^{\text{miss}}$ ) is a combination of two full Run 2  $t\bar{t} + E_{\text{T}}^{\text{miss}}$  analyses utilizing zero (0L) and two lepton (2L)  $t\bar{t}$  decay channels. This combination was performed for use in this  $h \rightarrow \text{inv.}$  combination effort. Both searches were designed to search for signs of supersymmetry through superpartners of the top quark, and here are interpreted in the SM context of  $h \rightarrow \text{invisible}$ .

The two lepton channel analysis [136] defines a signal region to target the production of a spin-0 mediator particle (here, interpreted as a SM Higgs) decaying directly to DM and produced in association with  $t\bar{t}$  decaying to  $b\bar{b}\ell\nu\ell\nu$  (see Table 6.2). The zero lepton channel analysis [137] focuses on events with  $E_{\text{T}}^{\text{miss}} > 250$  GeV and  $t\bar{t}$  decays resulting in lost leptons. The two signal regions used for the  $h \rightarrow \text{inv.}$  combination are optimized to different two-body supersymmetric decay scenarios.

For the 2L analysis, the dominant backgrounds of  $t\bar{t}$ ,  $t\bar{t}Z$ , and single-top decays are estimated through normalization factors computed from CRs that differ from the SR in lepton multiplicity among other variables. The 0L analysis defines CRs to extract normalization factors for  $V+\text{jets}$ ,  $t\bar{t}Z$ , and  $t\bar{t}$  backgrounds.

| Quantity                                      | 2L                       | 0L A                             | 0L B                             |
|-----------------------------------------------|--------------------------|----------------------------------|----------------------------------|
| $N_\ell$                                      | 2                        | 0                                | 0                                |
| $N_{\text{jet}}$                              | -                        | $\geq 4$                         | $\geq 4$                         |
| $N_b$                                         | $\geq 1$                 | $\geq 2$                         | $\geq 2$                         |
| $p_T(\ell)$                                   | $> (25, 20) \text{ GeV}$ | -                                | -                                |
| $p_T(\text{jet})$                             | -                        | $> (80, 80, 40, 40) \text{ GeV}$ | $> (80, 80, 40, 40) \text{ GeV}$ |
| $m_{\ell\ell}$                                | $> 20 \text{ GeV}$       | -                                | -                                |
| $m_{T, \text{min}}^{b, \text{min}}$           | -                        | $> 200 \text{ GeV}$              | -                                |
| $m_1^{R=1.2}$                                 | -                        | $> 120 \text{ GeV}$              | $> 120 \text{ GeV}$              |
| $m_{\text{leading}}^{R=0.8}$                  | -                        | $> 60 \text{ GeV}$               | -                                |
| $\Delta\phi_{\text{boost}}^1$                 | $< 1.5 \text{ rad}$      | -                                | -                                |
| $\Delta\phi_{\text{jet, MET}}$                | -                        | $> 0.4 \text{ rad}$              | $> 0.4 \text{ rad}$              |
| $E_T^{\text{miss}}$                           | -                        | $> 250 \text{ GeV}$              | $> 250 \text{ GeV}$              |
| $E_T^{\text{miss}}$ significance <sup>2</sup> | $> 12$                   | $> 25$                           | $> 14$                           |
| $m_{T2}^{\ell\ell}$ <sup>3</sup>              | $> 110$                  | $> 450 \text{ GeV}$              | $< 450 \text{ GeV}$              |
| $\Delta R(b_1, b_2)^4$                        | -                        | $> 1.0$                          | $> 1.4$                          |

TABLE 6.2. Event selection criteria for included  $t\bar{t} + E_T^{\text{miss}}$  signal regions

<sup>1</sup> Azimuthal angle between  $\mathbf{p}_T^{\text{miss}}$  and  $\mathbf{p}_{T, \text{boost}}^{\ell\ell}$  (vectorial sum of  $\mathbf{p}_T^{\text{miss}}$ ,  $\mathbf{p}_T(\ell_1)$ , and  $\mathbf{p}_T(\ell_2)$ ).

<sup>2</sup>  $E_T^{\text{miss}}$  sig. =  $\frac{E_T^{\text{miss}}}{\sqrt{\sigma_L^2(1-\rho_{LT}^2)}}$ ,  $\sigma_L$ : expected resolution of the total longitudinal momentum (relative to the direction of  $p_T^{\text{miss}}$ ) of all objects,  $\rho_{LT}$ : correlation factor between the longitudinal and transverse momentum resolutions for all objects.

<sup>3</sup> Stransverse mass, incorporating lepton momenta and missing momentum. For 0L, top candidates are used rather than leptons and a  $\chi^2$ -like penalty function is utilized ( $m_{T2, \chi^2}$ ).

<sup>4</sup> Only used in the 0L A bin in which two tops are fully reconstructed.

The main overlap between these two channels is in their  $t\bar{t}$  Z CRs, where the CR of 0L is nearly contained within the CR of 2L. For the  $t\bar{t}H$  combination, the higher statistics CR of 2L is used to normalize both analyses simultaneously. This introduces a higher uncertainty on the 0L extrapolation, but the sensitivity is not significantly altered as the  $t\bar{t}$  Z contribution to the SR of 0L is not dominant (12-26%). Theory uncertainties from this background are kept uncorrelated as they are computed differently for the two analyses. Experimental uncertainties are all correlated.

Combining the 2L and 0L channels improves the overall expected limit by nearly 15% relative over the more sensitive 2L channel (see Table 6.3), measuring  $\mathcal{B}_{h \rightarrow \text{inv}} < 0.40$  (0.360) obs (exp).

| Analysis    | Observed UL | Expected UL |
|-------------|-------------|-------------|
| 0L          | 0.94        | 0.639       |
| 2L          | 0.37        | 0.423       |
| Combination | 0.40        | 0.360       |

TABLE 6.3. Expected and observed ULs at 95% C.L. from the individual  $t\bar{t}H$  input analyses and their combination

### 6.2.3. Run 1 Combination

The Run 1 ATLAS  $h \rightarrow \text{inv.}$  combination utilized  $4.7 \text{ fb}^{-1}$  of  $pp$  collision data at  $\sqrt{s} = 7 \text{ TeV}$  and  $20.3 \text{ fb}^{-1}$  at  $\sqrt{s} = 8 \text{ TeV}$  [117]. This combination considers inputs from direct detection of  $\mathcal{B}_{h \rightarrow \text{inv}}$  through Higgs bosons produced through VBF and in association with a vector boson  $V$  where the vector boson decays leptonically ( $Z \rightarrow \ell\ell$ ) or hadronically ( $W/Z \rightarrow jj$ ).

Nuisance parameters between the three inputs were left uncorrelated with the exception of luminosity, scaling and resolution of jet energy calibration, parton shower modeling, and factorization/renormalisation scale and PDF uncertainties. This subset was treated as fully correlated.

All signal regions and data control regions are used to perform a maximum-likelihood fit with the  $\text{CL}_S$  technique, using the profiled likelihood ratio as a test statistic. This resulted in an observed (expected) upper limit of  $\mathcal{B}_{h \rightarrow \text{inv}} < 0.25$  (0.27) at 95% C.L.

#### 6.2.4. Other Considered Analyses

Other analyses were considered for inclusion in this work, but had not advanced fully enough to be included in a public result. Many are still considered for inclusion in the upcoming peer-reviewed paper.

The dominant Higgs production mode, ggF, can be considered if the signal contains ISR used for tagging. This signal consists of one energetic jet. In the time since the publication of this combination, the UL  $\mathcal{B}_{h \rightarrow \text{inv}} < 0.34$  (0.39) obs (exp) was offered as an interpretation of this channel [138].

The signal-over-background ratio of the VBF search can be improved by requiring an ISR photon, leading to a signature with two VBF-like jets and one photon. Though lower statistics than the standard VBF search, the additional photon leads to powerful background discrimination. This search was conducted for the first time in ATLAS and measured  $\mathcal{B}_{h \rightarrow \text{inv}} < 0.37$  (0.34) exp (obs) [139]. The result became public after the announcement of these combined results, so it was not included in this round.

Many studies are also underway to measure  $\mathcal{B}_{h \rightarrow \text{inv}}$  for Higgs produced with a vector boson. The  $Z(\ell\ell)H(\text{inv.})$  channel is very sensitive due to the two lepton signal.  $W(\ell\nu)H(\text{inv.})$  is being considered, though the presence of the neutrino complicates the MET measurement. There is also movement on  $V(\text{had.})H(\text{inv.})$ .

Some of these analyses contributed material for consideration during the optimization process of this combination effort, and these preliminary inputs are included in the discussion even though they were ultimately not included in the final result.

### 6.3. Shared MC Samples and Duplicated Event Numbers

The combination team coordinated the Monte Carlo production of Higgs signal sample files, so that all input analyses use the same Higgs signal samples. In the process of conducting optimization studies, it was noticed that some input ntuples (provided by the analyses) contained what appeared to be duplicated events. This was first noted by the VBF+MET analysis, and was seen in the VBF MC sample (DSID 346600) where  $\sim 1\%$  of events were duplicated. These events shared event numbers but had differing kinematics, such as jet multiplicity.

The problem was traced back to the derivations, and was also seen in the  $ttH$  semi-leptonic MC sample (DSID 346632). Both of these samples were produced with Powheg+Pythia8 event generation and showered with Pythia. They both also include a filter such that  $E_T^{\text{miss}} > 75$  GeV, which has a 50% efficiency.

In the VBF case, 1500 LHE files each contain 6000 events with matrix element level truth information. Each LHE file is used as input to create one showered file containing only 1000 events which respect the filter. A configuration file is generated for the showering, which defines this maximum number of final events and also defines the first event number to be used in the output. The Athena version used to shower the files was MCProd20.7.9.9.19. With this version, each LHE event is assigned an event number (incremental from the one preceding it where the first value for the file is set in the configuration), showered, and checked against the MET filter. If the MET threshold is not met, the event is discarded and the next LHE event is considered. This results in an output showered file with 1000 events but event numbers exceeding 1000 greater than the initial value, as half of the considered events do not pass the filter.



However, showering configuration was defined such that each new showered file began with an event number 1000 greater than the file before (resulting, in theory, in a continuous spread of numbers across the entire collection of showered files). Due to this, when multiple files are compared there is overlap because the event numbers truly contained within each file are not simply 1-1000.

Though event numbers are duplicated, the original LHE events are not. Analyses then can neglect this effect in their own studies, as true events are not duplicated, only event numbers are reused. For the combination, this complicated the task of defining overlap. To mitigate this, the additional variable of truth Higgs  $p_T$  is also considered which discriminates between different events that may happen to share the same event number.

#### 6.4. Orthogonality

The success of a bias-free combination requires that the contributing analyses consider different subsets of data events that best suit their analysis selections - that is to say, analyses must be as orthogonal to one another as possible. Each contributing analysis provided a collection of the events captured by their selections for both signal and control regions from shared Monte Carlo Higgs signal campaigns for MC<sup>5</sup> and data. Due to known issues with the generation of some samples (see Section 6.3), overlapping events are determined by comparing run number, event number, and the truth  $p_T$  of the simulated Higgs boson. Only if all three conditions match is the event identified as overlapping. This is done on an event-by-event basis, considering raw event yields (with no weighting or

---

<sup>5</sup>The sum of the three MC campaigns representing the three years of Run 2 data-taking is considered unless otherwise specified.

cross-section consideration). Interpretations of these yields in the context of signal compositions for each analysis is also considered.

The two Run 2 analyses used in the combination ( $\text{VBF} + E_{\text{T}}^{\text{miss}}$  and  $t\bar{t} + E_{\text{T}}^{\text{miss}}$ ) are orthogonal by design from  $b$ -tagged jet requirements. The  $b$ -tag requirement of the VBF analysis was included specifically for orthogonality with  $t\bar{t} + E_{\text{T}}^{\text{miss}}$ . These are orthogonal to Run 1 (different data sets).

The combination effort began with many Run 2 analyses being considered though only three of them ( $\text{VBF} + E_{\text{T}}^{\text{miss}}$ ,  $t\bar{t} + E_{\text{T}}^{\text{miss}}$  2L, and  $t\bar{t} + E_{\text{T}}^{\text{miss}}$  0L) enter this combination. The figures/tables presented below contain preliminary information about the monojet ( $H$  production through ggF) and  $\text{VBF} + \gamma$  analyses as well. Data is only considered for the public works ultimately used in the combination.

The overlap among analyses is first tested on signal MC samples<sup>6</sup>. No overlapping events are found between  $\text{VBF} + E_{\text{T}}^{\text{miss}}$  and  $t\bar{t} + E_{\text{T}}^{\text{miss}}$  in samples of Higgs production through VBF, ggF, or  $t\bar{t}H$ <sup>7</sup> (see Figure 6.1).

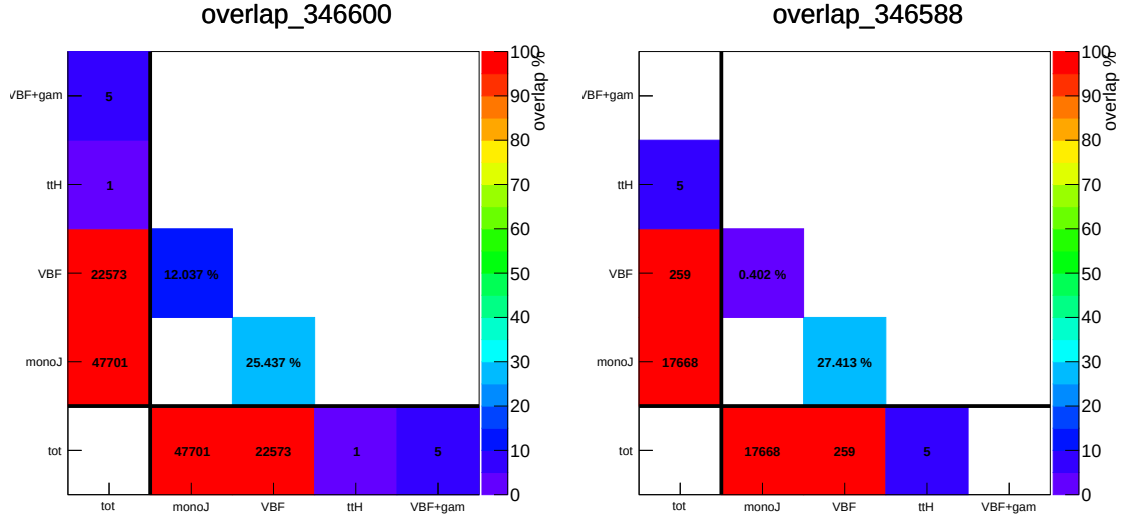
Overlap fraction values, designated with a %, show the fraction of events from the sample on the horizontal axis that are shared with the sample on the vertical axis. An absence of overlap in the diagonal implies that each sample has no duplicated events within itself, after both the event number and truth Higgs  $p_{\text{T}}$  are considered. These numbers result from the consideration of every event chosen for the analysis, in any signal or control region.

Overlap in data is also considered for the analyses included in the combination. The plots from data confirm the picture obtained by investigating

---

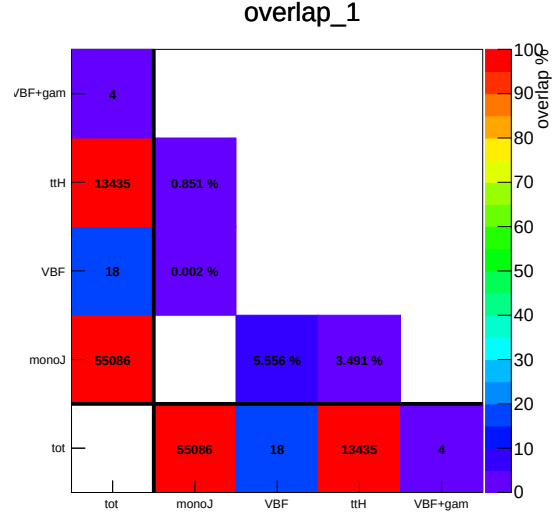
<sup>6</sup>Only the MC16e campaign is considered for  $\text{VBF} + E_{\text{T}}^{\text{miss}}$ .

<sup>7</sup>DSIDs 346632, 346633, and 346634



(a) VBF sample

(b) ggF sample



(c)  $ttH$  samples

FIGURE 6.1. Yields and overlap fraction of all analyses for signal samples

the MC for the signal processes;  $VBF + E_T^{\text{miss}}$  and  $t\bar{t} + E_T^{\text{miss}}$  analyses are orthogonal (see Figure 6.2).

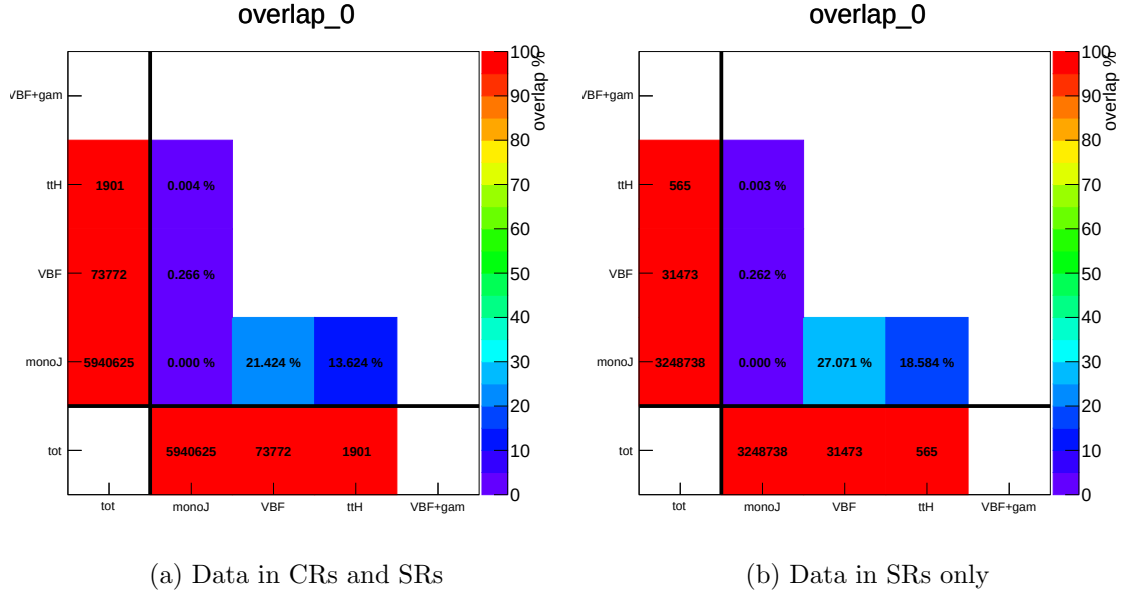


FIGURE 6.2. Yields and overlap fraction of all analyses in observed data events for (a) all SR and CR events, (b) for only events in the signal regions.

## 6.5. Impact on Limit

Before combining the individual analyses into the combination in order to set a combined limit, the sources of systematic uncertainties must be understood. Some sources can be correlated among the different analyses (and additionally between the different runs), but not all can. In order to help inform this correlation scheme, the impact that these NPs play on the limit of each individual analysis was considered. This was done by defining groups of similar NPs (such as JES or MET NPs) and recomputing the limit of each analysis while holding one set at its best-fit value and considering the improvement in the limit. This illustrates which NPs play the most influential role for each individual analysis, and hints toward what will impact the combination once correlations are considered.

NPs are grouped into categories based upon their source: jet energy resolution (JER), jet energy scale (JES), jet vertex tagging (JVT), missing transverse energy (MET), flavor tagging, lepton and photon identification/isolation/reconstruction, theoretical uncertainties associated with signal samples, theoretical uncertainties associated with background samples, data-driven uncertainties on backgrounds, normalization factors derived from control regions (dubbed as “Background Normalization”), MC statistics, and all others (containing mostly NPs for luminosity and pileup reweighting). An additional category considers a data-only scenario where all systematics are held fixed.

Each analysis is capable of setting an upper limit on  $\mathcal{B}_{h \rightarrow \text{inv}}$  with a full set of systematic uncertainties. This optimized POI value serves as the baseline for comparison, denoted  $\mu$ . The effect of the aforementioned sets of NPs on the POI  $\mu$  is computed by re-fitting the entire analysis while holding the NPs contained within the set to their best-fit values. All other NPs are allowed to float, as usual. The fit POI, denoted  $\mu_{\text{fixed\_set}}$  is then quadratically subtracted from  $\mu$  (see Table 6.4).

Due to residual correlations between categories, the quadratic sum of impacts can differ from the calculated impact from a full fit. The impact due to finite number of data events (“data statistical uncertainty”) is obtained by ignoring all NPs associated with systematic uncertainties and the floating background normalizations. The impact from all systematic uncertainties is estimated by quadratically subtracting the above impact and the one from floating background normalizations from the total impact. The impact from experimental sources and the impact related to the size of the MC sample (“MC statistical uncertainty”) are treated as separate categories.

| Source of uncertainty         | $\pm$ Uncertainty on $\mathcal{B}_{h \rightarrow \text{inv}}$ |         |
|-------------------------------|---------------------------------------------------------------|---------|
|                               | Run 2                                                         | Run 1+2 |
| Luminosity/PRW                | 0.002                                                         | 0.003   |
| Leptons+Photons               | 0.018                                                         | 0.015   |
| Jets                          | 0.023                                                         | 0.019   |
| Flavor tagging                | 0.002                                                         | 0.002   |
| MET                           | 0.008                                                         | 0.007   |
| V+jets modeling               | 0.011                                                         | 0.017   |
| Other backgr. modeling        | 0.015                                                         | 0.015   |
| Data-driven background        | 0.023                                                         | 0.019   |
| Signal modeling               | 0.004                                                         | 0.003   |
| MC statistics                 | 0.023                                                         | 0.021   |
| All experimental              | 0.041                                                         | 0.036   |
| All theory                    | 0.030                                                         | 0.030   |
| Total systematic uncertainty  | 0.051                                                         | 0.046   |
| Data statistics               | 0.019                                                         | 0.018   |
| Floating bkgd. norm.          | 0.031                                                         | 0.028   |
| Total statistical uncertainty | 0.037                                                         | 0.034   |
| Total uncertainty             | 0.063                                                         | 0.057   |

TABLE 6.4. Impact of groups of NPs on the best-fit POI value

The most relevant sources of uncertainty for the full combination are very similar to those of the Run 2 combination. One exception lies with  $V$ +jets modeling where the impact increases for the full combination. This is because the Run 1 combination relies heavily on MC simulations of this process due to the limited statistical power of the CRs, increasing the overall NP group impact.

## 6.6. Fitting and Results

Scans of the negative logarithmic profile likelihood (NLL) ratios,  $-2\Delta\ln(\Lambda(\mu; \theta))$ , in the best-fit signal strength  $\mu$  ( $\mathcal{B}_{h \rightarrow \text{inv}}$ ) illustrate the change in sensitivity as more inputs are included in the combination (see Figure 6.3) and check goodness of fit. Horizontal lines illustrate  $1\sigma$  and  $2\sigma$  error bands. The  $\text{CL}_S$

observed limit at 95% C.L. can be approximated off this plot where each line crosses  $-2\Delta\ln(\Lambda) = 3.84$  or  $1.96\sigma$ . These limits are shown in Figure 6.4 and Table 6.5.

The limit set by this combination,  $\mathcal{B}_{h\rightarrow\text{inv}} < 0.11$  (0.11), improves the sensitivity of the expected limit over the VBF result by a relative 13.7%. This is the strictest upper limit set on this process to date.

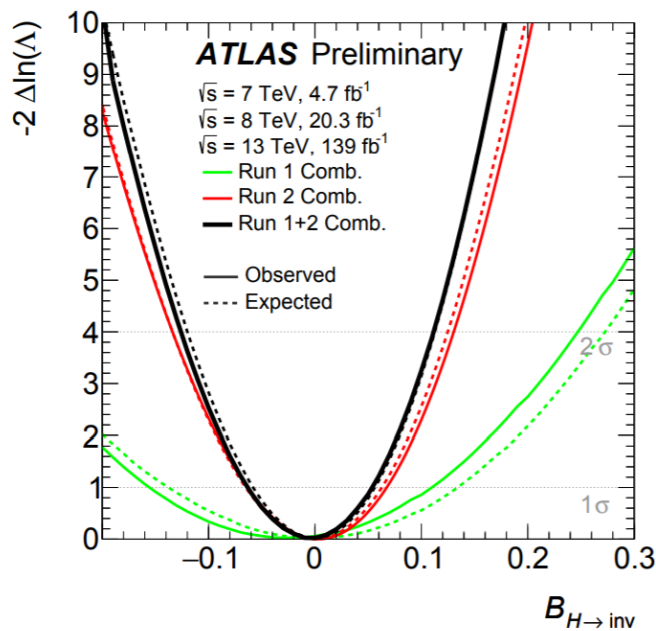


FIGURE 6.3. Observed and expected NLL scan of combinations

| Results     | Best Fit $\mu$          | Expected | +1 $\sigma$ | -1 $\sigma$ | +2 $\sigma$ | -2 $\sigma$ | Observed |
|-------------|-------------------------|----------|-------------|-------------|-------------|-------------|----------|
| ttH         | $0.04^{+0.20}_{-0.20}$  | 0.360    | 0.502       | 0.259       | 0.684       | 0.193       | 0.402    |
| Run 1       | $-0.02^{+0.14}_{-0.13}$ | 0.265    | 0.370       | 0.191       | 0.501       | 0.142       | 0.252    |
| VBF         | $0.00^{+0.07}_{-0.07}$  | 0.131    | 0.181       | 0.094       | 0.244       | 0.070       | 0.130    |
| comb. Run 2 | $0.00^{+0.06}_{-0.07}$  | 0.125    | 0.172       | 0.090       | 0.232       | 0.067       | 0.127    |
| combined    | $0.00^{+0.06}_{-0.06}$  | 0.113    | 0.156       | 0.081       | 0.209       | 0.060       | 0.111    |

TABLE 6.5. Summary of 95% CL limits on  $\mathcal{B}_{h\rightarrow\text{inv}}$  of the individual search regions

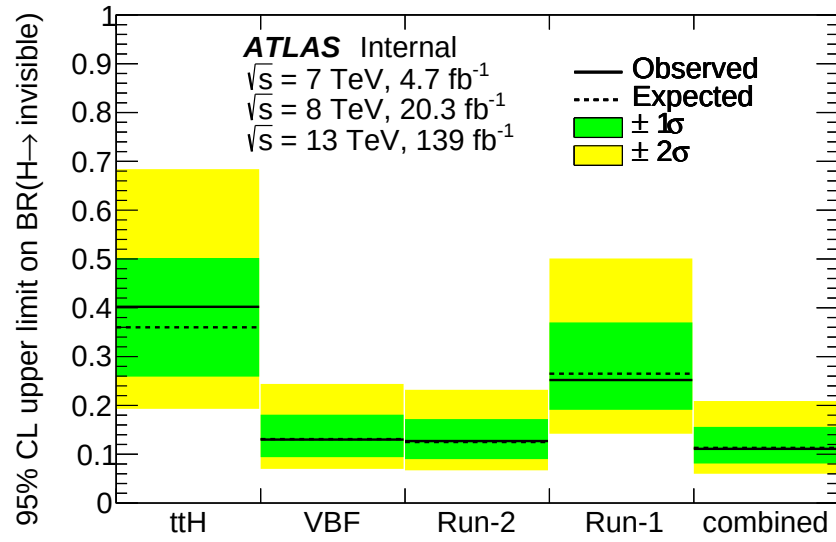


FIGURE 6.4. Observed and expected 95% CL limits for input analyses and combinations



## CHAPTER VII

### $h \rightarrow$ invisible IN FUTURE COLLIDERS

#### 7.1. Future Endeavors

The LHC has provided proton and heavy ion beams to experiments since 2009, and has plans to continue for another 15 years (see Figure 7.1). This inspires a look to the future and the next large particle collider. Concurrent with LHC operation, the high energy physics communities in different regions gather to organize reports of the prospects of new projects and technologies. These reports help guide the physics priorities within the field for the coming years.



