

A NOVEL SEARCH FOR EXOTIC DECAYS OF THE HIGGS
BOSON WITH THE ATLAS DETECTOR
AND ENHANCING THE PHYSICS POTENTIAL OF THE LARGE
HADRON COLLIDER AND ATOM INTERFEROMETERS WITH
NEW TECHNIQUES

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF PHYSICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Murtaza Safdari

August 2022



© 2022 by Murtaza Safdari. All Rights Reserved.
Re-distributed by Stanford University under license with the author.



This work is licensed under a Creative Commons Attribution-3.0 United States License.

<http://creativecommons.org/licenses/by/3.0/us/>

This dissertation is online at: <https://purl.stanford.edu/pp356hw3480>

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Ariel Schwartzman, Primary Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Patricia Burchat

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Lauren Tompkins

Approved for the Stanford University Committee on Graduate Studies.

Stacey F. Bent, Vice Provost for Graduate Education

This signature page was generated electronically upon submission of this dissertation in electronic format.

Abstract

Fundamental physics research aims to understand the theory of particle interactions, the Standard Model (SM) of particle physics being the current best theory of the electroweak and strong forces. Modern efforts seek to explain phenomena like the matter anti-matter asymmetry of the universe and the nature of dark matter using various experimental modalities such as terrestrial particle colliders like the Large Hadron Collider (LHC). The ATLAS detector on the LHC is conducting a diverse physics program of precision SM measurements and searches for Physics beyond the SM using deeply inelastic scattering products to study fundamental physics.

The original research presented here uses proton collision data from the ATLAS detector to search for an exotic decay mode of the Higgs boson coupling to a new light scalar field. Additionally, two research projects are presented to improve the performance of the ATLAS detector. The first introduces a new algorithm to improve the efficiency of locating interesting physics within saved events. The second improves the jet calibration procedure by enabling the use of gradient based regression with an innovative objective function along with a unified neural network based framework.

Additionally, a network of quantum sensors are in development to enhance the physics reach of modern detectors and expand the set of models of new physics that can be experimentally probed. One such technology is atomic gradiometer interferometric sensors, like the MAGIS-100 experiment, that utilize matter waves to search for ultralight bosonic dark matter. The research and development of a novel light field imaging device is presented here for the MAGIS-100 experiment, as part of a burgeoning collaboration between the high energy physics (HEP) and the atomic-molecular-optics (AMO) communities.

Dedicated to my mother's heroic efforts
getting me to school on time everyday.

Acknowledgements

Science is fundamentally a team sport. Likewise, the work presented in this thesis has been conducted with the assistance and guidance of many, who have helped shape my scientific approach and helped me mature as an academic researcher. In no uncertain terms, I am where I am thanks to the efforts of these remarkable individuals.

The most important of these is Prof Ariel Schwartzman, who has been my PhD advisor and has supervised me through all the projects discussed in this thesis. Ariel is a strong advocate for his students, and always has a wide array of interesting projects for his students to work on — The research environment he is able to create is truly exhilarating and intellectually stimulating. Working as his student, I found myself always able to turn to new projects that helped expand my horizons and pushed me outside my research comfort zone. I have also learned a great deal about scientific communication and effective presentation of results from Ariel, something that is so important in an academic career and yet is often overlooked as part of our university education. Under his guidance, I was able to work on several different sub-systems of the ATLAS detector and make meaningful contributions across the breadth of the experiment, and conduct a novel exotic Higgs decay search. Furthermore, it was due to Ariel that the novel HEP-AMO collaboration between SLAC and the MAGIS experiment came into being, and I was extremely fortunate to be at the start of this unique exchange and contribute to the cutting edge research that the SLAC MAGIS group has been engaged in ever since. It therefore goes without saying that I owe my academic achievements thus far to Ariel, and hope to do justice to his mentorship in the future as I continue my academic career.

I'd like to acknowledge my dissertation committee, chaired by Prof Paulo Somaini

and including Prof Lauren Tompkins, Prof Patricia Burchat, Prof Jason Hogan, and Ariel, and I'm grateful for my reading committee, comprising Ariel, Lauren, and Pat. Members of the committee have been greatly supportive of my research and I've enjoyed our conversations about extensions to several of the projects presented in this thesis as future work.

Dr Ben Nachman has been uniquely influential, and has advised several of the studies presented here. Ben has been an incredible collaborator and I have learned a lot from him about the various research topics as well as good research practices, particularly as they apply to machine learning applications in our research. Ben has also been fundamental in our search for the exotic Higgs decays, and has helped me learn and navigate the analysis procedures within large scientific collaborations like ATLAS. Dr Aviv Cukierman and Sanha Cheong, my fellow peers in the SLAC ATLAS group under Ariel, have also been critical to the projects I've conducted during my PhD. Aviv has also been a great role-model as the senior student, and has helped me understand how to be thorough and systematic as I conduct my research. Furthermore, Sanha has been a close collaborator on the SLAC MAGIS imaging projects, and his energy and enthusiasm has helped greatly advance several of our projects. I feel truly privileged to be part of such a wonderful PhD family, with collaborators that make conducting scientific inquiry pure joy.

Dr Valentina Cairo has had a huge influence on the vertexing project, and was crucial in getting the novel method incorporated within ATLAS. Her insightful guidance helped steer my research, and her advocacy helped me popularize the concepts within the wider collaboration. Valentina is also an exceptional supporter of early career scientists, and has been a strong mentor for me and helped me navigate my career decisions as I approached the end of my PhD.

Prof Jason Hogan of the MAGIS collaboration has been vital to the SLAC MAGIS effort, and it is with his help that we are able to conduct the various state-of-the-art imaging and reconstruction projects for the MAGIS experiment. His support has helped propel our research efforts and his willingness to explore new ideas has inspired creativity in our research output. Working on the novel MAGIS imaging concepts has been the highlight of my PhD research, and I hope to continue collaborating with

Jason and the MAGIS experiment in the future as we redefine the imaging paradigm in atom interferometer experiments.

The SLAC MAGIS team members, Ariel, Sanha, Dr Joseph Frisch, Dr Michael Kagan, Dr Sean Gasiorowski, Maxime Vandegar, have all been instrumental in setting up the new research group and developing novel imaging paradigms for the MAGIS experiment and beyond. The immense creativity of this collaboration is something all PhD students aspire to, and I have learned a great deal from the diverse expertise of our group members.

I would also like to acknowledge the exceptionally talented undergraduate student researchers that have worked with me on several of my PhD projects — Ivan Juarez-Reyes worked on the ATLAS low level multi-jet trigger study, and Will Michaels, Brion Ye, Patin Inkaew, and Jonah Ezekiel worked on the MAGIS experiment as part of the SLAC MAGIS group. These students went above and beyond their tasks and made tangible contributions to the various projects, and I constantly learn from their ingenuity and unwavering enthusiasm.

Several members of the wider ATLAS collaboration have enabled the execution and adoption of the techniques presented in this thesis. For the exotic Higgs search, the various conveners, sub-conveners, and the editorial board have been gracious in their support of the project and its timeline. The conveners and sub-conveners of the tracking and vertexing groups enabled the research into vertex selection algorithms and helped add that to the ATLAS code-base. Likewise, the conveners and sub-conveners of the jet and missing transverse energy group in ATLAS facilitated the jet calibration study, as well as the development of the di-gluon tagger in the context of the exotic Higgs search.

In addition to the above mentioned, senior members of the SLAC ATLAS group, including Prof Caterina Vernieri, Prof Su Dong, Dr Charles Young, have played a role in shaping my academic career and I greatly appreciate them for their guidance and mentorship. Caterina has in particular helped me chart a course for my academic career and I look forward to collaborating with her in the future. Dr Rainer Bartoldus and Dr Catrin Bernius have also assisted me in my initial investigation into topology aware low level multi-jet triggers for ATLAS. My fellow SLAC ATLAS PhD peers,

Nicole Hartman, Jannicke Pearkes, Rachel Smith, and Sanha have provided me the crucial camaraderie and support it takes to get through a graduate degree, and have helped keep me honest and clear my doubts through my various projects.

I'm also grateful for the administrative and computational support from the Stanford Physics department (Maria Frank, Nick Swan) and at SLAC (Adrienne Higashi, Wei Yang), which ensured me a smooth passage through my PhD with all the resources I needed to perform my research activities.

Prof Axel Kuhn, Prof Guy Wilkinson, Prof Philip Burrows, Dr Alan Merchant have all been vital in my academic development as tutors and project supervisors during my undergraduate degree at Oxford — It was under their excellent instruction that I developed my deep affinity for physics, and I shall be forever grateful for their support.

Finally I'd like to acknowledge my friends and family, who have sustained me all the way through my academic journey so far. Eileen Rivera has been a veritable bedrock through these years at Stanford and I hope we have many more. I've spent many an evening chatting with Danyal Mohaddes and Emeric Boigne about everything under the moon, and despite being scattered across the world now I hope we can continue this in some fashion. Additionally I'd like to acknowledge all the friends I've met at Stanford and earlier in my career at Oxford, who have helped me grow into my truest self and inspire me to be better everyday. I'm thankful for my brother, Mustafa Safdari, who had the foresight to encourage me to apply to Stanford for graduate school, and has been a strong supporter of my career decisions throughout my life. And of course, I am eternally grateful for the sacrifices that my parents, Nishat and Shabbir Husain Safdari, have made to get me to where I am today. Their unconditional love, support, and generosity has given me the strength to endure the toughest of challenges and inspires optimism and joy in me everyday. I also hope to one day fully explain the details and intricacies of my academic research to them, and I shall consider my job as a physicist incomplete till I have achieved this goal.

Contents

Abstract	iv
Acknowledgements	vi
1 Introduction	1
1.1 Units and Coordinate Systems	4
2 The Standard Model of Particle Physics	6
2.1 Introduction	6
2.2 Lagrange Formalism	8
2.3 Electroweak Interactions	10
2.3.1 Coupling to Fermions	14
2.3.2 Leptons	14
2.3.3 Quarks	18
2.4 Quantum Chromodynamics	22
2.5 Summary of Particles and Fields	26
2.6 Limitations and Exploring Physics Beyond the Standard Model . . .	30
3 The Large Hadron Collider	36
3.1 Introduction	36
3.2 The Accelerators	38
3.3 Collision Statistics	42
3.4 Past Campaigns and the High-Luminosity LHC	45
4 The ATLAS Detector	48
4.1 Introduction	48

4.2	Inner Detector	51
4.3	Calorimeters	54
4.4	Muon System	58
4.5	Triggers	61
4.6	Object Reconstruction	64
4.6.1	Track Reconstruction	64
4.6.2	Vertexing	68
4.6.3	Calorimeter Cell Clusters	72
4.6.4	Particle Flow Objects	73
4.6.5	Jets	74
4.6.5.1	Introduction and Definition	74
4.6.5.2	Calibration	77
4.6.5.2.1	Numerical Inversion	78
4.6.5.2.2	Small-R Jet Calibration	79
4.6.5.2.3	Large-R Jet Calibration	83
4.6.5.3	Pile-up Tagging	86
4.6.6	Electrons and Photons	87
4.6.7	Taus	88
4.6.8	Muons	89
4.6.9	Missing Transverse Momentum	90
4.7	Simulations	91
5	A Search for Exotic Higgs Boson Decays to Four Gluons Using the ATLAS Detector	94
5.1	Introduction	95
5.2	Event Samples	97
5.2.1	Data	97
5.2.2	Simulation	97
5.3	Event and object selection	99
5.3.1	Event Selection	100
5.3.1.1	Event requirements	100
5.3.1.2	Trigger requirements	101

5.3.1.3	Dimuon system requirements	102
5.3.2	Corrections to Monte Carlo samples	102
5.3.2.1	Pileup reweighting	102
5.3.2.2	Trigger scale factor	103
5.3.2.3	Muon scale factors	103
5.4	Analysis Strategy	104
5.4.1	Overview	104
5.4.2	Di-gluon tagging	106
5.4.3	Machine Learning for $H \rightarrow a\bar{a} \rightarrow gggg$	111
5.4.3.1	Architecture	112
5.4.3.2	Reweighter training	114
5.4.3.3	Tagger training	114
5.4.4	Signal tagger performance	119
5.4.4.1	$a \rightarrow g\bar{g}$ jet level tagger	119
5.4.4.2	$H \rightarrow a\bar{a} \rightarrow gggg$ event level tagger	122
5.5	Systematic Uncertainties	131
5.5.1	Track uncertainties	131
5.5.2	Pileup reweighting uncertainties	132
5.5.3	Muon uncertainties	133
5.5.3.1	Muon efficiency uncertainties	133
5.5.3.2	Muon calibration uncertainties	133
5.5.4	Modeling uncertainties	133
5.6	Results	134
5.6.1	Expected Limits	134
5.6.2	Blinded Analysis and Next Steps	136
5.7	Conclusion	137
6	Novel Hard-Scatter Primary Vertex Selection in the ATLAS	
	Experiment	138
6.1	Introduction	139
6.2	Datasets and Event Selection	143
6.3	Current Approach and its Limitations	145

6.4	New Technique	147
6.5	Experimental Challenges with the ATLAS ID geometry and Future Improvements	150
6.6	Conclusions	154
7	Calibration of Jets Assisted with Neural Networks in the ATLAS Experiment	155
7.1	Introduction	157
7.2	Parameterizing Detector Response Using Neural Networks	160
7.2.1	Mean	161
7.2.2	Median	162
7.2.3	Quantile	163
7.2.4	Mode	164
7.3	Simultaneous Jet Energy and Mass Calibrations with Neural Networks	169
7.3.1	Event Simulations	169
7.3.2	Methods	170
7.3.3	Results	172
7.4	Conclusions	175
8	Light Field Imaging for Cold Atom Clouds in the Matter-wave Atomic Gradiometer Interferometric Sensor	179
8.1	Introduction	180
8.2	Imaging requirements in atom interferometers	185
8.3	Optical design	186
8.3.1	Conventional imaging considerations and limitations	186
8.3.2	Light field imaging concept for atom interferometers	188
8.3.2.1	Geometry of a single folded view	188
8.3.2.2	Accommodating multiple views – Capturing the light field	190
8.3.2.3	Related work – Light field imaging	193
8.3.2.4	Implications for cold atom imaging	193
8.4	Experiment	195

8.4.1	Demonstrator Design	195
8.4.2	Imaging with Demonstrator	199
8.5	Simulation and reconstruction	200
8.6	Results	202
8.6.1	Simulation results	202
8.6.2	Captured data	202
8.6.3	Experimental results	207
8.6.4	Impact of misalignment and 3D printing resolution	214
8.7	Conclusions	216
9	Concluding Remarks	217
A	A Search for Exotic Higgs Boson Decays to Four Gluons Using the ATLAS Detector: Appendix	220
A.1	Signal Validation Plots	220
A.2	Reweighting Performance	223
A.2.1	Pythia Z + Jets Samples	223
A.2.2	Sherpa Z + Jets Samples	229
A.3	Inspecting closure and relative uncertainty on closures: Scanning DisCo λ Values	235
A.4	Statistical Analysis	239
A.5	Expected Limits on Signal Strength: Scanning DisCo λ Values	242
A.6	Impact and Pull Plots From the Fits During Limit Setting	244
B	Light Field Imaging for Cold Atom Clouds in the Matter-wave Atomic Gradiometer Interferometric Sensor: Appendix	248
B.1	Entocentric imaging - A geometric perspective	248
B.2	Normal angles for a folded view	251
B.3	Field of view for a folded view	254
B.4	Generated Image Grids	256
	Bibliography	258

List of Tables

2.1	Charges of the Standard Model lepton fields under U(1) hypercharge Y, SU(2) weak isospin T_3 , and the electromagnetic charge Q, related by $Q = T_3 + \frac{Y}{2}$	15
2.2	Lepton masses [1].	16
2.3	Charges of the Standard Model quark fields under U(1) hypercharge Y, SU(2) weak isospin T_3 , and the electromagnetic charge Q, related by $Q = T_3 + \frac{Y}{2}$	18
2.4	Quark masses [1].	19
5.1	The integrated luminosity of the data collected each year of the <i>Run 2</i> data-taking period. The luminosity shown here corresponds to the data that clears the ATLAS data quality checks [2].	97
5.2	Overview of the applied event selection.	101
A.1	Pythia $Z + \text{QCD}$ jet events non-closure, along with the relative uncertainty on the Sherpa and Pythia $Z + \text{QCD}$ jet events for the selected working points, where the working point selection procedure is described in Section 5.4.4.2. Also shows the Pythia-Sherpa $Z + \text{QCD}$ jet events cross generator non-closure, along with the signal efficiency / square root of the background efficiency.	235

List of Figures

2.1	Particle content (fundamental fermions and gauge bosons) of the Standard Model of particle physics. Figure sourced from CERN [3]. . . .	7
2.2	Summary of the Standard Model fermions and gauge bosons before and after the spontaneous symmetry breaking via the Higgs mechanism. Note that this figure uses the <i>alternate convention</i> for the hypercharge, wherein the electric charge is given as the simple sum of the weak isospin and the hypercharge — The value of the hypercharge shown in the figure are half the value shown in the corresponding sections of this thesis. Figure sourced from [4].	13
2.3	Strong coupling constant α_s measured as a function of probe energy scale Q . The experimentally determined values are compared with theoretical predictions that use the reference energy scale of the mass of the Z boson. Figure sourced from [5].	24
2.4	Summary of the Standard Model particles. For each particle, the corresponding mass, electric charge, and spin is shown. Figure sourced from [4].	26
2.5	Interaction vertex diagrams allowed in the Standard Model. Figure sourced from [4].	27
2.6	Summary of Standard Model cross-section measurements done by the ATLAS experiment at the LHC, compared to the Standard Model predictions. The branching factor corrections have been applied to the measurements for the respective theoretical expectation. Figure sourced from [6].	29

2.7	Limits on the branching ratio of exotic Higgs decays $h \rightarrow ss \rightarrow XYY$ via the coupling to new light (pseudo)scalar particles s , for a variety of final states searched with the ATLAS and CMS experiments at the LHC. The dashed horizontal line corresponds to the upper limit on the exotic Higgs branching ratio from global fits to the various Higgs properties. Figure sourced from [7].	31
2.8	Summary plot from various experiments of (a) the ultralight scalar dark matter coupling to the electron (d_{m_e}), and (b) the ultralight vector dark matter gauge coupling (g_{B-L}). Figure sourced from [8]. Of note here is the projected sensitivity of the MAGIS-100 experiment and its upgraded configuration as well as the MAGIS-km experiment [9].	34
3.1	Aerial photo of Geneva and nearby the France–Switzerland border area with an overlay graphic showing the LHC and the locations of the four primary experiments — ATLAS, CMS, LHCb, ALICE. Image sourced from Ref. [10], Maximilien Brice (CERN).	37
3.2	Layout of the CERN accelerator complex, August 2018. Image sourced from CERN Ref. [11].	39
3.3	Phase-space representation of a bunch in its stationary bucket along with its RF potential function (in red and labelled V). The bucket size along the Energy-Time axes are labelled as ΔE and Δt , and the bunch size along the two axes are labelled as ΔE_m and Δt_m . Figure sourced from Ref. [12].	40
3.4	Cross-sections for pp interactions as a function of the center-of-mass energy ($\sqrt{s}$). Vertical lines corresponding to the various center-of-mass energies at Tevatron [13] and the LHC are drawn. Figure sourced from [14].	43
3.5	(a) Luminosity vs μ during Run 1. (b) Integrated luminosity vs time during Run 1. Figure sourced from [15].	45
3.6	(a) Luminosity vs μ during Run 2. (b) Integrated luminosity vs time during Run 2. Figure sourced from [16].	46

4.1	A computer generated view of the ATLAS detector, with the overall dimensions and detector subsystems labeled. Figures of people are drawn on the cavern floor and between the muon wheels for scale. Figure sourced from CERN [17].	49
4.2	Diagrammatic overview of the interaction of various types of particles with the ATLAS subsystems. Dashed lines correspond to no interactions, and solid lines correspond to small or insignificant energy loss. Figure sourced from CERN [18].	50
4.3	(a) Computer generated cross-sectional view of the ATLAS ID. Image sourced from CERN [19]. (b) Single quadrant view of the layout of the ID geometry in Run 2. The bottom panel shows a zoomed in view of the pixel region of the ID. The IBL, shown in red in the bottom left panel, was added for Run 2. The origin of the horizontal axis is the nominal interaction point, and the vertical axis is the radius measured from the origin. Figure sourced from [20]	53
4.4	Computer generated image of the ATLAS Calorimeter. Image sourced from CERN [21]	55
4.5	Computer generated image of the ATLAS muon system. Image sourced from CERN [22]	58
4.6	The ATLAS trigger and data acquisition system in Run 2. Trigger and read-out related components are being shown in this figure. Figure sourced from [2]	62
4.7	Transverse cross-section of detectors showing the tracks of charged particles in a 2 T axial field. Figure on the left has various sections of the ATLAS ID barrel system (Section 4.2) shaded in different colours for illustrative purposes. Tracks with transverse momenta spanning two orders of magnitude are shown to visualize the effect on the curvature of the tracks. Figure sourced from [23].	66
4.8	Display of a pp event that ATLAS recorded in 2015. Hits in the various layers of the ID are shown along with the reconstructed tracks. Figure sourced from CERN [24].	67

4.9	(a) Vertex reconstruction efficiency versus number of associated tracks to the vertex, measured in low μ data and Monte Carlo (MC) simulations. (b) Histogram of Δz values computed over all pairs of vertices in same event, measured in high μ data and MC. Figures sourced from [25].	69
4.10	(a) Average number of reconstructed vertices versus the average number of $p\bar{p}$ interactions (μ), measured in Monte Carlo (MC) simulations. Note here the AMVF performance has been shown for the various types of reconstructed vertices, though only the comparison between AMVF and IVF inclusive efficiency is relevant here. (b) Histogram of Δz values computed over all pairs of vertices in same event, measured in MC. Figures sourced from [26].	70
4.11	The distribution of $\sum p_T^2$ from primary vertices, normalised to the same area. Various simulated hard-scatter processes are compared with a minimum bias sample from Run 1. Figure sourced from [27].	71
4.12	Particle flow algorithm flow chart. The algorithm starts with track selection and concludes when the energy of the selected tracks has been removed from the calorimeter. The algorithm outputs topo-clusters that have both been modified and unmodified by the action of the algorithm, along with the charged particle tracks. Figure sourced from [28].	73
4.13	ATLAS simulated event of a di-jet interaction (PYTHIA8). (a) The topo-clusters used as seeds in the y - ϕ plane, bins of size 0.04×0.03 . (b) Active area of the anti- k_t ($R = 1.0$) jets built using the topo-clusters. The colors indicate the p_T of the jets, and jets below 20 GeV are omitted. High energy hard-scatter jets (in red) are reconstructed with circular cachement areas, whereas the softer jets from the pile-up in the event have irregular boundaries. Figures sourced from [29]. . .	76
4.14	Schematic of the small- R jet calibration procedure. Figure sourced from [30].	79

4.15	Dependence of jet p_T on (A) the number of reconstructed primary vertices NPV and (b) the mean number of interactions per bunch crossing μ , before any correction, after the area subtraction, and after the residual linear correction is applied. The functions are plotted against the detector-level $ \eta $ of the jets, η^{det} . Figures sourced from [30].	80
4.16	Jet response for small-R jets, to be corrected with absolute MC calibration, and global sequential calibration. (a) Energy response versus E_T^{true} in bins of η^{det} , for absolute MC calibration. (a) Energy response versus η^{det} in bins of E_T^{true} , for absolute MC calibration. (c) p_T response versus the number of tracks associated with the jet, for GSC. (d) p_T response versus the track width, for GSC. Figures sourced from [30].	82
4.17	Schematic of the large-R jet calibration procedure. Figure sourced from [31].	84
4.18	Energy and mass responses prior to the large-R jet calibration. (a) Energy response versus η^{det} in bins of E_T^{true} . (b) Mass response versus η^{det} in bins of p_T^{true} , for jets with $m^{\text{true}} \approx 80$ GeV. Figures sourced from [31].	85
5.1	Invariant mass of the two leading jets in the event, m_{jj} , using reconstructed jets (left) and truth (particle-level) jets (right). Figures sourced from [32].	106
5.2	Number of tracks in the event associated to a given jet ($\Delta R < 0.6$), for signal $a \rightarrow gg$ jets (left), background QCD quark / gluon jets (right). Different p_T bins are shown from top to bottom. Figures sourced from [32].	107
5.3	ΔR between the two leading tracks within jets ($\Delta R < 0.6$), for signal $a \rightarrow gg$ jets (left), background QCD quark/gluon jets (right). Different p_T bins are shown from top to bottom. Figures sourced from [32].	109

5.4	ECF(2,1) evaluated using tracks in the event associated to a given jet ($\Delta R < 0.6$), for signal $a \rightarrow gg$ jets (left), background QCD quark / gluon jets (right). Different p_T bins are shown from top to bottom. Figures sourced from [32].	110
5.5	The jet substructure observable p_TD evaluated using tracks in the event associated to a given jet ($\Delta R < 0.6$), for signal $a \rightarrow gg$ jets (left), background QCD quark / gluon jets (right). p_TD varies slowly with p_T so only the nominal p_T bin is shown. Figures sourced from [32].	111
5.6	Architecture of the full event level tagger. The jet level tagger embeds each of the two leading jets from an event into a 2-D latent space, and uses the same set of weights to embed both jets. The time distributed layers embed each track into a high dimensional space using the same weights for each track. The Addition_Over_Tracks layer adds up the latent vectors across all the tracks, which results in a permutation invariant dimensional reduction across the tracks. Event level inputs and the final latent vectors from the two leading jets are used to derive the final output score for every event using fully connected layers. Figure sourced from [32].	113
5.7	Distributions for data and $Z + \text{QCD jets}$ Pythia MC background before and after variable- and high-dimensional reweighting. Events with m_{jj} inside the signal region contained within the two vertical lines in (a) were excluded during training of the reweighting classifier. Only the lowest p_T bin is shown in (b) that displays the largest correction from the reweighting scheme. Figures sourced from [32].	115
5.8	An illustration of the ABCD plane. Figure sourced from [33].	116
5.9	m_{jj} distribution for the signal $H \rightarrow aa \rightarrow gggg$ events, with $\mu_{m_{jj} m_a}$ (the average m_{jj} in the range 50–200 GeV/ c^2) shown as dashed lines of the corresponding colors.	116

5.10	Receiver Operating Characteristic (ROC) curves of the jet classifier derived from the jet level tagger module of the full signal - background classifier from Figure 5.6. The background jets are from the same Z +jets sample used to train the tagger. QCD gluon jet rejection (right) is better than QCD quark jet rejection (left), with performance degrading as the scalar mass increases due to worsening collimation and increasing number of tracks in the di-gluon jets. Here the ROC curves for two jet p_T bins is shown (top - bottom). The lines labeled ‘Random’ correspond to di-gluon efficiency = $1/\text{QCD rejection}$. Figure sourced from [32].	120
5.11	Ratio of the QCD gluon jet rejection to the quark jet rejection as a function of the di-gluon selection efficiency, for the lowest p_T bin. Figure sourced from [32]	121
5.12	Receiver Operating Characteristic (ROC) curves derived using ECF(2,1) as the discriminating variable (dotted line) versus ROC curves derived from the jet level tagger module of the full signal - background classifier (solid line). Here the QCD gluon jet rejection (right) and QCD quark jet rejection (left) are shown for the lowest p_T bin from Figure 5.10 for three representative mass points. Note that these curves show that the rejection achieved using ECF(2,1) is lower than the rejection using the jet classifier, particularly for the lower signal masses. Figure sourced from [32]	121
5.13	Receiver Operating Characteristic (ROC) curves of the full signal - background classifier from Figure 5.6. (top left) DisCo weight $\lambda = 0.02$, (top right) DisCo weight $\lambda = 10$, (bottom left) DisCo weight $\lambda = 50$, (bottom right) DisCo weight $\lambda = 125$. The lines labeled ‘Random’ correspond to the curve signal efficiency = $1/\text{background rejection}$	123

- 5.14 (Left) ABCD plane for Data where the region of tagger output > 0.9 and m_{jj} axis > 0.7 is masked to blind the signal region, (right) ABCD plane for Pythia $Z + \text{QCD jets}$ events. The Z -axis color scale is logarithmic and shows the normalized distribution of events in each bin. The scalar mass m_a goes from $0.5 \text{ GeV}/c^2$ to $2 \text{ GeV}/c^2$ from top to bottom as indicated on the plots. The axes of the ABCD plane are the neural network output and $1 - |1 - (m_{jj}/\mu_{m_{jj}|m_a})|$, where $\mu_{m_{jj}|m_a}$ is the average m_{jj} for the signal with mass m_a in the m_{jj} range $[50, 200] \text{ GeV}/c^2$ 124
- 5.15 (Left) ABCD plane for Data where the region of tagger output > 0.9 and m_{jj} axis > 0.7 is masked to blind the signal region, (right) ABCD plane for Pythia $Z + \text{QCD jets}$ events. The Z -axis color scale is logarithmic and shows the normalized distribution of events in each bin. The scalar mass m_a goes from $4 \text{ GeV}/c^2$ to $16 \text{ GeV}/c^2$ from top to bottom as indicated on the plots. The axes of the ABCD plane are the neural network output and $1 - |1 - (m_{jj}/\mu_{m_{jj}|m_a})|$, where $\mu_{m_{jj}|m_a}$ is the average m_{jj} for the signal with mass m_a in the m_{jj} range $[50, 200] \text{ GeV}/c^2$ 125
- 5.16 (Left) Signal efficiency divided by the square root of the background (Pythia $Z + \text{QCD jets}$ events) efficiency for a set of pairs of cuts along the ABCD axes, (middle) background (Pythia $Z + \text{QCD jets}$) closure relative uncertainty for a set of pairs of cuts along the ABCD axes, (right) background (Pythia $Z + \text{QCD jets}$) non-closure for a set of pairs of cuts along the ABCD axes. The scalar mass m_a goes from $0.5 \text{ GeV}/c^2$ to $2 \text{ GeV}/c^2$ from top to bottom as indicated on the plots. The closure (right) is limited to a range of $[40, 160]\%$, and the relative uncertainty on closures (middle) is limited to a range of $[0, 50]\%$; All values outside of these range are colored by the extreme colors on either side of the ranges. 127

- 5.17 (Left) Signal efficiency divided by the square root of the background (Pythia $Z + \text{QCD jets}$ events) efficiency for a set of pairs of cuts along the ABCD axes, (middle) background (Pythia $Z + \text{QCD jets}$) closure relative uncertainty for a set of pairs of cuts along the ABCD axes, (right) background (Pythia $Z + \text{QCD jets}$) non-closure for a set of pairs of cuts along the ABCD axes. The scalar mass m_a goes from 4 GeV/c² to 16 GeV/c² from top to bottom as indicated on the plots. The closure (right) is limited to a range of [40, 160]%, and the relative uncertainty on closures (middle) is limited to a range of [0, 50]%; All values outside of these range are colored by the extreme colors on either side of the ranges. 128
- 5.18 (Left) Signal efficiency divided by the square root of the background (Pythia $Z + \text{QCD jets}$ events) efficiency for a set of pairs of cuts along the ABCD axes, (middle) background (Sherpa $Z + \text{QCD jets}$) closure relative uncertainty for a set of pairs of cuts along the ABCD axes, (right) difference of the closures in the ABCD plane between Pythia and Sherpa $Z + \text{QCD jets}$ events for a set of pairs of cuts along the ABCD axes. The scalar mass m_a goes from 0.5 GeV/c² to 2 GeV/c² from top to bottom as indicated on the plots. The relative uncertainty on closures (middle) is limited to a range of [0, 50]%, and the difference in closures (right) is limited to a range of [-40, 40]%; All values outside of these range are colored by the extreme colors on either side of the ranges. 129

5.19	(Left) Signal efficiency divided by the square root of the background (Pythia $Z + \text{QCD jets}$ events) efficiency for a set of pairs of cuts along the ABCD axes, (middle) background (Sherpa $Z + \text{QCD jets}$) closure relative uncertainty for a set of pairs of cuts along the ABCD axes, (right) difference of the closures in the ABCD plane between Pythia and Sherpa $Z + \text{QCD jets}$ events for a set of pairs of cuts along the ABCD axes. The scalar mass m_a goes from 4 GeV/c ² to 16 GeV/c ² from top to bottom as indicated on the plots. The relative uncertainty on closures (middle) is limited to a range of [0, 50]%, and the difference in closures (right) is limited to a range of [-40, 40]%; All values outside of these range are colored by the extreme colors on either side of the ranges.	130
5.20	Expected upper limits on $\text{Br}(H \rightarrow aa) \times \text{Br}(a \rightarrow gg)^2$ across the range of signal a masses tested. Note here that the <i>Stats Only</i> line shows the expected limits when no sources of uncertainty are considered, fitting only to the Poisson counting experiment with the expected signal and background counts.	136
6.1	Hard-scatter vertex reconstruction and selection efficiency versus the average number of interactions μ . Figure sourced from [27].	139
6.2	Feynman diagram for Vector Boson Fusion initiated $H \rightarrow ZZ^* \rightarrow 4\nu$. Quarks radiated from the vector bosons form the characteristic two forward VBF jets. The tracks inside the VBF jets are the crucial handles that are used to identify the hard-scatter vertex due to the neutrinos in the final state not interacting with the detector.	140
6.3	Plot showing a single quadrant layout of the ITk. Strip detector active elements are shown in blue, and pixel detector active elements are shown in red. The origin of the horizontal axis is the nominal interaction point, and the vertical axis is the radius measured from the origin. Figure sourced from [34].	142

6.4	Simulated event display of an HL-LHC event in the ATLAS ITk detector. The event has an average number of collisions per bunch crossing of $\mu = 200$. The inset on the bottom left shows a longitudinal cross section where the different pp interactions can be seen as the origin of the corresponding tracks. Figure sourced from CERN [35].	142
6.5	Vertex selection efficiency for VBF Higgs invisible events versus the average local density of vertices at the primary vertex in the event. The default sumPT2 selection method (square) is compared with the new sumPTw selection (circle). Error bars are statistical errors. Best fit linear functions are drawn with 1-sigma error bands to visualize the error on the slope with respect to the pile-up density. (a) LHC Run 3 conditions, (b) HL-LHC Run 4 conditions. Note the higher efficiencies at Run 4 than Run 3 are due to the increased tracking acceptance of the ITk which reconstructs tracks inside forward hard-scatter jets. Figure sourced from [36].	146
6.6	$r - z$ event display plots from a single VBF Higgs invisible MC simulated Run 4 event. The top pane is centered (red ellipse) on the pile-up (PU) vertex incorrectly chosen by sumPT2, the bottom pane is centered on the reconstructed hard-scatter (HS) vertex chosen by sumPTw. All jets with $p_T > 30$ GeV are drawn as semi-transparent bars on top of the vertex of interest. In addition to the reconstructed vertices shown on the <i>Reco</i> axis, the true HS and PU vertex locations are also shown on the <i>Truth</i> axis on the lower half of each pane. Figure sourced from [37].	149

6.7	Vertex selection efficiency for VBF Higgs invisible events in Run 3 conditions, versus the average local density of vertices at the primary vertex in the event. (a) The sumPTw selection method on all events (circle) is compared with the sumPTw selection on high confidence events (cross). (b) The default sumPT2 selection (star) is compared with the sumPTw selection (cross) on high confidence events. Error bars are statistical errors. Best fit linear functions are drawn with 1-sigma error bands to visualize the error on the slope with respect to the pile-up density. Figure sourced from [36].	151
6.8	$r - z$ event display plots from a single VBF Higgs invisible MC simulated Run 4 event. The top pane is centered (red ellipse) on the pile-up (PU) vertex incorrectly chosen by sumPTw, the bottom pane is centered on the correct reconstructed hard-scatter (HS) vertex. All jets with $p_T > 30$ GeV are drawn as semi-transparent bars on top of the vertex of interest. In addition to the reconstructed vertices shown on the <i>Reco</i> axis, the true HS and PU vertex locations are also shown on the <i>Truth</i> axis on the lower half of each pane. Figure sourced from [37].	153
7.1	Schematic of the two steps involved in GenNI, also applicable to standard numerical inversion. The first step involves learning the detector response distribution with the first neural network L. This parameterized response is then inverted in the second step using the second neural network C. The inverse of the response learned by the second neural network undoes the action of the experiment (which is typified by the response distribution), hence the second step is labelled as the inversion of the experiment. This second neural network C is then the required calibration function that corrects the scale of the reconstructed observable X^{reco} to the true scale.	159

7.2	Loss landscapes for the MSE, MAE, and GK functions evaluated as the function value $\mathbf{v}_0 = f(\mathbf{x}_0)$ is swept through the range of the distribution $p(\mathbf{y} \mathbf{x}_0)$ for each of the functions, for the (a) Gamma Distribution with $k = 2$ and $\theta = 1$, (b) Crystal Ball Distribution with $\beta = 1$ and $m = 2$. The underlying distribution $p(\mathbf{y} \mathbf{x}_0)$ is shown with gray dots, and the loss landscapes for the MSE, MAE, and GK functions are offset subtracted to help visualize the minima. Additionally the true mean, median, and mode of the underlying distributions are plotted with vertical dashed lines. Figures sourced from [38].	166
7.3	Loss landscapes for the LGK loss function evaluated as the function value $\mathbf{v}_0 = f(\mathbf{x}_0)$ is swept through the range of the distribution $p(\mathbf{y} \mathbf{x}_0)$ for the Gamma Distribution with $k = 2$ and $\theta = 1$. Three values of alpha are used and each corresponding loss landscape, along with the corresponding minimum value, are shown. The mode of the Gamma distribution with $k = 2$ and $\theta = 1$ is $\mathbf{y} = 1$, while the median is $\mathbf{y} \approx 1.678$. Figure sourced from [38].	168
7.4	A comparison of the mode of the energy response as a function of truth energy for GenNI trained using MSE (top) and LGK (bottom) in bins of the truth jet mass using the nominal PYTHIA 8 MC sample. The left plots use GenNI sequentially while the right plots use it simultaneously for mass and energy. Data points off the scale of the plots may have their error bars within the visualization range, which may appear as marker-less vertical lines on the plot. Figure sourced from [39].	173
7.5	Energy (left) and mass (right) responses for the nominal PYTHIA 8 MC sample. From top to bottom, each row represents: no calibration, current ATLAS JES+JMS, and GenNI calibration with LGK loss. Data points off the scale of the plots may have their error bars within the visualization range, which may appear as marker-less vertical lines on the plot. Note that the top row has a different $\mathbf{y}$ -axis scale than the lower two rows. Figure sourced from [39].	174

7.6	Energy (left) and mass (right) responses for the SHERPA MC sample. The top row shows the uncalibrated energy and mass responses of the SHERPA MC sample. The middle and the bottom rows show the response differences between PYTHIA and SHERPA (PYTHIA 8 – SHERPA 2) samples after current ATLAS JES+JMS and GenNI calibration with LGK loss, respectively. Data points off the scale of the plots may have their error bars within the visualization range, which may appear as marker-less vertical lines on the plot. Figure sourced from [39].	176
7.7	Energy (left) and mass (right) responses for the HERWIG MC sample. The top row shows the uncalibrated energy and mass responses of the HERWIG MC sample. The middle and the bottom rows show the response differences between PYTHIA and HERWIG (PYTHIA 8 – HERWIG 7) samples after current ATLAS JES+JMS and GenNI calibration with LGK loss, respectively. Data points off the scale of the plots may have their error bars within the visualization range, which may appear as marker-less vertical lines on the plot. Figure sourced from [39].	177
8.1	Detector layout of the MAGIS-100 experiment under construction at the Fermi National Accelerator Laboratory. (a) Computer generated model of the full experiment with the 100 m vertical vacuum chamber. (b) Computer generated view of the laser lab room. (c) Computer generated model of one of the 17 the modular sections that constitute the full experiment, which contains a section of the pipe, vacuum pumps, magnetic shielding, etc. Figure sourced from [9].	181

8.2 (a) Simplified clock atom interferometer space-time diagram. The cold atom cloud manipulated by a series of laser pulses (wavy lines) after the initial launch of the cold atoms, throughout the duration of their free-fall under gravity. Blue (red) lines show the cold atom cloud trajectories in the ground (excited) state. (b) Simulated image of the interference pattern at the end of MAGIS-100 atom interferometer sequence. The image shows the two atom cloud populations corresponding to the two states at the end of the interferometer sequence. The interference fringes depicted here are due to the phase shear readout technique [40]. Figures sourced from [9]. 182

8.3 Schematic of a simplified clock gradiometer. Two cold atom cloud interferometers are manipulated with shared laser pulses, creating a symmetric pair of interferometers separated by the baseline of the experiment. (a) Initial positions of the two cold atom clouds. (b) Laser pulse sequence space-time diagram showing the creation of the two symmetric interferometers using the two starting cold atom clouds. Dark (light) gray lines show laser pulses traveling upwards (downwards). Blue (red) lines show the cold atom cloud trajectories in the ground (excited) state. Figure sourced from [9]. 183

8.4	(a) Circle of confusion in an ideal conventional single view entocentric imager relative to the size of a $100\text{ }\mu\text{m}$ feature, as a function of lens magnification for a range of lens f-numbers. The $100\text{ }\mu\text{m}$ object space feature dimensions are divided by the magnification to get the image space dimension, which is then used to divide the circle of confusion. The horizontal line is drawn for the points where the circle of confusion is 50% of the $100\text{ }\mu\text{m}$ feature on the imaging plane. (b) Circle of confusion in an ideal conventional single view entocentric imager relative to the diffraction limited spot size, as a function of lens magnification for a range of lens f-numbers. The circle of confusion sizes are computed using Eqn. B.1 with a $f = 62\text{ mm}$, though as seen from Eqn. B.2 the dependence on the focal length is not a dominant effect. The depth of field is set to 1 m to account for the typical size of the cold atom clouds in atom interferometers like MAGIS-100. Figure sourced from [41].	187
8.5	Percentage of light collected by an ideal conventional single view entocentric imager that is emanated by an on-axis point object in 4π , as a function of lens magnification for a range of lens F-numbers. The percentages are computed using Eqn. B.3. Figure sourced from [41].	187

8.6	Ray diagram of a single folded view. Here the object is placed at a distance x from the lens center, offset from the focal plane of the lens which is at a distance F from the lens. A flat mirror is introduced between the object and the focal plane, which produces a virtual object that lies on the focal plane; this virtual object is focused onto the sensor by the lens. The mirror normal angle ϕ is chosen such that the principal (central) ray emanating from the object (at angle θ) is reflected towards the center of the lens. This constrains the geometry of the mirror and determines the mirror normal angle as described in Appendix B.2. Note that the $\theta > 0$ view angle results in an image that sees the object as if it were rotated by angle θ about the optical axis. Note also that $\rho = 2\phi - \theta$. The marginal rays from the lens center to the mirror edges are shown in red, and corresponding field of view is highlighted on the focal plane. The half angle χ corresponds to the angular field of view and is determined following the logic presented in Appendix B.3. Figure sourced from [41].	189
-----	---	-----

8.7	Principal ray diagram in a setup with multiple view folding using mirrors. The object is shown as a black dot along the optical axis of the system. Each mirror is colored from blue to red proportionally to the magnitude of the polar view angle θ captured by the mirror. The principal rays are traced with the same color, and are shown to start at the location of the virtual object and end at the sensor. The focal plane is depicted here with a bounding box that corresponds to the size of the sensor divided by the lens magnification. It can be seen that for the given array of mirrors, there is some minimum distance between neighboring virtual objects centers. The density of virtual objects on the focal plane can be modified by increasing or decreasing the number of views and mirrors to appropriately account for the virtual object size. Note here that the view angles (θ, ψ) are sampled on the Fibonacci lattice as governed by Eqn. 8.1, but only those views (and corresponding mirrors) within the appropriate hemisphere, and which produce virtual objects that lie inside the field of view of the imaging apparatus, are shown. Figure sourced from [41].	191
8.8	Various views of the demonstrator CAD model. The teal disks are the mirrors. (a) Expanded view. The black parts to the left and the right of the mirrors are the front board and the back board, respectively. (b) Assembled view of the mirror holder. (c) A zoomed-in view, showing each mirror being pushed against a ring-like structure on the front board. (d) A zoomed-in view without the front board, showing each mirror being pushed by a cylindrical structure on the back board. Figure sourced from [41].	197
8.9	Laboratory photo of the demonstrator. A pair of thin (0.005") fibers are used to create a crosshair that defines the nominal target object position and holds the object in place. Figure sourced from [41]. . .	198
8.10	Image of the alignment grid of the demonstrator. The entire grid is focused within depth of field of 2 mm. Figure sourced from [41]. . .	198

8.11	Grid of views extracted from a simulated light field 60MP image of an atom cloud imaged over $10\text{ }\mu\text{s}$. Figure sourced from [41].	203
8.12	Data sampled from the learned 3D cloud model in gray, the learned 2D marginals in blue, and the contour lines of the target density in red (the 40, 68 and 95% contours are shown). Figure sourced from [41].	204
8.13	Grid of training views extracted from a 60MP image of the test object, taken in a single shot. Figure sourced from [41].	205
8.14	Comparison of subset of training (real) views to the corresponding subset of generated views from the learned NeuS reconstruction. Results are nearly identical between real captured images and their generated counterparts, with an average SSIM across training views of 0.98. Figure sourced from [41].	208
8.15	Reconstruction of the cube test object. The envelope dimension of the object is $(1\text{ mm})^3$, and the letter is $142\text{ }\mu\text{m}$ thick. (a-c): CAD design, (d-f): microscope images of a 3D printed object, (g-i): depth map of surface reconstructed using 77 views from the system. Different columns correspond to different views of the same objects. Note that the reconstructed results contain artifacts from the $120\text{ }\mu\text{m}$ fiber used to suspend the object, as well as from the limited range of viewing angles. Figure sourced from [41].	209
8.16	Images of reconstructed surface mesh. The top row (a-c) shows similar D, O, and E views as those presented in Figure 8.15. The bottom row (d-f) shows alternative views, where the first two columns show angles of the object that are not directly a D, O, or E lettering, emphasizing that a 3D model has been learned for the imaged object, while the third column shows the reconstruction of an object face that is never directly imaged by the system (the E face on the back). While quality is worse on this face (f) of the object, the available information from the views that are imaged is sufficient to infer the general structure of the object. Figure sourced from [41].	210

8.17	Interpolated views generated from the learned NeuS reconstruction. The range of view angles is between -55 and 55 degrees, and resulting views contain no explicit correspondence to those used for training. Image quality is generally high, demonstrating a learned 3D structure and light field. Figure sourced from [41].	212
8.18	Fourier shell correlation curve for the reconstructed D-O-E object. 77 extracted views are randomly split into two distinct halves, and each half is separately reconstructed. All reconstructions are trained for 140000 iterations, and the Fourier shell correlation is calculated by comparing logistic densities of learned width centered at the learned surfaces. Results shown are the mean and standard deviation calculated from running this procedure across 5 different random splits of the 77 views. Resolutions are quoted for two commonly used thresholds, with $71 \pm 17 \mu\text{m}$ resolution corresponding to the standard threshold used for split dataset comparisons (0.143). Results are qualitatively consistent with the size of reconstructed features. Figure sourced from [41].	213
8.19	Resolution of the system as a function of misalignment and 3D printing error. The resolutions were obtained by computing the FSC between the ground truth and the 3D reconstructed models of the test object. Without calibration of the simulator, displacements below 0.6 mm or 2° still allow to reconstruct $100 \mu\text{m}$ features. Results shown are the mean and standard deviation calculated from running the reconstruction algorithm with 5 different seeds. Figure sourced from [41].	215
9.1	Timeline showing the discoveries of various particles of the Standard Model of Particle Physics. Figure sourced from CERN [42]	217
A.1	Mass Validation Plots at Truth Level	221
A.2	Kinematic Validation Plots at Truth Level	222

A.3	Invariant mass of the two leading jets, m_{jj} . Distributions for data and $Z + \text{QCD jets}$ Pythia MC background before and after variable- and high-dimensional reweighting. Events with m_{jj} inside the signal region (Fig 5.7a) were excluded during training of the reweighting classifier, and are excluded in this plot for blinding.	223
A.4	Two point energy correlation $\text{ECF}(2,1)$. Distributions for data and $Z + \text{QCD jets}$ Pythia MC background before and after variable- and high-dimensional reweighting. Events with m_{jj} inside the signal region (Fig 5.7a) were excluded during training of the reweighting classifier, and are excluded in this plot for blinding.	224
A.5	Number of tracks in the jets. Distributions for data and $Z + \text{QCD jets}$ Pythia MC background before and after variable- and high-dimensional reweighting. Events with m_{jj} inside the signal region (Fig 5.7a) were excluded during training of the reweighting classifier, and are excluded in this plot for blinding.	225
A.6	Distributions for data and $Z + \text{QCD jets}$ Pythia MC background before and after variable- and high-dimensional reweighting. Events with m_{jj} inside the signal region (Fig 5.7a) were excluded during training of the reweighting classifier, and are excluded in this plot for blinding. Here kinematics of the leading jet are shown.	226
A.7	Distributions for data and $Z + \text{QCD jets}$ Pythia MC background before and after variable- and high-dimensional reweighting. Events with m_{jj} inside the signal region (Fig 5.7a) were excluded during training of the reweighting classifier, and are excluded in this plot for blinding. Here kinematics of the sub-leading jet are shown.	227
A.8	Distributions for data and $Z + \text{QCD jets}$ Pythia MC background before and after variable- and high-dimensional reweighting. Events with m_{jj} inside the signal region (Fig 5.7a) were excluded during training of the reweighting classifier, and are excluded in this plot for blinding. Here kinematics of the Z boson are shown, formed by adding the 4-vectors of the two compatible leptons in the event.	228

A.9	Invariant mass of the two leading jets, m_{jj} . Distributions for data and $Z + \text{QCD jets}$ Sherpa MC background before and after variable- and high-dimensional reweighting. Events with m_{jj} inside the signal region (Fig 5.7a) were excluded during training of the reweighting classifier, and are excluded in this plot for blinding.	229
A.10	Two point energy correlation $\text{ECF}(2,1)$. Distributions for data and $Z + \text{QCD jets}$ Sherpa MC background before and after variable- and high-dimensional reweighting. Events with m_{jj} inside the signal region (Fig 5.7a) were excluded during training of the reweighting classifier, and are excluded in this plot for blinding.	230
A.11	Number of tracks in the jets. Distributions for data and $Z + \text{QCD jets}$ Sherpa MC background before and after variable- and high-dimensional reweighting. Events with m_{jj} inside the signal region (Fig 5.7a) were excluded during training of the reweighting classifier, and are excluded in this plot for blinding.	231
A.12	Distributions for data and $Z + \text{QCD jets}$ Sherpa MC background before and after variable- and high-dimensional reweighting. Events with m_{jj} inside the signal region (Fig 5.7a) were excluded during training of the reweighting classifier, and are excluded in this plot for blinding. Here kinematics of the leading jet are shown.	232
A.13	Distributions for data and $Z + \text{QCD jets}$ Sherpa MC background before and after variable- and high-dimensional reweighting. Events with m_{jj} inside the signal region (Fig 5.7a) were excluded during training of the reweighting classifier, and are excluded in this plot for blinding. Here kinematics of the sub-leading jet are shown.	233
A.14	Distributions for data and $Z + \text{QCD jets}$ Sherpa MC background before and after variable- and high-dimensional reweighting. Events with m_{jj} inside the signal region (Fig 5.7a) were excluded during training of the reweighting classifier, and are excluded in this plot for blinding. Here kinematics of the Z boson are shown, formed by adding the 4-vectors of the two compatible leptons in the event.	234

A.15	Expected Limits. For each signal mass point considered, the expected limits from the single bin counting experiment are derived for every value of the DisCo weight hyperparameter λ tested.	243
A.16	Impact Plots. For each signal mass point considered, the impact of each modifier is individually determined by fixing its value to its best fit value plus/minus its fit error.	244
A.17	Impact Plots. For each signal mass point considered, the impact of each modifier is individually determined by fixing its value to its best fit value plus/minus its fit error.	245
A.18	Pull Plots. For each signal mass point considered, the pull plots for each modifier.	246
A.19	Pull Plots. For each signal mass point considered, the pull plots for each modifier.	247
B.1	Ray diagram for an ideal thin converging lens of aperture diameter A . Only the marginal rays are shown, where each line-style corresponds to rays starting from a unique object distance along the optical axis. The rays from object distance u are focused down to a point at image distance v . Rays starting at object distances F' (N') converge at a point closer (further) to the lens than v . In these cases, light spreads across the spot of diameter C when looking at the plane at distance v from the lens. This spot size corresponds to the geometric circle of confusion for the lens, and $F' - N'$ is the geometric depth of field of the lens about the nominal object distance u . Figure sourced from [41].	249

- B.2 Principal ray diagram of a single folded view. Here the object (smiley face) is placed at a distance x from the lens center, offset from the focal plane of the lens which is at a distance F from the lens. A flat mirror is introduced between the object and the focal plane, which produces a virtual object that lies on the focal plane; this virtual object is focused onto the sensor by the lens. The mirror normal angle ϕ is chosen such that the principal (central) ray emanating from the object (at angle θ) is reflected towards the center of the lens. This constrains the geometry of the mirror and determines the mirror normal angle as described in Appendix B.2. Note that the $\theta > 0$ view angle results in an image that sees the object as if it were rotated by angle θ about the optical axis. Note also that $\rho = 2\phi - \theta$. Figure sourced from [41]. 251
- B.3 Image formation geometry for a single folded view like Figure B.2. Here the black X marks the location of the lens center and the black dot the location of the object. The thick solid black vertical line on the right corresponds to the plane at which the lens is focused. The solid colored ellipses represent the family of ellipses that have the lens center and object as the two foci. The dashed circles are the corresponding circular directrices of the ellipses, centered at the lens center. The colored diamonds show the mirror center positions and the colored dots show the corresponding virtual object locations at the plane of focus. The dotted lines connect the lens center, mirror centers, and virtual object locations. Figure sourced from [41]. . . . 253
- B.4 Field of view as a function of the view angle θ determined by solving Eq. B.5. For this plot, an ideal lens of focal length 62.275 mm operating at a magnification of $m = 0.215$ is chosen, where the object distance is picked to allow the 55° view to lie at the edge of a full frame sensor. This is the same setup as described in section 8.3.2.2 and the object distance is picked following Eq. 8.2. The mirror here is a flat mirror of diameter 5 mm centered appropriately to satisfy the constraints on the principal rays. Figure sourced from [41]. 255

B.5	Views from the learned NeuS reconstruction corresponding to those used in training. Relative to the captured images shown in Figure 8.13, results are visually nearly identical, demonstrating reasonable closure of the model. Figure sourced from [41].	256
B.6	Generated grid of surface normals from the learned NeuS model, corresponding to the views used in training (and the generated images of Figure B.5). The colors correspond to the magnitude of the surface normal vectors when projected onto the corresponding 2D image plane. The learned 3D structure of the DOE object is apparent, as expressed via the meshes and depth maps of Figures 8.15 and 8.16. Figure sourced from [41].	257

Chapter 1

Introduction

The pursuit of knowledge about the universe and how it works lies at the core of all physics research, driven by the eternal flame of curiosity. As human science has progressed, nature is found only to present increasingly confounding mysteries as it is probed at smaller length scales and higher energy scales than ever before. Despite the challenges, many successful physical theories have been developed, tested, and refined over the course of scientific investigation. The current best understanding of the universe can be broadly split into reductionist or fundamental theories and emergent theories — fundamental theories aim to explain the behavior of the matter at the most fundamental levels, and emergent theories explain the behavior of matter at relatively macroscopic scales, ostensibly governed by laws that cannot be directly mapped to fundamental theories. The physics discussed in this thesis concerns fundamental theories of nature, more specifically the *Standard Model* of particle physics. The Standard Model encapsulates the various *field* theories that govern the most fundamental particles known to date (elementary fermions and gauge bosons), and describes to extremely high precision the various forces and interactions between fundamental particles. It is based up on the key insight that all interactions between particles are governed by quantum fields, and that particles themselves are the quanta or excitations of quantum fields.¹

¹A notable exception to the forces described by the SM is the force of gravitation, and reconciling gravitation with the fundamental theory of quantum fields is one of the most important outstanding problems in modern physics.

For its many successfully verified predictions, including the existence of the Higgs boson [43–48], there are many questions that are left unanswered by the Standard Model. One of the greatest challenges to the Standard Model come from Dark Matter and Dark Energy, phenomenon that are known to exist due to various astrophysical observations, and are yet unexplained by the Standard Model [49–54]. Additionally, the Standard Model offers no clear answers to a host of issues such as the presence of more visible matter than visible antimatter in the universe today, the number of quark and lepton families in the Standard Model totalling three, the apparent charge-parity-preserving nature of the strong force, the disparity between the strengths of the weak force and gravitational force.

To address the various challenges of the Standard Model, several hypothesis for physics beyond the Standard Model have been proposed and are under active investigation at the various experiments around the world. The study of matter at these fundamental scales is conducted using particle collider experiments along with astroparticle detection experiments. Collider experiments work by accelerating particles such as electrons, protons, or heavy nuclei to high energies and colliding them with each other or suitable fixed targets. The particle interactions that are facilitated in these colliders are guided by the energy of the colliding particles, and can undergo electromagnetic, weak, or strong force interactions. The leading collider experiment currently operational is the *Large Hadron Collider* (LHC) at CERN [55–58], which collides counter-propagating proton beams at a total center of mass energy of 13-14 TeV and can probe the quark-gluon interactions via *Quantum Chromodynamics* [59–62]. The ATLAS experiment [63] at the LHC is one of the experiments probing fundamental physics at these high energy scales in a quest to better characterize the Standard Model and explore extensions, and is relevant to the physics and methods discussed in this thesis. Astroparticle detectors come in several forms depending on the nature of the physics being targeted by the experiments. A new type of astroparticle detector based on matter wave interferometry called MAGIS-100 [9], which will be sensitive to interactions with the hypothesized ultralight Dark Matter, is of particular interest to this thesis. Ultralight Dark Matter, with particle masses ranging from 10^{-22} to 10^{-15} eV, is outside the reach of typical collider experiments. It

is therefore crucial to utilize innovative detection methodologies such as MAGIS-100 that unlock the physics reach to such Dark Matter candidates.

Looking for physics beyond the Standard Model is one of the most important endeavors of this time given the incompleteness and fine tuning inherent to the Standard Model. High energy colliders and astroparticle physics experiments have thus far not identified any physics beyond the Standard Model with defining significance. Notable, Dark Matter and Dark Energy continue to elude the experimental reaches of the experiments [64–68]. However, there are many signatures for physics beyond the Standard Model that are unexplored [69–73]. It is therefore incumbent upon the physics endeavor to expand its experimental reach to new signatures and scan the space of all possible models in the search for new physics that can explain the nature of Dark Matter and Dark Energy, and address the many issue of the Standard Model.

This is the driving motivation behind the original research presented in this thesis. A novel search for a hypothesized coupling of the Higgs boson to a new light scalar particle is shown in Chapter 5. Such couplings are well motivated as they potentially answer some of the outstanding issues of the Standard Model, however several final states of such a novel scalar coupling remain unexplored. The search presented in Chapter 5 tackles the scalar decaying to two Standard Model gluons — The full process, Higgs boson decaying to a pair of scalars to four gluons, has remained previously unexplored due to the challenging experimental signature of the process. The search is now conducted at ATLAS by means of advanced machine learning techniques that make it possible to identify the unique structure of the di-gluon decays from these scalar particles.

Additionally, the original research presented in Chapter 6 and Chapter 7 aims to boost the performance of the ATLAS experiment, in a quest to improve its sensitivity to new physics as it collects more data over the couple of the next few years. The projects focus on improving the efficiency of the ATLAS tracking system to identify physics of interest in Chapter 6, and on improving the calibration procedure for jets reconstructed in the ATLAS experiment in Chapter 7.

Furthermore, to expand the search for new physics beyond the Standard Model, the original research presented in Chapter 8 represents a unique collaboration between

the high energy physics and the atomic-molecular-optics communities. The chapter presents a state-of-the-art light field imaging system developed for the MAGIS-100 experiment, that will enable single-shot tomographic reconstruction of fluorescing cold atom clouds in atom interferometer experiments for the first time. As previously mentioned, the MAGIS-100 experiment is a new detection mechanism for ultralight Dark Matter, and presents an exciting opportunity to expand the search for new physics well beyond the reach of high energy particle collider experiments.

To help contextualize the research, Chapter 2 provides a brief overview of the Standard Model of particle physics, while Chapter 3 and Chapter 4 introduce the Large Hadron Collider and the ATLAS experiment. Additionally, Appendices are included at the end of this thesis to supplement the material shown in Chapter 5 and Chapter 8.

1.1 Units and Coordinate Systems

The system of units used to describe the accelerators and detectors in Chapter 3 and Chapter 4 is the SI unit system. Additionally, SI units are also used to describe the distances along the beam axis in Chapter 6, and the experiments and feature sizes in Chapter 8. All other units use the system of *natural* units, where $c = \hbar = 1$. The commonly used unit of energy is the $10^9 \text{ eV} = \text{GeV}$, and in natural units this also serves as the unit of mass and momentum, where c^{-2} for mass and inverse c^{-1} are implied. This also sets the unit of length and time as GeV^{-1} , where $\hbar c$ for length and $\hbar$ for time are implied.

As mentioned in Chapter 4, A right-handed coordinate system is used in ATLAS, where the origin is at the nominal interaction point in the center of the detector. The z -axis points along the beam direction, the x -axis points from the interaction point to the center of the LHC circular track, and the y -axis is orthogonal to the two and points upwards. Cylindrical coordinates (r, ϕ) are used in the transverse plane (X-Y), where ϕ is the azimuthal angle about the z -axis. Pseudorapidity η is defined using the polar angle θ as, $\eta = -\ln \tan(\theta/2)$. Angular distance in the $\eta - \phi$ plane is measured in units of $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

The momentum of a particle can be expressed as $\vec{p} = (p_x, p_y, p_z)$ or (p_T, η, ϕ) , where $p_T = \sqrt{p_x^2 + p_y^2}$. Here p_T is the transverse momentum of the particle, with angles η, ϕ as described above. The transform between the coordinate systems is:

$$p_x = p_T \cos \phi \quad (1.1)$$

$$p_y = p_T \sin \phi \quad (1.2)$$

$$p_z = p_T \sinh \eta \quad (1.3)$$

The pseudorapidity measures the angle of the particle to the z -axis, and accounts for Lorentz boosts along the z -axis. This can be seen with the full expression of η ,

$$\eta = -\ln \tan\left(\frac{\theta}{2}\right) = \frac{1}{2} \ln \left(\frac{|\vec{p}| + p_z}{|\vec{p}| - p_z} \right) = \tanh^{-1} \left(\frac{p_z}{|\vec{p}|} \right). \quad (1.4)$$

Hence z boosts (with magnitude β) can be accounted for as $(p_T, \eta + \tanh^{-1} \beta, \phi)$. Furthermore, $\Delta\eta$ between two particles is invariant for massless under boosts along the z -axis, motivating the choice of the angular distance metric. This logic can be generalized for massive particles using the rapidity, y ,

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right) = \tanh^{-1} \left(\frac{p_z}{E} \right) \quad (1.5)$$

$$\approx \eta - \frac{\cos \theta}{2} \left(\frac{m}{p_T} \right)^2, \quad (1.6)$$

where the second equation is approximated as a Maclaurin transform in m/p_T . Rapidity is also transformed additively under longitudinal boosts and thus is used to compute the z -boost invariant angular distance metric for massive particles. Note that the differences between η and y are not relevant for $m/p_T \ll 1$.

The four-momentum of a particle can be expressed as $\vec{p} = p^\mu = (E, \vec{p})$, where the mass of the particle is $m = \sqrt{E^2 - |\vec{p}|^2}$. Note here that the metric signature is conventionally taken to be $(+, -, -, -)$. The energy of the particle can be expressed in the cylindrical units as $E = \sqrt{m^2 + p_T^2} \cosh^2 \eta$. The transverse energy of the particle is also expressed as $E_T = \sqrt{m^2 + p_T^2}$.

Chapter 2

The Standard Model of Particle Physics

2.1 Introduction

The Standard Model of particle physics is the relativistic non-Abelian gauge quantum field theory [74, 75] that describes the interactions of the fundamental (matter) fermions and vector gauge bosons of the electromagnetic, weak, and strong forces. In the Standard Model, particles are excitations of quantum fields that in turn interact with one another through various forces by the means of the corresponding fields or gauge bosons, which are the force carrying quanta of the corresponding interactions. The internal symmetries of the various forces in the Standard Model guide the rich physics that the model describes via Lorentz and gauge invariance. The symmetry group of the combined model is $SU(3) \times SU(2) \times U(1)$, where the electroweak interactions are described by the $SU(2) \times U(1)$ symmetry group, and the strong force interactions are described by the $SU(3)$ symmetry group. The theory of the electroweak force will be detailed in Section 2.3, and that of strong force will be detailed in Section 2.4. Section 2.3 will also detail how the gauge bosons, and the fundamental fermions (quarks and leptons), get masses in the Standard Model via the spontaneous breaking of the $SU(2)$ symmetry due to the Higgs boson. It is important to note here that the Standard Model does not describe gravitation and breaks

down at the energy-length scales at which the force carrier of the gravitational force, the graviton, emerges — The Standard Model is hence an effective field theory. The following sections provide only a brief review of the Standard Model — The Author recommends a few texts for a more detailed discussion about the Standard Model, its motivations and justifications [76–78], along with detailed derivations of the various terms in the model [79–81]. Figure 2.1 shows all fermions and gauge bosons in the Standard Model, where the outer ring of particles are spin $1/2$ fermions, the inner ring are spin 1 gauge bosons, and the inside most particle is the scalar spin 0 Higgs boson. Note also that the upper half of the outer rings corresponds to quarks, and the lower half of the outer ring corresponds to the leptons in the Standard Model.

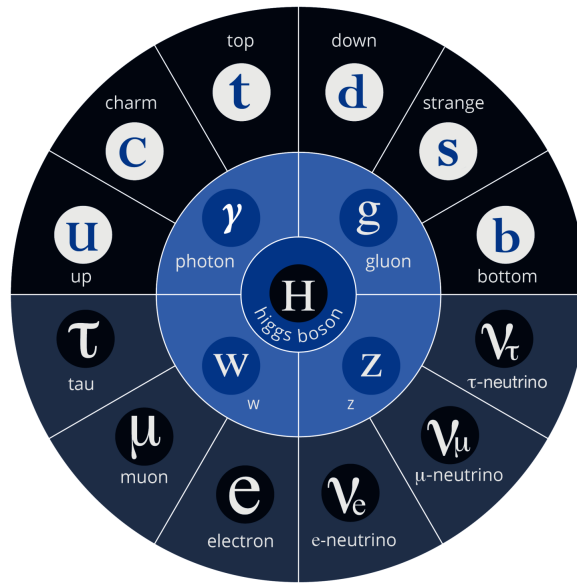


Figure 2.1: Particle content (fundamental fermions and gauge bosons) of the Standard Model of particle physics. Figure sourced from CERN [3].

Note that the Standard Model describes matter in terms of fundamental fermions, namely leptons, quarks, neutrinos. These spin $\frac{1}{2}$ particles follow the Dirac equation, which describes both massive and massless fermions. However, introducing massive fermions in this way violates the gauge invariance of the electroweak interaction. Instead, the mass of the fermions are introduced via the Higgs mechanism, which is also responsible for the masses of the gauge bosons of the electroweak interactions.

2.2 Lagrange Formalism

The Standard Model is fundamentally a perturbative quantum field theory that assumes that isolated or free particle states change only upon interacting with another particles. The solutions to the standard model Lagrangian then work by perturbing the solutions to the free field Lagrangian with the appropriate interaction take into account. Expressed simply, the Standard Model Lagrangian is broken into two parts, $\mathcal{L} = \mathcal{L}_{\text{free}} + \mathcal{L}_{\text{interactions}}$, where the $\mathcal{L}_{\text{interactions}}$ term is the perturbation on the free field $\mathcal{L}_{\text{free}}$. The free field solutions correspond to the suitable equations of motion for the particle being considered. Therefore, the Klein-Gordon equation, the Dirac equation, and the Proca equations apply to spin 0, spin $\frac{1}{2}$, and spin 1 particles accordingly. As noted before, the Lagrangian is both Lorentz invariant, as well as invariant under the $SU(3) \times SU(2) \times U(1)$ internal symmetry.

The free field Lagrangian for a particle is comprised of the mass terms (if any) as well as the kinetic (free propagation) term. As mentioned in the introduction, the mass of the fermions and the gauge bosons in the Standard Model come from the Higgs mechanism, which will be detailed in Section 2.3. The kinetic term for fermions, $i\bar{\psi}\gamma^\mu\partial_\mu\psi$, where ψ is the fermion Dirac field and γ^μ are the gamma or Dirac matrices, is folded into the covariant derivative for the electroweak and strong interactions into so-called *dynamical terms*. The kinetic terms for the gauge bosons rely on the definition of the field strength tensor in terms of the gauge fields F_μ^a ,

$$F_{\mu\nu}^a = \partial_\mu F_\nu^a - \partial_\nu F_\mu^a + gf^{abc}F_\mu^b F_\nu^c, \quad (2.1)$$

where (μ, ν) are Lorentz indices, a is the index of the gauge boson (> 1 if the symmetry group dictates > 1 gauge bosons), g is called the *coupling constant* and is representative of the strength of the interaction mediated by the gauge boson F_μ^a , and f^{abc} are the structure constants of the underlying symmetry group. The structure constants are a unique property of the underlying symmetry group and are given by the commutators of the generators of the group t_a , as $[t_a, t_b] = if^{abc}t_c$. Note that for Abelian groups, the structure constants are all zero, whereas they are non-zero for non-Abelian groups (totally anti-symmetric in the $SU(2)$ and $SU(3)$ case). The

kinetic term in the Lagrangian for the given field strength $F_{\mu\nu}^a$ is then,

$$\mathcal{L}_{\text{kinetic}} = -\frac{1}{4}F_{\mu\nu}^a F_a^{\mu\nu} \quad (2.2)$$

where the prefactor may depend upon the choice of convention, and here it comes from matching the energy of the field to its canonical value.

2.3 Electroweak Interactions

The electroweak force is chiral and results from the unification of the electromagnetic and weak forces under the *Glashow-Weinberg-Salam* electroweak unification model [82–85]. The force describes the action of four force mediating vector bosons, under the gauge symmetry group $U(2) \simeq SU(2) \times U(1)$. This was spurred on by the apparent participation of a charged current in neutron decays, the neutral current in electron-muon scattering as well as the participation of neutral and charged currents in electron-neutrino scattering — These currents, along with the previously known photon, are all related to one another and thus inspired the need for a unified theory with four force carrying bosons. In this theory, the symmetry group $SU(2)_W$ has three gauge bosons ($A_\mu^a = A_\mu^1, A_\mu^2, A_\mu^3$) and the conserved charge is called the *weak isospin*. Likewise, the symmetry group $U(1)_Y$ has one gauge boson (B_μ) and the conserved charge is called the *hypercharge*. In the theory of electroweak spontaneous symmetry breaking, a complex doublet breaks the $SU(2)$ symmetry to leave just the electromagnetic gauge symmetry $U(1)$ manifest [43–46, 86].

The kinetic term in the Lagrangian for the electroweak bosons is,

$$\mathcal{L}_{\text{kin, EW}} = -\frac{1}{4}W_{\mu\nu}^a W_a^{\mu\nu} - \frac{1}{4}B_{\mu\nu} B^{\mu\nu} \quad (2.3)$$

where $W_{\mu\nu}^a$ is the field strength tensor for the A_μ^a gauge bosons of the $SU(2)_W$ symmetry group, and $B_{\mu\nu}$ is the field strength tensor for the B_μ gauge bosons of the $U(1)_Y$ symmetry group. The dynamical term for the complex doublet $\vec{\phi}$ that interacts via the $SU(2)_W$ and $U(1)_Y$ interactions, along with the gauge invariant potential term that gives the field a vacuum expectation value ($\langle |\vec{\phi}| \rangle = v/\sqrt{2}$) is,

$$\mathcal{L}_{\text{Higgs}} = (D_\mu \vec{\phi})^\dagger (D_\mu \vec{\phi}) + m_h^2 (\vec{\phi}^\dagger \vec{\phi} - v^2/2)/2v^2 \quad (2.4)$$

$$\text{where, } D_\mu \vec{\phi} = \left[\partial_\mu - \frac{1}{2}ig_Y B_\mu - \frac{1}{2}ig_W \sigma^a A_\mu^a \right] \vec{\phi}. \quad (2.5)$$

Here g_W and g_Y are the coupling constants of the weak isospin and hypercharge respectively, and the factors of $1/2$ in the covariant derivative comes from the generators

of $SU(2)_W$ and the choice of the hypercharge $Y = 1$ for the scalar field. The potential terms include free parameters of the Higgs mechanism, m_h and v . Due to the $SU(2)$ invariance of the Lagrangian, a gauge can be chosen for real v that minimizes $\mathcal{L}_{\text{Higgs}}$ of the form $\vec{\phi}^T = (0, v + h)/\sqrt{2}$. This choice of gauge rotates away three of the four real scalar fields contained in $\vec{\phi}$ by choice of the $SU(2)$ gauge parameters, leaving only one real scalar field behind — This is called the Higgs field and corresponding boson h the Higgs boson. The Higgs boson describes small perturbations on the $\vec{\phi}$ field about the vacuum expectation value. This Higgs mechanism is evidently responsible breaking the $SU(2)$ symmetry of the electroweak $SU(2)_W \times U(1)_Y$ symmetry. Worth noting here is the fact that this symmetry breaking mechanism inserts an energy scale for the Standard Model proportional to the electroweak symmetry breaking mass scale — This makes it an effective field theory.

With the form of the field $\vec{\phi}$ that spontaneously breaks the symmetry, the kinetic term in $\mathcal{L}_{\text{Higgs}}$ can be re-examined to reveal the mass terms of the electroweak gauge bosons,

$$\mathcal{L}_{m^2} = \frac{v^2}{2} \frac{g_W^2}{4} ((A_\mu^1)^2 + (A_\mu^2)^2) + \frac{v^2}{2} \left(-\frac{g_W}{2} A_\mu^3 + \frac{g_Y}{2} B_\mu \right)^2. \quad (2.6)$$

This allows for the identification of the four gauge bosons in the broken theory.

$$W_\mu^+ \equiv \frac{A_\mu^1 - iA_\mu^2}{\sqrt{2}}, \quad W_\mu^- \equiv \frac{A_\mu^1 + iA_\mu^2}{\sqrt{2}} \quad (2.7)$$

$$Z_\mu \equiv \frac{g_W}{\sqrt{g_W^2 + g_Y^2}} A_\mu^3 - \frac{g_Y}{\sqrt{g_W^2 + g_Y^2}} B_\mu \quad (2.8)$$

$$A_\mu \equiv \frac{g_W}{\sqrt{g_W^2 + g_Y^2}} A_\mu^3 + \frac{g_Y}{\sqrt{g_W^2 + g_Y^2}} B_\mu \quad (2.9)$$

Here the charged W bosons correspond to the maximally parity violating $(V - A)$ charged currents, with masses $m_W = g_W v/2$. The Z boson corresponds to the neutral current, with mass $m_Z = v\sqrt{g_W^2 + g_Y^2}/2$. Finally, the electromagnetic photon A_μ is seen to remain massless after the symmetry breaking. Note that the fundamental unit of electric charge $e \equiv \frac{g_W g_Y}{\sqrt{g_W^2 + g_Y^2}}$, and that the Weinberg angle is defined as $\tan \theta_W = g_Y/g_W$.

The covariant derivative for any general $SU(2)_W$ doublet with hypercharge Y can now be expressed in terms of the broken theory bosons as,

$$D_\mu = \partial_\mu - i \frac{g_W}{\sqrt{2}} \frac{\sigma_1 + i\sigma_2}{2} W_\mu^+ - i \frac{g_W}{\sqrt{2}} \frac{\sigma_1 - i\sigma_2}{2} W_\mu^- \quad (2.10)$$

$$- i \left(\frac{\sigma_3}{2} g_W \cos \theta_W - Y \frac{1}{2} g_Y \sin \theta_W \right) Z_\mu - i \left(\frac{\sigma_3 + Y1}{2} \right) g_W \sin \theta_W A_\mu, \quad (2.11)$$

where the upper particle in the doublet has electric charge $Q_{\text{upper}} = \frac{1}{2} + \frac{Y}{2}$, and the lower particle has the electric charge $Q_{\text{lower}} = -\frac{1}{2} + \frac{Y}{2}$. Note that the $1/2$ comes from the charge of the particles under $SU(2)_W$, where the upper particle is taken to have the positive $1/2$ weak isospin and the lower particle is taken to have the negative $1/2$ weak isospin.

The scalar Higgs boson has hypercharge $Y = 1$, weak isospin $-1/2$ (lower particle in the doublet in the broken gauge $\vec{\Phi}$), and is electrically uncharged. The Higgs boson was discovered in 2012 [47, 48] at the Large Hadron Collider, and in the last decade both ATLAS and CMS have made numerous precision measurements of the various properties of the Higgs boson, finding striking consistency with the Standard Model predictions [87, 88]. Several experimentally observed properties [1] help set the parameters of the model: $\alpha_e = \frac{e^2}{4\pi} \approx 0.007297 \pm .15(\text{ppb})$, $m_Z = 91.1876 \pm 0.0021 \text{ GeV}$, $m_W = 80.377 \pm 0.012 \text{ GeV}$, $m_H = 125.25 \pm 0.17 \text{ GeV}$, $\sin^2 \theta_W = 0.23121 \pm 0.00004$ and $v \approx 246 \text{ GeV}$.

The following subsections discuss the coupling of the electroweak force to Standard Model fermions, and the Higgs mechanism imparting mass to the fermions. Figure 2.2 shows a schematic diagram depicting the effects of the spontaneous symmetry breaking from the Higgs mechanism on the Standard Model fermions and gauge bosons.

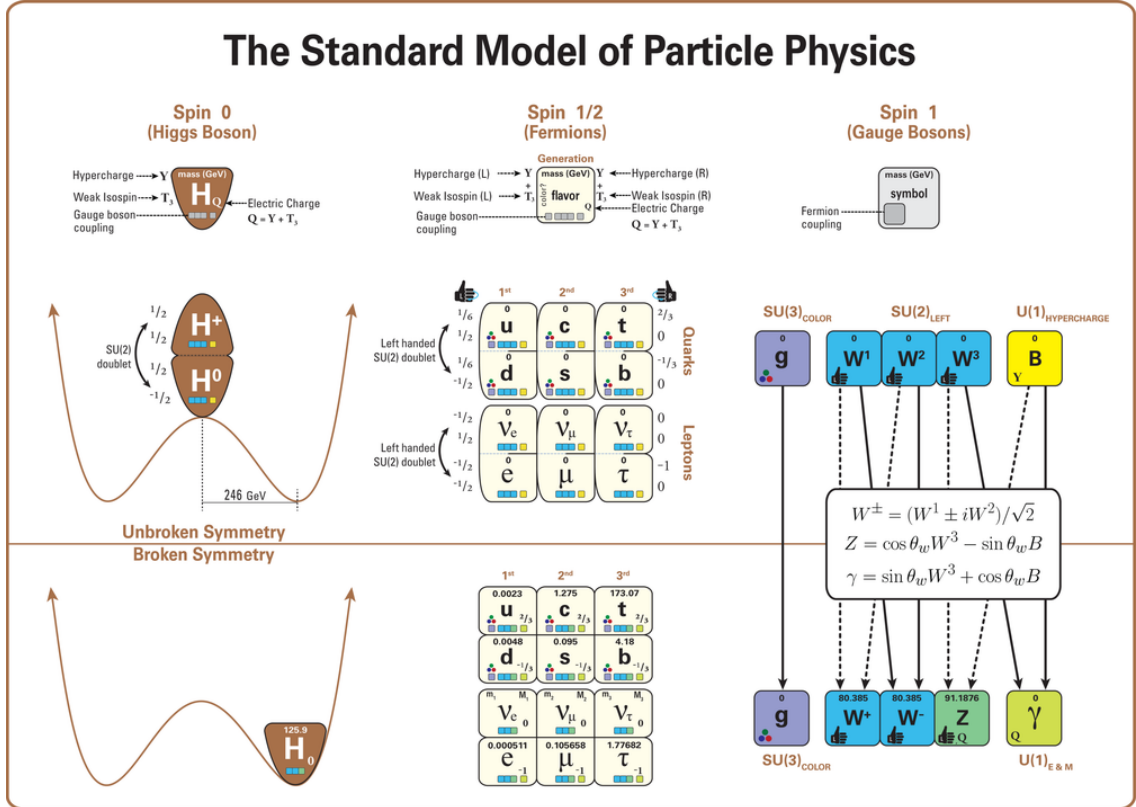


Figure 2.2: Summary of the Standard Model fermions and gauge bosons before and after the spontaneous symmetry breaking via the Higgs mechanism. Note that this figure uses the *alternate convention* for the hypercharge, wherein the electric charge is given as the simple sum of the weak isospin and the hypercharge — The value of the hypercharge shown in the figure are half the value shown in the corresponding sections of this thesis. Figure sourced from [4].

2.3.1 Coupling to Fermions

The weak force only interacts with left handed fermions and right handed anti-fermions — It is a chiral interaction. Note that the left and right chiral projections of a Dirac spinor (fermion field) can be projected using the operators $P_L = \frac{1}{2}(1 - \gamma_5)$ and $P_R = \frac{1}{2}(1 + \gamma_5)$, where γ_5 is the fifth gamma matrix. This means that right handed fermions and left handed anti-fermions are singlets under $SU(2)_W$ with zero weak isospin. In contrast, left handed fermions form doublets that interact with the $SU(2)_W$ component of the electroweak force. There are three generations of leptons and quarks in the Standard Model, each with corresponding left and right handed components that interact with the electroweak force correspondingly. The three lepton generations consist of electrons e , muons μ , and taus τ , and their corresponding neutrinos ν_e , ν_μ , ν_τ . The three generations of quarks consist of up-type quarks up u , charm c , and top t quark, and their corresponding down-type quarks down d , strange s , and bottom b quark. The left handed electron and neutrinos form the $SU(2)$ doublet for the leptons, and the left handed up and down type quarks form the $SU(2)$ doublet for the quarks.

2.3.2 Leptons

The lepton doublets and singlets that participate in electroweak interactions are,

$$l_L^i = \begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix}, \begin{pmatrix} \nu_{\mu L} \\ \mu_L \end{pmatrix}, \begin{pmatrix} \nu_{\tau L} \\ \tau_L \end{pmatrix} \quad (2.12)$$

$$l_R^i = e_R, \mu_R, \tau_R \quad (2.13)$$

$$\nu_R^i = \nu_{eR}, \nu_{\mu R}, \nu_{\tau R}, \quad (2.14)$$

Where the superscript indexes the lepton generation and the subscripts indicate the handedness of the underlying fermion fields. Note here that the doublet has been arranged suggestively such that the upper particle has isospin $1/2$ and the lower particle has isospin $-1/2$, provided that the electrons are known to have unit negative fundamental electric charge. The charges of the lepton fields under the $SU(2) \times U(1)$

Table 2.1: Charges of the Standard Model lepton fields under U(1) hypercharge Y , SU(2) weak isospin T_3 , and the electromagnetic charge Q , related by $Q = T_3 + \frac{Y}{2}$

Charge	e_L, μ_L, τ_L	e_R, μ_R, τ_R	$\nu_{eL}, \nu_{\mu L}, \nu_{\tau L}$	$\nu_{eR}, \nu_{\mu R}, \nu_{\tau R}$
SU(2) _W	$-1/2$	0	$1/2$	0
U(1) _Y	-1	-2	-1	0
Electric Charge	-1	-1	0	0

interactions are listed in Table 2.1. Note that the right handed neutrinos, called *sterile neutrinos*, are singlets under the SU(2) charge and have zero electric charge, leading them to have zero hypercharge, effectively rendering them completely inert to electroweak interactions.

From Eq. 2.10 it is clear that the electroweak interactions with the W bosons will lead to interactions that convert upper particles of the interacting doublet to the lower particle and vice versa. Note here that the particles stay within their generation — The generations are called *flavors* and the interactions are said to be in the flavor basis. The flavor basis Lagrangian for the lepton interactions with the W boson contains,

$$\mathcal{L} \supset \frac{g_W}{\sqrt{2}} \nu_{eL,F}^\dagger \bar{\sigma} \cdot W^+ e_{L,F} + \frac{g_W}{\sqrt{2}} \nu_{\mu L,F}^\dagger \bar{\sigma} \cdot W^+ \mu_{L,F} \quad (2.15)$$

$$+ \frac{g_W}{\sqrt{2}} \nu_{\tau L,F}^\dagger \bar{\sigma} \cdot W^+ \tau_{L,F} + \text{h.c.} \quad (2.16)$$

where the hermitian conjugate terms are implied, and the subscript F highlights the flavor basis of the interactions. The masses of the leptons can be explained in the same way as that for the W and Z bosons, via the scalar Higgs field $\vec{\phi}$ and expanding around the symmetry breaking vacuum expectation value. This requires Lagrangian

Table 2.2: Lepton masses [1].