FIGURE 7.1. LHC schedule for future running [26]

Europe summarizes their priorities in the European Strategy for Particle Physics [140] - a report that is updated roughly every seven years which provides recommendations to the community regarding how to best direct the field of high energy. The original document was published and adopted by the CERN council in 2006 [141], and the most recent update was publicized in 2020 [142]. Among other things, the report provides encouragement for the future projects that

are understood to most enhance the field. Elements of the report then inform governments and funding agencies of the consensus of the community to help organize global efforts.

The American equivalent of the European Strategy is the Particle Physics Project Prioritization Panel (P5) report [143] which is published by the Department of Energy. It is informed by the high energy physics community's multi-year long Particle Physics Community Planning Exercise [144], which culminates in a report published by the American Physical Society. The last P5 report was published in 2014 [145] and the next round is set to be compiled in 2022. This report also prioritizes future projects and defines the role that American physicists and institutions can play in realizing these global goals.

Planning in Asia is coordinated by individual countries, including Japan and China, and promoted throughout the region as a whole through the Asian Community for Future Accelerators (ACFA) [146].

Each of these reports are updated roughly every few years, but plan for ten to twenty years into the future of the field. They rely upon technical input from the high energy physics community to inform governments and funding agencies about the most promising future plans. Global coordination and collaboration is promoted by the International Committee for Future Accelerators (ICFA) [147].

A successful example of a shared area of support of these reports is a major upgrade to the LHC and detectors for Run 4 and beyond. Here, a period of High Luminosity running [148] is set to begin (see Section 7.2) which should culminate at the end of the LHC era<sup>1</sup> with 3000 fb<sup>-1</sup>.

---

<sup>1</sup>One possible extension beyond Run 5 is the High Energy LHC (HE-LHC), using the LHC tunnel for  $\sqrt{s}=27$  TeV  $pp$  collisions [149].

The community must also look further into the future, beyond the LHC. New projects require long time scales to account for project planning and construction, so they must be considered before the end of the LHC era. One project that has enjoyed continued support from both the European Strategy and the P5 report is the construction of the International Linear Collider (ILC) - an electron-positron collider [150]. Its lepton environment differs greatly from the hadron environment of the LHC and lends itself to precision measurements (see Section 7.4).

## 7.2. High Luminosity-LHC

Beyond Run 3, the LHC looks forward to the High Luminosity LHC (HL-LHC). This physics program involves long-term running of  $\sqrt{s} = 14$  TeV  $pp$  collisions (see Figure 7.1) with an average pileup  $\langle\mu\rangle$  of 140 (see Figure 7.2) in order to collect up to  $3 \text{ ab}^{-1}$  of data - nearly ten times more than the original LHC design - by the end of running in 2036. The increased luminosity can lead to higher precision PDF modeling and more precise direct measurements of Standard Model particles, including the Higgs boson. It can also provide constraints on BSM physics models 20-50% stronger than what can be achieved at the end of Run 3 [151].

In the HL-LHC era, the LHC machine itself and the accelerator complex (see Section 3.1) will receive a few upgrades. The increase in luminosity comes in part from an increase in the peak instantaneous luminosity from the design value of  $1 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$  to  $5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$  producing an integrated luminosity of nearly  $250 \text{ fb}^{-1}$  per year [152]. This will be done in part by compressing the beam more tightly and precisely by incorporating new technology for beam collimation at the collision points, and by introducing compact ultra-precise superconducting cavities

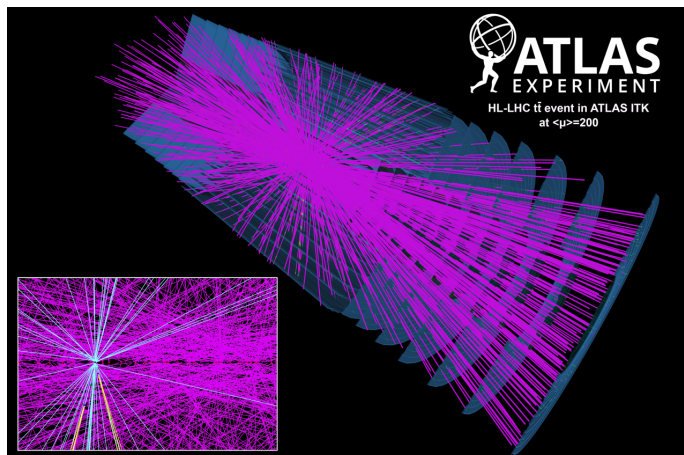


FIGURE 7.2. A simulated  $t\bar{t}$  HL-LHC event with two hundred reconstructed vertices [27]

for beam rotation. The LHC also plans to upgrade the superconducting magnets used for beam acceleration. This increased luminosity comes at the price of higher pileup, with  $\langle\mu\rangle = 140$ .

The LHC has outperformed every milestone set for it thus far. If this trend continues, the HL-LHC could outperform its goals as well if it can be run at the peak of its luminosity reach. This ultimate running would produce up to  $4 \text{ ab}^{-1}$  of data with a peak instantaneous luminosity of  $7.5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ , delivering up to  $350 \text{ fb}^{-1}$  each year at  $\langle\mu\rangle = 200$  [152].

Increasing the instantaneous luminosity also increases the pileup at each bunch crossing, and this high pileup environment creates its own challenges for detectors. Both software and hardware upgrade projects aim to increase the performance of ATLAS in this intense environment [153].

The subsystem most directly affected by this increase in pileup is the ID, so for the HL-LHC the entire ID will be replaced with the inner tracker (ITk) [154, 155]. Like the current ID, ITk will use layers of pixelated silicon sensors close to the IP to achieve high spatial resolution and strips of silicon further away from

the IP, which extend further forward than the current ID to  $|\eta| < 4$  with central barrels and forward rings. The all-silicon design adds precision but also longevity as ITk will receive nearly ten times the radiation dose of the ID in Runs 1-3.

A new forward detector will be introduced with the goal of reducing forward pileup jets. The High-Granularity Timing Detector (HGTD) [28] will sit from  $2.4 < |\eta| < 4.0$  and measure the timing of forward charged particles with a resolution of 30 ps (see Figure 7.3). This will allow for more precise vertex identification, leading to increased pileup rejection.

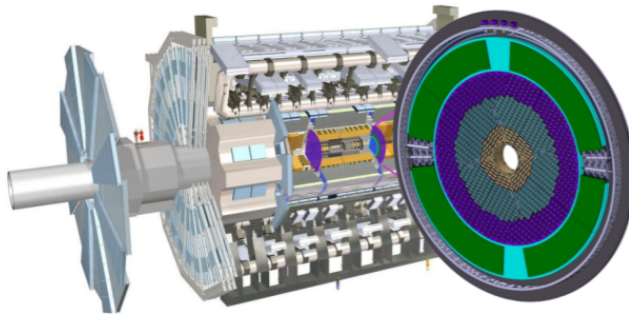


FIGURE 7.3. Placement of HGTD (green and purple disks) in the upgraded ATLAS detector for HL-LHC [28]

The calorimeters and muon system will receive upgraded electronics to ensure that timely readout can be maintained even as online pileup correction techniques are used and to protect against the increased radiation dosage.

The TDAQ system will also be upgraded for HL-LHC running, using a hardware-based Level-0 trigger and software-based Event Filter (EF) to reduce the rate of data to 10 kHz. Part of the increased rate is due to track-based triggers, made possible by ITk, which is processed on a devoted Hardware Tracker system that runs in parallel with the EF to reconstruct these tracks and further inform EF decisions. A new Global Trigger at Level-0 will perform algorithms on fine-

granularity calorimeter cells, protecting MET trigger performance in this extreme environment.

### 7.3. HL-LHC $h \rightarrow$ invisible **Projections**

The increased luminosity provided by the HL-LHC as well as its more challenging pileup environment will impact the measurement of  $h \rightarrow$  invisible. An investigation to project the VBF+MET  $h \rightarrow$  invisible search into the future uses the Run 2 VBF+MET analysis and statistics framework, but with samples that emulate the the HL-LHC environment. One way this done is by creating new simulated files that account for the HL-LHC environment. This is challenging for VBF+MET as the search has been plagued by large systematic errors from lacking Monte Carlo statistics in the past and therefore a re-simulation of samples is not feasible.

Another approach is to use existing truth-level samples and smear the truth objects to approximate their response at the HL-LHC. I implemented this truth-level smearing and created a comprehensive set of samples that emulate the HL-LHC environment. These will be used with the existing analysis framework to estimate the impact of the HL-LHC on the measurement and to identify potential problem areas to allow the analysis team to consider mitigation and re-optimization to refine the result.

A centrally-produced software package (`UpgradePerformanceFunctions`) is used to estimate the performance of the upgraded ATLAS detector in the HL-LHC. Parameterized estimates are applied separately to each truth-level object as derived from dedicated studies by combined performance and analysis teams. Resolutions, or smearing functions, and efficiency functions are provided

based upon the predicted detector response. A dictionary of pileup events is also included and overlaid to emulate  $\langle\mu\rangle = 200$ . The forward HGTD can be evoked to help with forward pileup mitigation, but samples including this detector have not yet been made.

Truth-level smearing is done at the first step of the analysis chain where variables relevant for the analysis are extracted from large xAOD simulation files and saved in more manageable “MiniNtuple” files. TRUTH3<sup>2</sup> derivations are used as inputs. The smeared quantities are then saved in the output in place of the original truth-level quantities. In this way, the remainder of the analysis chain is not disrupted.

Each truth-level object (electron, photon, muon, jet) is considered individually for smearing (see Appendix G for details of how each individual object is treated). The transverse momentum of all leptons and photons is smeared while the position is copied from truth information. Jets are treated separately, as there is a two-stage process that must occur. First, every truth-level jet is fully smeared by detector resolution functions. The energy, momentum, and position is all considered in the smearing. Then predefined pileup events are added these smeared jets such that  $\langle\mu\rangle = 200$ . In order to be considered by the analysis script, the smeared hard scatter or pileup jet must pass a  $p_T$  and  $\eta$  requirement. These conditions are necessary as only jets that satisfy a JVT requirement are passed back to the analysis software for further consideration (details in Appendix G), and the application of JVT is  $p_T$  and  $\eta$  dependent. The returned jet collection is then a combination of appropriate smeared hard scatter jets and pileup jets (see Figure 7.4).

---

<sup>2</sup>TRUTH3 is a derivation intended for analysis use that retains truth particle, jet, and summary information. It does not contain a full or partial truth record [156].

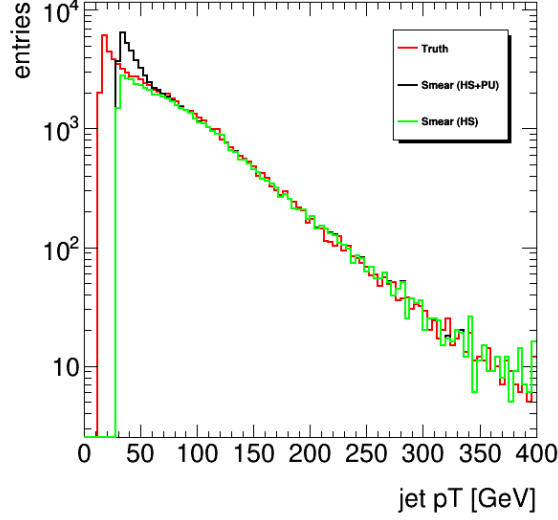


FIGURE 7.4. Smeared jet collection from VBF signal sample compared to truth

These smeared objects are used first for overlap removal (see Appendix E). Objects that pass the overlap removal requirements are then saved to the output trees, provided they pass additional  $p_T$  thresholds, in an effort to reduce the output file size by only retaining relevant objects.

Additional objects are built using these smeared physics objects.  $E_T^{\text{miss}}$  is built from smeared jets with  $p_T > 30$  GeV and smeared leptons with  $p_T > 5$  GeV to ensure self-consistency in the output, rather than using a smeared  $E_T^{\text{miss}}$  value. The soft  $E_T^{\text{miss}}$  term is approximated independently, as smeared tracks are not directly considered in this  $E_T^{\text{miss}}$  calculation. The inclusion of pileup jets adds confusion to the  $E_T^{\text{miss}}$  calculation leading to long tails in high- $E_T^{\text{miss}}$  for some samples (see Appendix G). The  $p_T$  threshold is increased to 50 GeV for forward jets ( $|\eta| > 3.8$ ), as this region is dominated by low- $p_T$  pileup jets. A jet-only  $E_T^{\text{jet, no JVT}}$  value is also calculated identically with only jets and no JVT requirement.



The smeared samples can then be used in the standard analysis flow to consider kinematic distributions and calculate the expected limit. The analysis could be further optimized by reconsidering kinematic cuts, such as the subleading jet  $p_T$  threshold and third- and fourth-jet criteria. At this time, these studies are ongoing.

The results for performance in this VBF channel can be compared to projected results for the HL-LHC [157]. VBF production alone (at 14 TeV with  $3 \text{ ab}^{-1}$ ) is projected to measure  $\mathcal{B}_{h \rightarrow \text{inv}} < 0.038$ . A statistical combination of searches from both ATLAS and CMS, as well as the inclusion of the  $VH$  production mode, could refine this projection to  $\mathcal{B}_{h \rightarrow \text{inv}} < 0.025$ .

#### 7.4. The International Linear Collider

The LHC collides hadrons (protons), but protons are not themselves fundamental particles. Their composite nature and the reliance upon PDFs limits what can be known about the initial state of interacting particles (such as the  $z$  momentum in the direction of the beamline) and leads to large rates of QCD multijet background. These problems are inherently mitigated by colliding leptons instead, themselves fundamental particles. Due to synchrotron radiation, the most ideal environment for colliding high-energy electron<sup>3</sup> beams is in a linear machine rather than the circular LHC. The clean collision environment of fundamental particles (see Figure 7.5) allows for precision work to be done.

One particle that benefits from precision work is the Higgs boson, as emphasized by the 2020 European Strategy update. This can be done by building a “Higgs Factory” - a collider that produces many Higgs bosons for precision

---

<sup>3</sup>Heavier muons have lower synchrotron radiation, but other technical difficulties (such as a finite lifetime).

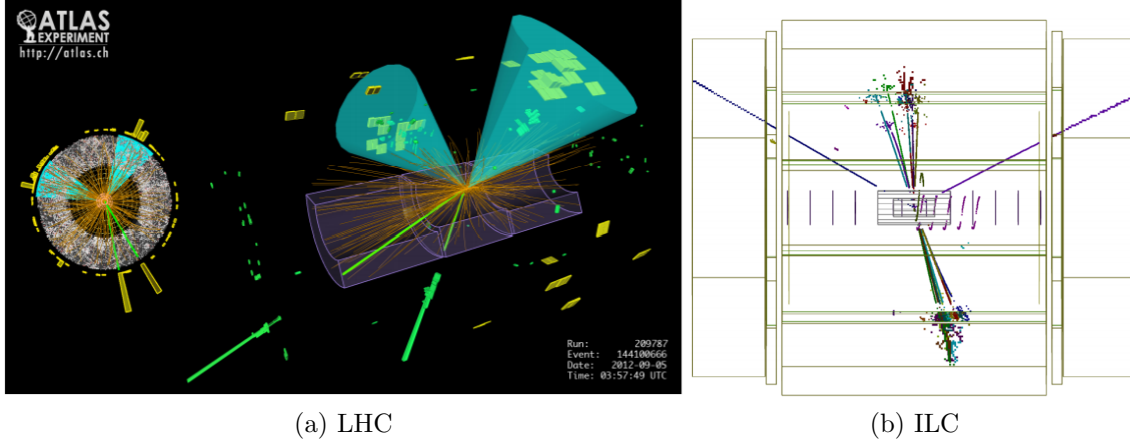


FIGURE 7.5. Simulated production of  $ZH \rightarrow \ell\ell b\bar{b}$  in the LHC ( $\ell = e$ ) with ATLAS in Run 1 (a) and in the ILC ( $\ell = \mu$ ) with ILD (b) [29, 30]

study. One proposed facility to address this need is the International Linear Collider (ILC).

The ILC is a future project with international support intended to be built in Japan and feature two independent detectors to record the interactions of polarized electron and positron beams. These beams will be accelerated with superconducting radio-frequency (SRF) accelerators [30] to center-of-mass collision energy of 250 GeV (ILC250), though the full physics program anticipates upgrades to  $\sqrt{s} = 500$  GeV and beyond. This energy requires linear tunnels each of a length of 10 m (see Figure 7.6) to house the linear accelerators. The SRF cavities will be made of nitrogen-doped niobium and can reliably achieve an acceleration gradient of 35 MV/m [32] - higher than the design of 31.5 MV/m.

The ILC will structure beams into trains each containing 1316 bunches. These trains interact at a rate of 5 Hz, with about 308 ns between bunches [31]. At the interaction point, the beams will be squeezed by magnets into ribbons with a cross sectional area of  $0.004 \mu\text{m}^2$  to increase the probability of particle

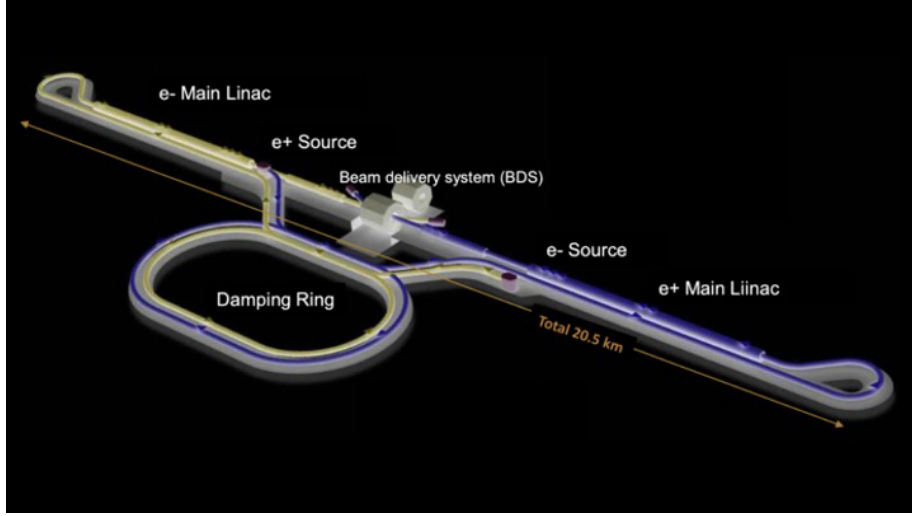


FIGURE 7.6. ILC layout [31]

interaction. These tightly focused beams allow for an instantaneous luminosity of  $0.75 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$  - a comparable to the LHC's Run 2.

The ILC only has one interaction point, but two approved detector designs are planned. These detectors, the International Large Detector (ILD) and Silicon Detector (SiD) [158] will share luminosity equally and operate with independent collaborations to provide cross-checks on physics measurements and increase opportunities for participation. Only one detector can collect data at a time, so a push-pull system will be employed to move one detector on its platform into the beamline as the other is moved out. ILD features a gas-based time projection chamber for tracking. SiD is the smaller of the two detectors, which is compensated for with a stronger magnetic field (5T compared to the 3.5 T field of ILD). It is designed with a silicon based tracker and electromagnetic calorimeter in order to be compact and cost effective (see Section 7.5). Both detectors are designed with high granularity in order to utilize particle-flow reconstruction [90] for precision measurements. The long quiet time between bunches allows the detector components to rest for nearly 200 ms between collisions in a quiescent

state that requires far less power. This “power-pulsing” creates diminished costs for power consumption and eliminates the need for cooling systems in the inner detectors.

The electron-positron annihilation interaction allows for a relatively large percentage of interactions to yield Higgs bosons, earning the Higgs Factory distinction. The ILC250 will optimize Higgs boson production through Higgsstrahlung ( $e^+e^- \rightarrow ZH$ ) [159]. The benefits of a lepton environment are immediately evident when considering electroweak precision measurements (see Figure 7.7), as the ILC is projected to refine the precision of, for example, SM Higgs boson couplings to vector bosons to  $< 0.5\%$  whereas the HL-LHC alone could only approach 2.5%. The Higgs can be studied more inclusively through agnostic recoil measurements off the  $Z$  produced in association, allowing full advantage to be taken of the luminosity without limiting study to single Higgs decay mechanisms.

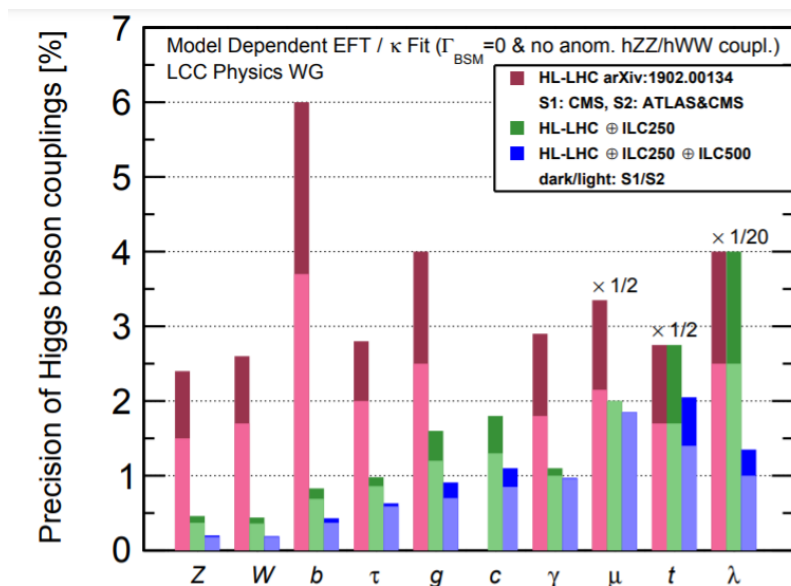


FIGURE 7.7. Projected precision of Higgs boson couplings of HL-LHC with and without ILC250 input [32]

The EM coupling constant is the same for all species of quarks and leptons, so  $e^+e^-$  annihilation produces pairs of all species at similar rates. This even includes many new particles from theories that preserve the electromagnetic coupling. New physics processes will have cross sections only a few orders of magnitude smaller than those of SM processes<sup>4</sup>, eliminating the need for a triggering system. This means that all events can be stored, and data is not biased by requiring notable SM signatures for triggering. It also means that all decay modes of particles are accessible offline - not only those used for triggering. This increases statistics and therefore the sensitivity of measurements. It also allows for the direct measurement of absolute branching ratios and total widths, rather than the comparisons and ratios done at the LHC. The lack of QCD background also allows for cleaner hadronic  $W$  and  $Z$  signatures, again bolstering statistics and precision.

At higher collision energies, the ILC can make precision measurements of other electroweak processes, such as the top Yukawa coupling or Higgs self-coupling. It also leads to a significant increase in the cross section for  $e^+e^- \rightarrow \nu\nu H$  where the  $H$  is produced through  $WW$  and  $ZZ$  fusion rather than Higgsstrahlung. This will allow for precise measurement of the  $HWW$  and  $HZZ$  couplings, which in turn will help refine the Higgs total width.

Precision studies of the SM carry BSM implications if any deviations from theory are discovered. In addition to Standard Model measurements though, the ILC can perform direct BSM searches as well. Lepton and hadron colliders are powerful when used in tandem, as hadron machines can push the energy frontier in search of new heavy particles and lepton machines can more precisely study

---

<sup>4</sup>At  $\sqrt{s}=500$  GeV, SM processes are  $\mathcal{O}(\text{pb})$  while new particle production processes are  $\mathcal{O}(10-100)$  fb [159].

what has been found. In the continued absence of new physics with the LHC, the ILC can also add to the search by considering, for example, extended Higgs sectors, supersymmetric models, and dark matter theories. It holds an advantage over the LHC by being able to probe BSM models that rely upon only electroweak couplings that occur at too low a rate in the LHC to be discernible over the large strong interaction background.

The ILC project has matured with decades of research and development work, and remains a viable option for the future of the field of high energy physics - especially with the continued international support provided by the 2020 European Strategy update [160] and 2014 P5 report. Its progress is reliant upon support from the international high energy community, and the interaction of scientists with governments and funding agencies around the world. As the proposed host country, Japan to date has conducted extensive internal study and discussion, resulting in the August 2020 announcement of the creation of an International Development Team that will consider early preparations for the construction of a Pre-Lab in Japan. This will serve as the predecessor of the ILC laboratory (see Figure 7.8), setting up the potential for ILC construction as early as 2026.



FIGURE 7.8. Anticipated timeline for the ILC [31]

## 7.5. SiD Detector Concept

One proposed detector concept, SiD, will feature high-granularity silicon tracking and calorimetry and is designed for particle flow reconstruction (see Figure 7.9). SiD is a multi-purpose detector design with  $4\pi$  solid angle coverage around the interaction point, envisioned to be applicable to a wide range of physics pursuits.

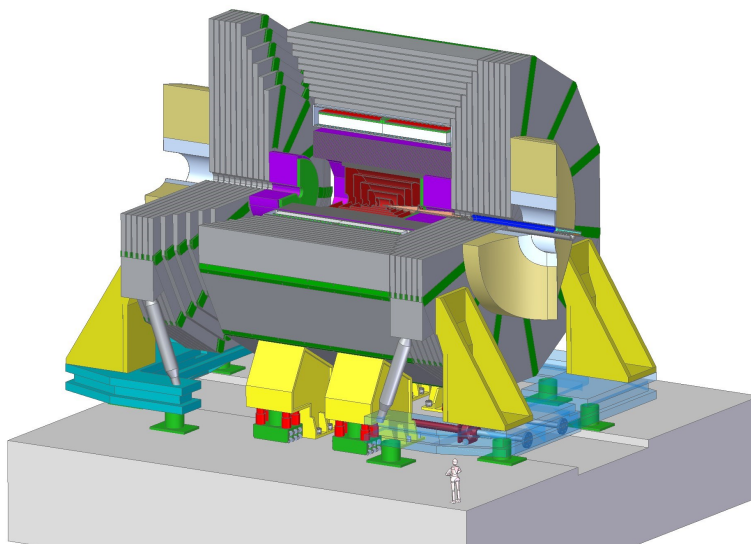


FIGURE 7.9. The SiD detector [30]

SiD is designed to be compact. This is possible in part by power-pulsing between bunch crossings, eliminating the need for much cooling infrastructure and reducing background rates between bunches. High resolution is reached by housing the granular tracking and calorimetry systems inside the solenoid, allowing for the compact design.

SiD will use silicon detectors for vertexing and tracking. The tracker's momentum resolution requirement is set by the necessity to identify Higgs bosons off recoiling  $Z$  bosons in ILC250. With  $\Delta p/p^2 = 5 \times 10^{-5} \text{ c/GeV}$  for charged tracks, reliable Higgs reconstruction is possible. The silicon-based design allows for precise

time-stamping of energy deposits. Five pixelated layers with  $20 \times 20 \mu\text{m}$  pixels comprise the central vertexing system, with five layers of silicon strips used for the central tracker. Disks are used for both systems in the endcap region.

Calorimeters must be granular enough to distinguish the two collimated jets from hadronic  $V$  decays in order to reconstruct recoiling Higgs bosons. This is achieved in the sampling electromagnetic calorimeter (ECal) with  $3.5 \times 3.5 \text{ mm}$  pixelated active silicon layers and a tungsten alloy used as an absorber. The first twenty layers of silicon and tungsten use 2.5 mm thick tungsten, and ten more layers in the back double the tungsten thickness to 5 mm for a total of thirty active silicon layers and  $26 X_0$  in the central region. In order to avoid projective cracks<sup>5</sup>, an overlapping trapezoidal design with 12-fold symmetry is used (see Figure 7.10).

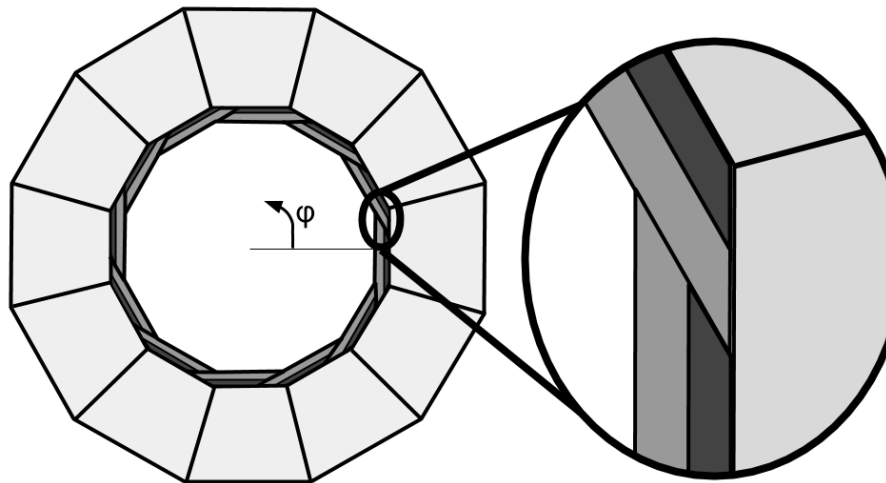


FIGURE 7.10. The SiD ECal (dark grays) and HCal (light gray) configuration, with overlapping trapezoidal modules with thin and thick absorber layers

---

<sup>5</sup>Projective cracks extend radially from the IP through the detector. In theory, a particle could be produced and traverse through the space between sensitive modules if the detector has projective cracks. This would lead to poor particle identification and reconstruction in those areas.



This design has been studied in detail to ensure that the development of EM showers in the overlapping region of neighboring trapezoidal modules is understood. I conducted studies to understand electromagnetic shower development in different areas of the detector - at normal incidence to all 30 layers and in these overlap regions. The nominal design of  $26 X_0$  largely contains electromagnetic showers without energy leakage into the hadronic calorimeter<sup>6</sup>, but the areas of overlap see a reduction of absorber length and therefore increased leakage that must be corrected for in reconstruction (see Appendix H for more details).

Prototypes of the design were tested with test beams. A nine-layer prototype collected data with a 12.1 GeV electron beam at SLAC in 2013. I designed an algorithm to clean the experimental data of underdeveloped showers<sup>7</sup> and soft photon radiation so that the data could be compared to a Geant4 simulation of the prototype setup. The simulation was also optimized to mimic experimental conditions, such as modeling events with multiple incident particles and simulating dead pixels noted during running. The cleaned data and optimized simulation output were also used to create a clustering algorithm to determine incident particle multiplicity. This algorithm was 98.5% effective in identifying two-electron events when the shower centers were separated by more than 1 cm. More details can be found in Appendix H.

The hadron calorimeter (HCal) is mounted directly to the ECal modules and has a mirroring 12-fold symmetry, though the trapezoidal modules do not overlap

---

<sup>6</sup>100 GeV incident electrons leak roughly 1% of their deposited energy in the hadronic calorimeter.

<sup>7</sup>Underdeveloped showers begin showering deeper in the detector than average. This leads to a deeper shower max and higher probability of shower leakage.

and contain projective cracks. This sampling calorimeter uses steel as an absorber and active layers of scintillating tiles for a depth of  $4.5\lambda$ .

The tracking system and calorimeters are enveloped in a 5 T magnetic field supplied by a superconducting solenoid. The strength of this field is necessary due to its compact size.

Forward calorimeters will complete the coverage in the high-radiation forward region and measure instantaneous luminosity with silicon and tungsten layers similar to the ECal.

Muons will likely scatter multiple times within the bulk of the detector before interacting with the muon system, so only centimeter-scale resolution is necessary. Nine layers of scintillator are proposed in the Technical Design Report [158].

## 7.6. $h \rightarrow$ invisible **Projections at the ILC**

Without QCD multijet background and operating as a Higgs factory, the ILC is ideal for measuring  $\mathcal{B}_{h \rightarrow \text{inv}}$  to a higher precision than that which can be achieved at the LHC.

The ILC250 will create Higgs bosons predominantly through Higgsstrahlung (see Figure 2.4), i.e. associated production of a  $Z$  boson. The clean ILC environment allows for an indirect Higgs search technique to be employed. The visible decay of the  $Z$  boson is first identified through leptonic or hadronic decays. Once the  $Z$   $p_T$  and mass are reconstructed, the Higgs can be identified as the invariant mass recoiling against the reconstructed  $Z$  as

$$m_{recoil}^2 = s + m_Z^2 - 2E_Z\sqrt{s} \ ,$$

where  $\sqrt{s}$  is the center of mass energy and  $m_Z$  and  $E_Z$  are the mass and energy of the reconstructed  $Z$  boson.

A search for  $h \rightarrow \text{inv.}$  utilizing MC data<sup>8</sup> was conducted with a simulation of the SiD detector. I optimized kinematic selections for the leptonic  $Z$  channel and was the first to consider the full potential performance of SiD considering the combination of leptonic and hadronic  $Z$  channels. The studies presented here should be considered preliminary, and are under continued development.

For both  $Z$  decay channels, a run scenario of  $1800 \text{ fb}^{-1}$  is considered. The beams are polarized  $(P_{e-}, P_{e+}) = (\mp 0.8, \pm 0.3)$  with  $900 \text{ fb}^{-1}$  collected for each polarization configuration<sup>9</sup>. Shorthand, these configurations are referred to by the particles left- or right-handedness (LR or RL).

Background sources are grouped by the number of fermions present. Backgrounds with 2-, 3-, and 4- fermions are relevant for these searches, as well as the inclusive SM decay of all Higgs produced through Higgsstrahlung. For orthogonality,  $h \rightarrow \text{inv.}$  events in the inclusive SM Higgs sample are not considered and instead are generated and included independently.

The branching ratio of the  $h \rightarrow \text{inv.}$  process is inflated to  $\mathcal{B}_{h \rightarrow \text{inv}} = 10\%$ . Yields of these signal events and the sum of background events for each kinematic requirement in the cutflow are used to calculate a significance,

$$\mathcal{S} = \frac{S}{\sqrt{S + B}} ,$$

---

<sup>8</sup>These samples do not include beam effects such as beamstrahlung or ISR photons.

<sup>9</sup>The polarization scheme for ILC250 plans to also include  $100 \text{ fb}^{-1}$  of  $(P_{e-}, P_{e+}) = (\pm 0.8, \pm 0.3)$  as well, for a total of  $2 \text{ ab}^{-1}$ . [161] =

where  $S$  and  $B$  are signal and background yields, respectively. From this significance, a rough estimate of the upper limit can be calculated where

$$95\% \text{ UL}[\%] = \frac{\text{SM } \mathcal{B}_{H \rightarrow \text{inv.}}[\%] \times 1.65}{\mathcal{S}} .$$

### 7.6.1. Leptonic $Z$

Dominant backgrounds for the leptonic  $Z$  channel include 2-fermion processes ( $e^+e^- \rightarrow \mu^+\mu^-$ ) and 4-fermion processes ( $e^+e^- \rightarrow ZZ \rightarrow \nu\bar{\nu}\ell\ell$  and  $e^+e^- \rightarrow WW \rightarrow \ell\nu\ell\nu$ ). All 2-, 3-, and 4- fermion background processes are considered, as well as inclusive SM Higgs decays. Contributing SM  $H$  processes are dominated by  $H \rightarrow b\bar{b}$  and  $H \rightarrow WW^*$  with additional contributions from  $H \rightarrow ZZ^*$ ,  $H \rightarrow c\bar{c}$ , and  $H \rightarrow \tau\tau$ . Samples are generated with the SiD fast simulation in Delphes with lepton isolation<sup>10</sup>.

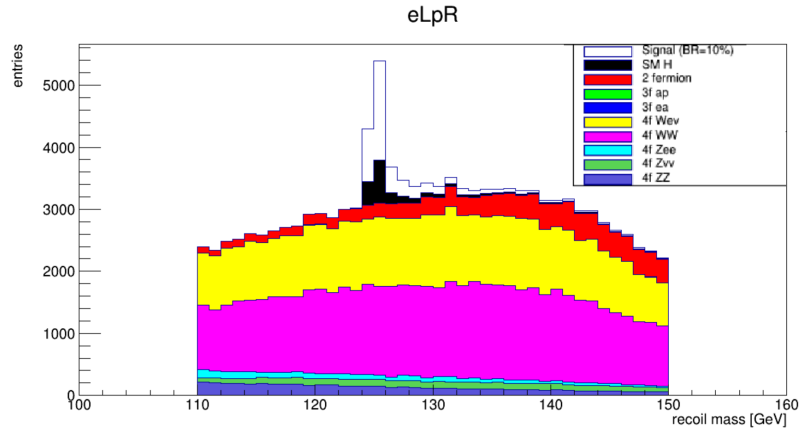
Lepton requirements are made to isolate and identify the visible  $Z$  decays. This requires two same-flavor-opposite-sign (SFOS) leptons with an invariant mass within 20 GeV of the  $Z$  mass and visible  $p_T$  between 20 – 70 GeV. The recoil mass must fall within a Higgs mass window,  $110 < m_{\text{recoil}} < 150$  GeV. A meager  $E_T^{\text{miss}}$  requirement of  $E_T^{\text{miss}} > 15$  GeV is also required to reduce the SM Higgs background (see Table 7.1). Beam polarization greatly impacts the prevalence of different background sources (see Figure 7.11). Signal efficiency ( $S/B$ ) for the combined polarizations is 0.05.

---

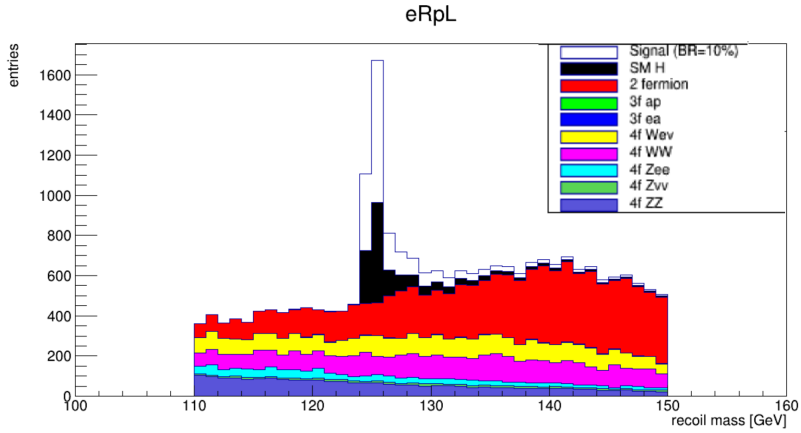
<sup>10</sup>Lepton isolation requirements are made for particle ID at the simulation stage, rather the analysis stage. Here, all leptons pass lepton isolation criteria - 12% (15%) relative  $E_T$  ( $p_T$ ) contained in a cone of radius  $R = 0.5$  around the reconstructed electron (muon).

| Cut                    | S (LR) | B (LR) | $\mathcal{S}$ | S (RL) | B (RL) | $\mathcal{S}$ |
|------------------------|--------|--------|---------------|--------|--------|---------------|
| All events             | 9.92e4 | 1.72e8 | 7.56          | 4.18e4 | 8.59e7 | 4.51          |
| MET > 15 GeV           | 7.12e4 | 3.75e7 | 11.6          | 3.15e4 | 8.42e6 | 10.8          |
| 2 leptons              | 5.13e3 | 1.68e6 | 3.95          | 2.28e3 | 2.70e5 | 4.37          |
| SFOS leptons > 10 GeV  | 4.87e3 | 1.41e6 | 4.09          | 2.16e3 | 1.90e5 | 4.94          |
| 75 <M_vis< 105 GeV     | 4.62e3 | 3.02e5 | 8.32          | 2.05e3 | 6.49e4 | 7.93          |
| 20 <pt_vis< 70 GeV     | 4.47e3 | 2.23e5 | 9.37          | 1.99e3 | 4.59e4 | 9.09          |
| 110 <m_recoil< 150 GeV | 4.41e3 | 1.17e5 | 12.6          | 1.96e3 | 2.17e4 | 12.7          |

TABLE 7.1. Cutflow for  $h \rightarrow \text{inv.}$  search with leptonic  $Z$  decays



(a)



(b)

FIGURE 7.11. Recoil mass distributions of events passing all kinematic criteria of Table 7.1

### 7.6.2. Hadronic $Z$

Dominant backgrounds for the leptonic  $Z$  channel include 4-fermion ( $e^+e^- \rightarrow ZZ \rightarrow \nu\bar{\nu}q\bar{q}$  and  $e^+e^- \rightarrow WW \rightarrow \ell\nu q\bar{q}$ ) and 2-fermion processes ( $e^+e^- \rightarrow q\bar{q}$ ). 3-fermion backgrounds play a larger role in this channel, accounting for roughly 30% of all background events. All 2-, 3-, and 4- fermion background processes are considered. Samples are generated with the SiD full simulation in Delphes with lepton isolation.

Jet requirements are made to isolate and identify the visible  $Z$  decays. This requires that the event contain no leptons and exactly two jets with an invariant mass within 20 GeV of the  $Z$  mass and visible  $p_T$  between 20 – 70 GeV. The recoil mass must fall within a Higgs mass window,  $110 < m_{\text{recoil}} < 150$  GeV. When possible, identical selections as the leptonic channel are chosen (see Table 7.2). The lepton veto leaves the two channels orthogonal. Signal efficiency ( $S/B$ ) for the combined polarizations is 0.04.

| Cut                                   | S (LR) | B (LR) | $\mathcal{S}$ | S (RL) | B (RL) | $\mathcal{S}$ |
|---------------------------------------|--------|--------|---------------|--------|--------|---------------|
| $20 < p_{T,\text{vis}} < 70$ GeV      | 1.25e4 | 7.71e6 | 4.48          | 8.84e3 | 1.07e6 | 8.53          |
| $75 < M_{\text{vis}} < 105$ GeV       | 1.16e4 | 1.79e6 | 8.63          | 8.21e3 | 3.14e5 | 14.5          |
| $N_{\text{jet}} == 2$                 | 1.16e4 | 1.79e6 | 8.63          | 8.21e3 | 3.14e5 | 14.5          |
| $-0.9 \leq \cos\theta_{jj} \leq -0.2$ | 1.08e4 | 8.68e5 | 11.5          | 7.65e3 | 1.78e5 | 17.7          |
| $110 < m_{\text{recoil}} < 150$ GeV   | 1.03e4 | 3.6e5  | 17            | 7.33e3 | 8.39e4 | 24.2          |

TABLE 7.2. Cutflow for  $h \rightarrow \text{inv.}$  search with hadronic  $Z$  decays

### 7.6.3. SiD Combination

Since the leptonic and hadronic channels are orthogonal, they can be simply considered in combination (see Table 7.3) [162]. Signal efficiency ( $S/B$ ) for the combined channels and polarizations is 0.16. This is an order of magnitude higher

than the  $S/B$  of 0.05 found in the ATLAS VBF+MET search. The highest significance, found by combining  $Z$  decay channels and polarization schemes, leads to an estimated  $\mathcal{B}_{h \rightarrow \text{inv}} \lesssim 0.3\%$ . This is comparable to  $\mathcal{B}_{h \rightarrow \text{inv}} < 0.0036$  at 95% C.L.<sup>11</sup>, estimated with a  $2 \text{ ab}^{-1}$  data set modeled with the expanded polarizaiton scheme [161]. This is two orders of magnitude stricter than the current LHC best limit  $\mathcal{B}_{h \rightarrow \text{inv}} < 0.11$  (0.11) (see Section VI) and one order of magnitude stricter than the anticipated HL-LHC performance of  $\mathcal{B}_{h \rightarrow \text{inv}} \lesssim 2.5\%$ .

|                 |          | S Yield | B Yield | $\mathcal{S}$ |
|-----------------|----------|---------|---------|---------------|
| $Z(\text{had})$ | eLpR     | 1.03e4  | 3.6e5   | 17            |
|                 | eRpL     | 7.33e3  | 8.39e4  | 24.2          |
| $Z(\text{lep})$ | eLpR     | 4.41e3  | 1.17e5  | 12.6          |
|                 | eRpL     | 1.96e3  | 2.17e4  | 12.7          |
| Combined        | eLpR     | 1.47e4  | 4.77e5  | 21.0          |
|                 | eRpL     | 9.29e3  | 1.06e5  | 27.4          |
|                 | Combined | 2.40e4  | 1.54e5  | 56.9          |

TABLE 7.3. Combined final yields of Tables 7.1 and 7.2 for  $1800 \text{ fb}^{-1}$  of LR/RL polarized beams

---

<sup>11</sup>Assuming SM Higgs total width.

## CHAPTER VIII

### SUMMARY

A non-technical summary of each chapter can be found below. I will introduce the scientific question I am interested in studying in Section 8.1. This is a search for a specific particle physics interaction. This particle physics pursuit requires an advanced experimental setup to collect data, explained in Section 8.2, and additional tools used to interpret the data, explained in Section 8.3. Once the data is interpreted in this way, we (particle physicists) can use it to draw physics conclusions. Section 8.4 overviews one attempt to measure the particle interaction of interest. This result is considered in a larger context in Section 8.5, where a second (and improved) measurement of the interaction is presented. Sections 8.4 and 8.5 use experimentally collected data, but possible new machines could improve the measurement of the interaction of interest even further. We can simulate their performance and the data that they would collect with computer models. Section 8.6 contains a third consideration of this physics process, but in the context of a future machine as considered with simulated data. The full work is summarized in Section 8.7.

#### **8.1. Standard Model Higgs Boson and Dark Matter**

After a century of research, scientists have created and refined the Standard Model (SM) of Particle Physics. This mathematical theory describes fundamental particles and the ways in which they interact with each other. The theory has proven very successful, as experimental testing continues to confirm its predictions.



However, some experimental observations cannot yet be explained by the SM - this means that the theory may be incomplete.

One such unexplained aspect is dark matter (DM) - a new particle not yet included in the SM whose existence is deduced from astronomical observation. Because little is known about the exact properties of this particle, many different theories exist. Particle physicists are responsible for performing experiments to prove or disprove these theories until the correct one is found.

One way SM can be expanded to include a new DM particle is by allowing the new DM particle to interact with the Higgs boson - the newest discovered particle in the SM. Higgs bosons are produced with high-energy particle colliders, so if a Higgs boson decayed to DM we can try to detect it. The signal it leaves in a detector would look identical to a process already allowed by the SM - Higgs decaying to invisible particles ( $h \rightarrow \text{inv.}$ ). Invisible particles are those that do not interact directly with the detector. By comparing how often the Higgs-to-invisible decay occurs experimentally to what is predicted from the theory, we can test the theory. If the measured frequency of decay (called the branching ratio, or  $\mathcal{B}_{h \rightarrow \text{inv.}}$ ) is larger than what the SM predicts, then a new process such as the Higgs decaying to DM could be contributing. This can then give us a hint to where new physics is hiding.

## 8.2. LHC and ATLAS

Collider experiments such as the type required for the above measurement are large, precise operations. The most powerful particle collider ever built is the Large Hadron Collider (LHC) near Geneva, Switzerland. This machine accelerates beams of protons to more than 99.9999% the speed of light as they travel around

27 km-long (16.8 mi) circular tunnels nearly 100 m (328 ft) underground. The beams are accelerated and directed with a series of magnets until they interact at collision points surrounded by detectors. Because there is no way to tune for one specific type of interaction, many different types of processes can occur.

One of these large detectors is ATLAS. The 7000 tonne (7716 ton) cylindrical detector is designed to detect as many types of particles as possible. Detection and identification requires distinct subsystems that are each designed to measure different specific characteristics of particles. When a particle passes through a subsystem, it deposits energy which is used as raw data. Individual particles can be identified by considering the position and intensity of energy deposits from all different subsystems. By creating a collection of particles measured by the detector, it is possible to reverse-engineer the physics process that likely occurred when the beams interacted.

Forty million collisions occur every second, which creates huge amounts of data. To reduce the amount of data managed and stored, each collision is considered in real time to decide whether there are noteworthy characteristics of the event worth storing and studying. From this large data set, events from different physics processes can be isolated and studied by the 3000 scientists in the ATLAS collaboration.

### **8.3. Simulation and Reconstruction**

One goal of particle physics is to test how well the SM agrees with data. This is done by comparing the data collected from beam collisions to simulated collision events that obey the SM. The experimental data and simulated data are considered identically - energy deposits in detector subsystems are used to identify

particles. This process is called “reconstruction”. Each type of particle interacts with the detector in a specific and unique way.

Not all particles that are produced when the beams collide deposit energy in the detector, though. Some decay or evolve first, and the products of the decay interact with the detector. This is the way in which “jets” are formed. A jet is a spray of particles in the detector that originated from one fundamental quark or gluon.

Some particles do not interact with the detector at all - they only carry energy away from the interaction. Their presence can still be deduced by considering all the energy deposits around the detector and looking for an imbalance. The “invisible” particle must have carried away the energy that balances out all other deposits. The energy carried away is called the missing transverse energy (MET). This is the signature that DM would leave if it was produced in a collision.

#### 8.4. VBF $h \rightarrow$ invisible

There are a number of different ways that the colliding protons can produce a Higgs boson, and oftentimes there are additional particles that result from the Higgs production. The most sensitive measurement of  $\mathcal{B}_{h \rightarrow \text{inv}}$  can be made at the LHC by identifying interactions that create a Higgs boson through the Vector Boson Fusion (VBF) mechanism. This process produces one Higgs boson and two jets. These jets are distinctive - there are expectations about where they should fall within the detector, for example - and are used to identify the event. When the Higgs decays invisibly, it leaves the trace of MET. Therefore, this analysis

looks for events where there are two VBF-like jets and MET from the invisible Higgs decay.

These events are selected out of the large data set by defining a set of requirements characteristic of the VBF  $h \rightarrow \text{inv.}$  process that each event must satisfy. Events that satisfy all the requirements populate a “signal region” enriched with signal events. For example, to fall into the signal region an event must have at least two jets and a relatively large amount of MET. However, there are processes other than  $h \rightarrow \text{inv.}$  that have similar detector signatures. These events still pass the requirements but contaminate the signal region with “background” events. Simulations modeling these specific background processes are used to estimate the degree of contamination of each of these background processes.

Sources of uncertainty on the measurement of events are also considered. Some uncertainties come from physical sources, such as the level of precision with which the detector can measure energy deposits. Others come from the precision of particle identification and reconstruction algorithms. Still others are associated with the simulation of physics processes and estimations that must be made in order to reduce computation complexity and run time. Finally, statistics play a large role in measurements so uncertainties related the mathematics of probability are also included.

By comparing what is measured in the data to what is estimated from the simulation, we can learn about  $\mathcal{B}_{h \rightarrow \text{inv.}}$ . Because the SM predicts a very low rate for  $h \rightarrow \text{inv.}$  decays -  $\mathcal{B}_{h \rightarrow \text{inv.}} \sim 0.001$ , where one in every 1,000 Higgs bosons produced is expected to decay invisibly -  $\mathcal{B}_{h \rightarrow \text{inv.}}$  is difficult to measure precisely. Instead, an upper limit on this value can be set. Until the value can be measured precisely, there is a possibility that dark matter processes contribute to  $\mathcal{B}_{h \rightarrow \text{inv.}}$ .