Lepton	e	μ	τ	ν_e	ν_μ	ν_τ
Mass [MeV]	0.511	105.7	1777	$< \mathcal{O}(10^{-6})$	$< \mathcal{O}(10^{-6})$	$< \mathcal{O}(10^{-6})$

terms that are $SU(2)_W \times U(1)_Y$ gauge and Lorentz invariant,

$$\mathcal{L} \supset -y_e \left(\nu_{eL}^\dagger, e_L^\dagger \right) \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix} e_R + -y_\mu \left(\nu_{\mu L}^\dagger, \mu_L^\dagger \right) \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix} \mu_R \quad (2.17)$$

$$+ -y_\tau \left(\nu_{\tau L}^\dagger, \tau_L^\dagger \right) \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix} \tau_R + \text{h.c.} \quad (2.18)$$

which results in mass terms like $-\frac{y_e v}{\sqrt{2}} e_L^\dagger e_R$ for coefficients y_i called *Yukawa couplings*. Note here that this is done in the mass basis, and can be re-expressed in terms of the lepton mass matrix M_l ,

$$\mathcal{L} \supset - \begin{pmatrix} e_L^\dagger, \mu_L^\dagger, \tau_L^\dagger \end{pmatrix} \begin{pmatrix} \frac{y_e v}{\sqrt{2}} & 0 & 0 \\ 0 & \frac{y_\mu v}{\sqrt{2}} & 0 \\ 0 & 0 & \frac{y_\tau v}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} e_R \\ \mu_R \\ \tau_R \end{pmatrix} + \text{h.c.} \quad (2.19)$$

In fact, this mass matrix formulation can be identically extended to give masses to the neutrinos with the neutrino mass matrix M_ν . The masses of the leptons and neutrinos are then given by the eigenvalues of the mass matrix, and the observed masses for the leptons are reported in Table 2.2. Since both the mass and flavor bases are complete, unitary transformations between the bases can be done via appropriate 3 unitary matrices, like

$$\begin{pmatrix} e_{L,F} \\ \mu_{L,F} \\ \tau_{L,F} \end{pmatrix} = U_l \begin{pmatrix} e_L \\ \mu_L \\ \tau_L \end{pmatrix}, \quad \begin{pmatrix} \nu_{eL,F} \\ \nu_{\mu L,F} \\ \nu_{\tau L,F} \end{pmatrix} = U_\nu \begin{pmatrix} \nu_{eL} \\ \nu_{\mu L} \\ \nu_{\tau L} \end{pmatrix}. \quad (2.20)$$

Here the two unitary matrices U_l and U_ν facilitate the transform, as seen by writing the flavor basis weak interaction terms in terms of the mass basis,

$$\mathcal{L} \supset \frac{g_W}{\sqrt{2}} \left(\nu_{eL}^\dagger, \nu_{\mu L}^\dagger, \nu_{\tau L}^\dagger \right) U_\nu^\dagger (\bar{\sigma} \cdot W^+) U_l \begin{pmatrix} e_L \\ \mu_L \\ \tau_L \end{pmatrix} + \text{h.c.} \quad (2.21)$$

From this, it can be identified that the lepton flavor matrix $V = U_\nu^\dagger U_l$. In the of massless neutrino scenario, the flavor matrix is just the identity matrix, however in the massive neutrino scenario the matrix is not trivial and hence leads to flavor mixing. Due to the experimental observation of neutrino oscillation (flavor changing in propagation) [1], a non trivial unitary transform can exist between the flavor and mass basis for neutrinos — This is the *Pontecorvo–Maki–Nakagawa–Sakata* matrix (PMNS matrix) [89]. The unitary transform can be generally represented as,

$$V = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta_{23} & \sin \theta_{23} \\ 0 & -\sin \theta_{23} & \cos \theta_{23} \end{pmatrix} \times \begin{pmatrix} \cos \theta_{13} & 0 & \sin \theta_{13} e^{i\delta} \\ 0 & 1 & 0 \\ -\sin \theta_{13} e^{i\delta} & 0 & \cos \theta_{13} \end{pmatrix} \times \begin{pmatrix} \cos \theta_{12} & \sin \theta_{12} & 0 \\ -\sin \theta_{12} & \cos \theta_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (2.22)$$

This is because despite there being nine degrees of freedom in a general 3×3 unitary matrix, the PMNS matrix can be fully described by four parameter — Five of the parameters can be absorbed by the lepton fields. The angles θ in the matrix correspond to the mixing angles between the different flavors, and the single complex phase is responsible for introducing charge-parity violation in the lepton sector (different behavior for particles and anti-particles).

Direct empirical measurements with neutrinos are challenging, but several experiments have made observations of oscillations between the neutrino flavors in both terrestrial and astroparticle experiments, the results of which are summarized in Refs. [1, 90, 91] — The first experimental detection of neutrino oscillation was awarded

Table 2.3: Charges of the Standard Model quark fields under $U(1)$ hypercharge Y , $SU(2)$ weak isospin T_3 , and the electromagnetic charge Q , related by $Q = T_3 + \frac{Y}{2}$

Charge	u_L, c_L, t_L	u_R, c_R, t_R	d_L, s_L, b_L	d_L, s_L, b_L
$SU(2)_W$	$1/2$	0	$-1/2$	0
$U(1)_Y$	$1/3$	$4/3$	$1/3$	$-2/3$
Electric Charge	$2/3$	$2/3$	$-1/3$	$-1/3$

the Nobel Prize in 2015 [92]. The best fit parameter values suggest strong flavor mixing for neutrinos, along with a bias towards CP-violation.

2.3.3 Quarks

The coupling of the electroweak force to quarks work analogously to the above discussion for leptons. The quark doublets and singlets that participate in electroweak interactions are,

$$q_L^i = \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \begin{pmatrix} c_L \\ s_L \end{pmatrix}, \begin{pmatrix} t_L \\ b_L \end{pmatrix} \quad (2.23)$$

$$q_{R,upper}^i = u_R, c_R, t_R \quad (2.24)$$

$$q_{R,lower}^i = d_R, s_R, b_R, \quad (2.25)$$

Where the superscript indexes the quark generation and the subscripts indicate the handedness of the underlying fermion fields. Note here that the doublet has been arranged suggestively such that the upper particle has isospin $1/2$ and the lower particle has isospin $-1/2$, provided that the protons are known to have unit positive fundamental electric charge. The charges of the quark fields under the $SU(2) \times U(1)$ interactions are listed in Table 2.3.

The flavor basis Lagrangian for the quark interactions with the W boson can be

Table 2.4: Quark masses [1].

Lepton	u	d	c	s	t	b
Mass [MeV]	2.2	4.7	1.28×10^3	96	173.1×10^3	4.18×10^3

compactly written as,

$$\mathcal{L} \supset \frac{g_W}{\sqrt{2}} \left(\mathbf{u}_{L,F}^\dagger, \mathbf{c}_{L,F}^\dagger, \mathbf{t}_{L,F}^\dagger \right) \bar{\sigma} \cdot \mathbf{W}^+ \begin{pmatrix} \mathbf{d}_{L,F} \\ \mathbf{s}_{L,F} \\ \mathbf{b}_{L,F} \end{pmatrix} + \text{h.c.} \quad (2.26)$$

where the hermitian conjugate terms are implied, and the subscript F highlights the flavor basis of the interactions. As in the case with leptons, the masses of the quarks can be explained via the scalar Higgs field $\vec{\phi}$ and expanding around the symmetry breaking vacuum expectation value. This requires Lagrangian terms that are $SU(2)_W \times U(1)_Y$ gauge and Lorentz invariant, which result in Lagrangian terms more compactly represented in the mass basis using the mass matrices as,

$$\mathcal{L} \supset - \left(\mathbf{u}_L^\dagger, \mathbf{c}_L^\dagger, \mathbf{t}_L^\dagger \right) \begin{pmatrix} \frac{y_{uV}}{\sqrt{2}} & 0 & 0 \\ 0 & \frac{y_{cV}}{\sqrt{2}} & 0 \\ 0 & 0 & \frac{y_{tV}}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \mathbf{u}_R \\ \mathbf{c}_R \\ \mathbf{t}_R \end{pmatrix} \quad (2.27)$$

$$- \left(\mathbf{d}_L^\dagger, \mathbf{s}_L^\dagger, \mathbf{b}_L^\dagger \right) \begin{pmatrix} \frac{y_{dV}}{\sqrt{2}} & 0 & 0 \\ 0 & \frac{y_{sV}}{\sqrt{2}} & 0 \\ 0 & 0 & \frac{y_{bV}}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \mathbf{d}_R \\ \mathbf{s}_R \\ \mathbf{t}_b \end{pmatrix} + \text{h.c.} \quad (2.28)$$

where the mass terms for the up- and down-type quarks can be identified along the diagonals of the matrices, and the coefficients y_i are corresponding Yukawa couplings. The various quark masses are reported in Table 2.4. The quark mass and flavor bases

can be related with unitary transformations for up- and down-type quarks,

$$\begin{pmatrix} \mathbf{u}_{L,F} \\ \mathbf{c}_{L,F} \\ \mathbf{t}_{L,F} \end{pmatrix} = \mathbb{U}_u \begin{pmatrix} \mathbf{u}_L \\ \mathbf{c}_L \\ \mathbf{t}_L \end{pmatrix}, \quad \begin{pmatrix} \mathbf{d}_{L,F} \\ \mathbf{s}_{L,F} \\ \mathbf{b}_{L,F} \end{pmatrix} = \mathbb{U}_d \begin{pmatrix} \mathbf{d}_L \\ \mathbf{s}_L \\ \mathbf{b}_L \end{pmatrix}. \quad (2.29)$$

Here the two unitary matrices $\mathbb{U}_u$ and $\mathbb{U}_d$ facilitate the transforms for the up- and down-type quarks respectively. The quark flavor matrix $\mathbf{V} = \mathbb{U}_u^\dagger \mathbb{U}_d$ is necessary to account for the difference between the bases, since the up- and down-type quarks all have masses, preventing this from being a trivial transformation. This quark flavor mixing matrix is called the *Cabibbo-Kobayashi-Masakawa* (CKM) matrix [93,94]. Like the case before, this unitary transform can be generally represented as,

$$\mathbf{F} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta_{23} & \sin \theta_{23} \\ 0 & -\sin \theta_{23} & \cos \theta_{23} \end{pmatrix} \times \begin{pmatrix} \cos \theta_{13} & 0 & \sin \theta_{13} e^{i\delta} \\ 0 & 1 & 0 \\ -\sin \theta_{13} e^{i\delta} & 0 & \cos \theta_{13} \end{pmatrix} \times \begin{pmatrix} \cos \theta_{12} & \sin \theta_{12} & 0 \\ -\sin \theta_{12} & \cos \theta_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (2.30)$$

where only four parameters are required to fully characterize the transform. The angles θ in the matrix correspond to the mixing angles between the different quark flavors, and the single complex phase is responsible for introducing charge-parity violation in the quark sector (different behavior for particles and anti-particles).

These matrix elements have been measured experimentally, with current values being [1] $\theta_{12} = 13.04 \pm 0.05^\circ$, $\theta_{13} = 0.201 \pm 0.011^\circ$, $\theta_{23} = 2.38 \pm 0.06^\circ$, and $\delta = 68.8 \pm 4.6^\circ$. Note here that the CKM matrix demonstrates that the largest mixing happens between the first two generations, but all the mixings are small as the matrix is close to diagonal — This is in contrast to the PMNS matrix which demonstrates much larger flavor mixing for neutrinos. Due to the non-zero complex phase term, $\mathbf{V}_{\text{CKM}}^* \neq \mathbf{V}_{\text{CKM}}$ — Quarks and anti-quarks couple differently to the W boson. Since charge-parity symmetry can transform between particle and anti-particles, this represents a violation of the CP symmetry. Therefore, along with being maximally parity violating,

the weak force also violates the CP symmetry for quarks. Since all theories with real observables must have charge-parity-time (CPT) symmetry to ensure Hermitivity of the Hamiltonian, the weak force can be equivalently thought to violate time reversal symmetry.

2.4 Quantum Chromodynamics

The *vector* theory of strong interactions describes the action of the force mediating bosons under the $SU(3)$ gauge symmetry group [59–62]. The force carrying vector gauge bosons for strong interactions are called gluons, and were motivated as a means to explain the binding of quarks inside hadrons. Gluons needed to have zero electromagnetic charge to ensure the existence of hadrons composed of all like-charged quarks, and were thus known to not interact with leptons. They were predicted to exist in electron-positron annihilation events [95], and were experimentally determined to be consistent with the spin 1 hypothesis [1]. The unbroken gauge invariance of the associated field theory also ensures the massless nature of the gluons, and conservation of the *color* charge — The color charge is the charge associated with the $SU(3)$ gauge symmetry. Quarks transform under the fundamental representation of the $SU(3)$ gauge group — They are spinors in 3 flavors that transform under the action of unitary 3×3 matrices (adjoint representation of $SU(3)$). The charge is called color and is indexed as (red, blue, green) in analogy to visible light colors where the sum of all colors results in colorless light (color charge zero in the strong force case). Therefore in the theory of strong interactions, the symmetry group $SU(3)$ has eight gauge bosons (A_μ^a , where the index a runs over the $\mathbf{8}$ -representation of $SU(3)$) and the conserved color charge — The theory is called Quantum Chromodynamics (QCD) to account for the color charges of the theory.

The kinetic term in the Lagrangian for the gluons is,

$$\mathcal{L}_{\text{kin, QCD}} = -\frac{1}{4} G_{\mu\nu}^a G_a^{\mu\nu} \quad (2.31)$$

where $G_{\mu\nu}^a$ is the field strength tensor for the A_μ^a gauge bosons of the $SU(3)$ symmetry group. The dynamical terms for the quark fermion fields can be written as,

$$\mathcal{L}_{\text{dyn, QCD}} = i\bar{\psi}_f \gamma^\mu D_\mu \psi_f \quad (2.32)$$

where γ^μ are the Gamma matrices, and the covariant derivative $D_\mu = \partial_\mu - ig_s A_\mu^a T^a$.

$T^a = \frac{\lambda^a}{2}$ are the generators of $SU(3)$ in adjoint representation ($= \lambda^a/2$, for the Gell-Mann matrices λ^a), and g_s is the strong coupling constant. The fermion field subscript corresponds to the color flavor of the fermion, and the term in the Lagrangian sums over all colors of the fundamental representation. Note here that since $SU(3)$ is non-Abelian, the full $\mathcal{L}_{\text{QCD}}$ as described by the above equations contains terms that are cubic and quartic in the gluon field A_μ^a — These terms represent self interactions of the gluon fields in the QCD theory. Additionally, the dynamical term allows for two quark + gluon interaction vertices that may or may not involve change in the color of the incoming and outgoing particles.

A vital aspect of QCD is the running of the strong coupling constant as the energy scale being probed is varied. The concept can be contextualized from a perturbation perspective as corrections to the strong coupling constant that depend on the energy scale the theory is being probed, resulting in an effective strong coupling constant. This effect is a consequence of the *renormalization* procedure that is necessary to ensure that the Stanford Model is consistent and predictive [96] — Renormalization is a systematic procedure for eliminating infinities in a quantum field theory by carefully taking differences of finite scales. This allows for calculations of observables such as the cross-sections of scattering process using path integral formulations (Feynman diagrams) despite there being potentially infinitely many (loop) diagrams contributing to the process. The energy dependence of the coupling constant is expressed using the β -function,

$$\beta(g) \equiv Q \frac{\partial g}{\partial Q}, \quad (2.33)$$

where Q is the energy scale the theory is being probed at, and g is the coupling constant. In QCD this can be computed using the one-loop quark-quark-gluon and gluon-gluon-gluon diagrams, and is found [97, 98] to be $\beta(g_s) = -\frac{g_s^3}{(4\pi)^2} (11 - \frac{2}{3}n_f)$. Here n_f is the number of active quark flavors at the energy scale Q , meaning $\beta(g_s)$ is always negative. While the term proportional to n_f has a classical analogue to electron-positrons pairs in vacuum screening the effect of the electromagnetic force (leading to a coupling constant that increases as the probe energy scale gets higher), the -11 term

produces an anti-screening effect that decreases the coupling constant g_s as the probe energies get higher, and dictates the overall behavior of the running of the strong coupling. Consequently, the strong force in QCD gets apparently weaker at higher energies, or as $Q \rightarrow \infty$, $g_s(Q) \rightarrow 0$. This phenomenon in QCD is called *asymptotic freedom* [99–101] since quarks interact increasingly weakly, behaving like free particles, as their energies get asymptotically high.¹ Figure 2.3 shows measurements for the running couple constant (as $\alpha_s = \frac{g_s^2}{4\pi}$) performed by the ATLAS and CMS experiments at the Large Hadron Collider.

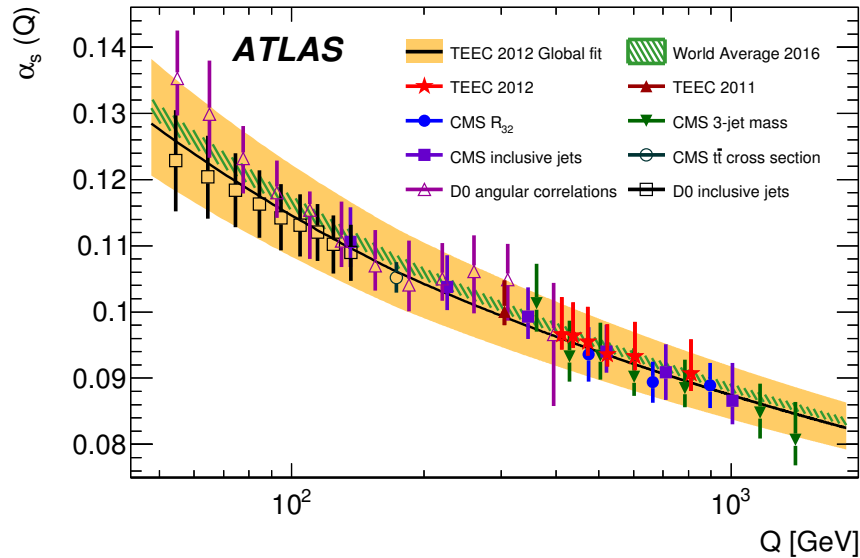


Figure 2.3: Strong coupling constant α_s measured as a function of probe energy scale Q . The experimentally determined values are compared with theoretical predictions that use the reference energy scale of the mass of the Z boson. Figure sourced from [5].

It is evident from Figure 2.3 that the strong coupling constant diverges at low energies and the underlying perturbative assumptions start to fail in this regime, rendering the theory non-perturbative. Due to the increasing interaction strength at low energies (or large distances), it takes greater energy to separate quarks from one another. In other words, in the low energy regime, quarks can form bound states

¹Note here that while the weak interaction SU(2) also has a negative β -function, it remains a short ranged force due to the massive force carrying gauge bosons [102].

through *confinement* [103]. Crucially, these bound states have overall zero color charge, since they would otherwise confine further with other colored objects at low energies — These bound states are called *hadrons*.²

Another consequence of the running of the strong coupling strength happens at the high energy scales, where the coupling diminishes with increasing energy, but logarithmically so. This means that the coupling is low and relatively constant over a range of energies, which implies that QCD is approximately scale invariant at high probe energies. This implies that at any sufficiently high probe energy, the QCD quark and gluon particle production must look alike; zooming in to lower energy scales should continue to reveal the same particle production and structure at all scales till the limits from the running strong coupling constant and the confinement from hadrons take over and cease further particle production. This fractal feature of QCD particle production implies that any quark is always associated by an arbitrary number of gluons that are roughly collinear to the quark direction of propagation. Similarly gluons can produce quark anti-quark pairs that in turn have the same collinear structure. Therefore any process that results in the production of a quark or gluon sees a stream of collimated particles associated to the quark or gluon produced in the process. This collimated stream of particles is called a *jet* [62, 100, 104, 105], and is a result of the *hadronization* that takes place due to the scale invariance of QCD and the running coupling constant.

²2 quark bound states are called *mesons*, 3 quark bound states are called *baryons*; 4 quark bound states are called *tetraquarks*, and 5 quark bound states are called *pentaquarks*. Bound states with more than 5 quarks have been theorized but have not been conclusively discovered to date.

2.5 Summary of Particles and Fields

Figure 2.4 shows the particle content of the Standard Model, which includes the fundamental spin 1/2 fermions, the spin 1 gauge bosons of the electroweak and strong force, and the scalar spin 0 Higgs boson. Note that the electroweak bosons shown in the figure are the bosons from the broken theory that arise as a result of the spontaneous symmetry breaking of the $SU(2)_W$ symmetry by the Higgs mechanism. Further recall that the masses for the fermions and electroweak gauge bosons also come from the Higgs mechanism. Additionally, the three color flavors for each quark and the eight gluon fields of $SU(3)_{QCD}$ are implicit.

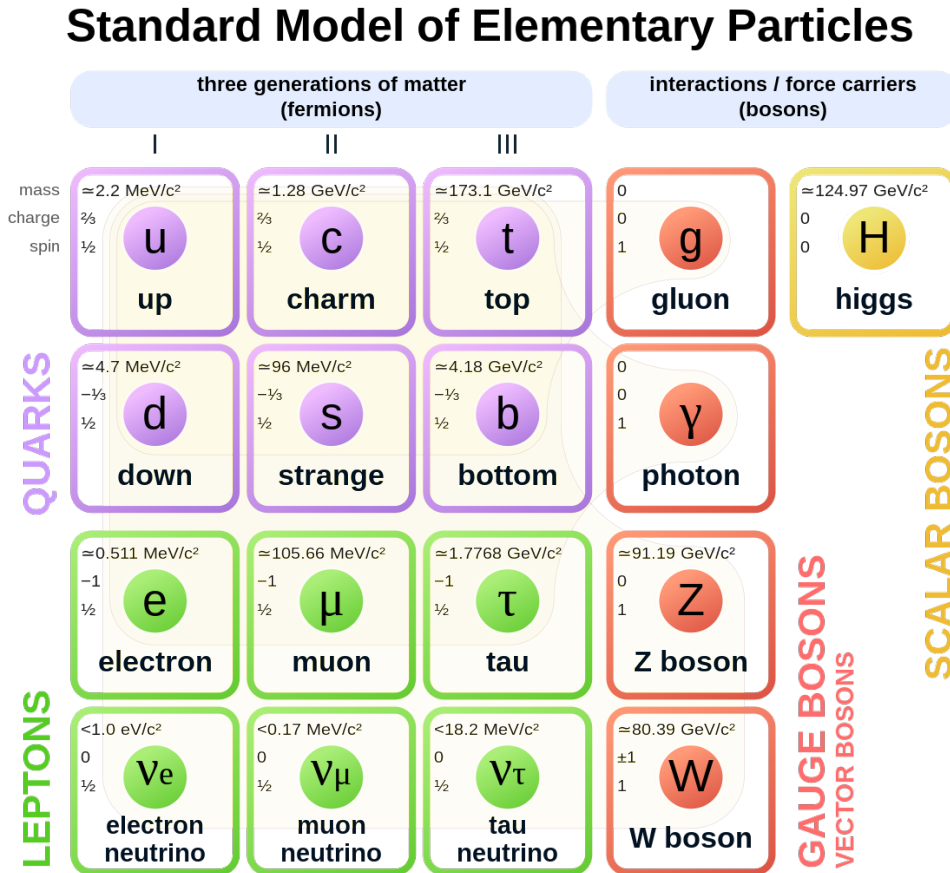


Figure 2.4: Summary of the Standard Model particles. For each particle, the corresponding mass, electric charge, and spin is shown. Figure sourced from [4].

Figure 2.5 shows all the allowed interaction vertex diagrams in the Standard Model. Starting with QCD, quark-quark-gluon and gluon self interaction vertices are allowed in the $SU(3)$ gauge invariant Lagrangian. In electroweak interactions, the electrically neutral Z boson couples to like fermions, where as the $W^\pm$ bosons converts between up-type and down-type left-handed fermions and right-handed anti-fermions. The massless and electrically neutral photon couples to all electrically charged particles. Additionally, due to the non-Abelian nature of the electroweak $SU(2)_W \times U(1)_Y$ gauge symmetry, self interaction terms between the gauge bosons of the electroweak theory are permitted. Finally the Higgs boson interacts with fermions via the Yukawa couplings, and interacts with the massive gauge bosons via the vacuum expectation value mechanism, resulting in effective mass terms for the particles.

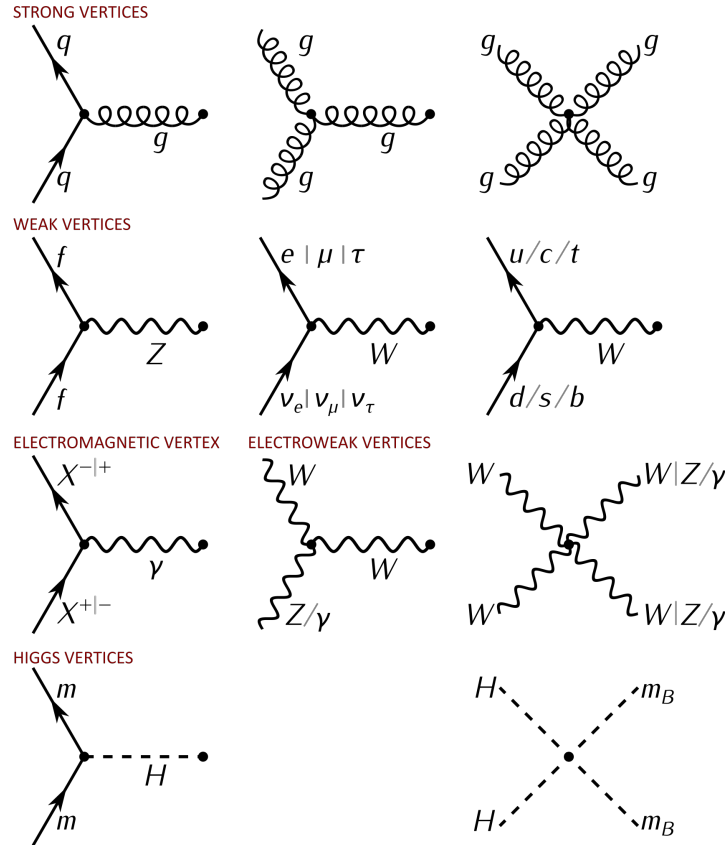


Figure 2.5: Interaction vertex diagrams allowed in the Standard Model. Figure sourced from [4].

An important consequence of massive up- and down-type quarks is the phenomenon of flavor mixing, which is described with the CKM matrix. Quarks can therefore undergo flavor changing interactions mediated by the electroweak force. Additionally, in the massless neutrino theory, no flavor mixing is expected in the lepton sector and the mass and flavor bases are expected to be the same. However the observation of neutrino oscillations suggests that neutrinos have non-zero mass, which then gives rise to lepton flavor mixing that can be described via the PMNS matrix.

Finally, the rich phenomenology of the SU(3) strong force gives rise to the running of the strong coupling strength as well as the scale invariance of QCD at high energies, which in turn leads to color confinement, hadronization and the formation of jets.

As is evident from the previous sections, there are several free parameters in the Standard Model. The Standard Model with massless neutrinos has 19 free parameters: 3 lepton masses + 6 quark masses; 3 CKM matrix angles + 1 CKM matrix complex phase; 3 coupling strengths of the U(1), SU(2), and SU(3) gauges; Higgs vacuum expectation value + Higgs mass; QCD vacuum angle (which is fine tuned to be close to zero due to no experimental observation of CP violation in QCD despite the corresponding term being allowed in the Lagrangian). Adding massive neutrinos to the Standard Model introduces an additional 7 parameters, that correspond to the 3 neutrino masses and the 4 parameters of the PMNS matrix. All the other Standard Model parameters can be expressed in terms of these free parameters. Values for these free parameters have been experimentally measured across an array of different experiments and are summarized in Ref. [1].

As mentioned in the introduction, the Standard Model is a Lorentz invariant and SU(3)×SU(2)×U(1) gauge invariant theory. This leads to the conservation of energy, momentum, angular momentum, as well as the color charge, weak isospin, hypercharge, and electric charge. However, the Standard Model also exhibits additional *accidental* continuous U(1) global symmetries. One such symmetry leads to the conservation of the *baryon number*, which implies that the number of quarks minus the number of antiquarks is a constant. Another such symmetry leads to the conservation of the total *lepton number*, which counts the number of leptons minus the number

of antileptons. It is interesting to note here that due to the flavor mixing of quarks and neutrinos, that individual quark or lepton flavor (or family) numbers are not conserved themselves.

The Standard Model of particle physics is one of the most precisely tested theories of the fundamental forces of nature. The various particles and gauge bosons that constitute the theory have been experimentally detected by looking at the production cross-sections and scattering distributions from a variety of physics processes [1]. Figure 2.6 shows a sampling of the cross-sections for some physics processes measured with the ATLAS detector at the Large Hadron Collider, and compares them with the theoretical predictions from Standard Model calculations — All measurements, across the wide range of $pp \rightarrow X$ final states and cross-sections spanning several orders of magnitude, show remarkable agreement with the predictions within uncertainties.

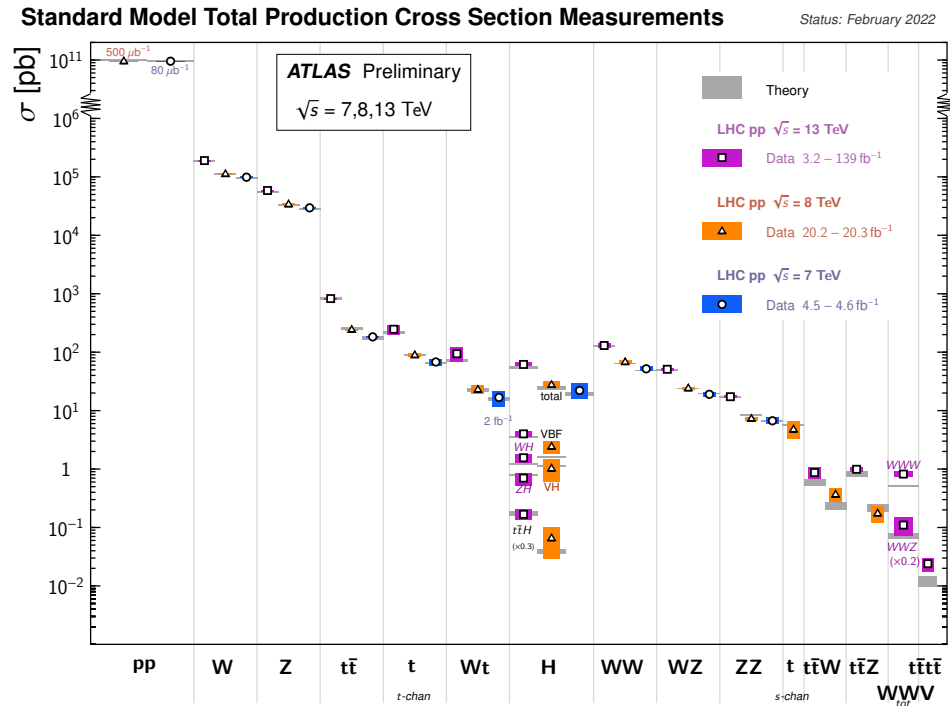


Figure 2.6: Summary of Standard Model cross-section measurements done by the ATLAS experiment at the LHC, compared to the Standard Model predictions. The branching factor corrections have been applied to the measurements for the respective theoretical expectation. Figure sourced from [6].

2.6 Limitations and Exploring Physics Beyond the Standard Model

The Standard Model has shown spectacular accuracy in explaining natural phenomenon, and has been experimentally verified to very high precision. However, there are several issues that limit its explanatory potential, and motivate new theories of physics beyond the Standard Model [106–111]. Some of the limitations and proposed new theories beyond the Standard Model will be discussed below, with an emphasis on the models that are relevant to the subject matter of the research presented in this Thesis.

One major limitation of the Standard Model is its inability to account for the force of gravitation, which is currently described by the theory of general relativity and has not been quantised into the Standard Model framework. The quantum theory of gravity would involve a force carrying gauge boson of spin 2, but the strength of the interaction would only start to dominate at the Plank scales ($\sim 10^{19}$ GeV) [112–115]. String theory [116] is a proposed theory of quantum gravity; work is underway to devise predictions that could be tested in scientific experiments.

Another limitation comes from the observation that the universe has a lot more matter in it than anti-matter [117] — The interactions of particles with the fundamental forces must have CP violating elements to explain the matter anti-matter asymmetry [118]. While the Standard Model does have CP violating interactions in the weak force, specifically in the CKM matrix, it is not sufficient to explain the magnitude of the asymmetry seen in the universe today. CP violation in the lepton sector of the weak force, from the PMNS matrix, is a promising source for additional asymmetry within the Standard Model. Additionally, CP violation in the strong force interactions would also help account for the asymmetry in the Standard Model and are in principle allowed within the framework of the model [119]; This is however not consistent with experimental observations [120], which limits the QCD vacuum angle $\theta_{\text{QCD}} < 10^{-9}$. Note here that the apparent absence of CP violation in strong force interactions constitutes the *strong CP problem*, which is itself another limitation of the Standard Model.

One proposed source of the matter-antimatter symmetry in the universe comes from increased production of matter relative to antimatter (*baryogenesis*) during the first-order electroweak phase transition [121] of the early universe. The transition involves the splitting of the weak and electromagnetic forces as the early universe cooled and happens continuously in the Standard Model, but could happen through a first or second-order phase transition. An important class of models of first-order transitions involve the coupling of the Standard Model Higgs to a new light singlet-like (pseudo)scalar particle (s) [122–124]. The mixing of the Standard Model Higgs with this new scalar results in exotic decays of the Higgs of the form $h \rightarrow ss \rightarrow XXYY$, which can be readily searched for at particle collider experiments [73]. Figure 2.7 shows the limits set on exotic Higgs decays of this form using the ATLAS and CMS experiments at the LHC. The analysis presented in Chapter 5 of this thesis concerns the search of an unexplored exotic Higgs decay of this form, $h \rightarrow ss \rightarrow gggg$.

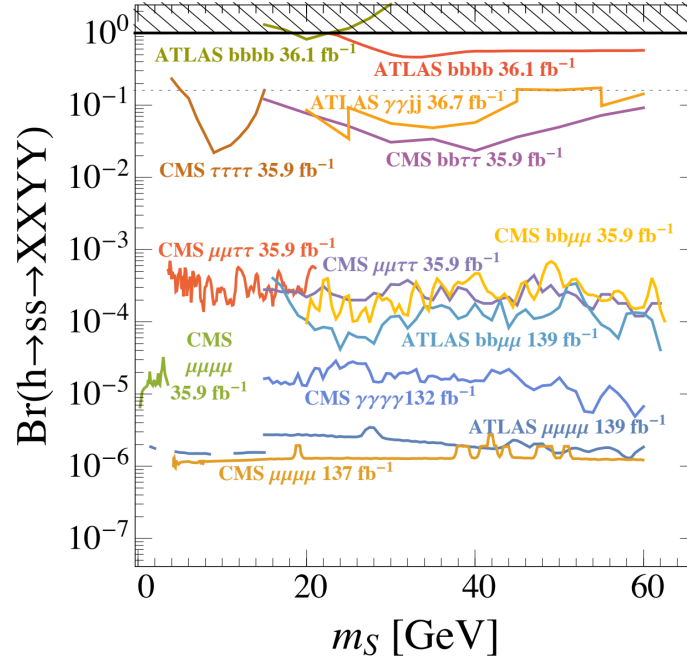


Figure 2.7: Limits on the branching ratio of exotic Higgs decays $h \rightarrow ss \rightarrow XXYY$ via the coupling to new light (pseudo)scalar particles s , for a variety of final states searched with the ATLAS and CMS experiments at the LHC. The dashed horizontal line corresponds to the upper limit on the exotic Higgs branching ratio from global fits to the various Higgs properties. Figure sourced from [7].

Another significant challenge to the Standard Model comes from *dark matter* [125–127]. Dark matter is matter postulated to explain various gravitational effects that have been observed that cannot be explained by visible matter. These effects include the need for additional matter to explain galaxy cluster formation [128–131], gravitational weak lensing [132], galaxy cluster merging [133], etc. The dark matter contribution to the total matter universe is enormous, and is estimated to be about 85% of all matter [134], with the rest comprising of matter from the Standard Model. Dark matter has the key characteristic of interacting with normal matter through the gravitational force, and not interacting with the electromagnetic force (hence the name dark matter).

Determining the composition of this dark matter constitutes one of the greatest challenges of modern physics, and promises to offer insights about the cosmology of the early universe and lead to the discovery of new physics and fundamental forces — The problem is being tackled from a variety of different perspectives. From the Standard Model, neutrinos are a valid candidate dark matter candidate since they are electrically neutral, but they are found to be unable to explain the observed dark matter properties [135]. Beyond the Standard Model, several proposed candidates and mechanisms could give rise to the observed dark matter properties. Weakly interactive massive particles (WIMPS) are a broad class of candidates that have the necessary properties to explain the observed dark matter effects. Particles from the *Supersymmetry* [136–148] theory (proposed also to explain the hierarchical difference between the Higgs mass and the Planck scale) or *axions* [149–155] (proposed also as a solution to the strong CP problem) are valid WIMP and dark matter candidates respectively that can resolve multiple issues with the Standard Model — new elementary particles that interact gravitationally and could form *cold dark matter*. Another interesting set of particles that can form part of the theory of dark matter are the light (pseudo)scalar particles introduced earlier to address the matter anti-matter asymmetry of the universe, as these particles open couplings to other potential dark sector particles — The search for the $h \rightarrow ss \rightarrow gggg$ final state presented in Chapter 5 is therefore relevant also as a search for a potential dark sector particle.

In addition to the above mentioned particles, ultralight bosonic dark matter [8] are

candidates for wave-like dark matter, since these bosons have high number densities to account for the observed dark matter effects. Of these candidates, particles with masses $< 10^{-15}$ eV are of particular interest due to their impact on large-scale structures and astrophysical measurements — These form coherent classically oscillating fields with frequencies set by the mass of the ultralight boson that can interact with matter in several interesting ways. Ultralight scalar bosons can couple to the Standard Model and induce oscillations of the fundamental constants of the theory such as the fine structure constant and in turn change the energy levels of atoms [156, 157] and lengths of solids [157, 158], as well as give rise to time-varying equivalence principle violating forces on bodies [159, 160]. Additionally, ultralight vector (gauge) bosons can couple to the Standard Model and produce new forces along with sourcing equivalence principle violating effects. These ultralight bosonic dark matter candidates also help explain the issue of naturalness of the Higgs, which is related to the hierarchical difference between the Higgs mass and the Planck scale.

The ultralight dark matter candidates discussed above are being searched using a vast array of different detection modalities and quantum sensors. Figure 2.8 shows the various current and projected limits for some example couplings of scalar and vector ultralight dark matter candidates across the range of relevant masses for these models. Atom interferometers have been used to make measurements of the fine structure constant [161, 162] and tests for new fundamental forces [163–166] — this can be exploited to search for ultralight dark matter candidates. This is because the energy levels of the atoms in the interferometer oscillate at the fundamental frequency of the dark matter candidate and can be detected in the phase of the interference fringes as well as the differential phase across separated interferometers in a gradiometer. Additionally, equivalence principle violating forces can be detected at such experiments by running co-located interferometers with different atomic species and taking the appropriate differential measurement [9]. Chapter 8 of this thesis concerns the development of a novel imaging technique for the MAGIS-100 [9] atom interferometer experiment, which is expected to be sensitive to ultralight dark matter across a range of candidate masses as seen from Figure 2.8.

The above discussion has summarized some of the challenges to the Standard

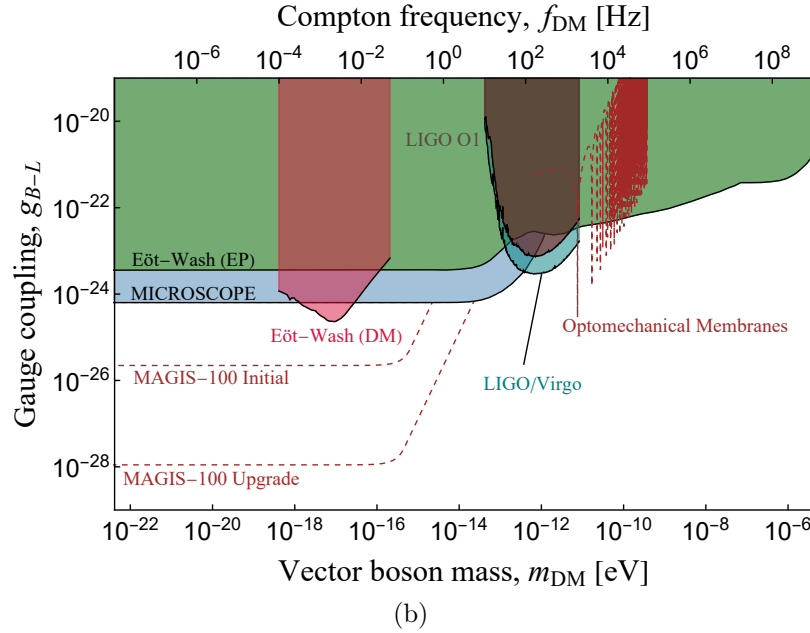
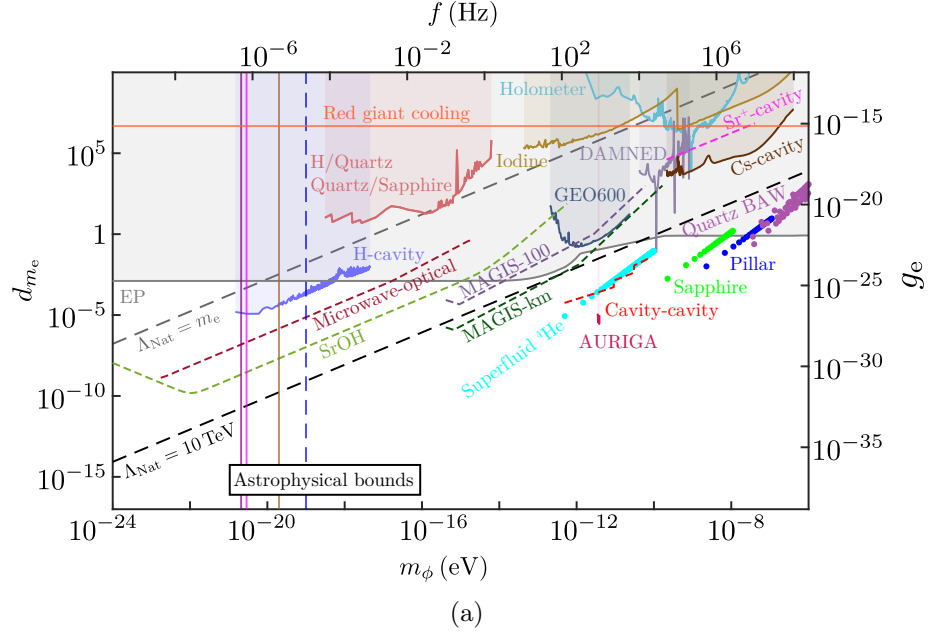


Figure 2.8: Summary plot from various experiments of (a) the ultralight scalar dark matter coupling to the electron (d_{m_e}), and (b) the ultralight vector dark matter gauge coupling (g_{B-L}). Figure sourced from [8]. Of note here is the projected sensitivity of the MAGIS-100 experiment and its upgraded configuration as well as the MAGIS-km experiment [9].

Model and potential physics beyond the Standard Model that can resolve the limitations. There are other limitations that were not mentioned here, such as the disparity between quark and lepton mass scales [167], and the possibility of neutrino masses originating from *Majorana* mass terms rather than the Dirac mass term descriptions on Section 2.3.2. Finally, the vast array of free parameters in the Standard Model also suggest the presence of a greater theory that may naturally explain the scales of the various Standard Model interactions and reveal the structure of dark matter, etc.

Of the many proposed physics models beyond the Standard Model discussed here, several candidate particles for new physics are accessible at the energy scales of the Large Hadron Collider. As mentioned previously, the search for an exotic decay of the Higgs boson presented in Chapter 5 is an attempt to generically uncover new physics that might be connected to the first-order electroweak phase transition of the early universe, and may be a part of the theory that describes the dark sector of matter. Going beyond the reaches of hadron colliders, the developments presented in Chapter 8 seek to improve the performance of quantum sensors that can aid in the search for new physics at ultralight mass scales.

The Author hopes that the interdisciplinary collaboration between communities that has made the research in this thesis possible grows with time, as searching for new physics across the vast phase space of all possible models of new physics will only get more challenging as theories for new physics continue to get ruled out by experimental observations — The quest for a more complete model of particle physics must continue, lest the immense potential of human curiosity and ingenuity be spent.

Chapter 3

The Large Hadron Collider

3.1 Introduction

The *Large Hadron Collider* [55–58] (LHC) is a two ring superconducting high energy particle collider located at the European Organization for Nuclear Research (CERN) in Geneva, Switzerland. It is the largest and most powerful particle accelerator and collider ever built, which started taking physics data in 2010 is expected to collect data till 2040. The LHC is housed within 3 m diameter tunnels dug 100 m under the ground, and covers a circular track of 27 km circumference. It was designed to facilitate head-on collisions¹ of two counter propagating high energy particle beams, with a total center-of-mass energy up to $\sqrt{s} = 14 \text{ TeV}$ ², at a few primary collision points along the accelerator. Four primary experiments currently operate around these collision points at the LHC: ATLAS [63], CMS [168], LHCb [169], and ALICE [170]. Additionally, several smaller and more specialized experiments operate around the other collision points [171–174]. Figure 3.1 shows the outline of the LHC along with the positions of the four primary experiments and where they are located in relation to the France–Switzerland border area near Geneva, Switzerland.

¹The beams collide at angles close to 200 mrad.

² s is the Mandelstam variable $s = (\mathbf{p} + \mathbf{p}')^2$, and $\mathbf{p}$ and $\mathbf{p}'$ are the four-momenta of the two colliding particles. With $E = E' \gg m = m'$, $\sqrt{s} = \sqrt{2E^2(1 - \cos\theta)} \approx 2E$ in the case of head-on collisions ($\theta \approx 180^\circ$), as is the case in the LHC.



Figure 3.1: Aerial photo of Geneva and nearby the France–Switzerland border area with an overlay graphic showing the LHC and the locations of the four primary experiments — ATLAS, CMS, LHCb, ALICE. Image sourced from Ref. [10], Maximilien Brice (CERN).

3.2 The Accelerators

The following describes the acceleration chain for protons before the long shutdown of the LHC in 2019-2020 which introduced several updates to the accelerators [175, 176]. The full acceleration chain is illustrated in Figure 3.2. The high energy particles first originate in a small tank containing Hydrogen gas. The Hydrogen molecules undergo ionization and the resulting protons are accelerated by the *Linear Accelerator 2* (Linac2)³ [178, 179] to 50 MeV ($\beta^4 \approx 5\%$). Protons are then injected in to the *Proton Synchrotron Booster* (PSB) [180, 181] (25 m in radius) and accelerated to 1.4 GeV ($\beta \approx 80\%$), and then into the *Proton Synchrotron* (PS) [180, 181] (100m in radius) which increases their energy to 25 GeV ($\beta \approx 99.9\%$). The final step in the acceleration chain comes at the *Super Proton Synchrotron* (SPS) [182] (7 km in circumference), which accepts the protons from the PS and accelerates them to 450 GeV before injecting the protons into the LHC (27 km in circumference). The LHC brings the counter propagating proton beams up to the nominal energies of 6.5 TeV ($\beta > 99.99999\%$).

The acceleration of the protons in these systems is achieved with radio-frequency cavities [183, 184] — the particles experience an accelerating potential at the vacuum gaps as they traverse the accelerator due to the synchronized oscillating potential. In linear accelerators, the particles pass through the equipment once and gain the requisite amount of energy. However in circular accelerators — cyclotrons, or the relativistic synchrotrons — the particles can cycle through the accelerating volume multiple times and gain energy with every revolution. It is important to note that as the energy of the particles increases, the perpendicular magnetic field strength required to maintain the particles' circular trajectory also increases [184]. This typically sets the total energy gain of a synchrotron as for a fixed radii the magnetic field strength of a given magnet can typically be varied only within a factor of 50, hence the full LHC accelerator chain contains the Linac2, PSB, PS, SPS, and finally the

³This first step was carried out using the Linac2 until the long shutdown 2019-2020, after which the Linac2 was replaced with the Linac4 [177]. The newer accelerator delivers particles at energies of 160 MeV. Furthermore, Linac4 accelerates H^- ions before stripping the extra electrons away using carbon foil, whereas the Linac2 accelerated protons directly.

⁴ $\beta = v/c$, where c is the speed of light and v is the speed of the particle in the inertial frame.

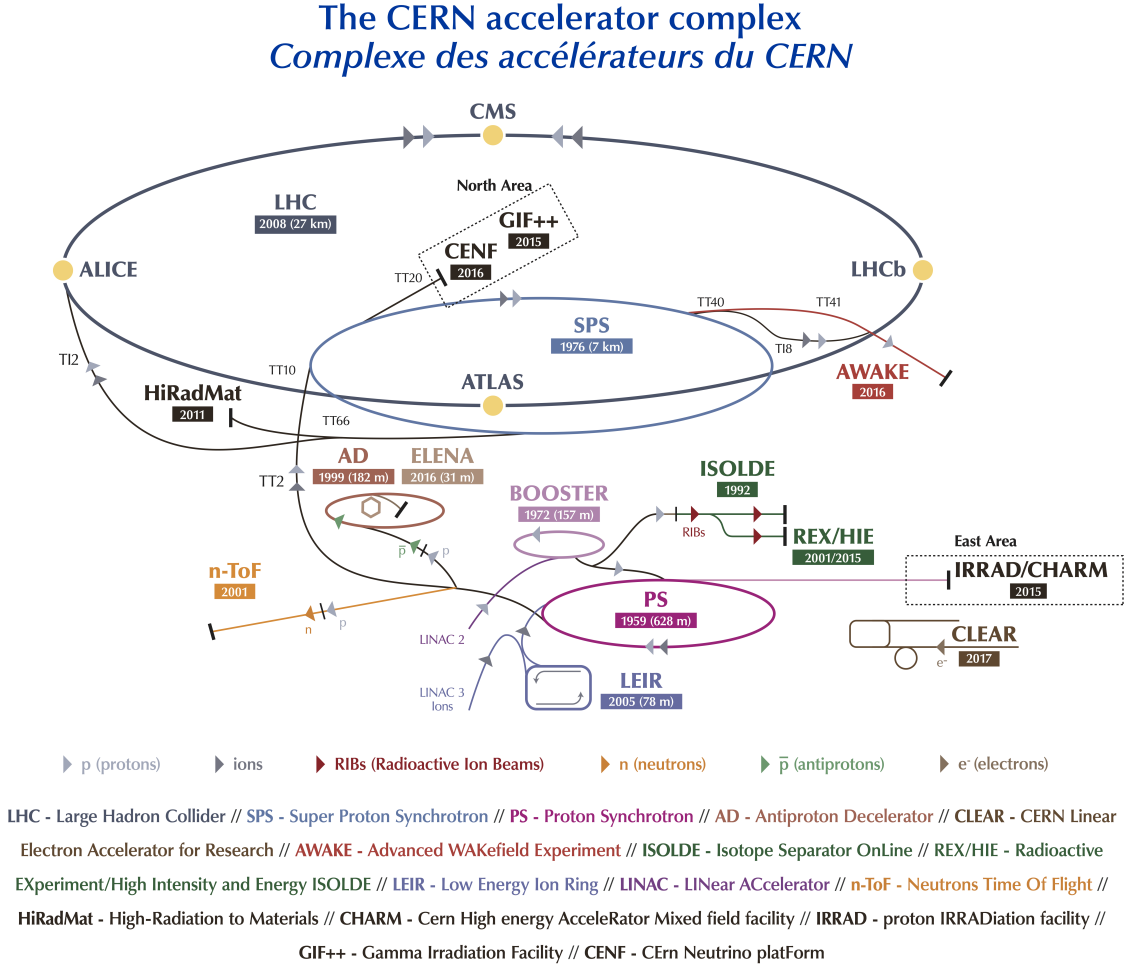


Figure 3.2: Layout of the CERN accelerator complex, August 2018. Image sourced from CERN Ref. [11].

LHC itself.

In the LHC eight cavities are used for the two beams, where each cavity operates at 4.5 K and delivers 2 MV at an RF frequency of 400 MHz. To ensure that the particles always experience an accelerating potential, the RF frequency in the synchrotrons is set at some integer multiple of the revolution frequency of the particles. This sets an upper limit on number of points along the circumference where the particles are in perfect synchrony with the acceleration. Regions of the circumference centered around these synchronous particles and with a size proportional to the corresponding RF wavelength are called *buckets*. Particles that are more or less energetic than the nominal energy will longitudinally oscillate around the synchronous particles inside their corresponding buckets [185]. This containment is a feature of the restorative forces resulting from the RF modulation of the accelerating potentials. The clumps of particles that form in the buckets as a result of the RF modulation are called *bunches*. This relation between the RF modulation, the bucket sizes and bunches is illustrated in Figure 3.3.

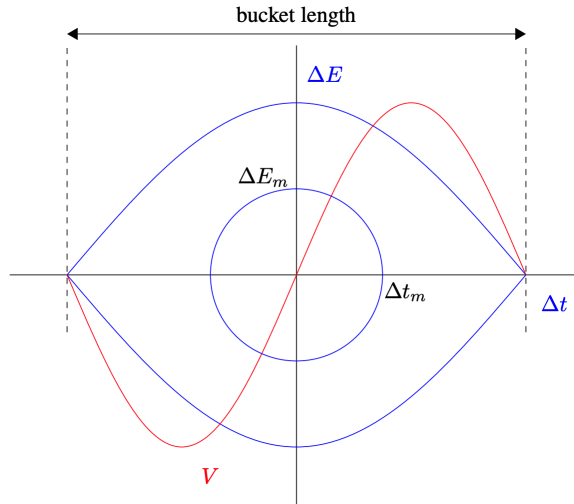


Figure 3.3: Phase-space representation of a bunch in its stationary bucket along with its RF potential function (in red and labelled V). The bucket size along the Energy-Time axes are labelled as ΔE and Δt , and the bunch size along the two axes are labelled as ΔE_m and Δt_m . Figure sourced from Ref. [12].

The maximum number of buckets that be simultaneously occupied in the LHC is

about 35640, assuming that the particles are traveling close to the speed of light. In practice, however, only 2808 buckets are occupied in the LHC, with particles arranged in bunch trains that have large bucket gaps between them. The empty buckets form an *abort gap* along the circumference of the accelerator and account for the time it takes to switch on the magnetic field needed to divert the particle beams from the LHC into the beam dump. Additionally, the bucket configuration also helps determine where the protons from the two counter-propagating beams will collide, since there are experiments at each of the many crossing points. Finally, only one-tenth of the available frequency is used in the LHC, resulting in consecutive bunches of about 10^{11} protons that are separated by 25 ns (corresponding to 40 MHz). This sets the bunch crossing and collision frequency that the various experiments are designed to record collision data at.

As mentioned before, perpendicular magnetic fields are required to maintain circular trajectories for charged particles inside synchrotrons. This is a consequence of the Lorentz force exerted by the magnetic field being perpendicular to the velocity of the charged particle, and hence providing the necessary centripetal force needed for circular motion [186]. These accelerating particles in-turn radiate (synchrotron radiation) [186], which requires constant input energy to maintain the beam at some steady energy. In the LHC, the magnetic fields required to accelerate the protons to 6.5 TeV and maintain their circular trajectories requires magnetic fields of strengths about 8 T. These large magnetic fields are generated using superconducting dipole magnets (simultaneously generating the fields for both beams) that utilize NbTi coils cooled down to 1.9 K. Synchrotron radiation scales as E^2/m^4 and thus is greatly suppressed for protons when compared to similarly energetic electrons. This radiation, however, poses a challenge for the superconducting magnets used in the LHC as they pose an additional cooling burden when the bunch sizes are very large. This is an important factor that constrains the number of protons per bunch to about 10^{11} .

3.3 Collision Statistics

The expected number of collisions, or interactions, that can be delivered is a key parameter in particle colliders. This is quantified using the (instantaneous) *Luminosity* $\mathcal{L}$ [187] and has units of inverse area per unit time. The rate of collisions can be found using the luminosity and the *cross-section* σ ,

$$\frac{dN}{dt} = \mathcal{L}\sigma. \quad (3.1)$$

The cross-section σ has units of area and is a measure of the probability of an interaction to occur as particles or excitations encounter localized effects such as density fluctuations or other particles. Cross-section is typically expressed in units of *barn*, where $1\text{b} = 10^{-24} \text{ cm}^2$. The maximum luminosity at the LHC is approximately $10^{34} \text{ cm}^{-2}\text{s}^{-1} = 10 \text{ nb}^{-1}\text{s}^{-1}$ [58]. This means that the rate of pp interactions at the LHC is about 10^9 s^{-1} , given that the total pp cross-section at 13 TeV is approximately 10^8 nb (Figure 3.4). It is important to recall here that protons in the LHC beams are focused into bunches separated by 25 ns. Hence these pp interactions occur as bunches from the counter propagating beams are collided at the crossing points, with detectors recording data from interactions occurring at each bunch crossing. Therefore the collision data is naturally quantized in units of *events*, where an event is the data generated from a single bunch crossing. Events at the LHC occur at the 40 MHz rate corresponding to the bunch spacing in the beams, and contain some average number of interactions per bunch crossing μ ; μ is thus a measure of the luminosity. From the numbers above it is clear that $\mu > 1$ at the LHC — greater than one pp interaction per bunch crossing leads to a phenomenon called *pile-up*.

The total number of interactions for a given process with cross-section $\sigma_{\text{pp} \rightarrow X}$ is $N = \sigma_{\text{pp} \rightarrow X} \int \mathcal{L} dt$, where the integral is done over some data taking period. $\int \mathcal{L} dt$ is called the integrated luminosity (expressed in units of inverse barns), and is a useful parameter for quantifying the statistics contained in the data collected over the data taking period. The full expression for the instantaneous luminosity is as

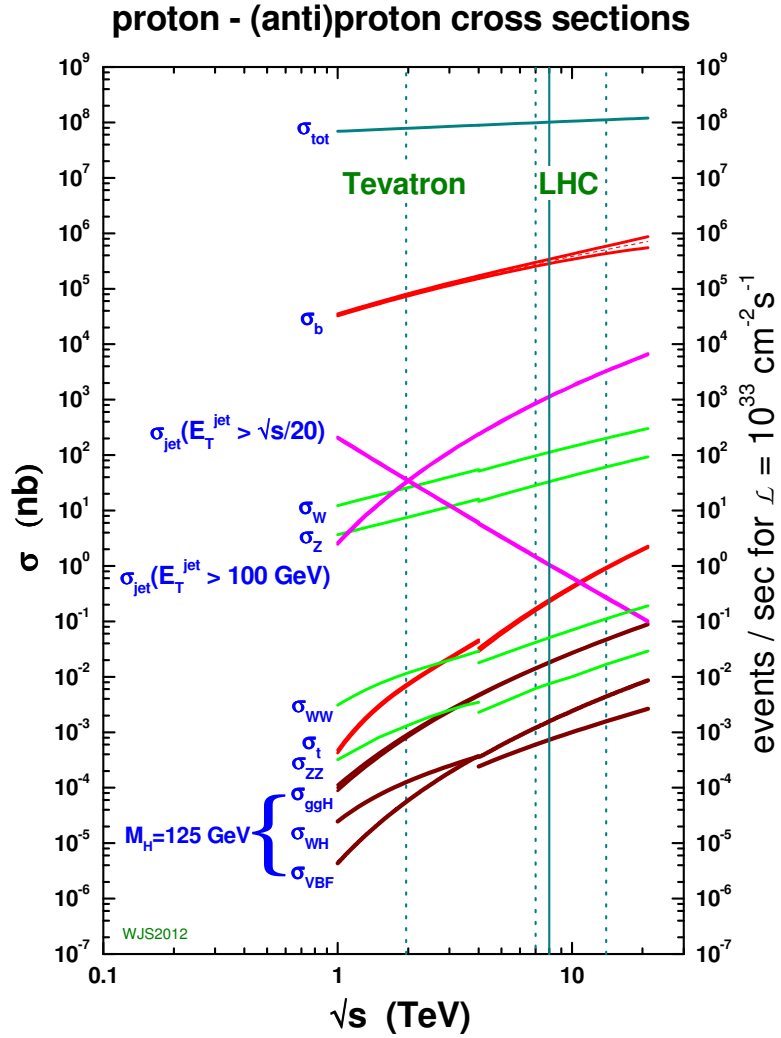


Figure 3.4: Cross-sections for pp interactions as a function of the center-of-mass energy ($\sqrt{s}$). Vertical lines corresponding to the various center-of-mass energies at Tevatron [13] and the LHC are drawn. Figure sourced from [14].

follows [185, 188, 189],

$$\mathcal{L} = \frac{N_b^2 n_b f_{\text{rev}} \gamma}{4\pi \epsilon^* \beta^*} F, \quad (3.2)$$

where N_b is the number of particles per bunch, n_b the number of bunches per beam, f_{rev} the frequency of revolution of the particles in the beam, γ is the relativistic factor of the particles, ϵ^* is the *normalized emittance*, β^* is the minimum of the *betatron function* β at the collision point, and F is an $\mathcal{O}(1)$ geometric factor for the angle of the beam crossing. The luminosity $\mathcal{L}$ is measured in several different ways using Van Der Meer Scans [190] and In-situ measurements [191, 192]⁵.

The transverse size of a bunch is given by $\sqrt{\beta\epsilon}$, where ϵ is the emittance and β is the betatron function. This is contextualized using the concept of bunches as discussed above. Not only do particles clump together longitudinally in bunches, they also undergo the oscillatory motion in the transverse direction to the beam direction, called *betatron oscillations*. This is due to presence of higher order magnets such as quadrupoles that are used in accelerators to focus the beam in the transverse directions [185, 188, 189]. For a single particle, this leads to an oscillatory motion in the transverse phase space (transverse position - momentum space) that traces an ellipse as the particle completes a revolution around the synchrotron. At any given point in the circumference of the accelerator, all the particles in a bunch populate the phase space with some density. Emittance ϵ is defined as the area of the ellipse that encloses a certain percentage of the particles in the transverse phase space, and the normalized emittance is defined as $\epsilon^* = \beta_{\text{rel}} \gamma \epsilon$ where β_{rel} is the relativistic velocity ratio in the direction of the beam. The β^* is the minimum of the betatron function, and is a measure of the distance from the point where the beam is focused to where its transverse size doubles. In the LHC, $\beta^* \approx 0.5$ m and $\epsilon^* \approx 3$ μm [193, 194]. Note that for $\sqrt{s} = 13$ TeV, $\beta_{\text{rel}} \gamma \approx 13000$, which gives $\epsilon \approx 0.2$ nm and the transverse size of the beam $\sqrt{\beta\epsilon} \approx 10$ μm .

⁵Here reporting measurements from the ATLAS experiment, where the luminosity is measured close to the interaction point using the total pp differential inelastic cross-section.

3.4 Past Campaigns and the High-Luminosity LHC

The LHC began its physics data collection at a center-of-mass energy of $\sqrt{s} = 7$ TeV in March 2010, and continued doing so through 2011. $\sqrt{s}$ was increased to 8 TeV for the 2012 data taking year, and the 2010 – 2012 data formed the first data campaign called *Run 1*. This run proved to be of historic significance as the ATLAS and CMS experiments were able to announce the discovery of the Higgs boson using the Run 1 data [47, 48].

The LHC luminosity, as measured by the ATLAS experiment, during Run 1 is shown against μ (average number of interactions per bunch crossing) and time in Figure 3.5.

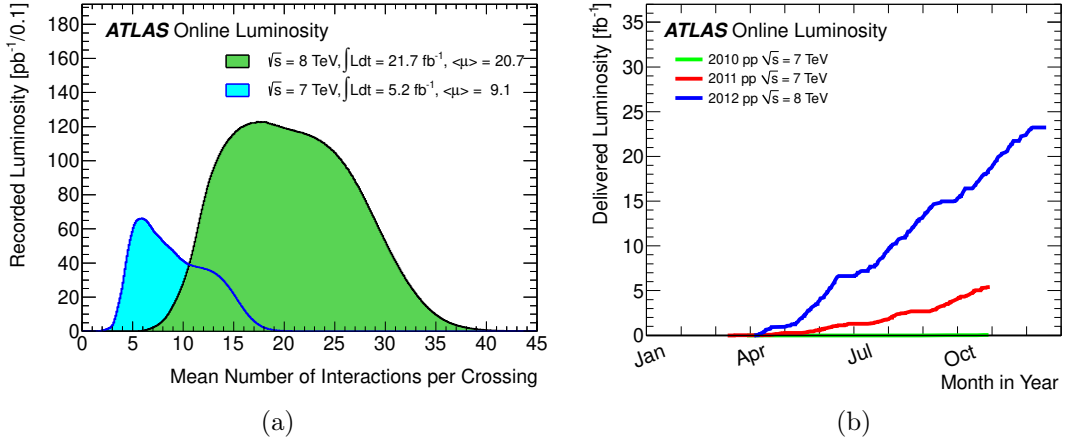


Figure 3.5: (a) Luminosity vs μ during Run 1. (b) Integrated luminosity vs time during Run 1. Figure sourced from [15].

The LHC resumed data taking for *Run 2*, after a long shutdown, with center-of-mass energy remained at 13 TeV from 2015 through 2018. The search for the novel Higgs boson decay presented in this Thesis (Chapter 5 uses this Run 2 dataset. Figure 3.6 shows the luminosity versus μ and time recorded by ATLAS during Run 2. Notably, the μ distributions at Run 2 were higher than in Run 1, with mode values close to $\langle \mu \rangle \approx 30$. The total integrated luminosity delivered to ATLAS during Run 2 was 156 fb^{-1} [16].

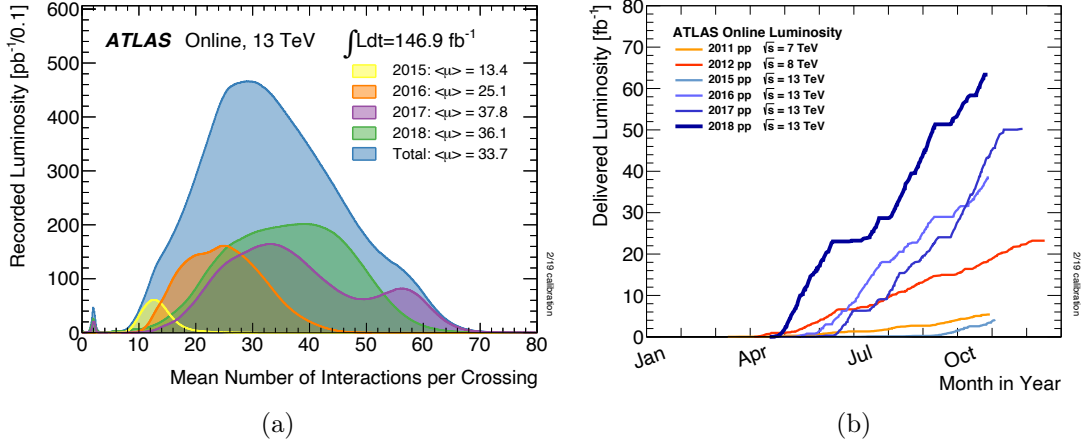


Figure 3.6: (a) Luminosity vs μ during Run 2. (b) Integrated luminosity vs time during Run 2. Figure sourced from [16].

At time of writing of this thesis, the LHC has restarted after another long shut-down, and is expected to start physics data taking in 2022. The new data taking campaign, *Run 3*, is expected to start in 2022 and go through the 2024 data taking period [195]. Run 3 targets a center-of-mass energy of 13.6 TeV and aims to deliver a total integrated luminosity of 250 – 300 fb^{-1} [196, 197] to ATLAS and CMS.

After Run 3, the LHC will shut down once again for major upgrades till until 2029. The next data campaigns for the LHC are planned from 2029 to 2032 during *Run 4*, and from 2035 to 2038 during *Run 5* [195]. During the planned long shutdown till 2029, the LHC will be transitioned to the *high-luminosity LHC* (HL-LHC) by means of upgrades to the entire accelerator chain [198, 199]. The upgrade, as the name suggests, aims to significantly increase the luminosity in the collider while keeping the same center-of-mass energy and is expected to deliver about 3000 fb^{-1} of data over the course of Run 4 and Run 5. The peak luminosity is targeted to be $5.0 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ by means of lower β^* and increased number of particles per bunch. Consequently, pile-up is expected to increase in the HL-LHC to $\langle\mu\rangle \approx 200$. This represents an order of magnitude increase over the Run 1/2/3 conditions in the LHC, and poses a challenge to all the detectors that operate at the interaction points around the LHC. This steep increase in pile-up will lead to technological and physics-reconstruction

related challenges that experiments will need to solve to unlock the physics potential of the increased statistics [200–204]. The work presented in this thesis in Chapter 6 and Chapter 7 directly relates to some of these challenges as they pertain to primary vertex reconstruction and jet calibrations in the ATLAS experiment.

Chapter 4

The ATLAS Detector

4.1 Introduction

As mentioned in Chapter 3, ATLAS [63] (*A Toroidal LHC ApparatuS*) is one of the four main experiments around the interaction points at the LHC. It is a general purpose particle detector with a close to 4π coverage (solid angle) around the interaction point.¹ ATLAS uses sophisticated hardware and software techniques to analyze the data generated from the pp interactions in LHC events and reconstructed the underlying physics interaction of interest.

The detector hardware is designed in several concentric layers (Figure 4.1) that each serve a unique reconstruction task and target particular types of particles. It has a set of tracking layers (Inner Detector, Section 4.2) that sit closest to the pp interaction point (IP) and are used to reconstruct the paths of charged particles through the layers. This is used to determine the locations of all the pp interactions that occurred in the event, along with measuring the momenta of the charged particles using the curvature of the tracks in its 2 T axial (solenoid generated) magnetic field.

¹A right-handed coordinate system is used in ATLAS, where the origin is at the nominal interaction point in the center of the detector. The z -axis points along the beam direction, the x -axis points from the interaction point to the center of the LHC circular track, and the y -axis is orthogonal to the two and points upwards. Cylindrical coordinates (r, ϕ) are used in the transverse plane (X-Y), where ϕ is the azimuthal angle about the z -axis. Pseudorapidity η is defined using the polar angle θ as, $\eta = -\ln \tan(\theta/2)$. Angular distance in the $\eta - \phi$ plane is measured in units of $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

The ID is surrounded by electromagnetic and hadronic calorimeters (Section 4.3) that are used to measure the energy of the particles. Finally, the titular toroidal magnetic field of 4 T along with accompanying trackers is used to detect muons and measure their momenta (Section 4.4). A diagrammatic overview of the interaction of various types of particles with the ATLAS subsystems is shown in Figure 4.2.

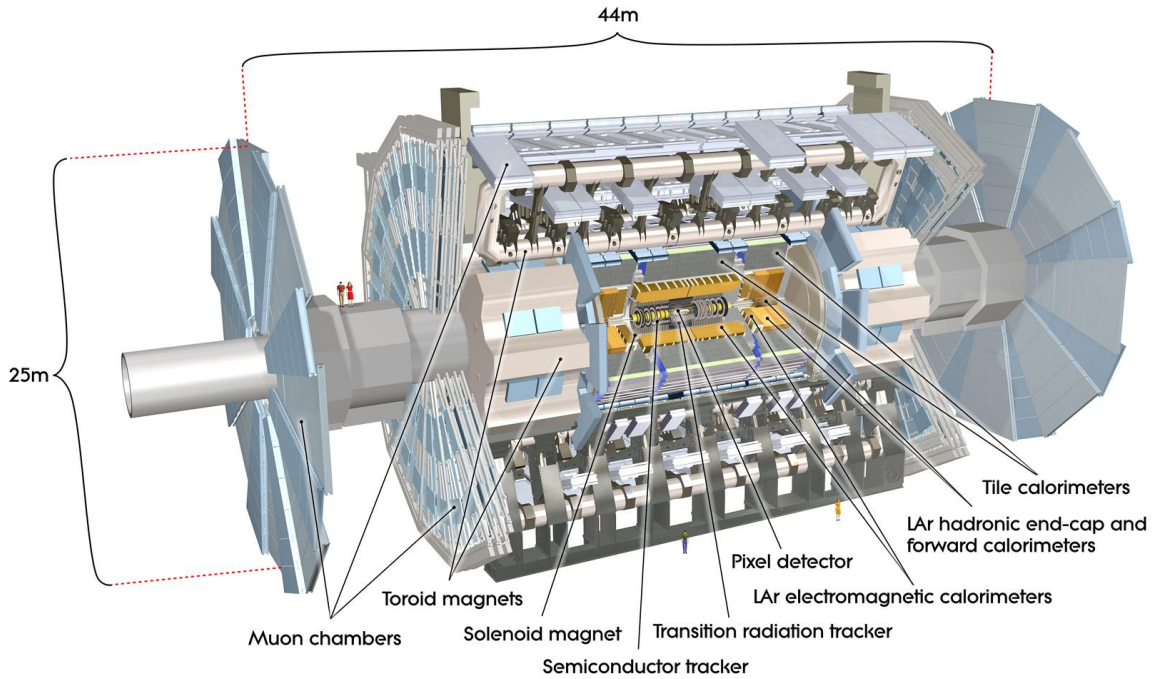


Figure 4.1: A computer generated view of the ATLAS detector, with the overall dimensions and detector subsystems labeled. Figures of people are drawn on the cavern floor and between the muon wheels for scale. Figure sourced from CERN [17].

A vital aspect of the experiment has to do with data rates and the detector's ability to save information from events. While the events at the LHC occur at 40 MHz, ATLAS is only able to save events at roughly 1 kHz corresponding to the maximum rate at which it can store information to disk. ATLAS utilizes a *trigger* system that lowers the data rate by monitoring the pp interactions and saving only a subset of events (Section 4.5).

Using the information returned by all the ATLAS components, reconstruction of

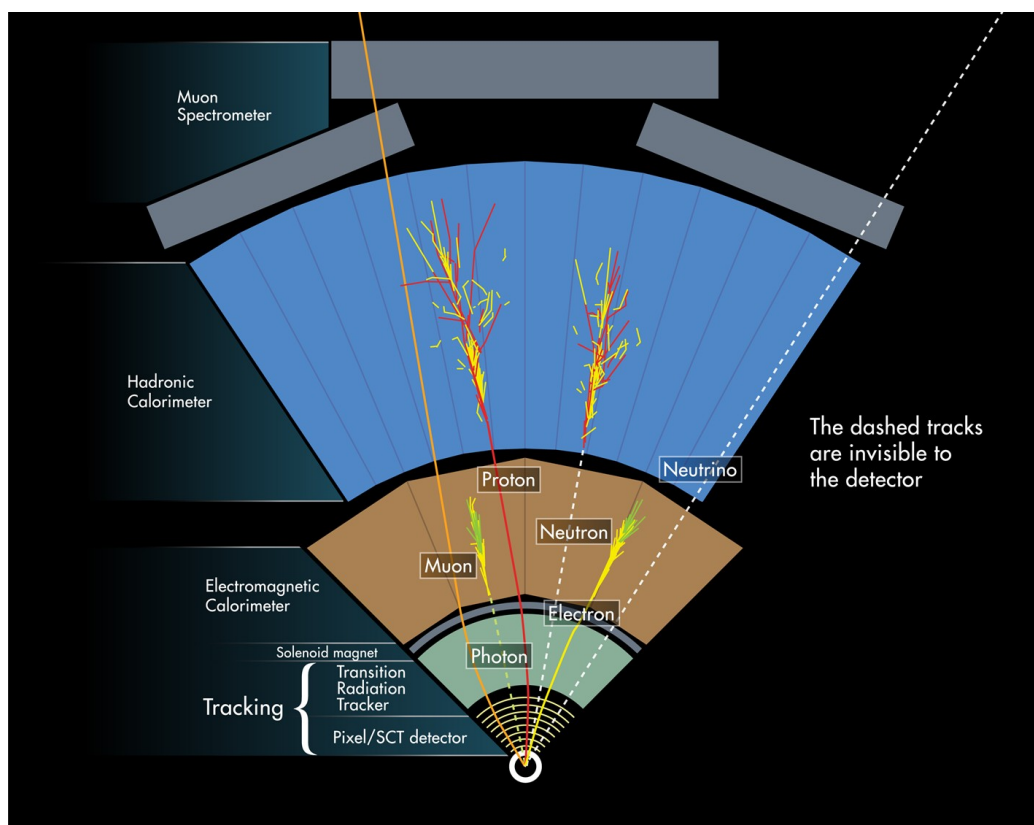


Figure 4.2: Diagrammatic overview of the interaction of various types of particles with the ATLAS subsystems. Dashed lines correspond to no interactions, and solid lines correspond to small or insignificant energy loss. Figure sourced from CERN [18].

individual (SM) particles that interacted with ATLAS is performed and their kinematics are recovered (Section 4.6).

Furthermore, a key element to the analysis of the data collected by ATLAS is the ability to simulate pp interactions under various hypotheses and understanding how these interact with the various detector elements. These physics interactions in pp collisions and their interactions with ATLAS are simulated using several different generators and a detailed particle-detector interaction simulation (Section 4.7).

4.2 Inner Detector

High energy pp collisions in ATLAS are expected to produce $\mathcal{O}(100)$ particles. By studying the secondary particles that are produced as these primary particles interact with the detector, information about the kinematics and type of the primary interacting particles can be learned. Additionally, a *passive* observation of these particles can be done so that the primary particles retain most of their energies as they traverse the detector. The ATLAS inner detector [205, 206] (ID), based on this concept, is a tracking detector that aims to help reconstruct the trajectories of charged particles as they cross this region of the detector. A set of observations of the particles trajectory are made as the charged particle interacts with successive layers of the inner detector. This trajectory is then fit to a high precision, from which the curvature of the track can be determined. Under an axial magnetic field that is perpendicular to the direction of transverse motion, the transverse curvature is related to the particle's transverse momentum, as²

$$\left(\frac{p_T}{\text{GeV}}\right) \approx 0.3 \left(\frac{B}{T}\right) \left(\frac{R}{m}\right), \quad (4.1)$$

where the axial magnetic Lorentz force has been equated with the centripetal force in the transverse plane. Here a single unit of charge (0.3 in natural units) is considered, and the transverse momenta p_T , magnetic field B , and trajectory transverse curvature R have been expressed in units of GeV, T, and m respectively. This constitutes a

²See Ref. [23] for details about unit conversions.

passive measurement since the magnetic field does no work, and the energy losses to the charged particles as they cross the various tracking layers is minimal. Note however this detection technique only works for charged particles, as neutral particles do not interact with the tracking volume.

The ID is made up of 3 distinct tracking technologies — silicon pixel detectors, strip semiconductor tracker (SCT), and transition radiation trackers (TRT). It has a solenoid provided 2 T axial magnetic field that curves the charged particle trajectories in the transverse direction [207]. The system is able to track charged particles with high precision for particles with $|\eta| < 2.5$. This central bias in the tracker is due to the lack of precision tracking layers in the forward ($|\eta| > 2.5$) region — this is planned to be addressed for the HL-LHC upgrade to ATLAS [208, 209]. A layout of the ID system with its various layers is shown in Figure 4.3.

The innermost layers of the ID are the pixel detectors and have the highest granularity of the tracking layers. It consists of layers of $50\text{ }\mu\text{m} \times 400\text{ }\mu\text{m}$ silicon pixels arranged in several cylindrical layers. An additional layer to the ones shown in Figure 4.3a was inserted at a radius of about 3 cm between Run 1 and Run 2 [210], as shown in Figure 4.3b. This layer (insertable B-layer or IBL) has smaller pixels with a size of $50\text{ }\mu\text{m} \times 250\text{ }\mu\text{m}$ each. Additionally the system also comprises of endcaps that are also covered in silicon pixels to complete the angular coverage on either side of the cylinders. These ionizing radiation detectors have a precision of about $10\text{ }\mu\text{m}$, and have approximately 92 million electronic channels.

The next layer out is the strip semiconductor tracker (SCT) which surrounds the pixel detector. This layer uses 6 million silicon strips as opposed to pixels, which provide precise positional information along the shorter dimension of the strip. The system is designed such that each charged particle interacts with at least 4 layers of the detector. This is accomplished by using 4 layers of double sided modules that have strips aligned axially on side, and offset by 40 mrad on the other side. The small offset in angles helps return an approximate measurement along the second dimensions as well. For the strips in the endcaps, the alignment is parallel to the radial direction. The SCT allows a measurement of the particle tracks with a precision of about $25\text{ }\mu\text{m}$.

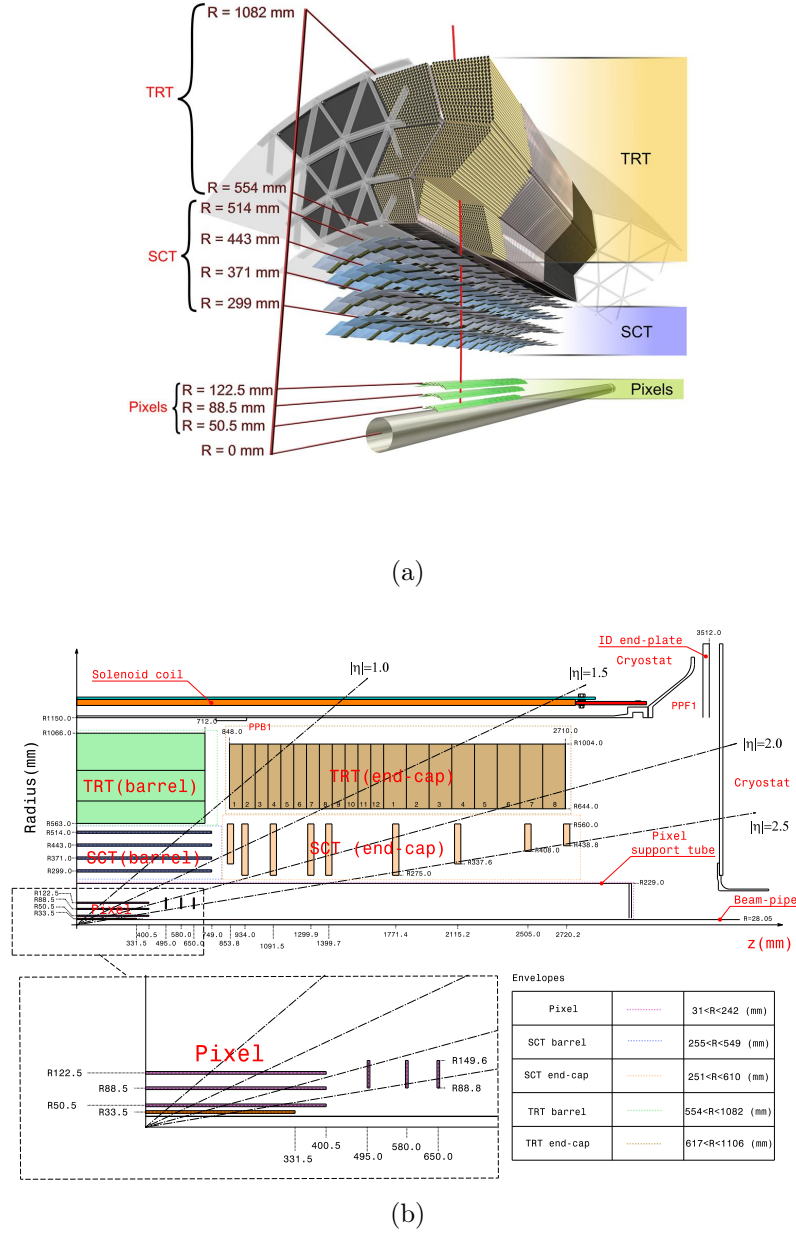


Figure 4.3: (a) Computer generated cross-sectional view of the ATLAS ID. Image sourced from CERN [19]. (b) Single quadrant view of the layout of the ID geometry in Run 2. The bottom panel shows a zoomed in view of the pixel region of the ID. The IBL, shown in red in the bottom left panel, was added for Run 2. The origin of the horizontal axis is the nominal interaction point, and the vertical axis is the radius measured from the origin. Figure sourced from [20]

The final layer of the ID is the transition radiation tracker (TRT). This is made using about 300,000 (4 mm diameter) thin-walled drift tubes (also called straws). These tubes are filled with a Xe-based gas mixture, and have a 30 μm gold plated tungsten wire in the center as the anode. The gas ionized as a charged particle travels through the tube — the signal is amplified by the potential difference (1530 V) between the center and the walls. Additionally, the transition radiation can distinguish between the types of particle that fly through the detector (like electrons, or pions, etc.). Each tube provides a drift time measurement which results in a positional precision of 170 μm per tube. The system is comprised of 50,000 tubes of length 144 cm stacked axially in the barrel, and 250,000 tubes stacked radially in the endcaps.

The signals from the different ID layers are combined in software to reconstruct the charged particle tracks through the ID and under the action of the axial magnetic field. This allows of the measurement of the charged particles charge as well as its momentum (Section 4.6.1). Additionally, the ID can measure the small amount of energy deposited per unit distance inside the tracking volume. This energy loss is a function of the mass of the particle, and thus can also be used to identify the particle type (pions, protons, kaons, etc.) [211].

4.3 Calorimeters

The calorimeter system in ATLAS [212] is an active measurement device — it stops the primary particles and measures the energy that these particles lost in the calorimeter. Owing to the multiple layers and fine granularity of the system, directional information about the particle can also be recovered. The system is able to perform measurements for both charged and neutral particles, unlike the ID. In fact, the association between measurements of the ID and the calorimeter system can help identify the type of particle and where it is charged or neutral. It is however unable to make measurements of muons and neutrinos, which are either not stopped in the calorimetric volume or do not interact with the calorimetric material. Likewise, the calorimeter has an angular reach of $|\eta| < 4.9$, which is also different (increased) when compared to the ID.

Calorimeters typically use alternating layers of absorbing or high-density *passive* material and detecting *active* material. The high-density absorptive passive material (like lead) is chosen such that it initiates the radiative/ionization/shower losses of the incoming particle's energy, which the active material then measures using ionization or scintillation. Ionization measurements rely on drifting and measuring the free electrons, whereas scintillation measurements use photo-multiplier tubes to sample the photon radiation.

The system consists of two major components: the liquid Argon (LAr) calorimeter [213], and the tile hadronic calorimeter [214]. Figure 4.4 shows the layout of the ATLAS calorimeter system. The LAr calorimeter is an electromagnetic calorimeter, which interacts electromagnetically with electrons and photons. The tile calorimeter interacts using nuclear and electromagnetic interactions, and thus stops hadronic particles.

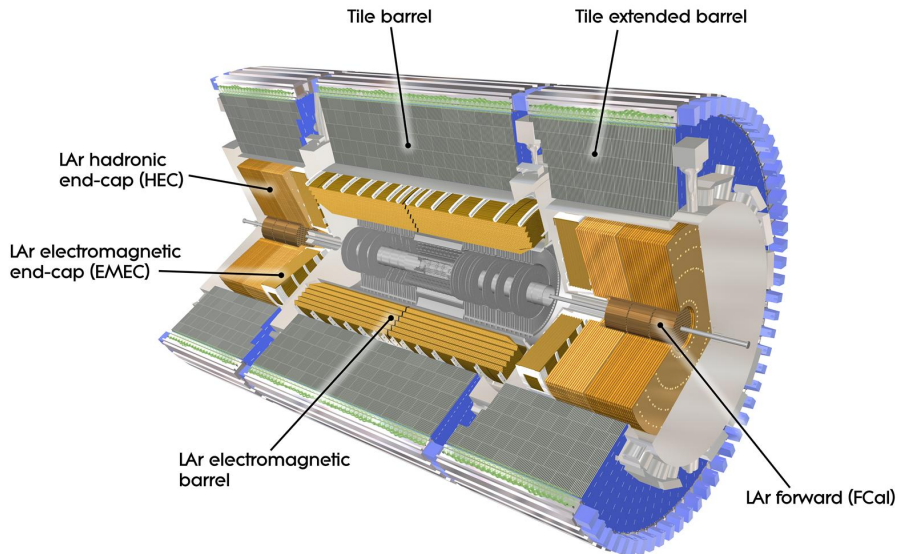


Figure 4.4: Computer generated image of the ATLAS Calorimeter. Image sourced from CERN [21]

Particles lose energy in calorimeters by means of cascading collisions and interactions. Photons and electrons that are stopped in the electromagnetic calorimeter

lose their energy due to electromagnetic showering. This is a phenomenon where high energy photons convert into electron positron pairs, and high energy electrons radiate photons (bremsstrahlung) as they decelerate due to material interactions [215]. Once the electromagnetic particles fall below a certain critical energy, ionization and other low energy effects dominate and the particles are ultimately absorbed. The high energy particles' loss of energy by electromagnetic showering can be described by,

$$\frac{dE}{dx} = -\frac{E}{X_0}, \quad (4.2)$$

where X_0 is the *radiation length*, and is a property of the calorimeter material. This implies that the energy loss by this showering results in an exponential decline in the particles' energy. Therefore, a calorimeter with a fixed depth that is many multiples of X_0 can successfully decelerate particles across a large range of energies to below the critical energy.

Showers initiated by nuclear processes are both deeper and transversely wider than showers initiated by electromagnetic particles [216]. Hadrons lose energy by means of ionization and nuclear (strong force) interactions — they interact with the material and create showers of pions, muons, etc. Their decay is characterized by a length scale called the *hadronic interaction length*. Typically λ is larger than X_0 for the same material [215]. For liquid Argon that is used in the ATLAS calorimeters, $\lambda \approx 6X_0$.

The ATLAS electromagnetic calorimeter uses liquid Argon as the active material, and lead as the passive material. It is split into a barrel component and endcaps, each having 3 layers in addition to a *presampling* layer. The presampling layer of the calorimeter is unique because it compensates for the loss of energy of the particles due to interactions with the ID. None of the other layers in the calorimeter system compensate for energy losses, hence reconstructed calorimeter objects require calibrations (Section 4.6.5.2). The system has a total angular reach of $|\eta| < 3.2$, and a thickness $> 22X_0$ throughout. These calorimeters (along with the LAr segments of the hadronic system) are housed in cryostats and cooled to -184°C . Since the barrel

of this system is the closest to the ID, it has fine segmentation to resolve electrons and photons with high precision. Additionally, small cells can help distinguish isolated photons from collinear photons (produced by a pion decay for example). The smallest cells in this system are in the barrel and are 0.025×0.025 in $\eta - \phi$ space.

The ATLAS tile hadronic calorimeter is composed of three sections: the barrel, the endcaps, and the forward calorimeters. Combined, the system has a total angular reach of $|\eta| < 4.9$, and has a thickness $\gtrsim 10\lambda$. The barrel section is made of 3 layers, and uses scintillating tiles as the active material, and steel as the passive material. The endcap sections are made of 4 layers each, and use liquid Argon as the active material, and copper as the passive material. Finally the forward sections are made of 3 layers each, and use liquid Argon as the active material, and copper as the passive material in the first layer and tungsten as the passive layer in the rest. This system has coarse granularity when compared to the LAr system; the smallest cells in the barrel and endcaps here are 0.1×0.1 in $\eta - \phi$ space. Furthermore in the forward region, the smallest cells are $3.0 \times 2.6 \text{ cm}^2$ in $X - Y$ space.³

An important quirk of the liquid argon cells used in ATLAS are their integration times. These cells take about 450 ns to collect the energy from ionization [217]. This is much longer than the bunch crossing time of 25 ns in the LHC, and presents a challenge in the read out of these cells. Cells are assigned to their correct bunch by fitting the ionization pulse shape. Errors in this assignment lead to erroneous energy deposits from one event adding to the energy deposits of a different event from a different bunch crossing — this phenomenon is called *out-of-time* pile-up⁴. In contrast to this, the tile cells in the hadronic system have a much faster integration time and do not introduce further pile-up [218].

³Measured here in spatial coordinates since the length scale starts diverging at very high pseudorapidities.

⁴*In-time* pile-up refers to other pp interactions in a given event (bunch crossing) that produce particles, tracks, energy deposits etc other than the pp interaction of interest in that event. Pile-up collectively refers to both out-of-time and in-time pile-up.

4.4 Muon System

Muons are leptons with a mass ~ 200 times the mass of the electron. Bremsstrahlung does not lead to significant energy losses for these heavy particles until their energy is above ~ 1000 GeV [216]. Therefore these particles do not undergo the same showering as the lighter electromagnetic particles and are able to traverse the calorimeter volume without significant interactions. Hence a specialized system is required to identify muons and measure their momenta, called the muon spectrometer [219]. This is the final radial component of the ATLAS detector. Figure 4.5 shows the layout of the ATLAS muon system.

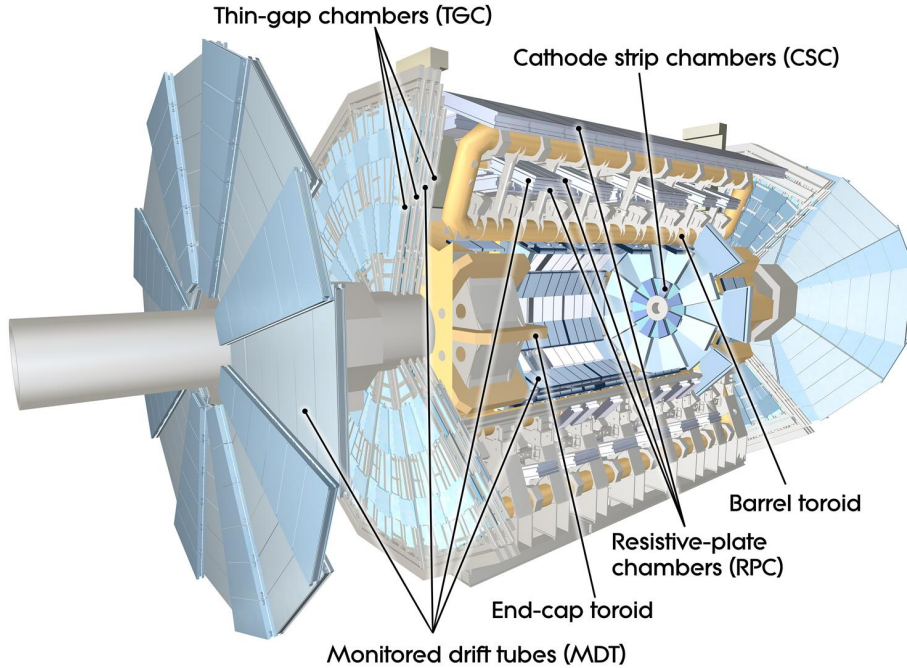


Figure 4.5: Computer generated image of the ATLAS muon system. Image sourced from CERN [22]

The muon system is made of 4000 muon chambers uses four distinct techniques: Thin Gap Chambers (TGC), Resistive Plate Chambers (RPC), Monitored Drift Tubes

(MDT), and Cathode Strip Chambers (CSC). It has an angular reach for muon tracking of $|\eta| < 2.7$, and for the muon triggering system of $|\eta| < 2.4$. The system is surrounded by a toroidal magnet which provides bending force along the axial direction of the barrel — eight toroids are arranged around the axis in the barrel about the beam axis [220], and two toroids are present in the endcaps [221].

The MDTs form the primary component of the muon system, and are responsible for the tracking of the muons. These drift tubes provide a resolution of $35\ \mu\text{m}$ in the z direction with no resolution in the radial direction, and cover an angular reach of $|\eta| < 2.7$. This choice of resolution is made to make sure the bending provided by the toroids are used to record the trajectory of the muons, which in turn is used to measure the muon momenta. This is done in conjunction with the hits in the ID (Section 4.2) for trajectory and momentum reconstruction. Like the LAr cells (Section 4.3), the MDTs also have read times that are longer than the bunch crossing time; the drift time in these tubes is $\sim 700\ \text{ns}$.

The CSCs provide positional measurements in the end-cap region ($2.0 < |\eta| < 2.7$) along both the z and the radial directions, with resolutions of $40\ \mu\text{m}$ and $5\ \text{mm}$ respectively. The measurements along both directions are needed to resolve tracks due to the higher particle flux and reduced bending from the magnets in the forward region of the muon system. These measurements are made possible because the CSCs are multi-wire proportional chambers with alternating layers of perpendicular strips.

The muon system also uses the RPCs and TGCs for triggering, in addition to the precise measurements from the MDTs and CSCs. These are meant to provide coarse but fast tracking information is used in the trigger (Section 4.5). The RPCs provide angular coverage in the barrel ($|\eta| < 1.05$) and the TGCs cover the endcap region $1.05 < |\eta| < 2.4$. the RPCs are gas filled parallel plate capacitors that are spaced with coarse granularity and provide a resolution of $10\ \text{mm}$ in both directions. These have shorter drift times than the MDTs and are hence useful for triggering. the TGCs are multi-wire proportional chambers like the CSCs, and have finer segmentation than the RPCs to better resolve tracks in the forward region due to higher particle flux and reduced bending from the magnets. The TGCs provide a resolution of approximately $2 - 7\ \text{mm}$ in both directions.

While the muon system is designed to primarily study muons, some extremely high energy hadrons can pass through the entire calorimeter system and reach the muon system. Additionally, fluctuations in the hadronic shower can also lead to some hadrons surviving out to the muon system. This hadronic leakage past the calorimeters is called *punch through* and can be measured using the muon system [222]. These measurements can then be used to correct the energy measurements derived from the calorimeter (Section 4.6.5.2).

4.5 Triggers

The ATLAS trigger system plays a critical role in data acquisition and the interplay between the hardware and software [2, 223, 224]. Due to the 40 MHz bunch crossing rate, the detector interacts with a large number of particles at a high rate. The full data generation rate is about 100 TB/s, on account of every event occupying a few MBs of disk space [225]. Since this is an unfeasible data rate, not every event can be saved by the ATLAS software infrastructure. To mitigate this, ATLAS employs a trigger system that saves a subset of all events that occur in the detector. The system picks events to save based on the presence of interesting physics topologies, as can be reconstructed from the various detector signals. This is accomplished using a two-tiered structure: the first is a hardware based trigger (Level 1 or L1) [226] that reduces the rate from 40 MHz to 100 kHz, and the second is a software based trigger (High Level Trigger or HLT) [227] that reduces the rate from 100 kHz to 1 kHz. Figure 4.6 shows the layout of the ATLAS trigger system

The L1 trigger uses information from two subsystems: the calorimeter (L1Calo) and the muon system (L1Muon). It also uses topological information from the L1Topo system that combines the inputs from L1Calo and L1Muon to compute event level information like transverse momentum imbalance. L1Calo uses coarsely segmented information from the electromagnetic and hadronic calorimeters (Section 4.3) to build trigger level calorimeter objects that correspond to electrons, photons, taus, jets, etc. The L1Muon system uses RPCs and TGCs (Section 4.4) to build muon candidates for the trigger decisions. The L1 trigger system does, however, lack any tracking related information since the computation time needed to reconstruct tracks from the hits in the ID (Section 4.2) exceeds the time the L1 system has to make decisions (roughly 2.5 μ s). The triggers at L1 look for a certain number of detector objects that have properties exceeding some threshold — these thresholds control the data rate that the L1 system passes on to the HLT.

The HLT system has a decision time of $\mathcal{O}(1 \text{ s})$, and can use reconstruction algorithms that are closer to offline⁵ algorithms for higher fidelity reconstruction than L1

⁵ATLAS reconstruction algorithms are split into two broad categories: *online* algorithms which are run in the trigger and are limited by the trigger resources, and *offline* algorithms which are run

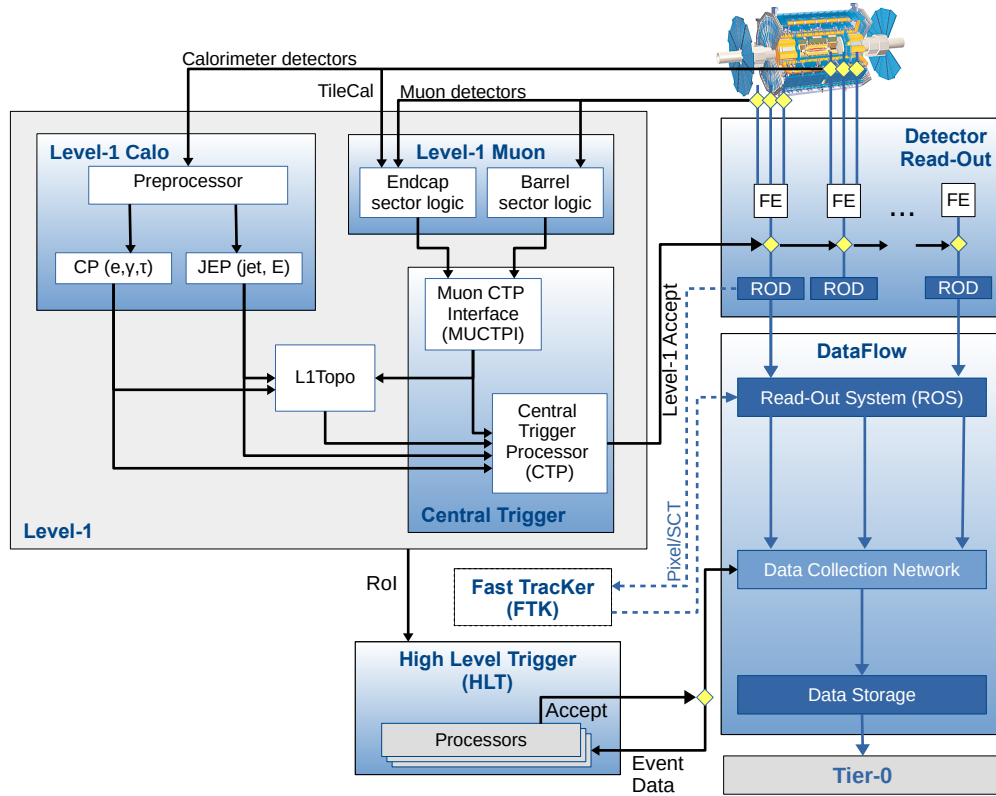


Figure 4.6: The ATLAS trigger and data acquisition system in Run 2. Trigger and read-out related components are being shown in this figure. Figure sourced from [2]

algorithms. It has access to tracks reconstructed using a fast tracking algorithm on hits in the tracker in regions of interest identified at L1. It also has access to higher granularity information from the calorimeters and the muon systems. Additionally, the HLT can look for heavy hadron decays as well as long lived particles since these manifest as vertices displaced from the beam axis and can be reconstructed using tracks. Consequently, the HLT has a high efficiency for most physics processes of interest as it can look for a myriad of topologies using the HLT objects.

The various triggers used are organised into a *trigger menu*, which contains the set of all the trigger objects and their corresponding thresholds. The menu is driven by the physics objectives of ATLAS and set based on the LHC delivered luminosity and pile-up [228–231]. The trigger system also utilizes redundancy to ensure that the efficiency of the system is high for processes of interest; An event can pass multiple triggers, but only one is required to save the event to disk. Despite its efficacy at reducing data rates, some triggers require a prescale to keep the data rates within acceptable limits — implemented by randomly rejecting a known fraction of events, where the known fraction is then used to correct for the observed luminosity. Furthermore, ATLAS conducts some special runs with exceptional trigger menus, such as the minimum-bias trigger which saves events with minimal selection criteria but with a very high prescaling.

The trigger system is evidently fundamental to every analysis, since the detector sensitivity to the physics process under considering is defined by the trigger efficiency for that process. For example, the analysis presented in Chapter 5 uses triggers that look for two muons that are compatible with a Z boson decay since the analysis uses Z boson associated production of the Higgs boson.

on events that have been saved to disk and are not subject to the resource limitations of the trigger.

4.6 Object Reconstruction

The next step in the ATLAS analysis pipeline, after an event passes the ATLAS triggers or is generated using a simulator, is to reconstruct physics objects. This is achieved by a variety of algorithms, some of which will be detailed in the upcoming sections. The reconstruction starts at the lowest levels, with tracks being reconstructed from hits in the ID (Section 4.6.1) and clusters being built from the calorimeter cells (Section 4.6.3). Tracks are clustered into their corresponding interaction points called vertices, and the primary hard-scatter vertex in the event is selected (Section 4.6.2). From these objects, electrons, photons, taus, and jets are reconstructed (Sections 4.6.6, 4.6.7, 4.6.5). Additionally muons are reconstructed from hits in the muon system along with the ID (Section 4.6.8). Finally, the information is combined to produce event-level observables such as missing transverse momentum to account for non-interacting particles such as neutrinos which are *invisible* to the detector (Section 4.6.9).

4.6.1 Track Reconstruction

Tracks are trajectories of charged particles in the detector. These are reconstructed using the hits in the different layers of the trackers in the ID system [232, 233]. Hits in the tracking layers (pixel layers and SCT layers) are treated as individual three dimensional space-points using the location of the hit and the geometry of the trackers.⁶ Sets of three collinear space-points from the pixel or SCT layers are then used to form track seeds, which are then grown into full track hypotheses (track candidates) using Kalman filtering to account for hits from successive layers [234] — this is an example of an *inside-out* algorithm. Each additional hit that gets incorporated to into the track candidate updates the track parameters and brings it closer to the true trajectory of the charged particle. Ambiguities in the hit-track matching can arise, particularly due to hits that may be compatible with more than one track candidate. These ambiguities are resolved using a machine learning model that can flag clusters

⁶Clustering is done to merge nearby hits in the layers from the same particle interacting with adjacent pixels, or to account for more than one particle interacting with nearby pixels.

likely to have arisen due to more than one track [235]. Track candidates that have their ambiguities resolved are then extended using the hits from the TRT layers of the ATLAS tracking system.

Once the track candidates are built, their properties can be studied in the context of the charged particles that corresponds to the tracks. As mentioned previously, the ID has a 2 T axial magnetic field that curves the trajectories of the charged particles. The curvature $\frac{1}{R}$ of the trajectory is related to the particle momentum p , charge $|q|$, and the axial magnetic field B as [184],

$$\frac{1}{R} = B \frac{|q|}{p}. \quad (4.3)$$

Therefore, the curvature of the track measures the q/p (charge - momentum ratio) for the particle, and direction of curvature relates to the sign of the charge of the particle. Furthermore, the distance of closest approach (impact parameter) of the track to the beam axis is a useful observable. It is split into the transverse impact parameter d_0 , and the longitudinal distance between the origin (nominal collision point) and the point of closest approach z_0 . d_0 is useful for heavy flavor hadron tagging [236–238], and z_0 is used to cluster tracks into vertices (Section 4.6.2). Additional track properties include the angular direction of the track after the ID, parameterized using the angular coordinates η and ϕ . Figure 4.7 shows the tracks of same-signed charged particles in an axial field like the ATLAS ID system.

Track candidates, built in the way described above, must pass a set of basic quality criteria designed to allow tracks that are likely to correspond to real charged particles that interacted with the ID and with the rest of the ATLAS detector. This criteria requires the track $p_T > 500$ MeV, since tracks with less p_T would not escape the ID volume, as visualized in Figure 4.6.1. Furthermore, tracks are required to have a minimum number of hits in the pixel and SCT layers, a maximum number of missing hits in the pixel and SCT layers, along with a maximum number of ambiguous hits [239]. Figure 4.8 shows a display from an event that ATLAS recorded in 2015, where the hits in the ID and the reconstructed tracks have been illustrated.

The resulting tracking efficiency of the ATLAS ID is $> 80\%$ [233], and the

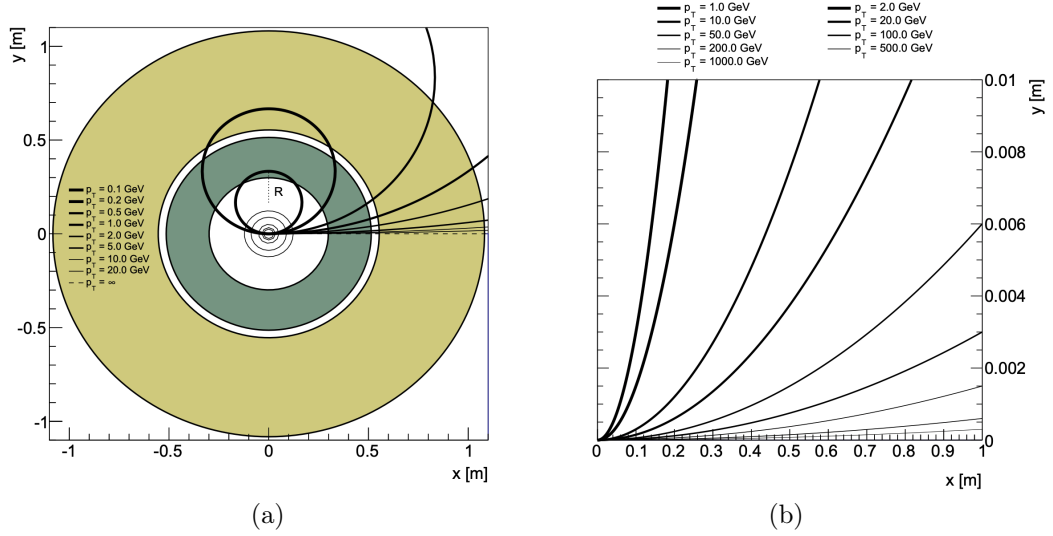


Figure 4.7: Transverse cross-section of detectors showing the tracks of charged particles in a 2 T axial field. Figure on the left has various sections of the ATLAS ID barrel system (Section 4.2) shaded in different colours for illustrative purposes. Tracks with transverse momenta spanning two orders of magnitude are shown to visualize the effect on the curvature of the tracks. Figure sourced from [23].

resolutions on the track parameters like momentum (about 1%), d_0 , and z_0 (less than 200 μm) achieve the system design goals across a range of track parameters [232, 239, 240]. It is important to note here that the momentum resolution of the tracks degrades as the track momentum increases, since this results in a less curved track — higher magnetic fields or larger lever arms (tracking depth) are needed to recoup the resolution loss at high track momenta [23].

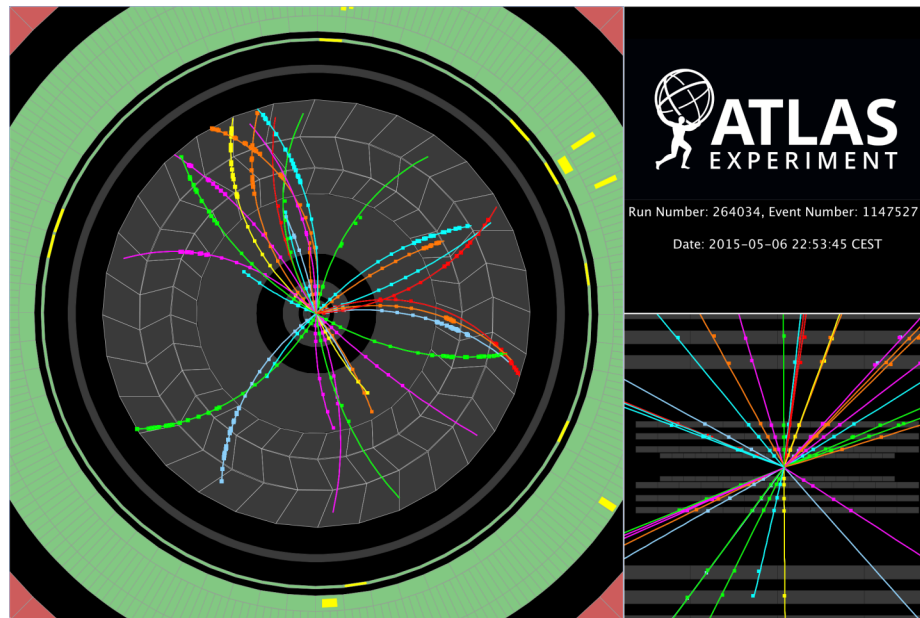


Figure 4.8: Display of a pp event that ATLAS recorded in 2015. Hits in the various layers of the ID are shown along with the reconstructed tracks. Figure sourced from CERN [24].

4.6.2 Vertexing

Vertices correspond to locations of particle interactions that produce daughter particles which interact with the detectors. Once tracks are reconstructed with the ATLAS ID (Section 4.6.1), vertex reconstruction can be done using the set of track candidates that pass quality criteria [25, 27, 241]. This reconstruction can be done by clustering tracks using their points of closest approach — the fundamental concept being that tracks originating from the same interaction all emanate from the same point in 3D space. Vertices can also be split into two broad categories: Primary vertices and secondary vertices. Primary vertices correspond typically to tree-level inelastic interactions (*hard-scatter*) or soft-QCD interactions that occur within the beamspot (Gaussian spread of protons in the beams) centered at the nominal collision point of the LHC beams. Note that interesting physics usually relates to interactions like tree-level inelastic collisions, which have a much smaller cross-section than soft-QCD interactions (See Figure 3.4). Secondary vertices correspond to decays or interaction points of long lived particles that are able to physically separate from the beam axis during their time of flight but can still be located within the ID volume — these provide a critical handle for heavy hadron tagging [242]. The performance of the vertex⁷ reconstruction in ATLAS is summarized in Refs [25, 27, 241, 243].

Primary vertex reconstruction is done by an iterative algorithm (Iterative Vertex Finder). The algorithm starts by finding the seed for the first vertex and then fitting for the position of this vertex by iteratively weighting the tracks and deriving the vertex position from the tracks. Incompatible tracks are then removed, and new vertex seeds are found with leftover tracks and fit one-by-one till all the tracks are associated to a vertex. The algorithm also rejects vertex candidates that have fewer than two associated tracks. When the iterative procedure has converged, a set of vertices with positions and covariances along with their corresponding set of associated tracks is returned. The performance of the algorithm can be gauged by studying the efficiency of reconstructing a vertex as a function of the number of tracks associated to it, as well as the distribution of Δz (distance along beam axis) between pairs of vertices.

⁷Unless otherwise stated, the terms *primary vertex* and *vertex* will be used interchangeably in the remainder of this text.

This is shown in Figure 4.9. Note here that the vertex reconstruction efficiency is $> 95\%$ for vertices with more than two tracks (Figure 4.9a). The algorithm also has a finite z resolution as it fails to reconstruct vertices that are closer than about 1 mm, resulting in *merged* vertices that have associated tracks from all the interactions that have been merged into one reconstructed vertex. Note also that the algorithm can result in *split* vertices where the tracks from the same interaction are reconstructed into two separate nearby vertices. The vertex merging and splitting effects can be visualized in Figure 4.9b, where the drop in the distribution corresponds to merging below 1 mm, and the spike close to $\Delta z = 0$ corresponds to splitting.

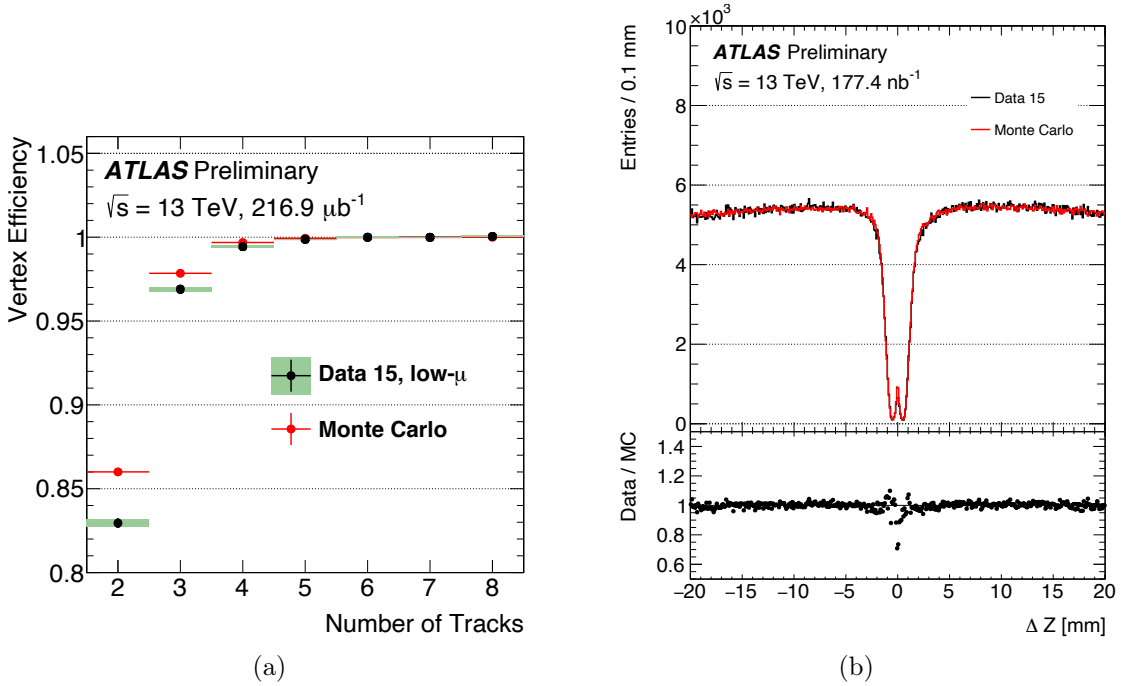


Figure 4.9: (a) Vertex reconstruction efficiency versus number of associated tracks to the vertex, measured in low μ data and Monte Carlo (MC) simulations. (b) Histogram of Δz values computed over all pairs of vertices in same event, measured in high μ data and MC. Figures sourced from [25].

As discussed previously (Section 3.3), due to pile-up the expected number of interactions $\mu > 1$ in every event, and therefore more than one primary vertices are reconstructed by the vertex finding algorithm. The total number of reconstructed

vertices in the events scales non-linearly with μ due to the aforementioned effect of vertex merging — the effect grows as the density of interactions grows in the regions of highest luminosities. The non-linearities can also be attributed to other reconstruction inefficiencies from several factors [27]. To address the challenges of the IVF algorithm, especially when $\mu \gg 1$, ATLAS has developed a novel adaptive vertex finding algorithm (Adaptive Multi Vertex Finder) [26] that has better reconstruction efficiency at high μ as well as better Δz resolution (Figure 4.10). The AMVF algorithm enables tracks to have weights in the vertex fits of more than one vertex. During the course of the convergence of the algorithm, tracks evolve to have non-trivial weights with only one compatible vertex. Furthermore, the vertex candidates adjust their positions to compete for the tracks. This new AMVF algorithm will be the default algorithm that ATLAS uses from Run 3 onwards.

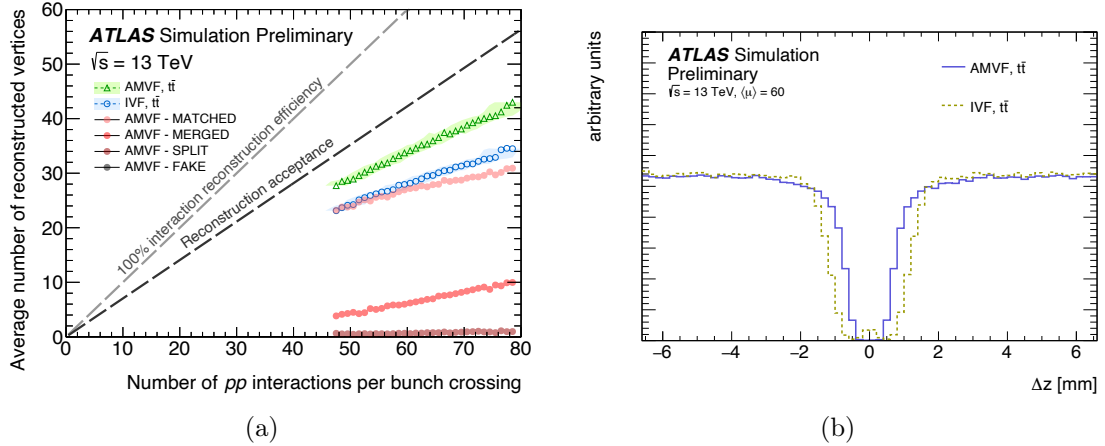


Figure 4.10: (a) Average number of reconstructed vertices versus the average number of pp interactions (μ), measured in Monte Carlo (MC) simulations. Note here the AMVF performance has been shown for the various types of reconstructed vertices, though only the comparison between AMVF and IVF inclusive efficiency is relevant here. (b) Histogram of Δz values computed over all pairs of vertices in same event, measured in MC. Figures sourced from [26].

Given that each event has multiple reconstructed vertices when pile-up is present, the hard-scatter vertex in the event needs to be identified since this is the interaction of interest that likely passed the trigger and led to the event being saved. The hard-scatter vertex is generally identified in ATLAS using the $\sum p_T^2$, which is the sum

of the track p_T squared over all the tracks associated to a vertex; the vertex with the maximum $\sum p_T^2$ in the event is identified as the hard-scatter primary vertex. All the other vertices and associated tracks are pile-up and typically vetoed in most physics analyses. This selection picks the hardest interaction in the event, likely to be the interesting physics interaction that passed the trigger. This is illustrated in Figure 4.11 where the $\sum p_T^2$ for various physics interactions is compared with the minimum-bias data collected during Run 1; Here minimum-bias data serves as a proxy for soft-QCD pile-up since in the absence of any biasing triggers the cross-section is dominated by soft-QCD interactions (Figure 3.4).

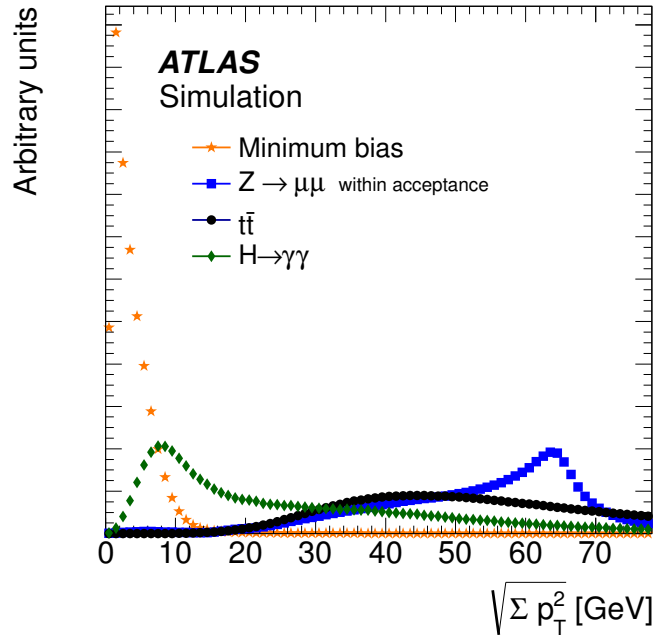


Figure 4.11: The distribution of $\sum p_T^2$ from primary vertices, normalised to the same area. Various simulated hard-scatter processes are compared with a minimum bias sample from Run 1. Figure sourced from [27].

Chapter 6 of this thesis covers an extension to the $\sum p_T^2$ selection criteria that is better suited for physics interactions with low visible track p_T in the detector, where the jets can be used to identify the hard-scatter vertex. Another exception to this selection criteria occurs in the analysis of the Higgs boson decays to diphotons and missing transverse momentum, where the hard-scatter interaction has very low visible

p_T but the diphoton objects can be used to identify the correct primary vertex [244, 245].

4.6.3 Calorimeter Cell Clusters

Energy deposits in the calorimeter cells (Section 4.3) are clustered together using a topological clustering algorithm [246] to recover the energy deposits of the single particles that likely sourced the shower and deposits in nearby cells. The clustering algorithm relies on an estimate of the noise in the cells. The expected noise in a given cell can come from electronic noise, as well as energy deposits from soft pile-up interactions. Therefore the expected noise in a cell depends on the location of the cell (calorimeter layer, η) and the expected number of interactions in the event (μ). For center-of-mass energy of 14 TeV and $\mu = 80$, this noise can be as low as 50 MeV for the electromagnetic calorimeter in the $|\eta| < 0.5$ region, and as high as 10 GeV in the forward calorimeter for $|\eta| > 4.5$ [247].

The clustering algorithm starts by assigning cells with energy exceeding 4σ above the expected noise σ as seeds. Then, it iteratively grows the clusters around the seed by including nearby⁸ clusters that have energy exceeding 2σ . Finally the cluster is grown further to include all nearby cells with positive energy.⁹ Additionally, the connected clusters can be split if there are local energy maxima within the clusters. This procedure results in the final set of 3D topologically connected clusters (*topo-clusters*). Each cluster is assumed to be massless when forming the four-momentum, and is assigned the total energy of the constituents and barycenter angular coordinates (η, ϕ).

The calibration used to correct the energy of each of the the topo-clusters is a local cluster weighting (LCW) scheme. Based on the electromagnetic versus hadronic shower likelihood of the entire cluster, a weighted average of the electromagnetic and hadronic corrections is applied to each cell of the cluster. The corrections account for energy lost in inactive materials, energy deposited in cells outside the cluster, and for the differences between the electromagnetic and hadronic calorimeter responses.

⁸Neighboring clusters if in the same layer, close-by in η - ϕ space otherwise

⁹Pulse-shape calibrations in the calorimeter cell read out can lead to cells with negative energy.

After the corrections (weighting) for the cells are determined, the entire topo-cluster four-momentum is recalculated using the LCW weighted cells.

4.6.4 Particle Flow Objects

Particle flow objects [28] are used at ATLAS to improve the low energy resolution of objects that utilize topo-clusters (Section 4.6.3). This is done by combining tracking information from the ID (Section 4.6.1) with the topo-clusters from the calorimeters. These particles then deposit energy in the calorimeter and may contribute to topo-clusters. The particle flow algorithm performs an overlap removal process, by removing the expected energy deposits of the charged particles from the topo-clusters. The schematic of the particle flow algorithm is shown in Figure 4.12. The algorithm iteratively removes the expected energy deposits of the selected tracks from the set of relevant topo-clusters on a cell-by-cell basis. The output of the algorithm consists of a set of charged particle tracks, along with the modified and unmodified topo-clusters. The primary benefit of the particle flow objects at low energies is driven by the enhanced momentum resolution for charged particle tracks in the ID as the curvature in the magnetic field is greater. At higher energies, the topo-clusters have superior energy resolution since the significance of the energy deposits is large compared to the stochastic noise floor. Therefore this combination of information from the ID and calorimeter systems results in objects that provide better overall energy resolutions — The combination of the information from these subsystems of the ATLAS detector is also relevant for the work presented in Chapter 6.

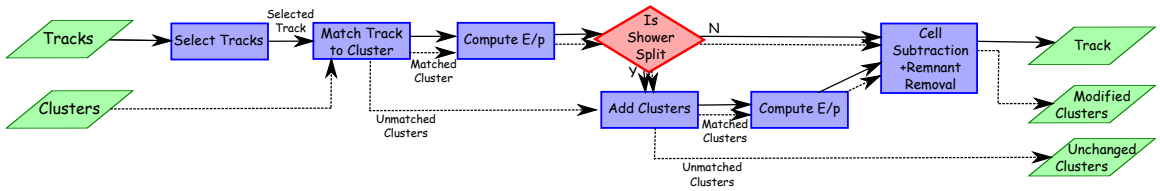


Figure 4.12: Particle flow algorithm flow chart. The algorithm starts with track selection and concludes when the energy of the selected tracks has been removed from the calorimeter. The algorithm outputs topo-clusters that have both been modified and unmodified by the action of the algorithm, along with the charged particle tracks. Figure sourced from [28].

4.6.5 Jets

4.6.5.1 Introduction and Definition

Jets are the collimated spray of particles detected in the ATLAS detector as a result of hadronization of particles as they travel to the detector. These could be sourced by quark or gluons, which are produced in $\mathbf{pp}$ collisions in large numbers; As these particles travel, color confinement leads to the fragmentation and hadronization of these partons. Furthermore, massive particles with finite lifetimes may also decay as they travel to the detector, resulting in the detection of their hadronic products rather than the massive particles themselves. It is crucial for the physics done at $\mathbf{pp}$ colliders that the properties of the initiating particles or partons are understood from these hadronic particles that ultimately interact with the detector. Consequently, clustered objects called *jets* are introduced to determine the energy-momentum of the initiating particles from the collimated spray of particles [62, 248–250]. These objects play a vital role in the analysis presented in Chapter 5, and a novel calibration technique for jets is introduced in Chapter 7.

Jets are built using sequential recombination algorithms [250] that group a set of seeds into an output set of jets. The seeds used by jet algorithms can be topoclusters (Section 4.6.3), tracks (Section 4.6.1), or particle flow objects (Section 4.6.4). In the case of the analysis presented in Chapter 5, jets reconstructed using particle flow objects are utilized. The seeds serve as unit particles¹⁰ with their transverse momentum p_T and angular positional coordinates (y, ϕ) .¹¹ Figure 4.13a illustrates the seeds in the y - ϕ plane for a simulated event.

The recombination algorithms cluster seeds into jets in an infrared and collinear safe manner (IRC safe) [251] — this is done to ensure that (potential theoretically unbound) soft emissions and (diverging in perturbative QCD) collinear splittings of particles do not affect the output of the algorithm. These algorithms work by iteratively combining seeds into a set of jets, based on a distance metric — each step

¹⁰Often assigned the mass of a pion for four-momentum operations.

¹¹Note the use of the rapidity $y = \frac{1}{2} \ln \frac{E+p_z}{E-p_z}$ rather than pseudorapidity η . Differences in rapidity are manifestly Lorentz invariant, whereas differences in pseudorapidity are invariant for massless particles only.

of the algorithm combines the closest seeds (or proto-jets) if the distance between them is below some defined threshold. The conventional k_t jet algorithms use the distance metric

$$d_{ij} = \min \left(p_{Ti}^{2q}, p_{Tj}^{2q} \right) \frac{\Delta R_{ij}^2}{R^2}, \quad \text{where } \Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2. \quad (4.4)$$

Here the exponential factor q , and denominator term R are parameters of the algorithm. Different values of q result in differently built jets; With the setting $q = 1$ the algorithm is called the k_t algorithm [252], with $q = 0$ it is called the Cambridge-Aachen algorithm [253], and with $q = -1$ it is called the anti- k_t algorithm [254]. Note that the anti- k_t algorithm is the default sequential recombination jet algorithm used by ATLAS, implemented using computational streamlining with the FastJet package [255]. The algorithm typically produces jets with a circular active area or *cachement area* [256] with radius R .¹² In the case where there are two nearby jets, the algorithm will produce a circular jet for the higher momentum cluster and a clipped cachement area for the other. The behaviour of the anti- k_t algorithm is illustrated in Figure 4.13b for a simulated event, where the circular cachement area and clipped jets can be seen.

The most widely used jets at ATLAS use the radius parameter $R = 0.4$ (called *small-R jets*), and $R = 1.0$ (called *large-R jets*). As mentioned above, the R parameter sets the size of the circular cone of energetic jets, and is chosen to appropriately capture the decays of the particle or parton of interest. For example, small- R jets are used for gluon or light quark jets, as these capture about 89% and 95% of the p_T of the initiating particle respectively [250].¹³ Similarly, the deposits in the electromagnetic calorimeter from electrons and photons are reconstructed as small- R jets, though in these cases the shower is typically contained within a smaller cone of radius $R = 0.2$ [257]. However for electrons and photons, the jets that overlap with the electron or photon objects are removed using discriminators that use various shower shape

¹²The area of a jet is estimated using *ghost association*, wherein the event is artificially populated with zero momentum ghost particles uniformly in angular space (y, ϕ) , and the set of ghost particles that get associated to a given jet are used to estimate the active area of the jet.

¹³Gluon jets are wider than light quark jets due to differences in color charge which affects the hadronization of these particles.

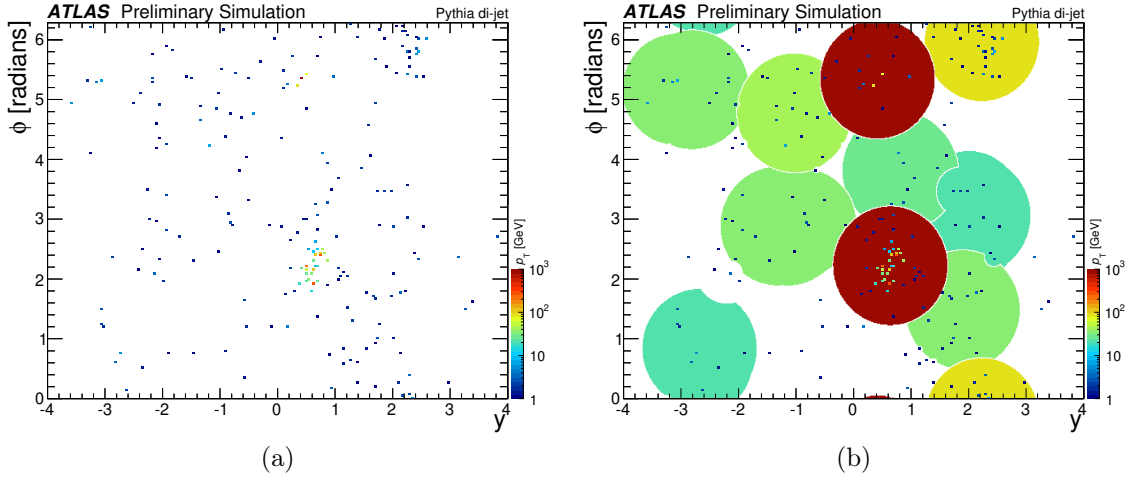


Figure 4.13: ATLAS simulated event of a di-jet interaction (PYTHIA8). (a) The topoclusters used as seeds in the y - ϕ plane, bins of size 0.04×0.03 . (b) Active area of the anti- k_t ($R = 1.0$) jets built using the topo-clusters. The colors indicate the p_T of the jets, and jets below 20 GeV are omitted. High energy hard-scatter jets (in red) are reconstructed with circular cachement areas, whereas the softer jets from the pile-up in the event have irregular boundaries. Figures sourced from [29].

information (Section 4.6.6).

Extending beyond the light objects discussed so far and to massive particles, the choice of the R parameter can be crucial to correctly reconstruct the momenta and mass of these particles. The angle between the decay products of such massive particles is approximately given by $\Delta R \sim \frac{2m}{p_T}$ [258]. Small- R jets are used for the reconstruction of jets resulting from the hadronic of the tau leptons (Section 4.6.7). Large- R jets are typically used for heavy particles jets such as those arising from W bosons, Z bosons, Higgs bosons, top quarks, bottom quarks, etc. Jets from heavy parton decays like bottom quarks contain additional information in the tracks and clusters inside the jet that help identify the displaced decay vertex of the heavy quarks and aid in the identification of heavy parton jets [236–238]. The analysis presented in Chapter 5 uses small- R jets in the study, since these are large enough to contain the decay products of the massive BSM particle for a wide range of the masses tested in the analysis.

4.6.5.2 Calibration

Once the jets are reconstructed using the algorithms described above, their energy-momenta need to be calibrated to account for the inaccuracies in the reconstruction [30, 31, 259, 260]. These inaccuracies in the jet properties can arise due to both in-time and out-of-time pile-up in the event as well as detector inefficiencies, defects, and non-linearities. The aim of jet calibration is to correct the energy-momenta of the reconstructed jets to match the corresponding properties of the *particle level* jets that would be formed if direct access was granted to the particles after the hadronization process.

To assist with that, the calibration procedure involves the use of simulations — there is direct access to the particle level jets (*truth jets*) since simulation grants access to the stable particles ($c\tau > 10$ mm) directly after the hadronization and before any detector effects have been applied. Reconstructed jets are matched to truth jets using a geometric criteria that requires the two objects to be within some maximum (η, ϕ) distance of each other $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

Additionally, the calibrations are performed using *numerical inversion* [260–262]. The procedure is described in detail in Chapter 8 of Ref. [262], and will be briefly discussed below and in Chapter 7 of this thesis.

The general calibration scheme for small- R and large- R jets is similar: the pile-up contribution to the jet energy is subtracted, the jet four-momentum is calibrated using Monte Carlo (MC) simulations (Section 4.7), and an in-situ correction is applied to account for data-simulation differences. The procedures for both small- R and large- R jet collections are also detailed below.

4.6.5.2.1 Numerical Inversion For a given task of mapping some reconstructed observable x^{reco} to its corresponding *true* scale x^{true} , calibration functions can be broadly learned in two distinct ways.

The first and most intuitive approach is to learn a mapping $f : x^{\text{reco}} \mapsto \hat{x}^{\text{reco}}$ such that $\hat{x}^{\text{reco}} = g(x^{\text{true}}|x^{\text{reco}})$, under the objective that $\hat{x}^{\text{reco}}/x^{\text{true}} \rightarrow 1$. Here the mapping f learns some measure of central tendency g of the condition distribution $(x^{\text{true}}|x^{\text{reco}})$, such that $\hat{x}^{\text{reco}}/x^{\text{true}} = 1$. Note here that the calibration function f becomes *prior-dependent* since it relies on the distribution $(x^{\text{true}}|x^{\text{reco}})$, which is subject to the physics idiosyncrasies of the sample or simulation used.

The second method learns to map $f : x^{\text{true}} \mapsto \hat{x}^{\text{true}}$ such that $\hat{x}^{\text{true}} = g(x^{\text{reco}}|x^{\text{true}})$ under the objective that $\hat{x}^{\text{true}}/x^{\text{reco}} \rightarrow 1$. This mapping is then inverted to give $f^{-1} : x^{\text{reco}} \mapsto \hat{x}^{\text{reco}}$, and results in a *response* that is defined as $R = g(\hat{x}^{\text{reco}}/x^{\text{true}}|x^{\text{true}})$. Note here that this method becomes dependent on the distribution $(x^{\text{reco}}|x^{\text{true}})$, unlike the previous method. This method of calibration is called numerical inversion.

Calibration using numerical inversion is the standard calibration method used at ATLAS, since detector effects are well understood and modeled accurately in simulation, and are generally applicable regardless of the physics process being simulated — these effects correspond to the distribution $(x^{\text{reco}}|x^{\text{true}})$. In contrast, the $(x^{\text{true}}|x^{\text{reco}})$ distribution does depend on the specifics of the physics process being simulated, and results in a calibration function that in principle is not guaranteed to *generally* achieve *closure* (defined as $R = 1$) [261, 262].

Additionally, the measure of central tendency g that is used for ATLAS calibrations is the *mode* of the distributions. This choice is made to ensure that the calibration function is unaffected by long tails or thresholds that often exist due to various experimental effects and constraints in the $(x^{\text{reco}}|x^{\text{true}})$ conditional distributions. It is interesting to note that despite the apparent validity of the numerical inversion method, the choice of g effects the theoretical best closure that the calibration can achieve — Only the median as the choice of central tendency leads to closure for any f , the mean and mode only lead to closures if f is linear [261, 262]. Consequently, the uncertainty (or *resolution*) of the calibrated scale using numerical inversion with the mode can grow if the underlying resolution of x^{reco} becomes larger.

This makes it essential to build the detector objects with good resolution to ensure that calibrating those objects will lead to an improvement in the accuracy and not degrade the precision of the observables.

Jet calibrations at ATLAS utilize numerical inversion by explicitly evaluating $g(\mathbf{x}^{\text{reco}}|\mathbf{x}^{\text{true}})$ in bins of $\mathbf{x}^{\text{true}}$ by fitting Gaussians to the core of the $(\mathbf{x}^{\text{reco}}|\mathbf{x}^{\text{true}})$ distributions, building a parametric map with the values from each bin and inverting the function to derive the calibration function. In cases where the calibration is also correcting for the dependence of the observable $\mathbf{x}$ on some parameter θ , the maps are parameterized and inverted by evaluating $g(\mathbf{x}^{\text{reco}}|\mathbf{x}^{\text{true}}, \theta)$ explicitly in bins of $\mathbf{x}^{\text{true}}$ and θ . Chapter 7 discusses a method which can perform numerical inversion without binning, by learning the parameterized $g(\mathbf{x}^{\text{reco}}|\mathbf{x}^{\text{true}})$ directly using neural networks.

4.6.5.2.2 Small-R Jet Calibration The schematic of the calibration procedure for small-R jets [30, 259, 260] is shown in Figure 4.14. The procedure begins by sub-

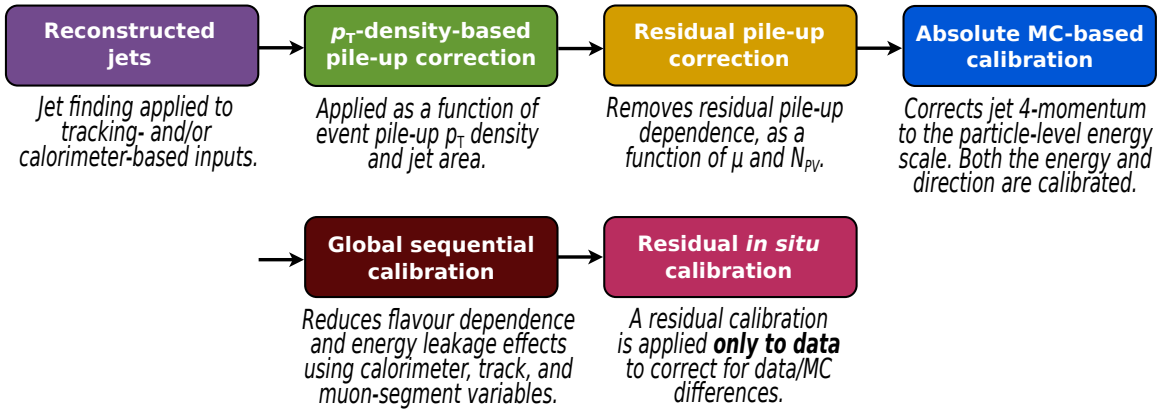


Figure 4.14: Schematic of the small-R jet calibration procedure. Figure sourced from [30].

tracting the soft pile-up contribution to the jet p_T by using an estimate of the assumed uniform pile-up p_T density over the whole event ρ . This is done by taking the median p_T/A over all the k_t jets in the event, where A is the area of a jet, and using this as an estimate of ρ . Note that the k_t algorithm finds soft jets as opposed to the anti- k_t

algorithm, and is an effective tool to measure the pile-up in the event. After this, the residual dependence of jet p_T on the number of reconstructed primary vertices (NPV) and the mean number of interactions per bunch crossing (μ) are corrected using a linear fit versus jet p_T and subtracted. Figure 4.15 shows the resulting dependence of jet p_T on NPV and μ at various points during the aforementioned three steps. The corrections applied at the end of this sequence can be summarized as follows [30], where p_T^{reco} is the uncalibrated jet p_T , p_T^{corr} is the corrected p_T , and α and β are the linear factors:

$$p_T^{\text{corr}} = p_T^{\text{reco}} - \rho A - \alpha(N_{PV} - 1) - \beta\mu. \quad (4.5)$$

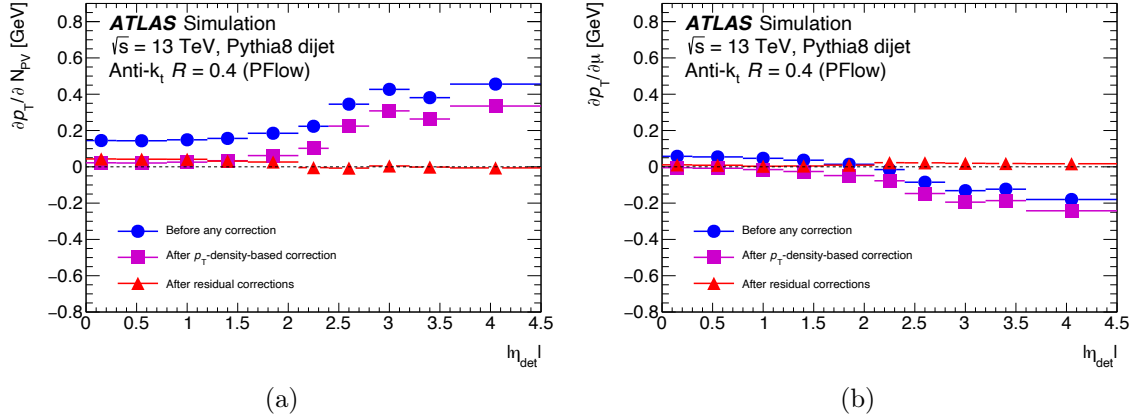


Figure 4.15: Dependence of jet p_T on (A) the number of reconstructed primary vertices NPV and (b) the mean number of interactions per bunch crossing μ , before any correction, after the area subtraction, and after the residual linear correction is applied. The functions are plotted against the detector-level $|\eta|$ of the jets, η^{det} . Figures sourced from [30].

The next step uses small- R anti- k_t jets in PYTHIA8 MC simulations (Section 4.7) of dijet events to perform corrections to the jet E_T and η , thus affecting the four-momentum of the jet. The jet energy response before the correction versus E_T^{reco} and η^{det} is shown in Figure 4.16. Truth jets are matched to reconstructed jets in simulations using a $\Delta R < 0.3$ threshold, while also requiring no truth (reco) jet within ΔR

of 1.0 (0.6) of another truth (reco) jet. Numerical inversion is utilized to perform a correction of the reconstructed jet energy (E_T^{reco}). It is performed by binning the jets in bins of true jet energy (E_T^{true}) and detector-level η (η^{det}), extracting the mode of the $E_T^{\text{reco}}/E_T^{\text{true}}$ distribution (response) by fitting a Gaussian to the core of the distribution, and parameterizing the response as a function of E_T^{reco} . Finally the function is inverted to obtain the jet calibration factor. Additionally, a correction is derived as the difference between the reconstructed and truth jet η parameterized as a function of E_T^{true} and η^{det} , wherein numerical inversion is used to get the correction in E_T^{reco} from E_T^{true} . Particle flow jets (PFlow) jets corrected with the η with the jet energy scale calibrations are considered to be at the PFlow+JES scale, as opposed to being at the PFlow scale prior to the jet calibration procedure.

Next, the global sequential calibration (GSC) applies a sequence of multiplicative factors to correct the jet p_T and reduce the dependence of the jet p_T on a set of jet features without changing the average jet energy response. These corrections for jet p_T are also derived using numerical inversion and parameterized as a function of p_T^{reco} . The features are six global jet observables such as the longitudinal structure of the energy depositions within the calorimeters, tracks associated with the jet, and the activity in the muon chambers behind a jet. For each observable, an independent jet four-momentum factor is determined as a function of p_T^{true} and η^{det} , and these corrections are applied sequentially to the jet four-momentum. Broadly, these features make the response similar for quark and gluon initiated jets, reduce the uncertainty due to the fragmentation model, account for dead material between the two calorimeter subsystems, and correct for the effect of hadronic showers punching through to the muon system. The p_T dependence on two of the six features of the GSC is shown in Figure 4.16.

The method discussed in Chapter 7 presents an alternative way to calibrate observables while simultaneously accounting for multiple features that effect the calibration. The novel method is free from binning and utilizes neural networks to perform the numerical inversion in a general all-encompassing single step.

The final step in the full calibration procedure involves studying the jet response with respect to a reference object, and correcting for data-simulation differences in

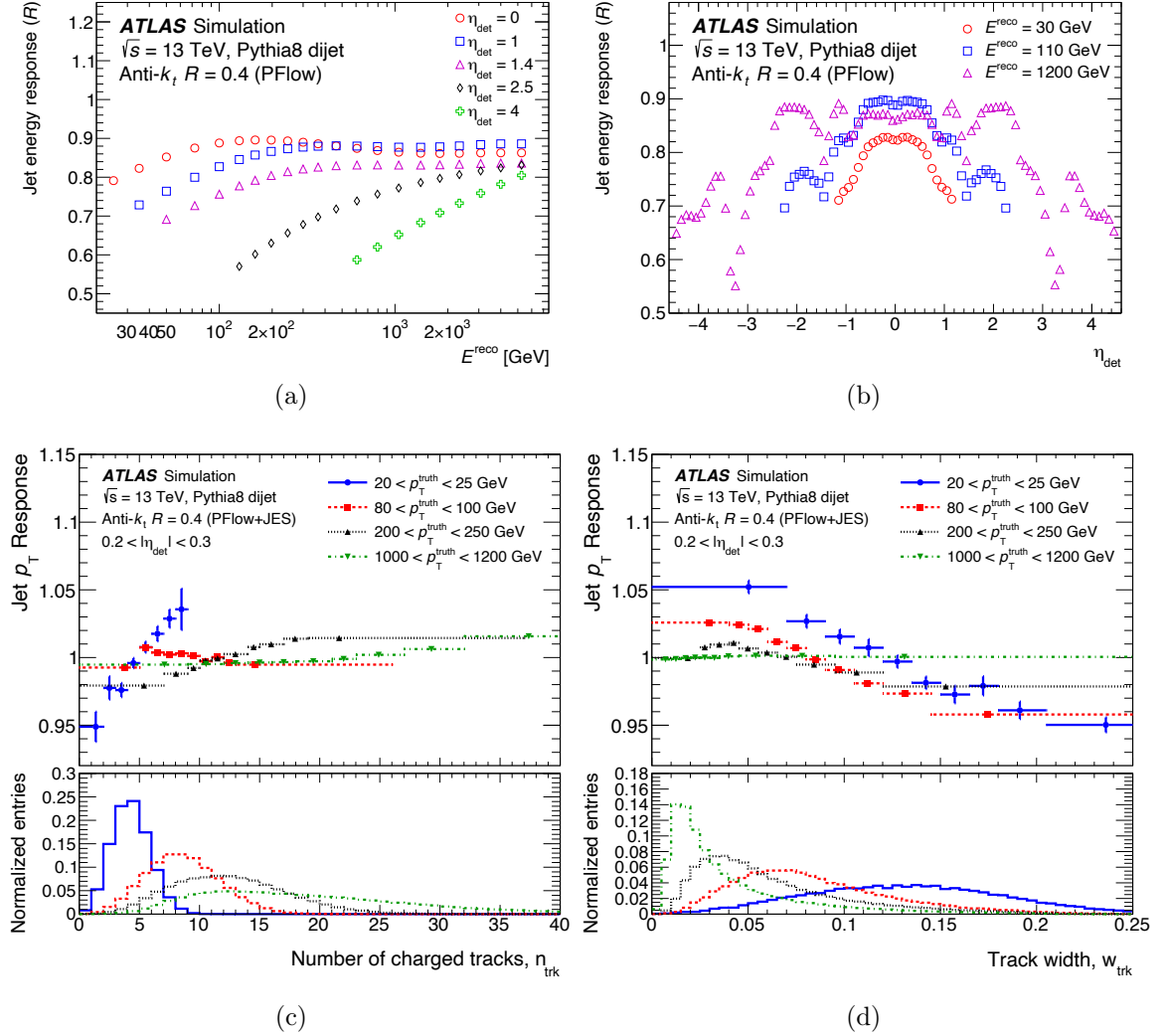


Figure 4.16: Jet response for small- R jets, to be corrected with absolute MC calibration, and global sequential calibration. (a) Energy response versus E_T^{true} in bins of η^{det} , for absolute MC calibration. (a) Energy response versus η^{det} in bins of E_T^{true} , for absolute MC calibration. (c) p_T response versus the number of tracks associated with the jet, for GSC. (d) p_T response versus the track width, for GSC. Figures sourced from [30].

that with in-situ corrections. The choice of the reference object varies with p_T scale being calibrated and uses the momentum balance in the event to compare with jets. In-situ corrections are also done sequentially, starting with corrections to make the energy of forward jets ($0.8 \leq |\eta^{\text{det}}| < 4.5$) match the scale of central jets ($|\eta^{\text{det}}| < 0.8$) using dijet events. $Z \rightarrow ee/\mu\mu$ objects are used for the $20 < p_T < 500$ GeV range, γ objects are used for the $36 < p_T < 950$ GeV range, and multijet balance is used to cover the range up to 2 TeV. These sequential corrections are proceeds applied with the numerical inversion approach.

The uncertainty of the energy scale calibration is evaluated using variations in the in-situ calibration. The various sources of uncertainty result in over 100 nuisance parameters that are reduced to about 30 parameters with *category reduction* — systematic are grouped in categories and correlations between categories in each jet η - p_T bin are evaluated and used to perform dimensional reduction. Furthermore, the jet energy resolution is quantified as the width of the Gaussian core fit to the response distribution, and is evaluated in dijet samples and with in-situ studies. Uncertainties on the energy resolution reduce to 13 nuisance parameters.

4.6.5.2.3 Large-R Jet Calibration The schematic of the calibration procedure for large-R jets [31] is shown in Figure 4.17. The procedure begins with large-R jets constructed using the LCW scale topo-clusters (Section 4.6.3). Large-R jets have a large cachement area owing to the large $R = 1.0$ anti- k_t radius parameter, and are susceptible to soft wide-angle radiation inside the jet that is primarily expected to impact the jet mass for these typically high p_T objects. The pile-up subtraction process (*grooming*) for large-R jets involves *trimming* the jets: The jet constituents are reclustered with the k_t algorithm with a radius parameter $R = 0.2$, and any of these soft pile-up dominated k_t sub-jets with a p_T fraction less than 0.05 with respect to the full jet p_T are discarded.

Next come the energy and mass calibrations for large-R jets, which are done sequentially in that order. The energy and mass responses for large-R jets before the calibration are illustrated in Figure 4.18. The energy calibration is derived in bins of η^{det} in simulation using numerical inversion. After this, the mass calibration is

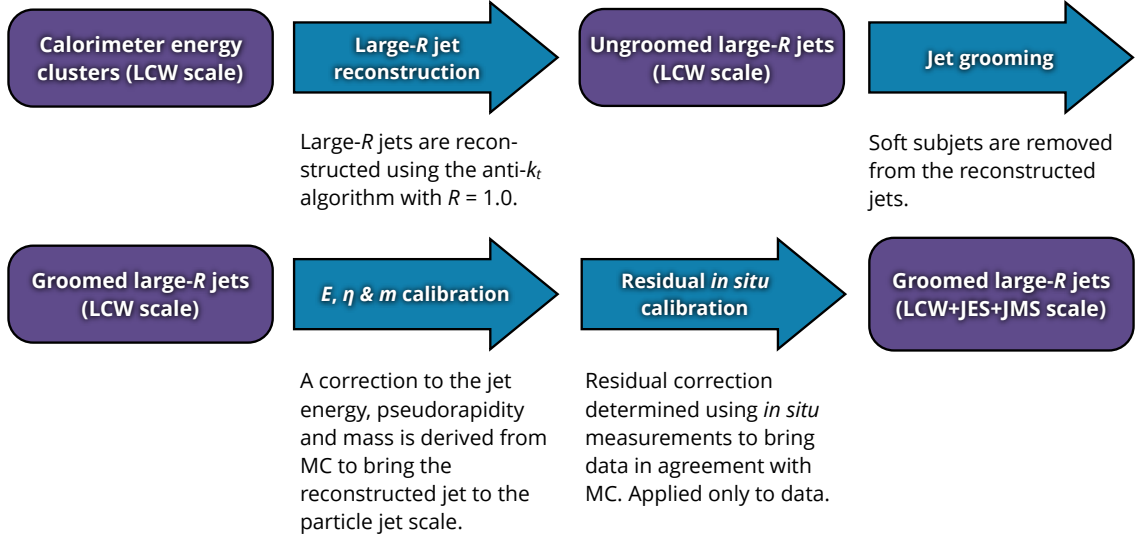


Figure 4.17: Schematic of the large- R jet calibration procedure. Figure sourced from [31].

also derived using numerical inversion, this time in bins of E_T^{true} and η^{det} . Numerical inversion is used here for similar reasons to the small- R jet calibrations, to avoid biasing the energy and mass calibrations by the simulation E_T^{true} and m^{true} distributions. Large- R jets corrected with the energy scale calibrations are considered to be at the LCW+JES scale, and further jets with the mass scale calibrations are considered to be at the LCW+JES+JMS scale, as opposed to being at the LCW scale prior to the jet calibration procedure. Additionally, the combined mass for large- R jets is composed of a weighted sum of the (independently calibrated) masses derived from the topo-clusters (m^{calo}) and the mass derived from the ghost-associated tracks (m^{track}) [263, 264].

The neural network assisted jet calibration procedure discussed in Chapter 7 is derived for large- R jets, and conducts the calibration of jet energy and mass simultaneously while also taking into account η^{det} as an observable that impacts the calibrations. The flexibility of the method to account for jet observables that impact the calibration can lead to the absolute MC based calibration step for large- R jets to also account for a variety of features like the GSC does for small- R jets.

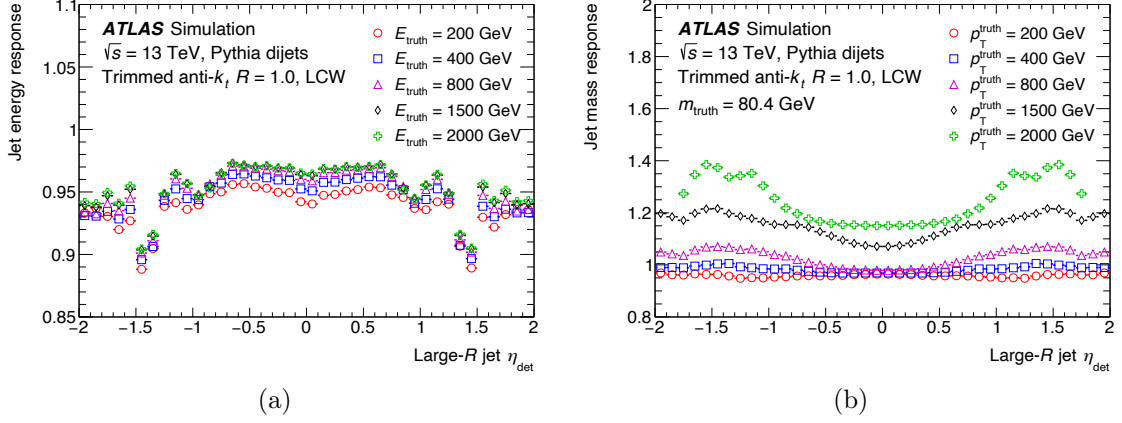


Figure 4.18: Energy and mass responses prior to the large-R jet calibration. (a) Energy response versus η_{det} in bins of $E_{\text{T}}^{\text{true}}$. (b) Mass response versus η_{det} in bins of $p_{\text{T}}^{\text{true}}$, for jets with $m^{\text{true}} \approx 80$ GeV. Figures sourced from [31].

Like the case of small-R jets, the large-R jet calibration process involves studying the jet p_{T} response with respect to a reference object, and correcting for data-simulation differences in that with in-situ corrections. In-situ corrections are also done sequentially, starting with corrections to make the energy of forward jets match the scale of central jets using dijet events. Z objects are used for the $200 < p_{\text{T}} < 500$ GeV range, γ objects are used for the $500 < p_{\text{T}} < 1000$ GeV range, and multijet balance is used to cover the range up to 2.5 TeV. These sequential corrections are proceeds applied with the numerical inversion approach.

Likewise in-situ corrections for the jet mass are also derived in two ways. Firstly, the ratio of the two jet masses m^{calo} to m^{track} can be compared in data and simulation and differences can be corrected. Secondly, *forward folding* [265] is used to compare the distribution between data and simulation and derive a correction based on any differences, using $t\bar{t}$ events where one hadronically decaying top quark large-R jet mass can be calibrated with the mass of the semileptonically decaying other top quark.

For large-R jets, the energy and mass resolution are not calibrated. Uncertainties are applied in analyses that use these objects to take data-simulation differences into account, and depends on the kinematics of the jets etc.

It is important to note here that ATLAS has recently introduced a new type

of jet input object, called a *unified flow object* (UFO) [266]. This object combines calorimeter and ID signals, like particle flow objects, to improve performance across a wide kinematic range. Consequently new large- R jet definitions are derived using these UFO objects, which improve on the ATLAS baseline large- R jet definition that uses LCW topo-clusters, as detailed in Ref. [266]. This novel jet collection will not be discussed here, and is not used in the study presented in Chapter 7, though the benefits of the method proposed in the chapter are equally applicable to the newer jet collection.

4.6.5.3 Pile-up Tagging

The jet reconstruction procedure can form jets that are entirely pile-up initiated [245, 267, 268]. These pile-up jets could be a result of local fluctuations in the soft particle noise from the other interactions, or from other hard-scatter interactions other than the interaction of interest. It is vital that these pile-up jets are excluded from analyses that are studying the physics from the hard-scatter interaction of interest. Therefore ATLAS uses a dedicated *jet vertex tagger* (JVT) that combines tracking, vertexing, and jet and variables in the event into a score that is indicative of the jet being associated to the hard-scatter primary interaction [245]. JVT is effective in the central region ($|\eta| < 2.5$), but suffers in the forward region due to the lack of tracking information beyond $|\eta| > 2.5$. However, forward QCD pile-up jets can be identified by using momentum conservation in the pile-up interaction that produced the QCD pile-up jet — the vector sum of the transverse momentum associated to the corresponding pile-up vertex will balance against the jet momentum. The *forward JVT* (fJVT) is the algorithm that is used by ATLAS to identify such forward QCD pile-up jets using the momentum balance topology [269, 270].

4.6.6 Electrons and Photons

Electrons and photons are reconstructed [257] using the ATLAS ID (Section 4.2) and calorimeter system (Section 4.3). In the electromagnetic calorimeter, electrons and photons have similar shower patterns as discussed in Section 4.3. These calorimeter signals are combined with tracks from the ID to reconstruct these particles.

For electrons, single tracks from the primary vertex are matched to a cluster of cells in the electromagnetic calorimeter. For photons, either no tracks or a pair of tracks starting from a secondary vertex in the ID are matched to clusters in the electromagnetic calorimeter.¹⁴ The topo-clusters used as seeds for this procedure are required to have more than half of their energy in the electromagnetic calorimeter versus the hadronic calorimeter, along with threshold on their energies. The topo-cluster seeds are grown into super-clusters of size $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ by combining nearby compatible clusters. The track-cluster matching is done by extrapolating the tracks to the calorimeter volume and checking for compatibility in the angular η - ϕ space.

The identification of electrons and photons is assisted by shower shape variables, and information from the TRT, that help determine if the deposits in the electromagnetic calorimeter, the transition radiation etc are consistent with electrons or photons. Multivariate likelihoods are formed by combining various these features and used to identify electrons [271], photons, and distinguish them from hadronic showers. Various electron/photon identification working points are derived based on different signal efficiency - background rejection levels, and are chosen by analysts based on the specifics of the corresponding physics analyses.

The efficiency of reconstructing electrons and photons are evaluated in data using $Z \rightarrow ee$ and $J/\psi \rightarrow ee$ events (for electrons [272]) and $Z \rightarrow ee\gamma$ and $Z \rightarrow \mu\mu\gamma$ events (for photons [273, 274]). The calibrations of the electrons and photons [275, 276] are also evaluated in data with these same samples. The data-simulation differences in the efficiency and calibration are corrected, and uncertainties on these *in-situ* corrections are applied as systematic uncertainties on simulated events.

¹⁴A photon, if it gets converted to an electron-positron pair within the ID volume, will have a signature of two tracks that can be matched to a secondary vertex.

4.6.7 Taus

Tau leptons decay hadronically roughly $2/3$ of the time, and to electrons or muons with neutrinos roughly $1/6$ of the time [1]. The leptonic decay is prompt — the lighter leptons travel to the detector from the primary interaction point. Therefore the tau leptonic channel has a signature in the detector that is identical to an electron or muon traveling to the detector from the interaction as the neutrino does not interact with the detector. Consequently, the hadronic channel is vital to tau reconstruction in the ATLAS detector.

In the hadronic decay channel, taus decays have one charged pion about 80% of the time, and three charged pions roughly 20% of the time. Thus tau reconstruction and identification in ATLAS [277–280] involves jets with uniquely narrow shower substructure. A boosted decision tree that combines various calorimeter and tracking features such as fraction of electromagnetic energy, jet mass, jet width, etc., is used to identify taus and discriminate from electrons and jets.

The tau jets are calibrated to the correct visible energy scale in a similar way to the usual jet calibration (Section 4.6.5.2) procedure. The tau identification and calibration are evaluated in data using $Z \rightarrow \tau\tau$ events (tag-and-probe), whereby the differences between data-simulation are corrected and uncertainties on the *in-situ* corrections are applied as systematic uncertainties on simulated events. Note here that tag-and-probe corresponds to identifying the event using the joint invariant mass of the two taus and one of the taus decaying to muons to identify the event, then measuring the efficiencies on the other tau.

4.6.8 Muons

As mentioned previously, muons do not interact with the calorimeter volume and are hence reconstructed from charged particle hits in the muon system [281, 282]. This reconstruction benefits from the fact that most charged particles decay within the calorimeter volume, with the exception of any leakage from the hadronic calorimeter, and therefore muon identification is very efficient. The muon charged particle track reconstruction makes use of tracks from both the ATLAS ID as well as the muon system layers. Track segments are first built in the muon layers and combined to form muon candidates, which can then be re-fit with candidates built from the ID hits — for muons that fall outside the angular acceptance of the ID, muon candidates are built solely from the muon system hits. The muon building procedure can happen using track candidates that built from the muon layers first, then associated with the candidates from the ID, or vice versa.

The issues of reconstructing muon candidates that do not correspond to real muons is an important one, and is dealt with in a variety of ways. The primary handle on muon identification is the compatibility of the track candidates from the ID and muon systems, such as consistency in the charge to momentum ratio, etc. Furthermore, determining whether the particle originates from the primary vertex or a secondary vertex is a good handle on muon that were produced from secondary decays. Together, these properties of the muon candidates can help distinguish true muons from cosmic muons or muons produced due to the decay of charged hadrons in the ID. The various muon properties are combined into several working points that can be used by the physics analysis teams to select muons with an appropriate muon reconstruction and identification efficiency. Additionally, analyses can also require muons to be isolated in the detector from other objects, such as electrons, photons, or jets.

As in the case for the other objects discussed so far, the reconstruction, identification, and isolation efficiencies for muons can be estimated in simulation samples and validated using the appropriate data sets. For muons, the $Z \rightarrow \mu\mu$ events are used — The join invariant mass of the two muons is used to identify the event along with one of the muons, and the second muon is tested for efficiency measurements (tag-and-probe). The differences between the muon identification efficiency between data

and the simulated samples is corrected for — Systematic uncertainties corresponding to this correction are provided to analyses.

4.6.9 Missing Transverse Momentum

In addition to all the objects that ATLAS can reconstruct, detailed in Sections 4.6.6 - 4.6.8, there are non-interacting particles like neutrinos that do not leave any signature in the detector. It is however vital that the momentum carried by the non-interacting (invisible) particles in an event is reconstructed to allow for the study of such events and derive physics inferences using the kinematics of the event. The reconstruction of this *missing* momentum [283, 284] — more specifically missing transverse momentum (E_T^{miss}) — relies on the conservation of momentum and the fact that the momentum in the transverse direction should add up to zero. Note here that the longitudinal direction is not useful since the exact balance of the longitudinal momentum in the primary interaction is not known.

E_T^{miss} is the magnitude of the vector $\vec{E}_T^{\text{miss}}$, which in turn is calculated as the negative sum of all the vector momenta of the reconstructed objects in the event. The total vector sum of momenta in the event is split into two broad terms: the *hard term*, and the *soft term*. The hard term comprises of the vector sum of all the reconstructed object momenta, and includes contributions from electrons, photons, muons, hadronically decaying taus, jets. The soft term comprises of the vector sum of all the tracks that are associated with the primary vertex, but not associated to any of the reconstructed objects listed above.

In the hard term, overlapping objects are vetoed by following a hierarchy of priority; The order is defined by the precision of the objects. As an example, jets are removed from the sum if they are found to be overlapping photons or electrons, and only small-R jets are used since they would be overlapping with large-R jets. The hard term is meant to account for the momenta originating from the hard scatter primary interaction, therefore pile-up rejection techniques are crucial in making sure that no objects from pile-up interactions are included in the sum. Furthermore, the calibration of each object improves the resolution of the hard term.

The soft term estimates the contribution to E_T^{miss} that comes from activity outside the standard reconstructed objects. This is estimated using all the tracks in the event that are not associated with any of the reconstructed objects but still come from the hard scatter primary vertex. This term is the subdominant term since by construction most of the activity is taken into account by the hard term. By adding the vector momenta of the unassociated hard scatter tracks, the soft term is accounting only for charged particles and ignoring the contribution from neutral particles. However, the neutral contribution to the sum is expected to cancel on average due to the azimuthal symmetry of the charge–neutral fluctuations. Since the soft term relies on tracks, it is robust to pile-up since tracks from vertices other than the hard scatter vertex can be readily vetoed.

Differences between data and simulation in E_T^{miss} for $Z \rightarrow e\bar{e}/\mu\bar{\mu}$ events are evaluated, and can be applied as systematic uncertainties in analyses that use E_T^{miss} .

4.7 Simulations

Simulations form a key component of the ATLAS analysis pipeline. While these do not constitute a physical component of the experiment, they are vital to most measurements performed with the detector. Simulations help tune the reconstruction algorithms to improve the reconstruction efficiency and benefit the physics yield of the detector. This is directly relevant to the reconstruction algorithm discussed in Chapter 6. Simulations also help us quantify the effect of the various measurement devices and how they relate to the properties of the interacting particles. This plays a crucial role in deriving the calibrations for the various object observables, and is the focus of Chapter 7. Additionally, simulated events are also used to understand how sensitive the detector is to particles and interactions following from new hypotheses. This is the case for the analysis presented shown in Chapter 5. Simulations can also be used to model the various backgrounds that might affect an analysis. While simulations are not used to estimate the background in the analysis presented in Chapter 5, they are used to correct for residual effects and estimate various sources of uncertainties.

Simulations of physics interactions are done using a series of Monte Carlo (MC) simulation programs that each target a step of the interaction-to-detection chain [285]. The simulation process is simplified since the length scales spanning the full simulation range from 10^{-20} m to $\mathcal{O}(10)$ m). Since physical laws *factorize*, simulating physics at a given length scale depends only on the physics at the next smaller length scale. Therefore the full simulation can be done sequentially, and the factorization can be realized using Markov Chain MC. Due to the relative rarity of *interesting* physics processes (Figure 3.4), only events from a specific tree-level **pp** interaction (*hard-scatter*) are simulated rather than the inclusive set of physics processes that occur during **pp** interactions.

The simulation chain starts with matrix element calculations that generate particles at the parton level that are the result of the interaction following the **pp** collision. These calculations first get the initial particles produced from the inelastic collision of the constituent partons in the colliding protons. Then the parton showering simulator generates the partons that are produced as the particles approach the hadronization scale $\mathcal{O}(\text{GeV})$ due to radiation. These generators typically specify what fixed order in α_s the calculations are being carried out perturbatively, like leading order (LO), next-to-leading order (NLO), etc. Common generators include PYTHIA [286, 287], SHERPA [288], and HERWIG [289, 290], and may use POWHEG-BOX [291, 292] (in the case of PYTHIA and HERWIG) to interface between the matrix element calculator and parton showering simulator. Analyses typically use one generator as the main physics simulator for the study, and one or more generators to estimate the theory modelling uncertainties. In the case of the analysis presented in Chapter 5, PYTHIA is used as the primary generator and SHERPA is used to form a two-point systematic [293] to estimate the background modelling uncertainties. Additionally HERWIG is used to form a two-point systematic to estimate the signal modelling uncertainty where possible. In the calibration study presented in Chapter 7, calibration results are presented for all three generators to understand the performance of the method as the theory modelling is varied.

The next step is to simulate fragmentation and hadronization — the process that converts partons to hadrons. The generators discussed above use phenomenological

models tuned to collected data, like in Ref. [294]). The generators are thus specified by their particular hadronization tune. It should be noted here that this reliance on data is necessary since fragmentation and hadronization are not fully described by fundamental theory yet.

While simulation of the hard-scatter interaction has been discussed so far, the aforementioned generators are also be used to generate the pile-up (concurrently occurring pp inelastic interactions) [295] which are then overlaid on top of the simulated hard-scatter interaction. This is done with the appropriate pile-up conditions (average number of interactions) expected in the events. Out-of-time pileup is modeled similarly with offsets in time to account the signal processing delays. Additionally the generators are also used to simulate the soft interactions of the partons in pp collisions (underlying event) using phenomenologically driven parton distribution functions [296].

Simulations of Heavy flavored hadrons are handled with specialized programs like EvtGen [297].

By this stage in the simulation chain, the final state particles that interact with the detector have been generated. The next step, therefore, is to simulate the interactions of these particles with the detectors and *digitize* the appropriate observables. This is done by processing the particles with the full ATLAS simulation toolkit, which is a modelling of the detector in GEANT [298]. This accurately models the response of the detector to these particles and generates detector signals in the same way that the detector would when recording real data. The resulting detector level information is used to reconstruct the properties of the particles (Section 4.6) so as to undo the effect of the detector and study the physics of the hard-scatter interactions of interest.

Chapter 5

A Search for Exotic Higgs Boson Decays to Four Gluons Using the ATLAS Detector

A novel search for exotic decays of the Higgs boson to four gluons via coupling to a new light (pseudo)scalar is presented in this Chapter. The final state targeted is $H \rightarrow \alpha\alpha \rightarrow gggg$, where the Higgs boson is produced in the ZH production channel, and the Z boson decays to two muons. The mass range for the (pseudo)scalar (α) searched in this analysis ranges from 0.5 GeV to 16 GeV, and the scalar is assumed to decay promptly to two gluons with a 100% branching ratio. A variety of machine learning techniques are used to perform this search, and a custom machine learning tagger has been developed by the Author to identify the signal $H \rightarrow \alpha\alpha \rightarrow gggg$ events, which is detailed in Ref. [32]. Additionally, the tagger uses an automated decorrelation method that enables the estimating of the background using the ABCD method. The analysis uses the full *Run 2* data collected by ATLAS during the 2015-2018 data taking period, amounting to an integrated luminosity of 139 fb^{-1} , and is targeted for a full publication in Fall 2022. In the Standard Model hypothesis, the analysis will set the first competitive upper limits for the branching ratio of the Higgs to four gluons via new scalar couplings for the light scalar masses (0.5 – 2 GeV).

5.1 Introduction

As introduced in Chapter 2, there are many limitations of the Standard Model that prompt searches for new physics beyond the Standard Model. At the Large Hadron Collider, these searches have focused on the Higgs boson after its discovery [47, 48] since carefully studying and characterizing the properties of the Higgs boson could reveal new BSM physics. A particularly interesting avenue of research focuses on rare and exotic decays of the Higgs boson, since existing constraints on the Higgs branching ratios still allow for non-trivial coupling of the Higgs boson to new physics — Global fits to the Higgs properties and decays constrain the Higgs branching ratio to undetected physics to less than 19% at 95% confidence level [299]. Therefore, BSM physics that couples to the Higgs with a branching ratio less than 19% is a viable target for searches for new physics.

Consistent with the above insight, there are several BSM physics models that lead to exotic decay modes of the Higgs bosons. These models aim to alleviate some of the issues with the SM discussed in Chapter 2, such as the matter anti-matter asymmetry of the universe, existence and constitution of dark matter, the naturalness of the Higgs boson mass, etc. Consequently, effort has been taken to systematically search for Higgs boson decays to a variety of final states, such as $H \rightarrow 4\ell$ for $\ell \in \{e, \mu\}$ [300–302], $H \rightarrow 2\mu 2\tau$ [303–305], $H \rightarrow 4\tau$ [304], $H \rightarrow 2b 2\mu$ [306, 307], $H \rightarrow 4b$ [308, 309], and more [310].