For example, imagine that the Higgs boson decays to dark matter and this increases the rate of  $h \rightarrow \text{inv.}$  processes such that  $\mathcal{B}_{h \rightarrow \text{inv}} = 0.01$ . Setting an upper limit where  $\mathcal{B}_{h \rightarrow \text{inv}} < 0.26$  only asserts that the true rate must be less than 0.26. The measurement would not be sensitive enough to see the DM contribution because  $0.01 < 0.26$ . Now imagine another search producing a measured upper limit of  $\mathcal{B}_{h \rightarrow \text{inv}} < 0.003$ . This limit says that the rate must be less than 0.003, so the theory that predicted  $\mathcal{B}_{h \rightarrow \text{inv}} = 0.01$  can now confidently be ruled out because  $0.003 < 0.01$ . This is still a limit though, and different DM theories still may contribute. However, this limit is very close to the SM predicted value of  $\mathcal{B}_{h \rightarrow \text{inv}} \sim 0.001$ . That means that though DM may contribute, it can only contribute a little bit - it cannot increase the SM rate more than the measured limit. Therefore, the strictest limits are those that are closest to the SM-predicted value of  $\mathcal{B}_{h \rightarrow \text{inv}}$ . Only when a measurement is made, stating the experimentally observed rate of  $\mathcal{B}_{h \rightarrow \text{inv}}$  directly, can a conclusive statement be made about whether or not DM contributes. If  $\mathcal{B}_{h \rightarrow \text{inv}} = 0.001$  is measured, the SM prediction is confirmed and no DM contributes. If  $\mathcal{B}_{h \rightarrow \text{inv}}$  is any value other than this SM prediction, then something else must be contributing.

Before the works that I contributed to (documented in this dissertation), the strictest limit set by ATLAS was  $\mathcal{B}_{h \rightarrow \text{inv}} < 0.25$ . This left a lot of room for DM to be hiding. Using all the available LHC data, an updated search for VBF - produced  $h \rightarrow \text{inv.}$  measured  $\mathcal{B}_{h \rightarrow \text{inv}} < 0.13$ . This is the strictest limit set to date by a single ATLAS analysis on this value. However, it is still only an upper limit meaning that there is still a possibility that DM contributes to the process.

### 8.5. $h \rightarrow$ invisible **Combination**

The measurement of  $\mathcal{B}_{h \rightarrow \text{inv}}$  can be improved by considering more data. The VBF analysis above considers only one Higgs production mechanism, but there are many ways that Higgs bosons can be produced at the LHC. A more refined result was obtained by considering Higgs bosons created in multiple different ways but all decaying invisibly, and combining these results together. This combination included the VBF result above and one additional Higgs boson production mechanism - creation with a top quark and top anti-quark. These two searches looked for  $h \rightarrow \text{inv.}$  processes with all the ATLAS data available. The combination additionally considered a previous combination of searches for  $h \rightarrow \text{inv.}$  using older data from when the LHC collided beams at a lower energy. This old combination measured the previously stated best limit of  $\mathcal{B}_{h \rightarrow \text{inv}} < 0.25$ .

The use of more data events selected by these different sources refines the measurement of  $\mathcal{B}_{h \rightarrow \text{inv}}$ , although the VBF result is the most sensitive input and drives this combination. The resulting upper limit of  $\mathcal{B}_{h \rightarrow \text{inv}} < 0.11$  is the strictest limit set to date by any experiment on the  $h \rightarrow \text{inv.}$  process.

### 8.6. $h \rightarrow$ invisible **in Future Colliders**

The LHC is scheduled to continue operating into the 2030s. After this, new future colliders are being considered for construction by the particle physics community and funding agencies. One strong contender is the International Linear Collider (ILC) which would collide fundamental particles (electrons and their anti-matter partner positrons) in a linear tunnel. This is distinct from the LHC which collides composite particles (protons) in a circular tunnel. Colliding fundamental particles in the ILC allows for different physics processes to occur with much lower

background rates, making precision measurements easier than they are at the LHC.

The ILC would be used as a “Higgs Factory,” producing nearly half a million Higgs bosons in the first phase of running. This will allow for a thorough study of the particle. Precision studies are also made possible with new detectors utilizing modern technologies, which are still being designed. My work helped to refine the design of one proposed ILC detector to improve its physics potential.

With many Higgs bosons being produced at the ILC,  $h \rightarrow \text{inv.}$  can be studied there as well. The results in this environment, or expected measured limits set on  $\mathcal{B}_{h \rightarrow \text{inv.}}$ , are expected to be much better than what can be measured at the LHC. Preliminary studies with simulated data confirm this, as I estimate an upper limit of  $\mathcal{B}_{h \rightarrow \text{inv.}} < 0.003$ . This is nearly one hundred times more precise than the best current LHC measurements ( $\mathcal{B}_{h \rightarrow \text{inv.}} < 0.11$ ), and begins to approach the SM-predicted value of  $\mathcal{B}_{h \rightarrow \text{inv.}} \sim 0.001$ . This higher level of precision reduces the space in which DM could be hiding.

## 8.7. Conclusion

One major goal of particle physics today is understanding the shortcomings of the Standard Model. One way that particle physicists attempt to do this is through precision measurements of Standard Model physics processes to see whether the theory makes correct predictions. If a measurement disagrees with the Standard Model prediction, it can give a hint to how the theory must be updated to correctly model Nature.

One open question that the Standard Model does not currently address is the presence and identity of dark matter. One possible way that dark matter could

interact with the known Standard Model particles is through the Higgs boson. In order to search for this process in particle colliders, physicists (myself included) search for events in the data where Higgs bosons are produced, and then decay to dark matter. Dark matter would not directly interact with the detector, so this “invisible” particle would carry energy away from the interaction.

I outlined three separate searches for the  $h \rightarrow \text{invisible}$  process. Two of them use data collected from the LHC with the ATLAS detector. These results are the best experimentally-observed limits that have been set on the  $h \rightarrow \text{inv.}$  process, measuring  $\mathcal{B}_{h \rightarrow \text{inv}} < 0.11$  when the Standard Model predicts  $\mathcal{B}_{h \rightarrow \text{inv}} \sim 0.001$ . The third search considers a potential future collider, the ILC, and uses computer-simulated data to model how well this new machine may measure  $\mathcal{B}_{h \rightarrow \text{inv}}$ . This study predicts that the ILC could measure  $\mathcal{B}_{h \rightarrow \text{inv}} < 0.003$ .

The  $h \rightarrow \text{invisible}$  process is an interesting channel in which to consider how dark matter may interact with Standard Model particles. Much work has been done to search for these events, and limits on how large the observed rate can be continue to decrease and approach the Standard Model value. For now though, only upper limits have been set and a direct measurement has not been made. This means that dark matter still may contribute to this process. With continued study, a measurement will be made that can definitively confirm the Standard Model predicted value or coax dark matter out of its hiding place.



## CHAPTER IX

## CONCLUSION

The Standard Model has stood up to rigorous experimental testing, yet there are known deficiencies such as the lack of a dark matter candidate. Therefore, the theory is an incomplete description of Nature. An extension through the Higgs portal allows new dark matter candidate particles to couple directly with the Higgs boson, resulting in a detector signature of invisible particles. This process is predicted through the Standard Model with  $\mathcal{B}_{h \rightarrow \text{inv}} \sim 0.001$ , but the addition of dark matter could inflate  $\mathcal{B}_{h \rightarrow \text{inv}}$  by up to 10%. Searches for this  $h \rightarrow \text{invisible}$  signature are a promising avenue for considering new physics possibilities in the electroweak sector.

This dissertation presented three searches for decays of Higgs bosons to invisible particles. First, a search Higgs bosons created through vector boson fusion was conducted using the full Run 2 data set of  $139 \text{ fb}^{-1}$  of  $\sqrt{s} = 13 \text{ TeV}$  proton–proton collision data collected by the ATLAS detector at the LHC. This search refined the measured upper limit  $\mathcal{B}_{h \rightarrow \text{inv}} < 0.13$  (0.13) obs (exp). This search was used as input to a statistical combination which also included results from the full Run 2 consideration of  $ttH$  production and the Run 1  $h \rightarrow \text{inv.}$  combination effort. Driven by the VBF channel, this combination found  $\mathcal{B}_{h \rightarrow \text{inv}} < 0.11$  (0.11) obs (exp). This is the strictest measured limit to date.

This measurement can also be considered in the cleaner environment of a future lepton collider, such as the International Linear Collider. With  $1.8 \text{ ab}^{-1}$  of simulated  $\sqrt{s} = 250 \text{ GeV}$  data, the proposed SiD detector is expected to measure  $\mathcal{B}_{h \rightarrow \text{inv}} \lesssim 0.003$ . Approaching SM predicted rate is key in probing its accuracy.

## APPENDIX A

### REFERENCE LIST OF ABBREVIATIONS

|                                                   |                                                    |
|---------------------------------------------------|----------------------------------------------------|
| 0L = zero lepton $t\bar{t}$ H+MET analysis        | IBL = insertable B-Layer (part of tracking system) |
| 2L = two lepton $t\bar{t}$ H+MET analysis         | ID = inner detector                                |
| BSM = Beyond the Standard Model                   | ILC = International Linear Collider                |
| C.L. = confidence level                           | IP = interaction point (of beam collision)         |
| CMS = Compact Muon Solenoid (experiment)          | ISR = initial state radiation                      |
| CR = control region                               | JER = jet energy resolution                        |
| CSC = cathode strip chamber (part of muon system) | JES = jet energy scale                             |
| DM = dark matter                                  | JVT = jet vertex tagging                           |
| ECal = electromagnetic calorimeter                | L1 = level one trigger                             |
| EM = electromagnetic                              | LAr = liquid argon electromagnetic calorimeter     |
| EWK = electroweak                                 | LHC = Large Hadron Collider                        |
| FCal = forward calorimeter                        | MC = Monte Carlo (simulation)                      |
| fJVT = forward jet vertex tagging                 | MDT = monitored drift tube (part of muon system)   |
| FSR = final state radiation                       | ME = matrix element                                |
| ggF = gluon-gluon fusion (Higgs production mode)  | MET = missing transverse energy                    |
| HCal = hadronic calorimeter                       | MIP = minimum ionizing particle                    |
| HEC = hadronic endcap calorimeter                 | MJ = QCD multijet background                       |
| HLT = high level trigger                          | MS = muon spectrometer                             |
| HS = hard scatter                                 |                                                    |

|                                                                      |                                                                                              |
|----------------------------------------------------------------------|----------------------------------------------------------------------------------------------|
| NLL = negative log likelihood                                        | SPS = Super Proton Synchrotron (part of LHC booster system)                                  |
| NP = nuisance parameter (in a fit)                                   | SR = signal region                                                                           |
| PDF = parton distribution function                                   | TDAQ = trigger and data acquisition                                                          |
| PFlow = particle flow (jet reconstruction technique)                 | TGC = thin gap chamber (part of muon system)                                                 |
| POI = parameter of interest                                          | TRT = transition radiation tracker                                                           |
| PS = Proton Synchrotron (part of LHC booster system) ; parton shower | TileCal = scintillating tile hadron calorimeter                                              |
| PSB = Proton Synchrotron Booster                                     | UL = upper limit                                                                             |
| PTV= $p_T$ of vector boson                                           | VBF = vector boson fusion, also called Vector Boson Scattering (VBS) (Higgs production mode) |
| PU = pileup                                                          | VBF+MET = search for invisibly decaying Higgs bosons produced by vector boson fusion         |
| PV = primary vertex                                                  | WCR = VBF+MET control region enriched with $W$ +jets background by requiring one lepton      |
| QCD = quantum chromodynamics                                         | WP = working point                                                                           |
| R&S = Rebalance and Smear                                            | ZCR = VBF+MET control region enriched with $Z$ +jets background by requiring two leptons     |
| VBF+MET multijet background estimation method                        |                                                                                              |
| RF = radio frequency                                                 |                                                                                              |
| RoI = region of interest                                             |                                                                                              |
| RPC = resistive plate chamber (part of muon system)                  |                                                                                              |
| SCT = semiconductor tracker                                          |                                                                                              |
| SFOS = same flavor, opposite sign (leptons)                          |                                                                                              |
| SM = Standard Model                                                  |                                                                                              |

## APPENDIX B

### ELECTROMAGNETIC CALORIMETRY

Electromagnetic showers are caused by cascades induced by bremsstrahlung (photon emission) and pair production ( $\gamma \rightarrow e^+e^-$ ) interactions with detector components. For low mass electrons, losses from bremsstrahlung dominate. The shape of electromagnetic showers depends in part on material properties of the calorimeter, measured as the radiation length,

$$X_0(\text{g/cm}^2) \approx \frac{716 \text{ g cm}^{-2} A}{Z(Z+1)\ln(287/\sqrt{Z})} ,$$

where  $Z$  and  $A$  are the atomic number (number of protons) and weight (number of protons+neutrons) of the material [75]. The radiation length is related to the average energy lost by an incident electron in some distance  $x$  by

$$\frac{dE}{dx} = -\frac{E}{X_0} ,$$

with total energy loss

$$E = E_0 e^{-x/X_0} .$$

Pair production occurs by the scaled  $\frac{9}{7}X_0$  [163] such that total energy loss is

$$E = E_0 e^{-7x/9X_0} .$$

The critical energy,

$$\epsilon = \frac{610 \text{ MeV}}{Z + 1.24} ,$$

occurs when ionization and brehmsstrahlung losses equal. This similar scale for electrons and photons means that the longitudinal behavior of their showers can be treated identically with the scale of radiation length. The longitudinal profile grows like

$$\frac{dE}{dt} = E_0 b \frac{(bt)^{a-1} e^{-bt}}{\Gamma(a)} ,$$

where  $t$  is the depth into the calorimeter ( $t = x/X_0$ ),  $E_0$  is the incident particle energy,  $a$  and  $b$  are fit parameters, and  $\Gamma(a) = (a-1)!$ . The shower develops a maximum number of secondary produced particles at a depth of

$$t_{\max} \sim \ln \frac{E_0}{\epsilon} + t_0 X_0 ,$$

where  $t_0 = (-)0.5$  for photons(electrons). After this point, produced particle energies fall below the critical energy and further multiplication dwindles.

The transverse extent of an electromagnetic shower is described by the Molière radius,

$$R_M \sim 21 \text{ MeV} \frac{X_0}{\epsilon(\text{MeV})} ,$$

which shows the average deflection of particles with energy  $\epsilon$  after one radiation length. 90% of an average shower is contained within one Molière radius, which is typically the order of a few centimeters.

Calorimeter performance can be measured by energy resolution, which is intrinsic to the detector design, material, and response. The sum of all ionization tracks from all particles in the shower (total track length) is proportional to the number of particles scaled by the detector radiation length, but shower development is stochastic and energy-dependent. An ideal calorimeter of infinite length and perfect symmetry and read out could have an intrinsic energy

resolution proportional to the square root of the total track length, but a realistic calorimeter has an energy resolution of

$$\frac{\sigma}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c ,$$

where  $\oplus$  is a quadratic sum and  $a$ ,  $b$ , and  $c$  are constants. The term scaled by  $a$  accounts for the stochastic effects previously mentioned. The term scaled by  $b$  accounts for noise in an experimental setting from electronics related to the readout chain, for example. The constant term  $c$  is independent of energy and could include instrumental effects such as detector geometry. Depending on the purpose of the detector and the typical energy of incident particles, different terms may dominate.

Calorimeters need to capture as much energy from the shower as possible. If a shower cannot be contained within the scope of the calorimeter, the shower is “leaked” to different detector components. This can degrade the energy resolution (see Appendix H).

## APPENDIX C

### TRIGGER SYSTEM EVOLUTION FROM RUN 1 TO RUN 2

The ATLAS Run 2 trigger system is considered in Section 3.2.7. Before this for Run 1, the trigger system was set up differently (see Figure C.1). Changes had to be made for Run 2 to account for the changed conditions (increased collision energy and average pileup, for example).

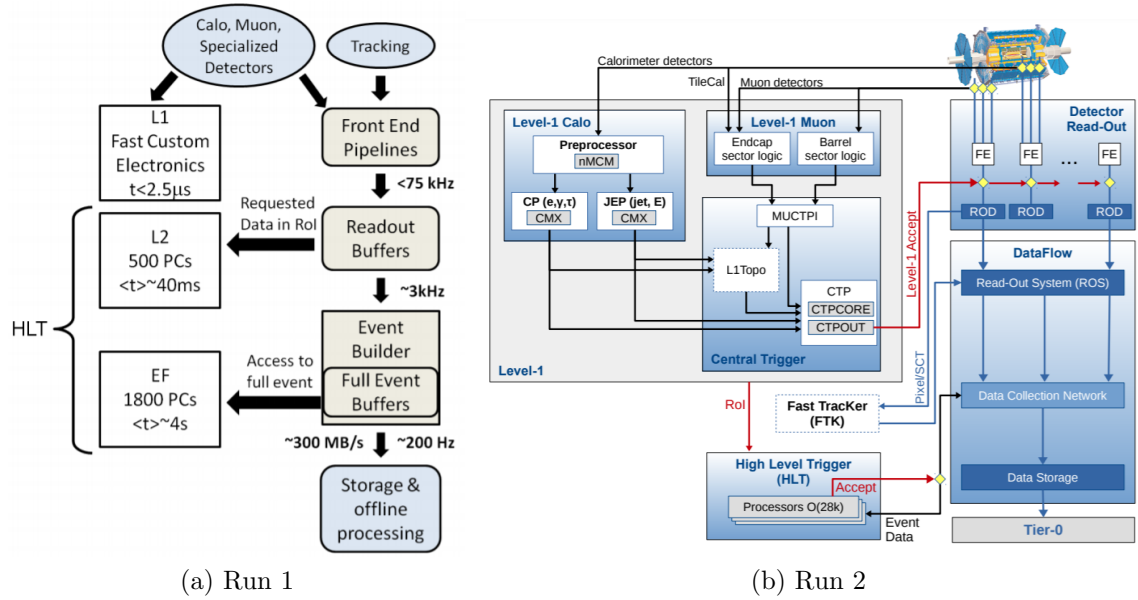


FIGURE C.1. Trigger systems for a) Run 1 [33] and b) Run 2 [15]

Upgraded L1 hardware and readout systems in Run 2 increased the allowed L1 rate from 70 kHz to 100 kHz [15]. Additional changes come from software developments. For example, a new topological processor, L1Topo, was added for Run 2. Along with upgrades of the central trigger processor, this allows for triggering on multiple objects with kinematic selections such as dijet pairs with a minimum invariant mass. Input from IBL was also included in the Run 2 trigger software.

The most significant changes are in the layout of the HLT. The HLT was separated into separate Level-2 and Event Filter (EF) processes in Run 1, which were combined into the unified Run 2 HLT. This reduced code duplication and increased flexibility. Algorithm optimization increased the performance of some triggers, more closely replicating offline reconstruction, and allowed for the output rate to increase from 400 Hz to 1 kHz.



## APPENDIX D

### FORWARD JET VERTEX TAGGING

Forward pileup jets may result from one QCD process (QCD pileup jets) or capture particles from multiple different interactions (stochastic jets). The different sources of pileup are treated differently.

Since stochastic jets are the result of multiple particles (with contributions of out-of-time pileup), the constituent calorimeter cells will be activated at different times. A jet timing value can be used to discriminate, as the average jet timing of hard scatter and QCD pileup jets are more centered around 0 with shorter tails than stochastic jets. Requiring that  $|t_{\text{jet}}| < 12$  ns rejects 20% of stochastic jets with only 1% loss of hard scatter jets [94].

Multiple contributing particles also means that the isotropic energy distribution of the jet cone differs from that of the hard scatter QCD pileup jet. Here, one initial parton that creates a dense energy core. This lacks in stochastic jet signatures.

The  $p_T$  distribution of a jet in the  $\Delta\eta/\Delta\phi$  plane is described as

$$f = \alpha + \beta\Delta\eta + \gamma\exp\left[-\frac{1}{2}\left(\frac{\Delta\eta}{0.1}\right)^2 - \frac{1}{2}\left(\frac{\Delta\phi}{0.1}\right)^2\right], \quad (\text{D.1})$$

which considers forward calorimeter towers. The constant and linear terms describe the stochastic contribution to the jet whereas the exponential addresses the underlying QCD jet. A QCD pileup jet would be expected to only have this Gaussian distribution, so the scaling  $\gamma$  parameter represents the contribution of a stochastic component. In a pileup-independent way,  $\gamma$  provides information about

the  $p_T$  contained within the core of a jet and can therefore discriminate between stochastic and QCD pileup jets. This is because stochastic jets with little structure and no core would have a low  $\gamma$  value close to zero.

QCD pileup jets can be considered separately, as they have the same shape and  $\gamma$  behavior as hard scatter jets. Every vertex is assumed to conserve energy in the transverse plane, so forward QCD jets can be identified if they balance the transverse energy left from other QCD jets in the central tracker region. This requires the distinction between stochastic and QCD pileup jets in the central region as well, which is done by identifying jets where  $R_{pT}^i > R_{pT}^0$  where  $i$  is the pileup vertex of origin. Stochastic jets, with contributions from multiple vertices, would have small  $R_{pT}$  values associated with any vertex. Central QCD pileup jets are then isolated at each vertex and balanced, if possible, with forward jets likely to have originated from that vertex.

These two elements inform a unified forward JVT (fJVT) discriminant for each jet from vertex  $i$ ,

$$\text{fJVT} = \frac{\langle \vec{p}_{T,i}^{\text{miss}} \rangle \cdot \vec{u}^{fj}}{\gamma}, \quad (\text{D.2})$$

where  $\langle \vec{p}_T^{\text{miss}} \rangle$  is the average missing transverse momentum from central jets associated with vertex  $i$  and  $\vec{u}^{fj}$  is a unit vector in the direction of the forward jet. In this way, the shape information contained in  $\gamma$  is incorporated into the vertex-by-vertex momentum balancing of QCD jets. With 85% efficiency for retaining hard scatter jets, using fJVT can identify 51% of forward pileup jets with  $\langle \mu \rangle = 35$ .

## APPENDIX E

### VBF+MET OVERLAP REMOVAL

Physics objects are reconstructed from energy deposits in the detector independently. To avoid the same energy cluster being used to reconstruct two different objects, overlap removal is performed. The position of electrons, muons, and jets are all compared and if reconstructed objects fall too close to each other then one is removed so the energy deposit is only used once (see Table E.1). The position difference is defined with rapidity,  $\Delta R = \sqrt{\Delta y^2 + \Delta \phi^2}$ . Lepton removal near jets involves a scaling by the lepton  $p_T$  to account for collimation of boosted objects. All jets are reconstructed with an anti-kt radius parameter of  $R = 0.4$ .

| Remove   | Keep     | Matching criteria                                              |
|----------|----------|----------------------------------------------------------------|
| electron | electron | shared inner detector track, electron with lower $p_T$ removed |
| muon     | electron | muon with calorimeter deposits and shared inner track          |
| electron | muon     | shared inner detector track                                    |
| photon   | electron | $\Delta R < 0.4$                                               |
| photon   | muon     | $\Delta R < 0.4$                                               |
| jet      | electron | $\Delta R < 0.2$                                               |
| electron | jet      | $\Delta R < \min(0.4, 0.04 + 10\text{GeV}/p_T^e)$              |
| jet      | muon     | number of tracks $< 3$ and $\Delta R < 0.2$                    |
| muon     | jet      | $\Delta R < \min(0.4, 0.04 + 10\text{GeV}/p_T^\mu)$            |
| photon   | jet      | $\Delta R < 0.4$                                               |

TABLE E.1. Overlap removal between objects for VBF+MET searches [25]

Overlap removal for the VBF+MET HL-LHC projection studies is slightly simplified (see Table E.2).

| Remove   | Keep     | Matching criteria |
|----------|----------|-------------------|
| jet      | electron | $\Delta R < 0.2$  |
| electron | jet      | $\Delta R < 0.4$  |
| jet      | muon     | $\Delta R < 0.2$  |
| muon     | jet      | $\Delta R < 0.4$  |
| jet      | neutrino | $\Delta R < 0.4$  |

TABLE E.2. Overlap removal of truth smeared objects for VBF+MET HL-LHC projection studies

## APPENDIX F

### VBF+MET YIELDS

Post-fit yields for the VBF+MET first look search are shown in Table F.1 and Table F.2 (separated by bins).

| Process    | SR              | $Z_{\ell\ell}$ | $W_{e\nu}$     | $W_{\mu\nu}$    | $W_{\ell\nu}$   | Fake- $e$ CR   |
|------------|-----------------|----------------|----------------|-----------------|-----------------|----------------|
| $Z$ strong | $6810 \pm 430$  | $1394 \pm 81$  | $48 \pm 11$    | $193 \pm 21$    | $241 \pm 23$    | $153 \pm 18$   |
| $Z$ EWK    | $2660 \pm 320$  | $634 \pm 75$   | $12 \pm 1$     | $41 \pm 2$      | $53 \pm 2$      | $26 \pm 2$     |
| $W$ strong | $3750 \pm 270$  | -              | $3530 \pm 230$ | $6730 \pm 390$  | $10260 \pm 610$ | $1760 \pm 140$ |
| $W$ EWK    | $1380 \pm 130$  | -              | $2140 \pm 210$ | $3770 \pm 370$  | $5910 \pm 570$  | $1120 \pm 120$ |
| Fake- $e$  | -               | -              | $239 \pm 62$   | -               | $239 \pm 62$    | $1190 \pm 180$ |
| Multijet   | $740 \pm 280$   | -              | -              | -               | -               | -              |
| Other      | $155 \pm 27$    | $37 \pm 27$    | $322 \pm 50$   | $395 \pm 60$    | $720 \pm 110$   | $57 \pm 7$     |
| Tot. bg.   | $15490 \pm 130$ | $2065 \pm 44$  | $6288 \pm 75$  | $11130 \pm 110$ | $17420 \pm 150$ | $4300 \pm 66$  |
| $H$ (VBF)  | $647 \pm 52$    | -              | -              | -               | -               | -              |
| $H$ (ggF)  | $90 \pm 43$     | -              | -              | -               | -               | -              |
| $H$ (VH)   | $0.81 \pm 0.14$ | -              | -              | -               | -               | -              |
| Data       | 15511           | 2050           | 6323           | 11095           | 17418           | 4293           |