Despite the extensive coverage of Higgs final states, one that is entirely unexplored is the all-hadronic $H \rightarrow 4g$ final state, which involves the decay of the Higgs boson to four gluons via coupling to a new scalar or pseudoscalar $\mathbf{a}$. This channel is backed by strong theoretical motivation [73], as briefly outlined in Section 2.6, and expanded upon below. In some two-Higgs doublet models (2HDM), the $\mathbf{a} \rightarrow gg$ decay can be the dominant mode in a variety of scenarios. For example, in the scenario where the $\mathbf{a}$ is very light ($m_{\mathbf{a}} < 2m_{\tau}$), this could be due to kinematic suppression to other decay modes. If the $\mathbf{a}$ is not light, the coupling to down-type quarks and charged leptons can still be suppressed if the ratio of the vacuum expectation values of the two Higgs doublets, $\tan \beta \ll 1$. In a separate model, the presence of a new BSM heavy vector

fermion that couples to the scalar could allow for $a \rightarrow \gamma\gamma$ or $a \rightarrow gg$ decays via quantum loop-induced processes. Additionally, several models that strongly motivate new BSM particles coupling with large branching ratios to di-photons (e.g. 750 GeV di-photons [311]) also motivate a large branching ratio to di-gluons in a different region of the model parameter space. Furthermore, models that do not have a di-photon component can still motivate the di-gluon decay, like models with warped extra dimensions [312]. It is therefore evident that the $4g$ channel is theoretically well-motivated, and requires a dedicated search.

From the perspective of the research presented here, the $4g$ channel is searched for in a model agnostic manner across a range of scalar masses. The phenomenology of the $H \rightarrow aa \rightarrow 4g$ final state depends strongly on the mass of the scalar, m_a . In the case of $m_a \gtrsim m_h/2$, the final state will have four resolved low momentum jets, which is especially hard to identify at the LHC given the enormous rate of low momentum jet production. Additionally, for $a \sim \text{few} \times m_\pi$, the scalar decays into exclusive final states with some small number of hadrons. Consequently, this analysis will target an intermediate mass regime, wherein the a is light enough to ensure that the Lorentz boost in the laboratory frame collimates the decay products into a single jet. Furthermore, the a decay is assumed to be prompt in this analysis.

Due to the final state containing two of these di-gluon jets, identifying these jets is a crucial part of this analysis and enables the identification of the $H \rightarrow aa \rightarrow gggg$ event as a whole. This is enabled by the use of deep neural network based on the particle flow networks [313, 314] that utilize every track inside the di-gluon jets to learn the substructure contained inside these unique objects.

To handle the issue with the trigger bandwidth that would be otherwise associated to a generic search with low momentum jets, this analysis makes use of the associated production ZH for triggering, where the Z decays into muons. The requirement that the two muons in the event reconstruct the Z boson mass mitigates some of the background that affects this analysis, leaving the Z +jets background as the dominant background that requires modeling — This is a difficult to background to model, which leads to a preference of data driven background estimation techniques. In this analysis, an extension of the data-driven ABCD method is used to estimate the

background rates [33, 315], assisted by machine learning.

The combination of the above listed techniques enable the search to achieve sensitivities to the four gluon decays with branching ratios that are near existing global fit bounds.

5.2 Event Samples

5.2.1 Data

The analysis presented here is performed using data from pp collisions delivered by the LHC with $\sqrt{s} = 13$ TeV to the ATLAS detector, in the 2015–2018 data taking period. The total integrated luminosity of the dataset is 139.0 fb^{-1} with an uncertainty of 1.7%, for data that clears the ATLAS data quality checks that ensures that the data was collected during nominal functioning of the equipment [2]. Note that data that clear these quality checks are part of the ATLAS Good Runs List. The total integrated luminosity of cleared data per year of data-taking is listed in Table 5.1.

Year	Integrated Luminosity (fb^{-1})
2015	3.21
2016	32.99
2017	44.31
2018	58.45
Total	139.0

Table 5.1: The integrated luminosity of the data collected each year of the *Run 2* data-taking period. The luminosity shown here corresponds to the data that clears the ATLAS data quality checks [2].

5.2.2 Simulation

Monte Carlo (MC) simulations are used to develop the analysis techniques and test the analysis selection methods. Details about MC simulations used in ATLAS can

be found in Section 4.7. The nominal background $Z \rightarrow \mu\mu + \text{QCD jets}$ and signal $H \rightarrow a\bar{a} \rightarrow gggg$ process are modeled using POWHEG-BOX v2 [291, 292, 316] integrated with PYTHIA 8.2 [287, 317] using the AZNLO [318] set of tuned parameters and the NNPDF3.0nlo Parton Distribution Function (PDF) set [319] for the signal and the CT10NLO PDF set [296] for the background. The matrix elements for the signal use GOSAM [320] and the MINLO procedure [321, 322] for merging with the parton shower. Systematic uncertainties of the background are evaluated with an alternative background sample using SHERPA 2.2.11 [323]. In this set-up, NLO-accurate matrix elements for up to two jets, and LO-accurate matrix elements for up to five jets in the five-flavour scheme are calculated with Comix [324]. The b - and c -quarks are treated as massless at matrix-element level and massive in the parton shower. Virtual QCD corrections for matrix elements at NLO accuracy are provided by OpenLoops [325–328]. The Hessian NNPDF3.0nnlo [329, 330] PDF set is used for the 2.2.11 configuration. The invariant mass of the charged leptons is required to be larger than 40 GeV for the SHERPA sample, and larger than 60 GeV for the POWHEG+PYTHIA sample. Additional details of the SHERPA 2.2.11 sample can be found in Ref. [331].

Six signal masses for the BSM scalar were evenly sampled in the range [0.5, 16] GeV on a logarithmic scale (powers of two), and simulated using AF3. Basic validation plots for these signals can be found in Appendix A.1.

All MC samples were processed through the full ATLAS detector simulation [332] pipeline based on Geant4 [298]. The pileup conditions matching the 2015–2018 data-taking period are simulated by overlaying minimum-bias events on the hard-scatter process. The minimum-bias are generated with PYTHIA 8.186 using the A3 set of tuned parameters [333] and the NNPDF2.3lo PDF set. Unless stated otherwise, background MC samples refer to $Z \rightarrow \mu\mu + \text{QCD jets}$ samples integrated with Pythia.

5.3 Event and object selection

The object reconstruction and jet labelling are described below:

PFlow Jets: Particle flow (PFlow) jets are constructed with the anti- k_t algorithm [334] as implemented in FastJet [255] using the radius parameter $R = 0.4$, using charged constituents associated with the primary vertex and neutral constituents as inputs [335]. The jet energy is then corrected for overlapping proton-proton collisions (pileup) and calibrated [30] using MC simulations and data. The jets used in this analysis are required to have transverse momentum $p_T > 20$ GeV and $|\eta| < 2.5$. The Jet Vertex Tagger (JVT) discriminant [336] is used to reject jets that do not originate from the hard-scatter vertex through the use of tracking and vertexing information for jets with $p_T < 60$ GeV and $|\eta| < 2.4$. Overlap removal between leptons and jets is applied using the standard overlap removal tool.

Truth jets: Truth jets in simulated events are made using truth information from the event record of the MC generator. Stable particles with proper lifetimes $c\tau_0 > 10$ mm are clustered into these jets using the same clustering algorithms and $R = 0.4$ as for detector-level jets. Particles such as muons and neutrinos which do not leave significant energy deposits in the calorimeter are excluded. Taus are included, however these are not tagged as quark or gluon jets and therefore do not affect the di-gluon jet tagging study presented here.

QCD Jet Quark-Gluon / Signal Di-gluon labeling: The labeling of the background QCD jets as originating from either a quark or gluon in simulation is done by geometrically matching the PFlow jets with truth jets. If the truth jet geometrically closest to the PFlow jet has $p_T > 10$ GeV and is found within $\Delta R = 0.3$ of the PFlow jet's direction, the PFlow jet is assigned the label of the truth jet. Truth jets are labeled using the highest energy parton within the jet cone from the MC event record. While quark/gluon labeling is not unique [337, 338], the simple procedure used here is useful for building an intuition for the performance of the di-gluon tagger (this information is not used for the final statistical analysis). In the signal samples, jets are labeled as di-gluon jets if they match within $\Delta R = 0.4$ with an a boson from the MC event record.

Tracks: Tracks are reconstructed from energy deposits, or hits, in the inner detector system [63] and are required to pass a quality selection: each track must have at least 7 hits in the silicon layers (pixel and silicon microstrip tracker, including dead pixels), less than three missing hits in the silicon layers, less than two missing hits in the pixel detector, less than two hits shared by multiple tracks, at least one hit in the pixel detector, $p_T > 500$ MeV and $|\eta| < 2.5$ [239]. Tracks are associated with the primary vertex, defined as the reconstructed vertex with largest $\sum p_T^2$ of the associated tracks [27]. Tracks are associated to jets if $\Delta R(\text{track}, \text{jet}) < 0.6$. This is larger than the jet clustering radius parameter to improve performance for higher mass α bosons.

5.3.1 Event Selection

The event selection is outlined in Table 5.2. It is primarily designed to reject background events in favour of the $(Z \rightarrow \mu\mu) + \text{jets}$ signal process. Following the selection, contributions from processes other than strong $Z + \text{jet}$ production are negligible^{1 2}. The various criteria are discussed in more detail below.

5.3.1.1 Event requirements

The first requirement is that the event must be a part of the ATLAS Good Run Lists (GRL), as discussed in Section 5.2.

Secondly, the events must pass event cleaning, which removes events that:

- Contain errors in the data collected in the LAr calorimeters
- Contain errors in the data collected in the tile calorimeters

¹The same event selection as Ref. [339] is followed, where this was explicitly validated, albeit at higher $p_T^{\mu\mu}$. Much of the text in this and other sections related to the setup is adapted from Ref. [339].

²Furthermore, since the background estimation here is data-driven to first order, the effects of background composition are weak. To second order, however, since the ABCD closure is used for $Z + \text{jets}$ as proxy for the MC background closure, a check could be performed to ensure that the closure is small for the non $Z + \text{jets}$ background processes. These processes are a few percent effect at most following Ref. [339] and even a 100% non-closure from them results in an approximately 2-5% relative uncertainty.

Event Selection	Description
Good Runs List	Event must be part of GRL
Event Cleaning	No LAr, tile calorimeter, or tracker errors. Event is complete.
Trigger	Single muon trigger: HLT_mu20_iloose_L1MU15_OR_HLT_mu50 (2015) HLT_mu26_ivarmedium_OR_HLT_mu50 (2016-18)
Primary Vertex	$N_{PV} \geq 1$
Muons	$N_{\text{muons}} \geq 2$ Opposite charges Pass TTVA recommendations for muons $81 \leq m_{\mu\mu}(\text{GeV}) \leq 101$
Jets	$N_{\text{jets}} \geq 2$ Pass JVT and overlap removal $p_T > 20 \text{ GeV}$ and $ \eta < 2.5$ $m_{jj} < 2\mu_{m_{jj} m_a}$, $\mu_{m_{jj} m_a}$ shown in Fig. 5.9

Table 5.2: Overview of the applied event selection.

- Contain errors in the silicon tracker data
- Are incomplete due to a restart of the trigger, timing and control system
- Fail the Loose event cleaning requirement. This is also applied to MC samples.

Additionally, it is required that the event in question have at least one primary vertex.

5.3.1.2 Trigger requirements

Events are required to pass an unprescaled single muon trigger, as outlined in Table 5.2. The recommended trigger is dependent on the year in which the data was taken (or the year of the corresponding Monte Carlo sample).

5.3.1.3 Dimuon system requirements

The event is required to have at least two muons of opposite charge, with a dilepton mass between 81 GeV and 101 GeV in order to ensure the muons originate from a Z boson.

Additionally, the muons must pass track-to-vertex association (TTVA) recommendations as follows:

- $|\mathbf{d}_0^{\text{BL}} \text{significance}| < 3$
- $|\Delta z_0^{\text{BL}} \sin \theta| < 0.5 \text{mm}$

where BL stands for beamline, and $\mathbf{d}_0$ and $\mathbf{z}_0$ represent the point along the track with the closest approach to the beamspot.

5.3.2 Corrections to Monte Carlo samples

5.3.2.1 Pileup reweighting

When a given MC event is produced, the average interactions per bunch crossing, μ , is an assigned parameter that determines how many pileup interactions are to be overlaid. It is important that the pileup modeling of the simulated samples correspond to data, therefore each MC event is assigned an event weight such that the μ distributions agree.

The `PileupReweightingTool` uses the μ and the MC event along with the luminosity of the corresponding data in the same run number to determine the pileup weight to be applied to the event. This requires that both the μ distribution for the specific MC sample is available to the tool, as well as the files used to calculate the luminosity for the data. As an unscaled trigger is used in this analysis, the centrally produced lumicalc files corresponding to the GoodRunsLists in Section 5.2.1 are used. Configuration files for the tool for each MC sample are also provided centrally. The pileup weight is then multiplied with the MC event weight in order to correct the μ distribution.

The tool also provides a scaling correction for data. The pileup simulation for a given μ value is known to be too hard, therefore the μ distribution of the dataset is scaled towards lower values to account for this.

5.3.2.2 Trigger scale factor

There are differing trigger efficiencies between data and MC simulation, thus a correction in the form of a scale factor applied to the event weight is necessary for MC events. The triggers used in this analysis are outlined in Section 5.3.1.2.

5.3.2.3 Muon scale factors

Muon scale factors related to reconstruction efficiency, isolation efficiency, and track-to-vertex (TTVA) efficiency are derived from J/ψ and Z events using the tag-and-probe method for muons in the central region. The resulting scale factors are applied to the event weight in the same way as the trigger scale factor.

5.4 Analysis Strategy

5.4.1 Overview

The analysis proceeds in the following steps:

1. Construct a parameterized reweighting function using the m_{jj} sideband to make the simulation match the data as close as possible. The reweighting is done using all of the tracks inside the jets, as this is the phase space that will be used to train the signal-versus-background classifier. See Section 5.4.3.2 for the impact of this reweighting.
2. Train a classifier to distinguish simulated signal from the reweighted background in step one. Even though the reweighting function is only constructed in the m_{jj} sideband, it can be interpolated into the signal region to be applied everywhere. The classifiers are trained to be independent of m_{jj} so that we can later apply the ABCD method. A hyperparameter λ trades off the classification and decorrelation performance. We train multiple classifiers, one for each signal mass point and one for a set of λ values.
3. For each mass point, select a classifier and classifier working point to maximize the expected sensitivity. In particular, we scan window cuts on m_{jj} and threshold cuts on the classifier so that the $S/\sqrt{B}$ in the signal region is highest. We only consider points where the ABCD closure (defining the non-signal regions by inverting the cuts on the classifier and m_{jj}) is sufficiently good.
4. Estimate the background in the signal region (A) using an ABCD method, using Monte Carlo to estimate the uncertainty from the non-closure. The non-signal regions (BCD) are used to estimate the number of background events in the signal region ($N'_A = N_B \times N_C / N_D$ where D is diametrically opposite to the signal region) which is then multiplied by the ABCD closure ($N_A^{MC \text{ bkgd}} / N'_A^{MC \text{ bkgd}}$) derived from the Monte Carlo background events.
5. Systematic uncertainties on the signal are estimated by varying some aspect

of the simulation (e.g. the tracking performance) and recomputing the signal yield.

6. Perform the final hypothesis tests, one per signal mass point.

The remaining subsections will describe steps (1) - (3) in more detail. Section 5.4.2 will highlight some of the key differences between quark/gluon jets and di-gluon jets. Then, Section 5.4.3 describes the machine learning for both (1) and (2). The tagger performance is presented in Section 5.4.4, which also includes how (3) is performed.

5.4.2 Di-gluon tagging

Di-gluon jets are expected to have distinct characteristics when compared to QCD quark / gluon jets. Firstly consider the fact that the signal jets are produced in pairs; the signal $H \rightarrow a\bar{a} \rightarrow gggg$ produces a pair of di-gluon jets which are connected by the Higgs resonance. This stands in contrast to the background QCD jets which are constrained by no such resonance. Figure 5.1 presents the invariant mass of the leading two jets using the MC simulation samples. All of the signal samples have an m_{jj} peak near the Higgs boson mass (125 GeV) while the background is smooth and steeply falling beyond about 50 GeV. In the limit of small m_a , a single $R = 0.4$ jet is slightly larger than needed to capture the a decay products and the peak is well localized just above 125 GeV. However, for larger a masses, the $R = 0.4$ jets are not able to capture all of the decay products and so the peak broadens and shifts to lower values. Properties of the radiation pattern within the $a \rightarrow gg$ jets are explored in further detail below.

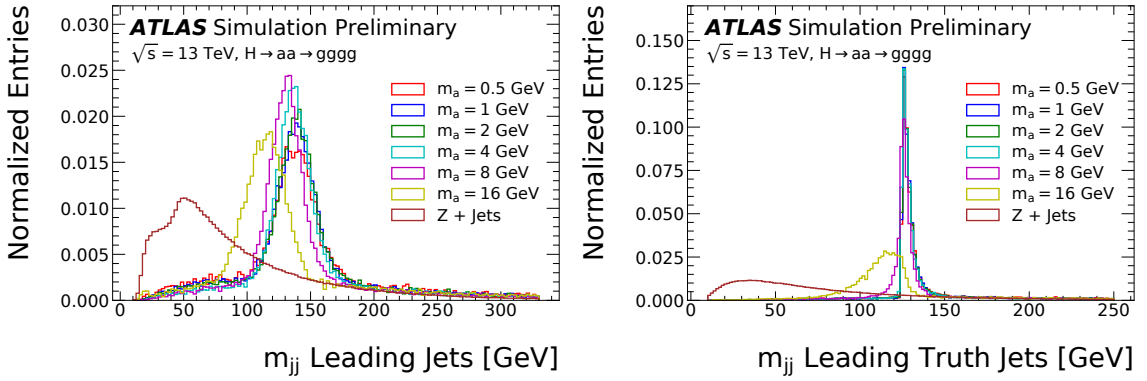


Figure 5.1: Invariant mass of the two leading jets in the event, m_{jj} , using reconstructed jets (left) and truth (particle-level) jets (right). Figures sourced from [32].

One of the most basic features of a jet is the number of constituents. Since the a boson is a massive color singlet, it is expected that the number of decay products is independent of its p_T and is instead only dependent on m_a . In contrast, the number of constituents inside a generic quark or gluon jet increases with the p_T , with gluon jets having more constituents than the quark jets due to the larger quantum charge under the strong force. Charged particles can be accurately measured inside jets

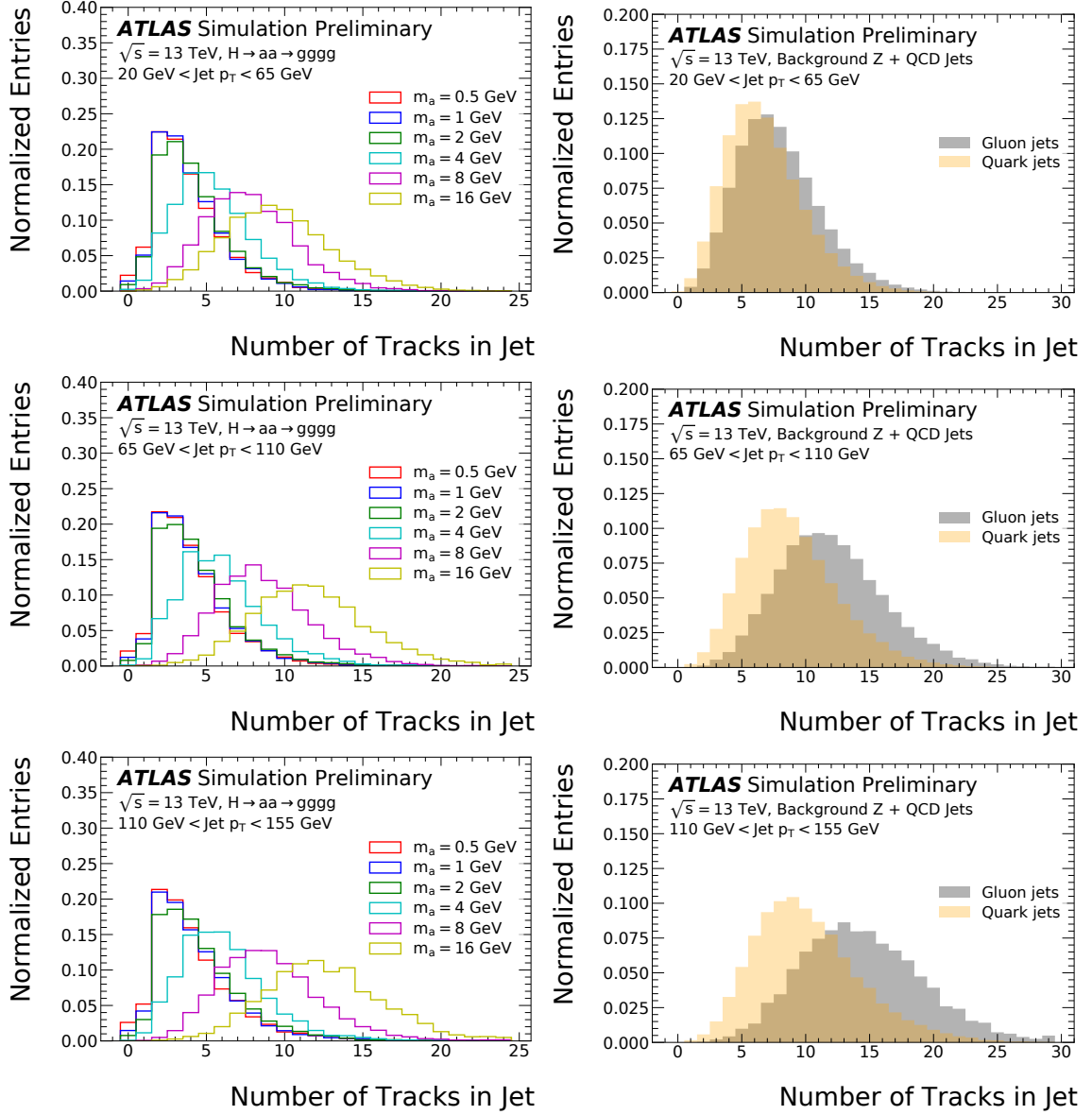


Figure 5.2: Number of tracks in the event associated to a given jet ($\Delta R < 0.6$), for signal $a \rightarrow gg$ jets (left), background QCD quark / gluon jets (right). Different p_T bins are shown from top to bottom. Figures sourced from [32].

and are used in this study as a proxy for all the constituents of the $\alpha \rightarrow gg$ jets. Figure 5.2 shows histograms of the number of tracks inside jets in three jet p_T bins. As expected, the spectra are nearly independent of p_T for the signal and shift to the right for the background. A jet from $m_\alpha = 16$ GeV has about twice as many tracks as a jet with $m_\alpha \lesssim 2$ GeV. When the α is sufficiently light, the number of hadron decay products does not depend strongly on m_α because the available phase space only allows for a small number of particles (mostly pions). Note also the slight rightwards shift for the $m_\alpha = 16$ GeV case with increasing jet p_T , where the shift highlights the limitation of using $R = 0.4$ jets for this high mass as the radius fails to capture all of the constituents and the shift is indicative of the collimation aiding in the capture of all the constituents within the $R = 0.4$ jet.

In addition to having a different number of constituents, $\alpha \rightarrow gg$ jets differ from generic quark and gluon jets in the angular distribution of the constituents. Figure 5.3 shows histograms of the ΔR between the leading two tracks within jets. Both signal and background jets become more collimated as the jet p_T increases, although the spread for the signal is typically smaller than for the background.

In addition to these features displayed above, it is also interesting to examine classical jet substructure observables to explore differences between the di-gluon and quark/gluon jets. Figure 5.4 shows the two point energy correlation function [340] for these jets. This metric is calculated using all the tracks associated to the jets, using the following formula,

$$\text{ECF}(2, 1) \equiv C_1 = \frac{\sum_{i,j}^{\text{tracks}} p_{T_i} \cdot p_{T_j} \cdot \Delta R_{ij}}{(\sum_k^{\text{tracks}} p_{T_k})^2}, \quad (5.1)$$

where ΔR_{ij} is the angular distance between tracks i and j .

Another well-studied observable shown in Figure 5.5 is p_TD [341], which probes the energy fractions carried by a jet's constituents independent of their geometric orientation:

$$p_TD = \frac{\sqrt{\sum_i^{\text{tracks}} p_{T_i}^2}}{\sum_j^{\text{tracks}} p_{T_j}}. \quad (5.2)$$

Figure 5.4 shows that signal di-gluon jets tend to have a lower $\text{ECF}(2,1)$ and

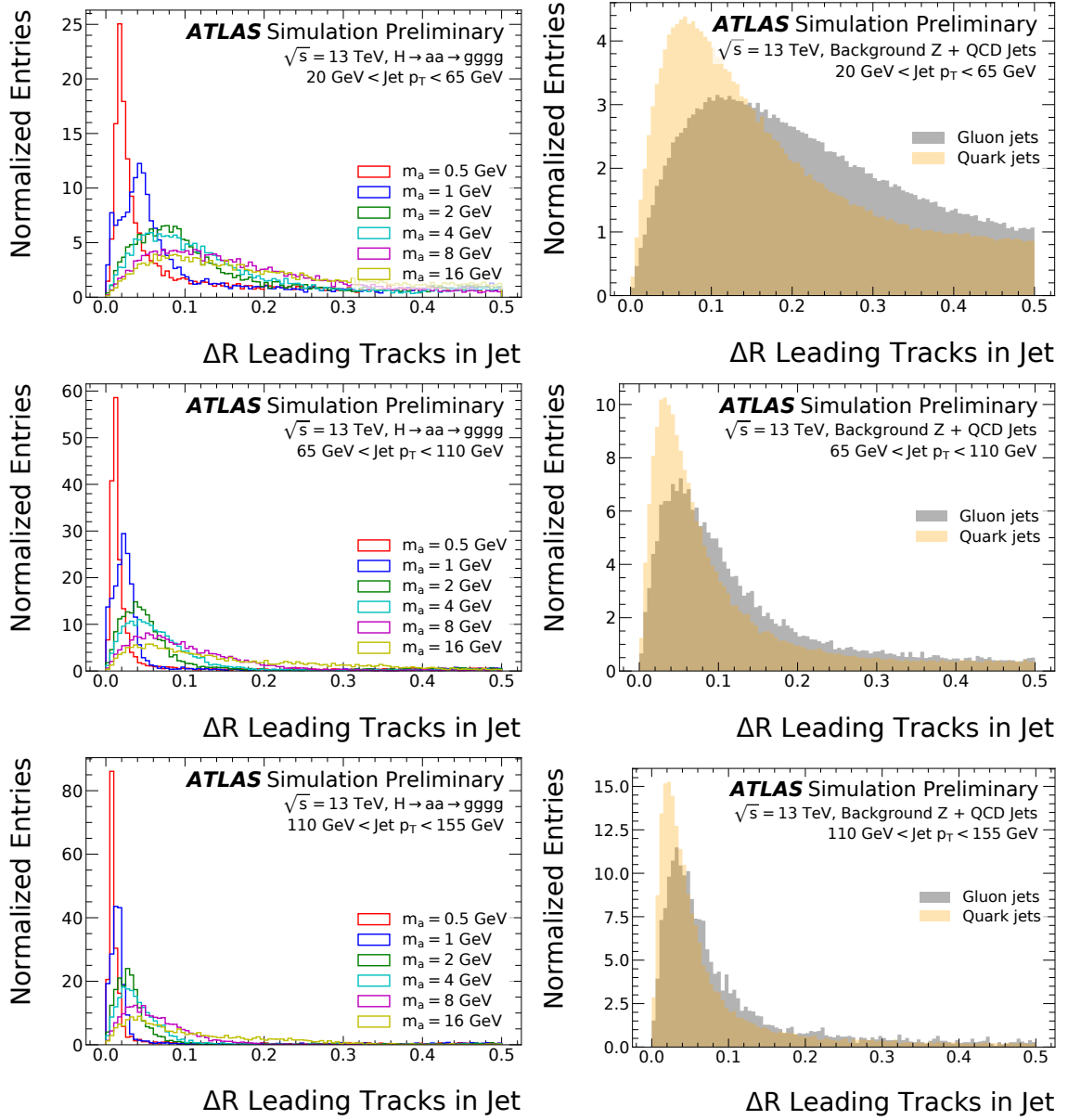


Figure 5.3: ΔR between the two leading tracks within jets ($\Delta R < 0.6$), for signal $a \rightarrow gg$ jets (left), background QCD quark/gluon jets (right). Different p_T bins are shown from top to bottom. Figures sourced from [32].

Figure 5.5 shows a roughly equal sharing of momentum by the leading constituents (the peak for the low-mass signals corresponds to $\sqrt{0.5^2 + 0.5^2} \sim 0.7$).

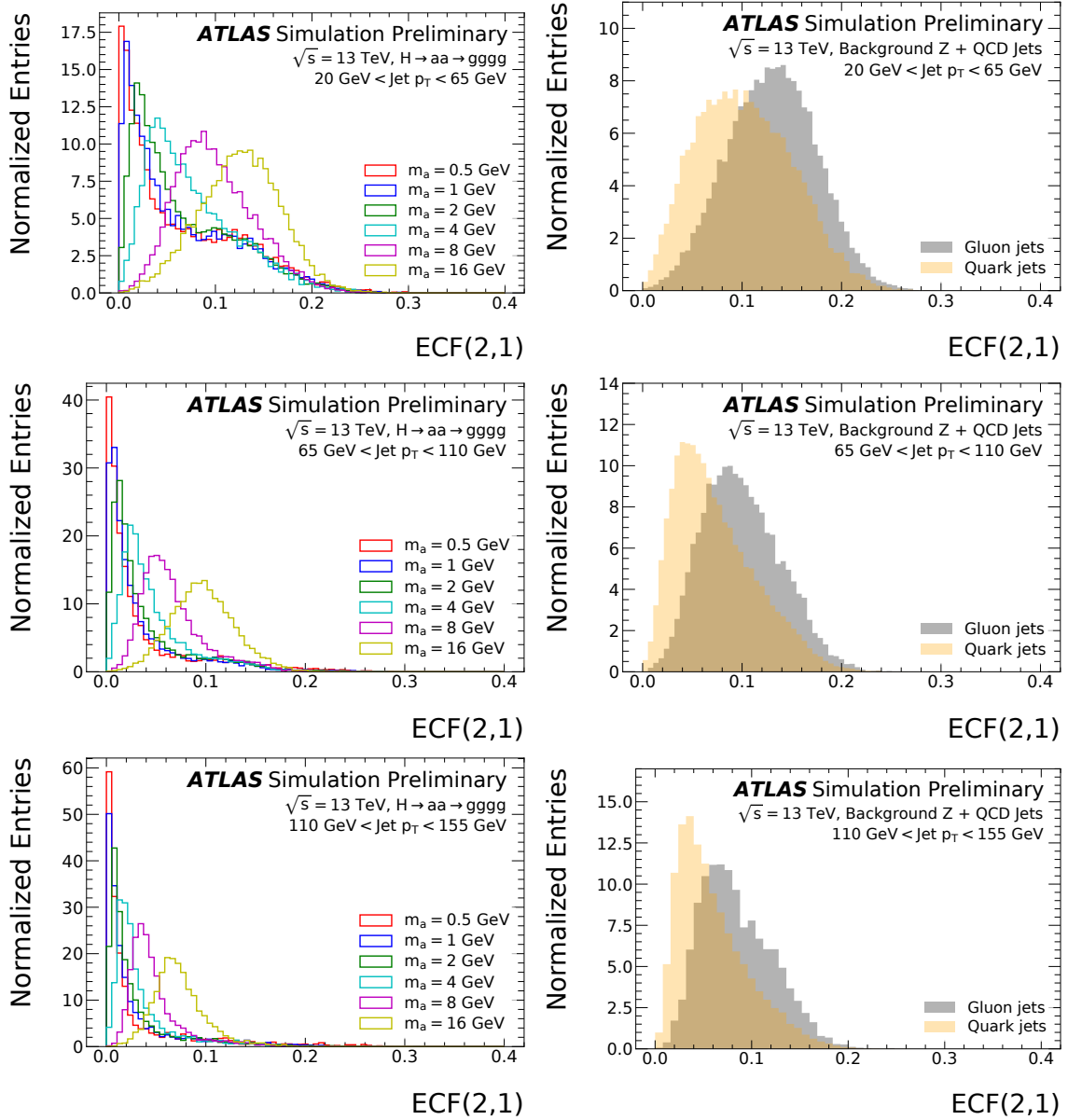


Figure 5.4: ECF(2,1) evaluated using tracks in the event associated to a given jet ($\Delta R < 0.6$), for signal $\alpha \rightarrow gg$ jets (left), background QCD quark / gluon jets (right). Different p_T bins are shown from top to bottom. Figures sourced from [32].

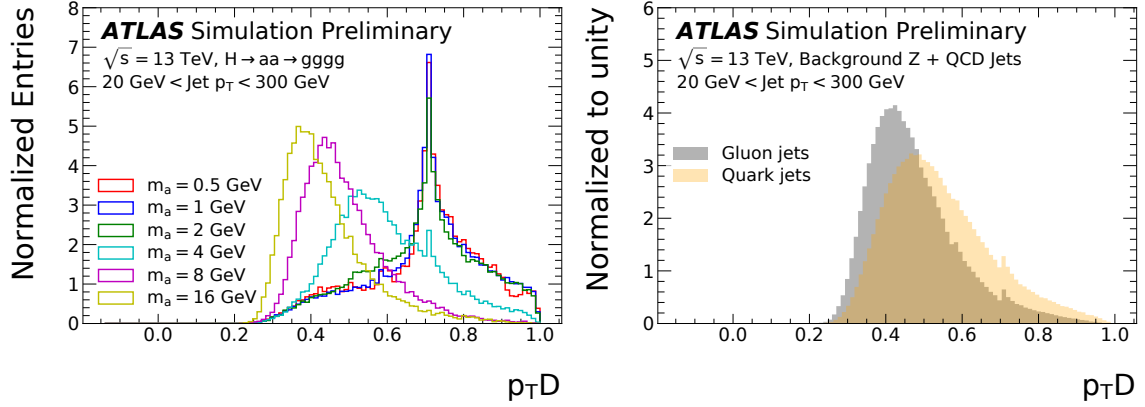


Figure 5.5: The jet substructure observable p_{TD} evaluated using tracks in the event associated to a given jet ($\Delta R < 0.6$), for signal $a \rightarrow gg$ jets (left), background QCD quark / gluon jets (right). p_{TD} varies slowly with p_T so only the nominal p_T bin is shown. Figures sourced from [32].

5.4.3 Machine Learning for $H \rightarrow aa \rightarrow gggg$

Machine learning is used in two different ways and both applications use the same neural network architecture (Section 5.4.3.1). The first application of machine learning is to implement the parameterized reweighting in (1) from Section 5.4.1. This makes use of the *Deep neural networks using Classification for Tuning and Reweighting* (DCTR) [342] protocol that is built on parameterized classifiers [343, 344] and make use of a well-known fact (see e.g. Ref. [345, 346]) that

$$\frac{f(x)}{1 - f(x)} \approx \frac{p(x|data)}{p(x|simulation)}, \quad (5.3)$$

for a classifier f trained using cross entropy to distinguish data from simulation. The second application of machine learning is to distinguish the signal from the reweighted background (using weights given by Eq. 5.3). Additional details are provided below.

5.4.3.1 Architecture

The neural network architecture is illustrated in Figure 5.6. It is comprised of two main components: a jet level tagger that takes in the variable length track inputs from each of the two leading jets and an event level tagger that uses global information. These two components are combined to return a single score per event.

The jet level tagger architecture is shown inside the dashed box in Figure 5.6. The tracks inside jets are represented as a variable-length set and are processed using a deep sets neural network [314], implemented as a particle flow network [313]. Each track is summarized by its $\mathbf{p}_T$ and $\Delta\mathbf{R}$ to the jet axis. The network operates by first embedding each track into a latent space and then summing all latent space vectors together. A fully connected network then acts on the summed vector. A similar setup is used for the Deep Sets based Neural Networks for Impact Parameter Flavour Tagging algorithm used by ATLAS to tag b-quark jets [347].

The event level tagger combines the information from the 2D latent space of both the leading jets from the event, as well as the information from the jet 4-vectors and the Z boson 4-vector obtained by adding the lepton 4-vectors from the event. This is done in a fully connected network that culminates in a single probability score for each event.

All networks are trained using KERAS [348] and TENSORFLOW [349] using the ADAM [350] optimization algorithm. A rectified linear unit (ReLU) activation function is used for all intermediate layers and a sigmoid is used for the final layer. The three fully connected layers in the jet level tagger and the event level tagger are comprised of 100 nodes each, and the latent space embedding layers in the jet level tagger are comprised of 100 nodes as well. None of these parameters were heavily optimized.

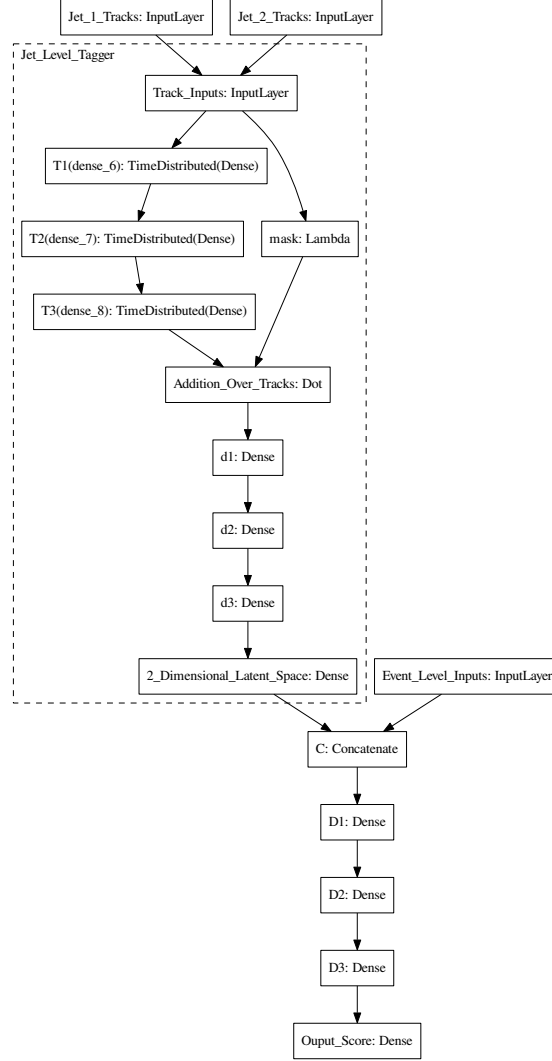


Figure 5.6: Architecture of the full event level tagger. The jet level tagger embeds each of the two leading jets from an event into a 2-D latent space, and uses the same set of weights to embed both jets. The time distributed layers embed each track into a high dimensional space using the same weights for each track. The Addition_Over_Tracks layer adds up the latent vectors across all the tracks, which results in a permutation invariant dimensional reduction across the tracks. Event level inputs and the final latent vectors from the two leading jets are used to derive the final output score for every event using fully connected layers. Figure sourced from [32].

5.4.3.2 Reweigher training

The reweighting function is learned using a well-known [345, 346] (also in high-energy physics [342, 343, 351–361]) fact that a classifier trained with the binary cross entropy loss function ($-\sum \mathbf{y}_i \log(f(\mathbf{x}_i))$ for labels $\mathbf{y}_i$ and inputs $\mathbf{x}_i$) learns to approximate a function monotonic with the ratio of class likelihoods, as mentioned at the top of Section 5.4.3. The features $\mathbf{x}$ here are all of the tracks inside the jets along with the jet and Z boson 4-vectors. While machine learning-based reweighting has been used before in data analyses [362, 363], this is the first high- and variable-dimensional reweighting [364] applied to collider data.

The performance of the reweighting is verified for the Pythia background sample in Figure 5.7. Figure 5.7a shows that the m_{jj} distribution is well-modeled after the reweighting, even in the region around the Higgs boson mass, which was excluded from the learning (N.B. only 5 fb^{-1} of 2018 data is used for validation to remain blinded to the signal region). Furthermore, since all tracks are used in the reweighting, it is expected that all high-level observables built from these tracks should also be well-modeled. This is illustrated in Figure 5.7b for the ECF(2,1) observable introduced earlier. A detailed set of plots showing the effects on the reweighting across a range of observables is shown in Appendix A.2, for reweighting trained of the nominal Pythia, sample as well as on the variation sample using the Sherpa generator that accounts for different fragmentation modelling. Plots shown in Section 5.4.2 are made using the reweighted background $Z + \text{QCD jets}$ MC events.

This reweighting is used for both training the signal versus background classifier and for evaluating its performance in the next sections.

5.4.3.3 Tagger training

During training, special attention is paid to ensure that the output score of the tagger is decorrelated with the invariant mass of the two leading jets in the event (m_{jj}) for the background $Z + \text{QCD jets}$ events³. This is done to ensure that the data driven

³Simultaneously training two classifiers that are decorrelated [33] was also explored. However, it was found to be technically and conceptually more effective to use the dijet mass as one of the observables and then let the neural network learn the second observable.

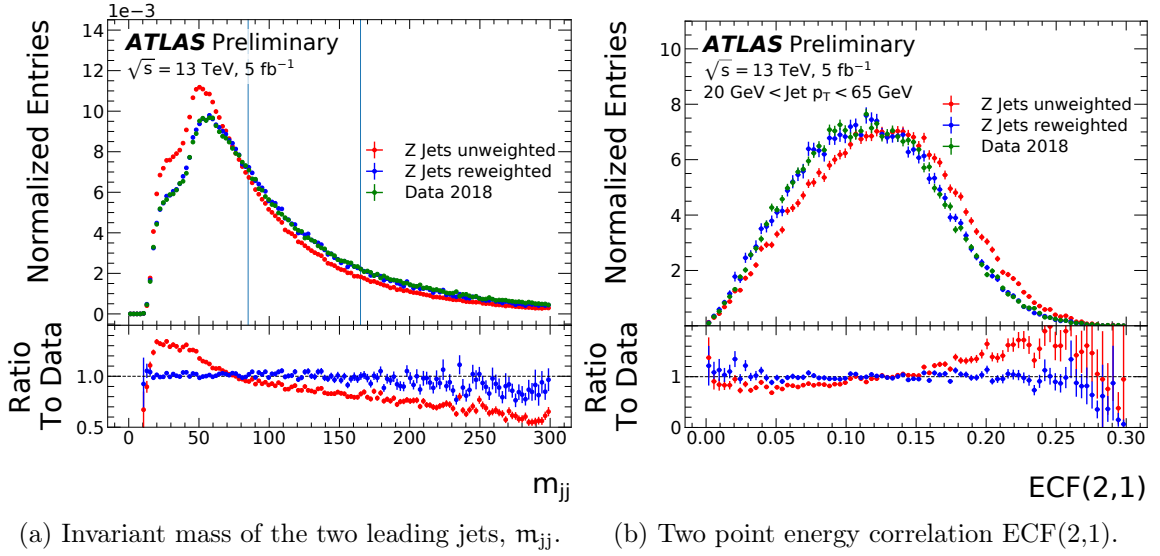


Figure 5.7: Distributions for data and $Z + \text{QCD jets}$ Pythia MC background before and after variable- and high-dimensional reweighting. Events with m_{jj} inside the signal region contained within the two vertical lines in (a) were excluded during training of the reweighting classifier. Only the lowest p_T bin is shown in (b) that displays the largest correction from the reweighting scheme. Figures sourced from [32].

ABCD background estimation strategy can be used to estimate the $Z + \text{QCD jets}$ events in the signal region when trying to find the signal $H \rightarrow aa \rightarrow gggg$ events in data. ABCD background estimation requires the two axes of the plane in which the four regions (A, B, C, and D) are defined to be independent for the background events, and in the case of this analysis the output of tagger along with m_{jj} form the two axes of the ABCD plane [33]. Then, one can estimate the background in the most signal-like region from the other three regions.

This is illustrated in Figure 5.8; the signal region is A and the number of background events $N_A \approx N_B N_C / N_D$. For this analysis, $f(x)$ is a neural network and $g = 1 - |1 - (m_{jj}/\mu_{m_{jj}|m_a})|$, where $\mu_{m_{jj}|m_a}$ is the average m_{jj} for the signal with mass m_a in the m_{jj} range $[50, 200]$ GeV/ c^2 . The $\mu_{m_{jj}|m_a}$ values are shown as dashed vertical lines in Figure 5.9. The observable g essentially turns a window cut around the peak of m_{jj} into a threshold cut so it is more easily visualized as a classical ABCD approach. Due to the event selection on m_{jj} , $0 \leq g \leq 1$. The signal tagger was also

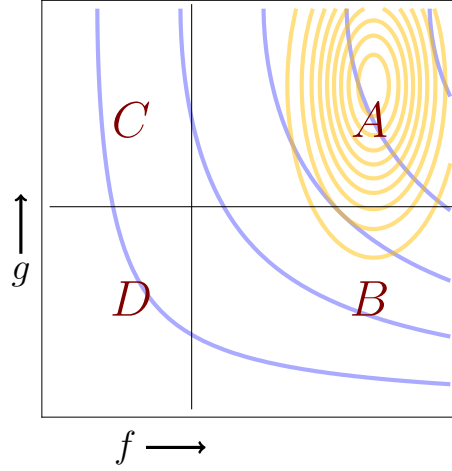


Figure 5.8: An illustration of the ABCD plane. Figure sourced from [33].

trained using g as the decorrelating variable against the tagger output but this did not result in any noticeable improvements in the speed/ stability of the training or the tagger performance. Therefore the signal tagger was trained using the invariant mass of the two leading jets as discussed above.

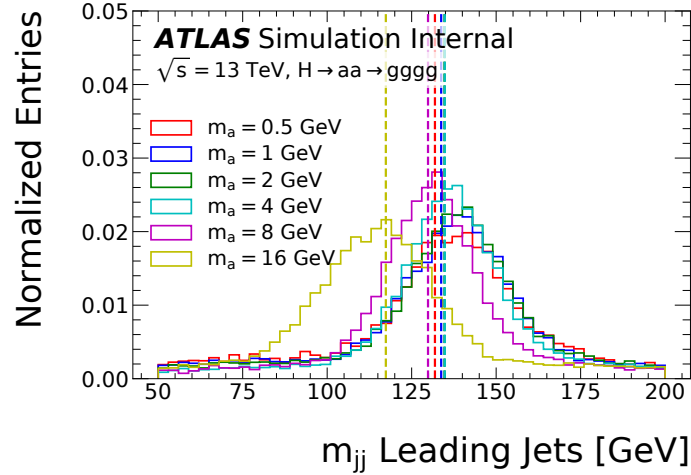


Figure 5.9: m_{jj} distribution for the signal $H \rightarrow aa \rightarrow gggg$ events, with $\mu_{m_{jj}|m_a}$ (the average m_{jj} in the range 50–200 GeV/ c^2) shown as dashed lines of the corresponding colors.

The independence of the tagger output and $\mathbf{m}_{jj}$ for the background class is enforced on the tagger by adding a penalty to the loss function that is proportional to a generalized correlation between the two distributions. The distance correlation metric [365–368] is used as the generalized correlation due to its ability to capture nonlinear dependence between distributions, and its differentiable form is expressed as a DisCo loss function [315] as follows,

$$\text{dCov}^2(X, Y) = \langle |X - X'| |Y - Y'| \rangle + \langle |X - X'| \rangle \langle |Y - Y'| \rangle - 2 \langle |X - X'| |Y - Y''| \rangle \quad (5.4)$$

$$\text{dCorr}^2(X, Y) = \frac{\text{dCov}^2(X, Y)}{\text{dCov}(X, X) \text{dCov}(Y, Y)} \quad (5.5)$$

$$\mathcal{L} = \mathcal{L}_{\text{classifier}}(\vec{\mathbf{y}}, \vec{\mathbf{y}}_{\text{true}}) + \lambda \text{dCorr}_{\mathbf{y}_{\text{true}}=0}^2(\vec{\mathbf{m}}_{jj}, \vec{\mathbf{y}}), \quad (5.6)$$

where $|\cdot|$ is the vector norm, (X, Y) , (X', Y') , (X'', Y'') are independent and identically distributed samples from the joint distribution (X, Y) . $\mathcal{L}$ is the training loss used in the tagger, where $\mathcal{L}_{\text{classifier}}$ is the binary cross entropy loss and the λ is a hyperparameter that is coarsely scanned for the best setting. Note here that $\mathbf{y}_{\text{true}} = 0$ selects the background events from the set of events that are being iterated over, ensuring that the $\mathbf{m}_{jj}$ - tagger output dependence is penalized only for the background class during training.

Additionally, care is also taken to ensure that the background $Z + \text{QCD jets}$ simulation events used to train the tagger do not deviate significantly from the data in regions that are expected to be devoid of signal. While this is not required to ensure reasonable background estimation since the ABCD method is data driven, it helps guarantee that the trained taggers are as close to optimal and have not learned any features that would be otherwise absent in data. This is done by reweighting the background events (see Section 5.4.3.2) using a classifier that is trained to distinguish between data and background before training the signal versus background classifier. The architecture of the reweighting classifier is identical to the main event tagger shown in Figure 5.6, where no distance correlation penalty is added to the loss function and instead $\mathbf{m}_{jj}$ is passed to the tagger as an event level input in order for it to learn the dependence on $\mathbf{m}_{jj}$ [343, 344] for interpolation into the signal region. The data

and background events are required to have $m_{jj} \leq 85$ GeV or $m_{jj} \geq 165$ GeV to ensure that the events come from the signal devoid regions of the phase space.

Furthermore, the tagger is trained using a background sample composed equally (by weights) of both Pythia and Sherpa $Z + \text{QCD jets}$ MC events. This is done to ensure that the tagger does not learn any quirks of a specific $Z + \text{QCD jets}$ MC generator that passed through the reweighting procedure described above. This has the additional benefit of reducing the differences between the predictions of the tagger for the two generators which reduces the uncertainty on the number of background events estimated in the signal region using the ABCD technique.

5.4.4 Signal tagger performance

5.4.4.1 $\alpha \rightarrow gg$ jet level tagger

Following the findings from Section 5.4.2, the tagger used to classify the signal $H \rightarrow \alpha\alpha \rightarrow gggg$ events and the background $Z + \text{QCD jets}$ events was designed and trained as described in Section 5.4.3. To gauge how effectively the tagger learns the differences between the signal di-gluon jets and the background QCD quark and gluon jets, the per-jet tagger is extracted from Section 5.4.3. A simple fully connected network is trained on node of the network just before being combined with the event information. This is required to transform the output into a classifier score. The resulting score is then used to make a Receiver Operator Characteristic (ROC) curve as shown in Figure 5.10. The ROC curves are made by scanning the discriminant with an optimized granularity to capture the variations in the distribution, and can result in scans stopping at some minimum signal efficiency for highly discriminating distributions. This procedure is described in greater detail in [32]. Interestingly, the performance for discriminating against quark jets (left plot of Figure 5.10) is lower than that for gluon jets (right plot of Figure 5.10), and both degrade for the higher scalar masses. This is further highlighted in Figure 5.11 where the ratio of gluon rejection to quark rejection is shown. This is consistent with the trends seen in Section 5.4.2 where the separation is greatest between the gluon and the di-gluon distributions for the lower scalar masses. The discriminating power degrades with higher masses, consistent with earlier sections that showed that the di-gluon jets using $R = 0.4$ start to resemble the background. For a di-gluon efficiency of 50%, a rejection of 100–1000 is achievable for the lower signal masses; in contrast, quark versus gluon tagging at the same efficiency can achieve a rejection of about 10 [369, 370]. Furthermore the performance of tagger is contrasted with the separation achieved from using the $\text{ECF}(2,1)$ observable in Figure 5.12, where it is apparent that at low mass, the tagger is able to learn a lot more about the differences between the QCD quark/gluon jets and the signal di-gluon jets than the simple substructure observables.

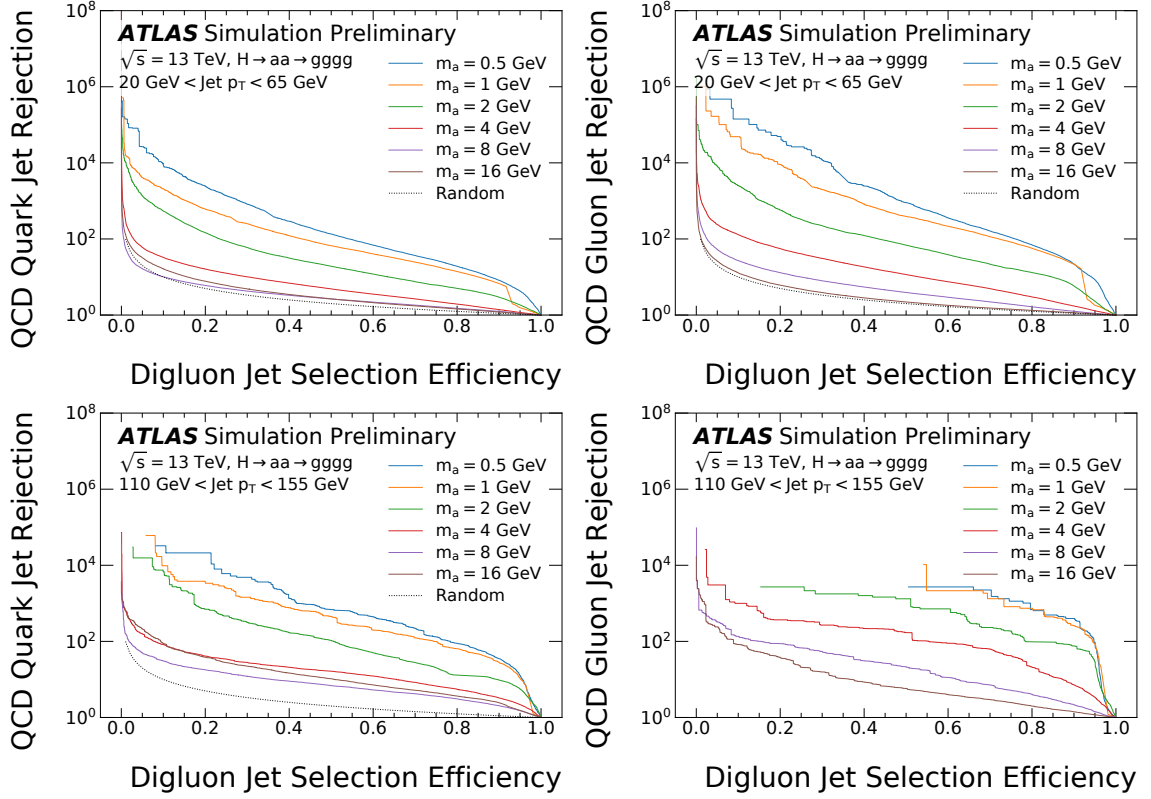


Figure 5.10: Receiver Operating Characteristic (ROC) curves of the jet classifier derived from the jet level tagger module of the full signal - background classifier from Figure 5.6. The background jets are from the same Z +jets sample used to train the tagger. QCD gluon jet rejection (right) is better than QCD quark jet rejection (left), with performance degrading as the scalar mass increases due to worsening collimation and increasing number of tracks in the di-gluon jets. Here the ROC curves for two jet p_T bins is shown (top - bottom). The lines labeled ‘Random’ correspond to di-gluon efficiency = $1/\text{QCD rejection}$. Figure sourced from [32].

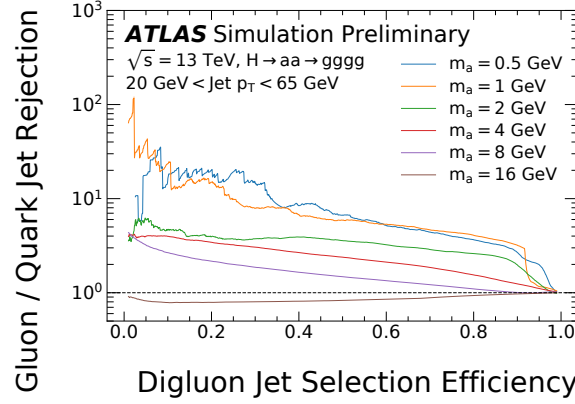


Figure 5.11: Ratio of the QCD gluon jet rejection to the quark jet rejection as a function of the di-gluon selection efficiency, for the lowest p_T bin. Figure sourced from [32]

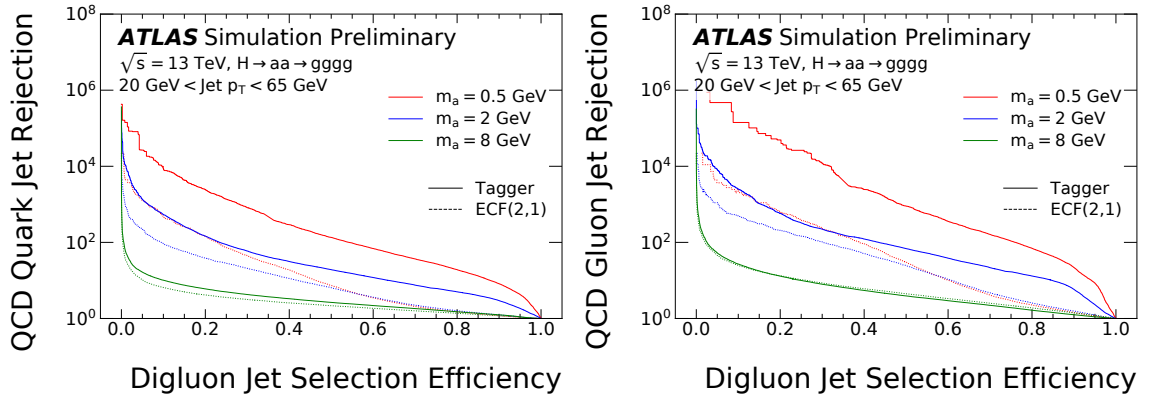


Figure 5.12: Receiver Operating Characteristic (ROC) curves derived using ECF(2,1) as the discriminating variable (dotted line) versus ROC curves derived from the jet level tagger module of the full signal - background classifier (solid line). Here the QCD gluon jet rejection (right) and QCD quark jet rejection (left) are shown for the lowest p_T bin from Figure 5.10 for three representative mass points. Note that these curves show that the rejection achieved using ECF(2,1) is lower than the rejection using the jet classifier, particularly for the lower signal masses. Figure sourced from [32]

5.4.4.2 $H \rightarrow a\bar{a} \rightarrow gggg$ event level tagger

The performance of the full signal tagger can be evaluated by composing the ROC curves as the thresholds along each of the two axes of the ABCD plane are varied. As described in Section 5.4.3.3, the tagger output along with (a transformed) m_{jj} form the two axes of the ABCD plane in this analysis.

ROC curves made using four values of the DisCo loss weight λ (Eqn. 5.6) are shown in Figure 5.13. These plots are made by first scanning a range of cuts along each of two axes in the ABCD plane, namely the m_{jj} axis and the tagger output axis. This results in a 2D grid of working points, each corresponding to a signal efficiency and background rejection. Working points that result in signal regions that produce a background (Pythia $Z + \text{QCD jets}$) non-closure greater than 25% are discarded. Furthermore, working points that result in statistically limited regions (A,B,C, or D) for background MC events are rejected by discarding points where the statistical relative error on the ABCD correction factor (or closure) is greater than 25%. This condition is imposed using events from both background MC generators (Pythia and Sherpa) separately. The remaining points are then binned by signal efficiency, and for each bin the maximum background efficiency is chosen as the statistic to plot. The resulting ROC curves for the different mass points are shown in Figure 5.13.

From Figure 5.13 it can be seen that high background rejection can be achieved for low scalar masses, with the discriminating power diminishing towards the higher scalar masses. This is consistent with the trends spotted in Section 5.4.2 and Section 5.4.4.1.

The performance of the decorrelation achieved using the DisCo loss during training and the background reweighting procedure described in Section 5.4.3.3 can also be studied from the structure of the ABCD plane. Figure 5.14 and Figure 5.15 shows the distribution of events in the ABCD plane for data as well as $Z + \text{QCD jets}$ MC events, using a logarithmic color scale for the Z -axis. This highlights the efficacy of the background reweighting procedure as a large chunk of the ABCD plane shows excellent agreement between data and $Z + \text{QCD jets}$ MC generated.

The final cuts along each axis are chosen such that the signal efficiency divided by the square root of the background efficiency is maximized for each signal mass tagger

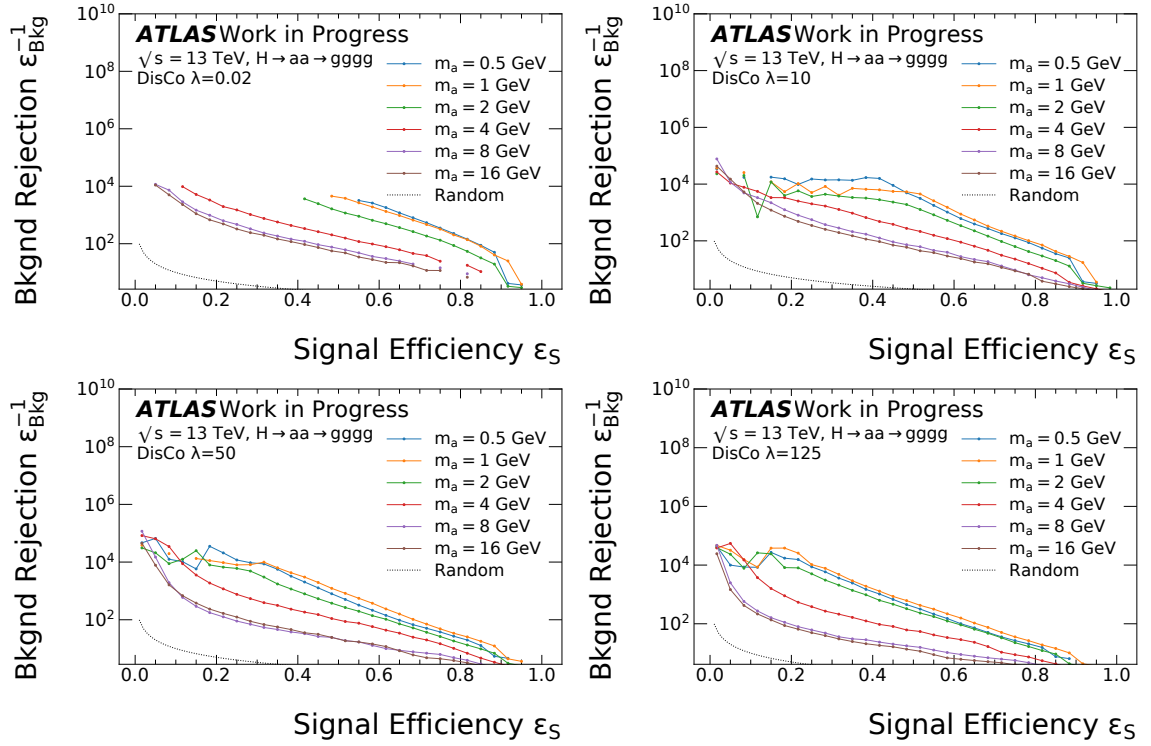


Figure 5.13: Receiver Operating Characteristic (ROC) curves of the full signal - background classifier from Figure 5.6. (top left) DisCo weight $\lambda = 0.02$, (top right) DisCo weight $\lambda = 10$, (bottom left) DisCo weight $\lambda = 50$, (bottom right) DisCo weight $\lambda = 125$. The lines labeled 'Random' correspond to the curve signal efficiency = $1/\text{background rejection}$.

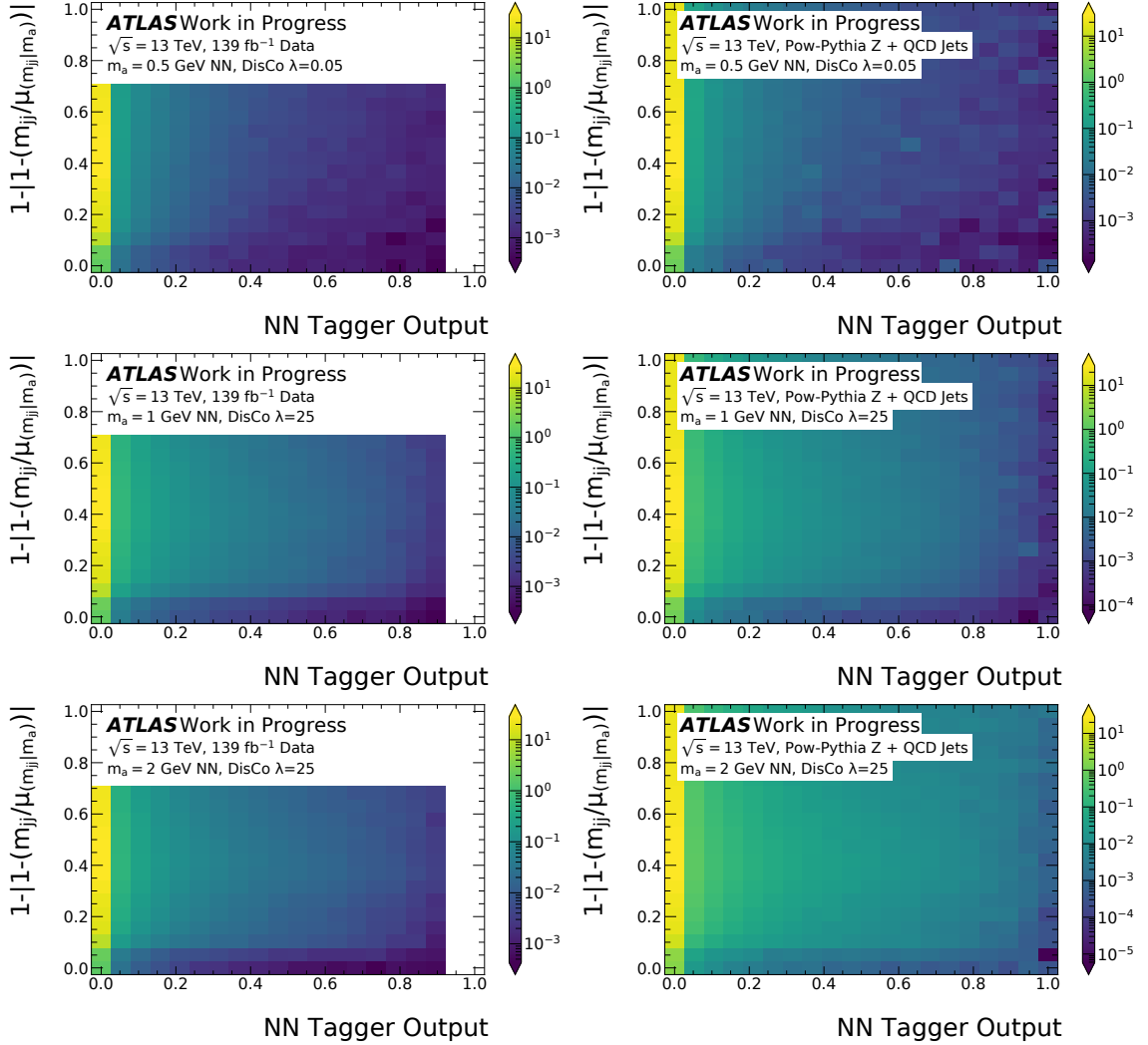


Figure 5.14: (Left) ABCD plane for Data where the region of tagger output > 0.9 and m_{jj} axis > 0.7 is masked to blind the signal region, (right) ABCD plane for Pythia Z + QCD jets events. The Z-axis color scale is logarithmic and shows the normalized distribution of events in each bin. The scalar mass m_a goes from 0.5 GeV/ c^2 to 2 GeV/ c^2 from top to bottom as indicated on the plots. The axes of the ABCD plane are the neural network output and $1 - |1 - (m_{jj}/\mu_{m_{jj}}|m_a)|$, where $\mu_{m_{jj}}|m_a$ is the average m_{jj} for the signal with mass m_a in the m_{jj} range $[50, 200]$ GeV/ c^2 .

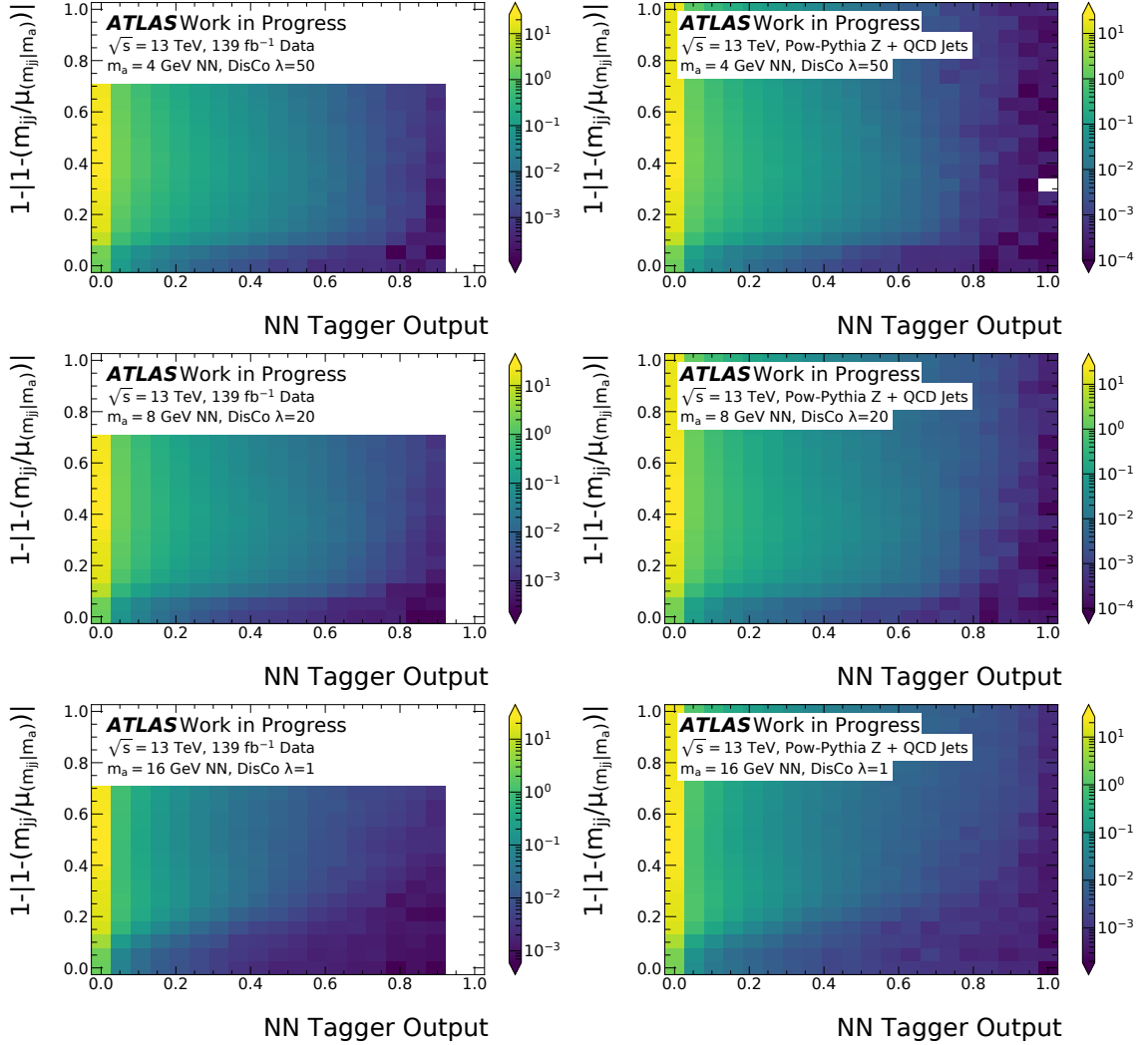


Figure 5.15: (Left) ABCD plane for Data where the region of tagger output > 0.9 and m_{jj} axis > 0.7 is masked to blind the signal region, (right) ABCD plane for Pythia Z + QCD jets events. The Z-axis color scale is logarithmic and shows the normalized distribution of events in each bin. The scalar mass m_a goes from 4 GeV/ c^2 to 16 GeV/ c^2 from top to bottom as indicated on the plots. The axes of the ABCD plane are the neural network output and $1 - |1 - (m_{jj}/\mu_{m_{jj}}|m_a)|$, where $\mu_{m_{jj}}|m_a$ is the average m_{jj} for the signal with mass m_a in the m_{jj} range $[50, 200]$ GeV/ c^2 .

individually. Additionally only cuts that result in signal regions that produce a background (Pythia $Z + \text{QCD jets}$) non-closure less than 25% are allowed. Furthermore as with the ROC curve, only cuts that result in the relative uncertainty on the ABCD closure of less than 25% for both the background MC generator events (Physics and Sherpa) are allowed. To ensure that the selection is immune to fluctuations as the ABCD cuts are scanned, an averaging kernel of size 3 is applied to the grid of 2D values to ensure that the thresholds on the nominal generator closure are applied pessimistically.

The values for λ are not heavily optimized and only a coarse grid search was performed on the range $[0.02, 400]$ to find the optimal values for each mass point. Based on this procedure, the $\lambda = 0.05$ best working point is used for the scalar mass $m_a = 0.5 \text{ GeV}/c^2$, the $\lambda = 25$ best working point is used for the scalar mass $m_a = 1 \text{ GeV}/c^2$, the $\lambda = 25$ best working point is used for the scalar mass $m_a = 2 \text{ GeV}/c^2$, the $\lambda = 50$ best working point is used for the scalar mass $m_a = 4 \text{ GeV}/c^2$, the $\lambda = 20$ best working point is used for the scalar mass $m_a = 8 \text{ GeV}/c^2$, and the $\lambda = 1$ best working point is used for the scalar mass $m_a = 16 \text{ GeV}/c^2$. See Appendix A.3 for a table listing the Pythia $Z + \text{QCD jets}$ events non-closures, and the Sherpa and Pythia $Z + \text{QCD jets}$ events relative uncertainty on the closures for all the lambda values tested. Also see Appendix A.5 for plots of the expected limits on the signal strength for the scanned range of λ values.

Figure 5.16 and Figure 5.17 show the final selected working points for each signal mass point as red crosses over three relevant 2D planes. The left plots show (on the Z -axis) the signal efficiency divided by the square root of the background efficiency for each pair of cuts in the ABCD plane scanned. The middle plots show (on the Z -axis) the background (Pythia $Z + \text{QCD jets}$) closure relative uncertainty for each pair of cuts scanned. And finally the plots on the right show (on the Z -axis) the background (Pythia $Z + \text{QCD jets}$) non-closure for each pair of cuts scanned. Note that the selected working points satisfy the thresholds on the background closures and live inside neighborhoods where the local closures are all under control, justifying the working point values and the averaging kernel approach.

Fig 5.18 and Fig 5.19 show the final selected working points for each signal mass

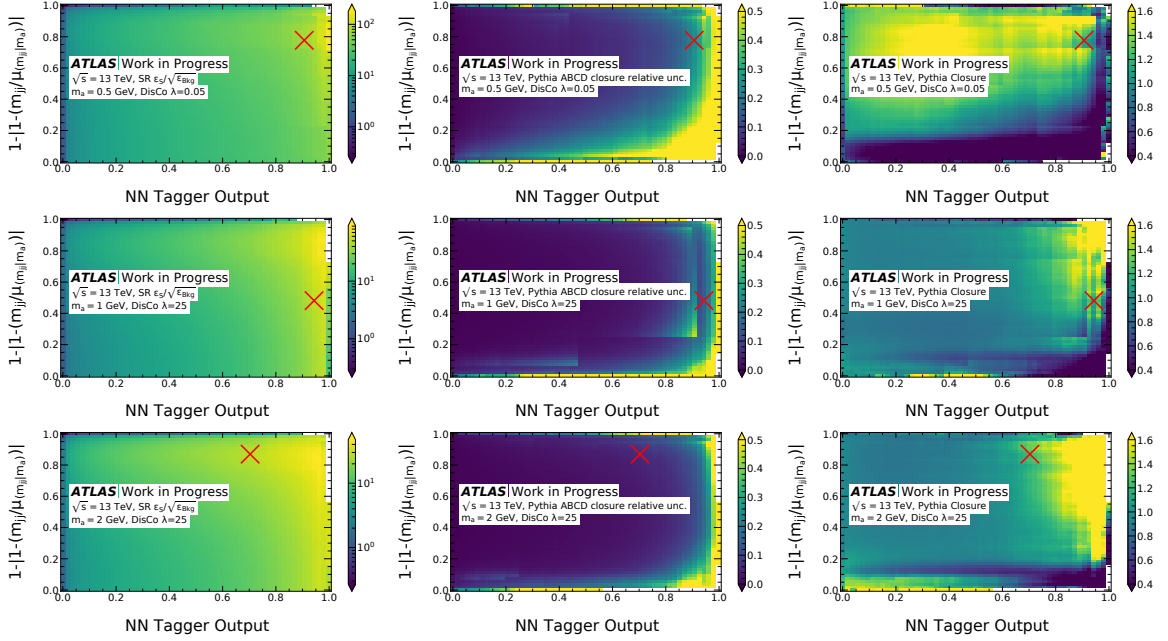


Figure 5.16: (Left) Signal efficiency divided by the square root of the background (Pythia Z + QCD jets events) efficiency for a set of pairs of cuts along the ABCD axes, (middle) background (Pythia Z + QCD jets) closure relative uncertainty for a set of pairs of cuts along the ABCD axes, (right) background (Pythia Z + QCD jets) non-closure for a set of pairs of cuts along the ABCD axes. The scalar mass m_a goes from 0.5 GeV/c² to 2 GeV/c² from top to bottom as indicated on the plots. The closure (right) is limited to a range of [40, 160]%, and the relative uncertainty on closures (middle) is limited to a range of [0, 50]%; All values outside of these range are colored by the extreme colors on either side of the ranges.

point as red crosses over similar three relevant 2D planes. The left plots show (on the Z axis) the signal efficiency divided by the square root of the background efficiency for each pair of cuts in the ABCD plane scanned. The middle plots show (on the Z axis) the background (Sherpa Z + QCD jets) closure relative uncertainty for each pair of cuts scanned. And finally the plots on the right show (on the Z axis) the background cross-generator (Pythia - Sherpa) closure difference for each pair of cuts scanned.

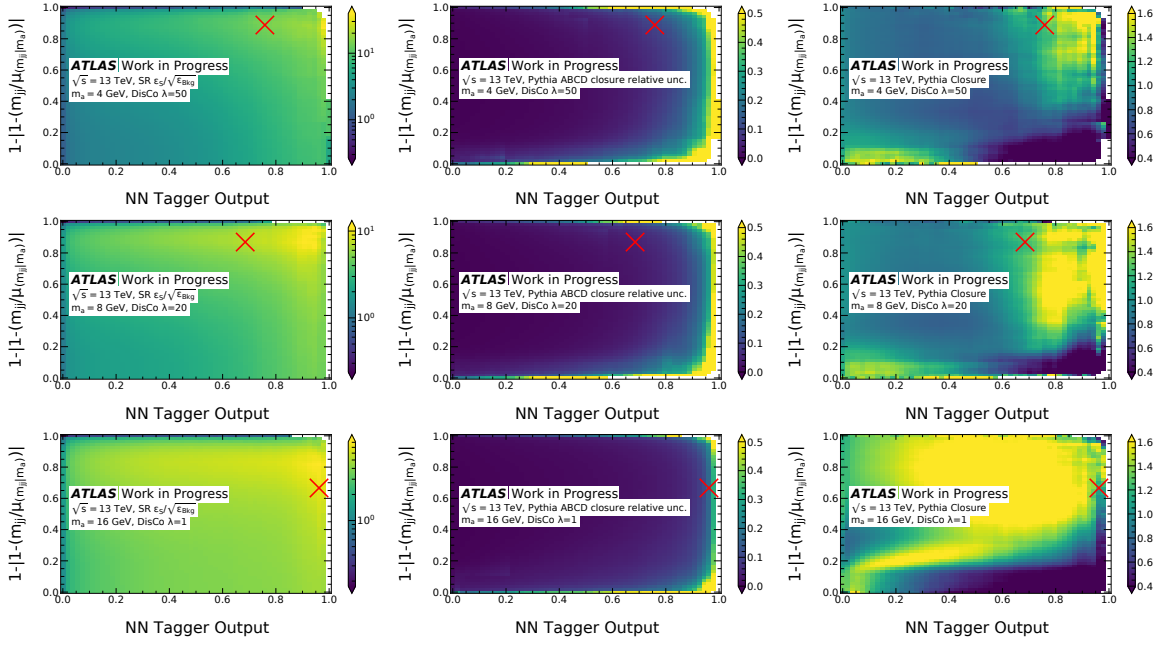


Figure 5.17: (Left) Signal efficiency divided by the square root of the background (Pythia Z + QCD jets events) efficiency for a set of pairs of cuts along the ABCD axes, (middle) background (Pythia Z + QCD jets) closure relative uncertainty for a set of pairs of cuts along the ABCD axes, (right) background (Pythia Z + QCD jets) non-closure for a set of pairs of cuts along the ABCD axes. The scalar mass m_a goes from 4 GeV/ c^2 to 16 GeV/ c^2 from top to bottom as indicated on the plots. The closure (right) is limited to a range of [40, 160]%, and the relative uncertainty on closures (middle) is limited to a range of [0, 50]%; All values outside of these range are colored by the extreme colors on either side of the ranges.

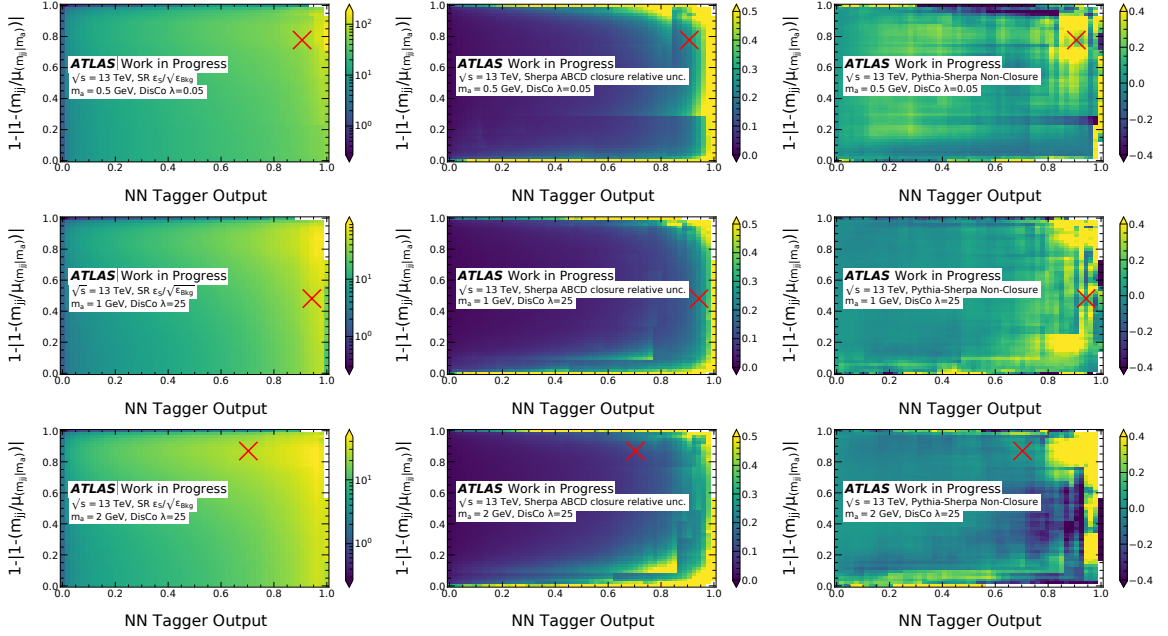


Figure 5.18: (Left) Signal efficiency divided by the square root of the background (Pythia $Z + \text{QCD jets}$ events) efficiency for a set of pairs of cuts along the ABCD axes, (middle) background (Sherpa $Z + \text{QCD jets}$) closure relative uncertainty for a set of pairs of cuts along the ABCD axes, (right) difference of the closures in the ABCD plane between Pythia and Sherpa $Z + \text{QCD jets}$ events for a set of pairs of cuts along the ABCD axes. The scalar mass m_a goes from 0.5 GeV/ c^2 to 2 GeV/ c^2 from top to bottom as indicated on the plots. The relative uncertainty on closures (middle) is limited to a range of $[0, 50]\%$, and the difference in closures (right) is limited to a range of $[-40, 40]\%$; All values outside of these range are colored by the extreme colors on either side of the ranges.

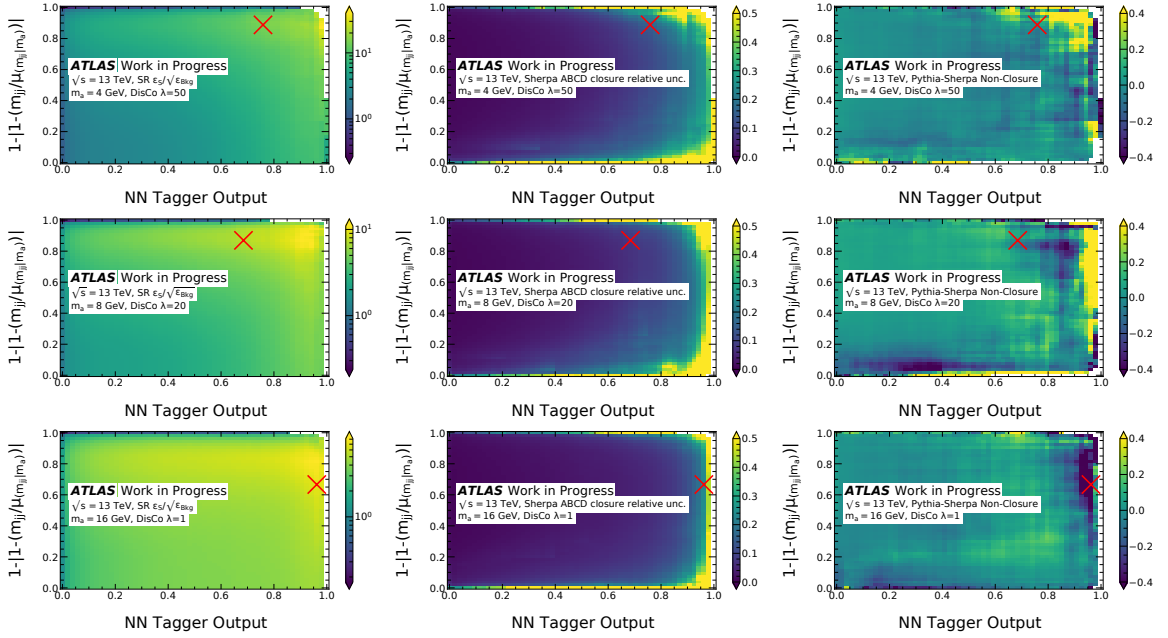


Figure 5.19: (Left) Signal efficiency divided by the square root of the background (Pythia $Z + \text{QCD jets}$ events) efficiency for a set of pairs of cuts along the ABCD axes, (middle) background (Sherpa $Z + \text{QCD jets}$) closure relative uncertainty for a set of pairs of cuts along the ABCD axes, (right) difference of the closures in the ABCD plane between Pythia and Sherpa $Z + \text{QCD jets}$ events for a set of pairs of cuts along the ABCD axes. The scalar mass m_a goes from 4 GeV/ c^2 to 16 GeV/ c^2 from top to bottom as indicated on the plots. The relative uncertainty on closures (middle) is limited to a range of $[0, 50]\%$, and the difference in closures (right) is limited to a range of $[-40, 40]\%$; All values outside of these range are colored by the extreme colors on either side of the ranges.

5.5 Systematic Uncertainties

There are two types of uncertainties to consider in this analysis: background uncertainties and signal uncertainties. The background is determined completely from data using the ABCD method and so we are only uncertain on the closure of the ABCD method. This is determined by comparing the closure with Pythia and Sherpa. For the signal, the analysis is completely dependent on simulation and so each component of the simulation needs to be varied to determine the impact on the final signal yield.⁴

5.5.1 Track uncertainties

The systematic uncertainties associated with the tracks include four global reconstruction efficiency related uncertainties, one to account for variations to the tracking efficiency when the track is inside a jet. Another uncertainty relates to the variations to the fake rate. And finally, the last uncertainty varies the p_T of the track based on the residual alignment uncertainties. These uncertainties can be seen as broadly falling into two categories: ones related to the reconstruction efficiency of the tracks, and ones related to the measurement of the track parameters for reconstructed tracks — these are detailed below.

Uncertainties in track reconstruction are driven by the material in the tracking layers and the uncertainty in the knowledge of the material. These can be estimated using simulations where the tracking material is varied, to study the resulting track reconstruction efficiency as a function of the track p_T and η [232, 233, 371]. The impact of this reconstruction efficiency can be estimated by randomly sampling tracks, dropping them with a probability that is a function of the track p_T and η and accounts for variations in the reconstruction due to the tracking material.

Likewise, track reconstruction inefficiency can arise due to tracks in regions of high density in the detectors being lost, such as in the cores of jets [372, 373]. The source of this can be traced to clusters in the detector that may be incorrectly identified

⁴The analysis currently implements the tracking and signal modeling uncertainties, which are expected to be dominant. Others uncertainties will be introduced into the analysis over the course of the full study.

as single particle clusters despite having contributions from multiple tracks, which is likely to happen in high density regions of the detector such as jets — the effect gets worse as the transverse momentum of the jets grow since this leads to increasing collimation of the charged particles. The effect of this reconstruction efficiency is also evaluated by randomly dropping tracks, this time as a function of the jet p_T with which the tracks are associated.

Track reconstruction efficiency is also impacted by the presence of fake tracks, or tracks that are formed by combining unrelated hits in the detector, often from multiple particle trajectories. The impact of fake tracks can be estimated by incorporating a dropping probability for fake tracks in the simulation samples, analogously to the methods described above for real tracks.

Variations in the reconstructed track p_T can arise due to misalignment of the tracker as well as the finite resolution of the tracker hits. The momentum resolution for isolated tracks can be studied using high quality $Z \rightarrow \mu\mu$ and $J/\psi \rightarrow \mu\mu$ events from captured data [281]. Correction factors on the track momentum are taken as systematic uncertainties on the resolution, and are parameterized by means of independent standard normal random variables. To estimate the impact of these uncertainties, the track momentum scale can be varied upwards and downwards by one standard deviation of the variation.

5.5.2 Pileup reweighting uncertainties

As discussed in section 5.3.2.1, a correction is applied to MC events such that the pileup modeling in the simulation corresponds to what is seen in data. As such, there are variations associated with this modeling that must be taken into account.

In addition, a scaling correction is applied to the data during pileup reweighting. Nominally, this scaling correction is equal to $1/1.03$, and the up and down variations are set to $1/1.072$ and $1/0.988$ respectively.

5.5.3 Muon uncertainties

5.5.3.1 Muon efficiency uncertainties

As summarized in section 5.3.2 there are several scale factors associated with muon reconstruction related to efficiencies in the trigger, reconstruction, isolation and track-to-vertex association. Hence, there are 2 systematic variations associated with the trigger efficiency, 2 for the isolation efficiency, 4 for the reconstruction efficiency, and 2 for the track to vertex association efficiency.

5.5.3.2 Muon calibration uncertainties

There are also systematic variations associated with the calibration applied to the muons. These will have an affect on the muon kinematics. The systematic variations include one for momentum variations due to measurements made in the inner detector, one for momentum effects related to the muon spectrometer, one for the momentum scale, and 2 systematics related to the measured sagitta value.

5.5.4 Modeling uncertainties

Due to the inability to calibrate the $a \rightarrow gg$ tagger with data, in addition to the tracking experimental uncertainties, modeling uncertainties related to fragmentation modeling need to be considered. This is evaluated using alternative signal simulation samples that introduce various modeling tunes to account for the differences and variations from modeling errors. In particular, for each mass point, we have signal samples that have all of the A14 tune variations (1u/d, 2u/d, 3au/d, 3bu/d, 3cu/d) as well as a set of Herwig7 samples. Due to a feature of the Herwig7 hadronization model, it is not possible to have $m_a < 4$ GeV. Therefore the 4 GeV uncertainty is used for $m_a < 4$ GeV.

5.6 Results

Using the tagger described in Section 5.4.3 and Section 5.4.4 the expected exclusion limits on the signal $\alpha \rightarrow gg$ production can be studied. Additionally, error bands around the expected limits can be established using the tagger along with samples of signal and background events where sources of uncertainties have been introduced. These uncertainties along with the nominally limit setting is described below.

5.6.1 Expected Limits

This analysis uses an ABCD background estimation strategy as described in Sec. 5.4.1 and Sec. 5.4.3.3. The signal region is treated as a single bin, and the signal and background events in the signal region form the single bin Poisson counting experiment for limit setting. The background counts in the signal region (N_A) are found by multiplying the ABCD estimate of the number of background events in the signal region ($N'_A = N_B \times N_C / N_D$ where D is diametrically opposite to the signal region) by the ABCD closure derived from the MC background events (closure = $N_A^{\text{MC bkgd}} / N'_A{}^{\text{MC bkgd}}$).

From the tagger performance and ROC curves as shown in Section 5.4.4.2 the cuts along the NN and the m_{jj} axes in the ABCD plane that define the signal region are found for each scalar mass point. These cuts result in a given signal efficiency and background efficiency on the MC samples used for training and evaluation purposes in this analysis (detailed in Section 5.2). To convert these efficiencies into expected number of events, multiplication by the expected production cross sections, branching ratios, and luminosity is required. For a full Run 2 analysis we assume the luminosity to be 139 fb^{-1} (see Section 5.2.1).

For the signal production, the production cross section of $Z \rightarrow \ell\ell + H$ is assumed to be 29.7 fb (Table 13 in [374]) (uncertainty on this is ignored for now). The limits are set on the product $\text{Br}(H \rightarrow \alpha\alpha) \times \text{Br}(\alpha \rightarrow gg)^2$, and under the assumption that $\alpha \rightarrow gg$ is dominant ($\text{Br}(\alpha \rightarrow gg) \sim 100\%$) this can be effectively treated as a limit on the $\text{Br}(H \rightarrow \alpha\alpha)$. The expected number of signal events in the signal region is then defined as the signal efficiency times the luminosity, times the $Z(\ell\ell)H$ production

cross section, times the fraction of events in the MC signal sample that pass the event selection cuts.

For the background events, the ABCD estimate derived for the full 139 fb^{-1} of Run 2 data is used. The ABCD closure factor applied to the signal counts is derived from the Powheg+Pythia MC background $Z \rightarrow \ell\ell + \text{QCD jets}$ events.

The statistical model used to set the upper limits on the $\text{Br}(H \rightarrow a\bar{a}) \times \text{Br}(a \rightarrow g\bar{g})^2$ includes two channels: one for the signal $H \rightarrow a\bar{a}$ events and one for the $Z \rightarrow \ell\ell + \text{QCD jets}$ background events. The two channels share a few correlated shape uncertainty modifiers, namely the track efficiency uncertainty, track fake uncertainty, track jet uncertainty, and the track p_T scale (See Sec. 5.5.1). The counts for these modifiers varied high and low are taken by zeroing out the tracks that fail the boolean checks or have their p_T varied higher or lower, and passing the modified inputs to the signal tagger to get the modified event scores. The background channel has an uncorrelated shape uncertainty modifier that accounts for the variation in background event counts, which is evaluated using the non-closure between the two MC generators (Pythia - Sherpa) in the bin. Additionally, the background channel has another uncorrelated shape uncertainty modifier that accounts for subdominant background processes, which is evaluated using the non-closure between the Pythia generated ($Z \rightarrow \mu\mu$) and ($t\bar{t} \rightarrow \mu\mu + X$) simulation samples in the bin. The signal channel also incorporates correlated shape uncertainty modifiers for A14 tune variations (1u/d, 2u/d, 3au/d, 3bu/d) and cross-generator differences using the Herwig7 samples (See Sec. 5.5.4). Signal strength ($\text{Br}(H \rightarrow a\bar{a}) \times \text{Br}(a \rightarrow g\bar{g})^2$) hypotheses are tested in the range 0-10 and the upper limits are extracted based on a series of CLs tests, these are shown in Fig. 5.20. Impact and pull plots for the fits are shown in Appendix A.6. Additionally, details about the statistical fitting procedure is described in Appendix A.4.

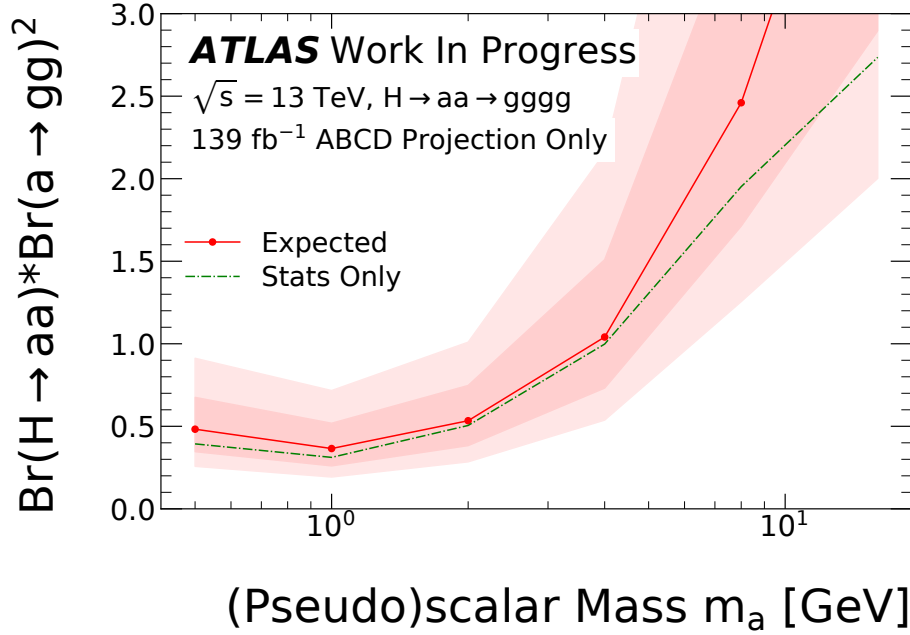


Figure 5.20: Expected upper limits on $\text{Br}(H \rightarrow aa) \times \text{Br}(a \rightarrow gg)^2$ across the range of signal a masses tested. Note here that the *Stats Only* line shows the expected limits when no sources of uncertainty are considered, fitting only to the Poisson counting experiment with the expected signal and background counts.

5.6.2 Blinded Analysis and Next Steps

This analysis is currently blinded as it is undergoing the editorial approval procedure within the ATLAS collaboration. It utilizes the di-gluon tagging technique and machine learning methods developed by the Author in Ref. [32], in collaboration with Ben Nachman and Ariel Schwartzman. There are a few sources of uncertainty that will be added to this analysis prior to full publication, such as jet calibration scale uncertainties, as well as the non-closure from non dominant backgrounds. Furthermore, aspects of the hyperparameter selection will be further optimized to improve the performance of the tagging methods, and potentially reduce the impact of certain sources of uncertainties. As mentioned in the Introduction, the analysis is targeting a full publication in Fall 2022, following the implementation of the above mentioned steps.

5.7 Conclusion

This chapter has presented the first search for $H \rightarrow a\bar{a} \rightarrow gg$. The range of the BSM scalar masses targeted boosts the decay products into a single small-radius jet, which is identified with a tagger designed to extract the unique radiation pattern from single $a \rightarrow gg$ jets.

A deep neural network based tagger makes use of every track inside candidate di-gluon jets to train a high dimensional classifier that can process a variable number of tracks per jet. Moreover the tagger uses an innovative technique to enforce the necessary decorrelation needed for background estimation. Additionally it uses simulated background events for training that have been reweighted to better match data using a variable- and high-dimensional reweighting scheme that consequently matches jet substructure distributions along with event level observables.

This search is sensitive to branching ratios of $H \rightarrow a\bar{a} \rightarrow gg$ of $\mathcal{O}(20\%)$ up to about 2 GeV, after which the sensitivity significantly degrades due to the large background and wider angle radiation. In the future, it would be exciting to explore the low(er) mass regime and to consider $a \rightarrow q\bar{q}'$ as well as displaced a decays. The all-hadronic decays of the new scalar to light quarks would share much of the same substructure as the di-gluon jets, and the decay to heavy quarks like charms would introduce the additional information from the secondary charm decay vertices.

Chapter 6

Novel Hard-Scatter Primary Vertex Selection in the ATLAS Experiment

This chapter discusses a new primary vertex selection algorithm that has been developed by the Author and collaborators (A. Schwartzman) for the ATLAS experiment [36,37]. This algorithm is designed to improve the efficiency for events with low visible p_T such as invisible Higgs boson decays in vector-boson fusion (VBF) production channel.¹ The technique utilizes information from calorimeter jets to supplement the information from the tracks, and mitigates the impact of pile-up vertex merging. For the limited tracker angular acceptance of ATLAS ($|\eta| < 2.5$) in Run 1/2/3 conditions, an event classification scheme is also proposed to handle events with jets that lie outside the ATLAS ID. The algorithm improves the average vertex selection efficiency by approximately 10%, under tight VBF Higgs invisible event selection cuts.

¹Referred to as VBF Higgs invisible events in the remainder of this text.

6.1 Introduction

Identifying the hard-scatter (HS) interaction amongst the set of reconstructed vertices in each event is vital to the physics analysis conducted using ATLAS data (Section 4.6.2). The default hard-scatter vertex selection is based on the principle of maximising $\sum \mathbf{p}_T^2$ (*sumPT2*), where the transverse momentum ($\mathbf{p}_T$) squared of the tracks associated to the vertex are summed. This selection criterion is likely to pick the interesting physics interaction that passed the trigger (See Figure 4.11), and bypass the other soft-QCD pile-up vertices in the event.

This selection criteria works reasonably well ($> 90\%$ efficiency) for final states that have hard physics objects within the tracker geometry (hence boosting *sumPT2*). This can be seen in Figure 6.1 where the vertex reconstruction efficiency times the selection efficiency is shown for the same final states as Figure 4.11). Note that the vertex reconstruction efficiency is close to 100% [27], and so Figure 6.1 is illustrative primarily of the vertex selection efficiency. The *sumPT2* selection has low selection

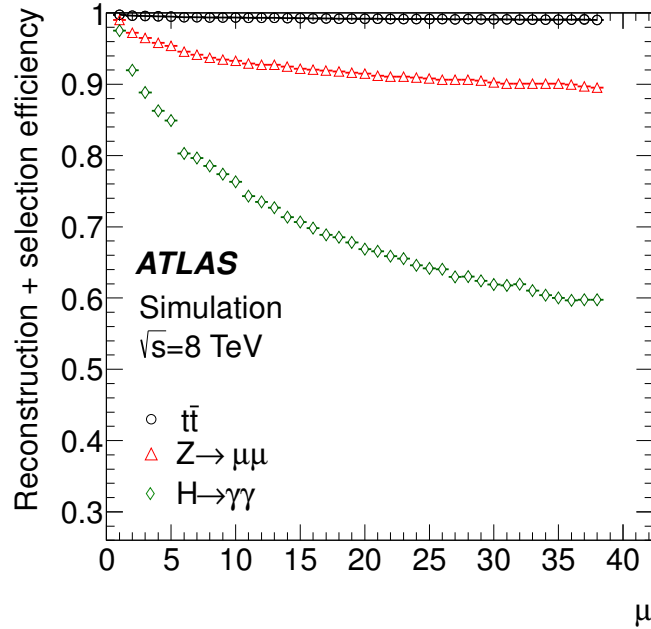


Figure 6.1: Hard-scatter vertex reconstruction and selection efficiency versus the average number of interactions μ . Figure sourced from [27].

efficiency in final states with low visible track p_T , where visible corresponds to tracks that are within the ID geometry. This can happen in final states that contain forward or low p_T jets, such as VBF Higgs invisible events (e.g. $H \rightarrow ZZ^* \rightarrow 4\nu$, see Figure 6.2) in vector-boson fusion (VBF). VBF Higgs invisible final state is chosen as

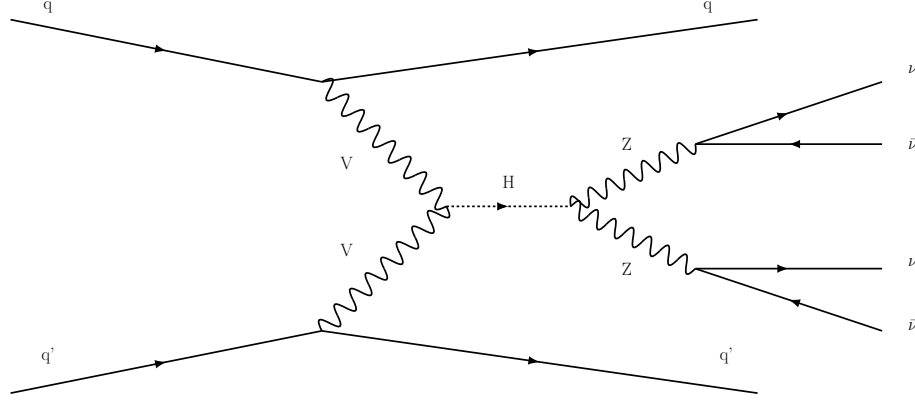


Figure 6.2: Feynman diagram for Vector Boson Fusion initiated $H \rightarrow ZZ^* \rightarrow 4\nu$. Quarks radiated from the vector bosons form the characteristic two forward VBF jets. The tracks inside the VBF jets are the crucial handles that are used to identify the hard-scatter vertex due to the neutrinos in the final state not interacting with the detector.

the case study since it is particularly sensitive to BSM couplings of the Higgs boson to DM, and is an important analysis in the ATLAS physics portfolio [375].

As discussed in Section 4.6.2, ATLAS uses a specialized diphoton vertex selection algorithm that identifies the hard-scatter vertex using the photons since the sumPT2 efficiency is low for final states such as $H \rightarrow \gamma\gamma + \text{MET}$ [244, 245]. The algorithm introduced here is distinct in that it does not use detector objects beyond tracks and calorimeter jets, and is hence generally applicable across final states other the particular case of VBF Higgs to invisible decays as considered here.

Two original ideas are introduced to address the challenges that limit sumPT2, and to improve the efficiency for VBF Higgs invisible events. Firstly, a new vertex selection algorithm is shown that integrates information from tracks with calorimeter jet information to counter the low visible track p_T issue. This algorithm improves the vertex selection efficiency by about 10%. Secondly, the concept of vertex confidence is introduced to tackle events with forward jets that lie outside the tracker. Events with

significant jet p_T outside the tracker angular acceptance ($|\eta| < 2.5$) are classified as low vertex confidence events. For VBF Higgs invisible events, the proposed algorithm achieves vertex selection efficiency of 97% for high confidence events. The remaining low confidence events are not assigned a hard-scatter vertex, though they still retain the VBF jet signature.

ATLAS will have an upgraded tracking system, the ATLAS Inner Tracker (ITk) [208, 209], for the HL-LHC era. One of the key benefits of the upgraded ITk will be its expanded angular reach of $|\eta| < 4$, illustrated in Figure 6.3. It will have improved longitudinal (along beam axis) impact parameter resolution for tracks when compared to the current ATLAS ID system, thanks to its 50 μm pixel pitch compared to the 250 – 400 μm pixel pitch in the ID system [34]. Furthermore the ITk strip detectors will have higher precision than the ID TRT system and will result in improved momentum resolution [34]. These planned upgrades for the ATLAS tracker will be essential to reconstruct and resolve tracks and vertices at the high luminosities ATLAS will experience at the HL-LHC — the average number of interactions per event are expected to be as high as $\mu = 200$ (Section 3.4). This is illustrated in Figure 6.4 where a simulated pp bunch crossing in HL-LHC conditions containing a $t\bar{t}$ interaction is shown within the volume of the ATLAS ITk system. The proposed selection algorithm leads to a gain in selection efficiency of approximately 10% in the Run 4 HL-LHC conditions.

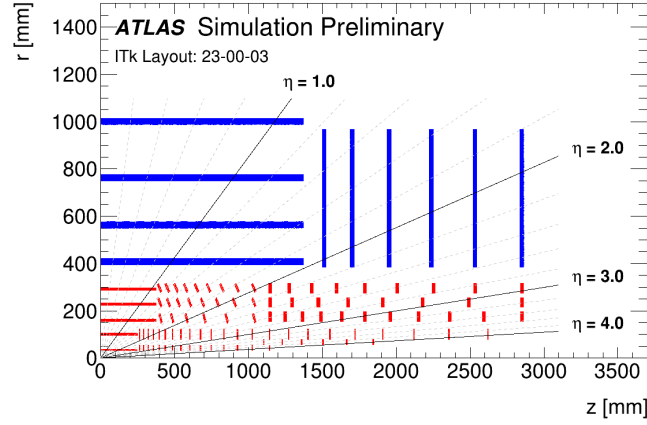


Figure 6.3: Plot showing a single quadrant layout of the ITk. Strip detector active elements are shown in blue, and pixel detector active elements are shown in red. The origin of the horizontal axis is the nominal interaction point, and the vertical axis is the radius measured from the origin. Figure sourced from [34].

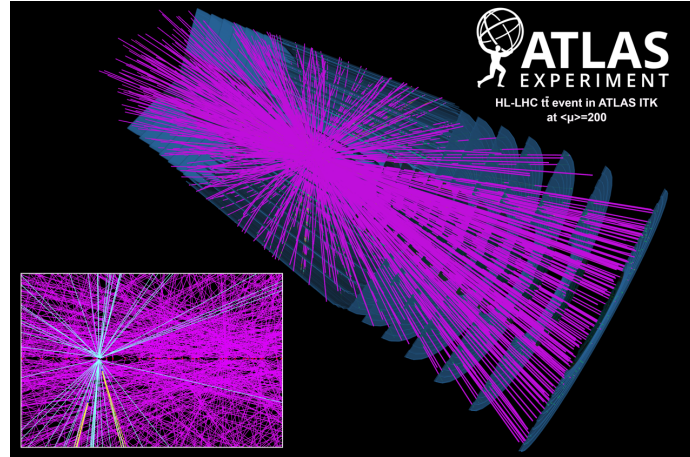


Figure 6.4: Simulated event display of an HL-LHC event in the ATLAS ITk detector. The event has an average number of collisions per bunch crossing of $\mu = 200$. The inset on the bottom left shows a longitudinal cross section where the different pp interactions can be seen as the origin of the corresponding tracks. Figure sourced from CERN [35].

6.2 Datasets and Event Selection

Samples of MC simulated $H \rightarrow ZZ^* \rightarrow 4\nu$ events in ATLAS Run 3 and Run 4 conditions are used for the study. The process is modeled using POWHEG-BOX v2 [291, 292, 316] integrated with PYTHIA 8.2 [287, 317] using the AZNLO [318] set of tuned parameters and the CT10NLO Parton Distribution Function (PDF) set [296].

The center-of-mass energy is set to be 13 TeV and 14 TeV for Run 3 and Run 4 respectively. The mean number of pp interactions μ is also set to 60 and 200 per bunch crossing for the two runs respectively. The ATLAS ITk geometry and reconstruction as described in Ref. [376] are used for the Run 4 sample.

Since this is a study of vertex selection efficiency, it is important to ensure that the theoretical maximum vertex selection efficiency is 100%. This can be done by imposing a physics motivated event selection scheme that allows only events which have an identifiable vertex of interest. This also presents a challenge to the study of vertex selection efficiency since typical physics driven event selection criteria are biased by the choice of the primary vertex to begin with. An example of this is the pile-up jet rejection criteria (Section 4.6.5.3) that requires jets to have a large fraction of their track p_T pointing to the selected hard-scatter vertex.² This is achieved by imposing the condition that jet $R_{p_T} > 0.1$, where R_{p_T} [377] is defined as:

$$R_{p_T} = \sum_{\text{tracks}} \frac{p_{T(\text{track})}}{p_{T(\text{jet})}} \theta_H(0.4 - \Delta R). \quad (6.1)$$

The sum here is only taken over tracks associated to the selected hard-scatter vertex, θ_H is the Heaviside function, ΔR is the angular distance between the track and jet in the $\eta - \phi$ space. This R_{p_T} criterion necessitates the selection of a hard-scatter vertex, and so biases the measurement of the vertex selection efficiency. For example in VBF Higgs invisible events where the selected vertex is not the correct hard-scatter interaction, the VBF jets will not pass the R_{p_T} criterion with respect to the incorrectly selected vertex. Since the VBF jets in these events are effectively discarded, these events will fail the VBF topology based selection and will not be part of the set of

²Track p_T here refers to the p_T of the tracks that are within the angular cone of the jet.

events that the vertex selection efficiency is evaluated over. This will therefore lead to an overestimation of the vertex selection efficiency since the set of events that are used to evaluate the efficiency are made up primarily of events where the VBF Higgs invisible interaction was correctly selected in the first place.

To overcome this, an unbiased event selection scheme is defined for VBF Higgs invisible events where every event is required to have a pair of jets that satisfy the VBF jet criteria. Therefore all identifiable VBF Higgs invisible events are considered in the evaluation of vertex selection efficiency regardless of the *a priori* choice of vertex. The VBF jet criteria used here is as follows:

- Leading (hardest p_T jet) jet has $p_T > 70$ GeV.
- Subleading (second hardest p_T jet) jet has $p_T > 50$ GeV.
- The two leading jets have $\Delta\eta > 3$.
- The two leading jets have $\Delta\phi < 2.6$.

6.3 Current Approach and its Limitations

As described in Section 4.6.2, the default method to identify the hard-scatter vertex in events is to select the vertex that maximizes the summed track transverse momentum squared (sumPT2),

$$\text{sumPT2} = \sum_{\text{tracks}} p_T^2. \quad (6.2)$$

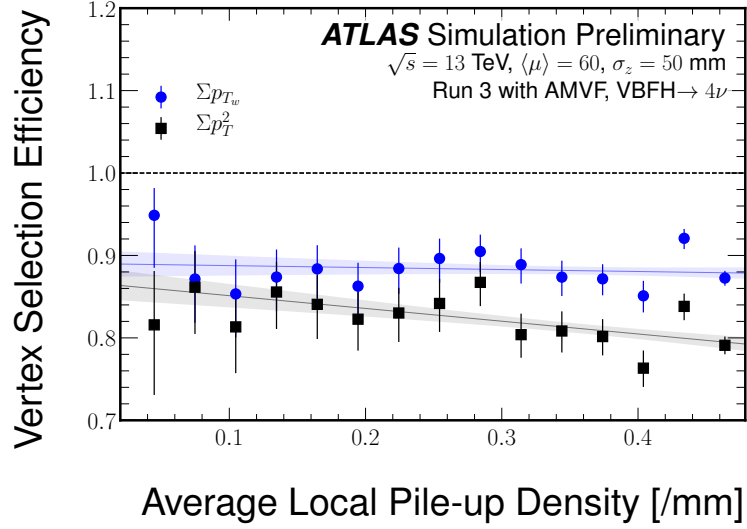
Besides the limitation from low visible p_T interactions, as shown in Figure 6.1 for $H \rightarrow \gamma\gamma$, this method is also vulnerable to pile-up vertex merging. This manifests as decreasing vertex selection efficiency as the number of pp interactions (μ) increases in the events, also visible in Figure 6.1. This is because the algorithm is preferentially selecting high p_T vertices that result from the merging of nearby pile-up vertices when the distance between the pile-up vertices is smaller than the vertex resolution (Figure 4.9b and Figure 4.10b).

Figure 6.5a illustrates the downward trend of the sumPT2 selection efficiency³ for VBF Higgs invisible events with increasing average local pile-up vertex density.⁴ This effect is exacerbated for VBF Higgs invisible events due to low visible p_T from the non-interacting particles in the final state. The issue of low visible p_T is highlighted in Figure 6.5b which shows that the efficiency of sumPT2 is (inclusively⁵) 88% in HL-LHC conditions where PU vertex merging is mostly addressed due to the improved resolution of the ITk. This residual inefficiency can be attributed to low visible p_T and not to pile-up vertex merging, since no significant downward trend with average local PU vertex density is seen for sumPT2 in Figure 6.5b.

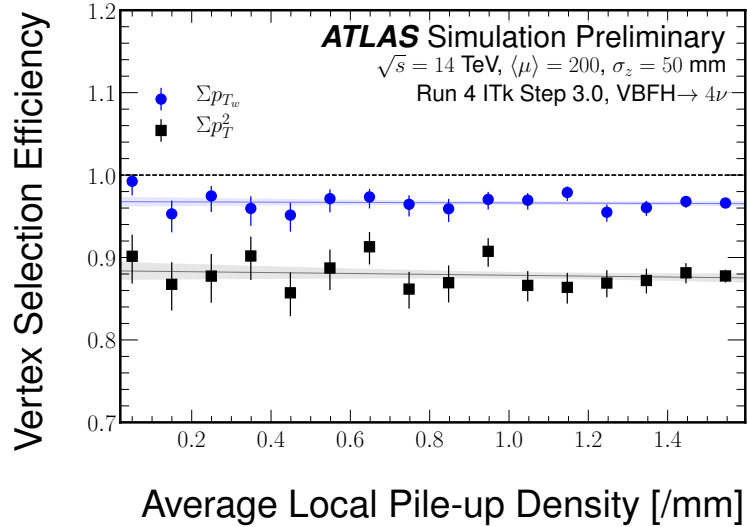
³A vertex is declared to have been chosen correctly if the Δz between the selected vertex and the true hard-scatter interaction location (known for simulations) along the beam axis is less than 1.6 mm. This threshold is motivated by the close-by vertex resolution from Figure 4.10b.

⁴The average local pile-up vertex density is the density of vertices evaluated at the location of the selected vertex, using a Gaussian vertex distribution with a width of 50 mm along the beam axis and centered at the nominal interaction point.

⁵Inclusive of local pile-up density in this context. This value tends to be very close to the selection efficiencies for the highest local pile-up densities, since the distribution of vertices is approximately a Gaussian density centered at the nominal interaction point.



(a)



(b)

Figure 6.5: Vertex selection efficiency for VBF Higgs invisible events versus the average local density of vertices at the primary vertex in the event. The default sumPT2 selection method (square) is compared with the new sumPTw selection (circle). Error bars are statistical errors. Best fit linear functions are drawn with 1-sigma error bands to visualize the error on the slope with respect to the pile-up density. (a) LHC Run 3 conditions, (b) HL-LHC Run 4 conditions. Note the higher efficiencies at Run 4 than Run 3 are due to the increased tracking acceptance of the ITk which reconstructs tracks inside forward hard-scatter jets. Figure sourced from [36].

6.4 New Technique

A novel vertex selection criterion is developed to address the limitations discussed in Section 6.3. Angular correlations between the tracks from the hard-scatter interaction and the high p_T hard-scatter jets in the event can be used to discriminate against pile-up vertices. Tracks from merged pile-up vertices are not strongly correlated with hard-scatter jets or with each other, because the tracks come from close-by independent interactions. Therefore by only selecting tracks from each vertex that are angularly correlated to hard jets, it is possible to resolve the substructure associated with every vertex and mitigate the impact of pile-up merging. Furthermore, the high p_T nature of hard-scatter jets in general can be leveraged to address the challenge of low visible p_T in events like VBF Higgs invisible.

Following these insights, a new criterion for hard-scatter vertex selection in VBF Higgs invisible events is proposed that utilizes track-jet correlations and picks the vertex that maximises $sumPTw$, an augmented version of the default summation:

$$sumPTw = \sum_{\text{tracks}} p_{T_w} \quad (6.3)$$

$$= \sum_{\text{tracks}} \frac{p_{T(\text{track})}^2 p_{T(\text{closest jet})}^2}{\Delta R} \theta_H(0.4 - \Delta R) \theta_H(p_{T(\text{jet})} - p_{T(\text{threshold})}). \quad (6.4)$$

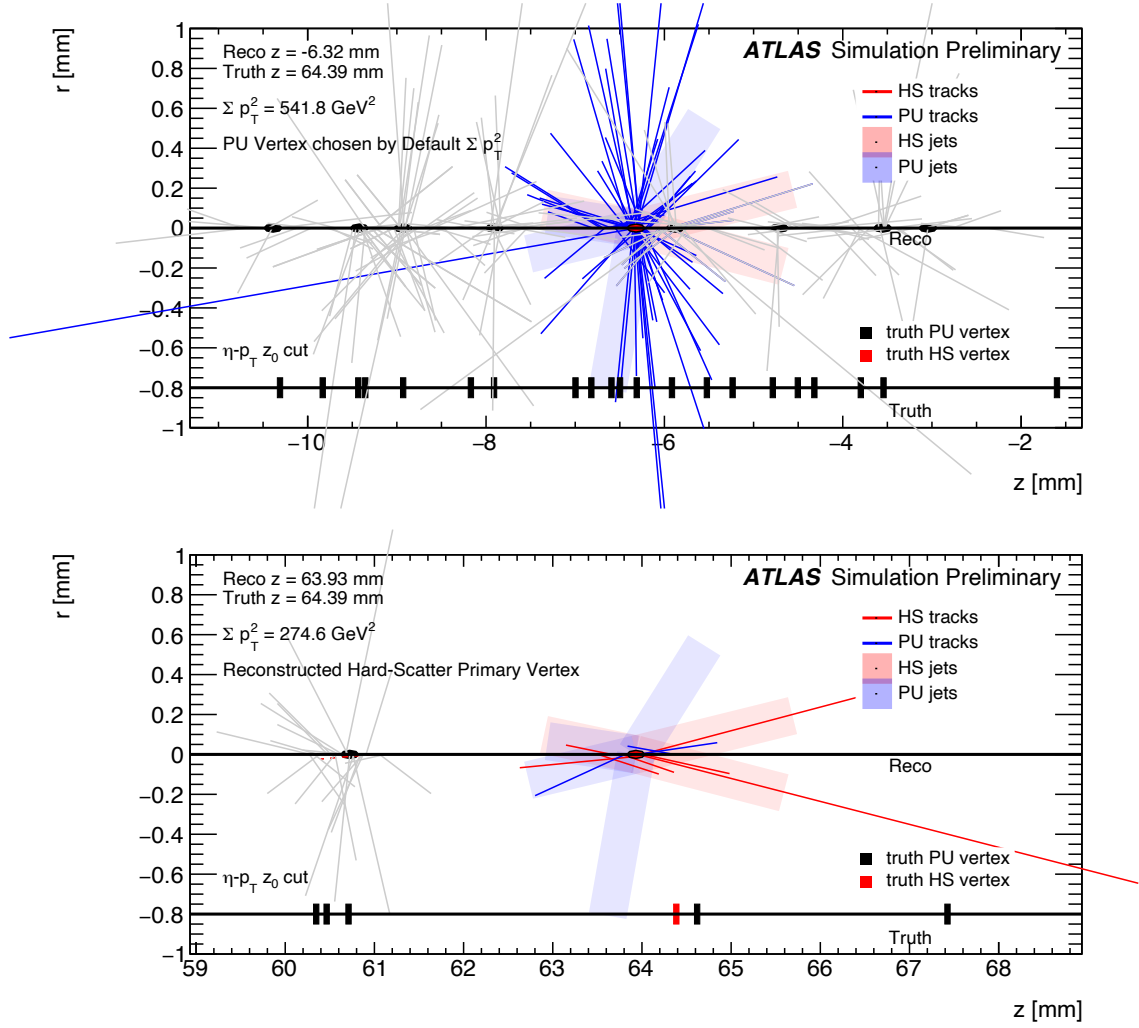
Here θ_H is the Heaviside function, and ΔR is the distance in $\eta - \phi$ space between a track and the closest jet with sufficient p_T . The jet p_T threshold used for this study is $p_{T(\text{threshold})} = 30$ GeV, and ensures that only hard (likely hard-scatter) jets are considered in the summation. The action of the Heaviside functions can be understood as follows: For every track, the closest jet with p_T greater than the threshold is found and the track is only considered if it lies within the specified angular cone of the jet. Furthermore, the use of the $p_{T(\text{closest jet})}$ introduces a bias weight term that increases the sum for hard-scatter interactions since only hard-scatter tracks will be close to hard jets in the $\eta - \phi$ space. Therefore $sumPTw$ is a measure of the total vertex $sumPT2$ inside hard jets, effectively projecting tracks onto hard jet axes and adding the weighted p_{T_w} as defined above.

Since merged PU vertices come from independent interactions, the tracks associated to these vertices are not correlated in the $\eta - \phi$ space in any direction, or with the hard-scatter interaction. Hence projecting tracks onto hard jet axes suppresses their otherwise large number of tracks, reducing their total $\mathbf{p}_T$ contribution.

Figure 6.5 shows the performance of the sumPTw criterion versus sumPT2. The sumPTw criterion displays a weaker dependence of the selection efficiency on pile-up density, as highlighted by the linear fit to the points on the plot in Figure 6.5a. This illustrates the robustness of sumPTw against the pile-up merging phenomenon. sumPTw improves the (inclusive) average selection efficiency from 80% to 88% in Figure 6.5a for the Run 3 scenario.

sumPTw also improves the (inclusive) average selection efficiency from 88% to 97% in Figure 6.5b for the Run 4 scenario with the ATLAS ITk. As discussed in Section 6.3, sumPT2 is primarily limited at Run 4 for VBF Higgs to invisible events to the low visible hard-scatter $\mathbf{p}_T$. From Figure 6.5b it can be seen that sumPTw largely mitigates this limitation thanks to the jet $\mathbf{p}_T$ bias weight term in the expression for $\mathbf{p}_{Tw}$ that boosts the summation for hard-scatter interactions. As a result sumPTw provides an approximately 10% relative improvement in hard-scatter primary vertex selection in both current LHC and HL-LHC scenarios.

The efficacy of this method is visually highlighted by means of an example from a particular Run 4 MC simulated event in Figure 6.6. In the figure, the top pane is centered at the pile-up vertex incorrectly chosen by the sumPT2 criterion, whereas the bottom pane is centered at the reconstructed hard-scatter vertex that is correctly chosen by the sumPTw criterion. Note that these are $r - z$ plane plots, where the cylindrical detector is unwrapped into a flat plane and so the azimuthal angular information is lost. Furthermore, the lengths of the lines are in arbitrary units and are illustrative of the $\mathbf{p}_T$ — they have no absolute scale of reference. The event display shows that the presence of a large number of pile-up tracks leads sumPT2 to pick the incorrect vertex, whereas sumPTw is able to pick the hard-scatter vertex due to the angular track-jet projection scheme.



(a)

Figure 6.6: $r - z$ event display plots from a single VBF Higgs invisible MC simulated Run 4 event. The top pane is centered (red ellipse) on the pile-up (PU) vertex incorrectly chosen by sumPT_2 , the bottom pane is centered on the reconstructed hard-scatter (HS) vertex chosen by sumPT_w . All jets with $p_T > 30$ GeV are drawn as semi-transparent bars on top of the vertex of interest. In addition to the reconstructed vertices shown on the *Reco* axis, the true HS and PU vertex locations are also shown on the *Truth* axis on the lower half of each pane. Figure sourced from [37].

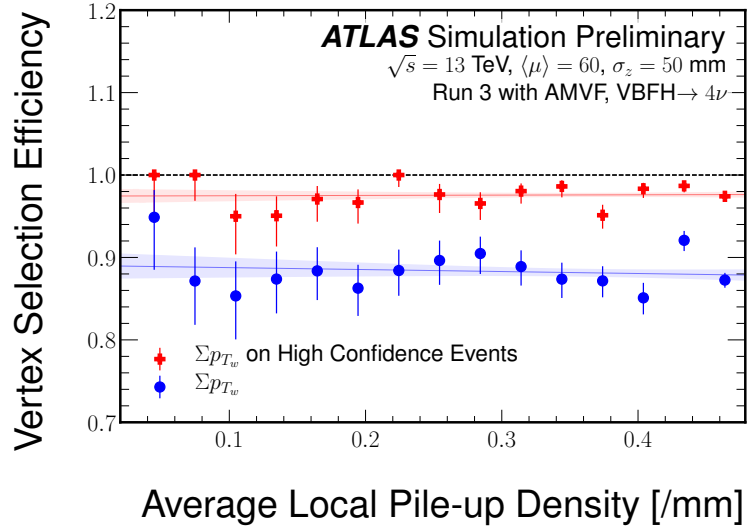
6.5 Experimental Challenges with the ATLAS ID geometry and Future Improvements

As seen in Figure 6.5, sumPTw achieves higher efficiency in HL-LHC conditions than in the Run 3 conditions. The improvement in the efficiency can be attributed to the expanded η acceptance of the ATLAS ITk (Figure 6.3). This is because VBF Higgs invisible events rely on tracks inside the VBF jets for vertex identification, and when any of these jets land outside the ID acceptance ($|\eta| < 2.5$) no tracks in them are reconstructed, leading to lower sumPT2 and sumPTw.

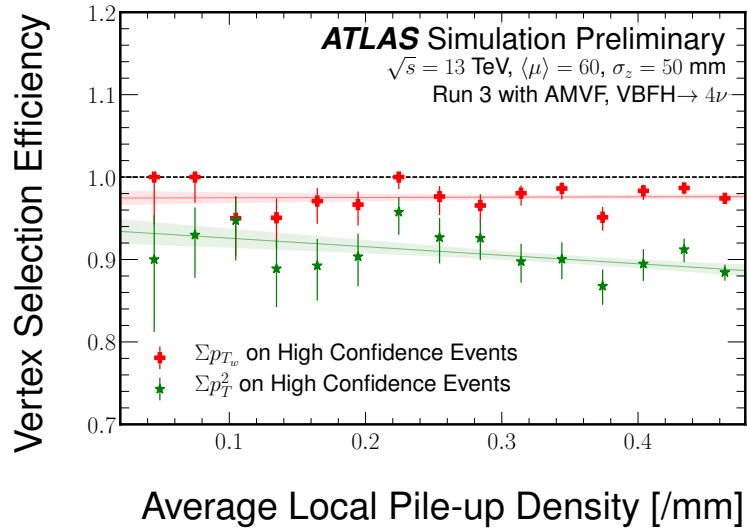
To account for this effect of missing forward tracks in Run 3, an event categorization scheme is proposed to flag events where significant jet $\mathbf{p}_T$ is outside angular acceptance of the tracker. Low confidence events are defined by the presence of at least one jet with $|\eta| > 2.1$ and $\mathbf{p}_T > 30$ GeV, and $R_{\mathbf{p}_T} < 0.1$ with respect to all the vertices in the event. High confidence events, by definition, have hard jets within the acceptance of the tracker. Therefore the hard-scatter vertices in high confidence events have all the tracks (inside jets) within the tracker acceptance. The split of events in the Run 3 VBF Higgs invisible MC simulated events that passed the VBF criteria was found to be 80% – 20% between the high and low confidence event categories respectively.

The efficiency of sumPTw on high confidence events as compared to all VBF Higgs invisible events in Run 3 is shown in Figure 6.7a. Note that the efficiency of sumPTw on high confidence events in Run 3 is comparable to the efficiency of sumPTw on all events at Run 4 with the ATLAS ITk detector in Figure 6.5b. This is because events in both cases are ensured to have hard-scatter vertices with tracks inside the tracker volume. The comparison of the efficiencies of sumPT2 and sumPTw on high confidence events is shown in Figure 6.7b. An improvement in the (inclusive) average selection efficiency from 88% to 97% is obtained with the proposed method.

Figure 6.7 justifies the claim that the inefficiency of sumPTw seen in Figure 6.5a versus Figure 6.5b is driven by events with hard-scatter jets outside the tracker volume. In high confidence events, the hard-scatter vertex can be identified with 97% efficiency. However, for low confidence events an attempt to identify the hard-scatter



(a)



(b)

Figure 6.7: Vertex selection efficiency for VBF Higgs invisible events in Run 3 conditions, versus the average local density of vertices at the primary vertex in the event. (a) The sumPTw selection method on all events (circle) is compared with the sumPTw selection on high confidence events (cross). (b) The default sumPT2 selection (star) is compared with the sumPTw selection (cross) on high confidence events. Error bars are statistical errors. Best fit linear functions are drawn with 1-sigma error bands to visualize the error on the slope with respect to the pile-up density. Figure sourced from [36].

vertex should not be made, even though the event still contains a VBF jet signature.

Another challenge to vertex selection algorithms come from the presence of hard-QCD interactions: Pile-up interactions with high p_T jets that come from a hard inelastic pp interaction other than the interaction of interest. These hard pile-up interactions may not be the interaction that triggers the event being saved, but could still be chosen by the vertex selection algorithm. An example of such an event is shown in Figure 6.8, where the top pane shows the $r-z$ plane centered at the pile-up vertex incorrectly chosen by sumPTw, and the bottom plane is centered at the correct reconstructed hard-scatter vertex. sumPTw in this example has chosen the hard-QCD interaction since it has a larger summed p_{T_w} due to tracks correlated with hard pile-up jets. This final limitation is a consequence of the topology-agnostic nature of the vertex selection algorithms discussed so far — sumPT2 and sumPTw evaluate sums over the tracks associated to every vertex with no regard to the topology of the physics associated to each vertex. This presents a path for improvement of vertex selection algorithms that can extend the developments presented here and combine them with topology discrimination. Furthermore, the sumPTw criterion can also be expanded to include other physics objects such as leptons, to extend the applicability of the method to final states with objects beyond tracks and jets.

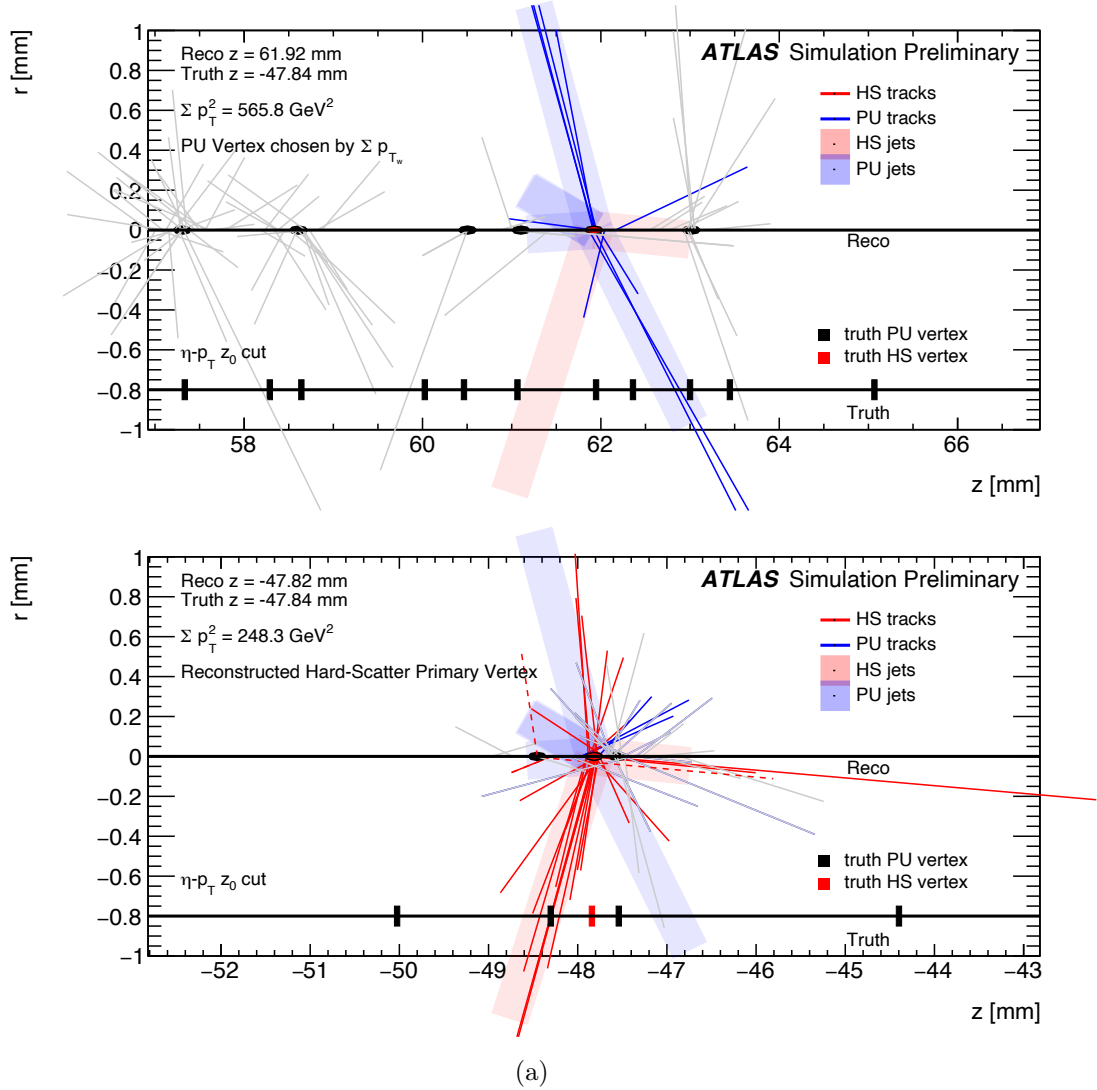


Figure 6.8: $r - z$ event display plots from a single VBF Higgs invisible MC simulated Run 4 event. The top pane is centered (red ellipse) on the pile-up (PU) vertex incorrectly chosen by sumPT_{Tw} , the bottom pane is centered on the correct reconstructed hard-scatter (HS) vertex. All jets with $p_T > 30$ GeV are drawn as semi-transparent bars on top of the vertex of interest. In addition to the reconstructed vertices shown on the *Reco* axis, the true HS and PU vertex locations are also shown on the *Truth* axis on the lower half of each pane. Figure sourced from [37].

6.6 Conclusions

In this chapter, a novel hard-scatter primary vertex selection algorithm developed by the Author for ATLAS is presented. This new algorithm combines information from both the ATLAS tracking and calorimeter systems, and effectively mitigates the impact of pile-up vertex merging on vertex selection efficiency. The method is, therefore, able to achieve higher selection efficiencies for final states with low visible track p_T , by projecting tracks onto hard jet axes.

The algorithm has been evaluated for the challenging invisible Higgs boson decay final state in the vector-boson fusion production channel. The selection efficiency of the proposed method is shown to be insensitive to the average local pile-up density (at the selected vertex) and improves the efficiency for events containing an invisible Higgs boson decay from 80% to 88% (inclusive) in Run 3 conditions. An event categorization scheme has also been introduced that classifies events into high or low confidence events based on the presence of high p_T jets outside the ATLAS Inner Detector acceptance. The new method achieves a selection efficiency of 97% (inclusive) on high confidence events.

The method has also been evaluated for the HL-LHC Run 4 era, assuming the ATLAS ITk detector volume and reconstruction performance. The algorithm shows an improvement in the efficiency of identifying invisible Higgs boson decay vertices from 88% to 97% (inclusive).

Several possible routes for improving vertex selection algorithms have also been discussed. These include making the vertex selection algorithm topology aware so that it can distinguish between hard-QCD interactions and other physics interactions of interest. Furthermore, extending the logic of the proposed technique to include physics objects like leptons can extend the improvement the method offers to other final states with low visible track p_T .

Chapter 7

Calibration of Jets Assisted with Neural Networks in the ATLAS Experiment

Jets are vital objects for the physics analysis of the data collected at pp colliders. As discussed in Section 4.6.5, jets in ATLAS are reconstructed using calorimeter objects and enhanced using tracking information from the ATLAS ID. The calibration of these reconstructed jets is required due to various detector effects that result in the energy-momentum scales of the jets being shifted from the true values that correspond to the properties of the particle-level jet. The calibration procedure for jets in ATLAS is composed of several steps that calibrate the jets using MC simulations and data, utilizing the numerical inversion technique (Section 4.6.5.2.1) to derive the corrective factors while remaining independent of the (prior) underlying spectrum. Additionally, ATLAS uses the global sequential calibration technique to mitigate the dependence of the energy-momentum responses with a set of jet features for small- R jets, and has derived a set of jet substructure related observables that improve the reconstruction of large- R jets [378]. The correction with respect to the set of *auxiliary* features in the GSC proceeds via a sequential application of numerical inversion for each feature. ATLAS has recently introduced the concept of *generalized numerical inversion*, which is a multivariate technique to perform numerical inversion using

neural networks [379]. This Chapter presents work extending generalized numerical inversion, and demonstrates simultaneous calibration of multiple observables. Specifically, the technique is demonstrated for large- R jet absolute MC-based calibration of the jet energy and mass simultaneously. Furthermore, the neural network calibration technique is made capable of learning a variety of measures of central tendency (mean, median, mode, quantiles) of the detector response, enabling the alignment of the method with ATLAS jet calibration objectives. ATLAS utilizes the mode of the detector response to calibrate jets (Section 4.6.5.2) — The ability of neural networks to learn the mode of the response is critical for the adoption of neural network based methods for jet calibrations. The mode learning procedure as described in this chapter is the first such technique known to the Author that enables gradient decent based regression to learn the mode of the conditional distribution. The methods presented in this Chapter are published in [38, 39] and were developed by the Author along with collaborators (S. Cheong, A. Cukierman, B. Nachman, A. Schwartzman) for the ATLAS experiment. The Author developed the mode learning objective function, as well as the architecture that enabled the simultaneous calibration of large- R jet energy and mass.

7.1 Introduction

Jets, as described in Section 4.6.5.1 are collimated streams of particles detected in hadron colliders as partons hadronize due to color confinement during transit towards the detector equipment. Due to the energy losses, detector imperfections and non-linearities, the reconstructed jet has properties that differ from the properties corresponding to the particle-level jet; Particle-level jet refers to the jet reconstructed by directly combining the four-momenta of the individual (detector stable) particles rather than the detector readouts. Jets are essential in a variety of precision Standard Model measurements, as well as in searches for physics beyond the Standard Model like the one presented in Chapter 5. It is therefore essential that the jet four-momenta be calibrated to match the corresponding particle-level jet four-momenta [30, 31, 259, 260]. The procedure for jet calibrations, as described in Section 4.6.5.2, fundamentally relies on *numerical inversion* [260–262].

Numerical inversion comprises of two steps, the first of which is learning the response of the detector. For observable X (for example jet energy, mass), some measure of central tendency (typically the mode) of the reconstructed observable X^{reco} is parameterized as a function of the true value of the observable X^{true} . This parameterization can be further extended as a function of other parameters θ that affect the scale of the observable X but are not calibrated themselves. Given that the detector is responsible for the mapping $(X^{\text{true}}, \theta) \mapsto X^{\text{reco}}$, this parameterization learns the response of the detector — The detector response is captured by the probability distribution $(X^{\text{reco}}|X^{\text{true}}, \theta)$ and the parameterization learns some measure of central tendency of this distribution. The parameterization can be obtained with MC simulations of the experiment, or with data where X^{true} is measured independently and to a high precision. The parameterized function is then inverted, and this forms the calibration function — the inverted function maps $X^{\text{reco}} \mapsto \hat{X}^{\text{reco}}$.

Conventional numerical inversion relies on deriving the parameterization of the detector response by building look-up tables in bins of $(X^{\text{true}}, \theta)$, or by assuming a functional form of the response $(X^{\text{reco}}|X^{\text{true}}, \theta)$ and performing unbinned fits. The former becomes unfeasible as the calibration task gets multi-dimensional, and the later

results in biased parameterizations due to imperfect assumptions. The conventional approach is used to calibrate one observable at a time, taking at most two auxiliary variables into account — This is evident in the calibration procedures described in Section 4.6.5.2.2 and Section 4.6.5.2.3 where several sequential steps are taken to calibrate the jet four-momenta. A novel unbinned and assumption-free method of implementing numerical inversion, called generalized numerical inversion (GenNI), was introduced by ATLAS as an improvement to the standard numerical inversion [379]. This method is capable of calibrating several observables simultaneously, while correcting for the effects of several auxiliary parameters at the same time.

GenNI executes the two steps in numerical inversion using neural networks. The first neural network, L , learns the continuous multi-dimensional parameterization of the detector response, $L : (X^{\text{true}}, \theta) \mapsto X^{\text{reco}}$; Here θ , X^{true} , and X^{reco} can be multi-dimensional. The function L learns some central tendency of the $(X^{\text{reco}}|X^{\text{true}}, \theta)$ distribution, set by the loss function used to train the neural network; This is described in detail in Section 7.2. Given that L is a deterministic function, the distribution $(X^{\text{true}}|L(X^{\text{true}}, \theta), \theta)$ is a Dirac δ centered at X^{true} , provided the learned response is a monotonic function of X^{true} . Consequently, the second neural network, $C : (L(X^{\text{true}}, \theta), \theta) \mapsto X^{\text{true}}$, learns the inversion of the first neural network and has trivial training due to the lack of spread or local minima in the training distributions.¹ C is the final learned calibration function, and the calibrated value of the observed X^{reco} is $\hat{X}^{\text{reco}} = C(X^{\text{reco}}, \theta)$. The general schematic of the two steps in GenNI, which also apply to standard numerical inversion, are shown in Figure 7.1.

The work presented here extends the GenNI formalism to learn various measures of central tendency, as detailed in Section 7.2. Of the measures of central tendency that neural networks are shown to be capable of learning in Section 7.2, the mode stands out as a vital breakthrough since it enables the adoption of neural network based calibration procedures for ATLAS. Furthermore, the work shown here also extends the GenNI formalism to simultaneously calibrate multiple features of large- R ($R = 1.0$) jets: the energy E and the mass m , as will be detailed in Section 7.3

¹Care must be taken when θ is strongly correlated with X^{true} , since C can learn X^{true} using θ and without actually inverting the learned response function L . This is, however, outside the scope of the work presented here.

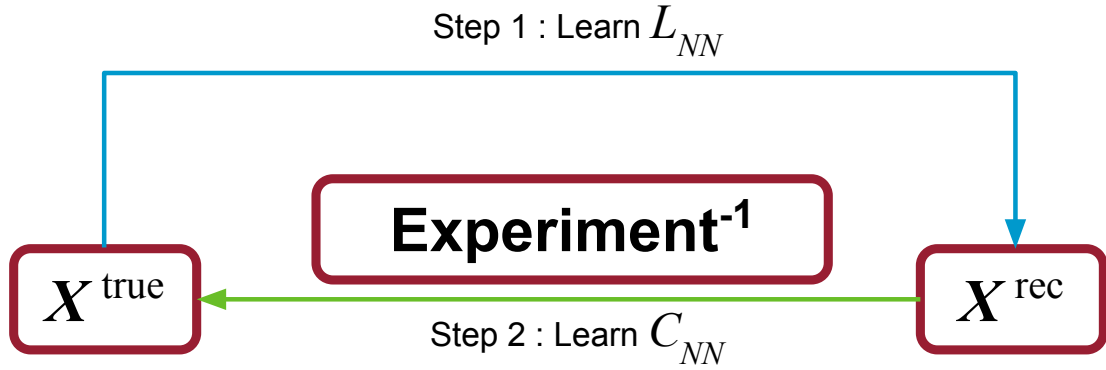


Figure 7.1: Schematic of the two steps involved in GenNI, also applicable to standard numerical inversion. The first step involves learning the detector response distribution with the first neural network L . This parameterized response is then inverted in the second step using the second neural network C . The inverse of the response learned by the second neural network undoes the action of the experiment (which is typified by the response distribution), hence the second step is labelled as the inversion of the experiment. This second neural network C is then the required calibration function that corrects the scale of the reconstructed observable X^{reco} to the true scale.

7.2 Parameterizing Detector Response Using Neural Networks

For this section, simplified notation will be used to convey the mathematical concepts being presented. The measurement process takes random variable X and produces the output random variable Y — This is analogous to the detector mapping some true observable to its reconstructed value. Due to noise from marginalizing unmeasured quantities and quantum mechanical sources, $Y|X$ is a random variable characterized by the probability distribution $p(y|x)$, where x and y are instances of X and Y . Note that this is the detector response function introduced earlier $p(y|x)$. In principle, X and Y can be multi-dimensional and not share the same physical dimensions — Here Y will be assumed to be one dimensional, as is the most common case, to simplify the notation, acknowledging that the logic presented here extends identically to the multi-dimensional case. The function $f(x)$ represents the parameterization of the detector response to the input x , and represents some central tendency of the response distribution $p(y|x)$.

Neural networks that map $f : X \mapsto Y$ are trained by minimizing some objective or loss function ℓ :

$$\begin{aligned} f &= \operatorname{argmin}_{f'} \mathbb{E}[\ell(Y, f'(X))] \\ &= \operatorname{argmin}_{f'} \iint \ell(y, f'(x)) p(x, y) dx dy, \end{aligned} \tag{7.1}$$

where $\mathbb{E}[\cdot]$ is the expected value function. Loss functions used as objectives for the neural network training are required to be bounded below and usually have a unique global minimum, like the mean squared error (MSE) loss. Sufficiently flexible neural networks will traverse the optimisation phase space and reach the condition Eq. 7.1, provided the dataset used to train the network is a sufficiently large sample from $p(x, y)$.

To extract some properties of the loss function ℓ , consider a neighborhood $V(x')$ around some $x' \in \{x\}$, in the support of the training dataset. Under the assumption that $p(x, y)$ and $f' \equiv f(x')$ vary slowly over $V(x')$ (roughly constant) for a small

enough neighborhood $V(\mathbf{x}')$, the converged neural network will satisfy:

$$\begin{aligned}
0 &= \frac{\partial}{\partial f'} \mathbb{E}[\ell(\mathbf{Y}, f(\mathbf{X}))] \\
&= \frac{\partial}{\partial f'} \left[\int_{V(\mathbf{x}')} \mathbf{p}(\mathbf{x}) \int_{-\infty}^{\infty} \ell(\mathbf{y}, f(\mathbf{x})) \mathbf{p}(\mathbf{y}|\mathbf{x}) \, d\mathbf{y} \, d\mathbf{x} \right] \\
&\approx \left[\int_{V(\mathbf{x}')} \mathbf{p}(\mathbf{x}) \, d\mathbf{x} \right] \times \left[\int_{-\infty}^{\infty} \frac{\partial \ell(\mathbf{y}, f')}{\partial f'} \mathbf{p}(\mathbf{y}|\mathbf{x}') \, d\mathbf{y} \right] \\
&\propto \mathbb{E} \left[\frac{\partial \ell(\mathbf{y}, f')}{\partial f'} \middle| \mathbf{x}' \right].
\end{aligned} \tag{7.2}$$

For a converged neural network, Eq. 7.2 is true for all $\mathbf{x}' \in \{\mathbf{x}\}$ across the support $\mathbf{p}(\mathbf{x})$, since ℓ is bounded from below. Therefore the neural network generally satisfies,

$$\mathbb{E} \left[\frac{\partial \ell(\mathbf{y}, f(\mathbf{x}))}{\partial f(\mathbf{x})} \middle| \mathbf{x} \right] = 0 \tag{7.3}$$

which enables the engineering of the loss function ℓ such that $f(\mathbf{x})$ can learn specific summary statistics for the parameterization of the response. The following sections cover some of the typical summary statistics that can be learned using standard loss functions, culminating with the mode which requires the use of the novel *Leaky Gaussian Kernel* objective function. As has been discussed previously, learning the mode of the response is crucial for the task of jet calibrations in ATLAS. The CMS experiment has recently investigated the use of the Huber loss [380] to reduce the sensitivity to outliers [381]. The Huber loss interpolates between the mean and median of the response distribution, but is still biased by the asymmetry of the distribution. This work, therefore, focuses on learning the mode directly using the newly developed loss function.

7.2.1 Mean

The mean of the response $\mathbf{p}(\mathbf{y}|\mathbf{x})$ can be learned with the mean squared error (MSE) loss function or simply the squared error (SE) loss, $\ell_{\text{SE}}(\mathbf{y}, f(\mathbf{x})) \equiv (\mathbf{y} - f(\mathbf{x}))^2$. As per

Eq. 7.3,

$$\mathbb{E} \left[\frac{\partial \ell_{\text{SE}}(\mathbf{y}, \mathbf{f}(\mathbf{x}))}{\partial \mathbf{f}(\mathbf{x})} \middle| \mathbf{x} \right] = 2 (\mathbf{f}(\mathbf{x}) - \mathbb{E}[\mathbf{y}|\mathbf{x}]) = 0. \quad (7.4)$$

Hence the learned parameterization $\mathbf{f}(\mathbf{x}) = \mathbb{E}[\mathbf{y}|\mathbf{x}]$ is exactly the mean of the response distribution $\mathbf{p}(\mathbf{y}|\mathbf{x})$. This is one of the most common objective functions used in neural network driven regression, and produces stable convergence to the mean of the response using gradient descent.

7.2.2 Median

Another commonly task in regression is to learn the median of the response. This is especially useful in scenarios where robustness is needed with respect to the outliers in the response distribution. This can be achieved using the mean absolute error (MAE) loss function or simply the absolute error (AE) loss, $\ell_{\text{AE}}(\mathbf{y}, \mathbf{f}(\mathbf{x})) \equiv |\mathbf{y} - \mathbf{f}(\mathbf{x})|$. As per Eq. 7.3,

$$\begin{aligned} 0 &= \mathbb{E} \left[\frac{\partial \ell_{\text{AE}}(\mathbf{y}, \mathbf{f}(\mathbf{x}))}{\partial \mathbf{f}(\mathbf{x})} \middle| \mathbf{x} \right] \\ &= - \int_{-\infty}^{\mathbf{f}(\mathbf{x})} \mathbf{p}(\mathbf{y}|\mathbf{x}) \, d\mathbf{y} + \int_{\mathbf{f}(\mathbf{x})}^{\infty} \mathbf{p}(\mathbf{y}|\mathbf{x}) \, d\mathbf{y} \\ &= \mathbb{E} [\text{sgn}(\mathbf{y} - \mathbf{f}(\mathbf{x})) | \mathbf{x}]. \end{aligned} \quad (7.5)$$

where $\text{sgn}(z)$ is the signum function defined by:

$$\text{sgn}(z) \equiv \begin{cases} 1 & \text{if } z > 0 \\ 0 & \text{if } z = 0 \\ -1 & \text{if } z < 0. \end{cases}$$

Hence the learned parameterization $\mathbf{f}(\mathbf{x})$ is exactly the median of the response distribution $\mathbf{p}(\mathbf{y}|\mathbf{x})$, such that it integrates equally on their side of $\mathbf{f}(\mathbf{x})$. Like the MSE, the MAE is a commonly used objective function that produces stable convergence to the median on the response using gradient descent.

7.2.3 Quantile

Like the median, which is the 50% quantile, it is interesting to learn any arbitrary quantile of the response distribution, often to quantify the width of the distribution using an interquantile range. These measures enjoy the same outlier insensitivity that applies to the median, and are hence used widely in statistics. Additionally, quantiles can help determine the limits on the output of the detector. Arbitrary quantiles of $p(y|x)$ can be learned using the the generalized absolute error (GAE) [382, 383] function:

$$\ell_{\text{GAE}, q}(y, f(x)) = \begin{cases} q|y - f(x)| & y \geq f(x) \\ (1 - q)|y - f(x)| & y < f(x) \end{cases} \quad (7.6)$$

where $q \in [0, 1]$ is the quantile of the response learned using $\ell_{\text{GAE}, q}$. This loss function can be made smooth [384, 385], learn multiple quantiles at once [386], among several other possible extensions. As per Eq. 7.3,

$$\begin{aligned} 0 &= \mathbb{E} \left[\frac{\partial \ell_{\text{GAE}}(y, f(x))}{\partial f(x)} \middle| x \right] \\ &= -(1 - q) \int_{-\infty}^{f(x)} p(y|x) dy + (q) \int_{f(x)}^{\infty} p(y|x) dy \\ \frac{1 - q}{q} &= \frac{\int_{f(x)}^{\infty} p(y|x) dy}{\int_{-\infty}^{f(x)} p(y|x) dy}. \end{aligned} \quad (7.7)$$

Hence the learned parameterization $f(x)$ is exactly the $100q\%$ quantile of the response distribution $p(y|x)$.

7.2.4 Mode

The mode is another measure of central tendency that is insensitive to outliers, like the median and quantiles. In addition to this, the mode is also insensitive to truncations in the distribution so long as the truncations do not affect the maxima of the distribution. These properties are particularly useful for jet calibration where the jet energy measurements have non-Gaussian tails from detector effects such as inactive material and punch-through beyond the calorimeter. Additionally, jet reconstruction algorithms typically have a minimum transverse energy threshold to suppress the stochastic and pile-up jets, which are usually at lower energy scales than the interesting hard-scatter jets. Therefore jet energy-momenta distributions suffer from the effects of both outliers and truncation, making the use of the mode crucial for jet calibrations. However the downside of using the mode is that it is not guaranteed to be unique, unlike the mean, median, or quantiles of the distribution. Analytically, it is clear that the mode of the response can be learned using the 0-1 loss, or $\ell(y, f(x)) = -\delta(y - f(x))$ for the Dirac delta function δ . This can be seen from Eq. 7.3,

$$-\int_{-\infty}^{\infty} \frac{\partial \delta(y - f(x))}{\partial f(x)} p(y|x) dy = - \left. \frac{\partial p(y|x)}{\partial y} \right|_{y=f(x)} = 0.$$

It should be noted here that the Dirac delta function as defined above will select all turning points, not just the maxima of the distribution — the tie is broken, however, along the optimization path. This loss function can however not be used with gradient based optimizers, hence a *surrogate* loss function is required. Surrogates functions for linear regression have been studied [387, 388] — However, as far as the Author is aware, this concept has not been deployed for use with neural networks.

The negative Gaussian kernel (GK) can serve as a simple surrogate for the Dirac delta-function,

$$\ell_{\text{GK}}(y, f(x); h) \equiv -\frac{1}{\sqrt{2\pi}h} \exp\left(-\frac{(y - f(x))^2}{2h^2}\right)$$

where the hyper-parameter h is called the *bandwidth* of the kernel, and controls how finely the kernel can estimate the mode and accommodate small sample sizes.

The use of this negative kernel function to find the mode of the distribution is

in concert with the idea of kernel density estimation. In kernel density estimation, the sum of kernels centered at some finite set of points produces an estimate for the probability distribution that the points were sampled from. Likewise, the sum of negative Gaussian kernels in a small neighborhood of some point $\mathbf{x}$ will estimate $-\mathbf{p}(\mathbf{y}|\mathbf{x})$ up to some normalization and offset. It follows that $\mathbf{f}(\mathbf{x}) = \operatorname{argmax}_{\mathbf{y}} \mathbf{p}(\mathbf{y}|\mathbf{x})$ will minimize the mean training loss across the support of the training dataset.

Figure 7.2 demonstrates the efficacy of the ℓ_{GK} loss for the Gamma [389] and Crystal Ball [390] distributions, for a fixed $\mathbf{x} = \mathbf{x}_0$. The figure shows the actual loss landscapes evaluated using $\mathbb{E}[\ell(\mathbf{Y}, \mathbf{f}(\mathbf{x})) | \mathbf{X} = \mathbf{x}_0]$, where $\mathbf{f}(\mathbf{x}_0)$ is swept through the range of values $\mathbf{p}(\mathbf{y}|\mathbf{x}_0)$ — MSE, MAE, and GK loss functions are used to illustrate the differences between the summary statistics that minimize each. Note that the Gamma distribution is as follows, for parameters $\mathbf{k}$, θ and the Gamma function Γ :

$$\mathbf{p}_{\text{Gamma}}(\mathbf{x}) = \mathbf{x}^{\mathbf{k}-1} \frac{e^{-\mathbf{x}/\theta}}{\theta^{\mathbf{k}} \Gamma(\mathbf{k})}.$$

Note also the Crystal Ball (CB) distribution, for parameters $\beta > 0$ and $\mathbf{m} > 1$, is:

$$\mathbf{p}_{\text{CB}}(\mathbf{x}) \equiv \begin{cases} \mathbf{N} \exp(-\mathbf{x}^2/2) & \text{if } \mathbf{x} > -\beta \\ \mathbf{N} \mathbf{A} (\mathbf{B} - \mathbf{x})^{-\mathbf{m}} & \text{if } \mathbf{x} \leq -\beta, \end{cases}$$

where $\mathbf{A} = (\mathbf{m}/|\beta|)^{\mathbf{m}} \exp(-\beta^2/2)$, $\mathbf{B} = (\mathbf{m}/|\beta|) - |\beta|$, $\mathbf{N}$ is the normalization factor.

It is evident from Figure 7.2 that the MSE and MAE are minimized at the mean and median of the $\mathbf{p}(\mathbf{y}|\mathbf{x})$ distribution respectively, as discussed in Section 7.2.1 and Section 7.2.2. Furthermore, it can be seen that the GK function produces a loss landscape that estimates the inverted distribution $\mathbf{p}(\mathbf{y}|\mathbf{x})$, albeit a smooth version that is unable to recover the sharp cut off seen in the Gamma Distribution. For both Gamma and Crystal Ball distributions, the GK loss results in a minimum that is unbiased unbiased with respect to the mode of the underlying $\mathbf{p}(\mathbf{y}|\mathbf{x})$ distribution — The GK loss will be minimized at the mode of the $\mathbf{p}(\mathbf{y}|\mathbf{x})$ distribution provided that $\mathbf{p}(\mathbf{y}|\mathbf{x})$ is smooth around the mode. The examples shown in Figure 7.2 also motivate the choice of the mode as the summary statistic to characterize the detector response, given such

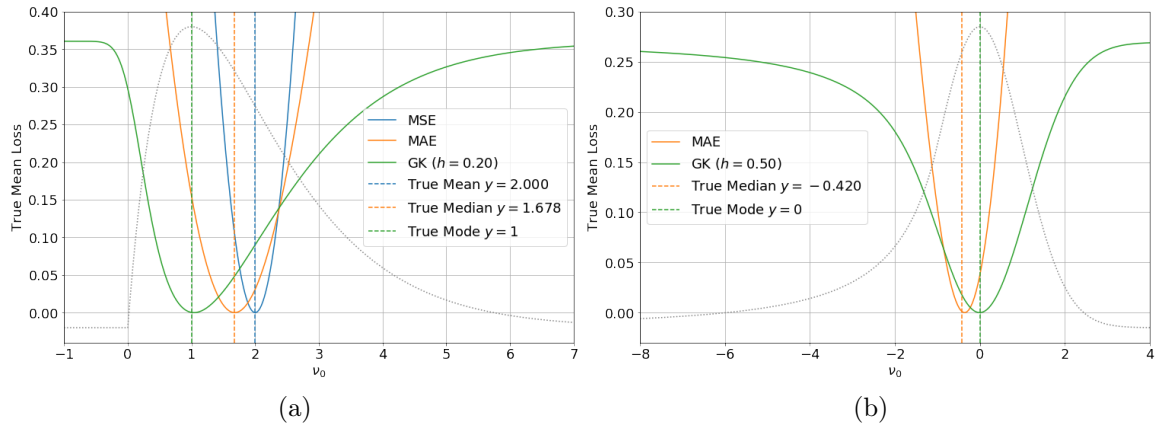


Figure 7.2: Loss landscapes for the MSE, MAE, and GK functions evaluated as the function value $\nu_0 = f(x_0)$ is swept through the range of the distribution $p(y|x_0)$ for each of the functions, for the (a) Gamma Distribution with $k = 2$ and $\theta = 1$, (b) Crystal Ball Distribution with $\beta = 1$ and $m = 2$. The underlying distribution $p(y|x_0)$ is shown with gray dots, and the loss landscapes for the MSE, MAE, and GK functions are offset subtracted to help visualize the minima. Additionally the true mean, median, and mode of the underlying distributions are plotted with vertical dashed lines. Figures sourced from [38].

asymmetric detector response distributions that may not even have statistics like the mean well-defined (as is the case for the Crystal Ball distribution).

As with the Dirac delta function, the GK loss is also not an ideal objective function for use with gradient based optimization due to vanishing gradients sufficiently far away from the mode of the response. This makes converging to the optimal solution challenging, and requires modification to the GK loss that can mitigate the vanishing gradients. This can be achieved by introducing *leakiness* to the negative Gaussian kernel, as defined by the leaky Gaussian kernel (LGK) loss:

$$\ell_{\text{LGK}}(\mathbf{y}, f(\mathbf{x}); \mathbf{h}, \alpha) = \ell_{\text{GK}}(\mathbf{y}, f(\mathbf{x}); \mathbf{h}) + \alpha |\mathbf{y} - f(\mathbf{x})|, \quad (7.8)$$

where the hyper-parameter α introduces the leakiness to the kernel, and sets a bound on the gradient when the function is far from the optimal value. It is required that $\alpha \ll 1$ such that the optimal solution to Eq. 7.2 is found to be close to the mode of $p(\mathbf{y}|\mathbf{x})$. Fig. 7.3 illustrates the effect of the hyper-parameter α by visualizing the loss landscape for some values of α and the Gamma distribution with $k = 2$ and $\theta = 1$. For small α and despite orders of magnitude separating the tested values, the minimum of the LGK loss is found to be close to the mode of the $p(\mathbf{y}|\mathbf{x})$ distribution at $\mathbf{y} = 1$. As α grows, however, the LGK kernel tends towards the absolute error $|\mathbf{y}_1 - \mathbf{y}_2|$ loss and hence the minimum of the function tends towards the median of the distribution at $\mathbf{y} \approx 1.678$ (See Figure 7.2 for the median). Therefore, for sufficiently small α , the leakiness introduced in the LGK loss function aids in the training of neural networks and convergence to the mode of the response distribution, with only a small α -suppressed bias.

The GenNI calibration procedure can be executed with any of the above stated objective functions for both the neural networks L and C. As discussed in Section 7.1, the first neural network, L, learns the detector response distribution through a summary statistic as set by the objective function used to train the network. Using the LGK loss as described here, the GenNI method can learn the mode of the jet response, hence achieving parity with the standard numerical inversion procedure used by ATLAS. This is demonstrated, along with the first demonstration of simultaneous

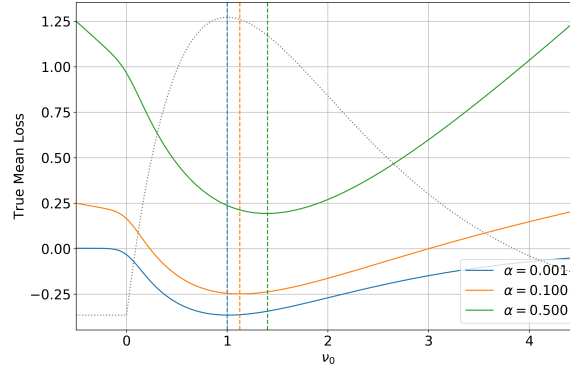


Figure 7.3: Loss landscapes for the LGK loss function evaluated as the function value $\nu_0 = f(x_0)$ is swept through the range of the distribution $p(y|x_0)$ for the Gamma Distribution with $k = 2$ and $\theta = 1$. Three values of α are used and each corresponding loss landscape, along with the corresponding minimum value, are shown. The mode of the Gamma distribution with $k = 2$ and $\theta = 1$ is $y = 1$, while the median is $y \approx 1.678$. Figure sourced from [38].

jet energy and mass calibrations for large- R jets, in the Sections below.

7.3 Simultaneous Jet Energy and Mass Calibrations with Neural Networks

The calibration scheme for large- R jets in ATLAS is summarized in Section 4.6.5.2.3. First, a jet energy scale (JES) correction is derived from MC simulations by fitting the reconstructed jet energy response $E_T^{\text{reco}}/E_T^{\text{true}}$ in bins of truth (particle-level) jet energy, and then applying numerical inversion to the reconstructed jet energies. The second step is a jet mass scale (JMS) correction, derived by fitting the reconstructed jet mass response $m^{\text{reco}}/m^{\text{true}}$ across different bins of truth jet energy and truth jet mass and applying numerical inversion. One challenge with this sequential JES+JMS approach is that the jet energy response depends on the jet mass, and therefore the jet energy may be biased across bins of jet mass. This challenge is mitigated using the GenNI framework to simultaneously calibrate the energy and mass scales, while also simultaneously accounting for auxiliary features such as the jet η^{det} . Additionally, the Leaky Gaussian Kernel (LGK) loss function defined above in Section 7.2.4 is used to learn the mode of jet response distributions.

7.3.1 Event Simulations

Studies documented here utilize several MC simulated samples. The nominal dijet event sample is simulated using PYTHIA 8.230 [286, 287, 317] with the NNPDF2.3 leading order (LO) [391] set of parton distribution functions (PDF) and the A14 set of tuned parameters [294]. Additional samples are simulated using SHERPA 2.2.5 [392] with the CT14NNLO PDF set [393] and the default cluster hadronization model and HERWIG 7.1.3 [289, 290, 394] with the MMHT2014NLO PDF set [395] and the default angularly-ordered parton shower with the default cluster hadronization [396]. Additional details of these samples may be found in Ref. [397].

The generated events are passed through a simulation of the ATLAS detector [332] in GEANT 4 [298], which encapsulates all the particle-detector interactions and digitizes the event. Pile-up interactions are modeled by overlaying minimum-bias pp collisions on the generated samples. These minimum-bias pp collisions are generated

using PYTHIA 8 with the A3 set of tuned parameters [333] and NNPDF2.3 LO PDF set.

With the events fully simulated, the large- R jet reconstruction proceeds as summarized in Section 4.6.5.1. The jets are formed from LCW clusters and the anti- k_t jet algorithm [254] with radius parameter $R = 1.0$. Jets are also trimmed using $R = 0.2$ subjets and a momentum fraction of 5%, as described in Section 4.6.5.2.3. Truth jets are formed from detector-stable simulated particles, excluding muons and neutrinos. Reconstructed jets are geometrically matched to truth jets using the ΔR distance metric; all jets with a reconstructed-truth match are considered. Reconstructed jets with truth matched energy > 200 GeV and mass > 60 GeV, and the detector pseudorapidity $0.0 \leq |\eta^{\text{det}}| \leq 0.5$ are considered for this study.

7.3.2 Methods

Following the two step numerical inversion scheme, as described earlier, the first step in the GenNI procedure is learning the neural network $L : (E^{\text{true}}, m^{\text{true}}) \mapsto (E^{\text{reco}}, m^{\text{reco}})$, which learns the response of the detector. The second step then learns network $C : L(E^{\text{true}}, m^{\text{true}}) \mapsto (E^{\text{true}}, m^{\text{true}})$, which inverts the learned response of the detector. The calibrated values of E^{reco} and m^{reco} are finally given by $C(E^{\text{reco}}, m^{\text{reco}})$, essentially undoing the effect of the detector and bringing the scale of the jet energy and mass simultaneously to their calibrated scales. The detector pseudorapidity, η_{det} , is used as an auxiliary variable in each of the above neural networks.

The loss function used is the LGK loss from Section 7.2.4, written here explicitly in terms of the predicted and target values of the variable,

$$\text{Loss}(x^{\text{target}}, x^{\text{pred}}) = -\exp\left(-\frac{(x^{\text{target}} - x^{\text{pred}})^2}{2\alpha}\right) + \beta|x^{\text{target}} - x^{\text{pred}}|,$$

where x^{target} is the target value given by the training set, x^{pred} is the neural network predicted value, and α and β are tunable hyper-parameters. For the simultaneous calibration of E and m , the total neural network loss is the sum of the individual

LGK losses for each of the observables,

$$\text{Loss}(\mathbf{E}^{\text{target}}, \mathbf{m}^{\text{target}}; \mathbf{E}^{\text{pred}}, \mathbf{m}^{\text{pred}}) = \text{Loss}(\mathbf{E}^{\text{target}}, \mathbf{E}^{\text{pred}}) + \text{Loss}(\mathbf{m}^{\text{target}}, \mathbf{m}^{\text{pred}}). \quad (7.9)$$

To help the neural network reach convergence in fewer iterations, the energies and masses are independently and inclusively standardized using the global mean and standard deviation. The networks are trained with Keras [398] and TensorFlow [349] using the Adam [350] optimization algorithm. The L network uses two hidden layers, with 60 and 40 nodes respectively. All intermediate layers use the Leaky ReLU [399] activation function. The loss used is the total loss as defined in Equation 7.9, with the values of α and β set to 10^{-3} after a coarse hyperparameter scan. The C network uses two hidden layers, with 60 and 2 hidden nodes, respectively. Since the C network is learning a typically single-valued function, the loss function is not as important and thus the MSE is used for faster convergence.

To accommodate differences in the performance of the calibration across multiple features, care must be taken to ensure that the learning of one feature calibration does not adversely impact that of another. A split network architecture is used when training L to address this problem. The first intermediate layer with 60 nodes has access to all the input parameters, and the successive layers are split in two such that the back propagation gradient flow from the output nodes flows through one half of the split intermediate layers. In this way, while gradient descent tries to minimize the overall loss across all the parameters being calibrated, gradients from each parameter output node have access to a unique portion of the training network (half of the intermediate layer with 40 nodes), ensuring that each parameter is calibrated sufficiently well without being affected by the calibration of a different parameter.

Additionally, in order to ensure that the networks learn the response equally well across the phase space, the energy and mass spectra of the training set are flattened by assigning weights of $1/p(\mathbf{E}^{\text{true}}, \mathbf{m}^{\text{true}})$. At the evaluation stage of the networks, the physical spectra of jets are used.

The standard sequential calibration method and GenNI method are evaluated

using identical metrics. Specifically, the response of the two methods is characterized by extracting the mode of the distribution in bins of the truth observables: In a given bin of truth x ($x = E$ or m), the distribution of the response $x^{\text{reco}}/x^{\text{truth}}$ is fit with a Gaussian function over a restricted range around the peak of the distribution defined by the distribution mean and standard deviation. Assuming that the core of the response distribution is Gaussian, the mean of the Gaussian fit gives an unbiased and stable estimator of the mode of the response within the given bin of truth x . The results of the full comparisons are discussed below in Section 7.3.3.