TABLE F.1. Observed and expected background event yields with associated uncertainties in the SR and CRs after the likelihood fit. A dash (-) indicates that the corresponding yield is  $< 0.5$  events.

| Process                                   | $N_{\text{jet}} = 2,  \Delta\phi_{jj}  < 1, m_{jj}$ bins     |                      |                      |                      |                    |
|-------------------------------------------|--------------------------------------------------------------|----------------------|----------------------|----------------------|--------------------|
|                                           | 0.8–1.0 TeV<br>Bin 1                                         | 1.0–1.5 TeV<br>Bin 2 | 1.5–2.0 TeV<br>Bin 3 | 2.0–3.5 TeV<br>Bin 4 | >3.5 TeV<br>Bin 5  |
| $Z$ strong                                | $1\,089 \pm 82$                                              | $1\,220 \pm 110$     | $355 \pm 58$         | $188 \pm 45$         | $8 \pm 4$          |
| $Z$ EWK                                   | $165 \pm 45$                                                 | $339 \pm 82$         | $199 \pm 46$         | $209 \pm 36$         | $27 \pm 8$         |
| $W$ strong                                | $582 \pm 62$                                                 | $770 \pm 150$        | $170 \pm 35$         | $82 \pm 22$          | $5 \pm 3$          |
| $W$ EWK                                   | $68 \pm 18$                                                  | $148 \pm 34$         | $104 \pm 21$         | $121 \pm 18$         | $32 \pm 7$         |
| Multijet                                  | $120 \pm 48$                                                 | $140 \pm 57$         | $44 \pm 18$          | $30 \pm 13$          | $3 \pm 2$          |
| Other                                     | $15 \pm 3$                                                   | $24 \pm 4$           | $13 \pm 3$           | $12 \pm 3$           | $3 \pm 1$          |
| Total Bkg                                 | $2\,040 \pm 44$                                              | $2\,647 \pm 53$      | $884 \pm 28$         | $641 \pm 25$         | $79 \pm 8$         |
| $H$ ( $\mathcal{B}_{\text{inv}} = 0.13$ ) | $64 \pm 10$                                                  | $135 \pm 15$         | $87 \pm 9$           | $89 \pm 11$          | $16 \pm 2$         |
| Data                                      | 2 065                                                        | 2 639                | 890                  | 633                  | 76                 |
| Data/Bkg                                  | $1.01 \pm 0.02$                                              | $1.00 \pm 0.02$      | $1.01 \pm 0.03$      | $0.99 \pm 0.04$      | $0.97 \pm 0.10$    |
| $\beta_W$                                 | $1.07 \pm 0.26$                                              | $1.03 \pm 0.20$      | $1.02 \pm 0.19$      | $1.05 \pm 0.18$      | $1.16 \pm 0.29$    |
| $\beta_Z$                                 | $1.20 \pm 0.30$                                              | $1.11 \pm 0.24$      | $1.01 \pm 0.22$      | $1.11 \pm 0.21$      | $0.80 \pm 0.27$    |
|                                           | $N_{\text{jet}} = 2, 1 <  \Delta\phi_{jj}  < 2, m_{jj}$ bins |                      |                      |                      |                    |
|                                           | 0.8–1.0 TeV<br>Bin 6                                         | 1.0–1.5 TeV<br>Bin 7 | 1.5–2.0 TeV<br>Bin 8 | 2.0–3.5 TeV<br>Bin 9 | >3.5 TeV<br>Bin 10 |
| $Z$ strong                                | $727 \pm 61$                                                 | $1\,250 \pm 130$     | $424 \pm 68$         | $193 \pm 51$         | $21 \pm 8$         |
| $Z$ EWK                                   | $116 \pm 30$                                                 | $430 \pm 100$        | $266 \pm 61$         | $332 \pm 49$         | $55 \pm 11$        |
| $W$ strong                                | $386 \pm 50$                                                 | $688 \pm 95$         | $234 \pm 45$         | $108 \pm 42$         | $5 \pm 2$          |
| $W$ EWK                                   | $55 \pm 15$                                                  | $216 \pm 45$         | $146 \pm 28$         | $177 \pm 22$         | $43 \pm 8$         |
| Multijet                                  | $71 \pm 28$                                                  | $132 \pm 53$         | $37 \pm 15$          | $24 \pm 10$          | -                  |
| Other                                     | $10 \pm 2$                                                   | $19 \pm 4$           | $8 \pm 2$            | $8 \pm 3$            | $5 \pm 2$          |
| Total Bkg                                 | $1\,365 \pm 36$                                              | $2\,728 \pm 52$      | $1\,115 \pm 31$      | $842 \pm 28$         | $129 \pm 10$       |
| $H$ ( $\mathcal{B}_{\text{inv}} = 0.13$ ) | $22 \pm 3$                                                   | $76 \pm 7$           | $51 \pm 4$           | $56 \pm 6$           | $11 \pm 2$         |
| Data                                      | 1 362                                                        | 2 730                | 1 132                | 836                  | 133                |
| Data/Bkg                                  | $1.00 \pm 0.02$                                              | $1.00 \pm 0.02$      | $1.02 \pm 0.03$      | $0.99 \pm 0.03$      | $1.03 \pm 0.08$    |
| $\beta_W$                                 | $1.02 \pm 0.24$                                              | $1.08 \pm 0.20$      | $1.01 \pm 0.19$      | $0.96 \pm 0.16$      | $0.98 \pm 0.24$    |
| $\beta_Z$                                 | $1.19 \pm 0.28$                                              | $1.08 \pm 0.24$      | $0.92 \pm 0.20$      | $0.96 \pm 0.17$      | $1.36 \pm 0.34$    |
|                                           | $2 < N_{\text{jet}} < 5$<br>$m_{jj} > 3.5$ TeV<br>Bin 11     |                      |                      |                      |                    |
|                                           |                                                              |                      |                      |                      |                    |
| $Z$ strong                                | $1\,330 \pm 200$                                             |                      |                      |                      |                    |
| $Z$ EWK                                   | $530 \pm 160$                                                |                      |                      |                      |                    |
| $W$ strong                                | $720 \pm 120$                                                |                      |                      |                      |                    |
| $W$ EWK                                   | $267 \pm 66$                                                 |                      |                      |                      |                    |
| Multijet                                  | $139 \pm 57$                                                 |                      |                      |                      |                    |
| Other                                     | $38 \pm 8$                                                   |                      |                      |                      |                    |
| Total Bkg                                 | $3\,019 \pm 56$                                              |                      |                      |                      |                    |
| $H$ ( $\mathcal{B}_{\text{inv}} = 0.13$ ) | $132 \pm 24$                                                 |                      |                      |                      |                    |
| Data                                      | 3 015                                                        |                      |                      |                      |                    |
| Data/Bkg                                  | $1.00 \pm 0.02$                                              |                      |                      |                      |                    |
| $\beta_W$                                 | $0.94 \pm 0.22$                                              |                      |                      |                      |                    |
| $\beta_Z$                                 | $1.04 \pm 0.28$                                              |                      |                      |                      |                    |

TABLE F.2. Yields of data, signal and major backgrounds in the SR for each bin after the likelihood fit. Minor backgrounds from  $t\bar{t}$ ,  $VV$ ,  $VVV$ , and VBF  $H \rightarrow W^+W^- / \tau^+\tau^-$  are summed up as “Other”. A dash (-) indicates that the corresponding yield is less than 0.5 events.

## APPENDIX G

### HL-LHC SMEARING OF TRUTH-LEVEL OBJECTS FOR VBF $h \rightarrow$ invisible

#### G.1. Electrons

Truth electrons are pulled from truth particle containers in the TRUTH3 xAOD. Each electron is individually considered for smearing, which is done with resolution functions based on the truth electron's energy and  $\eta$  position. Dressed values are used for these quantities rather than the truth values, as the dressed value more closely represents what is measured in the detector. The `looseElectron` working point is used.

The smeared energy is pulled from a Gaussian distribution around the origin with a width defined by the energy resolution of the detector in a given  $E_T$  and  $\eta$  slice. The smear amount is added to the true dressed energy and this smeared energy value is returned. The smeared  $p_T$  is calculated, as the ratio of  $p_T/E$  should be the same for the dressed and smeared electron. All other electron values ( $\eta/\phi$  position, for example) are unchanged from the dressed truth electron. Smeared values are used for overlap removal consideration.

An efficiency is then calculated as a function of smeared electron  $p_T$  and  $\eta$ . This indicates the offline identification efficiency, and is  $\eta$ -dependent from the detector geometry. To judge whether the efficiency is passed, a number is generated at random from zero to one. The identification efficiency as calculated from the function must exceed the random generated number. To be saved to the output tree, the electron must: have smeared  $p_T > 5$  GeV, pass overlap removal

and the efficiency requirement, be detector stable (`status==1`), and originate from a  $W$  or  $Z$  boson.

The smearing should widen the distribution when compared to truth (see Figure G.1). The decreased yield of smeared electrons is due to the efficiency requirement, which for the HL-LHC is roughly 80%.

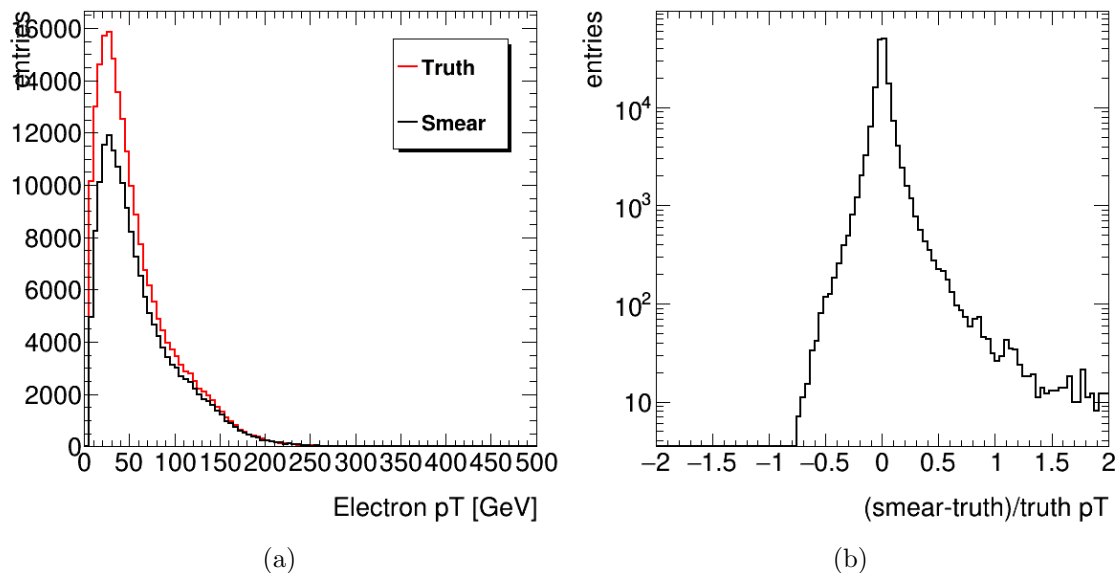


FIGURE G.1. Truth and smeared electron  $p_T$  distributions from a  $W \rightarrow e\nu$  sample with a) MAXHTPTV 140-280 and b) smearing resolution

## G.2. Photons

Truth photons are pulled from truth particle containers in the TRUTH3 xAOD. Each photon is individually considered for smearing, which is done with resolution functions based on the truth photon's full four-momentum. The `tightPhoton` working point is used, as it is the only one fully incorporated into the resolution functions. A default noise scaling value of 0.375 is defined, which



adds additional smearing to the resolution function due to intrinsic and pileup noise sources.

The smeared energy is pulled from a Gaussian distribution around the origin with a width defined by the energy resolution of the detector in a given  $E$  and  $\eta$  slice. The energy resolution uses the noise term to define an error associated with noise and sampling error terms that inform the the energy-scaled relative resolution. The smear amount is added to the true energy and only photons with smeared  $p_T > 20$  GeV are accepted. Studies of the expected performance of ATLAS [164] define additional smearing for the photon's  $\phi$  and  $\eta$  values. All these smeared values constitute a new smeared momentum four-vector for the photon, which is then returned.

An offline identification efficiency is then calculated as a function of smeared photon  $p_T$ . To judge whether the efficiency is passed, a number is generated at random from zero to one. The identification efficiency as calculated from the function must exceed the random generated number. To be saved to the output tree, the photon must have smeared  $p_T > 5$  GeV and pass the efficiency requirement.

The photon efficiency cut dramatically decreases the statistics when compared to truth (see Figure G.2 a)). This is the most evident for low- $p_T$  photons (see Figure G.2 b)), which is expected as offline identification efficiency increased with photon  $p_T$ .

Photons are not used at this stage of the analysis, but could be used by the VBF+MET+ $\gamma$  search which uses shared analysis software.

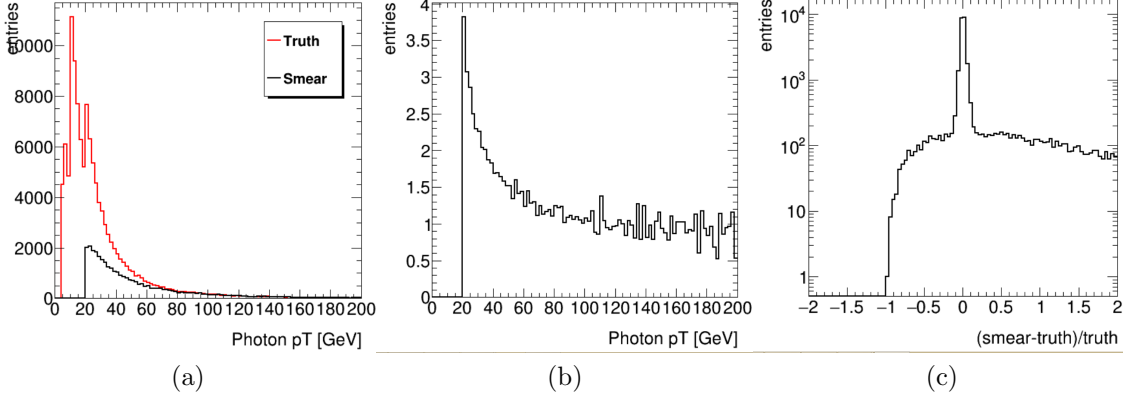


FIGURE G.2. Truth and smeared photon a)  $p_T$  distributions, b) ratio of distributions, and c) smearing resolution (c) from a VBF  $Z \rightarrow \gamma\gamma$  sample

### G.3. Muons

Truth muons are pulled from truth particle containers in the TRUTH3 xAOD. Each muon is individually considered for smearing, which is done with resolution functions based on the truth muon's  $p_T$  and  $\eta$  position. Dressed values are used for these quantities rather than the truth values, as the dressed value more closely represents what is measured in the detector. The `looseMuon` working point is used.

Resolution functions smear the ratio  $q/p_T$  by comparing the momentum resolution of the different aspects of the muon system - the ID and the muon spectrometer (MS). These resolutions are computed separately as functions of the truth muon's  $p_T$  and  $\eta$  and, when appropriately central, combined for a full muon resolution. As with electrons, this resolution is used as the width of a Gaussian centered at the origin, and the  $q/p_T$  smear is pulled from this distribution. The smear amount is added to the true dressed  $q/p_T$  ratio and this smeared  $p_T$  value is extracted and returned. All other muon values ( $\eta/\phi$  position, for example) are

unchanged from the dressed truth muon. Smeared values are used for overlap removal consideration.

An efficiency is then calculated as a function of smeared muon  $p_T$  and  $\eta$ . This indicates the offline identification efficiency, and is  $\eta$ -dependent from the detector geometry. To judge whether the efficiency is passed, a number is generated at random from zero to one. The identification efficiency as calculated from the function must exceed the random generated number. To be saved to the output tree, the muon must: have smeared  $p_T > 5$  GeV, pass overlap removal and the efficiency requirement, be detector stable (`status==1`), and originate from a  $W$  or  $Z$  boson.

The smearing behaves appropriately at low- $p_T$  (see Figure G.3 a)) where the majority of signal muons used in control region definitions lie. Some muons though are smeared to unphysically high  $p_T$  values (see Figure G.3 b)). This results from forward truth muons with  $2.7 < |\eta| < 4$  (see Figure G.4), where only ITk defines the  $q/p_T$  resolution. For 0.17% of true muons in this forward range, a large resolution close to the original  $q/p_T$  value is returned which leaves an unphysically large  $p_T$  upon inversion of the ratio. Though unphysical, this behavior is confirmed by muon experts and analysis-specific mitigation will be necessary.

#### G.4. Neutrinos

Neutrinos exist in the truth record but their attributes are not considered by the analysis software as their signature is not discernible with the detector. They are only used for overlap removal (see Appendix E).

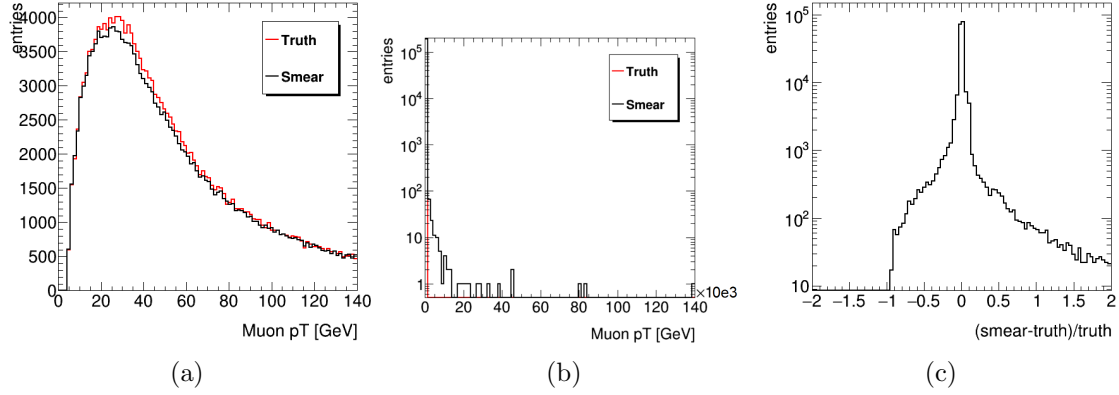


FIGURE G.3. Truth and smeared muon  $p_T$  distributions for a)  $p_T < 140$  GeV, b) the full smeared  $p_T$  range, and c) smearing resolution from a  $W \rightarrow \mu\nu$  sample with MAXHTPTV 140-280

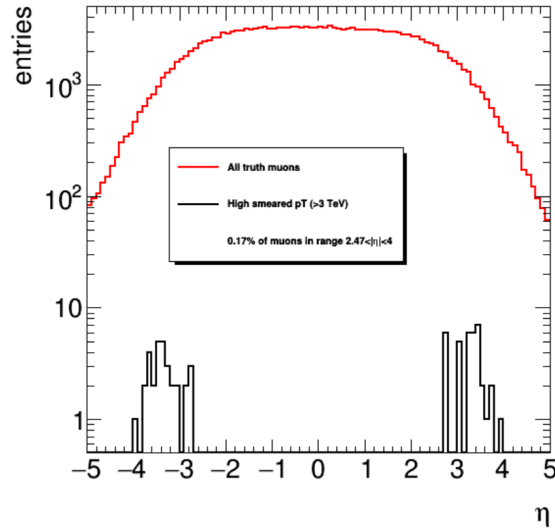


FIGURE G.4.  $\eta$  distribution of truth muons compared to muons with smeared  $p_T > 3$  TeV

## G.5. Jets

Truth jets are pulled from truth particle containers in the TRUTH3 xAOD. Dressed anti-kt particle flow jets with a radius parameter of  $R = 0.4$  are used. Each jet is individually considered for smearing, which is done with Gaussian distributions centered at the origin with a width determined by jet energy

resolution functions based on the the truth jet's  $p_T$  and  $\eta$  position. This smearing value is then applied to the full four-momentum vector of the jet, not only  $p_T$ . An additional boolean variable is created identifying the jet as a smeared hard scatter jet.

Additional pileup jets are also included at this step to emulate  $\langle\mu\rangle = 200$ . These are pulled from a pre-defined dictionary of pileup jets all with  $p_T > 20$  GeV. The hard scatter boolean variable is set to false, identifying the jet as pileup, and the smeared hard scatter jets and pileup jets are combined together into a single jet collection. Before these jets are approved for analysis consideration, they must pass a JVT requirement. This is done separately for smeared hard scatter and pileup jets.

A pseudo-JVT value is calculated for each smeared jet with  $30 < p_T < 100$  GeV and  $|\eta| < 3.8$  (the extent of ITk). This is informed by HGTD information, if the subdetector system is engaged. An efficiency is calculated in  $p_T$  bins and must be passed for the jet to be retained. Pileup suppression is assumed to be constant with a flat efficiency of 0.2% within the ITk range ( $|\eta| < 3.8$ ).

These JVT requirements lend themselves to kinematic restrictions on the jets as well. All smeared and pileup jets must have  $p_T > 30$  GeV and  $|\eta| < 3.8$  for JVT consideration. The VBF signal is characterized by forward jets, so the  $|\eta|$  restriction for returned jets was loosened to  $|\eta| < 4.5$  to allow for more signal. However, this extends beyond the range considered by JVT. Pileup suppression is not assumed beyond the ITk extend (see Figure G.5), leading to large forward pileup contamination. JVT is extended to 200 GeV and the efficiency used at 100 GeV is extended as a conservative estimate (see Figure G.6) to eliminate the discontinuity in the spectrum seen from the increase in pileup.

This all results in one smeared jet collection the combines smeared hard scatter jets and overlaid pileup events that all pass JVT and satisfy  $p_T > 30$  GeV and  $|\eta| < 4.5$  (see Figure G.7). The impact of pileup contamination is largest in the lowest  $Z \rightarrow \nu\nu$  MAXHTPTV slice of 0-70, so the modeling seen here is the poorest of all slices. The impact of pileup is evident.

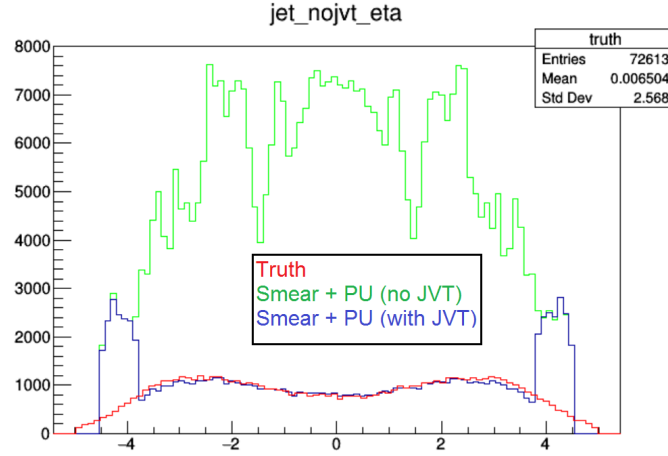


FIGURE G.5. Jet  $\eta$  distribution for truth and smeared jet collections with and without JVT consideration for VBF  $h \rightarrow \text{inv.}$  signal

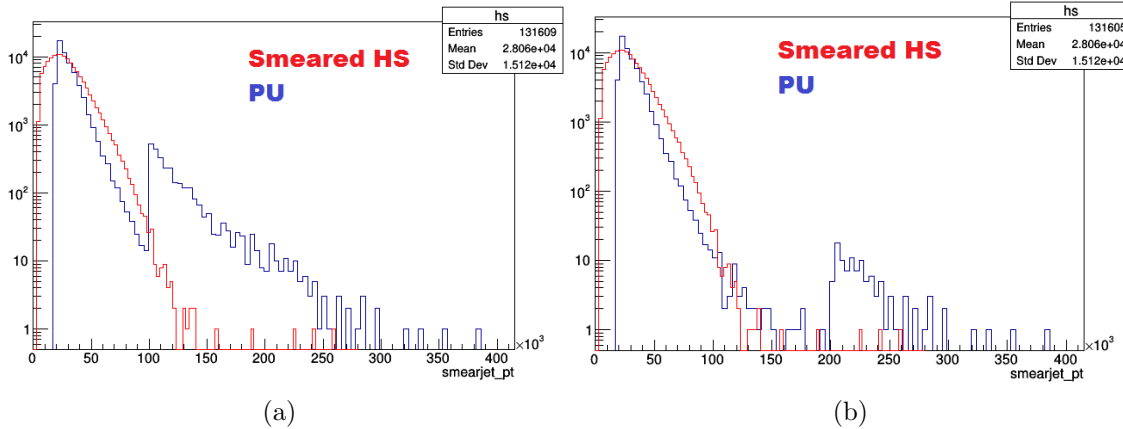


FIGURE G.6. Smeared and pileup jet  $p_T$  distributions for  $W \rightarrow e\nu$  MAXHTPTV 0-70 sample with JVT applied up to a) 100 GeV and b) 200 GeV

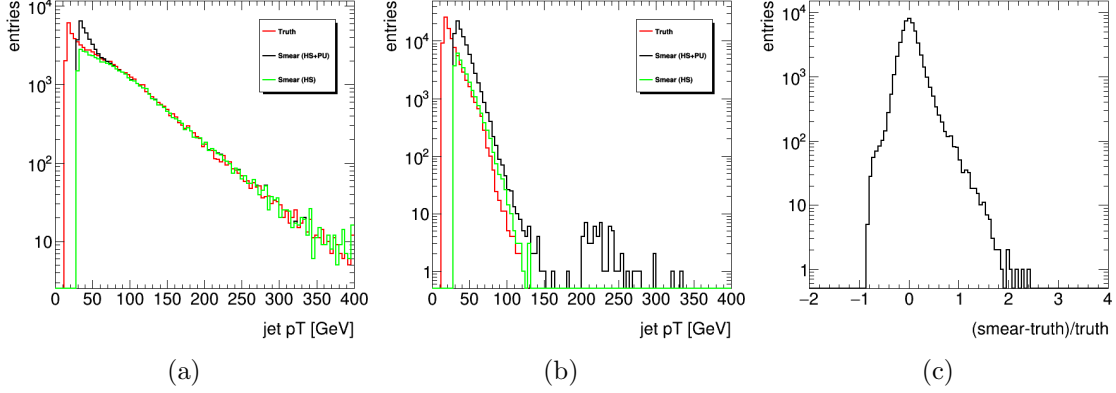


FIGURE G.7. Truth and smeared jet  $p_T$  distributions for a) VBF  $h \rightarrow \text{inv.}$  signal, b)  $Z \rightarrow \nu\nu$  sample with MAXHTPTV 0-70, and c) smearing resolution for VBF  $h \rightarrow \text{inv.}$  signal

## G.6. Missing Transverse Energy

For self-consistency,  $E_T^{\text{miss}}$  is calculated from smeared objects rather than smeared directly from a truth value. Standard  $E_T^{\text{miss}}$  (see Section 4.2.4) is the negative vector sum of physics objects and tracks originating from the primary vertex. A smeared  $E_T^{\text{miss}}$  value is calculated as the negative vector sum of all smeared jets with  $p_T > 30$  GeV. The increased pileup in the forward region beyond  $|\eta| = 3.8$  skews the  $E_T^{\text{miss}}$  distribution toward larger values, so as pileup mitigation a tighter constraint of  $p_T > 50$  GeV is placed upon jets with  $|\eta| > 3.8$ . This value was chosen by optimizing  $S/\sqrt{B}$  where  $S$  is the yield of events with  $E_T^{\text{miss}} > 200$  GeV (signal-like) from a VBF  $h \rightarrow \text{inv.}$  sample and  $B$  is the yield of events with  $E_T^{\text{miss}} > 200$  GeV from a  $Z \rightarrow \nu\nu$  MAXHTPTV 0-70 background sample (see Table G.1).