7.3.3 Results

Figure 7.4 shows the calibration performance for both the sequential calibration and simultaneous GenNI approach, using the MSE loss as well as the LGK loss. For this comparison, the standard sequential approach is replicated by repeated sequential application of GenNI to calibrate one observable at a time, which makes the calibration procedure identical to the standard sequential numerical inversion approach. There is an overall bias in the response when using MSE compared with the LGK, which is due to the difference between the mean and mode of the response distribution. Furthermore, the residual mass dependence of the energy response is nearly eliminated by using a simultaneous approach over a sequential approach, as demonstrated by the better closure in the bottom right plot of Figure 7.4.

A detailed comparison of the standard sequential calibration (JES+JMS) and the simultaneous GenNI calibration is presented in Figure 7.5. Raw (uncalibrated) jet energy and mass responses for the nominal PYTHIA 8 sample are shown in the top row. Prior to any jet calibration, both the energy and mass response show a strong energy dependence. Due to the local topo-cluster energy calibration, the overall energy scale is within 10% of unity across the entire energy range. The overall scale is significantly improved after the standard sequential calibration, although there is a significant non-closure (i.e., response different from 1) at low energy and a 1-2% spread over the range of truth jet masses. The GenNI approach nearly eliminates the residual mass dependence of the energy response and also improves the response at

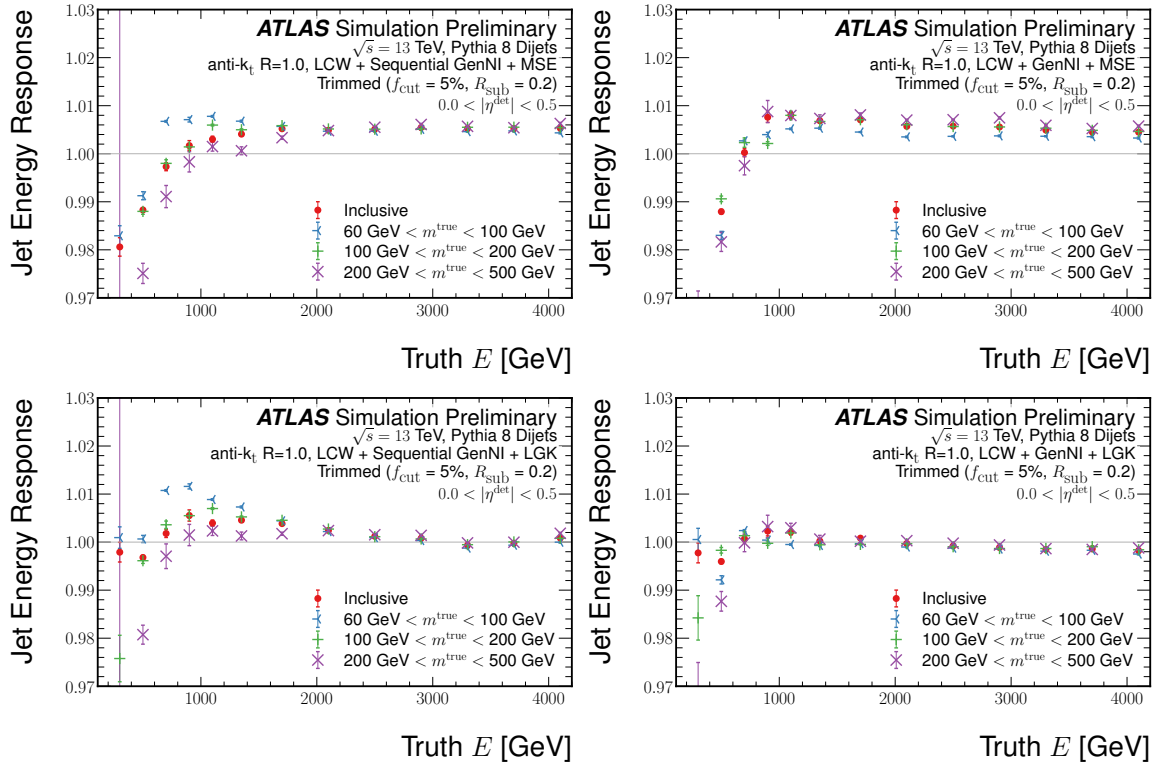


Figure 7.4: A comparison of the mode of the energy response as a function of truth energy for GenNI trained using MSE (top) and LGK (bottom) in bins of the truth jet mass using the nominal PYTHIA 8 MC sample. The left plots use GenNI sequentially while the right plots use it simultaneously for mass and energy. Data points off the scale of the plots may have their error bars within the visualization range, which may appear as marker-less vertical lines on the plot. Figure sourced from [39].

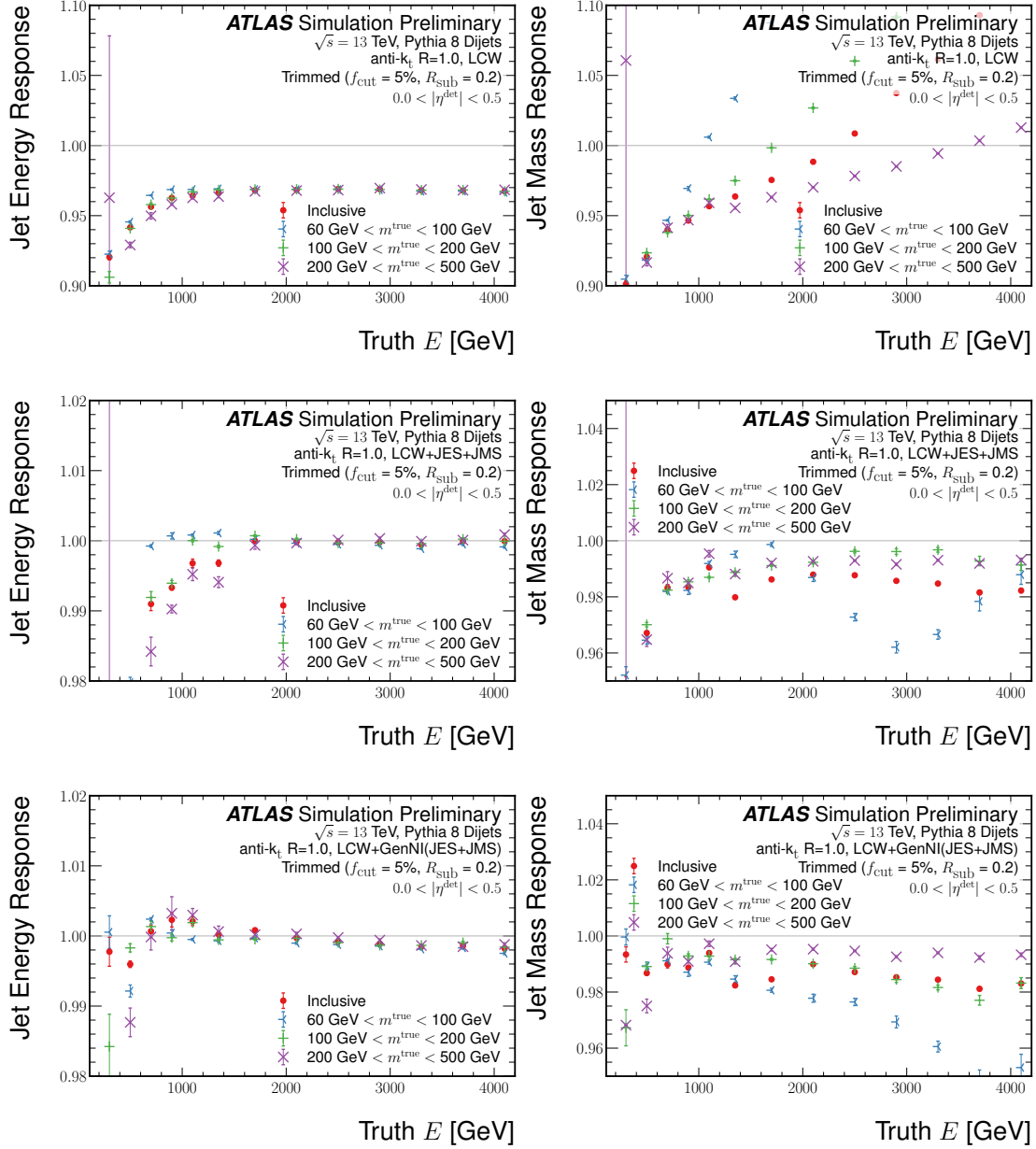


Figure 7.5: Energy (left) and mass (right) responses for the nominal PYTHIA 8 MC sample. From top to bottom, each row represents: no calibration, current ATLAS JES+JMS, and GenNI calibration with LGK loss. Data points off the scale of the plots may have their error bars within the visualization range, which may appear as marker-less vertical lines on the plot. Note that the top row has a different y-axis scale than the lower two rows. Figure sourced from [39].

low energy. The energy and mass resolutions (not reproduced here) are nearly the same between the the traditional simultaneous calibration procedure and the GenNI procedure.

The results presented in Figure 7.4 and Figure 7.5 use the PYTHIA 8 MC simulation for both training and evaluating the calibration. Figure 7.6 and Figure 7.7 instead show the performance of the calibration trained on PYTHIA, but tested with SHERPA and HERWIG, respectively. The figures show the difference in the responses of PYTHIA and SHERPA/HERWIG respectively. The overall calibration closure for the alternative MC simulations is typically no worse for GenNI compared with the standard method. These results suggest that the systematic uncertainties from the neural network based method will be no larger than the standard method and that the gains in closure are robust.

7.4 Conclusions

The work presented in this chapter extends the GenNI framework for neural network based jet calibrations to simultaneously calibrate multiple features. A brief review of how various summary statistics can be learned by machine learning methods depending on the choice of the objective function is shown. Additionally, a novel objective function is proposed that can enable the learning of the mode of the response distribution via gradient decent. The Author believes that this is the first demonstration of mode learning neural networks, and is a key development for the parameterization of the detector response where the mode is the typical summary statistic of choice due to insensitivity to outliers and thresholds. The GenNI calibration is shown to use the mode of the response instead of the mean using the LGK loss function. This development enables the use of the GenNI framework for calibrating jet observables in ATLAS.

The illustration of the multi-variate power of the GenNI technique is done by simultaneously calibrating the energy and mass of large-radius jets in ATLAS simulations. The residual mass-dependence of the energy calibration is reduced and is robust to variations in the MC simulation. These techniques will be useful tools for

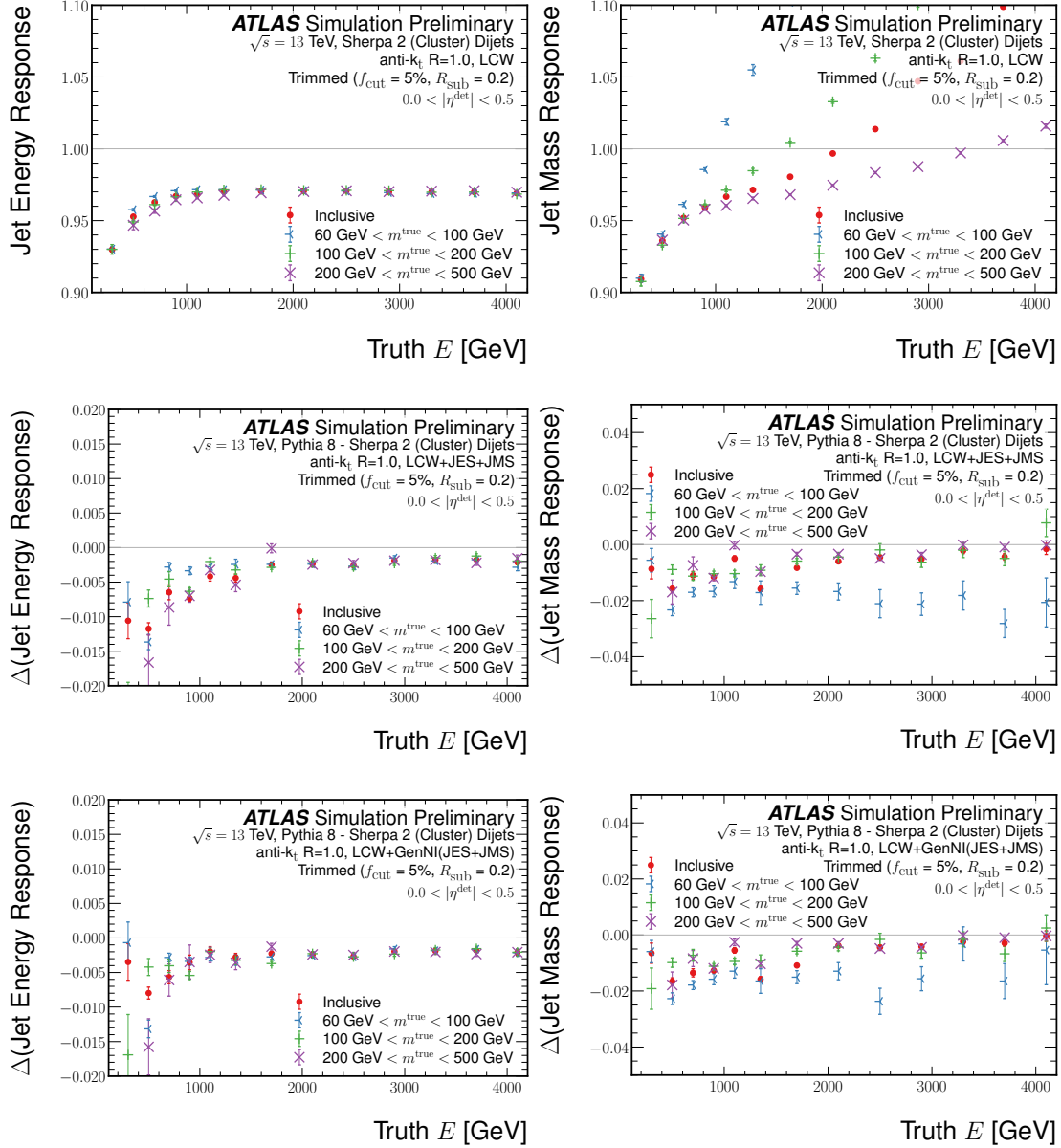


Figure 7.6: Energy (left) and mass (right) responses for the SHERPA MC sample. The top row shows the uncalibrated energy and mass responses of the SHERPA MC sample. The middle and the bottom rows show the response differences between PYTHIA and SHERPA (PYTHIA 8 – SHERPA 2) samples after current ATLAS JES+JMS and GenNI calibration with LGK loss, respectively. Data points off the scale of the plots may have their error bars within the visualization range, which may appear as marker-less vertical lines on the plot. Figure sourced from [39].

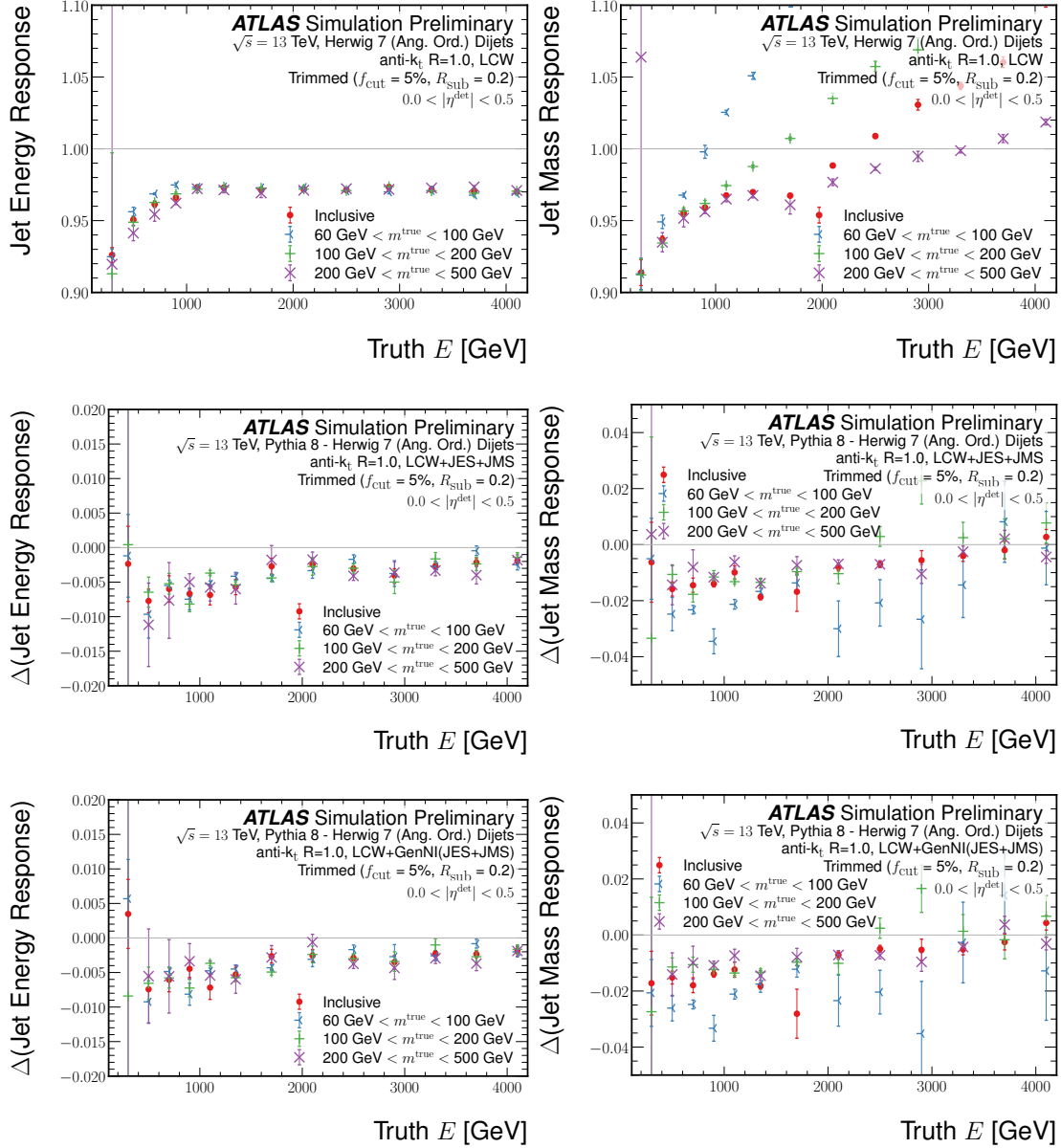


Figure 7.7: Energy (left) and mass (right) responses for the HERWIG MC sample. The top row shows the uncalibrated energy and mass responses of the HERWIG MC sample. The middle and the bottom rows show the response differences between PYTHIA and HERWIG (PYTHIA 8—HERWIG 7) samples after current ATLAS JES+JMS and GenNI calibration with LGK loss, respectively. Data points off the scale of the plots may have their error bars within the visualization range, which may appear as marker-less vertical lines on the plot. Figure sourced from [39].

further calibration developments in jet reconstruction and beyond. In particular, additional features can be easily added to this streamlined GenNI approach where only two neural networks are required instead of many sequential steps.

This development of the GenNI multi-variate jet calibration framework, along with the mode learning loss function, has enabled studies into general multi-variate jet calibration techniques by ATLAS. The Author believes that these promising inquiries have the potential to improve the ATLAS jet calibration performance for jets with interesting substructure, such as the jets found in the analysis presented in Chapter 5 — The ability to make the calibration learn from jet substructure observables opens up previously inaccessible handles to improve the reconstruction fidelity of ATLAS. This is therefore, in the Author’s opinion, an exciting space for future developments in jet reconstruction techniques.

Chapter 8

Light Field Imaging for Cold Atom Clouds in the Matter-wave Atomic Gradiometer Interferometric Sensor

This chapter introduces an innovative super-aperture spatially multiplexed light field imaging device (Ref. [41], adapted here under open access) that has been developed by the Author, along with collaborators (S. Cheong, J.C. Frisch, S. Gasiorowski, J.M. Hogan, M. Kagan, A. Schwartzman, M. Vandegar), for the MAGIS-100 experiment [9]. The imaging device concept has also been granted a provisional patent (No. 63/364,799) by the United States Patent and Trademark Office. The device is able to capture multiple views of an object in a single data-taking shot, collecting more light than a conventional lens operating at the same depth of field and magnification. The system has a projected 3D resolution of $100\text{ }\mu\text{m}$ for objects of size 1 mm^3 , suitable for the reconstruction of fluorescing cold atom cloud densities in atom interferometers. The 3D reconstruction has been demonstrated with a 3D-printed prototype, able to reconstruct $100\text{ }\mu\text{m}$ surface features on a micro-3D-printed 1 mm^3 fluorescing opaque test object. The system has been developed with a focus on the imaging of cold atom cloud densities in atom interferometer experiments like MAGIS-100. The system presents a new technique for imaging the cold atom clouds by enhancing the light collected and providing the information needed to perform 3D reconstruction in

a single shot. It can therefore provide more information to characterize the cold atom cloud densities than conventional imaging. The Author contributed to this work by designing the optics for this imaging modality, along with developing the concept for the experimental prototype. Additionally the Author also assisted collaborators in the simulation and object reconstruction by advising on the appropriate optical constraints needed to be incorporated into the setup.

8.1 Introduction

As introduced in Chapter 2.6, atomic sensors present a unique opportunity to expand the search for physics beyond the Standard Model [400–403]. Specifically, these sensors can be sensitive to ultralight dark matter which has particle masses in the range of 10^{-22} eV – 10^{-12} eV [8]. These masses are below the reach of high energy particle colliders like the LHC, and instead these dark matter candidates manifest as oscillating waves that require different detection mechanisms. Atom interferometers are one such promising experimental techniques that can be sensitive to ultralight dark matter [9]. Additionally, scaled up atom interferometers are quantum sensors that can also be sensitive to *mid-band* gravitational waves in the (0.03 Hz – 3 Hz) frequency range. It is therefore imperative to develop these detection technologies to maturity, and harness the full information gathering ability of this experimental modality.

MAGIS-100 [9] is a path-finder long baseline atom interferometer experiment that is being built to develop and test state-of-the-art atom interferometry techniques and data processing methods for the next generation of quantum sensors.¹ In MAGIS-100, atom interferometry is done using a 100 m vertical baseline ultra-high vacuum tube, encased in a magnetic shield. The experiment has a series of connection nodes along its length, equipped with view-ports for atom manipulation as well as for imaging the fluorescing atom clouds at the end of the interferometer sequence. Cold atom clouds 1 mm^3 in size are launched from three different atom sources located at the bottom,

¹A quantum sensor is a detection apparatus that relies on quantum superposition, entanglement, and related phenomenon.

middle, and top of the experiment. A layout of the MAGIS-100 experiment along with a view of the laser hall is shown in Figure 8.1.

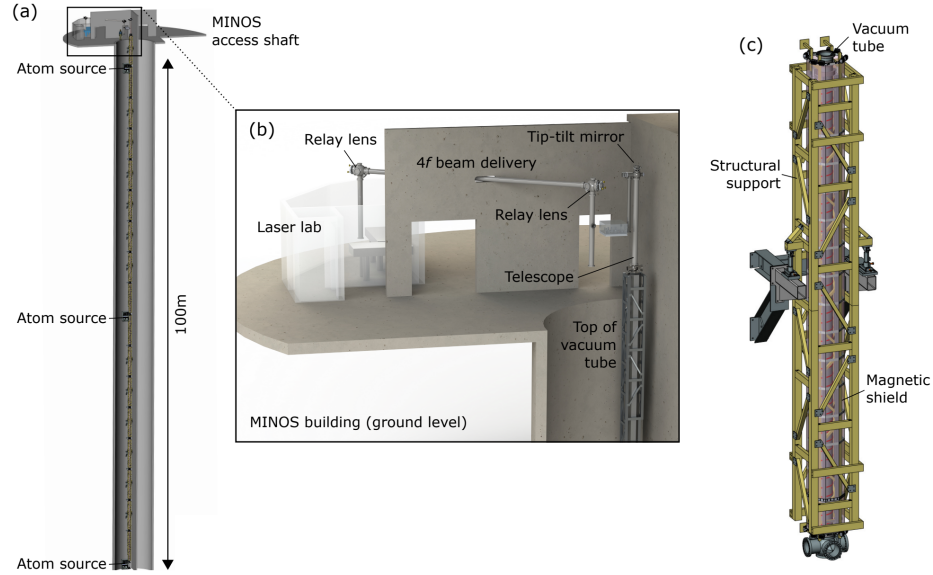


Figure 8.1: Detector layout of the MAGIS-100 experiment under construction at the Fermi National Accelerator Laboratory. (a) Computer generated model of the full experiment with the 100 m vertical vacuum chamber. (b) Computer generated view of the laser lab room. (c) Computer generated model of one of the 17 the modular sections that constitute the full experiment, which contains a section of the pipe, vacuum pumps, magnetic shielding, etc. Figure sourced from [9].

MAGIS-100 operates as a gradiometer by making differential measurements across multiple atom interferometers that are separated (vertically) by some baseline. Atom interferometers rely on ensembles of cold atoms that can be switched between stable internal states using pulses of light. These experiments use the phase accumulated across delocalized atom clouds in a superposition state to produce an interference pattern in the atom cloud densities. In MAGIS-100, the interferometer sequence is driven by laser pulses that travel along the length of the experiment. The laser pulse sequence serves to transfer momentum to the atoms, splitting the atomic wavefunction into two paths that subsequently interfere, realizing a Mach-Zehnder interferometer geometry. Figure 8.2a shows a simplified atom interferometer sequence made using laser pulses that generate the momentum superposition state which then accumulates

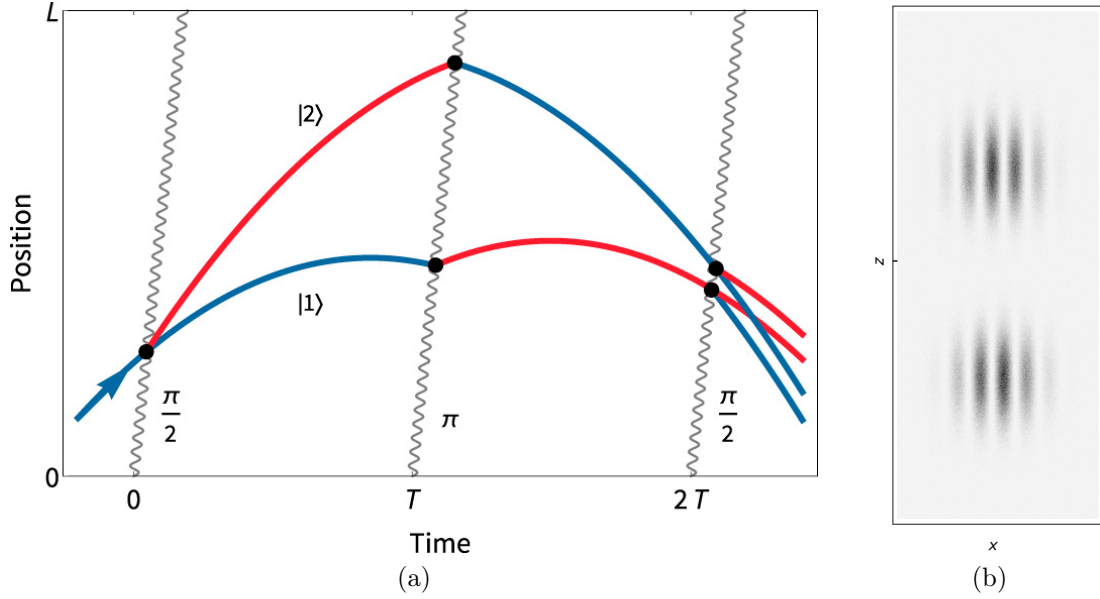


Figure 8.2: (a) Simplified clock atom interferometer space-time diagram. The cold atom cloud manipulated by a series of laser pulses (wavy lines) after the initial launch of the cold atoms, throughout the duration of their free-fall under gravity. Blue (red) lines show the cold atom cloud trajectories in the ground (excited) state. (b) Simulated image of the interference pattern at the end of MAGIS-100 atom interferometer sequence. The image shows the two atom cloud populations corresponding to the two states at the end of the interferometer sequence. The interference fringes depicted here are due to the phase shear readout technique [40]. Figures sourced from [9].

a phase difference across the two clock states.² At the end of the interferometer pulse sequence, the final interference pattern consists of two output ports, one for each of the final interfering atomic states. Each atom from the initial ensemble lands in one of the two ports with a probability determined by the phase accumulated during passage through the interferometer. Figure 8.2b shows a simulated image of the cold atom clouds at the end of the interferometer sequence, where the cold atom clouds in the two states can be seen offset from one another.

Two (or more) atom interferometers sharing a common laser can form a matter-wave gradiometer, as shown in the simplified schematic in Figure 8.3. The difference in the phase of the interference fringes across the two (or more) interferometers forms

²(Quasi-)stable atomic energy states.

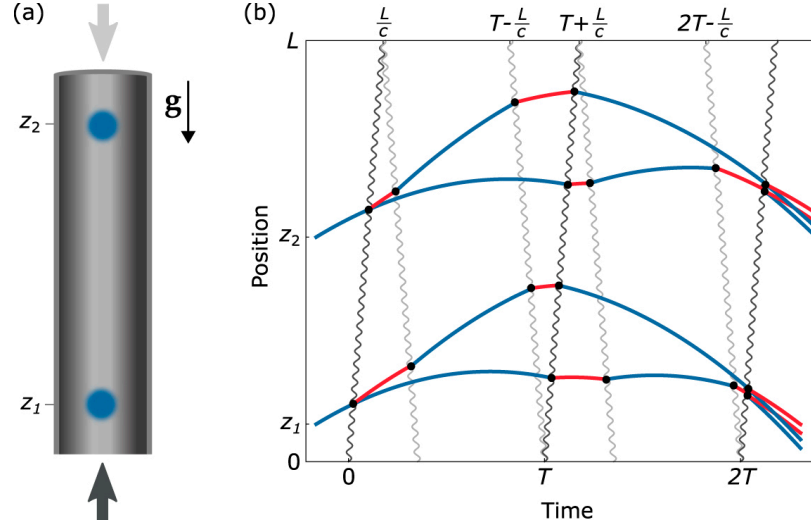


Figure 8.3: Schematic of a simplified clock gradiometer. Two cold atom cloud interferometers are manipulated with shared laser pulses, creating a symmetric pair of interferometers separated by the baseline of the experiment. (a) Initial positions of the two cold atom clouds. (b) Laser pulse sequence space-time diagram showing the creation of the two symmetric interferometers using the two starting cold atom clouds. Dark (light) gray lines show laser pulses traveling upwards (downwards). Blue (red) lines show the cold atom cloud trajectories in the ground (excited) state. Figure sourced from [9].

a key physics observable of a matter-wave gradiometer. This phase difference is proportional to the excitation energy of the clock (interference) transition and the baseline length between the interferometers — this is what makes matter-wave gradiometers suitable for searches for ultralight dark matter, since such dark matter can modulate the atomic energy levels (See Section 2.6). This approach also benefits from common-mode noise rejection since the interferometers all share the same laser, greatly improving the sensitivity of these experiments.

As described above, operation cycle of the MAGIS-100 experiment consists of launching the cold atoms from any subset of the three atom sources, executing the atom interferometer sequence while the atom clouds are experiencing free fall, and imaging the matter-wave interference pattern for each of the launched cold atom clouds. The process of imaging the interference pattern is destructive since the cold atoms are excited into fluorescence using a dedicated fluorescence laser pulse which

induces diffusive heating of the cold atom clouds the longer it interacts with the atoms [9, 404]. The interference pattern is imaged using a sensor through the viewports in the experiment tube. This process is repeated periodically, acquiring a time series of interference pattern measurements which are the basis for the physics analysis conducted at the experiment [9].

Consequently, harnessing the full physics potential of atom interferometer experiments such as MAGIS-100 requires advanced imaging techniques to maximise the light collection from the fluorescing cold atom clouds, while simultaneously retaining structural information about the cold atom cloud densities, in a single shot. Therefore, the novel light field imaging system presented here utilizes plane mirrors to image an array of unique orientations of the object on a single imaging sensor in a single shot. The system multiplexes angular information onto the spatial dimensions of the sensor, and is able to capture directional information about the light emitted from the object; thus capturing its light field. The use of plane mirrors also enables the system to capture more light than conventionally allowed by the aperture of the imaging lens, while still maintaining approximately the same imaging characteristics for each unique orientation imaged. The system thus presents a realization of the maximum-information-capture single shot imaging paradigm for atom interferometers, wherein both light collection can be maximised and structural information preserved.

This image acquisition technique opens up new analysis methods. The angular multiplexing allows the system to image the object from a variety of unique orientations, this enabling 3D tomographic reconstruction of the cold atom cloud densities. The captured data can also be used for 3D function fitting and analysis. Furthermore, this information could enable the quantification of vital systematic uncertainties, such as the laser wavefront aberration phase errors [9]. Additionally, the increased light gathering ability of the system could allow for the cold atoms clouds to be imaged with shorter fluorescence pulses, reducing the distortions to the cloud densities from diffusive heating.

While the focus here is on the MAGIS-100 Experiment, the light field imaging approach presented here is applicable to any scientific or industrial application that deals with imaging small objects and suffers from light-limiting single shot constraints.

8.2 Imaging requirements in atom interferometers

Sr atoms will be used in MAGIS-100, with typical Gaussian density atom clouds of $\mathcal{O}(10^6)$ atoms produced with a volume of about 1 mm^3 . Phase shear readout (PSR) technique [40] is used to imprint a spatial fringe pattern on the two output ports (Figure 8.2b), wherein the fringes that appear in the two ports are offset by 180 degrees due to unitarity. MAGIS-100 is expected to produce PSR interference fringes with a wavelength of $\mathcal{O}(100 \text{ }\mu\text{m})$ over the 1 mm^3 atom clouds.

An important physics observable in atom interferometers is the phase of the interference pattern that is imprinted on the atom cloud density profiles. This sets the primary constraint on the imaging system used to image the atom clouds as they fluoresce, since the system needs to resolve the fringe pattern clearly enough to recover phase variations. Furthermore, each of the Sr atoms at MAGIS-100 will scatter 10^8 photons per second when excited by the imaging light pulse at 461 nm. The imaging light pulse will be turned on for a duration between $10 \text{ }\mu\text{s} - 100 \text{ }\mu\text{s}$, where the upper limit is set by the requirement that diffusive heating from the laser does not average out the interference pattern. The lower limit on the imaging time is driven by the amount of light collected by the imaging system, and can be pushed lower for a system with large light capturing ability.

The above features set the minimal requirements and parameters for any imaging system that will be deployed to image the atom interferometer fringe patterns at MAGIS-100. From the perspective of imaging a single atom cloud, these are listed below:

- Minimum imaging resolution capable of distinguishing $\mathcal{O}(100 \text{ }\mu\text{m})$ features.
- Sufficient field of view and depth of field to accommodate the $\mathcal{O}(1 \text{ mm}^3)$ atom cloud envelope
- Maximize the light collected from the atom cloud to minimize imaging time and improve the signal-to-noise ratio.

8.3 Optical design

8.3.1 Conventional imaging considerations and limitations

A conventional, or entocentric, imaging system constitutes a camera and a finite conjugate (simple or compound) lens focused on the object being imaged. The imaging requirements described in Section 8.2 for imaging in the MAGIS-100 experiment can be used to design a simple conventional imaging system, following the geometric model of image formation (reviewed with relevant details in Appendix B.1). For this study a CMOS sensor with a pixel size of $3\text{ }\mu\text{m}$ is considered with a total sensor size that is arbitrarily large such that it can accommodate the full 1 mm^3 atom cloud image on the sensor. Therefore the problem reduces to a choice of appropriate parameters of the entocentric lens that is used in the conventional imaging of the atom clouds.

A scan of lens magnifications m and f-number N is presented in Figure 8.4 where the impact of various combinations can be seen in terms of the resulting circle of confusion for point source imaging. The depth of field is set to be 1 mm per Section 8.2. Note here that the circle of confusion size is not reported as a raw length or number of pixels, but rather as a fraction of the size a $100\text{ }\mu\text{m}$ feature takes on the imaging plane. This helps contextualize the size of the circle of confusion in terms of the minimum feature size the system is designed to capture. Figure 8.4 also shows the size of the geometric circle of confusion relative to the diffraction limited spot size. Here the diffraction spot size is approximated using the Fraunhofer limit, i.e. the first zero of the Airy Disk ($2.44\lambda N$) [405] where N is the lens f-number and the wavelength $\lambda = 461\text{ nm}$ is used which is the Sr imaging wavelength.

Figure 8.4 helps constrain the choice of lens parameters to a space where the circle of confusion is less than half the size of the $100\text{ }\mu\text{m}$ feature (below the horizontal line). Note that this is a pessimistic analysis that assumes the atom cloud density to be uniform in the full 1 mm^3 volume, whereas the atom clouds are expected to have a Gaussian density profile. Nevertheless it helps isolate a part of the lens parameter space that can be exploited for high fidelity imaging of the atom clouds. The fraction of light collected from an on-axis isotropic point source by lenses in this space reaches a maximum of approximately 0.3%, as shown in Figure 8.5.

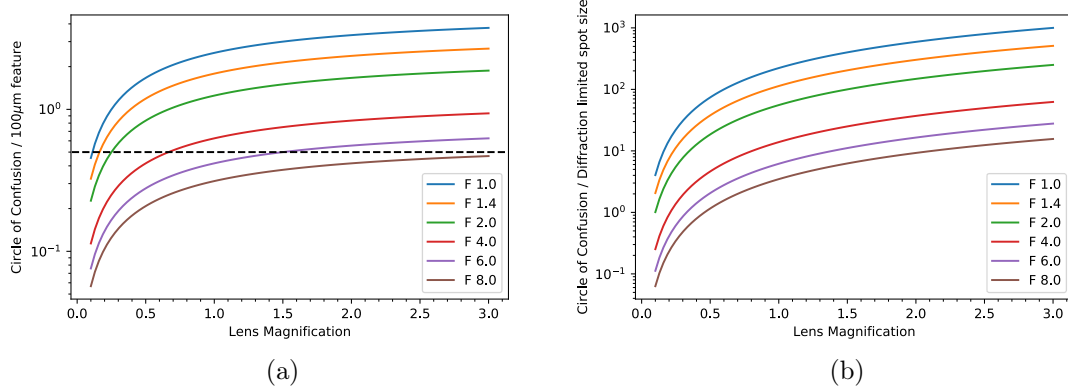


Figure 8.4: (a) Circle of confusion in an ideal conventional single view entocentric imager relative to the size of a $100\text{ }\mu\text{m}$ feature, as a function of lens magnification for a range of lens f-numbers. The $100\text{ }\mu\text{m}$ object space feature dimensions are divided by the magnification to get the image space dimension, which is then used to divide the circle of confusion. The horizontal line is drawn for the points where the circle of confusion is 50% of the $100\text{ }\mu\text{m}$ feature on the imaging plane. (b) Circle of confusion in an ideal conventional single view entocentric imager relative to the diffraction limited spot size, as a function of lens magnification for a range of lens f-numbers. The circle of confusion sizes are computed using Eqn. B.1 with a $f = 62\text{ mm}$, though as seen from Eqn. B.2 the dependence on the focal length is not a dominant effect. The depth of field is set to 1 m to account for the typical size of the cold atom clouds in atom interferometers like MAGIS-100. Figure sourced from [41].

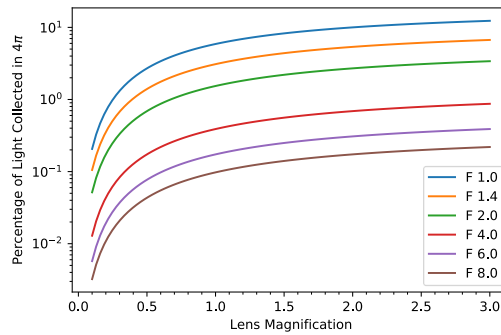


Figure 8.5: Percentage of light collected by an ideal conventional single view entocentric imager that is emanated by an on-axis point object in 4π , as a function of lens magnification for a range of lens F-numbers. The percentages are computed using Eqn. B.3. Figure sourced from [41].

This presents a challenge in utilizing conventional entocentric imaging for atom interferometry; prioritizing light collection jeopardizes the ability of the imaging system to resolve the features of the interference fringes and vice versa. However, the target object here being only 1 mm^3 opens up the possibility to collect more light on the same sensor by utilizing the available pixels and folding in more optical paths into the same imaging system. Furthermore, collecting light from a larger set of angles around the object captures more of the light field, which in turn provides more information to reconstruct the object from the captured image. This is described in detail below.

8.3.2 Light field imaging concept for atom interferometers

8.3.2.1 Geometry of a single folded view

The concept of the light field imager presented here relies on folding optical paths using flat mirrors. Figure 8.6 illustrates the key properties of this folding method for a single view.³ For a given lens and sensor, the focal plane of the imaging apparatus is some well defined surface which can be assumed to be flat for the case of an ideal lens. Rather than placing the object at the focal plane directly, the object is placed closer to the imaging apparatus, and a flat mirror is used to create a virtual copy of the object that sits at the focal plane of the lens. Note that the relative angle between the mirror and optical axis creates a virtual object that is rotated with respect to the imaging apparatus. Light from this virtual object is then focused onto the sensor, along with the light from the object directly that forms a defocused spot on the center of the sensor. Since the focused image on the sensor can be mapped to a specific view angle, this setup preserves both the intensity and angular information of the light emanating from the object — capturing the light field.

The principal ray passing through the center of the lens is highlighted as the dashed line in Figure 8.6. For a given view angle θ , the constraints that the virtual object must sit on the focal plane, and that the principal ray must pass through the center of the lens, uniquely determine the mirror normal angle ϕ , and the position of

³This text uses the word “view” to refer to mirror folded views.

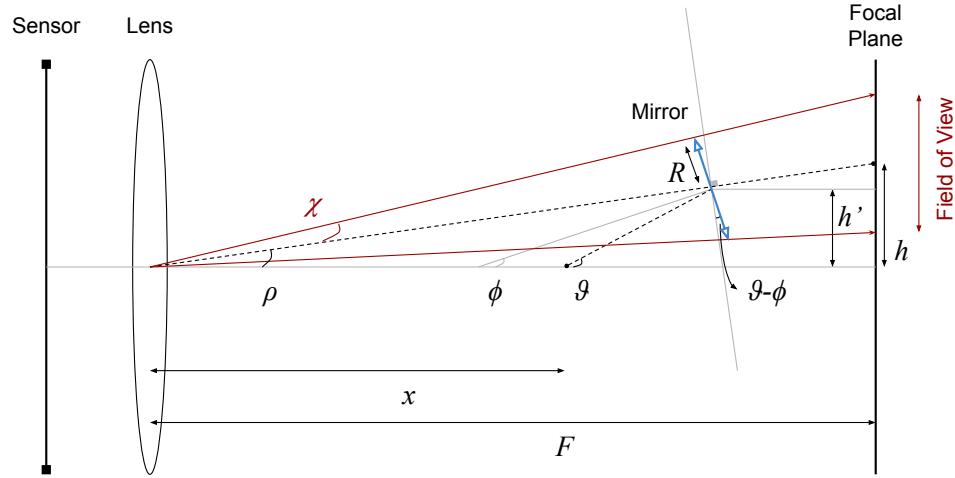


Figure 8.6: Ray diagram of a single folded view. Here the object is placed at a distance x from the lens center, offset from the focal plane of the lens which is at a distance F from the lens. A flat mirror is introduced between the object and the focal plane, which produces a virtual object that lies on the focal plane; this virtual object is focused onto the sensor by the lens. The mirror normal angle ϕ is chosen such that the principal (central) ray emanating from the object (at angle θ) is reflected towards the center of the lens. This constrains the geometry of the mirror and determines the mirror normal angle as described in Appendix B.2. Note that the $\theta > 0$ view angle results in an image that sees the object as if it were rotated by angle θ about the optical axis. Note also that $\rho = 2\phi - \theta$. The marginal rays from the lens center to the mirror edges are shown in red, and corresponding field of view is highlighted on the focal plane. The half angle χ corresponds to the angular field of view and is determined following the logic presented in Appendix B.3. Figure sourced from [41].

the mirror center. This is detailed more explicitly in Appendix B.2.

The point of reflection of the principal rays that pass through the lens center can be treated as the center of the mirror, and a symmetrically shaped mirror can ensure that the light cone emanating from the object reflecting off the mirror stays symmetric about the center. Note that the distance between the virtual object and lens increases by a factor of $1/\cos \rho$ as the virtual object rises above the optical axis. Furthermore, the perpendicular area of the lens aperture (as seen by the virtual object) decreases by a factor of $\cos \rho$. Therefore, when the lens is the light-limiting aperture, this leads to a drop in the light collected from the object as the view gets increasingly off-axis. This effect can be mitigated by using mirrors that function as the light-limiting apertures, and increasing the relevant mirror dimension by the appropriate factor of $1/\cos(\theta - \phi)$ for each mirror. Alternatively, if mirrors of fixed radii are used that form the light-limiting apertures, then a drop in light intensity collected at the lens is expected from similar factors of $\cos(\theta - \phi)$.

The field of view for a single folded view can be geometrically determined by tracing marginal rays from the center of the lens to the edges of the mirror, as shown in Figure 8.6. Under the assumption that the mirror is symmetric about its center, this can be done following the details presented in Appendix B.3. The mirror sizes can be chosen based on the field of view to account for the object of interest. Therefore the system is flexible in its ability to accommodate objects of various scales.

8.3.2.2 Accommodating multiple views – Capturing the light field

The single view geometry can be readily extended to a multi view geometry given that the image of each virtual object occupies a small portion of the imaging sensor. This can be visualized by mapping the image sensor dimensions onto the focal plane of the lens, as shown in Figure 8.7. In this setup, the view angles (θ, ψ) are first sampled on a Fibonacci lattice as follows,

$$(\theta_i, \psi_i) = \left(\arccos \left[1 - 2 \frac{i}{N} \right], 2\pi \frac{i}{GR} \right) \quad (8.1)$$

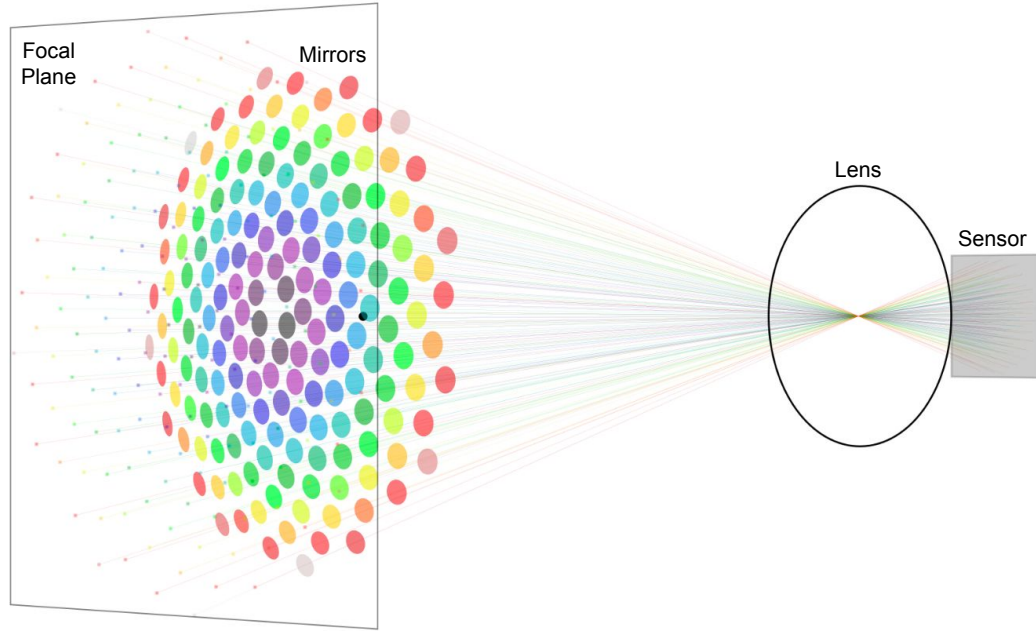


Figure 8.7: Principal ray diagram in a setup with multiple view folding using mirrors. The object is shown as a black dot along the optical axis of the system. Each mirror is colored from blue to red proportionally to the magnitude of the polar view angle θ captured by the mirror. The principal rays are traced with the same color, and are shown to start at the location of the virtual object and end at the sensor. The focal plane is depicted here with a bounding box that corresponds to the size of the sensor divided by the lens magnification. It can be seen that for the given array of mirrors, there is some minimum distance between neighboring virtual objects centers. The density of virtual objects on the focal plane can be modified by increasing or decreasing the number of views and mirrors to appropriately account for the virtual object size. Note here that the view angles (θ, ψ) are sampled on the Fibonacci lattice as governed by Eqn. 8.1, but only those views (and corresponding mirrors) within the appropriate hemisphere, and which produce virtual objects that lie inside the field of view of the imaging apparatus, are shown. Figure sourced from [41].

where $GR = \frac{1+\sqrt{5}}{2}$, $i = (1/2 + j)$ where j is an integer from 0 to N , and N is the total number of points being sampled on the Fibonacci lattice.⁴ Here θ_i correspond to angles with respect to the optical axis, and ψ_i correspond to the azimuthal angles. For Figure 8.7 $N = 800$ angles are sampled on the unit sphere from the lattice, but not all angles are displayed; Only angles in one hemisphere, which upon folding by mirrors produce images that fall inside the sensor, are retained. Flat circular mirrors with equal radii and corresponding normal angles are appropriately positioned around the object. Note that $N = 800$ corresponds to a coarse enough sampling such that none of the mirrors overlap. The sensor dimensions are shown on the focal plane to help visualize how the dense packing of the sensor can work. As seen in Figure 8.7, the virtual objects are not expected to overlap provided the object dimensions are smaller than the inter-virtual object distance. Moreover it can be seen that, for this view angle sampling and mirror size, no overlap of ray bundles occurs as the light reaches each mirror.

In Figure 8.7 the maximum view angle shown corresponds to $\theta = 55^\circ$, as angles beyond that result in virtual objects that lie outside the field of view of the imaging apparatus. Here the camera has a full frame sensor (36 mm $\times$ 24 mm), with a f/1.2 lens of focal length 62.275 mm focused for a magnification of $m = 0.215$. The object distance x from the lens is tuned such that the $\theta = 55^\circ$ virtual object lies at the edge of the viewable area. More generally (refer to Figure B.2),

$$x = F - \frac{h^*}{\tan \phi^*}, \quad \text{where } \phi^* = \left(\arctan \frac{h^*}{F} + \theta^* \right) / 2 \quad (8.2)$$

If h^* corresponds to the maximum height above the optical axis that can be imaged by the lens focused at distance F ,⁵ then to allow a maximum view angle of θ^* , Eqn. 8.2 determines the corresponding object distance x .

As mentioned in Section 8.3.2.1, the defocused image of the object will be formed

⁴It should be noted that the Fibonacci lattice is not the most efficient sampling of points on the surface of a sphere, but it returns points with reasonable density and uniformity for the purpose of this discussion — The view sampling can be trivially switched to a denser lattice if the application requires achieving the maximum packing density of views.

⁵This is distinct from the lens focal length which can be defined as the distance behind the lens where light from infinity is focused at.

in the center of the sensor. To avoid overlapping this defocused image with in focus images produced by folded views, no mirrors are placed close to the optical axis, as shown in Figure 8.7.

8.3.2.3 Related work – Light field imaging

A comprehensive review of light field acquisition and reconstruction techniques can be found in Refs. [406] and [407]. As presented in Ref. [406], light field acquisition can be split into three main categories: multi-sensor capture, time-sequential capture and multiplexed imaging. Multi-sensor capture involves the use of multiple cameras that each capture one part/direction of the light field. Time-sequential capture involves the use of one sensor that is used to capture different parts of the light field by displacing the sensor and taking images for each direction separately. Finally, multiplexed imaging captures the light field by multiplexing the directions into the spatial or frequency domain.

The multiple folded view light field imaging concept presented in this text falls into the spatially multiplexed light field imaging category. However, the folding geometry allows a conventional lens to capture light field directions that exceed what is permitted by the lens' aperture. Therefore the angular span captured by this technique exceeds typical spatial multiplexed imaging methods and is in line with the light gathering abilities of multi-sensor imaging devices. Hence the folded view geometry presented here is a hybrid of the two acquisition modalities, and better described as *super-aperture spatially multiplexed imaging*.

8.3.2.4 Implications for cold atom imaging

It is evident from the above discussion that this multi view imaging device is able to capture the light intensity emanating from the object and factorize it into different directions, effectively capturing the light field of the object. This gain in information, along with the gain in light collection, is crucial to the usefulness of the image captured by this device in reconstruction tasks such as those described in Section 8.6. This opens up the potential for enhanced light collection, one shot 3D reconstruction and

characterization of small scale objects that are dim, fragile, or ephemeral - such as the cold atom clouds seen in atom interferometers.

The total light collected by the light field imaging device is dictated by the array of mirrors and the lens used to image the plane of virtual objects. In the limit of very small mirrors, the total light collected by the systems tends to 50%. In this regime each view has a severely reduced field of view and collects a small fraction of the total light. Note that the maximum is 50% since the mirrors are arranged to cover one hemisphere around the object.

In the limit of large mirrors, each view collects as much light as the lens' aperture maximally allows. Assuming the imaging apparatus is tuned for atom cloud imaging as described in Section 8.3.1, the light field imager can collect approximately m times⁶ the light that the conventional imager collects, where m is the number of views folded with the mirrors. This offers a potential order of magnitude gain in the light collected depending on the size and number of mirrors used. This is done while simultaneously ensuring that the depth of field of the system is acceptable since the same parameters for the imaging apparatus are used as discussed in Section 8.3.1. Each view has an effective f-number greater than or equal to the single-view conventional setup and thus necessarily has a greater or equal depth of field. An experimental realization of this idea is presented below in Section 8.4 where $n = 111$.

⁶Here the $\cos^4\theta$ factor of image plane irradiance for an off-axis (by angle θ) point source [408] is ignored as it is a slowly varying function over the angles concerned. A complete treatment of the light collection from all the off-axis views will result in a gain in light collection slightly less than the factor of m quoted here.

8.4 Experiment

8.4.1 Demonstrator Design

To validate the imaging concept presented in Section 8.3, a 3D printed demonstrator has been made to achieve multi view imaging and demonstrate the 3D reconstruction capabilities of the information recorded by this imaging concept. This section presents the specifics of the design parameters of this demonstrator.

To accommodate constraints from realistic magneto-optical traps and lasers in atom interferometer experiments, the closest mirrors in the demonstrator are required to be at least 3 cm away from the nominal object position. Furthermore, since cameras typically lie outside the vacuum chambers in such experiments, the lens in the demonstrator is required to be at least 15 cm away from the nominal object position. These constraints set the maximum view angle captured by the demonstrator to be $\theta = 55^\circ$ for a 35 mm full-frame sensor.

5 mm mirrors are used in the demonstrator, since these were the smallest commercially available mirrors at the time of building the system. These mirrors are large enough to allow for a sufficient field-of-view for an $\mathcal{O}(1\text{ mm})$ target object, while being small enough to pack the range of solid angles efficiently and accommodate up to $n = 111$ views. Views (mirrors) are not populated in the central region of the array of mirrors to avoid overlapping images from those mirrors with the defocused direct (head-on) image of the object. The results in this paper are produced with a demonstrator holding $n = 90$ mirrors.

The QHY600 astrophotography camera was used for this study, which has a full-frame back-illuminated sensor with a small pixel size ($3.76\text{ }\mu\text{m}$) and 16-bit analog-to-digital bit depth. For the lens, priority was placed on having a wide range of f-numbers, while targeting reasonable magnification and object distance constraints, explained in Section 8.3. The commercially available high performance Nikkor Z 58mm f/0.95 S lens was chosen as it satisfies all the constraints discussed above. Furthermore, this lens has one of the largest lens apertures with good and well-documented off-axis imaging performance.

The array of mirrors are held place at the correct positions and orientations by

two 3D-printed boards as shown in Figure 8.8a. The two boards, with the mirrors sandwiched in between, are tightened using bolts and nuts along the edges of the demonstrator. The *front board* has an array of small rings that define the mirror positions and angles, while the *back board* has an array of cylinders that push and apply pressure to the back of the mirrors to pin them in place (See Figures 8.8c and 8.8d). The thickness of the front board is 1 mm and is thinner than the thickness of the mirrors which is 1.5 mm, ensuring that the cylindrical knobs from the back board are pushing against the rear surface of the mirrors. The front board has angle brackets around its edges, with dimensions such that the front-most plane defined by the four corners of the bracket defines the nominal imaging target position along the optical axis ($x = 0$). The corners of the angle brackets are engraved with grooves, and thin (0.005") fibers are used to place the target object at the precise nominal object position using these grooves (See Figure 8.9).

The demonstrator is 3D-printed using stereolithography (SL) technology with 0.004" Z-resolution and 0.010" XY-resolution, and the optical axis of the demonstrator is aligned as the Z-axis during 3D printing. This sets the positional and the angular resolutions of the mirror holes on the front board. Consequently, the resolutions degrade with increasing view angles. For the largest view angle $\theta = 55^\circ$, the mirror surface normal is at 35° from the optical axis (Z-axis in manufacturing), and this normal surface has a combined resolution of 160 nm. With 5 mm diameter mirrors, this corresponds to an angular resolution of 1.8° . Image-level comparisons with simulations confirm that the errors are smaller than these expectations.

To facilitate optical alignment, the front board has a 12 cm $\times$ 12 cm square alignment grid of 1 mm wide and 1 mm deep lines printed on the flat surface of the front board. The demonstrator is aligned perpendicularly to the optical axis by ensuring that the entire alignment grid is simultaneously in focus while operating the lens at a low depth of field. The effective depth of field of the camera and lens operating at f/1.1 is found to be approximately 2 mm for the alignment grid. Consequently alignment using the grid and the lens at f/1.1 aligns the demonstrator within $\pm 1/60$ rad ($\pm 0.95^\circ$) from the exact perpendicular plane. See Figure 8.10 for an image of the front board alignment grid after the alignment procedure has been completed.

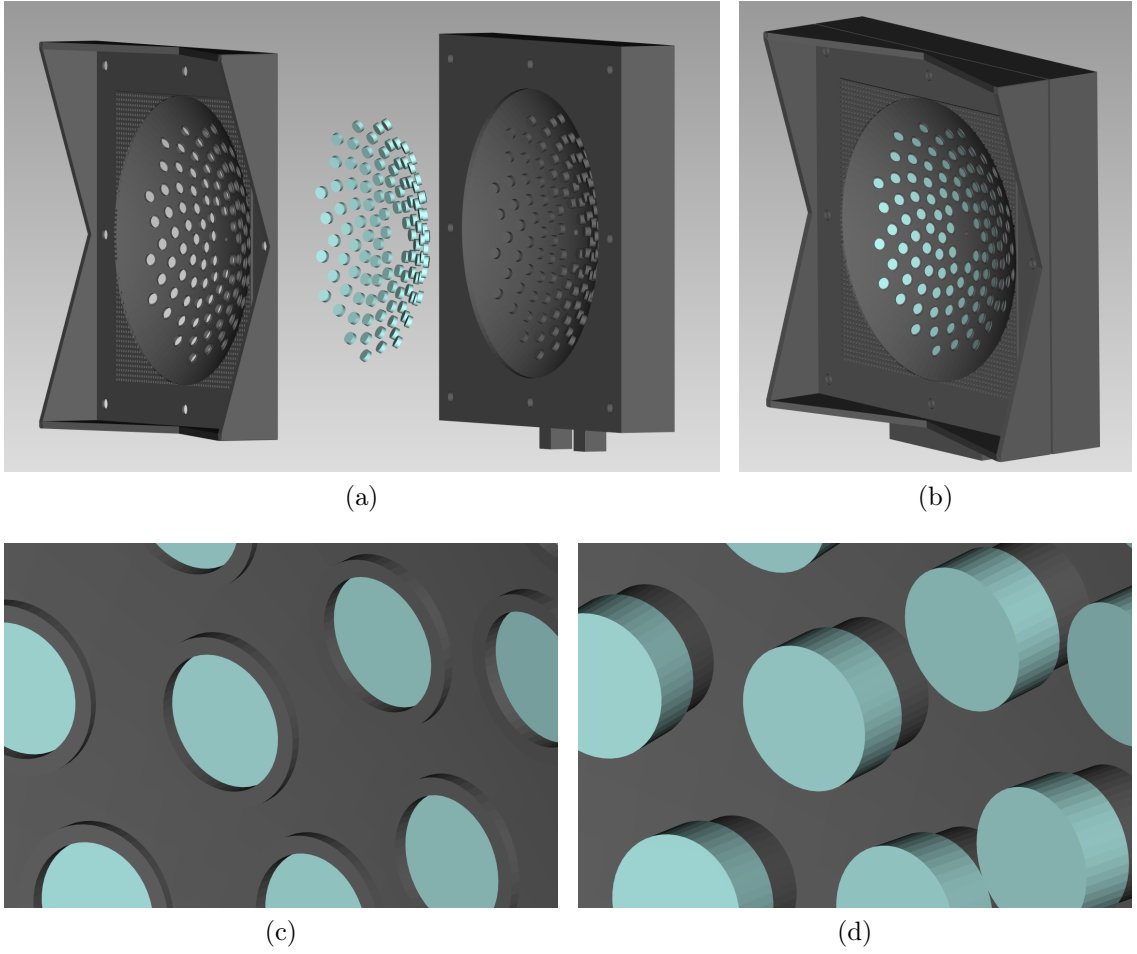


Figure 8.8: Various views of the demonstrator CAD model. The teal disks are the mirrors. (a) Expanded view. The black parts to the left and the right of the mirrors are the front board and the back board, respectively. (b) Assembled view of the mirror holder. (c) A zoomed-in view, showing each mirror being pushed against a ring-like structure on the front board. (d) A zoomed-in view without the front board, showing each mirror being pushed by a cylindrical structure on the back board. Figure sourced from [41].

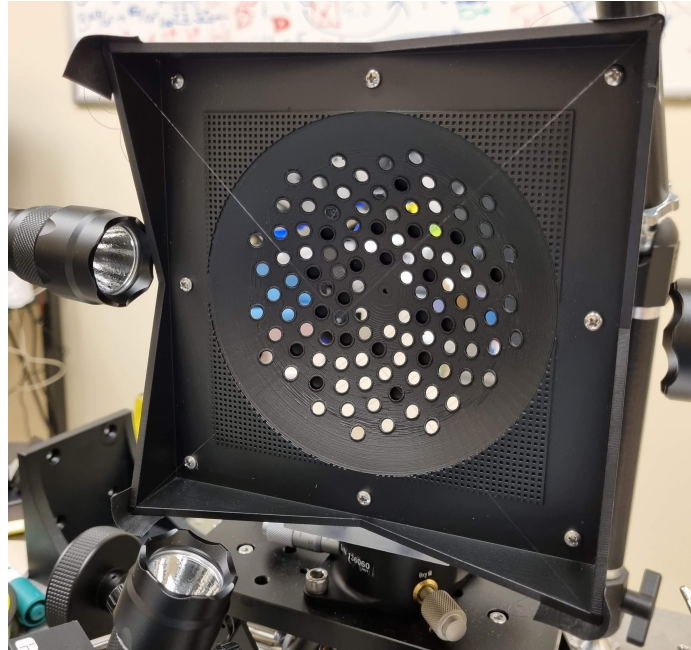


Figure 8.9: Laboratory photo of the demonstrator. A pair of thin (0.005") fibers are used to create a crosshair that defines the nominal target object position and holds the object in place. Figure sourced from [41].

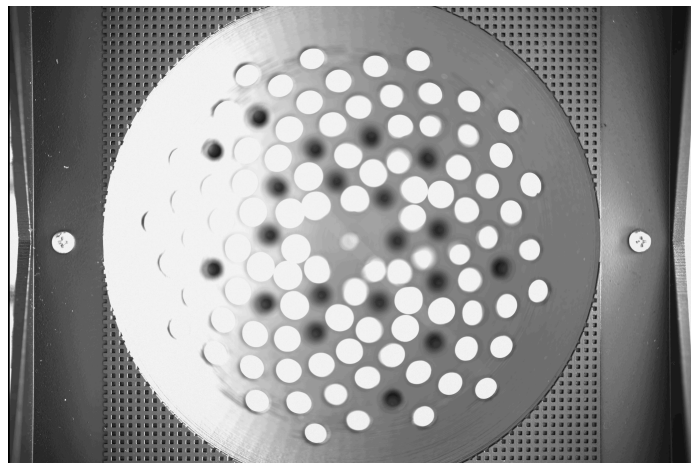


Figure 8.10: Image of the alignment grid of the demonstrator. The entire grid is focused within depth of field of 2 mm. Figure sourced from [41].

8.4.2 Imaging with Demonstrator

To test the imaging system and validate the multi view imaging concept, the demonstrator is used to take images and perform 3D reconstructions of a test object. The test object is a framed cubic shape that is $\mathcal{O}(1\text{ mm})$ in overall size and has $\mathcal{O}(100\text{ }\mu\text{m})$ feature, and it spells out the letters D, O, and E along each of the three orthogonal projections (See Figure 8.15 for the orthogonal projections of the test object, and Figure 8.13 for views of the test object from an array of different orientations). The test object is 3D printed with projection micro-stereolithography (PμSL) technology, which enables precise manufacturing of micron-scale features.

The test object is manufactured using HTL resin [409], which fluoresces broadly over 430 nm – 550 nm when illuminated with 405 nm light. Images of the test object are captured by illuminating the object using 405 nm LED lights, with a 450 nm filter placed in front of the camera sensor. The object is illuminated from five different directions: top, bottom, left, right, and through the central hole in the demonstrator. While this arrangement does not illuminate the object uniformly and the object does not irradiate uniformly over 4π solid angle, the reconstruction method is able to capture the view-angle dependent light field.

8.5 Simulation and reconstruction

For analysis and reconstruction using the system described in Sections 8.3 and 8.4, a differentiable optics and light transport simulation has been developed, called **GradOptics**, in which this imaging system has been implemented [410]. The framework uses geometric optics and automatically simulates effects such as depth of field. It can also incorporate point spread functions / modulation transfer functions in the imaging process if specified. The framework is fully differentiable, allowing for gradient-based reconstruction of the simulation inputs and gradient-based calibration. A review of physically-based rendering frameworks, which have informed this framework design, can be found in Ref. [411].

The reconstruction here is done using the views generated by the optics as described in Section 8.3. These views are formed by extracting pixel patches from the full image captured by the camera, where the extraction is achieved using a patch-finding algorithm. The individual views can then be matched with corresponding views from the simulator, with the goal of the reconstruction being to produce a 3D density function that produces simulated views that match the views extracted from real image data. In this work, the 3D densities that correspond to the object of interest are represented using neural networks, leveraging benefits from recent developments in neural scene representation [412, 413] and, more broadly, implicit neural representation [414, 415]. Implicit neural representations provide greater flexibility and tighter integration with the simulator than standard reconstruction methods, such as filtered back projection [416].

Reconstructions of the 3D printed target object (Section 8.4.2) are performed using the views generated by the demonstrator (Section 8.4.1), to assess the capabilities of the demonstrator and the underlying imaging concept. Note here that the 3D printed target object has well defined surfaces and is semi-opaque. Therefore, the NeuS [417] surface reconstruction method is used to reconstruct the object from the views captured by the demonstrator. This represents a surface via two neural networks: a signed distance function (SDF) and a view dependent color. The SDF neural network takes a position as input and outputs a density at that point in space,

similar to recent methods in volumetric rendering [411, 413, 418]. The learned density peaks about the 0 value of the SDF, and this manifold can be taken as the learned surface of the object. The color neural network takes both position, and angle at which the point in question is being viewed at, as inputs and outputs the corresponding color for that point in space. The neural network predicting view dependent color is needed due to the non-uniform illumination of the object. The weights of the neural network are optimized with gradient descent as the simulated views and real data views converge, and the resulting the neural networks model the density and color of the target volume. Note that the optics simulation of the physical demonstrator is used for the generation of views at each training stage during the learning process. While the simulation parameters are kept fixed to nominal values during training, co-optimization of the parameters and the object model is an interesting direction for future work.

In addition to performing reconstruction using the physical demonstrator, reconstruction studies are also performed in simulation using simulated atom cloud densities. These studies are done to assess the efficacy of the system in the planned atom interferometer application. For the reconstruction of the atom cloud densities, the problem reduces to that of tomographic reconstruction. Like the SDF network, the volume in the tomographic case is modeled with a neural network that takes a position as input and outputs a density at that point in space. This work also utilizes SIREN [414], which is a recently proposed architecture for modeling 3D densities that uses neural networks with sinusoidal activation functions to learn high quality volumetric signals. Tomographic reconstruction is also used for studying the tolerance of the system to misalignment and 3D printing errors.

8.6 Results

8.6.1 Simulation results

The imaging of an atom cloud with a sine-wave density modulation fringe pattern is simulated to demonstrate the tomographic capabilities of the light field device in atom interferometers. The number of atoms is sampled from a Poisson distribution with $\lambda = 10^6$ atoms, where each atom emits 10^8 photons per second. Figure 8.11 shows 110 (out of 111) views extracted from a simulated light field image, using an exposure time of $10 \mu\text{s}$.

For reconstruction, the views from Figure 8.11 are used to train the network as described in Section 8.5. Figure 8.12 shows the learned density after training the neural network for 5,000 epochs. It can be observed that the input views acquired in a single shot with the multi view device allow successful reconstruction of the atom cloud density without using domain knowledge or regularization.

8.6.2 Captured data

Images captured by the demonstrator are 9600×6422 pixels, each with a set of embedded views of the imaged object. As mirrors within the constructed imaging apparatus may differ slightly from the expected position due to imperfect construction or alignment, mirrors in simulation and observed data must be matched before per-view comparisons can be performed. A patch finding algorithm is run using tools in OpenCV [419] to extract pixel patches corresponding to each view, and the centroids of these patches are used for comparison. The centroids from the extracted views are compared to the positions of each view in the simulator corresponding to the central ray from the corresponding mirror for the given view — The centroid closest to the central ray point on the sensor from a given mirror is identified as the view corresponding to that mirror. The set of views used for the reconstruction results presented here are shown in Figure 8.13. Out of the 90 views instrumented in the demonstrator, 86 views are found by the patch finding, of which 82 views are available for reconstruction once an upper limit of 350 pixels is enforced on the distance between

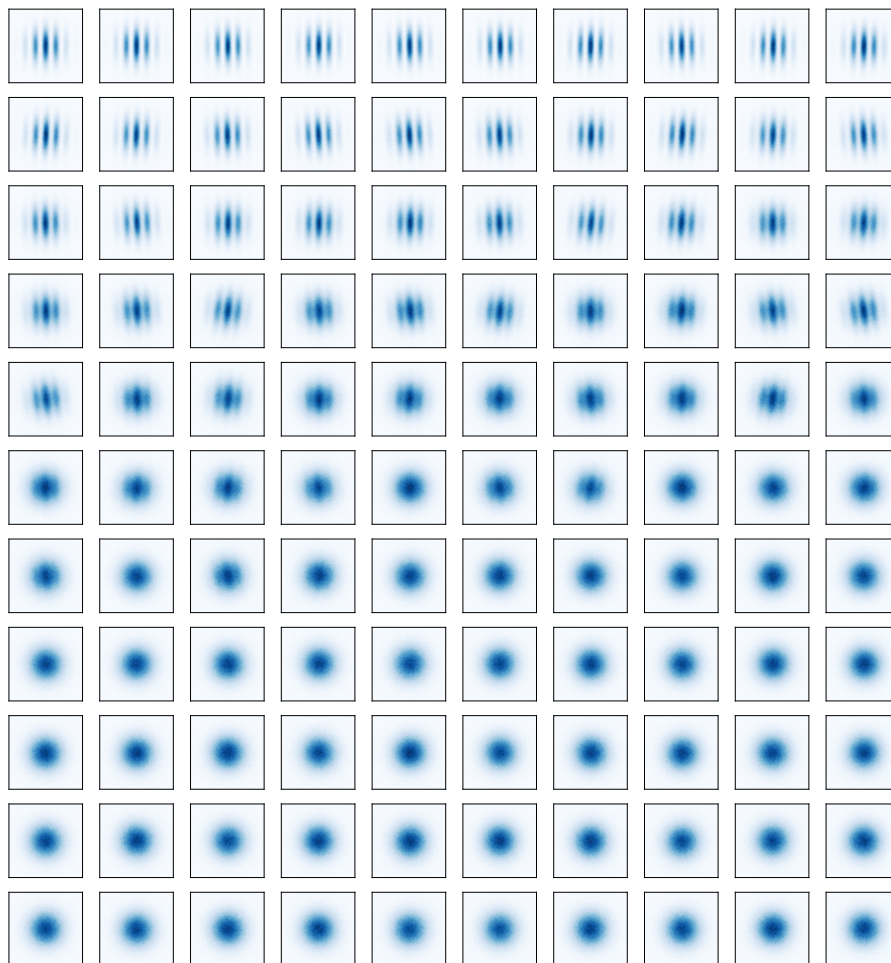


Figure 8.11: Grid of views extracted from a simulated light field 60MP image of an atom cloud imaged over $10\text{ }\mu\text{s}$. Figure sourced from [41].

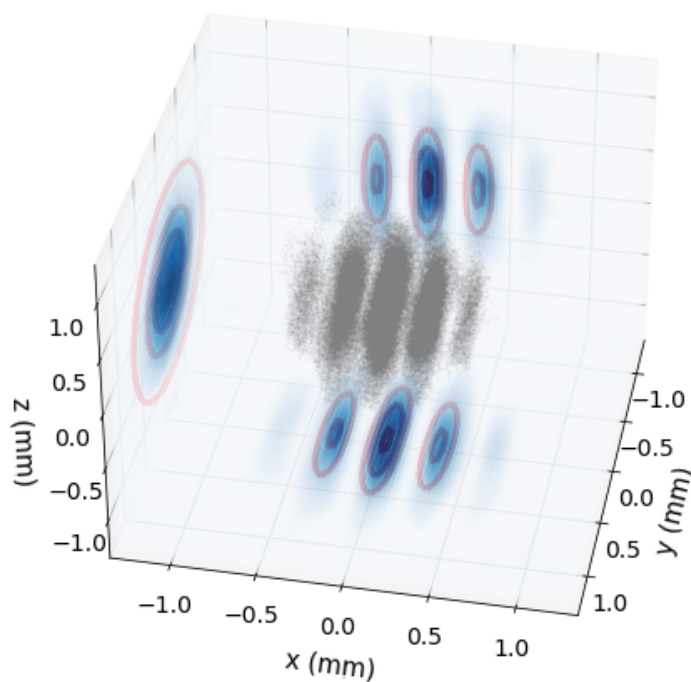


Figure 8.12: Data sampled from the learned 3D cloud model in gray, the learned 2D marginals in blue, and the contour lines of the target density in red (the 40, 68 and 95% contours are shown). Figure sourced from [41].

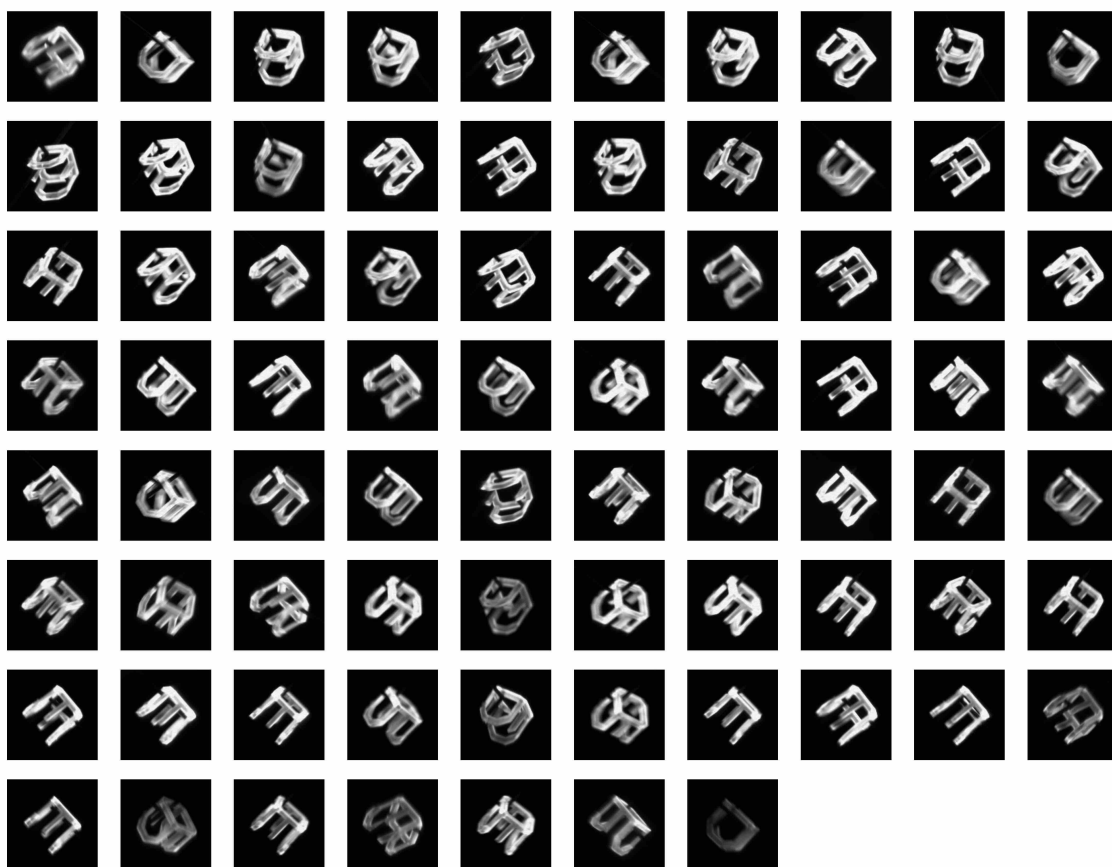


Figure 8.13: Grid of training views extracted from a 60MP image of the test object, taken in a single shot. Figure sourced from [41].

the real view centroid and the simulated central ray position for the corresponding mirror. Reconstruction quality is improved by discarding an additional 5 views which are poorly centered⁷ or poorly illuminated. Note that the issue of illumination of the mounting fiber arises due to the peculiarity of this demonstration setup, and the impacted views are recoverable with a different object mounting technique or further post-processing. The views which are poorly illuminated or are lost to the patch finding algorithm are likely the result of individual mirror misalignments, and may be improved with a more sophisticated mirror mounting or higher precision construction of the device.

⁷Due to shifting in the bounding box due to illumination of the suspending fiber, which moves the centroid.

8.6.3 Experimental results

For the reconstruction, the extracted patches from the captured image data are assumed to be at the corresponding nominal positions from simulation and centered at the expected corresponding view position. These patches, with the correspondingly identified mirrors and pixel positions, are used as the inputs for the reconstruction methods described in Section 8.5. Patches of 130×130 pixels are used for the results presented here. The NeuS reconstruction is trained for 100000 iterations on the 77 views shown in Figure 8.13, where a training iteration sweeps across randomized views, and uses a sampling of 512 pixels from a view. No explicit masking, i.e. supervised information about the object location in each view, is used and the distinction between the object in the foreground and the dark background is learned directly from data. The background is modeled by NerF++ [420] following the standard NeuS implementation, and network and training settings closely follow those used in Ref. [417]. Full hyperparameter details can be found in the appendix of Ref. [41].

The quality of the learned NeuS model can be estimated from Figure 8.14, where a subset of real data views are compared with the corresponding views generated from the learned NeuS model. Results are nearly identical between the real and generated images, with an average SSIM⁸ [421] across training views of 0.98 (1 is perfectly identical, 0 is no similarity), demonstrating the good convergence of this trained model. The full grid of generated training views is available in Appendix B.4.

Since the goal here is to reconstruct a three dimensional object, it is useful to analyze the learned 3D surface, which is represented here by the zero level set of the learned SDF. To do so, the learned SDF is sampled on a regular grid of 64^3 points within a box of 2.02 mm per side centered at the nominal object position, with this sampling grid chosen to roughly match the anticipated resolution. A marching cubes algorithm within the trimesh software [422] is then used to construct a triangular mesh of the learned surface. Orthographic depth maps of the three D-O-E letters from this

⁸The structural similarity index measure (SSIM) is an image similarity metric that uses structural correlations between images, rather than evaluating absolute errors as in the case of the mean squared error (MSE) or peak signal to noise ratio (PSNR). This is done by evaluating metrics over regions of each image being compared, rather than using a simple pixel-by-pixel difference metric.



Figure 8.14: Comparison of subset of training (real) views to the corresponding subset of generated views from the learned NeuS reconstruction. Results are nearly identical between real captured images and their generated counterparts, with an average SSIM across training views of 0.98. Figure sourced from [41].

learned surface are compared to the corresponding orthographic projections from the CAD model and surface images from a microscope in Figure 8.15. Figure 8.16 shows the surface mesh from a variety of different orientations, demonstrating the true 3D structure of the learned surface. The object is reconstructed with high fidelity, with very clear D-O-E lettering as well as more minute structure, such as the flaring out of the legs of the E. The reconstruction also manages to recover the $120\text{ }\mu\text{m}$ fiber used to suspend the D-O-E object in front of the device, a notable artifact present relative to the CAD and microscope images. The reconstruction is primarily limited in regions where there is minimal information from the training views, such as the E face on the back (from the mirror perspective) or the right side of the O, as shown in Figure 8.16 — This is expected behavior due to the limited angular span of the demonstrator. Despite this, however, much of the structure is still inferred from the views captured by the device.

Information about the learned 3D surface is fully contained in the SDF, and can be visualized using colorless meshes, as done in Figure 8.16. It is also interesting to analyze the learned color as an insight into the learned light field. This can be done by sampling views from the learned model that correspond to a diverse array of angles — Here views are sampled on an even grid of angles between -55 and 55 degrees measured both vertically and horizontally from the optical axis. This set of views is shown in Figure 8.17. Note that the center of the grid corresponds to the center of the device and is correspondingly the brightest, while the edges of the grid correspond to regions on the outer edge/outside of the device, and the learned

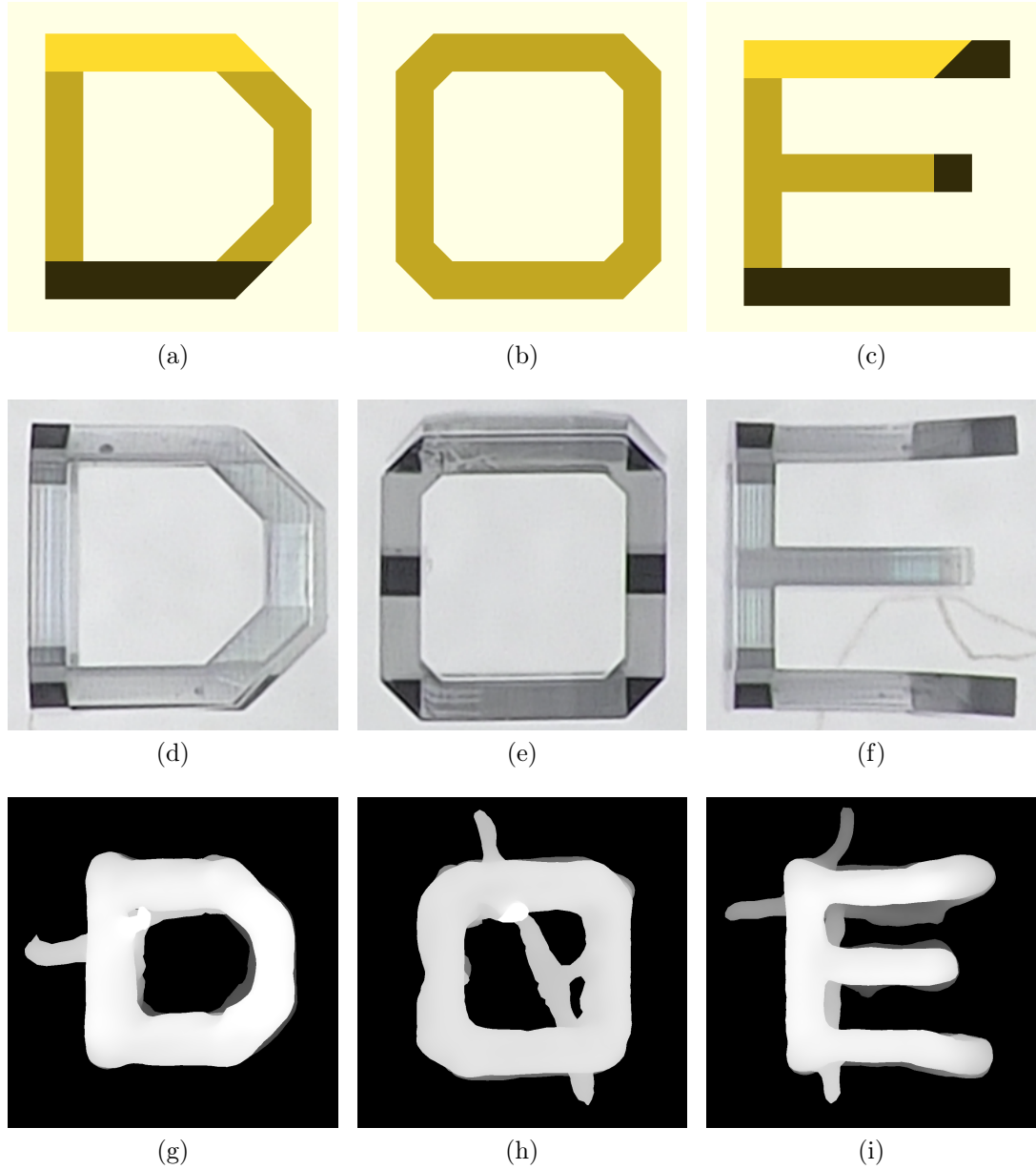


Figure 8.15: Reconstruction of the cube test object. The envelope dimension of the object is $(1\text{ mm})^3$, and the letter is $142\text{ }\mu\text{m}$ thick. (a-c): CAD design, (d-f): microscope images of a 3D printed object, (g-i): depth map of surface reconstructed using 77 views from the system. Different columns correspond to different views of the same objects. Note that the reconstructed results contain artifacts from the $120\text{ }\mu\text{m}$ fiber used to suspend the object, as well as from the limited range of viewing angles. Figure sourced from [41].

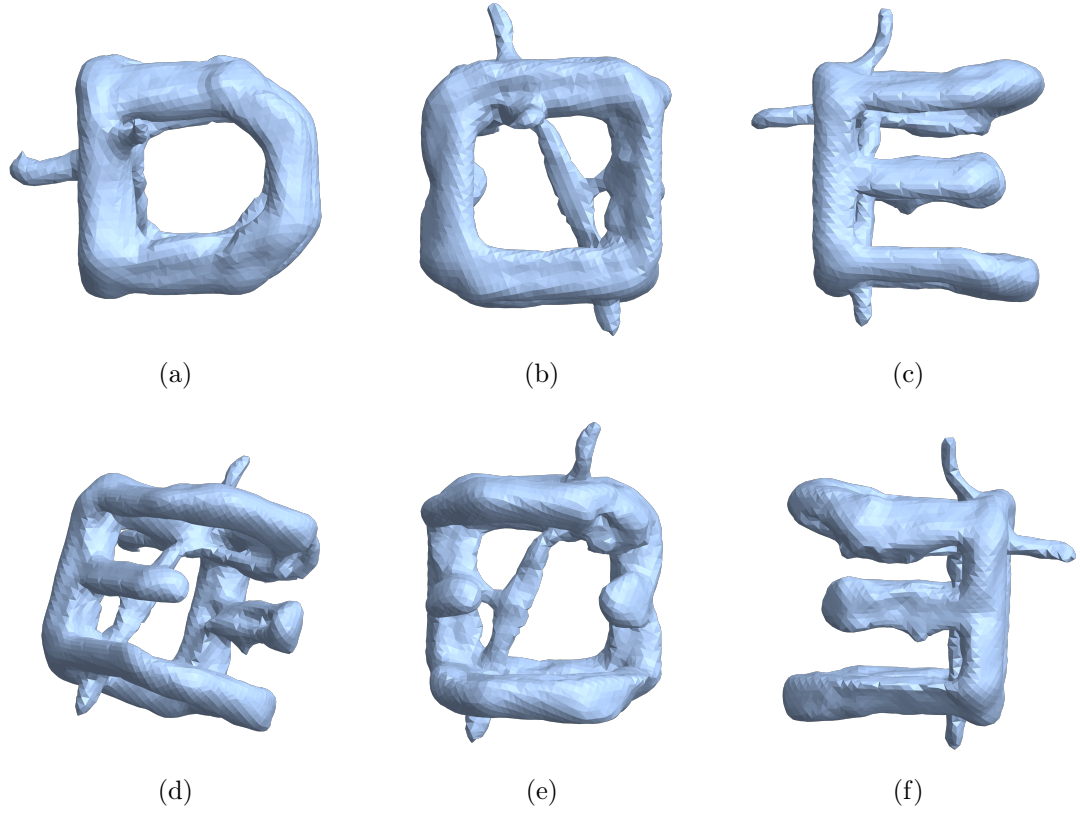


Figure 8.16: Images of reconstructed surface mesh. The top row (a-c) shows similar D, O, and E views as those presented in Figure 8.15. The bottom row (d-f) shows alternative views, where the first two columns show angles of the object that are not directly a D, O, or E lettering, emphasizing that a 3D model has been learned for the imaged object, while the third column shows the reconstruction of an object face that is never directly imaged by the system (the E face on the back). While quality is worse on this face (f) of the object, the available information from the views that are imaged is sufficient to infer the general structure of the object. Figure sourced from [41].

intensity is correspondingly dimmer. This drop in brightness from central views to the edge views is expected since the mirrors form light limiting apertures subject to the cosine drop off, compounded further by the cosine-to-the-fourth drop in illumination expected from off-axis imaging, as discussed in Section 8.3.2.1. The effect is also exacerbated by the non-uniform illumination used in this demonstrator. Therefore, the quality of the views in Figure 8.17 and their illumination characteristics suggest that the reconstruction has learned the appropriate 3D structure of the object as well as its light field.

The reconstructed features shown in the Figures 8.16 and 8.17 provide a qualitative measure of resolution for the system with the chosen reconstruction algorithm. The resolution of the system is further quantified using the Fourier shell correlation (FSC) [423] metric, which compares the correlation between two volumes as a function of spatial frequencies — This metric is used extensively in several other contexts, including cryogenic electron microscopy. Pure signal, or perfect reconstruction, corresponds to perfect correlation or a FSC at 1 for all spatial frequencies, while pure noise corresponds to no correlation, or a FSC at 0 for all spatial frequencies. Since the orientation of the imaged test object is not known exactly, making a comparison to ground truth is difficult in this case. Therefore the standard procedure of *split-halves* comparison is followed — the set of training views is randomly split into two halves, and the correlation is calculated between the reconstructions from the two independent halves. Figure 8.18 shows the resulting Fourier shell correlation curve, with error bars derived from running this procedure with various random splits. The standard threshold⁹ for such split halves comparisons is 0.143 [424], corresponding to a resolution of $71 \pm 17 \mu\text{m}$ for the system and the chosen dataset and reconstruction method.¹⁰ This resolution, as well as the regime of significant correlation shown by the full curve, is compatible with the qualitative results shown above, and meets the design specifications of the demonstrator.

⁹Thresholds determine the spatial frequency at which the transition from signal to noise occurs.

¹⁰The threshold of 0.5 is used when the correlation is computed between ground truth and reconstruction, which is not applicable here, but the threshold is shown on Figure 8.18 for completeness.

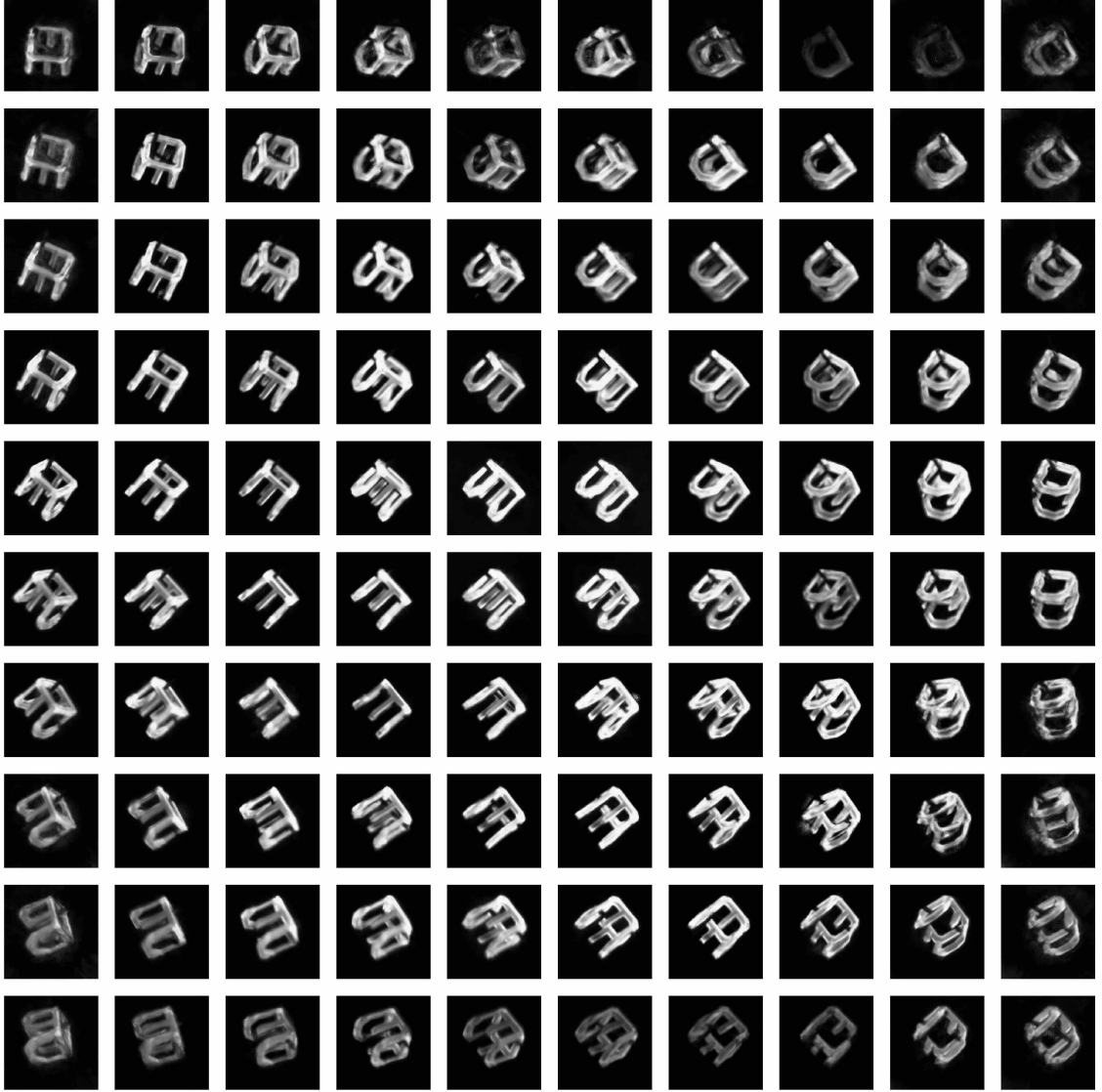


Figure 8.17: Interpolated views generated from the learned NeuS reconstruction. The range of view angles is between -55 and 55 degrees, and resulting views contain no explicit correspondence to those used for training. Image quality is generally high, demonstrating a learned 3D structure and light field. Figure sourced from [41].

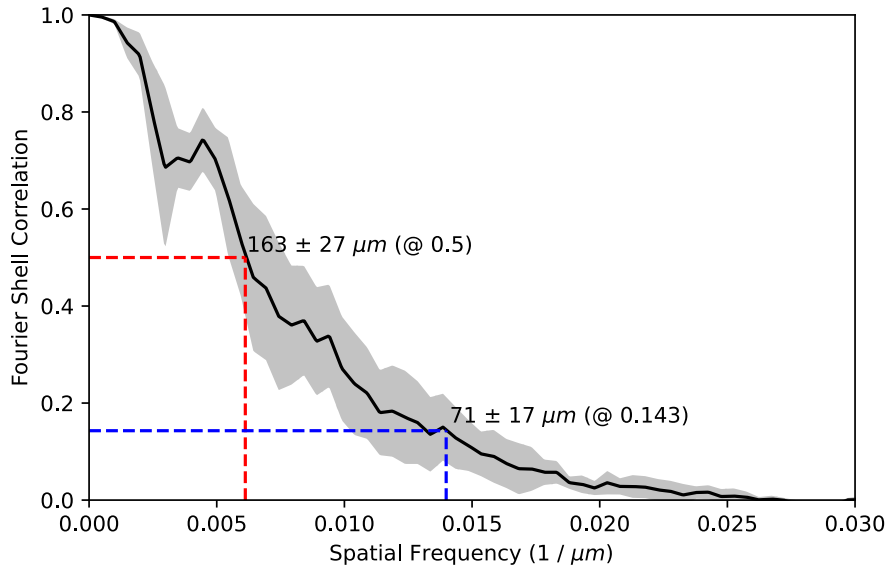


Figure 8.18: Fourier shell correlation curve for the reconstructed D-O-E object. 77 extracted views are randomly split into two distinct halves, and each half is separately reconstructed. All reconstructions are trained for 140000 iterations, and the Fourier shell correlation is calculated by comparing logistic densities of learned width centered at the learned surfaces. Results shown are the mean and standard deviation calculated from running this procedure across 5 different random splits of the 77 views. Resolutions are quoted for two commonly used thresholds, with $71 \pm 17 \mu\text{m}$ resolution corresponding to the standard threshold used for split dataset comparisons (0.143). Results are qualitatively consistent with the size of reconstructed features. Figure sourced from [41].

8.6.4 Impact of misalignment and 3D printing resolution

While the simulator introduced above assumes a perfectly calibrated device, the instrument is imperfect due to misalignment and 3D printing errors. Results from Section 8.6.3, however, suggest that in the demonstrator these effects are small enough such that they do not adversely impact the reconstruction, and no calibration of the simulator is necessary.¹¹ The tolerance of the (uncalibrated) system to such errors can be gauged with synthetic images of the test object generated with artificially introduced displacements. The impact that these displacements have on the quality of the reconstruction is studied. Five effects are considered:

- Displacement of the entire demonstrator along the optical axis Δx_{dome} .
- Stochastic displacement of the mirrors along the optical axis $\Delta x_{\text{mirrors}}$.
- Displacement of the entire demonstrator in the plane perpendicular the optical axis $\Delta y_{\text{dome}}, \Delta z_{\text{dome}}$.
- Stochastic displacement of the mirrors in the plane perpendicular the optical axis $\Delta y_{\text{mirrors}}, \Delta z_{\text{mirrors}}$.
- Stochastic displacement of the orientation of the mirrors $\Delta\theta, \Delta\phi$.

For each of the above effects, an imperfect device is simulated and used to generate a light field image of the test object. Note here that the stochastic displacements account for the 3D printing errors that can effect the mirrors individually. The image is then used to perform tomographic reconstruction with a simulator that does not account for the displacement effects (an uncalibrated simulator). Finally, the quality (FSC) of the reconstruction as a function of the magnitude of the introduced displacements is reported (Figure 8.19). Also note here that tomographic simulation and reconstruction of the test object is being performed, as opposed to the assumption of an opaque test object in the previous section. This is done to understand the tolerance of the system with the targeted atom cloud application, while retaining the

¹¹A detailed calibration procedure will be explored in a future publication.

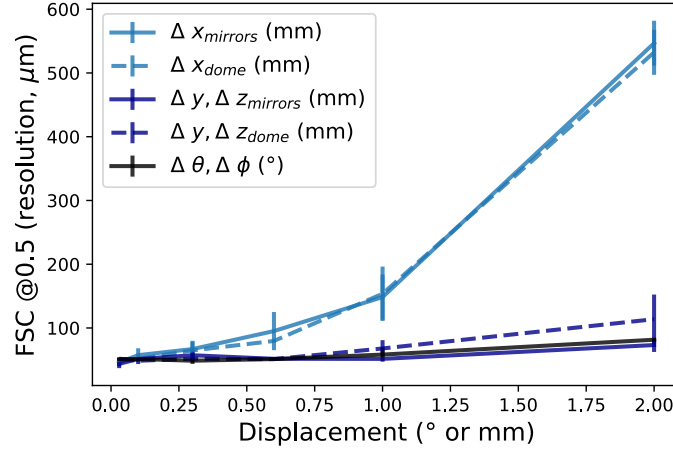


Figure 8.19: Resolution of the system as a function of misalignment and 3D printing error. The resolutions were obtained by computing the FSC between the ground truth and the 3D reconstructed models of the test object. Without calibration of the simulator, displacements below 0.6 mm or 2° still allow to reconstruct 100 μm features. Results shown are the mean and standard deviation calculated from running the reconstruction algorithm with 5 different seeds. Figure sourced from [41].

rich spectrum of features of the test object. The reconstruction quality is reported here as the spatial resolution of the reconstructed object using the FSC. The resolution at which the FSC between the ground truth model and the learned model crosses the threshold value of 0.5 [424] is reported here, since the correlation is evaluated between reconstruction and ground truth and unlike the case with the real image data. It is observed that the displacements of the mirror orientations and displacements in the plane perpendicular to the optical axis does not considerably impair the resolution of the system; 100 μm features can still be reconstructed with a 2° or 2 mm deviation. However, displacements along the optical axis are much more critical and should be carefully minimized when designing the system — This is intuitively expected as such displacements affect the magnification of the image, along with displacing the image on the sensor, and as a result have a larger impact on the resulting reconstruction.

8.7 Conclusions

Atomic sensor, such as atom interferometers, offer an abundance of scientific opportunities to look for physics beyond the standard model and make record breaking quantum measurements. To harness the full capabilities of such systems, new imaging technologies are required that can maximizing light collection while keeping a high depth of field, and that can facilitate 3D imaging reconstruction.

In this chapter, the design and prototype of a novel single-shot light field imaging device has been presented, that is suitable for the MAGIS-100 experiment. Results using a 3D-printed target demonstrate that the system can achieve its design requirements. This super-aperture spatially multiplexed light field imaging device can have broad applicability in scientific and industrial applications that can benefit from single-shot 3D imaging of a small low light emitter. Examples of such applications include 3D inspection of industrial parts, 3D-printing inspection lines, optical particle tracking, X-ray phase contrast photography, etc.

This device has several key advantages compared to other light field systems. It provides high light collection, large solid angle, good resolution, field of view, and depth of field, all in a single-shot, and uses commercially available cameras and lenses. It is also, in essence, a 3D microscope capable of characterizing an object that may be very dim in a single shot. This is a fundamentally different way to do 3D microscopy than structured light microscopy with multiple shots or varying lighting.

There are several directions of future research that can be conducted to advance the imaging modality presented here. The planned next publication will involve work to adapt this imaging device to capture images of Sr atoms in a magneto-optical trap, and perform tomographic reconstruction using the captured data. This will be the first demonstration of single-shot tomography for cold atom clouds. Furthermore, there are several planned projects to advance the optics simulator introduced in this Chapter. These include a detailed calibration procedure that uses captured data and differentiable programming to learn the misalignments in the experimental setup. These will also include an enhancement to the simulator setup that can take wave optics into account naturally include the effects from point spread functions, etc.

Chapter 9

Concluding Remarks

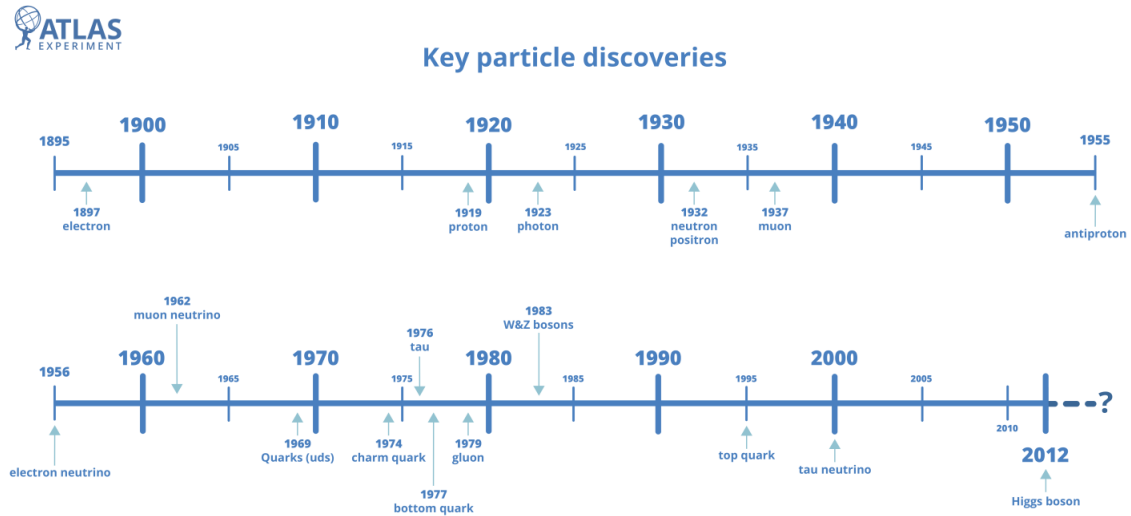


Figure 9.1: Timeline showing the discoveries of various particles of the Standard Model of Particle Physics. Figure sourced from CERN [42]

The Standard Model of particle physics is one of the most precisely tested theories of humankind, describing the electromagnetic, weak, and strong forces of nature and interactions between the fundamental fermions that constitute matter. A timeline of the various discoveries of Standard Model particles is shown in Figure 9.1. With the discovery of the Higgs boson in 2012, the Standard Model particle zoo is complete.

Despite its many successes, the Standard Model does not account for a number of experimental observations, as accounted in Section 2.6. Therefore, the quest to find physics beyond the Standard Model is one of the most significant undertakings of modern physics. This is complemented by the vast diversity of modern experimental modalities that are capable of conducting this search over a large range of energies and models. Foremost in this effort are the experiments at the Large Hadron Collider, which will continue to see proton collisions at unprecedented energies and luminosities, making them ideal to conduct a plethora of searches and precision measurements.

The ATLAS experiment at the LHC is expanding its program of searches for new physics, and preparing to harness the potential of increased statistics at the LHC. Chapter 5 of this thesis presents a novel search conducted with the ATLAS experiment, looking for exotic decays of the Higgs boson, via coupling to a new (pseudo)scalar field, to four gluons. With this foray into extended all-hadronic final states of the Higgs coupling to new scalars, ATLAS is expanding its active program to scan the various possible decay channels of new scalar fields. The di-gluon tagger used in this analysis shows that the substructure information within the scalar decay jets contain information sufficient to distinguish these jets from the standard quark/gluon background. This presents the tantalizing prospect of utilizing similar techniques to look for di-quark decays of the scalars, where the quarks could be light quarks or charm quarks. Additionally, these methods can also be extended to searches for boosted Standard Model Higgs to di-gluon decays, where the two gluons are captured within the same jet in the detector.

To improve the physics yield of the data that is collected by the ATLAS detector, various detector performance enhancements are being developed by the collaboration. The research presented in this Chapter 6 and Chapter 7 highlight two of such detector performance projects that promise improvements across the detector. Chapter 6 presents a vertex selection algorithm that can improve the physics yield of challenging final states like VBF Higgs to invisible decays by $\sim 10\%$ by combining the information from the tracking and calorimetric detectors. Furthermore the research also identifies a unique event classification scheme that can save events with useful physics that would otherwise be discarded due to incorrectly chosen vertices. Chapter 7 presents

a machine learning framework for jet calibrations at ATLAS that is demonstrated to calibrate multiple observables at once while accounting for calibration affecting parameters. Additionally the research introduces an innovative objective function that enables the use of gradient based learners, such as neural networks, for calibrations at ATLAS by allowing the methods to learn the mode of the conditional response distributions. This unlocks the potential for further machine learning driven innovations in regression tasks across the various detector subsystems.

While proton colliders remain fertile soil in the hunt for new physics, various new quantum sensing experiments are now being developed to greatly expand the reach of searches for new physics — Several of these technologies are focused on determining the nature of dark matter. Atomic gradiometer interferometric sensing is one such novel technology, aimed at looking for ultra light bosonic dark matter. This experimental technique is also an active site for collaboration between the high energy physics (HEP) and the atomic-molecular-optics (AMO) communities, leading to unique interdisciplinary research efforts. Chapter 8 presents the results of one such effort, developing a super-aperture spatially multiplexed light field imaging concept for fluorescing cold atom clouds. The technique enhances the light gathering capability of a conventional lens limited by few windows into the active volume of the experiment. The technique also enables the 3D reconstruction of the atom cloud densities directly from a single shot since directional information of the light emanating from the atom clouds is preserved. The concept therefore represents a paradigm shift in the imaging of atom clouds in atom interferometer experiments. The applications of this light field imaging concept even extend beyond atom interferometers, aiding in single shot light field capture for any system with small (relative to the imaging apparatus) objects of interest.

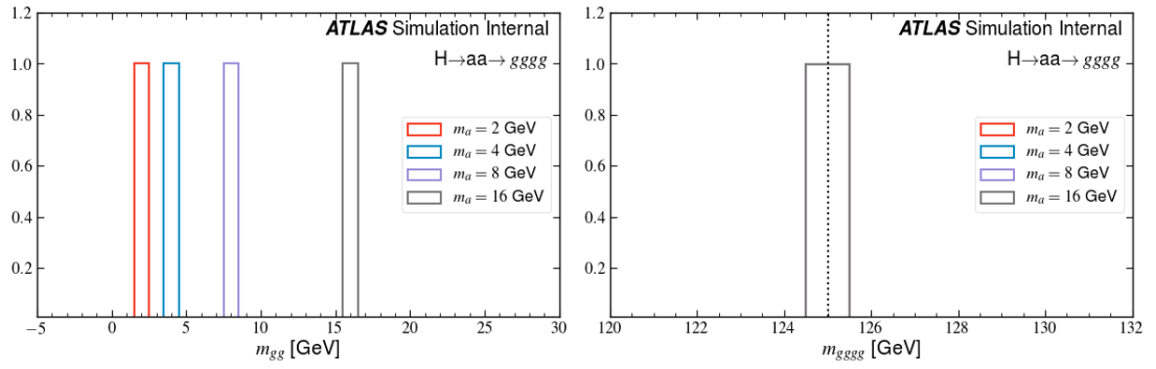
The various developments presented in this thesis will aid in the search for new physics, both at the LHC and at atom interferometer experiments — As such, they represent small steps in the forward march of the science of fundamental forces and interactions of the universe.

Appendix A

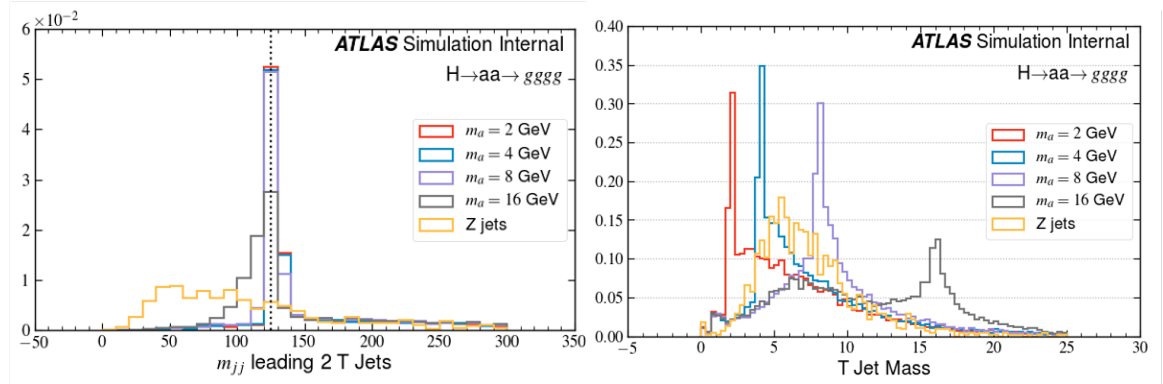
A Search for Exotic Higgs Boson Decays to Four Gluons Using the ATLAS Detector: Appendix

A.1 Signal Validation Plots

Figure [A.1](#) and Figure [A.2](#) show validation plots using truth (particle-level) MC samples. These validate the parameters of the signal simulations used to generate the samples used for the analysis.



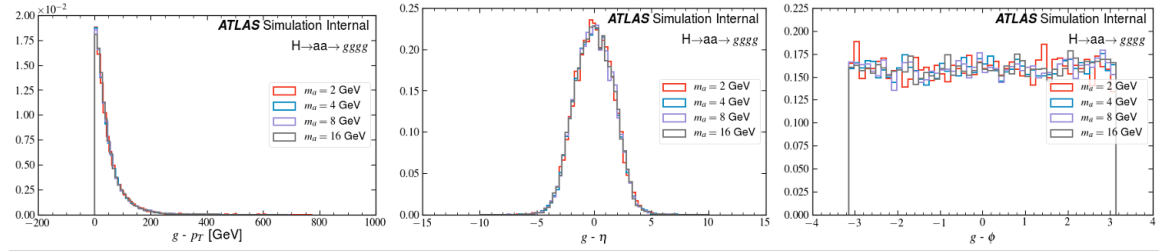
(a) Mass of the digluons and the joint mass of the full system of four gluons



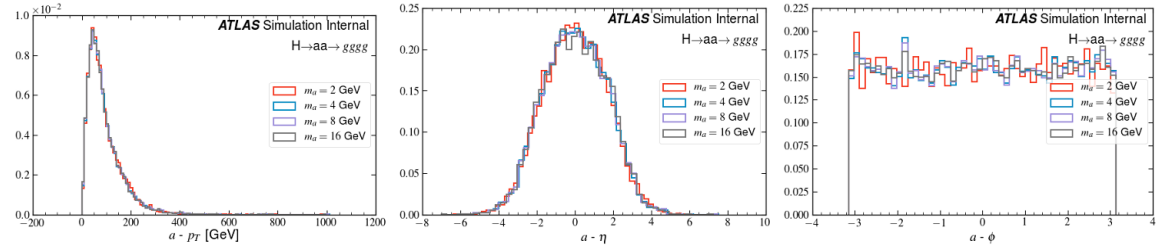
(b) Joint mass of the two leading truth jets

(c) Mass of the leading truth jets

Figure A.1: Mass Validation Plots at Truth Level



(a) Kinematics of the scalars



(b) Kinematics of the gluons

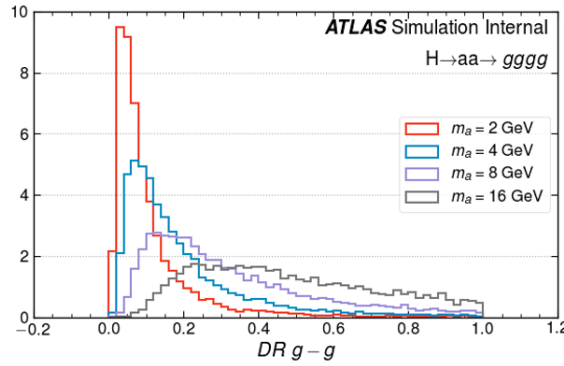
(c) Distance between the two gluons from a scalar decay in $\eta - \phi$ space

Figure A.2: Kinematic Validation Plots at Truth Level

A.2 Reweighting Performance

The following plots show the performance of the reweighting network trained on the nominal Pythia samples, and trained on the variation Sherpa samples. Note here that the performance is evaluated over the full 139fb^{-1} dataset, and therefore the signal mass region ($86 < m_{jj} < 165$) is blinded — the interpolation ability of the reweighting is demonstrated in Figure 5.7.

A.2.1 Pythia Z + Jets Samples

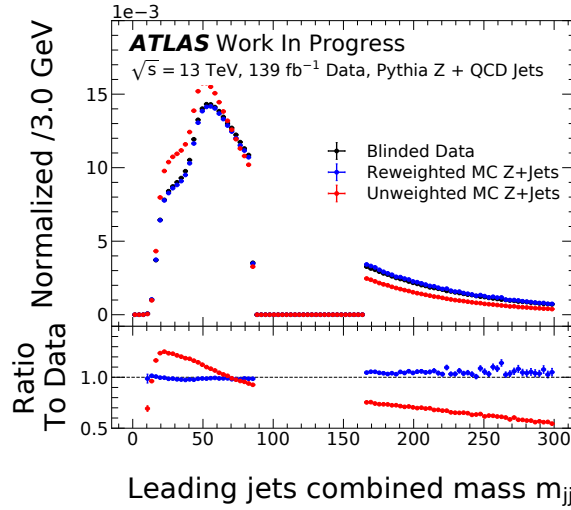


Figure A.3: Invariant mass of the two leading jets, m_{jj} . Distributions for data and Z + QCD jets Pythia MC background before and after variable- and high-dimensional reweighting. Events with m_{jj} inside the signal region (Fig 5.7a) were excluded during training of the reweighting classifier, and are excluded in this plot for blinding.

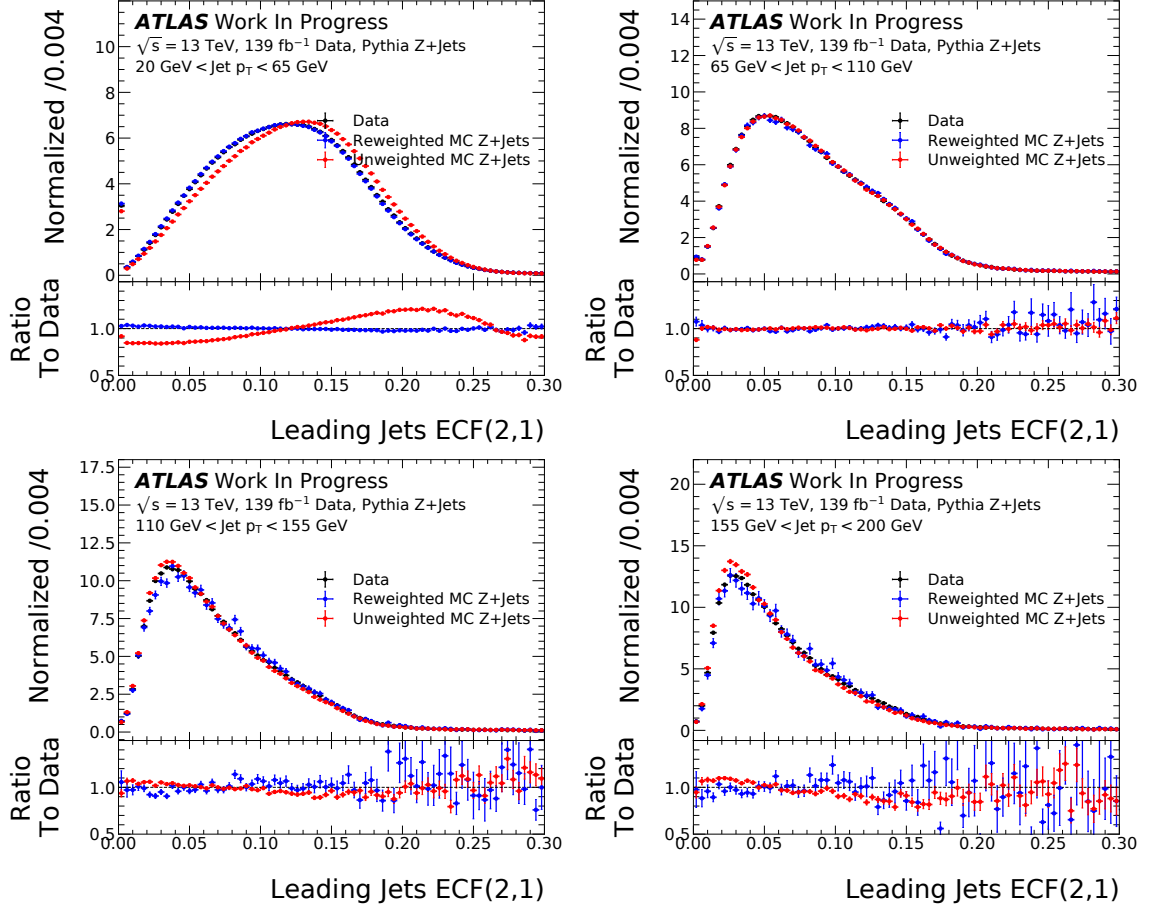


Figure A.4: Two point energy correlation $ECF(2,1)$. Distributions for data and Z + QCD jets Pythia MC background before and after variable- and high-dimensional reweighting. Events with m_{jj} inside the signal region (Fig 5.7a) were excluded during training of the reweighting classifier, and are excluded in this plot for blinding.

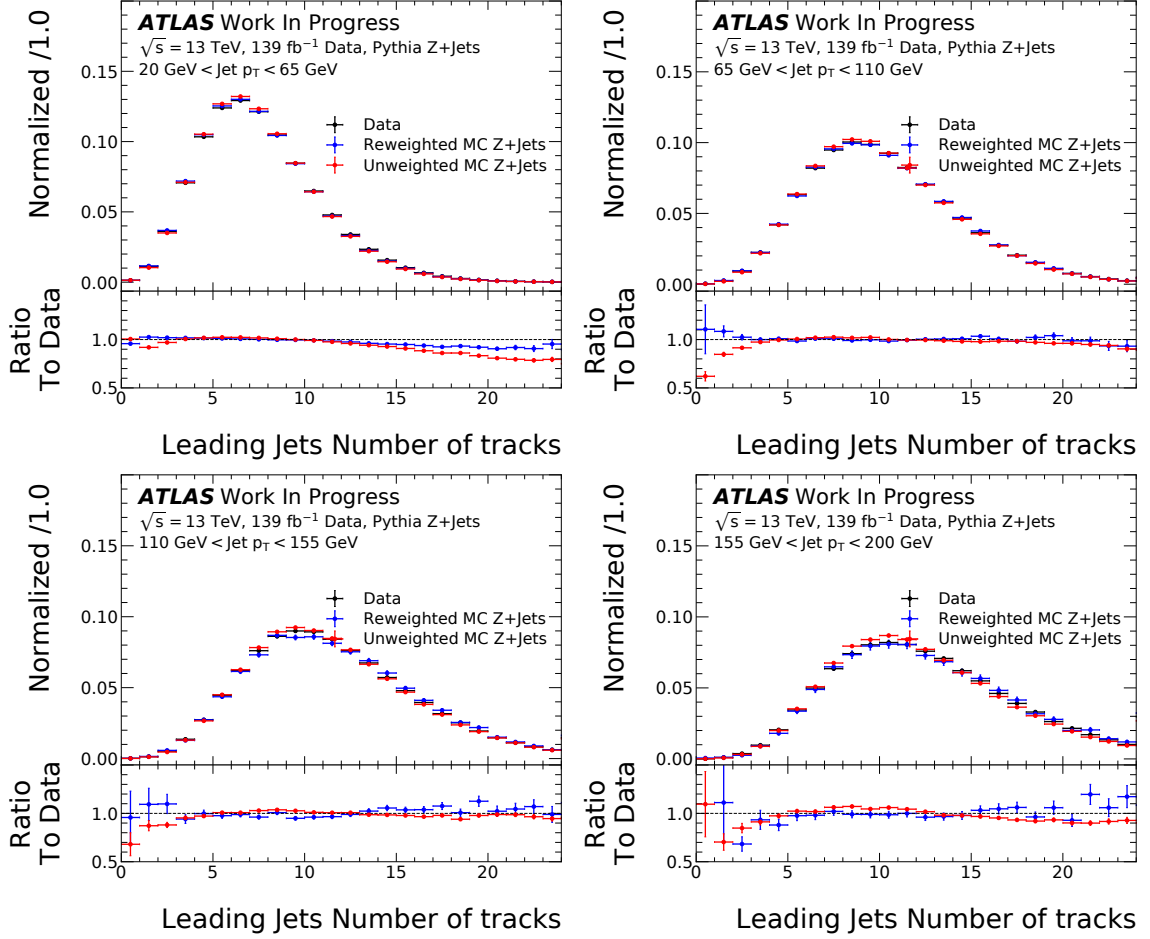


Figure A.5: Number of tracks in the jets. Distributions for data and Z + QCD jets Pythia MC background before and after variable- and high-dimensional reweighting. Events with m_{jj} inside the signal region (Fig 5.7a) were excluded during training of the reweighting classifier, and are excluded in this plot for blinding.

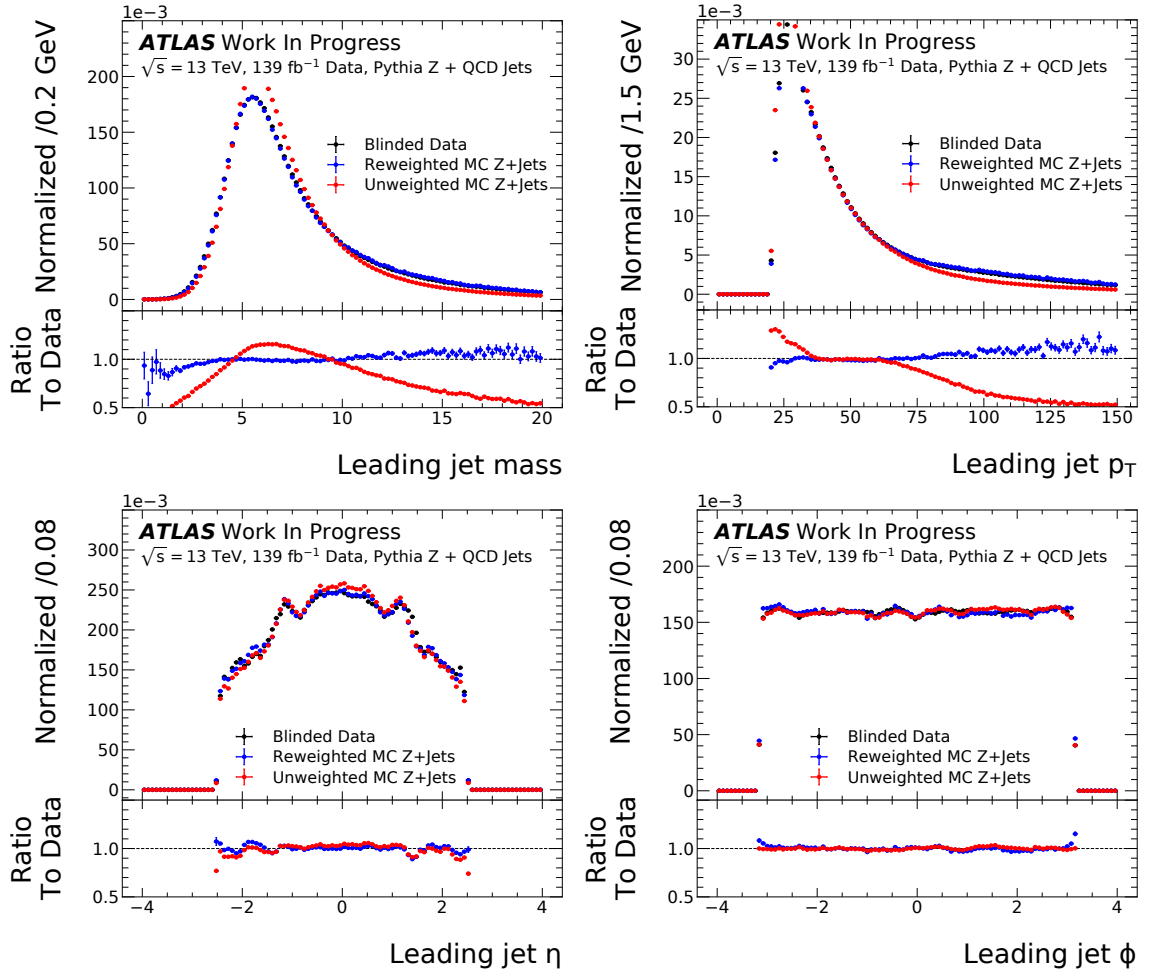


Figure A.6: Distributions for data and Z + QCD jets Pythia MC background before and after variable- and high-dimensional reweighting. Events with m_{jj} inside the signal region (Fig 5.7a) were excluded during training of the reweighting classifier, and are excluded in this plot for blinding. Here kinematics of the leading jet are shown.

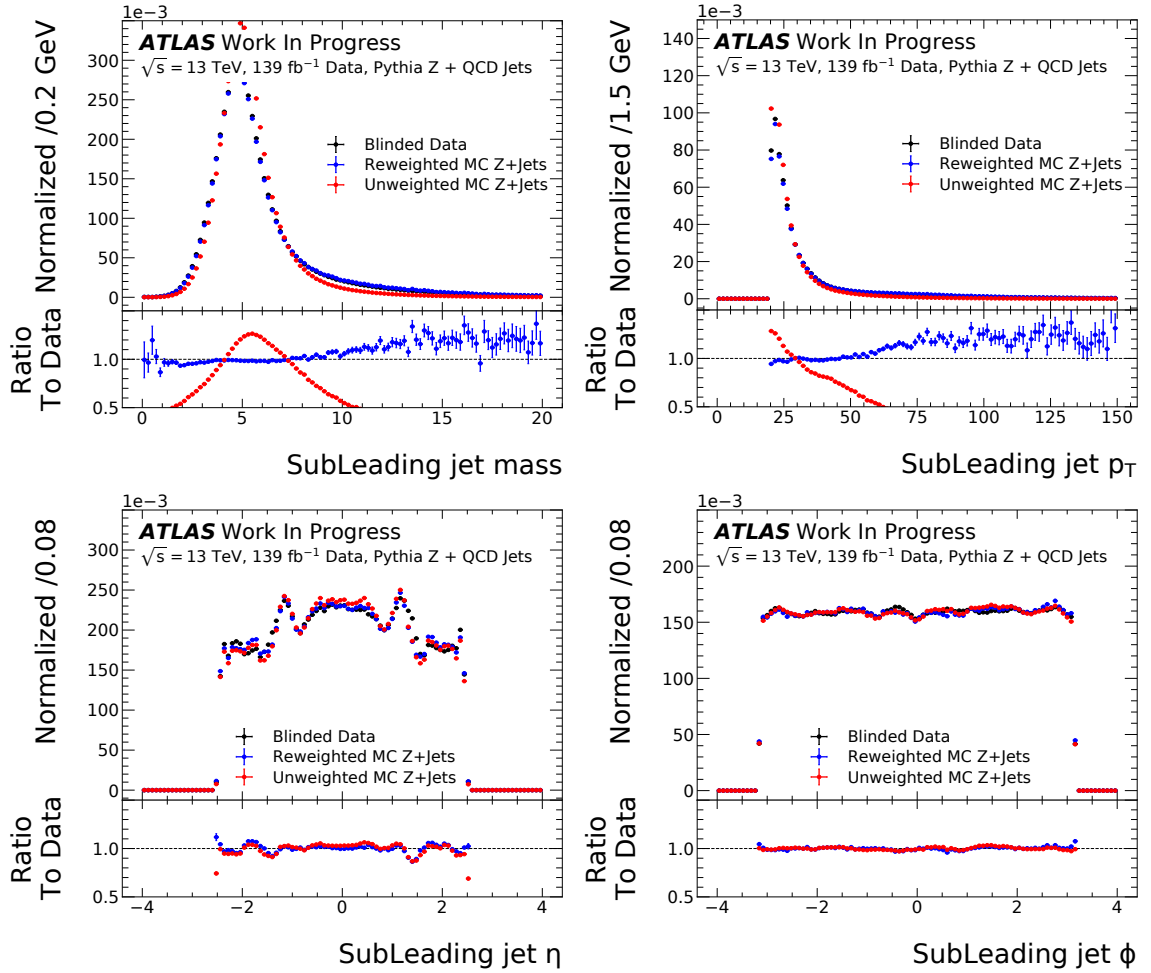


Figure A.7: Distributions for data and Z + QCD jets Pythia MC background before and after variable- and high-dimensional reweighting. Events with m_{jj} inside the signal region (Fig 5.7a) were excluded during training of the reweighting classifier, and are excluded in this plot for blinding. Here kinematics of the sub-leading jet are shown.

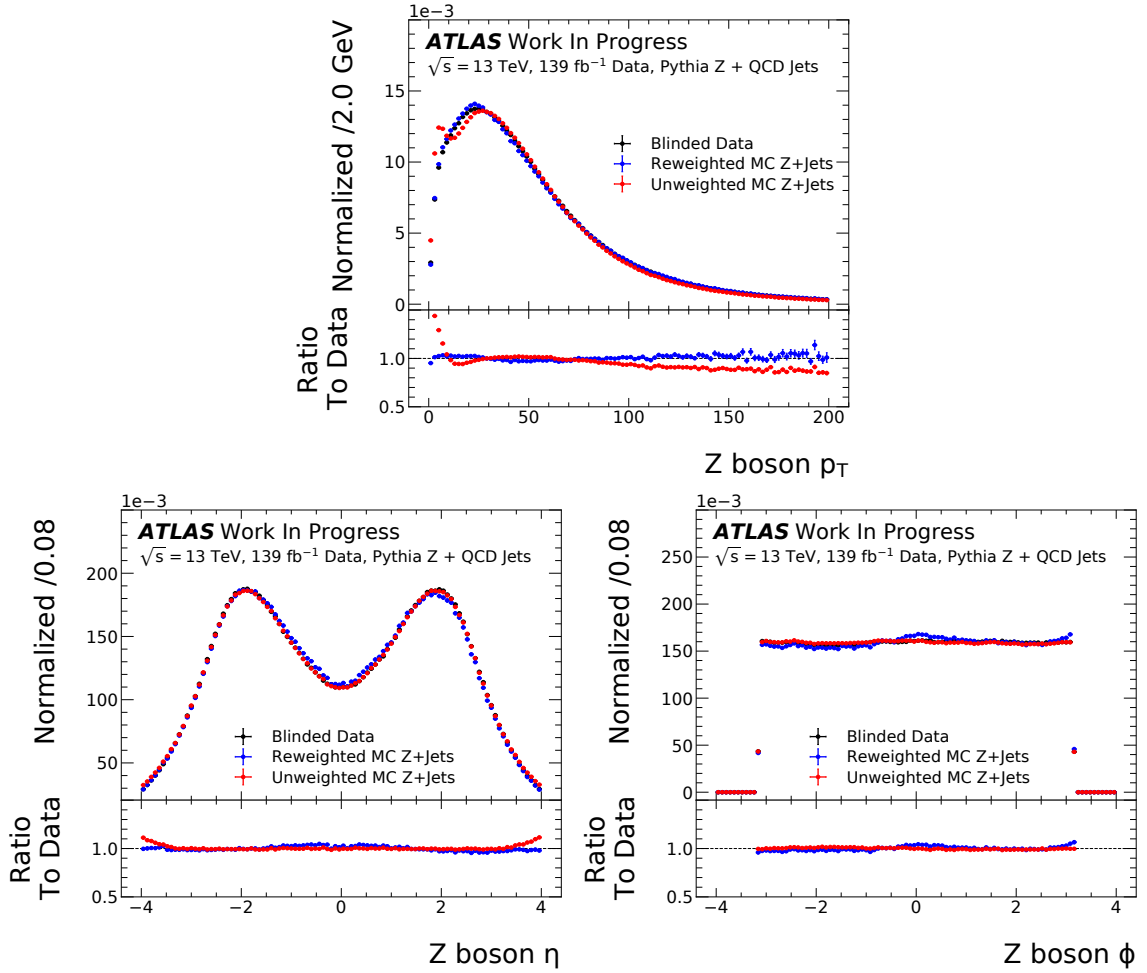


Figure A.8: Distributions for data and Z + QCD jets Pythia MC background before and after variable- and high-dimensional reweighting. Events with m_{jj} inside the signal region (Fig 5.7a) were excluded during training of the reweighting classifier, and are excluded in this plot for blinding. Here kinematics of the Z boson are shown, formed by adding the 4-vectors of the two compatible leptons in the event.

A.2.2 Sherpa Z + Jets Samples

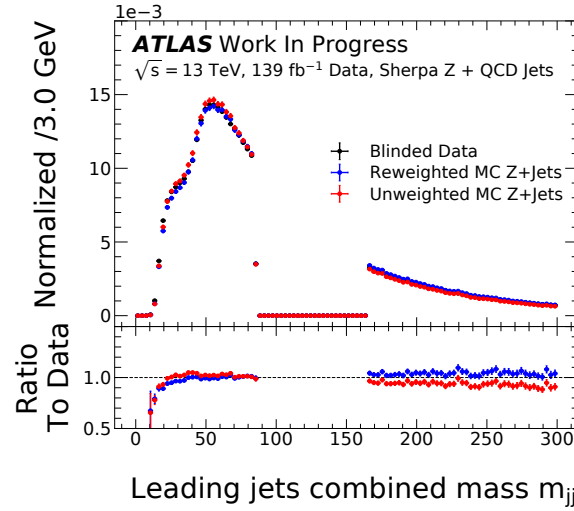


Figure A.9: Invariant mass of the two leading jets, m_{jj} . Distributions for data and Z + QCD jets Sherpa MC background before and after variable- and high-dimensional reweighting. Events with m_{jj} inside the signal region (Fig 5.7a) were excluded during training of the reweighting classifier, and are excluded in this plot for blinding.

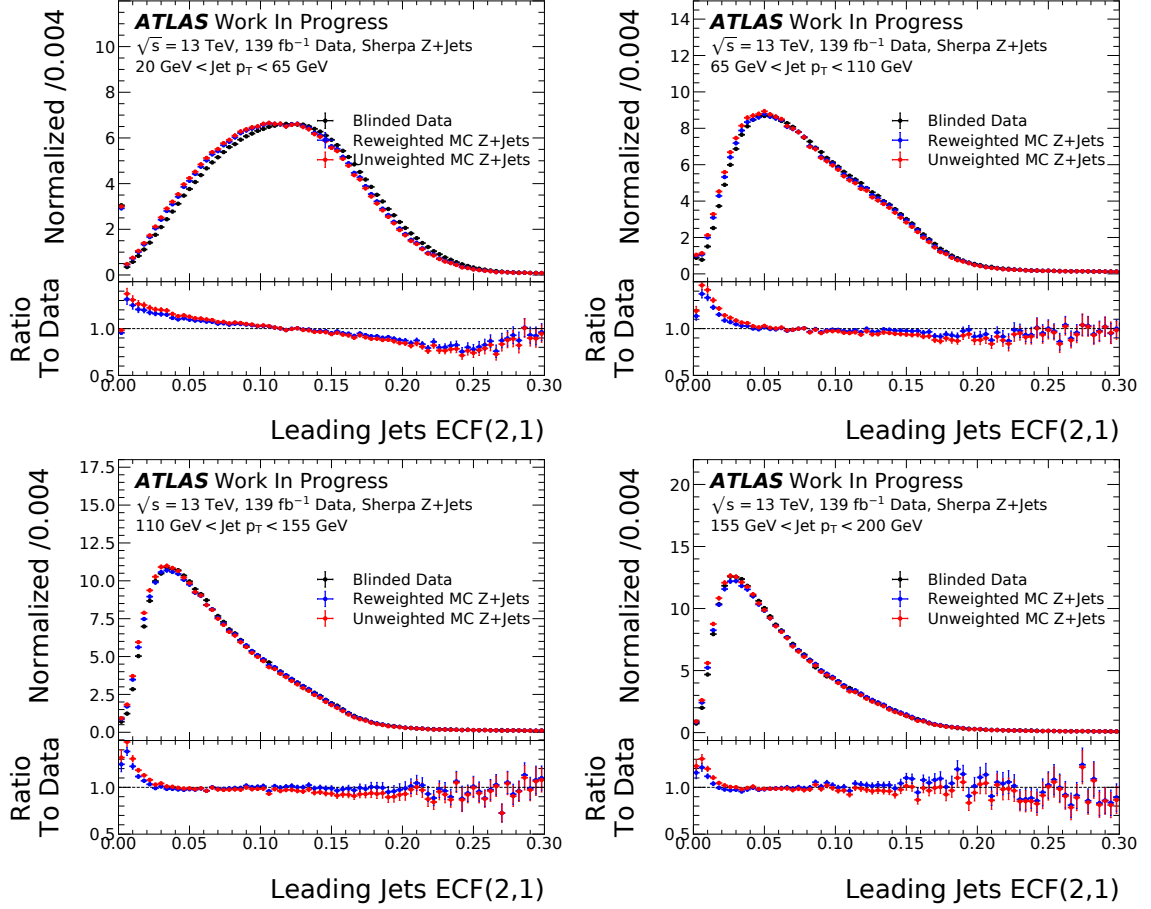


Figure A.10: Two point energy correlation $ECF(2,1)$. Distributions for data and Z + QCD jets Sherpa MC background before and after variable- and high-dimensional reweighting. Events with m_{jj} inside the signal region (Fig 5.7a) were excluded during training of the reweighting classifier, and are excluded in this plot for blinding.

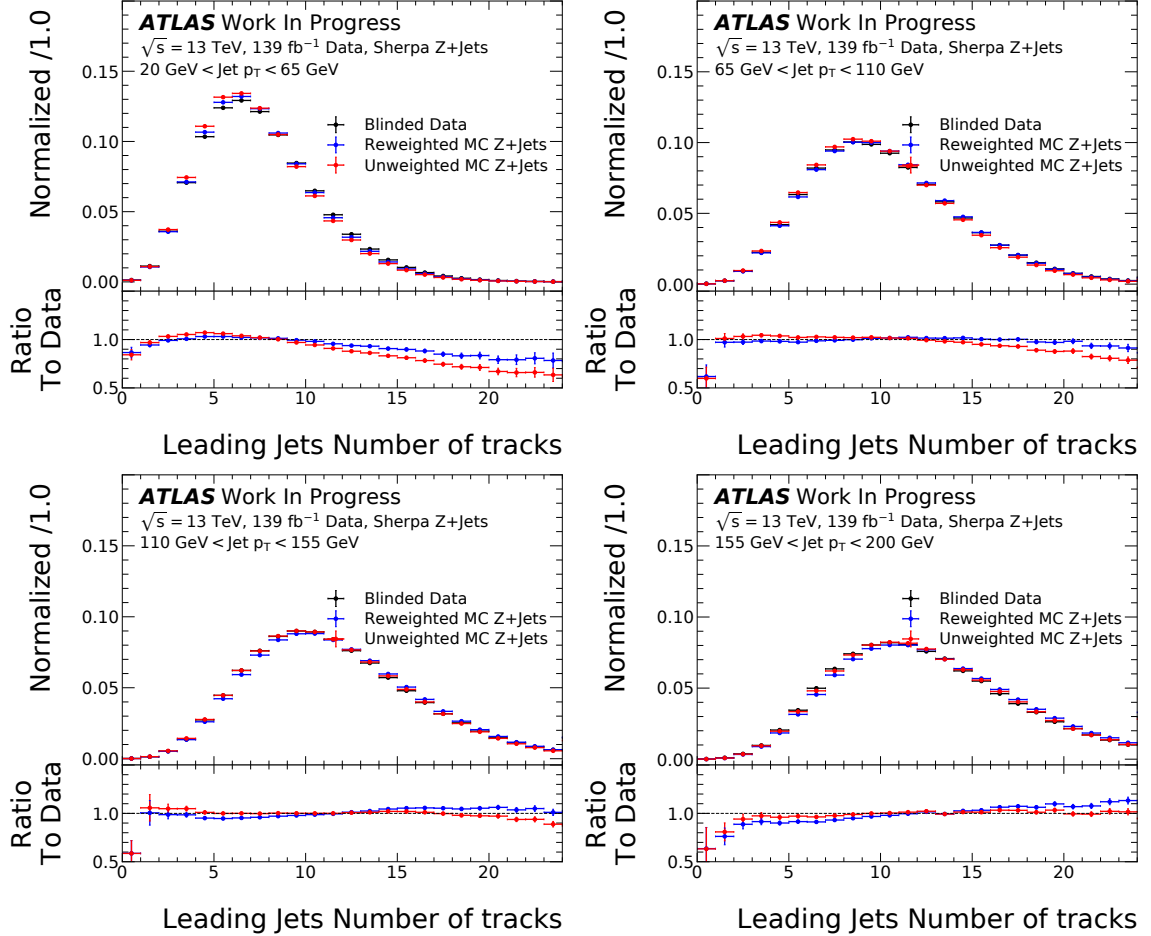


Figure A.11: Number of tracks in the jets. Distributions for data and Z + QCD jets Sherpa MC background before and after variable- and high-dimensional reweighting. Events with m_{jj} inside the signal region (Fig 5.7a) were excluded during training of the reweighting classifier, and are excluded in this plot for blinding.

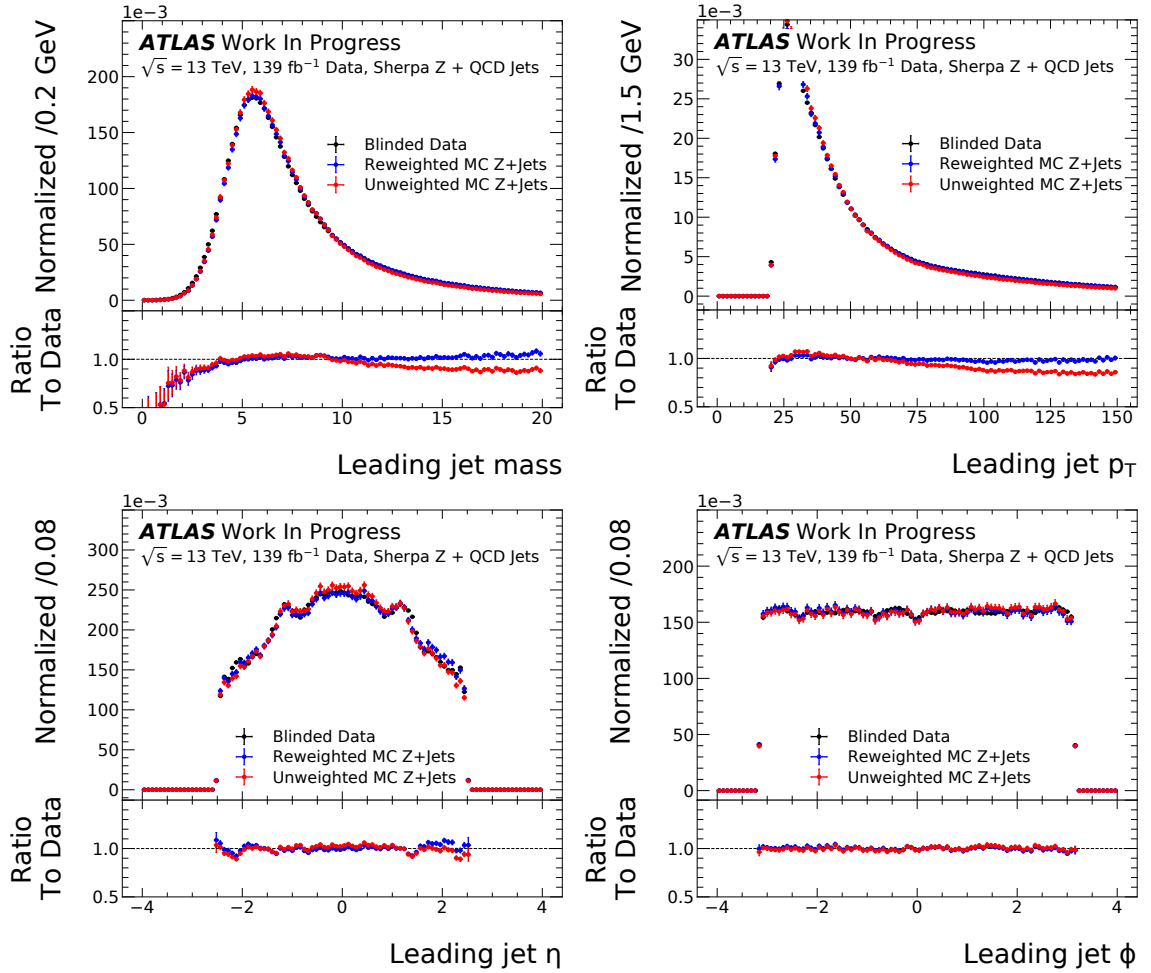


Figure A.12: Distributions for data and $Z + \text{QCD jets}$ Sherpa MC background before and after variable- and high-dimensional reweighting. Events with m_{jj} inside the signal region (Fig 5.7a) were excluded during training of the reweighting classifier, and are excluded in this plot for blinding. Here kinematics of the leading jet are shown.

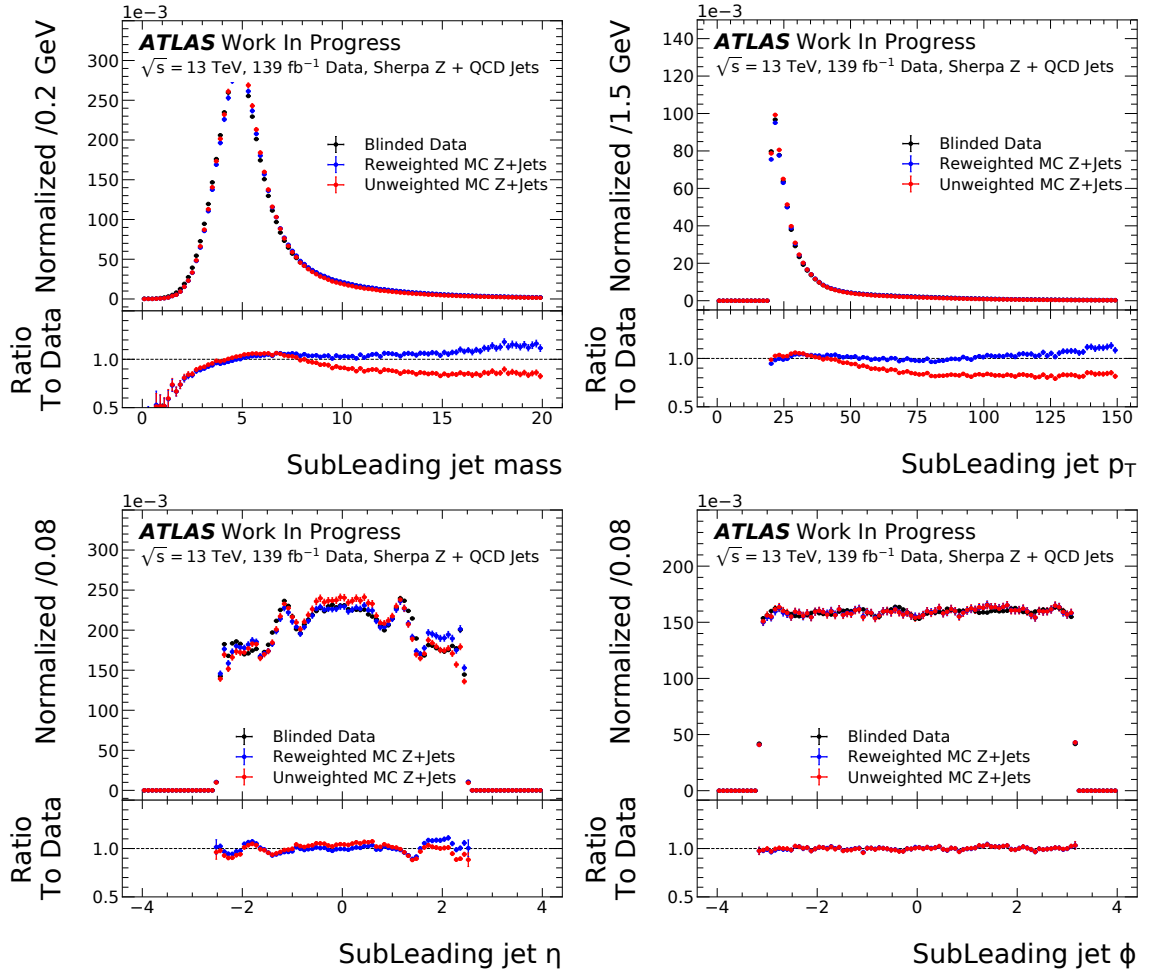


Figure A.13: Distributions for data and $Z + \text{QCD jets}$ Sherpa MC background before and after variable- and high-dimensional reweighting. Events with m_{jj} inside the signal region (Fig 5.7a) were excluded during training of the reweighting classifier, and are excluded in this plot for blinding. Here kinematics of the sub-leading jet are shown.

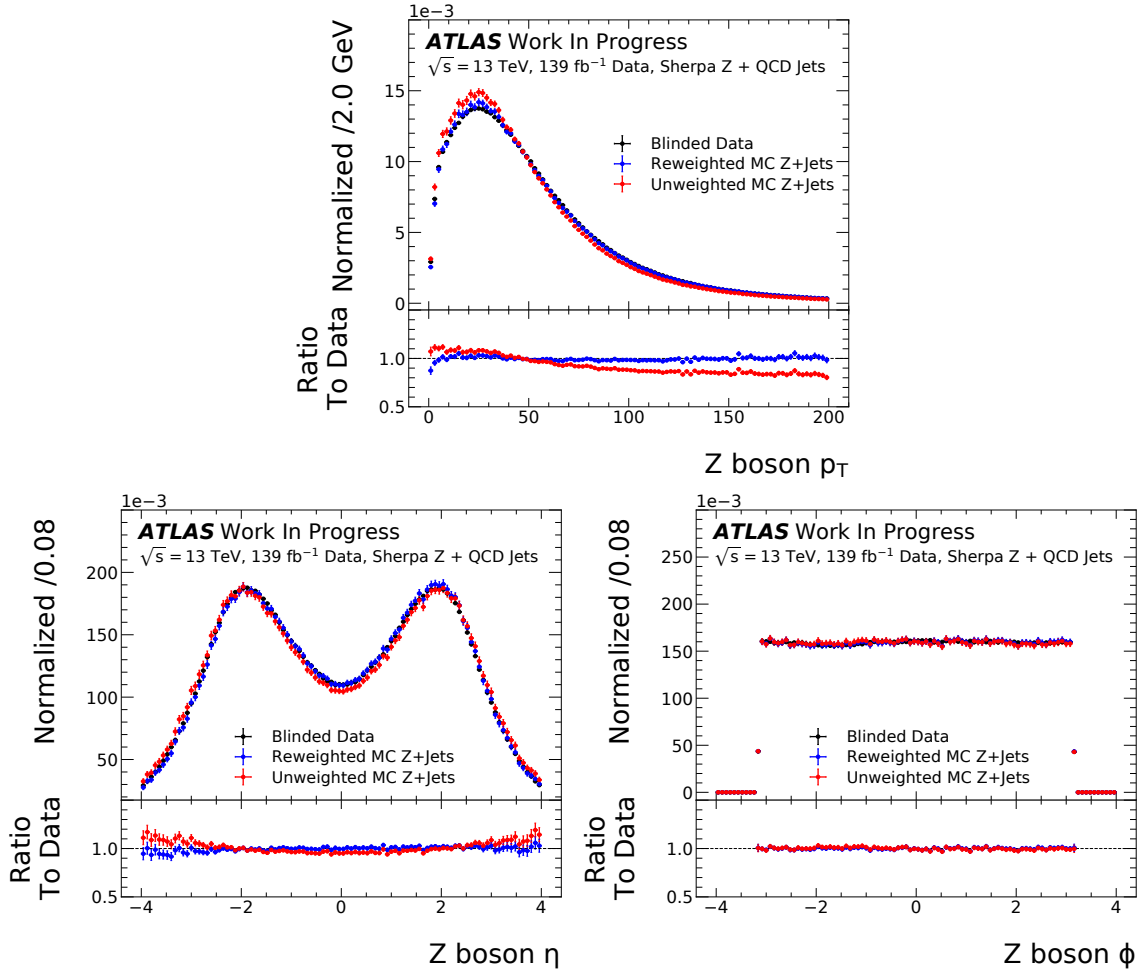


Figure A.14: Distributions for data and $Z + \text{QCD jets}$ Sherpa MC background before and after variable- and high-dimensional reweighting. Events with m_{jj} inside the signal region (Fig 5.7a) were excluded during training of the reweighting classifier, and are excluded in this plot for blinding. Here kinematics of the Z boson are shown, formed by adding the 4-vectors of the two compatible leptons in the event.

A.3 Inspecting closure and relative uncertainty on closures: Scanning DisCo λ Values

Table A.1: Pythia $Z + \text{QCD jet}$ events non-closure, along with the relative uncertainty on the Sherpa and Pythia $Z + \text{QCD jet}$ events for the selected working points, where the working point selection procedure is described in Section 5.4.4.2. Also shows the Pythia-Sherpa $Z + \text{QCD jet}$ events cross generator non-closure, along with the signal efficiency / square root of the background efficiency.

Scalar Mass [GeV]	λ	Signal efficiency / root Background efficiency	Pythia $Z + \text{QCD jet}$ events non-closure	Relative uncertainty on Pythia $Z + \text{jets}$ events closure	Relative uncertainty on Sherpa $Z + \text{jets}$ events closure	Pythia-Sherpa cross generator non-closure
0.5	0.02	31.34862	0.15913	0.23315	0.18437	0.54810
0.5	0.05	36.75652	0.22739	0.10976	0.24616	0.13619
0.5	0.1	31.52849	0.18529	0.24789	0.20199	0.74772
0.5	0.2	29.27676	0.20297	0.24311	0.18461	0.41230
0.5	1	30.81388	0.07037	0.23092	0.18417	0.46015
0.5	5	30.51597	0.11710	0.22505	0.23308	0.53966
0.5	10	51.48139	0.19988	0.10181	0.23721	0.26021
0.5	15	38.92430	0.22114	0.06227	0.23649	0.05688
0.5	20	24.89044	0.23375	0.03518	0.12297	0.02096
0.5	25	27.86632	0.23919	0.05675	0.23262	0.04975
0.5	50	32.65496	0.23367	0.15537	0.24024	0.14415
0.5	100	28.65627	0.22562	0.11283	0.24200	0.21661
0.5	125	25.70978	0.14163	0.06241	0.23862	0.03387
0.5	400	0.62400	0.19257	0.01202	0.04531	0.00917

Continuation of Table A.1						
1.0	0.02	32.81437	0.00913	0.24004	0.22759	0.57815
1.0	0.05	35.70603	0.15556	0.23182	0.23997	0.70805
1.0	0.1	34.64056	0.19281	0.23250	0.20006	0.68172
1.0	0.2	37.46765	0.11982	0.24972	0.24461	1.09970
1.0	1	36.65450	0.16933	0.22924	0.20515	0.86213
1.0	5	23.36622	0.22968	0.03150	0.09601	0.04020
1.0	10	34.98427	0.21105	0.05143	0.17383	0.08872
1.0	15	30.22498	0.24124	0.05211	0.19741	0.03754
1.0	20	50.91546	0.18116	0.10700	0.22880	0.49643
1.0	25	43.26605	0.16134	0.23060	0.23424	0.17014
1.0	50	30.18148	0.24151	0.05775	0.21468	0.25982
1.0	100	39.29202	0.06404	0.09406	0.24637	0.31529
1.0	125	33.40463	0.22631	0.09564	0.24381	0.18445
1.0	400	4.05814	0.08602	0.00737	0.03690	0.07464
2.0	0.02	24.23125	0.16448	0.14274	0.21854	0.17275
2.0	0.05	22.32424	0.06384	0.15369	0.23235	0.20895
2.0	0.1	24.30636	0.03014	0.16593	0.24165	0.39501
2.0	0.2	23.37149	0.07479	0.14883	0.24606	0.29167
2.0	1	21.37550	0.04631	0.17202	0.24186	0.34933
2.0	5	20.90921	0.04632	0.16362	0.23125	0.33908
2.0	10	21.83731	0.22256	0.03889	0.12782	0.04400
2.0	15	26.31008	0.21274	0.05075	0.16604	0.08144
2.0	20	26.15627	0.23757	0.05363	0.17229	0.10189
2.0	25	27.27974	0.22661	0.06055	0.19558	0.10411
2.0	50	21.74269	0.23298	0.18386	0.24586	0.47174

Continuation of Table A.1						
2.0	100	19.21120	0.10094	0.09026	0.24516	0.35665
2.0	125	21.06288	0.06879	0.11697	0.24870	0.30377
2.0	400	15.15520	0.22675	0.04371	0.18161	0.22398
4.0	0.02	10.39405	0.16199	0.13624	0.23998	0.05772
4.0	0.05	10.18152	0.10672	0.13192	0.24427	0.21438
4.0	0.1	11.07623	0.20250	0.16228	0.24598	0.27160
4.0	0.2	9.93719	0.14008	0.14367	0.18856	0.44696
4.0	1	9.73455	0.15181	0.16165	0.24747	0.38465
4.0	5	9.18066	0.24643	0.02036	0.04721	0.04985
4.0	10	11.09430	0.22548	0.03032	0.05935	0.01018
4.0	15	14.29781	0.23985	0.05838	0.12497	0.21344
4.0	20	15.01202	0.21499	0.05762	0.12223	0.15706
4.0	25	12.97501	0.24495	0.06806	0.14917	0.03765
4.0	50	14.19522	0.21900	0.10504	0.24646	0.10489
4.0	100	10.55747	0.20436	0.08506	0.22434	0.27036
4.0	125	10.52355	0.06895	0.11124	0.23789	0.09239
4.0	400	9.02648	0.24157	0.10236	0.23653	0.24686
8.0	0.02	6.08034	0.07601	0.16378	0.24742	0.00249
8.0	0.05	6.70101	0.11426	0.13401	0.22585	0.24120
8.0	0.1	6.09378	0.09103	0.13146	0.24519	0.18487
8.0	0.2	6.11895	0.18339	0.13530	0.24376	0.19400
8.0	1	6.29911	0.14152	0.15926	0.22675	0.17123
8.0	5	6.88520	0.24251	0.01696	0.04462	0.04214
8.0	10	6.46006	0.24939	0.02530	0.06153	0.06818
8.0	15	6.71964	0.23019	0.03178	0.08227	0.06663

Continuation of Table A.1						
8.0	20	7.12050	0.21444	0.04388	0.14067	0.06343
8.0	25	7.29050	0.24304	0.06543	0.16138	0.19519
8.0	50	5.57829	0.06792	0.14241	0.24095	0.32212
8.0	100	1.93609	0.02405	0.00300	0.01622	0.02440
8.0	125	2.05397	0.02121	0.00294	0.01624	0.01460
8.0	400	0.00000	1.54982	0.41093	1.02325	5.78079
16.0	0.02	5.14464	0.08285	0.08297	0.21402	0.14643
16.0	0.05	4.86919	0.14033	0.09779	0.24621	0.14954
16.0	0.1	5.05823	0.24513	0.08282	0.23431	0.04477
16.0	0.2	5.03098	0.19942	0.07087	0.21663	0.28965
16.0	1	4.96305	0.15633	0.09505	0.24451	0.30426
16.0	5	5.33465	0.24191	0.01039	0.05449	0.06147
16.0	10	5.10176	0.24191	0.02774	0.08981	0.00335
16.0	15	5.30305	0.10379	0.09806	0.22718	0.65423
16.0	20	3.21425	0.24876	0.01246	0.06113	0.00700
16.0	25	5.45567	0.04794	0.09734	0.23107	0.29795
16.0	50	3.18532	0.18348	0.06255	0.21791	0.21179
16.0	100	1.98427	0.01291	0.00284	0.01694	0.00163
16.0	125	1.78850	0.03947	0.00249	0.01490	0.02597
16.0	400	0.00000	0.56376	0.68165	nan	inf

A.4 Statistical Analysis

The analysis presented in Chapter 5 uses a frequentist single bin counting experiment (in the signal ABCD bin) to determine if the observed data is consistent with the Standard Model hypothesis, or with the BSM scenario of the $h \rightarrow aa \rightarrow gggg$ decays. The expected number of signal events (μ_s) in the signal region is determined directly by using the efficiency of the ABCD selection on the signal MC simulation samples (ϵ_s), described as,

$$\mu_s = \mathcal{L} \sigma_{Z(\mu\mu)H} \text{Br}(H \rightarrow aa) \text{Br}(a \rightarrow gg)^2 \epsilon_s$$

where $\mathcal{L}$ is the integrated luminosity, $\sigma_{Z(\mu\mu)H}$ is the cross section of $Z(\rightarrow \mu\mu)H$ production, and $\text{Br}(\cdot)$ refers to the corresponding branching ratio. The branching ratios can be folded into the parameter of interest for the statistical fit,

$$\mu = \text{Br}(H \rightarrow aa) \text{Br}(a \rightarrow gg)^2.$$

Therefore the expected number of signal events can be parameterized as $\mu_s = Q\mu$, where $Q = \mathcal{L} \sigma_{Z(\mu\mu)H} \epsilon_s$.

The expected number of background events (μ_b) in the signal region is determined from data, first by estimating the number of data events in the signal region (A), as $N'_A = N_B \times N_C / N_D$ where D is diametrically opposite to the signal region A. This N'_A is then multiplied by a corrective factor derived from the nominal MC simulation background sample that accounts for the non-closure in the ABCD bins, $N_A^{\text{MC bkgd}} / N_A'^{\text{MC bkgd}}$. Putting these factors together,

$$\mu_b = \frac{N_B^{\text{data}} \times N_C^{\text{data}}}{N_D^{\text{data}}} \left(\frac{N_A^{\text{MC bkgd}} \times N_D^{\text{MC bkgd}}}{N_B^{\text{MC bkgd}} \times N_C^{\text{MC bkgd}}} \right).$$

Given the expected number of signal and background counts in the signal regions, it is important to note that the Standard Model hypothesis is the background-only scenario, when $\mu = 0$.

The observed data events in the signal region, using the same selection criteria used

to define the ABCD regions in the MC simulation samples, is N_A^{data} . This observed count in the bin can be composed of signal and background events, requiring a fit to determine the parameter of interest (μ) that maximises the likelihood [425].

The full likelihood can be represented in terms of the parameter of interest (μ) and the nuisance parameters ($\bar{\theta}$), where the nuisance parameters are required to encapsulate the various sources of uncertainties that affect the predictions of the signal and background counts in the bin. The nuisance parameters are constrained by auxiliary measurements interpreted as one standard deviation variations from the nominal estimates — These are represented as Gaussian constraints and account for the various sources of systematic uncertainties. Additionally, Poisson constraints are also added to account for the statistical uncertainties on the predictions. Thus, the likelihood can be represented, in a simplified manner hiding the differences in the sources of systematics between the signal and background predictions, as follows

$$\mathcal{L}(\mu, \bar{\theta}) = \mathcal{P}(N_A^{\text{data}} | \mu_s(\mu, \bar{\theta}) + \mu_b(\bar{\theta})) \prod_i \mathcal{X}_i(\tilde{\theta}_i | \theta_i)$$

where $\mathcal{P}$ is the Poisson function, and the constraints $\mathcal{X}_i$ are either Poisson or Gaussian constraints based on whether the nuisance parameter θ_i is a statistical or systematic uncertainty respectively.

These likelihoods are built using the *pyhf* [426, 427] pythonic implementation of the HistFactory framework [428], and maximum likelihood fits can be performed with the parameters unconstrained or some of them constrained. The profile likelihood that describes the consistency of the data with some value μ of the parameter of interest is,

$$\lambda(\mu) = \begin{cases} \frac{\mathcal{L}(\mu, \hat{\hat{\theta}}(\mu))}{\mathcal{L}(\hat{\mu}, \hat{\hat{\theta}})} & \text{if } \hat{\mu} \geq 0 \\ \frac{\mathcal{L}(\mu, \hat{\hat{\theta}}(\mu))}{\mathcal{L}(0, \hat{\hat{\theta}}(0))} & \text{if } \hat{\mu} < 0. \end{cases}$$

Here $(\hat{\mu}, \hat{\hat{\theta}})$ correspond to the unconstrained fit to the likelihood, whereas $\hat{\hat{\theta}}$ is the fit to the nuisance parameters for a fixed hypothesis μ . This profile likelihood is used to form the test statistic $\tilde{q}_\mu$ that applies for the case where μ is required to be

non-negative,

$$\tilde{q}_\mu = \begin{cases} -2 \ln \lambda(\mu) & \text{if } \hat{\mu} \leq \mu \\ 0 & \text{if } \hat{\mu} > \mu. \end{cases}$$

Given the test statistic, the p-values for the background-only hypothesis (p_0) and the signal+background hypothesis (p_μ) can be determined as follows,

$$p_\mu = \int_{\tilde{q}_\mu^{\text{obs}}}^{\text{inf}} f(\tilde{q}_\mu | \mu, \bar{\theta}(\mu)) d\tilde{q}_\mu$$

$$p_0 = 1 - \int_{\tilde{q}_\mu^{\text{obs}}}^{\text{inf}} f(\tilde{q}_\mu | 0, \bar{\theta}(0)) d\tilde{q}_\mu.$$

Here $f(\cdot)$ is the probability distribution of the test statistic, which is a non-central chi-squared distribution in the large sample limit, and $\tilde{q}_\mu^{\text{obs}}$ is the observed value of the test statistic for the given parameter value μ . Note here that p_0 and p_μ are defined as per convention, and can be interpreted as the probability to correctly interpret the result as being consistent with the background-only hypothesis or as being consistent with the signal+background hypothesis, respectively.

Besides knowing the best fit values, confidence intervals on the value of μ can be established using the p-values, using the *CLs* method [429], wherein hypotheses μ are rejected if the confidence level is below some threshold α (typically 5%):

$$CL_s = \frac{p_\mu}{1 - p_0} < \alpha.$$

The expected limits and derived prior to un-blinding using an Asimov dataset [430] created using the background-only component where all the parameters are zeroed. Figure 5.20 shows the expected limits for the $h \rightarrow aa \rightarrow gggg$ analysis made using the background-only hypothesis as per the Asimov dataset.

A.5 Expected Limits on Signal Strength: Scanning DisCo λ Values

The degeneracy of the λ hyperparameter results in similar limits are set across a range of λ values. The hyperparameter scan is also susceptible to neural network stochasticity from random initializations during training which results in some scans failing to learn as effectively as possible, resulting in working points translating into diverging limits from the neighboring λ values. This means optimizing over λ for the best working points does not necessarily result in a value of λ that is optimal for a given mass point, rather it points to the best trained network among a family of networks that would otherwise have comparable performance.

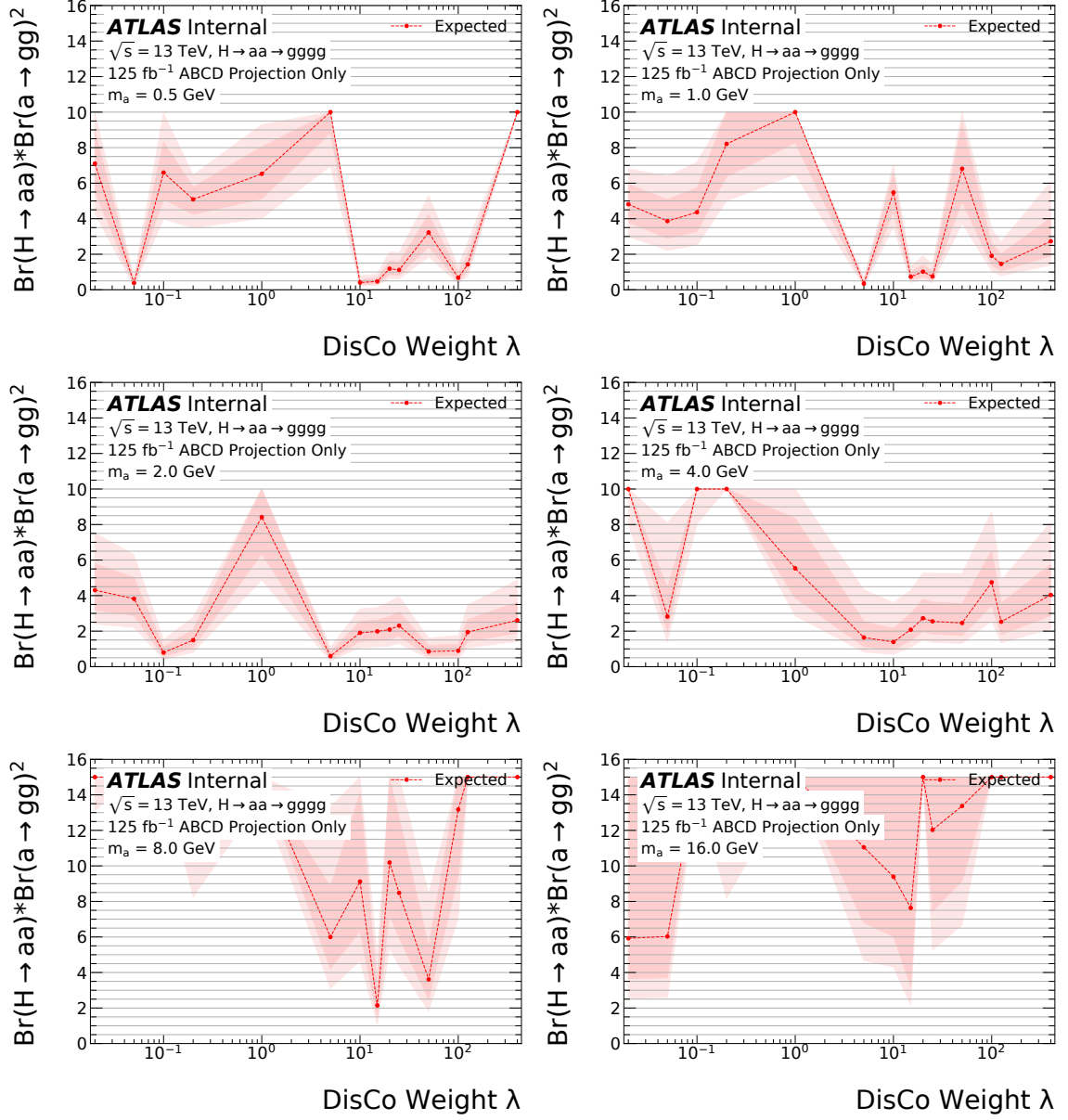


Figure A.15: Expected Limits. For each signal mass point considered, the expected limits from the single bin counting experiment are derived for every value of the DisCo weight hyperparameter λ tested.