Tracks are not smeared and added individually to the smeared  $E_T^{\text{miss}}$  as jets are. Instead, an approximation for this soft  $E_T^{\text{miss}}$  term is made based upon [165]. This value is pileup dependent, so first a  $\mu$  value is chosen at random between

| Central<br>jet $p_T$ [GeV] | Forward<br>jet $p_T$ [GeV] | $S$              | $B$                | $S/\sqrt{B}$ | $Z\nu\nu$<br>Resolution RMS |
|----------------------------|----------------------------|------------------|--------------------|--------------|-----------------------------|
| 20                         | 20                         | $3040 \pm 55.1$  | $13492 \pm 116.2$  | 26.17        | 0.319                       |
| 20                         | 30                         | $3305 \pm 57.5$  | $10297 \pm 101.5$  | 32.57        | 0.305                       |
| 20                         | 40                         | $3340 \pm 57.8$  | $7395 \pm 86.0$    | 45.26        | 0.295                       |
| 20                         | 50                         | $3698 \pm 60.8$  | $5447 \pm 73.8$    | 50.11        | 0.288                       |
| 30                         | 30                         | $3032 \pm 55.06$ | $11427 \pm 106.90$ | 28.36        | 0.320                       |
| 30                         | 50                         | $2900 \pm 53.85$ | $7896 \pm 88.86$   | 32.64        | 0.303                       |

TABLE G.1. Optimization of forward ( $|\eta| > 3.8$ ) jet  $p_T$  requirement in  $E_T^{\text{miss}}$  definition

190-210. Soft  $E_T^{\text{miss}}$  components in  $x$  and  $y$  are independently approximated as a random number from a Gaussian distribution centered at the origin with a width of  $10/\sqrt{2}$  GeV which is scaled by  $0.033 \text{ GeV} * \mu$ . In this way, a scalar value of soft  $E_T^{\text{miss}}$  is found (see Figure G.8) and a vector quantity can be added to the full  $E_T^{\text{miss}}$  calculation to address the soft term.

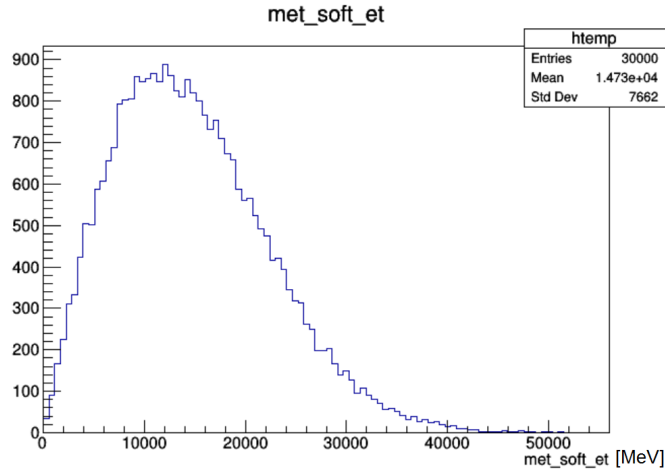


FIGURE G.8. Approximated soft  $E_T^{\text{miss}}$  distribution for VBF  $h \rightarrow \text{inv.}$  signal

The truth  $E_T^{\text{miss}}$  distribution, as taken as the non-interacting  $E_T^{\text{miss}}$  from the truth container for each event, can be replicated with this calculation method using truth objects (see Figure G.9). Therefore, calculating a smeared  $E_T^{\text{miss}}$  in



this manner should be reliable. The presence of pileup leads to degraded  $E_T^{\text{miss}}$  modeling in some samples, even with the stricter forward  $p_T$  requirement.

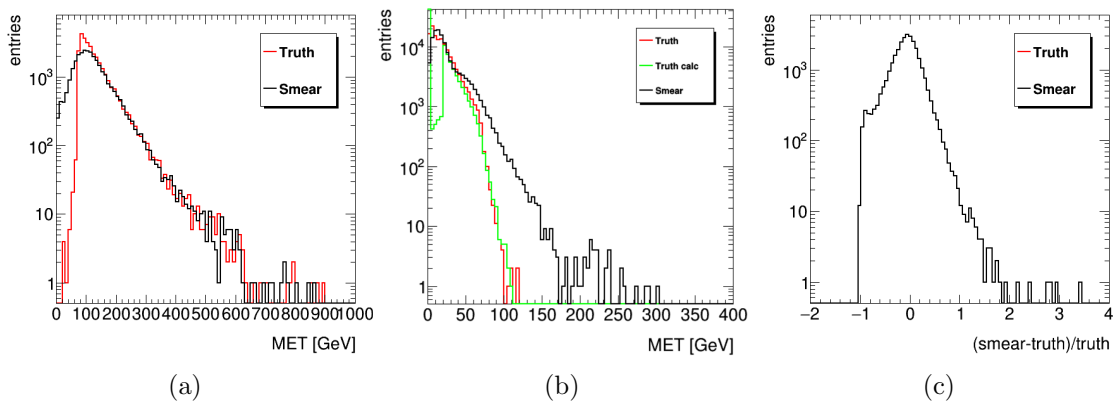


FIGURE G.9. Truth and smeared  $E_T^{\text{miss}}$  distributions for a) VBF  $h \rightarrow \text{inv.}$  signal, b)  $Z \rightarrow \nu\nu$  sample with MAXHTPTV 0-70, and c) smearing resolution for VBF  $h \rightarrow \text{inv.}$  signal

## G.7. Missing Transverse Energy with Leptons

A  $E_T^{\text{miss}}$  variant is calculated which includes electrons and muons so direct  $E_T^{\text{miss}}$  comparisons can be made between the VBF+MET control and signal regions.  $E_T^{\text{miss}}$  is calculated as above, and leptons with  $1 \text{ GeV} < p_T^\ell$  that pass all other criteria as explained above are included in the calculation.

## G.8. Missing Transverse Energy with Jets

Another  $E_T^{\text{miss}}$  variant ( $E_T^{\text{jet, no JVT}}$ ) is calculated from only jets, with no regard for their JVT score. A separate smeared jet collection is created that disregards the JVT requirement and returns any jet satisfying only the  $p_T$  and  $\eta$  requirements.  $E_T^{\text{jet, no JVT}}$  is then calculated identically to  $E_T^{\text{miss}}$  with only the jets in this collection (see Figure G.10).

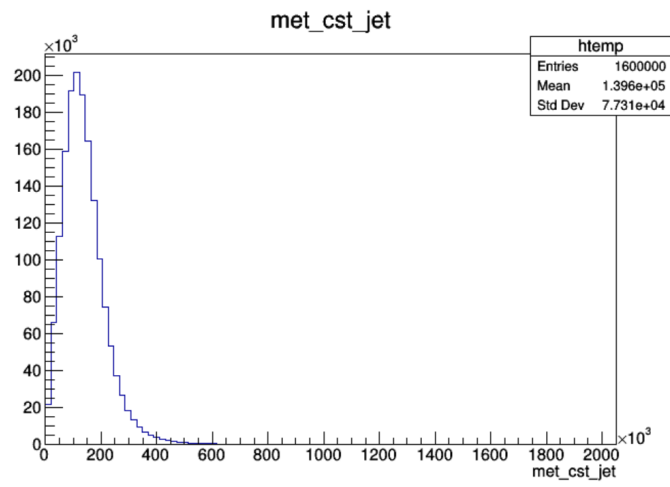


FIGURE G.10. Calculated  $E_T^{\text{jet, no JVT}}$  distribution from smeared quantities for VBF  $h \rightarrow \text{inv. signal}$

## APPENDIX H

### SID ELECTROMAGNETIC CALORIMETER STUDIES AND OPTIMIZATION

Four studies of the my original work considering the SiD electromagnetic calorimeter (ECal) are summarized here. More details can be found in [166, 167, 168, 169]. Additional work collecting data at the DESY test beam with ECal and tracker sensors outfitted with KPiX ASIC chips is summarized.

#### H.1. Backscatter and Shower Spreading

As electromagnetic showers develop, they spread through the calorimeter (see Appendix B). Some fraction of the shower is deflected at larger angles resulting in backscatter (shower components disconnected from the bulk) and shower spreading. To reduce the role that these played in shower identification, only pixel deposits from the SiD detector simulation within an optimized 0.2 radians of the true particle trajectory were considered. This cone definition acted similarly to a radius parameter in jet reconstruction. 97.5% or more of the total shower deposits are contained within this cone over a range of incident particle energies. Lower energy showers produce more backscatter and shower spreading signal.

Further details are provided in [166].

#### H.2. Geometry

The proposed geometry of the SiD ECal with twelve overlapping trapezoidal modules (see Figure 7.10) leads to detector properties asymmetric in  $\phi$ . Nearly 30% of the total  $\phi$  extent contains overlapping modules. This geometry must be

very well understood for shower reconstruction and particle identification, as the signal from multiple modules must be considered with the correct physical relation.

Within the general overlap areas  $\phi \in [(4.03 + 30n)^\circ, (15 + 30n)^\circ]$  (where  $n = 0, 1, 2, \dots, 11$ ) is an area where an incident MIP would only encounter thin layers of tungsten absorber ( $\phi \in [(8.79 + 30n)^\circ, (10.14 + 30n)^\circ]$ ). The total absorber extent and sampling frequency (see Figure H.1) are exaggerated even further in this “thin overlap” region.

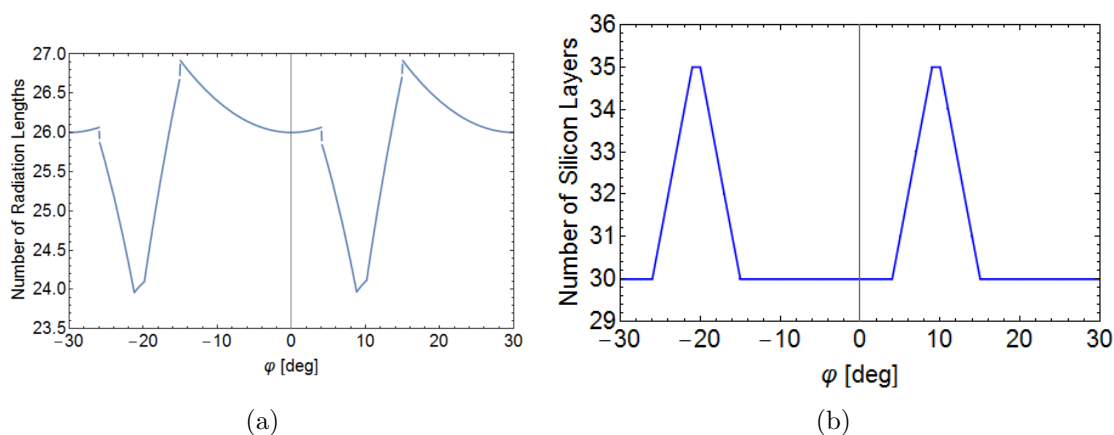


FIGURE H.1. a) Absorber depth and b) number of sampling layers of SiD as a function of  $\phi$ . The overlap region falls within  $[4, 15]$

The effect of these asymmetries is evident from a scan in  $\phi$  of incident particles in the SiD detector simulation. Electromagnetic showers that developed through the overlap region experience larger leakage (shower development beyond the ECal into the following hadron calorimeter), with up to 2% of shower energy for 100 GeV incident electrons being deposited in the hadron calorimeter (HCal) compared to  $\sim 1.1\%$  at normal incidence (see Figure H.2).

These effects could be reduced by increasing the small angle of each ECal trapezoidal module. This would reduce the total fraction of the calorimeter containing overlapping modules without allowing for projective cracks. These

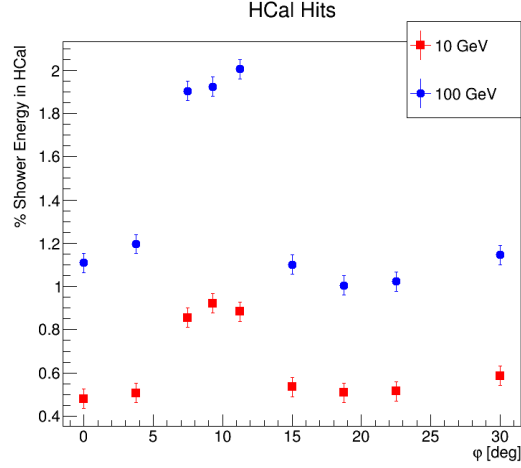


FIGURE H.2. Fraction of electromagnetic shower energy deposited beyond the ECal in the HCal

studies also informed an update to the ECal design, reducing the extent of the first ECal timing silicon layer to only the inner edge of the calorimeter. Original designs imagined this layer to be built along the entire module, but in the overlap region this adds an unnecessary additional layer of complication as it falls within the bulk of the shower.

Further details are provided in [167].

### H.3. Shower Calibration

The previous studies addressed ECal design but the detector response must also be understood. As a first step, response to particles at normal incidence to the detector ( $\theta = 90^\circ$ ,  $\phi = 0^\circ$ ) was considered. These studies will need to be considered as a function of  $\phi$  for full detector understanding.

Higher energy incident particles have higher associated amounts of leakage, or the component of a shower that is not contained within the ECal and is deposited in the HCal (see Figure H.2 ). In a full reconstruction setting, HCal

information associated with the shower could be utilized to inform calibration, but studies here were conducted with the intention of understanding ECal performance and thus only ECal signal was considered. The loss of this information affects energy resolution and calibration ability.

Leakage could simply be estimated as the energy deposited in the last two ECal layers, though on average this underestimated the leakage when compared to HCal deposits. Without HCal calibration there is large associated uncertainty on this measurement, so HCal energy deposits are only used here as a reference for leakage. To be more accurate, shower shape information from the ECal was used and extrapolated into the HCal region. Similar results were produced with linear and gamma distribution ( $y = Ax^\alpha e^{-\beta x}$ ) extrapolations. The leakage was then defined as the integral under the extrapolated curve in the HCal region. This extrapolation was tested on bulk distribution, but can also be done on an event-by-event basis.

The SiD design requires energy resolution around  $20\%/\sqrt{E}$  for particle flow. Energy resolution was calculated for a scan of incident photon energies as  $\sigma/\mu$ , where  $\sigma$  and  $\mu$  are respectively the standard deviation and mean of a Gaussian fit to deposited energy at each energy point. The scaled energy resolution scaled this value by  $\sqrt{E}$ , and is not constant as a function of simulated particle energy  $E$  (see Figure H.3 blue curve). The degraded performance at high incident photon energies results from leakage, so on an event-by-event basis the leakage was calculated with linear extrapolation and added back to the total measured energy. After this correction, the scaled energy resolution at high energy improved by 2.7% (see Figure H.3 red curve).

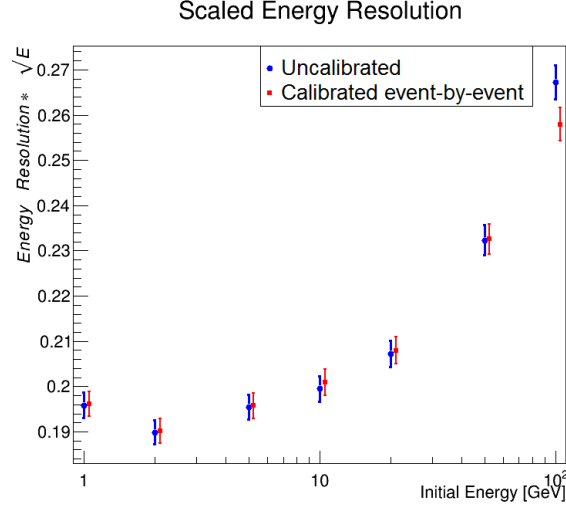


FIGURE H.3. Scaled energy resolution of the ECal barrel with and without leakage correction

The raw deposits made in the detector represent only a small fraction ( $< 2\%$ ) of the total shower energy, as much is lost in the absorber layer. I developed a calibration technique which utilizes detector input to produce the expected incident particle energy which accounts for leakage. Standard calibration simply scaled the measured deposits. This did not account for leakage, so a nonlinear term was included,

$$E_{\text{calib}} = E_{\text{deposit}}(a + bE_{\text{deposit}}^{1/2}) , \quad (\text{H.1})$$

where  $a$  and  $b$  are constants derived from a simultaneous fit over different incident particle energies. The linear term dominated calibration, but at larger incident energy the nonlinear term's contribution increased (see Figure H.4). The nonlinear (or “1/2”) term behaved identically to all other leakage estimation techniques (see Figure H.5), confirming that it accounted for leakage.

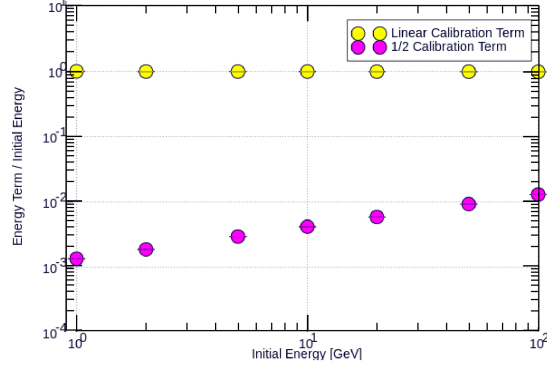


FIGURE H.4. Contribution of linear and nonlinear terms of Eq. H.1 to calibrated energy

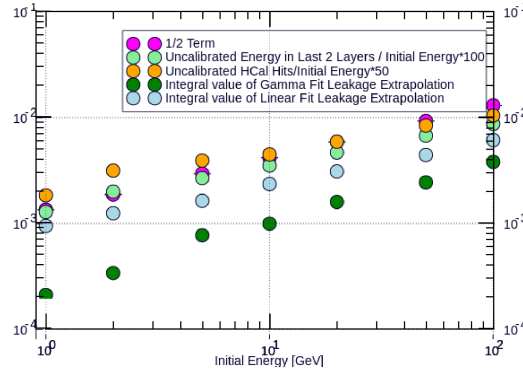


FIGURE H.5. Estimated leakage by different techniques for a scan of incident photon energies

The inclusion of the nonlinear term avoided degradation in calibration as a function of incident particle energy (see Figure H.6) without introducing large error that would jeopardize energy resolution (see Figure H.7).

Further details are provided in [168].

#### H.4. Test Beam Studies

In 2013, an ECal prototype was tested at the SLAC test beam facility. Nine ECal wafers outfitted with KPix ASIC chips were layered with 5 mm tungsten alloy layers, and this stack configuration was exposed to a 12.1 GeV electron beam.



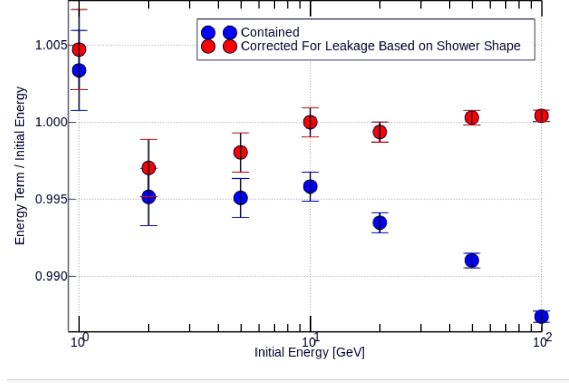


FIGURE H.6. Performance of linear (blue) and nonlinear (red) techniques

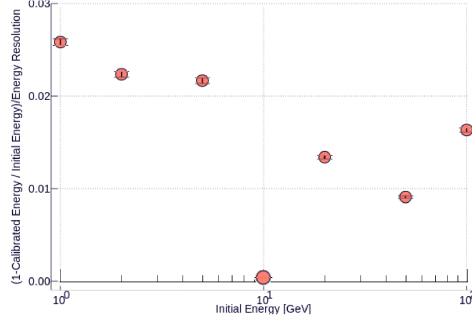


FIGURE H.7. Uncertainty in calibration per energy resolution

With an even number of silicon and tungsten layers, the stack could be inverted such that the beam encountered either a silicon or tungsten layer first. All data was collected with internal KPiX triggering and an ILC-like 5 Hz trigger rate.

This data was cleaned before analysis. Parasitic couplings in the KPiX traces cause all pixels to saturate, creating “monster events” that had to be removed from the dataset. A large number of events with very low deposited energy remained (see Figure H.8 a)), which I removed with the creation of a categorization algorithm. A strict cut would remove both showers with low deposited energy and the contaminating soft photon emission and under-developed showers. Discrimination was determined upon a ratio of deposited charge in each

layer, weighted by depth in the calorimeter,

$$R = \frac{\sum_h L_h^2 C_h}{\sum_h C_h} ,$$

where  $C$  is the deposited charge and  $L$  is the layer number for every recorded pixel hit  $h$ . Soft emission had low  $R$  values with deposits in few layers compared to showers that developed through the entire prototype detector. A cut on  $R$  was then applied to eliminate these contaminating events (see Figure H.8 b)).

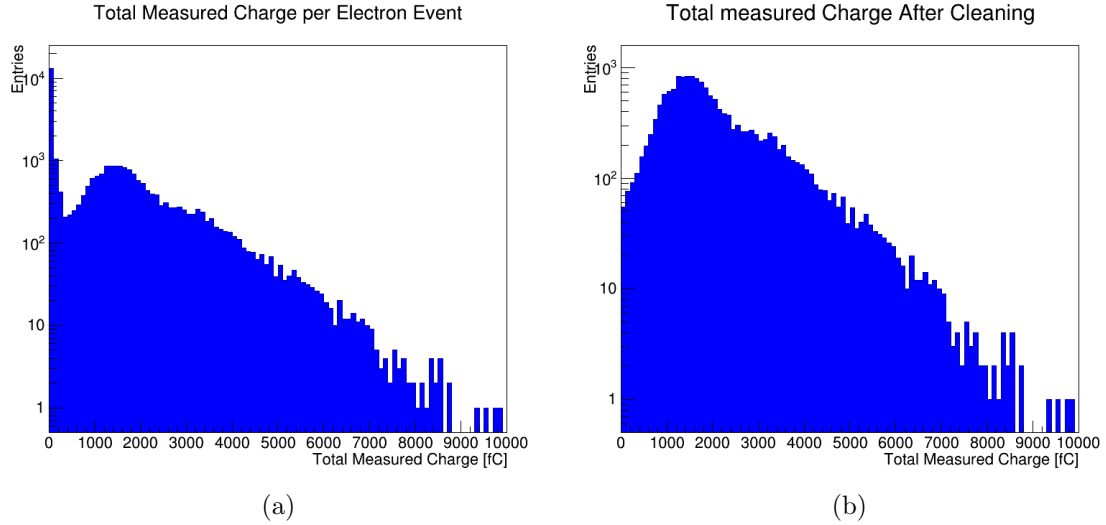


FIGURE H.8. Data collected with the SiD ECal prototype a) before and b) after cleaning

This cleaned data could then be compared to simulated data from a Geant4 prototype simulation. This was optimized to mimic experimental conditions, including randomly distributed dead pixels and overlaying electron events to model test beam events with multiple incident particles. I modeled this test beam structure with a Poisson distribution and scaled through a fit to the data itself. This lead to agreement between cleaned test beam data and Geant4 simulation for both silicon-first and tungsten-first configurations (see Figure H.9). The early

showering induced with a tungsten layer in front required an additional smearing parameter to cover additional experimental uncertainties. This agreement is also seen in each individual layer (illustrated in [169]).

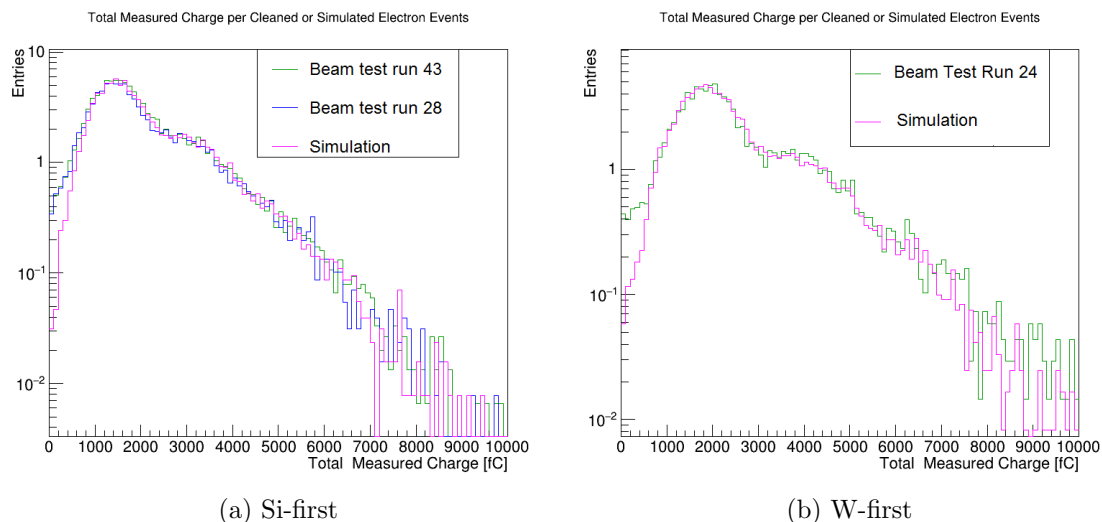


FIGURE H.9. Test beam data runs and simulated data match for a) silicon-first and b) tungsten-first configurations

Using these cleaned and optimized data sets, I developed an algorithm to consider the efficiency of shower separation. This rudimentary clustering algorithm identified the core of each shower by identifying the pixel in each layer with the most deposited charge. The algorithm examined each layer of each event and determined local maxima of charge deposits. The pixels are labeled identically in each layer. An incident particle is identified if the same pixel location was a local maximum in at least four layers. The algorithm vetoes events in which the incident particle fell on the border of two pixels, leaving nearly equal deposits in both. This simple method tended to undercount the true number of incident particles, with 90.3% of simulated events counted correctly. 82.6% of simulated events with two incident particles were correctly identified, with an average efficiency of

98.5% if the shower maxima were separated by more than 1 cm (see Figure H.10). The  $13 \text{ mm}^2$  hexagonal pixels have a long axis of 4.5 mm, so efficiency at low separation is inherently limited.

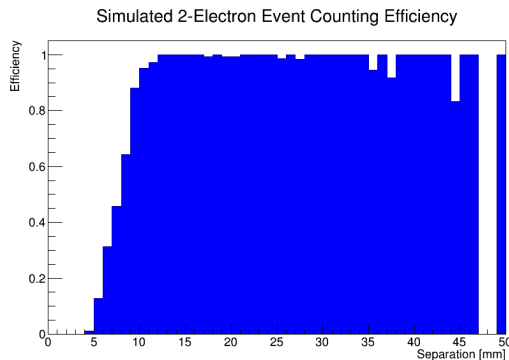


FIGURE H.10. Efficiency of incident particle counting algorithm for simulated events with two incident particles

The tendency to undercount was noted to a higher degree when used with the test beam data as well (see Figure H.11 a)). This could account for the discrepancy in shape between identified two-particle events in simulation and data (see Figure H.11 b)). 4 mm bins are used to approximate pixel size.

Further details are provided in [169].

## H.5. KPiX Test Beam

In 2019, ECal wafers and tracker sensors fitted with KPiX ASIC chips were utilized in a data collection campaign at the DESY test beam facility. Local interest was also motivated by the development of LYCORIS, a beam telescope for the DESY test beam, featuring the KPiX chip [170]. I assisted with the initial test of KPiX readout boards and remote access, focusing on ensuring that basic KPiX functions such as triggering and beam synchronization could be achieved in

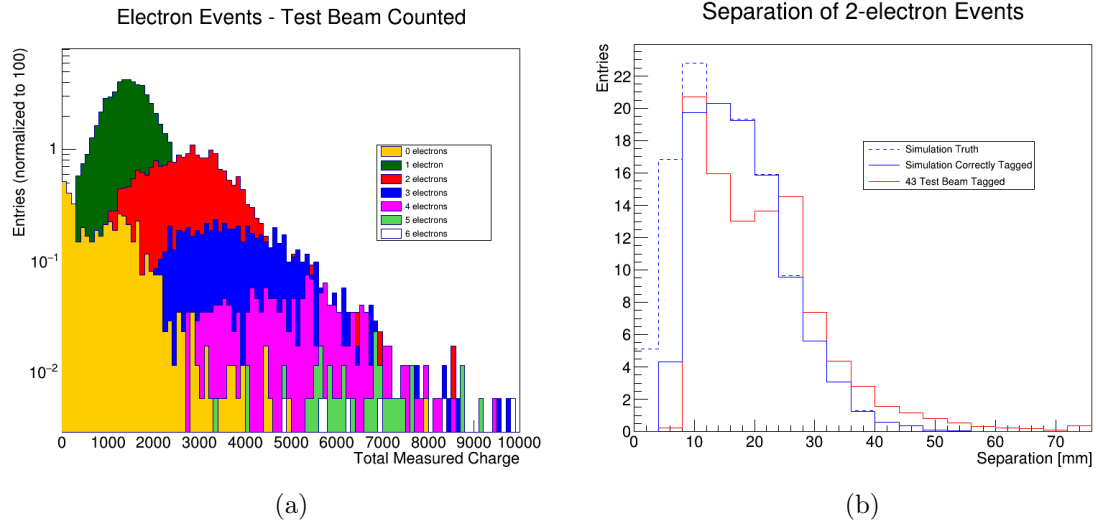


FIGURE H.11. a) Counting algorithm applied to the full test beam data set and b) comparison of identified two-particle events between simulation and data

the test beam environment. Initial calibration data was gathered, which informed further LYCORIS studies.

## REFERENCES CITED

- [1] Hitoshi Murayama and Kurt Risselmann. Doe explains... the standard model of particle physics. <https://www.energy.gov/science/doe-explainsthe-standard-model-particle-physics>.
- [2] John Ellis. Higgs Physics. (KCL-PH-TH-2013-49. KCL-PH-TH-2013-49. LCTS-2013-36. CERN-PH-TH-2013-315):117–168. 52 p, Dec 2013. doi:10.5170/CERN-2015-004.117. URL <https://cds.cern.ch/record/1638469>. 52 pages, 45 figures, Lectures presented at the ESHEP 2013 School of High-Energy Physics, to appear as part of the proceedings in a CERN Yellow Report.
- [3] *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*. CERN Yellow Reports: Monographs. Oct 2016. doi:10.23731/CYRM-2017-002. URL <https://cds.cern.ch/record/2227475>. 869 pages, 295 figures, 248 tables and 1645 citations. Working Group web page: <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWG>.
- [4] C. Grojean. Higgs Physics. (arXiv:1708.00794):143–158. 12 p, Aug 2017. doi:10.5170/CERN-2016-005.143. URL <https://cds.cern.ch/record/2243593>. 12 pages, contribution to the CERN in the Proceedings of the 2015 CERN-Latin-American School of High-Energy Physics, Ibarra, Ecuador, 4 - 17 March 2015.
- [5] D.M. Asner, T. Barklow, C. Calancha, K. Fujii, N. Graf, H.E. Haber, A. Ishikawa, S. Kanemura, S. Kawada, M. Kurata, A. Miyamoto, H. Neal, H. Ono, C. Potter, J. Strube, T. Suehara, T. Tanabe, J. Tian, J. Tsumura, S. Watanuki, G. Weiglein, K. Yagyu, and H. Yokoya. ILC Higgs White Paper. Technical Report arXiv:1310.0763, Oct 2013. URL <https://cds.cern.ch/record/1605366>. Comments: 143 pages, 77 figures, Snowmass 2013.
- [6] Oleg Brandt. The speaker’s guide to the dark matter (in our galaxy). <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/ASpeakersGuideToTheDM>.
- [7] Tongyan Lin. Dark matter models and direct detection. *PoS*, 333:009, 2019. doi:10.22323/1.333.0009.
- [8] Cinzia De Melis. The CERN accelerator complex. Complexe des acclrateurs du CERN. Jan 2016. URL <https://cds.cern.ch/record/2119882>. General Photo.

- [9] ATLAS Experiment Public Results. ATLAS Stand-AloneEvent Displays, Aug 2012. URL [https://twiki.cern.ch/twiki/bin/viewauth/AtlasPublic/EventDisplayStandAlone#Events\\_with\\_pile\\_up](https://twiki.cern.ch/twiki/bin/viewauth/AtlasPublic/EventDisplayStandAlone#Events_with_pile_up).
- [10] ATLAS Experiment Public Results. LuminosityPublicResultsRun2, July 2015. URL <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2>.
- [11] The ATLAS Collaboration. The ATLAS experiment at the CERN large hadron collider. *Journal of Instrumentation*, 3(08):S08003–S08003, aug 2008. doi:10.1088/1748-0221/3/08/s08003. URL <https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08003>.
- [12] Fabian Kuger. Signal Formation Processes in Micromegas Detectors and Quality Control for large size Detector Construction for the ATLAS New Small Wheel, Aug 2017. URL <https://cds.cern.ch/record/2277011>. Presented 31 Jul 2017.
- [13] Karolos Potamianos. The upgraded pixel detector and the commissioning of the inner detector tracking of the atlas experiment for run-2 at the large hadron collider, 2016.
- [14] *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1996. URL <https://cds.cern.ch/record/331061>.
- [15] M. Aaboud, G. Aad, B. Abbott, J. Abdallah, O. Abdinov, B. Abeloos, R. Aben, O. S. AbouZeid, N. L. Abraham, and et al. Performance of the atlas trigger system in 2015. *The European Physical Journal C*, 77(5), May 2017. ISSN 1434-6052. doi:10.1140/epjc/s10052-017-4852-3. URL <http://dx.doi.org/10.1140/epjc/s10052-017-4852-3>.
- [16] Martin zur Nedden. The Run-2 ATLAS Trigger System: Design, Performance and Plan. Technical Report ATL-DAQ-PROC-2016-039, CERN, Geneva, Dec 2016. URL <https://cds.cern.ch/record/2238679>.
- [17] Chaowaroj Wanotayaroj. Search for a Scalar Partner of the Top Quark in the Jets+MET Final State with the ATLAS detector, Nov 2016. URL <https://cds.cern.ch/record/2242196>. Presented 25 Oct 2016.
- [18] Joao Pequeno. Event Cross Section in a computer generated image of the ATLAS detector. Mar 2008. URL <https://cds.cern.ch/record/1096081>.
- [19] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. The anti-ktjet clustering algorithm. *Journal of High Energy Physics*, 2008(04):063063, Apr 2008. ISSN 1029-8479. doi:10.1088/1126-6708/2008/04/063. URL <http://dx.doi.org/10.1088/1126-6708/2008/04/063>.

- [20] Tagging and suppression of pileup jets with the ATLAS detector. Technical Report ATLAS-CONF-2014-018, CERN, Geneva, May 2014. URL <https://cds.cern.ch/record/1700870>.
- [21] Per Ola Hansson Adrian. The ATLAS b-Jet Trigger. Technical Report arXiv:1111.4190, CERN, Geneva, Nov 2011. URL <https://cds.cern.ch/record/1397942>. Comments: 4 pages, 6 figures, conference proceedings for PIC2011.
- [22] S. Chekanov. Boosted jets in high-granularity calorimeter at a 100 tev pp collider. [https://indico.cern.ch/event/579660/contributions/2575089/attachments/1495811/2327298/B00ST2017\\_hcal.pdf](https://indico.cern.ch/event/579660/contributions/2575089/attachments/1495811/2327298/B00ST2017_hcal.pdf).
- [23] Andrew J. Larkoski, Gavin P. Salam, and Jesse Thaler. Energy correlation functions for jet substructure. *Journal of High Energy Physics*, 2013(6), Jun 2013. ISSN 1029-8479. doi:10.1007/jhep06(2013)108. URL [http://dx.doi.org/10.1007/JHEP06\(2013\)108](http://dx.doi.org/10.1007/JHEP06(2013)108).
- [24] Georges Aad et al. Performance of the missing transverse momentum triggers for the ATLAS detector during Run-2 data taking. *JHEP*, 08:080, 2020. doi:10.1007/JHEP08(2020)080.
- [25] ATLAS Collaboration. Search for invisible Higgs boson decays with vector boson fusion signatures with the ATLAS detector using an integrated luminosity of  $139\text{fb}^{-1}$ . Technical report, 2020. URL <http://cdsweb.cern.ch/record/2715447>.
- [26] Longer term LHC schedule. <https://lhc-commissioning.web.cern.ch/schedule/LHC-long-term.htm>.
- [27] ATLAS Collaboration. Simulated HL-LHC collision event in the ATLAS detector. General Photo, May 2019. URL <https://cds.cern.ch/record/2674770>.
- [28] Technical Design Report: A High-Granularity Timing Detector for the ATLAS Phase-II Upgrade. Technical Report CERN-LHCC-2020-007. ATLAS-TDR-031, CERN, Geneva, Jun 2020. URL <https://cds.cern.ch/record/2719855>.
- [29] Search for the Standard Model Higgs boson in produced in association with a vector boson and decaying to bottom quarks with the ATLAS detector. Technical Report ATLAS-CONF-2012-161, CERN, Geneva, Nov 2012. URL <http://cds.cern.ch/record/1493625>.



- [30] Ties Behnke, James E. Brau, Brian Foster, Juan Fuster, Mike Harrison, James McEwan Paterson, Michael Peskin, Marcel Stanitzki, Nicholas Walker, and Hitoshi Yamamoto. The international linear collider technical design report - volume 1: Executive summary, 2013.
- [31] Pushpa Bhat. Beginning a new phase toward preparation for the International Linear Collider, November 2020. URL <https://news.fnal.gov/2020/11/beginning-new-phase-toward-preparation-for-international-linear-collider/>.
- [32] Philip Bambade, Tim Barklow, Ties Behnke, Mikael Berggren, James Brau, Philip Burrows, Dmitri Denisov, Angeles Faus-Golfe, Brian Foster, Keisuke Fujii, Juan Fuster, Frank Gaede, Paul Grannis, Christophe Grojean, Andrew Hutton, Benno List, Jenny List, Shinichiro Michizono, Akiya Miyamoto, Olivier Napoly, Michael Peskin, Roman Poeschl, Frank Simon, Jan Strube, Junping Tian, Maksym Titov, Marcel Vos, Andrew White, Graham Wilson, Akira Yamamoto, Hitoshi Yamamoto, and Kaoru Yokoya. The international linear collider: A global project, 2019.
- [33] G. Aad, B. Abbott, J. Abdallah, A. A. Abdelalim, A. Abdesselam, O. Abdinov, B. Abi, M. Abolins, H. Abramowicz, and et al. Performance of the atlas trigger system in 2010. *The European Physical Journal C*, 72(1), Jan 2012. ISSN 1434-6052. doi:10.1140/epjc/s10052-011-1849-1. URL <http://dx.doi.org/10.1140/epjc/s10052-011-1849-1>.
- [34] P.A. Zyla *et al.* (Particle Data Group). Review of particle physics. <https://academic.oup.com/ptep/article/2020/8/083C01/5891211>, August 2020.
- [35] J. J. Thomson M.A. F.R.S. Xl. cathode rays. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 44(269):293–316, 1897. doi:10.1080/14786449708621070. URL <https://doi.org/10.1080/14786449708621070>.
- [36] Peter W. Higgs. Broken symmetries and the masses of gauge bosons. <https://link.aps.org/doi/10.1103/PhysRevLett.13.508>, October 1964.
- [37] P.W. Higgs. Broken symmetries, massless particles and gauge fields. *Physics Letters*, 12(2):132 – 133, 1964. ISSN 0031-9163. doi:[https://doi.org/10.1016/0031-9163\(64\)91136-9](https://doi.org/10.1016/0031-9163(64)91136-9). URL <http://www.sciencedirect.com/science/article/pii/0031916364911369>.
- [38] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. <https://link.aps.org/doi/10.1103/PhysRevLett.13.321>, August 1964.

- [39] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global conservatoin laws and massless particles.  
<https://link.aps.org/doi/10.1103/PhysRevLett.13.585>, November 1964.
- [40] The nobel prize in physics 2012.  
<urlhttps://www.nobelprize.org/prizes/physics/2013/summary/>, 2013.
- [41] ATLAS Collaboration. Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc. *Physics Letters B*, 716(1):1 – 29, 2012. ISSN 0370-2693.  
doi:<https://doi.org/10.1016/j.physletb.2012.08.020>. URL <http://www.sciencedirect.com/science/article/pii/S037026931200857X>.
- [42] CMS Collaboration. Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc. *Physics Letters B*, 716(1):30 – 61, 2012. ISSN 0370-2693. doi:<https://doi.org/10.1016/j.physletb.2012.08.021>. URL <http://www.sciencedirect.com/science/article/pii/S0370269312008581>.
- [43] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to quantum field theory*. Addison-Wesley, Reading, USA, 1995. ISBN 978-0-201-50397-5.
- [44] Mark Srednicki. *Quantum Field Theory*. Cambridge Univ. Press, Cambridge, 2007. URL <https://cds.cern.ch/record/1019751>.
- [45] Virginia Trimble. Existence and nature of dark matter in the universe. *Annual Review of Astronomy and Astrophysics*, 25(1):425–472, 1987.  
doi:10.1146/annurev.aa.25.090187.002233. URL <https://doi.org/10.1146/annurev.aa.25.090187.002233>.
- [46] Gianfranco Bertone, Dan Hooper, and Joseph Silk. Particle dark matter: evidence, candidates and constraints. *Physics Reports*, 405(5):279 – 390, 2005. ISSN 0370-1573. doi:<https://doi.org/10.1016/j.physrep.2004.08.031>.  
URL <http://www.sciencedirect.com/science/article/pii/S0370157304003515>.
- [47] N. Aghanim, Y. Akrami, F. Arroja, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, and et atl. Planck2018 results. *Astronomy & Astrophysics*, 641:A1, Sep 2020. ISSN 1432-0746.  
doi:10.1051/0004-6361/201833880. URL <http://dx.doi.org/10.1051/0004-6361/201833880>.
- [48] Eder Izaguirre and Brian Shuve. Multilepton and lepton jet probes of sub-weak-scale right-handed neutrinos. *Physical Review D*, 91(9), May 2015. ISSN 1550-2368. doi:10.1103/physrevd.91.093010. URL <http://dx.doi.org/10.1103/PhysRevD.91.093010>.

- [49] Brian Patt and Frank Wilczek. Higgs-field portal into hidden sectors. 5 2006.
- [50] Abdelhak Djouadi, Oleg Lebedev, Yann Mambrini, and Jrmie Quevillon. Implications of lhc searches for higgs-portal dark matter. *Physics Letters B*, 709(1):65 – 69, 2012. ISSN 0370-2693. doi:<https://doi.org/10.1016/j.physletb.2012.01.062>. URL <http://www.sciencedirect.com/science/article/pii/S0370269312001037>.
- [51] Abdelhak Djouadi, Adam Falkowski, Yann Mambrini, and Jeremie Quevillon. Direct Detection of Higgs-Portal Dark Matter at the LHC. *Eur. Phys. J.*, C73(6):2455, 2013. doi:10.1140/epjc/s10052-013-2455-1.
- [52] Robert E. Shrock and Mahiko Suzuki. Invisible decays of higgs bosons. *Physics Letters B*, 110(3):250 – 254, 1982. ISSN 0370-2693. doi:[https://doi.org/10.1016/0370-2693\(82\)91247-3](https://doi.org/10.1016/0370-2693(82)91247-3). URL <http://www.sciencedirect.com/science/article/pii/0370269382912473>.
- [53] David Curtin, Rouven Essig, Stefania Gori, Prerit Jaiswal, Andrey Katz, Tao Liu, Zhen Liu, David McKeen, Jessie Shelton, Matthew Strassler, and et al. Exotic decays of the 125 gev higgs boson. *Physical Review D*, 90(7), Oct 2014. ISSN 1550-2368. doi:10.1103/physrevd.90.075004. URL <http://dx.doi.org/10.1103/PhysRevD.90.075004>.
- [54] Shinya Kanemura, Shigeki Matsumoto, Takehiro Nabeshima, and Nobuchika Okada. Can wimp dark matter overcome the nightmare scenario? *Physical Review D*, 82(5), Sep 2010. ISSN 1550-2368. doi:10.1103/physrevd.82.055026. URL <http://dx.doi.org/10.1103/PhysRevD.82.055026>.
- [55] D. S. Akerib et al. Results from a search for dark matter in the complete LUX exposure. *Phys. Rev. Lett.*, 118(2):021303, 2017. doi:10.1103/PhysRevLett.118.021303.
- [56] E. Aprile et al. Dark Matter Search Results from a One Ton-Year Exposure of XENON1T. *Phys. Rev. Lett.*, 121(11):111302, 2018. doi:10.1103/PhysRevLett.121.111302.
- [57] Xiangyi Cui et al. Dark Matter Results From 54-Ton-Day Exposure of PandaX-II Experiment. *Phys. Rev. Lett.*, 119(18):181302, 2017. doi:10.1103/PhysRevLett.119.181302.
- [58] P. Agnes et al. Low-Mass Dark Matter Search with the DarkSide-50 Experiment. *Phys. Rev. Lett.*, 121(8):081307, 2018. doi:10.1103/PhysRevLett.121.081307.

- [59] A. Albert, B. Anderson, K. Bechtol, A. Drlica-Wagner, M. Meyer, M. Snchez-Conde, L. Strigari, M. Wood, T. M. C. Abbott, F. B. Abdalla, and et al. Searching for dark matter annihilation in recently discovered milky way satellites withfermi-lat. *The Astrophysical Journal*, 834(2):110, Jan 2017. ISSN 1538-4357. doi:10.3847/1538-4357/834/2/110. URL <http://dx.doi.org/10.3847/1538-4357/834/2/110>.
- [60] Lyndon R Evans and Philip Bryant. LHC Machine. *JINST*, 3:S08001. 164 p, 2008. doi:10.1088/1748-0221/3/08/S08001. URL <https://cds.cern.ch/record/1129806>. This report is an abridged version of the LHC Design Report (CERN-2004-003).
- [61] The CMS Collaboration. The CMS experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08004–S08004, aug 2008. doi:10.1088/1748-0221/3/08/s08004. URL <https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08004>.
- [62] The LHCb Collaboration. The LHCb detector at the LHC. *Journal of Instrumentation*, 3(08):S08005–S08005, aug 2008. doi:10.1088/1748-0221/3/08/s08005. URL <https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08005>.
- [63] The ALICE Collaboration. The ALICE experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08002–S08002, aug 2008. doi:10.1088/1748-0221/3/08/s08002. URL <https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08002>.
- [64] Linear accelerator 2. Sep 2012. URL <https://cds.cern.ch/record/1997427>.
- [65] Richard Scrivens and Maurizio Vretenar. Linac2: The tale of a billion-trillion protons. Dec 2018. URL <http://cds.cern.ch/record/2654723>.
- [66] Giorgio Brianti. The CERN synchrotrons. 1997. doi:10.5170/CERN-1997-004.29. URL <https://cds.cern.ch/record/340514>.
- [67] Pulling together: Superconducting electromagnets. <https://home.cern/science/engineering/pulling-together-superconducting-electromagnets>.
- [68] M. Aaboud, G. Aad, B. Abbott, J. Abdallah, O. Abdinov, B. Abeloos, R. Aben, O. S. AbouZeid, N. L. Abraham, and et al. Luminosity determination in pp collisions at  $\sqrt{s} = 8$  tev using the atlas detector at the lhc. *The European Physical Journal C*, 76(12), Nov 2016. ISSN 1434-6052. doi:10.1140/epjc/s10052-016-4466-1. URL <http://dx.doi.org/10.1140/epjc/s10052-016-4466-1>.

- [69] Luminosity determination in  $pp$  collisions at  $\sqrt{s} = 13$  TeV using the ATLAS detector at the LHC. Technical Report ATLAS-CONF-2019-021, CERN, Geneva, Jun 2019. URL <https://cds.cern.ch/record/2677054>.
- [70] Fu-Hu Liu, Ya-Hui Chen, Ya-Qin Gao, and Er-Qin Wang. On current conversion between particle rapidity and pseudorapidity distributions in high energy collisions. *Advances in High Energy Physics*, 2013, 2013. doi:10.1155/2013/710534. URL <https://doi.org/10.1155/2013/710534>.
- [71] *ATLAS inner detector: Technical Design Report, 1*. Technical Design Report ATLAS. CERN, Geneva, 1997. URL <https://cds.cern.ch/record/331063>.
- [72] M Capeans, G Darbo, K Einsweiler, M Elsing, T Flick, M Garcia-Sciveres, C Gemme, H Pernegger, O Rohne, and R Vuillermet. ATLAS Insertable B-Layer Technical Design Report. Technical Report CERN-LHCC-2010-013. ATLAS-TDR-19, Sep 2010. URL <https://cds.cern.ch/record/1291633>.
- [73] Bartosz Mindur. Atlas transition radiation tracker (trt): Straw tubes for tracking and particle identification at the large hadron collider. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 845:257 – 261, 2017. ISSN 0168-9002. doi:<https://doi.org/10.1016/j.nima.2016.04.026>. URL <http://www.sciencedirect.com/science/article/pii/S0168900216301905>. Proceedings of the Vienna Conference on Instrumentation 2016.
- [74] Boris Dolgoshein. Transition radiation detectors. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 326(3):434 – 469, 1993. ISSN 0168-9002. doi:[https://doi.org/10.1016/0168-9002\(93\)90846-A](https://doi.org/10.1016/0168-9002(93)90846-A). URL <http://www.sciencedirect.com/science/article/pii/016890029390846A>.
- [75] Christian Wolfgang Fabjan and F Gianotti. Calorimetry for Particle Physics. *Rev. Mod. Phys.*, 75(CERN-EP-2003-075):1243–1286. 96 p, Oct 2003. doi:10.1103/RevModPhys.75.1243. URL <https://cds.cern.ch/record/692252>.
- [76] *ATLAS tile calorimeter: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1996. URL <https://cds.cern.ch/record/331062>.
- [77] G. Aad, B. Abbott, J. Abdallah, S. Abdel Khalek, O. Abdinov, R. Aben, B. Abi, M. Abolins, O. S. AbouZeid, and et al. Electron and photon energy calibration with the atlas detector using lhcb run 1 data. *The European Physical Journal C*, 74(10), Oct 2014. ISSN 1434-6052. doi:10.1140/epjc/s10052-014-3071-4. URL <http://dx.doi.org/10.1140/epjc/s10052-014-3071-4>.

- [78] *ATLAS muon spectrometer: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1997. URL <https://cds.cern.ch/record/331068>.
- [79] *ATLAS magnet system: Technical Design Report, 1*. Technical Design Report ATLAS. CERN, Geneva, 1997. URL <https://cds.cern.ch/record/338080>.
- [80] ATLAS Outreach. ATLAS Fact Sheet : To raise awareness of the ATLAS detector and collaboration on the LHC. 2010. URL <https://cds.cern.ch/record/1457044>.
- [81] M. Abolins et al. The ATLAS Data Acquisition and High Level Trigger system. *JINST*, 11(06):P06008, 2016. doi:10.1088/1748-0221/11/06/P06008.
- [82] ATLAS Collaboration. The ATLAS Level-1 Calorimeter Trigger. Technical Report ATL-DAQ-PUB-2008-001. ATL-COM-DAQ-2008-002, CERN, Geneva, Jan 2008. URL <http://cds.cern.ch/record/1080560>.
- [83] The ATLAS collaboration. Operation of the atlas trigger system in run 2. *Journal of Instrumentation*, 15(10):P10004P10004, Oct 2020. ISSN 1748-0221. doi:10.1088/1748-0221/15/10/p10004. URL <http://dx.doi.org/10.1088/1748-0221/15/10/P10004>.
- [84] Richard D. Ball, Valerio Bertone, Stefano Carrazza, Christopher S. Deans, Luigi Del Debbio, Stefano Forte, Alberto Guffanti, Nathan P. Hartland, Jos I. Latorre, Juan Rojo, and et al. Parton distributions with lhc data. *Nuclear Physics B*, 867(2):244289, Feb 2013. ISSN 0550-3213. doi:10.1016/j.nuclphysb.2012.10.003. URL <http://dx.doi.org/10.1016/j.nuclphysb.2012.10.003>.
- [85] The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie-Landshoff diffractive model. Technical Report ATL-PHYS-PUB-2016-017, CERN, Geneva, Aug 2016. URL <https://cds.cern.ch/record/2206965>.
- [86] Geant4a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003. ISSN 0168-9002. doi:[https://doi.org/10.1016/S0168-9002\(03\)01368-8](https://doi.org/10.1016/S0168-9002(03)01368-8). URL <https://www.sciencedirect.com/science/article/pii/S0168900203013688>.
- [87] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. The catchment area of jets. *Journal of High Energy Physics*, 2008(04):005–005, apr 2008. doi:10.1088/1126-6708/2008/04/005. URL <https://doi.org/10.1088/1126-6708/2008/04/005>.

- [88] G. Aad, B. Abbott, J. Abdallah, O. Abdinov, R. Aben, M. Abolins, O. S. AbouZeid, H. Abramowicz, H. Abreu, and et al. Topological cell clustering in the atlas calorimeters and its performance in lh run 1. *The European Physical Journal C*, 77(7), Jul 2017. ISSN 1434-6052. doi:10.1140/epjc/s10052-017-5004-5. URL <http://dx.doi.org/10.1140/epjc/s10052-017-5004-5>.
- [89] M. Aaboud, G. Aad, B. Abbott, J. Abdallah, O. Abdinov, B. Abeloos, S. H. Abidi, O. S. AbouZeid, N. L. Abraham, and et al. Jet reconstruction and performance using particle flow with the atlas detector. *The European Physical Journal C*, 77(7), Jul 2017. ISSN 1434-6052. doi:10.1140/epjc/s10052-017-5031-2. URL <http://dx.doi.org/10.1140/epjc/s10052-017-5031-2>.
- [90] ALEPH Collaboration. Performance of the aleph detector at lep. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 360(3):481 – 506, 1995. ISSN 0168-9002. doi:[https://doi.org/10.1016/0168-9002\(95\)00138-7](https://doi.org/10.1016/0168-9002(95)00138-7). URL <http://www.sciencedirect.com/science/article/pii/0168900295001387>.
- [91] Stephen D. Ellis and Davison E. Soper. Successive combination jet algorithm for hadron collisions. *Physical Review D*, 48(7):31603166, Oct 1993. ISSN 0556-2821. doi:10.1103/physrevd.48.3160. URL <http://dx.doi.org/10.1103/PhysRevD.48.3160>.
- [92] Matteo Cacciari and Gavin P. Salam. Pileup subtraction using jet areas. *Physics Letters B*, 659(1):119–126, 2008. ISSN 0370-2693. doi:<https://doi.org/10.1016/j.physletb.2007.09.077>. URL <https://www.sciencedirect.com/science/article/pii/S0370269307011094>.
- [93] M. Aaboud, G. Aad, B. Abbott, J. Abdallah, O. Abdinov, B. Abeloos, S.H. Abidi, O.S. AbouZeid, N.L. Abraham, H. Abramowicz, and et al. Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at s=13tev with the atlas detector. *Physical Review D*, 96(7), Oct 2017. ISSN 2470-0029. doi:10.1103/physrevd.96.072002. URL <http://dx.doi.org/10.1103/PhysRevD.96.072002>.
- [94] ATLAS Collaboration. Identification and rejection of pile-up jets at high pseudorapidity with the atlas detector. *The European Physical Journal C*, 77(9), Sep 2017. ISSN 1434-6052. doi:10.1140/epjc/s10052-017-5081-5. URL <http://dx.doi.org/10.1140/epjc/s10052-017-5081-5>.
- [95] Optimisation of the ATLAS *b*-tagging performance for the 2016 LHC Run. Technical Report ATL-PHYS-PUB-2016-012, CERN, Geneva, Jun 2016. URL <https://cds.cern.ch/record/2160731>.

- [96] Simone Marzani, Gregory Soyez, and Michael Spannowsky. Looking inside jets. *Lecture Notes in Physics*, 2019. ISSN 1616-6361.  
doi:10.1007/978-3-030-15709-8. URL  
<http://dx.doi.org/10.1007/978-3-030-15709-8>.
- [97] Amanda Lynn Steinhebel. JetAttrs : HLT Jet Hypothesis for Jet Moments. Technical Report ATL-COM-DAQ-2019-001, CERN, Geneva, Dec 2018. URL  
<https://cds.cern.ch/record/2652981>.
- [98] Herjuno Nindhito, Steven Schramm, Will Kalderon, and Alex Martyniuk. Jet trigger qt recap, study of inclusive anti-qcd jet trigger in atlas experiment.  
[https://indico.cern.ch/event/762214/contributions/3163074/attachments/1735808/2807666/181017\\_JunoUpdate.pdf](https://indico.cern.ch/event/762214/contributions/3163074/attachments/1735808/2807666/181017_JunoUpdate.pdf).
- [99] Amanda Lynn Steinhebel. Jetattrs : Hlt jet hypothesis for jet moments. Technical Report ATL-COM-DAQ-2019-001, CERN, Geneva, Dec 2018. URL  
<https://cds.cern.ch/record/2652981>.
- [100] Grigore Tarna. Electron identification with the ATLAS detector. Technical Report ATL-PHYS-PROC-2017-173, CERN, Geneva, Sep 2017. URL  
<https://cds.cern.ch/record/2286383>.
- [101] Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data. Technical Report ATLAS-CONF-2016-024, CERN, Geneva, Jun 2016. URL  
<https://cds.cern.ch/record/2157687>.
- [102] G. Aad, B. Abbott, J. Abdallah, S. Abdel Khalek, O. Abdinov, R. Aben, B. Abi, M. Abolins, O. S. AbouZeid, and et al. Electron and photon energy calibration with the atlas detector using lhc run 1 data. *The European Physical Journal C*, 74(10), Oct 2014. ISSN 1434-6052.  
doi:10.1140/epjc/s10052-014-3071-4. URL  
<http://dx.doi.org/10.1140/epjc/s10052-014-3071-4>.
- [103] Electron and photon energy calibration with the ATLAS detector using data collected in 2015 at  $\sqrt{s} = 13$  TeV. Technical Report ATL-PHYS-PUB-2016-015, CERN, Geneva, Aug 2016. URL  
<https://cds.cern.ch/record/2203514>.
- [104] M. Aaboud, G. Aad, B. Abbott, O. Abdinov, B. Abeloos, D. K. Abhayasinghe, S. H. Abidi, O. S. AbouZeid, N. L. Abraham, and et al. Measurement of the photon identification efficiencies with the atlas detector using lhc run 2 data collected in 2015 and 2016. *The European Physical Journal C*, 79(3), Mar 2019. ISSN 1434-6052.  
doi:10.1140/epjc/s10052-019-6650-6. URL  
<http://dx.doi.org/10.1140/epjc/s10052-019-6650-6>.



- [105] G. Aad, B. Abbott, J. Abdallah, O. Abdinov, B. Abeloos, R. Aben, M. Abolins, O. S. AbouZeid, N. L. Abraham, and et al. Muon reconstruction performance of the atlas detector in protonproton collision data at  $\sqrt{s}=13$  tev. *The European Physical Journal C*, 76(5), May 2016. ISSN 1434-6052. doi:10.1140/epjc/s10052-016-4120-y. URL <http://dx.doi.org/10.1140/epjc/s10052-016-4120-y>.
- [106] Performance of missing transverse momentum reconstruction for the ATLAS detector in the first proton-proton collisions at  $\sqrt{s}=13$  TeV. Technical Report ATL-PHYS-PUB-2015-027, CERN, Geneva, Jul 2015. URL <https://cds.cern.ch/record/2037904>.
- [107] The ATLAS transverse-momentum trigger performance at the LHC in 2011. Technical Report ATLAS-CONF-2014-002, CERN, Geneva, Feb 2014. URL <https://cds.cern.ch/record/1647616>.
- [108] S Heinemeyer, editor. *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working Group*. CERN Yellow Reports: Monographs. CERN, Geneva, Jul 2013. doi:10.5170/CERN-2013-004. URL <https://cds.cern.ch/record/1559921>. Comments: 404 pages, 139 figures, to be submitted to CERN Report. Working Group web page: <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSections>.
- [109] Eilam Gross. LHC Statistics for Pedestrians. 2008. doi:10.5170/CERN-2008-001.205. URL <https://cds.cern.ch/record/1099994>.
- [110] Glen Cowan. Statistical methods for particle physics lecture 3: asymptotics; asimov data set. [https://www.pp.rhul.ac.uk/~cowan/stat/weizmann15/cowan\\_weizmann15\\_3.pdf](https://www.pp.rhul.ac.uk/~cowan/stat/weizmann15/cowan_weizmann15_3.pdf), 2015. Accessed: 202125-02.
- [111] G. Aad, B. Abbott, J. Abdallah, O. Abdinov, B. Abeloos, R. Aben, O. S. AbouZeid, N. L. Abraham, H. Abramowicz, and et al. Measurements of the higgs boson production and decay rates and constraints on its couplings from a combined atlas and cms analysis of the lh pp collision data at  $\sqrt{s}=7$  and 8 tev. *Journal of High Energy Physics*, 2016(8), Aug 2016. ISSN 1029-8479. doi:10.1007/jhep08(2016)045. URL [http://dx.doi.org/10.1007/JHEP08\(2016\)045](http://dx.doi.org/10.1007/JHEP08(2016)045).
- [112] Combination of searches for invisible Higgs boson decays with the ATLAS experiment. Technical Report ATLAS-CONF-2020-052, CERN, Geneva, Oct 2020. URL <http://cds.cern.ch/record/2743055>.

- [113] A combination of measurements of Higgs boson production and decay using up to  $139 \text{ fb}^{-1}$  of proton–proton collision data at  $\sqrt{s} = 13 \text{ TeV}$  collected with the ATLAS experiment. Technical Report ATLAS-CONF-2020-027, CERN, Geneva, Aug 2020. URL <http://cds.cern.ch/record/2725733>.
- [114] G. Aad, B. Abbott, J. Abdallah, O. Abdinov, R. Aben, M. Abolins, O. S. AbouZeid, H. Abramowicz, H. Abreu, and et al. Search for invisible decays of the higgs boson produced in association with a hadronically decaying vector boson in pp collisions at  $\sqrt{s} = 8 \text{ tev}$  with the atlas detector. *The European Physical Journal C*, 75(7), Jul 2015. ISSN 1434-6052. doi:10.1140/epjc/s10052-015-3551-1. URL <http://dx.doi.org/10.1140/epjc/s10052-015-3551-1>.
- [115] G. Aad, T. Abajyan, B. Abbott, J. Abdallah, S. Abdel Khalek, O. Abdinov, R. Aben, B. Abi, M. Abolins, O. S. AbouZeid, and et al. Search for invisible decays of a higgs boson produced in association with azboson in atlas. *Physical Review Letters*, 112(20), May 2014. ISSN 1079-7114. doi:10.1103/physrevlett.112.201802. URL <http://dx.doi.org/10.1103/PhysRevLett.112.201802>.
- [116] G. Aad, B. Abbott, J. Abdallah, O. Abdinov, R. Aben, M. Abolins, O. S. AbouZeid, H. Abramowicz, H. Abreu, and et al. Search for invisible decays of a higgs boson using vector-boson fusion in pp collisions at  $\sqrt{s} = 8 \text{ tev}$  with the atlas detector. *Journal of High Energy Physics*, 2016(1), Jan 2016. ISSN 1029-8479. doi:10.1007/jhep01(2016)172. URL [http://dx.doi.org/10.1007/JHEP01\(2016\)172](http://dx.doi.org/10.1007/JHEP01(2016)172).
- [117] Georges Aad et al. Constraints on new phenomena via Higgs boson couplings and invisible decays with the ATLAS detector. *JHEP*, 11:206, 2015. doi:10.1007/JHEP11(2015)206.
- [118] M. Aaboud, G. Aad, B. Abbott, O. Abdinov, B. Abeloos, D. K. Abhayasinghe, S. H. Abidi, O. S. AbouZeid, N. L. Abraham, and et al. Search for dark matter in events with a hadronically decaying vector boson and missing transverse momentum in pp collisions at  $\sqrt{s} = 13 \text{ tev}$  with the atlas detector. *Journal of High Energy Physics*, 2018(10), Oct 2018. ISSN 1029-8479. doi:10.1007/jhep10(2018)180. URL [http://dx.doi.org/10.1007/JHEP10\(2018\)180](http://dx.doi.org/10.1007/JHEP10(2018)180).

- [119] M. Aaboud, G. Aad, B. Abbott, O. Abdinov, B. Abeloos, S.H. Abidi, O.S. AbouZeid, N.L. Abraham, H. Abramowicz, H. Abreu, and et al. Search for an invisibly decaying higgs boson or dark matter candidates produced in association with a z boson in pp collisions at s=13 tev with the atlas detector. *Physics Letters B*, 776:318337, Jan 2018. ISSN 0370-2693. doi:10.1016/j.physletb.2017.11.049. URL <http://dx.doi.org/10.1016/j.physletb.2017.11.049>.
- [120] ATLAS Collaboration. Search for invisible higgs boson decays in vector boson fusion at s=13 tev with the atlas detector. *Physics Letters B*, 793:499519, Jun 2019. ISSN 0370-2693. doi:10.1016/j.physletb.2019.04.024. URL <http://dx.doi.org/10.1016/j.physletb.2019.04.024>.
- [121] ATLAS Collaboration. Combination of searches for invisible Higgs boson decays with the ATLAS experiment. *Phys. Rev. Lett.*, 122(23):231801, 2019. doi:10.1103/PhysRevLett.122.231801.
- [122] ATLAS Collaboration. Search for new phenomena in events with an energetic jet and missing transverse momentum in  $pp$  collisions at  $\sqrt{s} = 13$  tev with the atlas detector, 2021.
- [123] The ATLAS transverse-momentum trigger performance at the LHC in 2011. Technical Report ATLAS-CONF-2014-002, CERN, Geneva, Feb 2014. URL <https://cds.cern.ch/record/1647616>.
- [124] Jason Gallicchio and Matthew D. Schwartz. Quark and gluon tagging at the lhc. *Physical Review Letters*, 107(17), Oct 2011. ISSN 1079-7114. doi:10.1103/physrevlett.107.172001. URL <http://dx.doi.org/10.1103/PhysRevLett.107.172001>.
- [125] Christopher Frye, Andrew J. Larkoski, Jesse Thaler, and Kevin Zhou. Casimir meets poisson: improved quark/gluon discrimination with counting observables. *Journal of High Energy Physics*, 2017(9), Sep 2017. ISSN 1029-8479. doi:10.1007/jhep09(2017)083. URL [http://dx.doi.org/10.1007/JHEP09\(2017\)083](http://dx.doi.org/10.1007/JHEP09(2017)083).
- [126] Quark versus Gluon Jet Tagging Using Charged Particle Multiplicity with the ATLAS Detector. Technical Report ATL-PHYS-PUB-2017-009, CERN, Geneva, May 2017. URL <http://cds.cern.ch/record/2263679>.
- [127] Luminosity determination in  $pp$  collisions at  $\sqrt{s} = 13$  TeV using the ATLAS detector at the LHC. Technical Report ATLAS-CONF-2019-021, CERN, Geneva, Jun 2019. URL <https://cds.cern.ch/record/2677054>.

- [128] G. Avoni, M. Bruschi, G. Cabras, D. Caforio, N. Dehghanian, A. Floderus, B. Giacobbe, F. Giannuzzi, F. Giorgi, P. Grafstrm, V. Hedberg, F. Lasagni Manghi, S. Meneghini, J. Pinfold, E. Richards, C. Sbarra, N. Semprini Cesari, A. Sbrizzi, R. Soluk, G. Uccelli, S. Valentinetti, O. Viazlo, M. Villa, C. Vittori, R. Vuillermet, and A. Zoccoli. The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS. *Journal of Instrumentation*, 13(07):P07017–P07017, jul 2018. doi:10.1088/1748-0221/13/07/p07017. URL <https://doi.org/10.1088/1748-0221/13/07/p07017>.
- [129] Internal release 21 jet calibration twiki. <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JVTCalibration>, 2019. URL <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JVTCalibration>.
- [130] Leif Lnnblad and Stefan Prestel. Matching tree-level matrix elements with interleaved showers. *Journal of High Energy Physics*, 2012(3), Mar 2012. ISSN 1029-8479. doi:10.1007/jhep03(2012)019. URL [http://dx.doi.org/10.1007/JHEP03\(2012\)019](http://dx.doi.org/10.1007/JHEP03(2012)019).
- [131] Sm higgs production cross sections at 13 tev. <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageAt13TeV>, 2019.
- [132] Wouter Verkerke and David P. Kirkby. The RooFit toolkit for data modeling. *eConf*, C0303241:MOLT007, 2003.
- [133] Lorenzo Moneta, Kevin Belasco, Kyle Cranmer, Sven Kreiss, Alfio Lazzaro, Danilo Piparo, Gregory Schott, Wouter Verkerke, and Matthias Wolf. The roostats project, 2011.
- [134] Kyle Cranmer, George Lewis, Lorenzo Moneta, Akira Shibata, and Wouter Verkerke. HistFactory: A tool for creating statistical models for use with RooFit and RooStats. Technical Report CERN-OPEN-2012-016, New York U., New York, Jan 2012. URL <https://cds.cern.ch/record/1456844>.
- [135] Alexander L. Read. Presentation of search results: The CL(s) technique. *J. Phys. G*, 28:2693–2704, 2002. doi:10.1088/0954-3899/28/10/313.
- [136] Search for new phenomena in events with two opposite-charge leptons, jets and missing transverse momentum in  $pp$  collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector. Technical Report ATLAS-CONF-2020-046, CERN, Geneva, July 2020. URL <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2020-046/>.

- [137] ATLAS Collaboration. Search for a scalar partner of the top quark in the all-hadronic  $t\bar{t}$  plus missing transverse momentum final state at  $\sqrt{s}=13$  tev with the atlas detector, 2020.
- [138] ATLAS Collaboration. Search for new phenomena in events with an energetic jet and missing transverse momentum in  $pp$  collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector. Technical Report arXiv:2102.10874, CERN, Geneva, Feb 2021. URL <https://cds.cern.ch/record/2752675>.
- [139] Searches for electroweak production of two jets in association with a Higgs boson decaying fully or partially to invisible particles, including a final state photon using proton-proton collisions at 13 TeV with the ATLAS detector. Technical Report ATLAS-CONF-2021-004, CERN, Geneva, Mar 2021. URL <http://cds.cern.ch/record/2758212>.
- [140] European Strategy for Particle Physics. <https://europeanstrategy.cern/home>, .
- [141] European Strategy for Particle Physics. [https://home.cern/sites/home.web.cern.ch/files/2018-08/Strategy\\_Brochure.pdf](https://home.cern/sites/home.web.cern.ch/files/2018-08/Strategy_Brochure.pdf), .
- [142] 2020 Update of the European Strategy for Particle Physics. <https://home.cern/sites/home.web.cern.ch/files/2020-06/2020%20Update%20European%20Strategy.pdf>, .
- [143] About Particle Physics. <https://www.usparticlephysics.org/>.
- [144] Welcome to Snowmass. <https://snowmass21.org/>.
- [145] Building For Discovery: Strategic Plan for U.S. Particle Physics in the Global Context. [https://www.usparticlephysics.org/wp-content/uploads/2018/03/FINAL\\_P5\\_Report\\_053014.pdf](https://www.usparticlephysics.org/wp-content/uploads/2018/03/FINAL_P5_Report_053014.pdf).
- [146] Asian Committee for Future Accelerators. URL <http://www.acfa-forum.net/about/index.html>.
- [147] International Committee for Future Accelerators. URL <https://icfa.fnal.gov/>.
- [148] Project Schedule. <https://project-hl-lhc-industry.web.cern.ch/content/project-schedule>, .
- [149] A. Abada et al. HE-LHC: The High-Energy Large Hadron Collider: Future Circular Collider Conceptual Design Report Volume 4. *Eur. Phys. J. ST*, 228 (5):1109–1382, 2019. doi:10.1140/epjst/e2019-900088-6.

- [150] What is the ILC? <https://linearcollider.web.cern.ch/ILC.html>.
- [151] Andrea Dainese, Michelangelo Mangano, Andreas B Meyer, Aleandro Nisati, Gavin Salam, and Mika Anton Vesterinen. Report on the Physics at the HL-LHC, and Perspectives for the HE-LHC. Technical Report CERN-2019-007, Geneva, Switzerland, 2019. URL <https://cds.cern.ch/record/2703572>.
- [152] HiLumi in a Nutshell. <https://project-hl-lhc-industry.web.cern.ch/content/hilumi-nutshell>, .
- [153] Expected performance of the ATLAS detector at the High-Luminosity LHC. Technical Report ATL-PHYS-PUB-2019-005, CERN, Geneva, Jan 2019. URL <https://cds.cern.ch/record/2655304>.
- [154] Technical Design Report for the ATLAS Inner Tracker Strip Detector. Technical Report CERN-LHCC-2017-005. ATLAS-TDR-025, CERN, Geneva, Apr 2017. URL <https://cds.cern.ch/record/2257755>.
- [155] Technical Design Report for the ATLAS Inner Tracker Pixel Detector. Technical Report CERN-LHCC-2017-021. ATLAS-TDR-030, CERN, Geneva, Sep 2017. URL <https://cds.cern.ch/record/2285585>.
- [156] Monte Carlo Truth in the xAOD. [https://indico.cern.ch/event/472469/contributions/1982685/attachments/1222751/1789718/truth\\_tutorial.pdf](https://indico.cern.ch/event/472469/contributions/1982685/attachments/1222751/1789718/truth_tutorial.pdf).
- [157] M. Cepeda et al. Report from Working Group 2: Higgs Physics at the HL-LHC and HE-LHC. *CERN Yellow Rep. Monogr.*, 7:221–584, 2019. doi:10.23731/CYRM-2019-007.221.
- [158] Ties Behnke, James E. Brau, Philip N. Burrows, Juan Fuster, Michael Peskin, Marcel Stanitzki, Yasuhiro Sugimoto, Sakue Yamada, and Hitoshi Yamamoto. The international linear collider technical design report - volume 4: Detectors, 2013.
- [159] Howard Baer, Tim Barklow, Keisuke Fujii, Yuanning Gao, Andre Hoang, Shinya Kanemura, Jenny List, Heather E. Logan, Andrei Nomerotski, Maxim Perelstein, Michael E. Peskin, Roman Pschl, Jrgen Reuter, Sabine Riemann, Aurore Savoy-Navarro, Geraldine Servant, Tim M. P. Tait, and Jaehoon Yu. The international linear collider technical design report - volume 2: Physics, 2013.
- [160] ICFA applauds the 2020 Update of the European Strategy for Particle Physics. [https://icfa.fnal.gov/wp-content/uploads/ICFA\\_Statement\\_2020ESPPUpdate.pdf](https://icfa.fnal.gov/wp-content/uploads/ICFA_Statement_2020ESPPUpdate.pdf), August 2020.

- [161] Keisuke Fujii, Christophe Grojean, Michael E. Peskin, Tim Barklow, Yaunng Gao, Shinya Kanemura, Hyungdo Kim, Jenny List, Mihoko Nojiri, Maxim Perelstein, Roman Poeschl, Juergen Reuter, Frank Simon, Tomohiko Tanabe, James D. Wells, Jaehoon Yu, Junping Tian, Taikan Suehara, Marcel Vos, Graham Wilson, James Brau, and Hitoshi Murayama. Tests of the standard model at the international linear collider, 2019.
- [162] Amanda Steinhebel, Jim Brau, and Chris Potter.  $H \rightarrow$ invisible at the ilc with sid, 2021.
- [163] James E. Brau, John A. Jaros, and Hong Ma. Advances in calorimetry. *Annual Review of Nuclear and Particle Science*, 60(1):615–644, 2010. doi:10.1146/annurev.nucl.012809.104449. URL <https://doi.org/10.1146/annurev.nucl.012809.104449>.
- [164] ATLAS Collaboration. *Expected performance of the ATLAS experiment: detector, trigger and physics*. CERN, Geneva, 2009. URL <https://cds.cern.ch/record/1125884>.
- [165] Measurement prospects for VBF  $H \rightarrow WW^{(*)} \rightarrow e\nu\mu\nu$  production with  $3 \text{ ab}^{-1}$  of HL-LHC  $pp$ -collisions. Technical Report ATL-PHYS-PUB-2016-018, CERN, Geneva, Aug 2016. URL <https://cds.cern.ch/record/2209092>.
- [166] Amanda L Steinhebel and James Brau. Backscatter and Electromagnetic Shower Spreading in the SiD ECal. <https://pages.uoregon.edu/asteinhe/SiDNotes/backscatter/backscatter.pdf>, March 2018.
- [167] Amanda L Steinhebel and James Brau. SiD ECal Geometry. <https://pages.uoregon.edu/asteinhe/SiDNotes/geometry/geometry.pdf>, March 2018.
- [168] Amanda L Steinhebel and James Brau. Summary of Studies of SiD ECal Manual Calibration. <https://pages.uoregon.edu/asteinhe/SiDNotes/calibrationLeakageStudies/calibLeakage.pdf>, March 2018.
- [169] Amanda L Steinhebel and James Brau. Summary of Studies of the Response of a Prototype SiD ECal. <https://pages.uoregon.edu/asteinhe/SiDNotes/testBeamStudies/testBeam.pdf>, March 2018.
- [170] James Brau, Martin Breidenbach, Dietrich R. Freytag, Claus Kleinwort, Uwe Kraemer, Benjamin A. Reese, Sebastiaan Roelofs, Marcel Stanitzki, Amanda Steinhebel, Dimitra Tsionou, and Mengqing Wu. Lycoris – a large-area, high resolution beam telescope, 2021.