A.6 Impact and Pull Plots From the Fits During Limit Setting

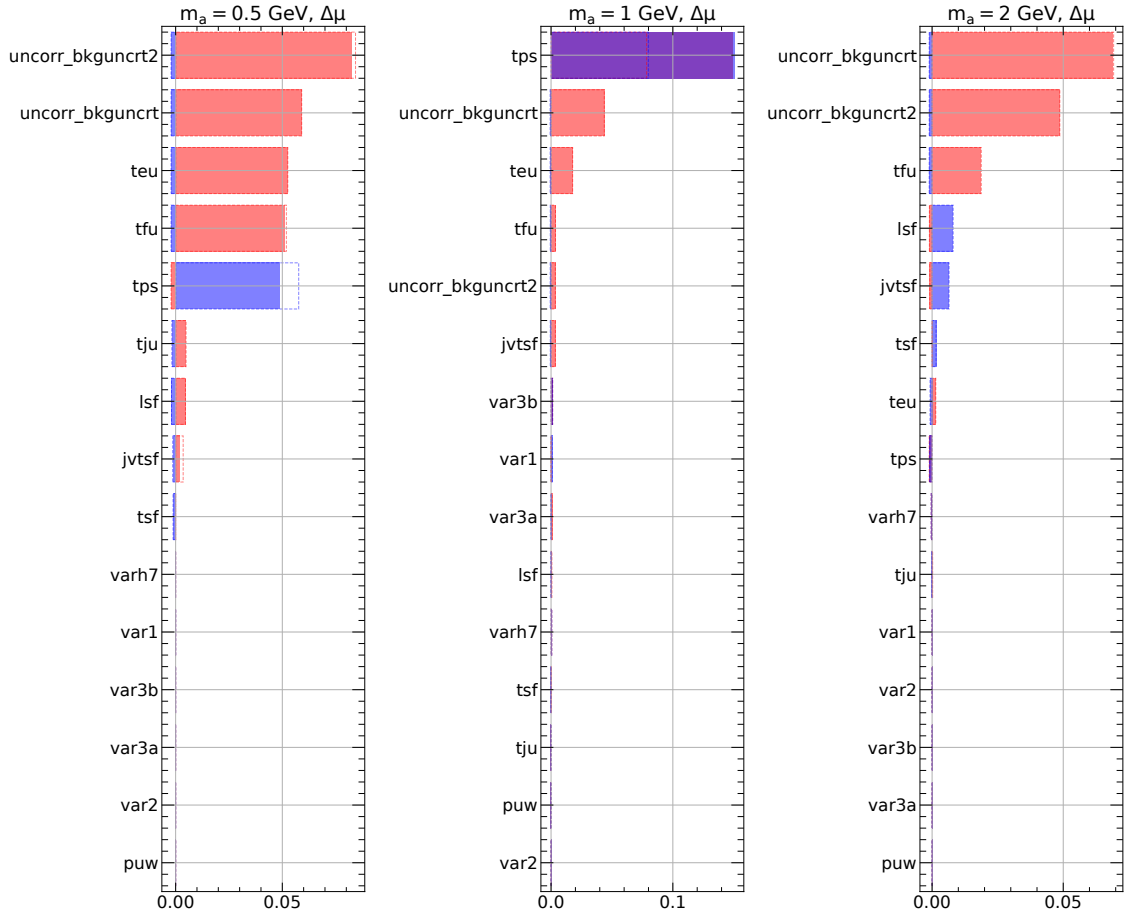


Figure A.16: Impact Plots. For each signal mass point considered, the impact of each modifier is individually determined by fixing its value to its best fit value plus/minus its fit error.

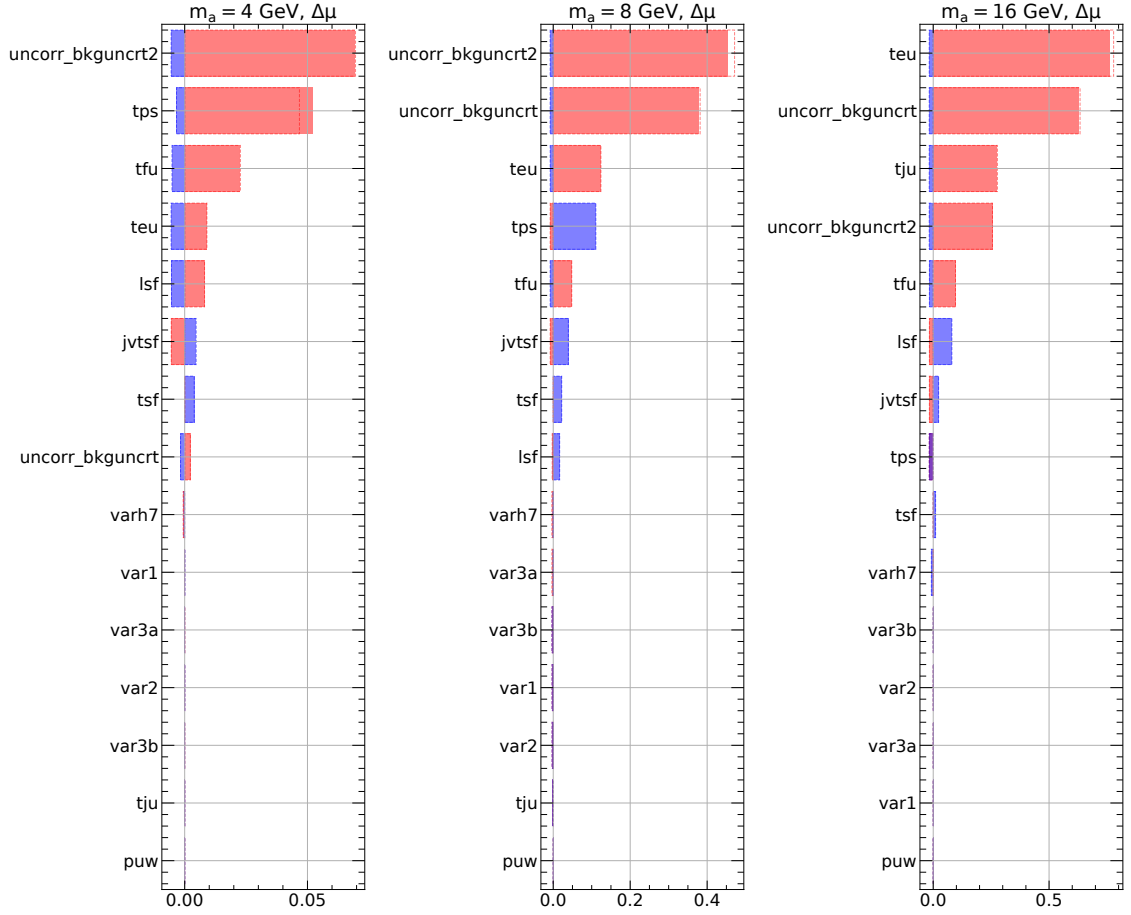


Figure A.17: Impact Plots. For each signal mass point considered, the impact of each modifier is individually determined by fixing its value to its best fit value plus/minus its fit error.

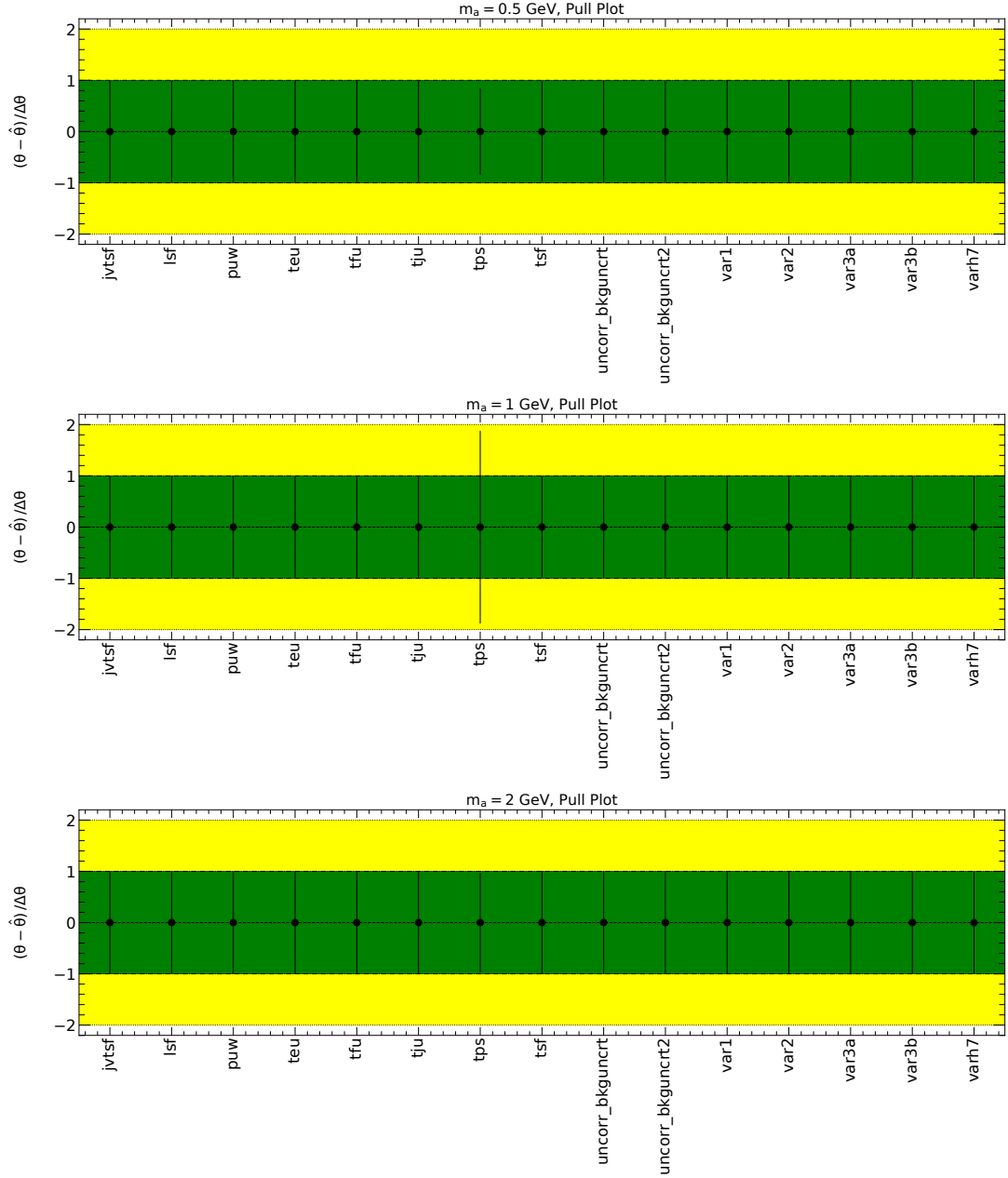


Figure A.18: Pull Plots. For each signal mass point considered, the pull plots for each modifier.

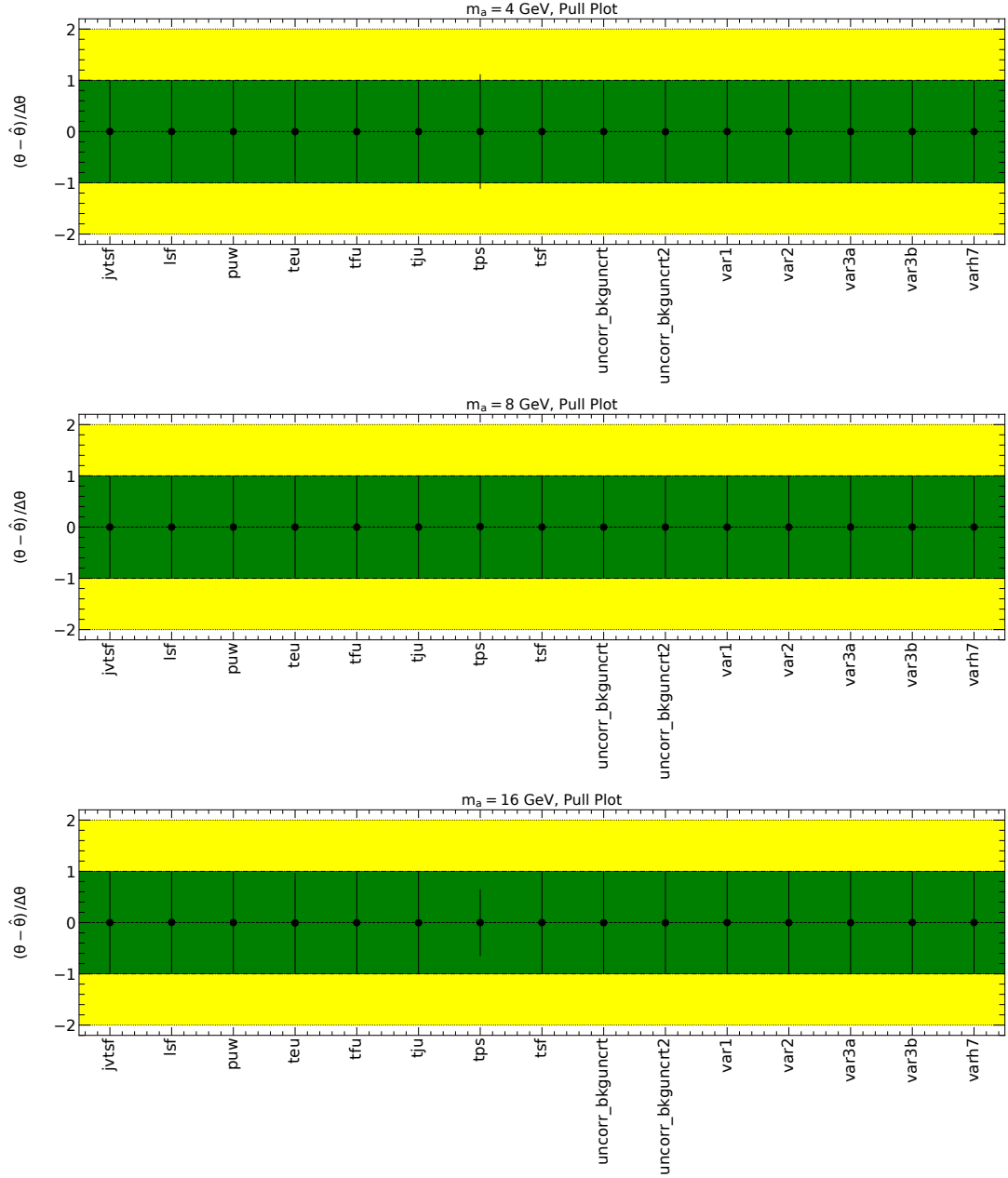


Figure A.19: Pull Plots. For each signal mass point considered, the pull plots for each modifier.

Appendix B

Light Field Imaging for Cold Atom Clouds in the Matter-wave Atomic Gradiometer Interferometric Sensor: Appendix

B.1 Entocentric imaging - A geometric perspective

A brief review of entocentric image formation is provided here from a purely geometric point of view. Using an ideal thin converging lens as the focusing optic one can visualize the key image formation properties that are relevant to the work presented in Chapter 8. Figure B.1 shows ideal image formation for such a simple optic of focal length f , where the marginal rays are traced through the system for point sources located at distances u , F' , and N' from the lens. A point source at the nominal object distance of u from the lens is imaged by the lens to a point at a distance v from the lens on the image plane side of the optic. The point at object distance F' (N') is focused closer (further) to the lens, hence results in a defocused image when imaged at the image distance of v from the lens.

Using the lensmaker's equation [405] in the thin lens paraxial approximation, the object distance F' (N') can be determined such that the spot size (circle of confusion)

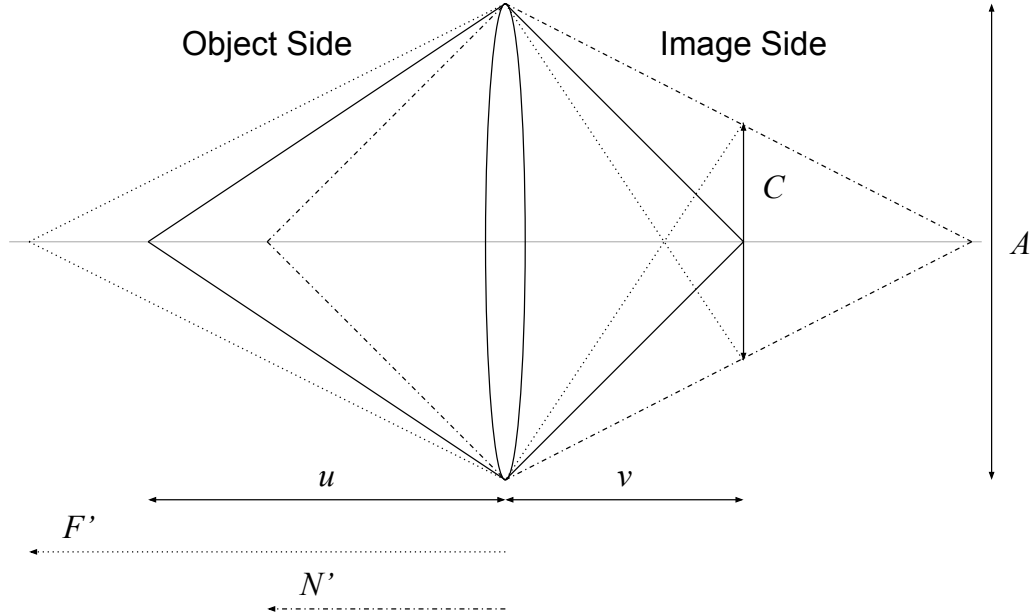


Figure B.1: Ray diagram for an ideal thin converging lens of aperture diameter A . Only the marginal rays are shown, where each line-style corresponds to rays starting from a unique object distance along the optical axis. The rays from object distance u are focused down to a point at image distance v . Rays starting at object distances F' (N') converge at a point closer (further) to the lens than v . In these cases, light spreads across the spot of diameter C when looking at the plane at distance v from the lens. This spot size corresponds to the geometric circle of confusion for the lens, and $F' - N'$ is the geometric depth of field of the lens about the nominal object distance u . Figure sourced from [41].

diameter at the plane of focus v is equal to C . Note the convention here that all distances measured from the lens are positive. This depends on the lens aperture (diameter = A as shown) and can be solved for using the similarity of triangles, resulting in:

$$F' = \frac{fuA}{fA - C(u - f)}, \quad N' = \frac{fuA}{fA + C(u - f)}$$

and

$$\text{DoF} = F' - N' = \frac{2fuAC(u-f)}{(fA)^2 - C^2(u-f)^2} \quad (\text{B.1})$$

$$\approx 2CN \left(1 + \frac{1}{m}\right)^2 \quad (\text{B.2})$$

where DoF is the geometric depth of field, and is the range of object distances over which the circle of confusion resulting from the geometric defocus of the lens is less than some value C . The final approximation is made in the limit that the circle of confusion C is significantly less than u , f , and A , and using the lens f-number $N = f/A$ and the lens magnification $m = f/(u-f)$ where positive m corresponds to an inverted image. In this limit, DoF is independent of the lens focal length and can be interpreted as function of N and m for a given C .

This simple model can also provide a geometric estimate of the total light collected from an isotropically fluorescing point source that is being focused by the lens. Assuming that the lens shown in Figure B.1 has uniform light collection over its aperture of diameter A , the fraction of light collected by the lens is (also a function of N and m),

$$\frac{1}{4\pi} \times 2\pi \left(1 - \cos \left(2 \arctan \frac{A}{2u}\right)\right) = \frac{1}{2} \left(1 - \cos \left(2 \arctan \frac{m}{2N(1+m)}\right)\right) \quad (\text{B.3})$$

B.2 Normal angles for a folded view

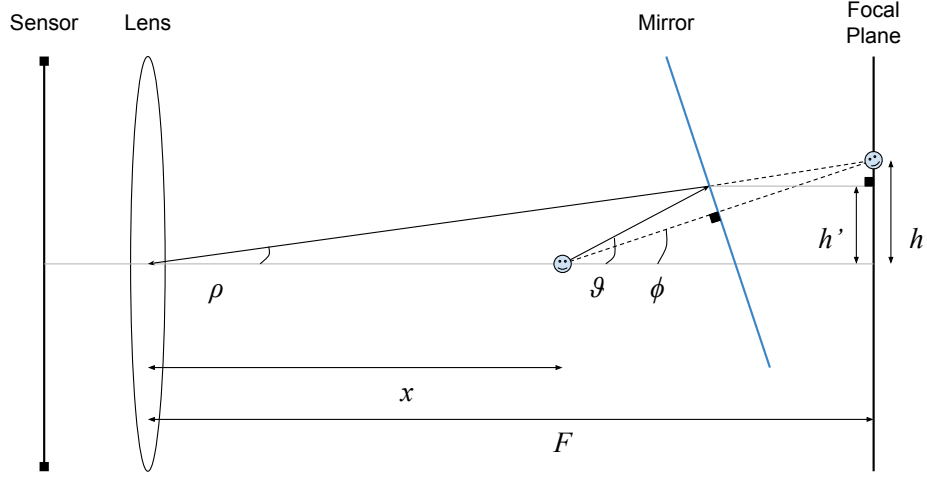


Figure B.2: Principal ray diagram of a single folded view. Here the object (smiley face) is placed at a distance x from the lens center, offset from the focal plane of the lens which is at a distance F from the lens. A flat mirror is introduced between the object and the focal plane, which produces a virtual object that lies on the focal plane; this virtual object is focused onto the sensor by the lens. The mirror normal angle ϕ is chosen such that the principal (central) ray emanating from the object (at angle θ) is reflected towards the center of the lens. This constrains the geometry of the mirror and determines the mirror normal angle as described in Appendix B.2. Note that the $\theta > 0$ view angle results in an image that sees the object as if it were rotated by angle θ about the optical axis. Note also that $\rho = 2\phi - \theta$. Figure sourced from [41].

Figure B.2 is used as the reference for the following equations. We start by requiring that the height of the point of reflection h' is consistent when derived from the lens center perspective and from the object's perspective. Note that $r = F/x$ is used for simplification, along with the fact that $\rho = 2\phi - \theta$ which is shown

geometrically in Figure B.2.

$$\begin{aligned}
\frac{\frac{F-x}{2\cos\phi}\sin\theta}{\cos(\theta-\phi)} &= h' \\
h' &= x \left[\frac{1}{\tan\rho} - \frac{1}{\tan\theta} \right]^{-1} \\
\frac{(F-x)\sin\theta}{2\cos\phi\cos(\theta-\phi)} &= \frac{x\tan\theta\tan\rho}{\tan\theta - \tan\rho} \\
\frac{(r-1)\sin\theta}{\cos\theta + \cos\rho} &= \frac{\tan\theta\tan\rho}{\tan\theta - \tan\rho} \\
(r-1)\sin\theta(\tan\theta - \tan\rho) &= \tan\rho\sin\theta + \sin\rho\tan\theta \\
(r-1)\sin\theta\tan\theta &= \sin\rho \left[\frac{\sin\theta}{\cos\rho} + \frac{\sin\theta(r-1)}{\cos\rho} + \tan\theta \right] \\
(r-1)\tan\theta &= \sin\rho \left[\frac{r}{\cos\rho} + \frac{1}{\cos\theta} \right] \\
\text{Squaring, } (r-1)^2\sin^2\theta\cos^2\rho &= r^2\cos^2\theta + \cos^2\rho + 2r\cos\theta\cos\rho \\
&\quad - r^2\cos^2\theta\cos^2\rho - \cos^4\rho - 2r\cos\theta\cos^3\rho \\
\text{Rearranging, } 0 &= \cos^4\rho(-1) - \cos^3\rho(2r\cos\theta) \\
&\quad + \cos^2\rho[1 - r^2\cos^2\theta - (r-1)^2\sin^2\theta] \\
&\quad + \cos\rho(2r\cos\theta) + (r^2\cos^2\theta)
\end{aligned} \tag{B.4}$$

This equation is solved numerically for $\cos\rho$ for specific values of r and θ . Note that solving for ρ fully constrains the mirror parameters ϕ and h' .

Alternatively, the system can be geometrically solved using the properties of ellipses. Figure B.3 shows the construction that be used to find the mirror normal angles and mirror center positions. For any point on the plane of focus where a virtual image of the object is to be formed, a unique circular directrix from the family of ellipses will pass through that point. The point of intersection of the line between the lens center and that point on the plane of focus with the ellipse identifies the location of the corresponding mirror center. The normal to the ellipse at the intersection point corresponds to the normal angle of the mirror. This procedure also determines the mirror parameters given a non-flat plane of focus, as it is identically applicable.

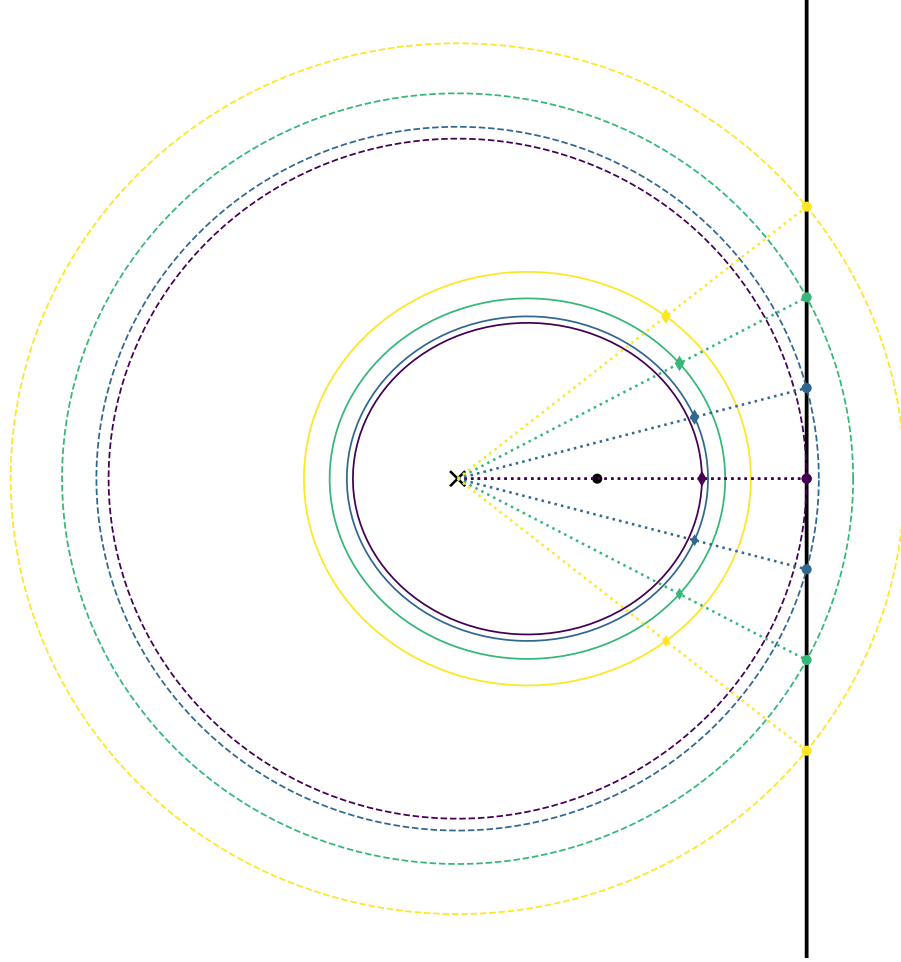


Figure B.3: Image formation geometry for a single folded view like Figure B.2. Here the black X marks the location of the lens center and the black dot the location of the object. The thick solid black vertical line on the right corresponds to the plane at which the lens is focused. The solid colored ellipses represent the family of ellipses that have the lens center and object as the two focii. The dashed circles are the corresponding circular directrices of the ellipses, centered at the lens center. The colored diamonds show the mirror center positions and the colored dots show the corresponding virtual object locations at the plane of focus. The dotted lines connect the lens center, mirror centers, and virtual object locations. Figure sourced from [41].

B.3 Field of view for a folded view

Figure 8.6 is used as the reference for the following equations. We start by finding the mirror radius projected onto the perpendicular to the central ray direction, l . Then, the distance between the lens center and the center of the mirror, D , can be found. These two quantities are then combined to determine χ , which is used to determine the field of view, FoV.

$$\begin{aligned}
 l &= R \cos(\theta - \phi) \\
 D &= \frac{h'}{\sin \rho} = \frac{(F - x) \sin \theta}{2 \cos \phi \sin \rho \cos(\theta - \phi)} \\
 \tan \chi &= \frac{l}{D} = \frac{2R \cos^2(\theta - \phi) \cos \phi \sin \rho}{(F - x) \sin \theta} \\
 \text{FoV} &= F [\tan(\rho + \chi) - \tan(\rho - \chi)]
 \end{aligned} \tag{B.5}$$

The variation of the field of view with the view angle θ for a particular set of parameters is shown in Figure B.4. There are two competing effects that define the field of view of the system. Firstly, the mirror gets increasingly inclined towards the optical axis as view angle gets larger, and this results in a reduced field of view footprint due to a $\cos(\theta - \phi)$ factor. Secondly, the mirrors move further away from the axis as the view angle gets larger to satisfy the constraint for the virtual objects to land on the focal plane - this results in a larger field of view footprint. Therefore, the field of view variation as shown in Figure B.4 depicts a minima and a diverging field of view as θ approaches 90° . The mirror sizes can be chosen based on the field of view to account for the object of interest. Hence, the system is flexible in its ability to accommodate objects of various scales.

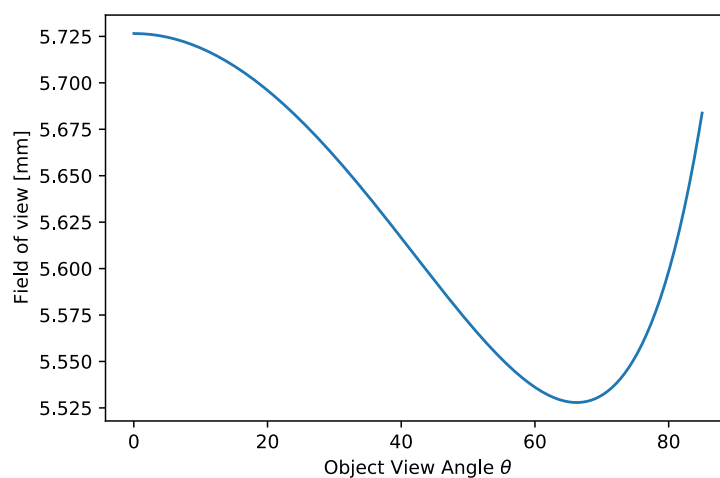


Figure B.4: Field of view as a function of the view angle θ determined by solving Eq. B.5. For this plot, an ideal lens of focal length 62.275 mm operating at a magnification of $m = 0.215$ is chosen, where the object distance is picked to allow the 55° view to lie at the edge of a full frame sensor. This is the same setup as described in section 8.3.2.2 and the object distance is picked following Eq. 8.2. The mirror here is a flat mirror of diameter 5 mm centered appropriately to satisfy the constraints on the principal rays. Figure sourced from [41].

B.4 Generated Image Grids

Figure B.5 shows the full grid of 77 views generated from the NeuS model, corresponding to the 77 views used for training shown in Figure 8.13. As discussed in Figure 8.14, the real and generated views are visually nearly identical, demonstrating reasonable closure of the model.

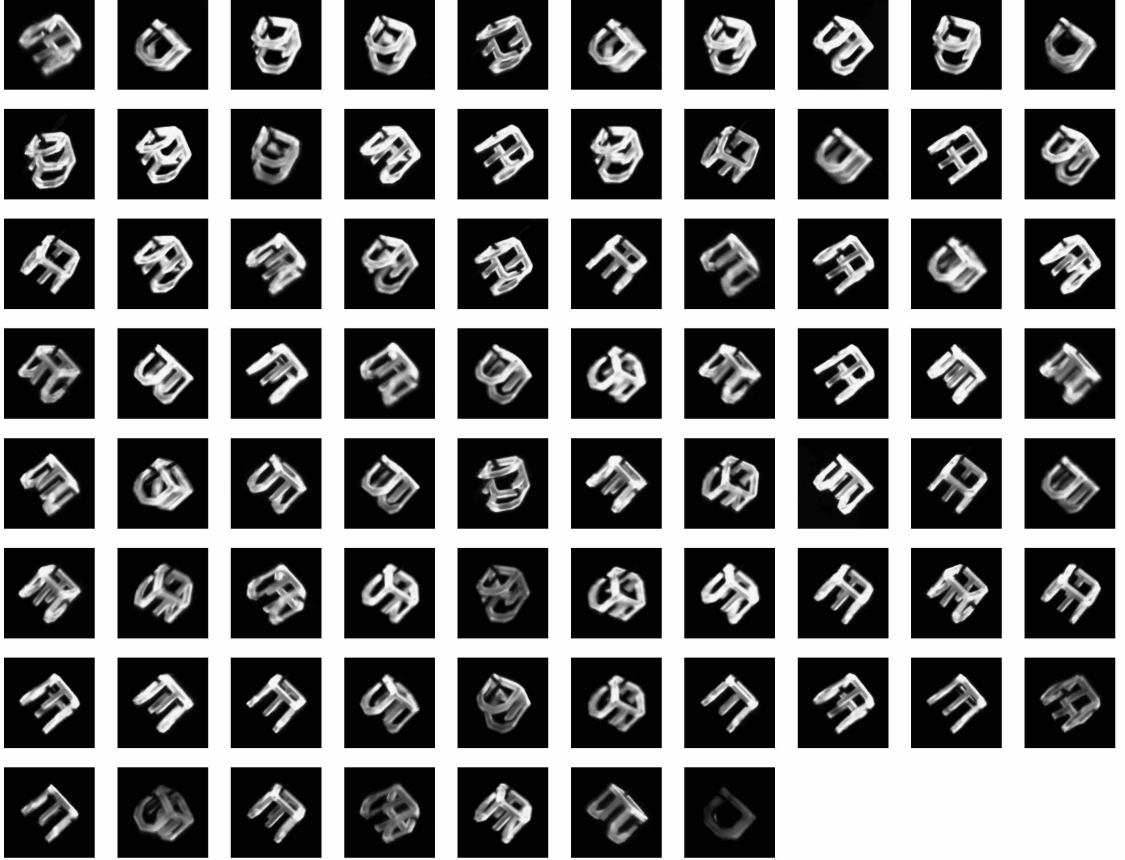


Figure B.5: Views from the learned NeuS reconstruction corresponding to those used in training. Relative to the captured images shown in Figure 8.13, results are visually nearly identical, demonstrating reasonable closure of the model. Figure sourced from [41].

Figure B.6 shows a grid of generated images corresponding to the same 77 training views, with information about the normal vectors of the learned surface. Colors correspond to the magnitude of these normal vectors when projected into the 2D image

plane for each view. The learned 3D structure is evident, with clearly learned letter and thread structure, as already shown by Figures 8.15 and 8.16. Note that there are sometimes small shifts in apparent location of the colored objects of Figure B.5 relative to the projections of the surface in Figure B.6, demonstrating the way that NeuS deals with misalignments – the learned surface aggregates the consistent 3D structure across all of the views, and remaining misalignments for individual views are handled via coloration of that learned surface.



Figure B.6: Generated grid of surface normals from the learned NeuS model, corresponding to the views used in training (and the generated images of Figure B.5). The colors correspond to the magnitude of the surface normal vectors when projected onto the corresponding 2D image plane. The learned 3D structure of the DOE object is apparent, as expressed via the meshes and depth maps of Figures 8.15 and 8.16. Figure sourced from [41].

Bibliography

- [1] **Particle Data Group** Collaboration, R. L. Workman *et. al.*, *Review of Particle Physics, Progress of Theoretical and Experimental Physics* **2022** (08, 2022).
- [2] **ATLAS** Collaboration, *Operation of the ATLAS trigger system in Run 2*, *JINST* **15** (Oct, 2020) P10004. 60 p.
- [3] **ATLAS** Collaboration, “ATLAS Standard Model.” <https://cds.cern.ch/record/2765709>, May, 2021. General Photo.
- [4] Wikipedia contributors, “Mathematical formulation of the standard model — Wikipedia, the free encyclopedia.” https://en.wikipedia.org/wiki/Mathematical_formulation_of_the_Standard_Model, 2022.
- [5] M. Aaboud, G. Aad, B. Abbott *et. al.*, *Determination of the strong coupling constant α_s from transverse energy–energy correlations in multijet events at $\sqrt{s} = 8$ TeV using the ATLAS detector*, *Eur. Phys. J. C* **77** (2017), no. 12 872 [[1707.02562](https://arxiv.org/abs/1707.02562)].
- [6] **ATLAS** Collaboration, *Standard Model Summary Plots February 2022*, *ATL-PHYS-PUB-2022-009* (Mar, 2022).
- [7] M. Carena, J. Kozaczuk, Z. Liu, T. Ou, M. J. Ramsey-Musolf, J. Shelton, Y. Wang and K.-P. Xie, “Probing the electroweak phase transition with exotic higgs decays.” <https://arxiv.org/abs/2203.08206>, 2022.

- [8] D. Antypas, A. Banerjee, C. Bartram, M. Baryakhtar, J. Betz, J. J. Bollinger, C. Boutan, D. Bowring, D. Budker, D. Carney, G. Carosi, S. Chaudhuri, S. Cheong, A. Chou, M. D. Chowdhury, R. T. Co, J. R. C. López-Urrutia, M. Demarteau, N. DePorzio, A. V. Derbin, T. Deshpande, M. D. Chowdhury, L. Di Luzio, A. Diaz-Morcillo, J. M. Doyle, A. Drlica-Wagner, A. Droster, N. Du, B. Döbrich, J. Eby, R. Essig, G. S. Farren, N. L. Figueroa, J. T. Fry, S. Gardner, A. A. Geraci, A. Ghalsasi, S. Ghosh, M. Giannotti, B. Gimeno, S. M. Griffin, D. Grin, D. Grin, H. Grote, J. H. Gundlach, M. Guzzetti, D. Hanneke, R. Harnik, R. Henning, V. Irsic, H. Jackson, D. F. J. Kimball, J. Jaeckel, M. Kagan, D. Kedar, R. Khatiwada, S. Knirck, S. Kolkowitz, T. Kovachy, S. E. Kuenstner, Z. Lasner, A. F. Leder, R. Lehnert, D. R. Leibrandt, E. Lentz, S. M. Lewis, Z. Liu, J. Manley, R. H. Maruyama, A. J. Millar, V. N. Muratova, N. Musoke, S. Nagaitsev, O. Noroozian, C. A. J. O'Hare, J. L. Ouellet, K. M. W. Pappas, E. Peik, G. Perez, A. Phipps, N. M. Rapidis, J. M. Robinson, V. H. Robles, K. K. Rogers, J. Rudolph, G. Rybka, M. Safdari, M. Safdari, M. S. Safronova, C. P. Salemi, P. O. Schmidt, T. Schumm, A. Schwartzman, J. Shu, M. Simanovskaia, J. Singh, S. Singh, M. S. Smith, W. M. Snow, Y. V. Stadnik, C. Sun, A. O. Sushkov, T. M. P. Tait, V. Takhistov, D. B. Tanner, D. J. Temples, P. G. Thirolf, J. H. Thomas, M. E. Tobar, O. Tretiak, Y. D. Tsai, J. A. Tyson, M. Vandegar, S. Vermeulen, L. Visinelli, E. Vitagliano, Z. Wang, D. J. Wilson, L. Winslow, S. Withington, M. Wooten, J. Yang, J. Ye, B. A. Young, F. Yu, M. H. Zaheer, T. Zelevinsky, Y. Zhao and K. Zhou, “New horizons: Scalar and vector ultralight dark matter.” <https://arxiv.org/abs/2203.14915>, 2022.
- [9] M. Abe *et. al.*, *Matter-wave atomic gradiometer interferometric sensor (MAGIS-100)*, *Quantum Science and Technology* **6** (jul, 2021) 044003.
- [10] Maximilien Brice (CERN), “Aerial view of CERN.” <https://supernova.eso.org/exhibition/images/cern-aerial-cc/>. Aerial Photo.

- [11] E. Mobs, “The CERN accelerator complex - August 2018. Complexe des accélérateurs du CERN - Août 2018.”
<https://cds.cern.ch/record/2636343>, Aug, 2018. General Photo.
- [12] T. Ferrand, H. Klingbeil and H. Damerau, “Synchronization of Synchrotrons for bunch-to-bucket Transfers.” <https://cds.cern.ch/record/2053285>, Sep, 2015.
- [13] *Design Report Tevatron 1 project, FERMILAB-DESIGN-1984-01* (1984).
- [14] W. Stirling, “Parton luminosity and cross section plots, private communication.”
<http://www.hep.ph.ic.ac.uk/~wstirlin/plots/plots.html>, 2012.
- [15] “Luminosity Public Results.” <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults>.
- [16] “Luminosity Public Results Run 2.” <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2>.
- [17] J. Pequenaio, “Computer generated image of the whole ATLAS detector.”
<https://cds.cern.ch/record/1095924>, Mar, 2008.
- [18] J. Pequenaio and P. Schaffner, “How ATLAS detects particles: diagram of particle paths in the detector.” <https://cds.cern.ch/record/1505342>, Jan, 2013.
- [19] J. Pequenaio, “Computer generated image of the ATLAS inner detector.”
<https://cds.cern.ch/record/1095926>, Mar, 2008.
- [20] **ATLAS** Collaboration, M. Aaboud *et. al.*, *Study of the material of the ATLAS inner detector for run 2 of the LHC*, *Journal of Instrumentation* **12** (dec, 2017) P12009. 71 p.
- [21] J. Pequenaio, “Computer Generated image of the ATLAS calorimeter.”
<https://cds.cern.ch/record/1095927>, Mar, 2008.

- [22] J. Pequenaio, “Computer generated image of the ATLAS Muons subsystem.” <https://cds.cern.ch/record/1095929>, Mar, 2008.
- [23] B. Nachman, *Investigating the Quantum Properties of Jets and the Search for a Supersymmetric Top Quark Partner with the ATLAS Detector*. PhD thesis, Sep, 2016. Presented 2016.
- [24] **ATLAS** Collaboration, “ATLAS event at 900 GeV - 6 May 2015 - Run 264034 Evt 11475271.” <https://cds.cern.ch/record/2014666>, May, 2015. General Photo.
- [25] **ATLAS** Collaboration, *Vertex Reconstruction Performance of the ATLAS Detector at $\sqrt{s} = 13$ TeV*, ATL-PHYS-PUB-2015-026 (Jul, 2015).
- [26] **ATLAS** Collaboration, *Development of ATLAS Primary Vertex Reconstruction for LHC Run 3*, ATL-PHYS-PUB-2019-015 (Apr, 2019).
- [27] M. Aaboud, G. Aad, B. Abbott *et. al.*, *Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton–proton collisions at the LHC*, *Eur. Phys. J. C* **77** (2017), no. 5 332 [[1611.10235](#)].
- [28] **ATLAS** Collaboration, M. Aaboud *et. al.*, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, *Eur. Phys. J.* **C77** (2017), no. 7 466 [[1703.10485](#)].
- [29] “Jet 2012 Monte Carlo Event Displays.” <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/JetEtmisApprovedB00ST2014EventDisplays>, Aug, 2014.
- [30] **ATLAS** Collaboration, G. Aad, B. Abbott *et. al.*, *Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ tev with the atlas detector*, *The European Physical Journal C* **81** (2021), no. 8 689.
- [31] **ATLAS** Collaboration, M. Aaboud, G. Aad, B. Abbott *et. al.*, *In situ calibration of large-radius jet energy and mass in 13 TeV proton-proton*

- collisions with the ATLAS detector*, *Eur. Phys. J.* **C79** (2019), no. 2 135 [[1807.09477](#)].
- [32] **ATLAS** Collaboration, *Digluon Tagging using $\sqrt{s} = 13$ TeV pp Collisions in the ATLAS Detector*, *ATL-PHYS-PUB-2021-027* (2021).
- [33] G. Kasieczka, B. Nachman, M. D. Schwartz and D. Shih, *Automating the ABCD method with machine learning*, *Phys. Rev. D* **103** (2021), no. 3 035021 [[2007.14400](#)].
- [34] **ATLAS** Collaboration, *Expected tracking and related performance with the updated ATLAS Inner Tracker layout at the High-Luminosity LHC*, *ATL-PHYS-PUB-2021-024* (Jul, 2021).
- [35] **ATLAS** Collaboration, “Simulated HL-LHC collision event in the ATLAS detector.” <https://cds.cern.ch/record/2674770>, May, 2019. General Photo.
- [36] **ATLAS** Collaboration, M. Safdari, A. Schwartzman, V. M. Cairo, N. E. Pettersson, M. Goblirsch-Kolb, G. R. Lee and D. Su, *Rescuing VBF Higgs Invisible Events with Novel Vertex Selection*, *ATL-PHYS-PROC-2020-042* (Jun, 2020).
- [37] **ATLAS** Collaboration, M. Safdari, A. Schwartzman, V. M. Cairo and G. R. Lee, *Primary Vertex Selection in VBF Higgs to Invisibles at the HL-LHC with the ATLAS Experiment*, *ATL-PHYS-PROC-2019-044* (Jun, 2019) [[1910.07681](#)].
- [38] S. Cheong, A. Cukierman, B. Nachman, M. Safdari and A. Schwartzman, *Parametrizing the Detector Response with Neural Networks*, *JINST* **15** (2020), no. 01 P01030 [[1910.03773](#)].
- [39] **ATLAS** Collaboration, *Simultaneous Jet Energy and Mass Calibrations with Neural Networks*, *ATL-PHYS-PUB-2020-001* (Jan, 2020).

- [40] A. Sugarbaker, S. M. Dickerson, J. M. Hogan, D. M. S. Johnson and M. A. Kasevich, *Enhanced atom interferometer readout through the application of phase shear*, *Phys. Rev. Lett.* **111** (Sep, 2013) 113002.
- [41] S. Cheong, J. C. Frisch, S. Gasiorowski, J. M. Hogan, M. Kagan, M. Safdari, A. Schwartzman and M. Vandegar, *Novel light field imaging device with enhanced light collection for cold atom clouds*, *JINST* **17** (2022) P08021 [[2205.11480](#)].
- [42] E. Ward, “Key Particle Discoveries Timeline.”
<https://cds.cern.ch/record/2665175>, Mar, 2019. General Photo.
- [43] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, *Phys.Rev.Lett.* **13** (1964) 321–323.
- [44] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, *Phys.Rev.Lett.* **13** (1964) 508–509.
- [45] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, *Phys.Lett.* **12** (1964) 132–133.
- [46] G. Guralnik, C. Hagen and T. Kibble, *Global Conservation Laws and Massless Particles*, *Phys.Rev.Lett.* **13** (1964) 585–587.
- [47] **ATLAS** Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett. B* **716** (2012) 1 [[1207.7214](#)].
- [48] **CMS** Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Phys. Lett. B* **716** (2012) 30 [[1207.7235](#)].
- [49] G. Bertone, D. Hooper and J. Silk, *Particle dark matter: evidence, candidates and constraints*, *Physics Reports* **405** (2005), no. 5 279–390.
- [50] W. Hu, R. Barkana and A. Gruzinov, *Fuzzy cold dark matter: The wave properties of ultralight particles*, *Phys. Rev. Lett.* **85** (Aug, 2000) 1158–1161.

- [51] M. Y. Khlopov, B. A. Malomed and Y. B. Zeldovich, *Gravitational instability of scalar fields and formation of primordial black holes*, *Monthly Notices of the Royal Astronomical Society* **215** (08, 1985) 575–589.
- [52] W. H. Press, B. S. Ryden and D. N. Spergel, *Single mechanism for generating large-scale structure and providing dark missing matter*, *Phys. Rev. Lett.* **64** (Mar, 1990) 1084–1087.
- [53] V. Sahni and L. Wang, *New cosmological model of quintessence and dark matter*, *Phys. Rev. D* **62** (Oct, 2000) 103517.
- [54] V. Trimble, *Existence and nature of dark matter in the universe*, *Annual Review of Astronomy and Astrophysics* **25** (1987), no. 1 425–472.
- [55] “LHC Guide.” <https://cds.cern.ch/record/2255762>, Mar, 2017.
- [56] **LHC Study Group** Collaboration, T. S. Pettersson and P. Lefèvre, *The Large Hadron Collider: conceptual design*, CERN-AC-95-05-LHC (Oct, 1995).
- [57] L. R. Evans, *The Large Hadron Collider Project*, LHC-Project-Report-53. CERN-LHC-Project-Report-53 (Sep, 1996) 10 p.
- [58] *LHC Machine*, *JINST* **3** (2008) S08001.
- [59] M. Gell-Mann, *Quarks*, *Acta Phys. Austriaca Suppl.* **9** (1972) 733–761.
- [60] H. Fritzsch, M. Gell-Mann and H. Leutwyler, *Advantages of the Color Octet Gluon Picture*, *Phys. Lett. B* **47** (1973) 365–368.
- [61] H. Fritzsch and M. Gell-Mann, *Current algebra: Quarks and what else?*, eConf **C720906V2** (1972) 135–165 [[hep-ph/0208010](#)].
- [62] A. Ali and G. Kramer, *Jets and QCD: A Historical Review of the Discovery of the Quark and Gluon Jets and its Impact on QCD*, *Eur. Phys. J. H* **36** (2011) 245–326 [[1012.2288](#)].

- [63] **ATLAS** Collaboration, G. Aad *et. al.*, *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003.
- [64] A. Boveia and C. Doglioni, *Dark matter searches at colliders*, *Annual Review of Nuclear and Particle Science* **68** (2018), no. 1 429–459.
- [65] **XENON** Collaboration, E. Aprile *et. al.*, *Dark matter search results from a one ton-year exposure of xenon1t*, *Phys. Rev. Lett.* **121** (Sep, 2018) 111302.
- [66] F. Kahlhoefer, *Review of lhc dark matter searches*, *International Journal of Modern Physics A* **32** (2017), no. 13 1730006.
- [67] **PandaX-4T** Collaboration, Y. Meng *et. al.*, *Dark matter search results from the pandax-4t commissioning run*, *Phys. Rev. Lett.* **127** (Dec, 2021) 261802.
- [68] T. A. Porter, R. P. Johnson and P. W. Graham, *Dark matter searches with astroparticle data*, *Annual Review of Astronomy and Astrophysics* **49** (2011), no. 1 155–194.
- [69] N. Blinov, N. Craig, M. J. Dolan, J. de Vries, P. Draper, I. G. Garcia, B. Lillard and J. Shelton, “Snowmass white paper: Strong cp beyond axion direct detection.” <https://arxiv.org/abs/2203.07218>, 2022.
- [70] M. Carena, J. Kozaczuk, Z. Liu, T. Ou, M. J. Ramsey-Musolf, J. Shelton, Y. Wang and K.-P. Xie, “Probing the electroweak phase transition with exotic higgs decays.” <https://arxiv.org/abs/2203.08206>, 2022.
- [71] M. Zaazoua, , L. Truong, K. A. Assamagan, F. Fassi, and and, *Higgs portal vector dark matter interpretation: Review of EffectiveField theory approach and ultraviolet complete models*, *Letters in High Energy Physics* **2022** (2022).
- [72] T. G. Rizzo, “Portal matter and dark sector phenomenology at colliders.” <https://arxiv.org/abs/2202.02222>, 2022.
- [73] D. Curtin *et. al.*, *Exotic decays of the 125 GeV Higgs boson*, *Phys. Rev. D* **90** (2014), no. 7 075004 [[1312.4992](https://arxiv.org/abs/1312.4992)].

- [74] C. N. Yang and R. L. Mills, *Conservation of isotopic spin and isotopic gauge invariance*, *Phys. Rev.* **96** (Oct, 1954) 191–195.
- [75] Nobel Media AB 2020, “The nobel prize in physics 1957.”
<https://www.nobelprize.org/prizes/physics/1957/summary/>, 2020.
- [76] G. Barr, Devenish.R., R. Walczak and T. Weidberg, *Particle Physics in the LHC Era*. Oxford University Press, 2016.
- [77] A. J. Larkoski, *Elementary Particle Physics: An Intuitive Introduction*. Cambridge University Press, 2019.
- [78] S. Weinberg, *The Quantum theory of fields. Vol. 1: Foundations*. Cambridge University Press, 6, 2005.
- [79] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*. Addison-Wesley, Reading, USA, 1995.
- [80] M. D. Schwartz, *Quantum Field Theory and the Standard Model*. Cambridge University Press, 3, 2014.
- [81] B. R. Martin and G. Shaw, *Particle Physics, 2nd Edition*. 1998.
- [82] S. L. Glashow, *The renormalizability of vector meson interactions*, *Nucl. Phys.* **10** (1959) 107–117.
- [83] A. Salam, *Weak and Electromagnetic Interactions*, *Conf. Proc. C* **680519** (1968) 367–377.
- [84] S. Weinberg, *A Model of Leptons*, *Phys. Rev. Lett.* **19** (1967) 1264–1266.
- [85] Nobel Media AB 2020, “The nobel prize in physics 1979.”
<https://www.nobelprize.org/prizes/physics/1979/summary/>, 2020.
- [86] Nobel Media AB 2020, “The nobel prize in physics 2013.”
<https://www.nobelprize.org/prizes/physics/2013/summary/>, 2020.

- [87] **ATLAS** Collaboration, *A detailed map of higgs boson interactions by the atlas experiment ten years after the discovery*, *Nature* **607** (2022), no. 7917 52–59.
- [88] **CMS** Collaboration, *A portrait of the higgs boson by the cms experiment ten years after the discovery*, *Nature* **607** (2022), no. 7917 60–68.
- [89] Z. Maki, M. Nakagawa and S. Sakata, *Remarks on the Unified Model of Elementary Particles*, *Progress of Theoretical Physics* **28** (11, 1962) 870–880.
- [90] I. Esteban, M. C. Gonzalez-Garcia, M. Maltoni, T. Schwetz and A. Zhou, *The fate of hints: updated global analysis of three-flavor neutrino oscillations*, *Journal of High Energy Physics* **2020** (2020), no. 9 178.
- [91] “NuFIT 5.1.” www.nu-fit.org, 2021.
- [92] A. Taroni, *Nobel prize 2015: Kajita and mcdonald*, *Nature Physics* **11** (2015), no. 11 891–891.
- [93] N. Cabibbo, *Unitary symmetry and leptonic decays*, *Phys. Rev. Lett.* **10** (Jun, 1963) 531–533.
- [94] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, *Progress of Theoretical Physics* **49** (02, 1973) 652–657.
- [95] J. R. Ellis, M. K. Gaillard and G. G. Ross, *Search for Gluons in $e^+ e^-$ Annihilation*, *Nucl. Phys. B* **111** (1976) 253. [Erratum: *Nucl.Phys.B* 130, 516 (1977)].
- [96] G. 't Hooft, *Renormalization of Massless Yang-Mills Fields*, *Nucl. Phys. B* **33** (1971) 173–199.
- [97] H. D. Politzer, *Reliable perturbative results for strong interactions?*, *Phys. Rev. Lett.* **30** (Jun, 1973) 1346–1349.
- [98] D. J. Gross and F. Wilczek, *Ultraviolet behavior of non-abelian gauge theories*, *Phys. Rev. Lett.* **30** (Jun, 1973) 1343–1346.

- [99] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Nonabelian Gauge Theories*, *Phys. Rev. Lett.* **30** (1973) 1343–1346.
- [100] H. Politzer, *Reliable Perturbative Results for Strong Interactions?*, *Phys. Rev. Lett.* **30** (1973) 1346–1349.
- [101] D. J. Gross, *How to Test Scaling in Asymptotically Free Theories*, *Phys. Rev. Lett.* **32** (1974) 1071.
- [102] D. Kazakov, *Beyond the standard model: In search of supersymmetry*, in *2000 European School of High-Energy Physics*, pp. 125–199, 8, 2000. [hep-ph/0012288](#).
- [103] K. G. Wilson, *Confinement of Quarks*, .
- [104] S. Drell, D. J. Levy and T.-M. Yan, *A Theory of Deep Inelastic Lepton-Nucleon Scattering and Lepton Pair Annihilation Processes. 3. Deep Inelastic electron-Positron Annihilation*, *Phys. Rev. D* **1** (1970) 1617–1639.
- [105] R. Blankenbecler, S. J. Brodsky and J. Gunion, *Inclusive Processes at High Transverse Momentum*, *Phys. Lett. B* **42** (1972) 461–465.
- [106] J. D. Lykken, *Beyond the Standard Model*, 5, 2010. [1005.1676](#).
- [107] H. M. Lee, *Lectures on Physics Beyond the Standard Model*, 7, 2019. [1907.12409](#).
- [108] J. Ellis, *Physics Beyond the Standard Model*, *Nucl. Phys. A* **827** (2009) 187C–198C [[0902.0357](#)].
- [109] J. Ellis, *Outstanding questions: Physics beyond the Standard Model*, *Phil. Trans. Roy. Soc. Lond. A* **370** (2012) 818–830.
- [110] T. Virdee, *Beyond the standard model of particle physics*, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **374** (08, 2016) 20150259.

- [111] E. Halkiadakis, G. Redlinger and D. Shih, *Status and Implications of Beyond-the-Standard-Model Searches at the LHC*, *Ann. Rev. Nucl. Part. Sci.* **64** (2014) 319–342 [[1411.1427](#)].
- [112] C. Rovelli, *Notes for a brief history of quantum gravity*, in *9th Marcel Grossmann Meeting on Recent Developments in Theoretical and Experimental General Relativity, Gravitation and Relativistic Field Theories (MG 9)*, pp. 742–768, 6, 2000. [gr-qc/0006061](#).
- [113] A. Ashtekar and R. Geroch, *Quantum theory of gravitation*, *Reports on Progress in Physics* **37** (Oct, 1974) 1211–1256.
- [114] A. Ashtekar, M. Reuter and C. Rovelli, *From General Relativity to Quantum Gravity*, [1408.4336](#).
- [115] R. H. Kraichnan, *Special-relativistic derivation of generally covariant gravitation theory*, *Phys. Rev.* **98** (May, 1955) 1118–1122.
- [116] S. Mukhi, *String theory: a perspective over the last 25 years*, *Class. Quant. Grav.* **28** (2011) 153001 [[1110.2569](#)].
- [117] L. Canetti, M. Drewes and M. Shaposhnikov, *Matter and Antimatter in the Universe*, *New J. Phys.* **14** (2012) 095012 [[1204.4186](#)].
- [118] A. Sakharov, *Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe*, *Sov. Phys. Usp.* **34** (1991), no. 5 392–393.
- [119] M. Dine, *TASI lectures on the strong CP problem*, in *Theoretical Advanced Study Institute in Elementary Particle Physics (TASI 2000): Flavor Physics for the Millennium*, pp. 349–369, 6, 2000. [hep-ph/0011376](#).
- [120] C. Baker *et. al.*, *An Improved experimental limit on the electric dipole moment of the neutron*, *Phys. Rev. Lett.* **97** (2006) 131801 [[hep-ex/0602020](#)].
- [121] D. E. Morrissey and M. J. Ramsey-Musolf, *Electroweak baryogenesis*, *New Journal of Physics* **14** (dec, 2012) 125003.

- [122] S. Profumo, M. J. Ramsey-Musolf and G. Shaughnessy, *Singlet higgs phenomenology and the electroweak phase transition*, *Journal of High Energy Physics* **2007** (aug, 2007) 010–010.
- [123] J. Kozaczuk, M. J. Ramsey-Musolf and J. Shelton, *Exotic higgs boson decays and the electroweak phase transition*, *Physical Review D* **101** (jun, 2020).
- [124] M. Carena, Z. Liu and Y. Wang, *Electroweak phase transition with spontaneous z_2 -breaking*, *Journal of High Energy Physics* **2020** (aug, 2020).
- [125] V. Trimble, *Existence and Nature of Dark Matter in the Universe*, *Ann. Rev. Astron. Astrophys.* **25** (1987) 425–472.
- [126] J. Silk *et. al.*, *Particle Dark Matter: Observations, Models and Searches*. Cambridge Univ. Press, Cambridge, 2010.
- [127] A. G. Riess, “Dark matter.”
<https://www.britannica.com/science/dark-matter>.
- [128] F. Zwicky, *On the Masses of Nebulae and of Clusters of Nebulae*, *Astrophysical Journal* **86** (Oct, 1937) 217.
- [129] V. C. Rubin and J. Ford, W. Kent, *Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions*, *Astrophysical Journal* **159** (Feb, 1970) 379.
- [130] K. C. Freeman, *On the Disks of Spiral and S0 Galaxies*, *Astrophysical Journal* **160** (Jun, 1970) 811.
- [131] T. van Albada, J. N. Bahcall, K. Begeman and R. Sancisi, *The Distribution of Dark Matter in the Spiral Galaxy NGC-3198*, *Astrophys. J.* **295** (1985) 305–313.
- [132] D. Clowe, A. Gonzalez and M. Markevitch, *Weak lensing mass reconstruction of the interacting cluster 1E0657-558: Direct evidence for the existence of dark matter*, *Astrophys. J.* **604** (2004) 596–603 [[astro-ph/0312273](https://arxiv.org/abs/astro-ph/0312273)].

- [133] M. Markevitch, A. Gonzalez, D. Clowe, A. Vikhlinin, L. David, W. Forman, C. Jones, S. Murray and W. Tucker, *Direct constraints on the dark matter self-interaction cross-section from the merging galaxy cluster 1E0657-56*, *Astrophys. J.* **606** (2004) 819–824 [[astro-ph/0309303](#)].
- [134] **Planck** Collaboration, N. Aghanim *et. al.*, *Planck 2018 results. VI. Cosmological parameters*, [1807.06209](#).
- [135] S. D. White, C. Frenk and M. Davis, *Clustering in a Neutrino Dominated Universe*, *Astrophys. J.* **274** (1983) L1–L5.
- [136] J. Wess and J. Bagger, *Supersymmetry and supergravity*. World Publishing Corporation, 2009.
- [137] J. Wess, *SUPERSYMMETRY AND SUPERGRAVITY*, in *22nd International Conference on High-Energy Physics*, p. II.84, 7, 1984.
- [138] J. D. Lykken, *Introduction to supersymmetry*, in *Theoretical Advanced Study Institute in Elementary Particle Physics (TASI 96): Fields, Strings, and Duality*, pp. 85–153, 6, 1996. [hep-th/9612114](#).
- [139] F. Quevedo, S. Krippendorff and O. Schlotterer, *Cambridge Lectures on Supersymmetry and Extra Dimensions*, [1011.1491](#).
- [140] M. J. Strassler, *An Unorthodox introduction to supersymmetric gauge theory*, in *Theoretical Advanced Study Institute in Elementary Particle Physics (TASI 2001): Strings, Branes and EXTRA Dimensions*, pp. 561–638, 9, 2003. [hep-th/0309149](#).
- [141] M. E. Peskin, *Duality in supersymmetric Yang-Mills theory*, in *Theoretical Advanced Study Institute in Elementary Particle Physics (TASI 96): Fields, Strings, and Duality*, pp. 729–809, 2, 1997. [hep-th/9702094](#).
- [142] N. Seiberg, *Electric - magnetic duality in supersymmetric nonAbelian gauge theories*, *Nucl. Phys. B* **435** (1995) 129–146 [[hep-th/9411149](#)].

- [143] K. A. Intriligator and N. Seiberg, *Lectures on Supersymmetry Breaking*, *Class. Quant. Grav.* **24** (2007) S741–S772 [[hep-ph/0702069](#)].
- [144] B. A. Dobrescu and K. T. Matchev, *Light axion within the next-to-minimal supersymmetric standard model*, *JHEP* **09** (2000) 031 [[hep-ph/0008192](#)].
- [145] U. Ellwanger, J. F. Gunion, C. Hugonie and S. Moretti, *Towards a no lose theorem for NMSSM Higgs discovery at the LHC*, [hep-ph/0305109](#).
- [146] R. Dermisek and J. F. Gunion, *Escaping the large fine tuning and little hierarchy problems in the next to minimal supersymmetric model and $h \rightarrow aa$ decays*, *Phys. Rev. Lett.* **95** (2005) 041801 [[hep-ph/0502105](#)].
- [147] S. Chang, R. Dermisek, J. F. Gunion and N. Weiner, *Nonstandard Higgs Boson Decays*, *Ann. Rev. Nucl. Part. Sci.* **58** (2008) 75–98 [[0801.4554](#)].
- [148] D. E. Morrissey and A. Pierce, *Modified Higgs Boson Phenomenology from Gauge or Gaugino Mediation in the NMSSM*, *Phys. Rev.* **D78** (2008) 075029 [[0807.2259](#)].
- [149] R. D. Peccei and H. R. Quinn, *CP Conservation in the Presence of Instantons*, *Phys. Rev. Lett.* **38** (1977) 1440–1443.
- [150] V. Baluni, *CP Violating Effects in QCD*, *Phys. Rev. D* **19** (1979) 2227–2230.
- [151] R. Crewther, P. Di Vecchia, G. Veneziano and E. Witten, *Chiral Estimate of the Electric Dipole Moment of the Neutron in Quantum Chromodynamics*, *Phys. Lett. B* **88** (1979) 123. [Erratum: *Phys.Lett.B* 91, 487 (1980)].
- [152] R. Peccei and H. R. Quinn, *Constraints Imposed by CP Conservation in the Presence of Instantons*, *Phys. Rev. D* **16** (1977) 1791–1797.
- [153] S. Weinberg, *A New Light Boson?*, *Phys. Rev. Lett.* **40** (1978) 223–226.
- [154] F. Wilczek, *Problem of Strong P and T Invariance in the Presence of Instantons*, *Phys. Rev. Lett.* **40** (1978) 279–282.

- [155] G. G. Raffelt, *Axions in astrophysics and cosmology*, in *30th Rencontres de Moriond: Euroconferences: Dark Matter in Cosmology, Clocks and Tests of Fundamental Laws*, pp. 159–168, 1995. [hep-ph/9502358](#).
- [156] A. Arvanitaki, J. Huang and K. Van Tilburg, *Searching for dilaton dark matter with atomic clocks*, *Phys. Rev. D* **91** (2015), no. 1 015015 [[1405.2925](#)].
- [157] Y. V. Stadnik and V. V. Flambaum, *Searching for dark matter and variation of fundamental constants with laser and maser interferometry*, *Phys. Rev. Lett.* **114** (2015) 161301 [[1412.7801](#)].
- [158] Y. V. Stadnik and V. V. Flambaum, *Enhanced effects of variation of the fundamental constants in laser interferometers and application to dark-matter detection*, *Phys. Rev. A* **93** (Jun, 2016) 063630.
- [159] P. W. Graham, D. E. Kaplan, J. Mardon, S. Rajendran and W. A. Terrano, *Dark Matter Direct Detection with Accelerometers*, *Phys. Rev. D* **93** (2016), no. 7 075029 [[1512.06165](#)].
- [160] A. Hees, O. Minazzoli, E. Savalle, Y. V. Stadnik and P. Wolf, *Violation of the equivalence principle from light scalar dark matter*, *Phys. Rev. D* **98** (2018), no. 6 064051 [[1807.04512](#)].
- [161] R. Bouchendira, P. Cladé, S. Guellati-Khélifa, F. m. c. Nez and F. m. c. Biraben, *New Determination of the Fine Structure Constant and Test of the Quantum Electrodynamics*, *Phys. Rev. Lett.* **106** (Feb, 2011) 080801.
- [162] R. H. Parker, C. Yu, W. Zhong, B. Estey and H. Müller, *Measurement of the fine-structure constant as a test of the Standard Model*, *Science* **360** (2018) 191 [[1812.04130](#)].
- [163] G. Rosi, G. D’Amico, L. Cacciapuoti, F. Sorrentino, M. Prevedelli, M. Zych, Č. Brukner and G. M. Tino, *Quantum Test of the Equivalence Principle for Atoms in Coherent Superposition of Internal Energy States*, *Nature Communications* **8** (Jun, 2017) 15529.

- [164] D. Schlippert, J. Hartwig, H. Albers, L. L. Richardson, C. Schubert, A. Roura, W. P. Schleich, W. Ertmer and E. M. Rasel, *Quantum Test of the Universality of Free Fall*, *Phys. Rev. Lett.* **112** (May, 2014) 203002.
- [165] P. Asenbaum, C. Overstreet, M. Kim, J. Curti and M. A. Kasevich, *Atom-Interferometric Test of the Equivalence Principle at the 10^{-12} Level*, *Phys. Rev. Lett.* **125** (Nov, 2020) 191101.
- [166] J. Hartwig, S. Abend, C. Schubert, D. Schlippert, H. Ahlers, K. Posso-Trujillo, N. Gaaloul, W. Ertmer and E. M. Rasel, *Testing the Universality of Free Fall with Rubidium and Ytterbium in a Very Large Baseline Atom Interferometer*, *New Journal of Physics* **17** (mar, 2015) 035011.
- [167] Z.-z. Xing, *Quark Mass Hierarchy and Flavor Mixing Puzzles*, *Int. J. Mod. Phys. A* **29** (2014) 1430067 [[1411.2713](#)].
- [168] **CMS** Collaboration, *The CMS experiment at the CERN LHC*, *JINST* **3** (2008) S08004.
- [169] **LHCb** Collaboration, J. Alves, A. Augusto *et. al.*, *The LHCb Detector at the LHC*, *JINST* **3** (2008) S08005.
- [170] **ALICE** Collaboration, K. Aamodt *et. al.*, *The ALICE experiment at the CERN LHC*, *JINST* **3** (2008) S08002.
- [171] J. L. Feng, I. Galon, F. Kling and S. Trojanowski, *Forward search experiment at the lhc*, *Phys. Rev. D* **97** (Feb, 2018) 035001.
- [172] **MoEDAL** Collaboration, J. Pinfold *et. al.*, *Technical Design Report of the MoEDAL Experiment*, .
- [173] **LHCf** Collaboration, O. Adriani *et. al.*, *The LHCf detector at the CERN large hadron collider*, *Journal of Instrumentation* **3** (aug, 2008) S08006–S08006.
- [174] **TOTEM** Collaboration, G. Anelli *et. al.*, *The TOTEM experiment at the CERN large hadron collider*, *Journal of Instrumentation* **3** (aug, 2008) S08007–S08007.

- [175] “Accelerator upgrades during LS2.”
<https://home.cern/press/2022/accelerator-upgrades-during-ls2>, 2022.
- [176] “LHC upgrades during LS2.”
<https://home.cern/press/2022/lhc-upgrades-during-ls2>, 2022.
- [177] M. Vretenar, J. Vollaïre, R. Scrivens, C. Rossi, F. Roncarolo, S. Ramberger, U. Raich, B. Puccio, D. Nisbet, R. Mompo, S. Mathot, C. Martin, L. A. Lopez-Hernandez, A. Lombardi, J. Lettry, J. B. Lallement, I. Kozsar, J. Hansen, F. Gerigk, A. Funken, J. F. Fuchs, N. Dos Santos, M. Calviani, M. Buzio, O. Brunner, Y. Body, P. Baudrenghien, J. Bauche and T. Zickler, *Linac4 design report*, vol. 6 of *CERN Yellow Reports: Monographs*. CERN, Geneva, 2020.
- [178] D. J. Warner, *Project study for a new 50 MeV linear accelerator for the C. P. S*, CERN-MPS-LINP-73-1 (Oct, 1973).
- [179] *New Linac + ‘old’ Booster = many protons. Nouveau Linac + «Booster» = multiplication des protons*, *CERN Bulletin* (1978) 1–2.
- [180] K. Hanke, *Past and present operation of the cern ps booster*, *International Journal of Modern Physics A* **28** (2013), no. 13 1330019
[\[https://doi.org/10.1142/S0217751X13300196\]](https://doi.org/10.1142/S0217751X13300196).
- [181] S. Gilardoni, D. Manglunki, J.-P. Burnet, C. Carli, M. Chanel, R. Garoby, M. Giovannozzi, S. Hancock, H. Haseroth, K. Hübner, D. Küchler, J. Lewis, A. Lombardi, M. Martini, S. Maury, E. Métral, D. Möhl, G. Plass, L. Rinolfi, R. Scrivens, R. Steerenberg, C. Steinbach, M. Vretenar and T. Zickler, *Fifty years of the CERN Proton Synchrotron: Volume 2*. CERN Yellow Reports: Monographs. Sep, 2013.
- [182] T. Mastoridis and P. Baudrenghien, *Cern’s super proton synchrotron 200 mhz cavity regulation upgrade: Modeling, design optimization, and performance estimation*, *Phys. Rev. Accel. Beams* **25** (Feb, 2022) 021002.

- [183] “CERN Accelerators.” <https://home.cern/science/accelerators>.
- [184] S. Humphries, *Principles of charged particle acceleration*. J. Wiley, 1986.
- [185] E. Wilson, *An Introduction to Particle Accelerators*. Oxford University Press, 2001.
- [186] J. D. Jackson, *Classical Electrodynamics*. Wiley, 3 ed., 1999.
- [187] D. Griffiths, *Introduction to Elementary Particles*. Wiley, 2 ed., 2008.
- [188] B. Holzer, *Introduction to Particle Accelerators and their Limitations*, [1705.09601](#).
- [189] H. Wiedemann, *Particle Accelerator Physics*. Springer Berlin, Heidelberg, 2007.
- [190] S. van der Meer, *Calibration of the effective beam height in the ISR*, *CERN-ISR-PO-68-31*, *ISR-PO-68-31* (1968).
- [191] **ATLAS** Collaboration, G. Aad *et. al.*, *Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC*, *The European Physical Journal C* **73** (aug, 2013).
- [192] G. Avoni *et. al.*, *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, *JINST* **13** (2018), no. 07 P07017.
- [193] S. Papadopoulou, F. Antoniou, I. Efthymiopoulos, M. Hostettler, G. Iadarola, N. Karastathis, S. Kostoglou, Y. Papaphilippou and G. Trad, *Monitoring and modelling of the LHC emittance and luminosity evolution in 2018*, *Journal of Physics: Conference Series* **1350** (nov, 2019) 012011.
- [194] N. Aquilina, M. Giovannozzi, M. Lamont, N. Sammut, R. Steinhausen, E. Todesco and J. Wenninger, *Tune variations in the large hadron collider*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **778** (2015) 6–13.

- [195] “LHC Commissioning - Longer Term LHC Schedule.”
<http://lhc-commissioning.web.cern.ch/lhc-commissioning/schedule/LHC-long-term.htm>.
- [196] S. Fartoukh, S. Kostoglou, M. Solfaroli Camillocci, G. Arduini, H. Bartosik, C. Bracco, K. Brodzinski, R. Bruce, X. Buffat, M. Calviani, F. Cerutti, I. Efthymiopoulos, B. Goddard, G. Iadarola, N. Karastathis, A. Lechner, E. Metral, N. Mounet, F.-X. Nuiry, P. S. Papadopoulou, Y. Papaphilippou, B. Petersen, T. H. B. Persson, S. Redaelli, G. Rumolo, B. Salvant, G. Sterbini, H. Timko, R. Tomas Garcia and J. Wenninger, *LHC Configuration and Operational Scenario for Run 3*, CERN-ACC-2021-0007 (Nov, 2021).
- [197] “First report to the LMC on Run III machine configurations (p-p runs).”
https://lhc-commissioning.web.cern.ch/2020/LCR3report_2019_1.pdf.
- [198] “High-Luminosity LHC.”
<https://home.cern/science/accelerators/high-luminosity-lhc>.
- [199] J. Boyd, *LHC Run-2 and Future Prospects*, in *2019 European School of High-Energy Physics*, 1, 2020. [2001.04370](#).
- [200] B. Schmidt, *The high-luminosity upgrade of the LHC: Physics and technology challenges for the accelerator and the experiments*, *Journal of Physics: Conference Series* **706** (apr, 2016) 022002.
- [201] Vankov, Peter, *Atlas upgrade for the hl-lhc: meeting the challenges of a five-fold increase in collision rate*, *EPJ Web of Conferences* **28** (2012) 12069.
- [202] J. Nielsen, *Physics prospects for ATLAS at the HL-LHC*, *Journal of Physics: Conference Series* **1690** (dec, 2020) 012156.
- [203] S. Kuehn, *Impact of the hl-lhc detector upgrades on the physics program of the atlas and cms experiments*, 2021.
- [204] J.-R. Vlimant, *HL-LHC tracking challenge. 2nd IML Machine Learning Workshop*, .

- [205] **ATLAS** Collaboration, *ATLAS inner detector: Technical Design Report, 1.* Technical design report. ATLAS. CERN, Geneva, 1997.
- [206] **ATLAS** Collaboration, S. Haywood, L. Rossi, R. Nickerson and A. Romaniouk, *ATLAS inner detector: Technical Design Report, 2.* Technical design report. ATLAS. CERN, Geneva, 1997.
- [207] **ATLAS** Collaboration, *ATLAS central solenoid: Technical Design Report.* Technical design report. ATLAS. CERN, Geneva, 1997.
- [208] **ATLAS** Collaboration, *Technical Design Report for the ATLAS Inner Tracker Pixel Detector, CERN-LHCC-2017-021, ATLAS-TDR-030* (Sep, 2017).
- [209] **ATLAS** Collaboration, *Technical Design Report for the ATLAS Inner Tracker Strip Detector, CERN-LHCC-2017-005, ATLAS-TDR-025* (Apr, 2017).
- [210] **ATLAS** Collaboration, 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, CERN-LHCC-2010-013, ATLAS-TDR-19* (Sep, 2010).
- [211] **ATLAS** Collaboration, *dE/dx measurement in the ATLAS Pixel Detector and its use for particle identification, ATLAS-CONF-2011-016* (Mar, 2011).
- [212] **ATLAS** Collaboration, *ATLAS calorimeter performance: Technical Design Report.* Technical design report. ATLAS. CERN, Geneva, 1996.
- [213] **ATLAS** Collaboration, *ATLAS liquid-argon calorimeter: Technical Design Report.* Technical design report. ATLAS. CERN, Geneva, 1996.
- [214] **ATLAS** Collaboration, *ATLAS tile calorimeter: Technical Design Report.* Technical design report. ATLAS. CERN, Geneva, 1996.
- [215] A. Lechner, *Particle interactions with matter, CERN Yellow Rep. School Proc.* **5** (2018) 47. 22 p.

- [216] E. Halkiadakis, “Particle interactions with matter (part ii) and particle id.” http://www.physics.rutgers.edu/~evahal/talks/tasi09/TASI_day3_school.pdf, Aug, 2009.
- [217] H. Abreu *et. al.*, *Performance of the electronic readout of the ATLAS liquid argon calorimeters*, *JINST* **5** (2010) P09003.
- [218] **ATLAS** Collaboration, A. Solodkov, *Performance of the ATLAS Tile Calorimeter*, *ATL-TILECAL-PROC-2015-024* (Dec, 2015).
- [219] **ATLAS** Collaboration, *ATLAS muon spectrometer: Technical Design Report*. Technical design report. ATLAS. CERN, Geneva, 1997.
- [220] **ATLAS** Collaboration, J. P. Badiou, J. Beltramelli, J. M. Baze and J. Belorgey, *ATLAS barrel toroid: Technical Design Report*. Technical design report. ATLAS. CERN, Geneva, 1997.
- [221] **ATLAS** Collaboration, *ATLAS end-cap toroids: Technical Design Report*. Technical design report. ATLAS. CERN, Geneva, 1997.
- [222] **ATLAS** Collaboration, G. Aad *et. al.*, *Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV*, *The European Physical Journal C* **73** (2013), no. 3 2304.
- [223] **ATLAS** Collaboration, M. Aaboud *et. al.*, *Performance of the ATLAS Trigger System in 2015*, *Eur. Phys. J. C* **77** (2017), no. 5 317 [[1611.09661](#)].
- [224] **ATLAS** Collaboration, A. Martinez, *The Run-2 ATLAS Trigger System*, *J. Phys. Conf. Ser.* **762** (2016), no. 1 012003.
- [225] **ATLAS** Collaboration, A. Buckley, T. Eifert, M. Elsing, D. I. Gillberg, K. Koeneke, A. Krasznahorkay, E. Moyse, M. Nowak, S. Snyder and P. van Gemmeren, *Implementation of the ATLAS Run 2 event data model*, *ATL-SOFT-PROC-2015-003* (May, 2015).

- [226] **ATLAS** Collaboration, *ATLAS level-1 trigger: Technical Design Report*. Technical design report. ATLAS. CERN, Geneva, 1998.
- [227] **ATLAS** Collaboration, P. Jenni, M. Nessi, M. Nordberg and K. Smith, *ATLAS high-level trigger, data-acquisition and controls: Technical Design Report*. Technical design report. ATLAS. CERN, Geneva, 2003.
- [228] **ATLAS** Collaboration, *2015 start-up trigger menu and initial performance assessment of the ATLAS trigger using Run-2 data, ATL-DAQ-PUB-2016-001* (Mar, 2016).
- [229] **ATLAS** Collaboration, *Trigger Menu in 2016, ATL-DAQ-PUB-2017-001* (Jan, 2017).
- [230] **ATLAS** Collaboration, *Trigger Menu in 2017, ATL-DAQ-PUB-2018-002* (Jun, 2018).
- [231] **ATLAS** Collaboration, *Trigger menu in 2018, ATL-DAQ-PUB-2019-001* (Oct, 2019).
- [232] **ATLAS** Collaboration, *Track Reconstruction Performance of the ATLAS Inner Detector at $\sqrt{s} = 13$ TeV, ATL-PHYS-PUB-2015-018* (Jul, 2015).
- [233] **ATLAS** Collaboration, M. Aaboud *et. al.*, *Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC Run 2*, *Eur. Phys. J. C* **77** (2017), no. 10 673 [[1704.07983](#)].
- [234] R. Fruhwirth, *Application of Kalman filtering to track and vertex fitting*, *Nucl. Instrum. Meth. A* **262** (1987) 444–450.
- [235] **ATLAS** Collaboration, G. Aad *et. al.*, *A neural network clustering algorithm for the ATLAS silicon pixel detector*, *JINST* **9** (2014) P09009 [[1406.7690](#)].
- [236] **ATLAS** Collaboration, G. Aad *et. al.*, *Performance of b -Jet Identification in the ATLAS Experiment*, *JINST* **11** (2016), no. 04 P04008 [[1512.01094](#)].

- [237] **ATLAS** Collaboration, M. Aaboud *et. al.*, *Measurements of b-jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s} = 13$ TeV*, *JHEP* **08** (2018) 089 [[1805.01845](#)].
- [238] **ATLAS** Collaboration, G. Aad *et. al.*, *ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **79** (2019), no. 11 970 [[1907.05120](#)].
- [239] **ATLAS** Collaboration, *Early Inner Detector Tracking Performance in the 2015 data at $\sqrt{s} = 13$ TeV*, *ATL-PHYS-PUB-2015-051* (Dec, 2015).
- [240] **ATLAS** Collaboration, G. Aad *et. al.*, *The ATLAS Inner Detector commissioning and calibration*, *Eur. Phys. J. C* **70** (2010) 787–821 [[1004.5293](#)].
- [241] **ATLAS** Collaboration, F. Meloni, *Primary vertex reconstruction with the ATLAS detector*, *JINST* **11** (2016), no. 12 C12060.
- [242] **ATLAS** Collaboration, *Secondary vertex finding for jet flavour identification with the ATLAS detector*, *ATL-PHYS-PUB-2017-011* (Jun, 2017).
- [243] S. Boutle, D. Casper, B. Hooberman, K. Grimm, B. Gui, G. Lee, J. Maurer, A. Morley, S. P. Griso, B. Petersen, K. Prokofiev, L. Shan, D. Shope, A. Wharton, B. Whitmore and M. Z. and, *Primary vertex reconstruction at the ATLAS experiment*, *Journal of Physics: Conference Series* **898** (oct, 2017) 042056.
- [244] **ATLAS** Collaboration, G. Aad *et. al.*, *Measurements of the Higgs boson inclusive and differential fiducial cross-sections in the diphoton decay channel with pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Journal of High Energy Physics* **2022** (2022), no. 8 27.
- [245] K. G. Tomiwa, *The performance of diphoton primary vertex reconstruction methods in $h \rightarrow \gamma \gamma + \text{met}$ channel of ATLAS experiment*, *Journal of Physics: Conference Series* **889** (sep, 2017) 012010.

- [246] **ATLAS** Collaboration, G. Aad *et. al.*, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, *Eur. Phys. J. C* **77** (2017) 490 [[1603.02934](#)].
- [247] “Public Liquid-Argon Calorimeter Plots on Upgrade.” <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LArCaloPublicResultsUpgrade>.
- [248] R. Kogler *et. al.*, *Jet Substructure at the Large Hadron Collider: Experimental Review*, *Rev. Mod. Phys.* **91** (2019), no. 4 045003 [[1803.06991](#)].
- [249] A. J. Larkoski, I. Moult and B. Nachman, *Jet Substructure at the Large Hadron Collider: A Review of Recent Advances in Theory and Machine Learning*, *Phys. Rept.* **841** (2020) 1–63 [[1709.04464](#)].
- [250] G. P. Salam, *Towards Jetography*, *Eur. Phys. J. C* **67** (2010) 637–686 [[0906.1833](#)].
- [251] J. E. Huth *et. al.*, *Toward a standardization of jet definitions*, in *1990 DPF Summer Study on High-energy Physics: Research Directions for the Decade (Snowmass 90)* Snowmass, Colorado, June 25-July 13, 1990, pp. 0134–136, 1990.
- [252] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*, *Phys. Rev. D* **48** (1993) 3160–3166 [[hep-ph/9305266](#)].
- [253] Y. L. Dokshitzer, G. D. Leder, S. Moretti and B. R. Webber, *Better jet clustering algorithms*, *JHEP* **08** (1997) 001 [[hep-ph/9707323](#)].
- [254] M. Cacciari, G. P. Salam and G. Soyez, *The anti- k_t jet clustering algorithm*, *JHEP* **04** (2008) 063 [[0802.1189](#)].
- [255] M. Cacciari, G. P. Salam and G. Soyez, *FastJet User Manual*, *Eur. Phys. J. C* **72** (2012) 1896 [[1111.6097](#)].
- [256] M. Cacciari, G. P. Salam and G. Soyez, *The Catchment Area of Jets*, *JHEP* **04** (2008) 005 [[0802.1188](#)].

- [257] **ATLAS** Collaboration, G. Aad *et. al.*, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data*, *JINST* **14** (2019), no. 12 P12006 [[1908.00005](#)].
- [258] J. Shelton, *TASI Lectures on Jet Substructure*, in *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: Searching for New Physics at Small and Large Scales (TASI 2012): Boulder, Colorado, June 4–29, 2012*, pp. 303–340, 2013. [1302.0260](#).
- [259] **ATLAS** Collaboration, M. Aaboud *et. al.*, *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Rev. D* **96** (2017), no. 7 072002 [[1703.09665](#)].
- [260] **ATLAS** Collaboration, *Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV*, *Eur. Phys. J. C* **73** (2013) [[1112.6426](#)].
- [261] A. Cukierman and B. Nachman, *Mathematical properties of numerical inversion for jet calibrations*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **858** (jun, 2017) 1–11.
- [262] A. R. Cukierman, *Searches for New Physics Using Jets with the ATLAS Detector*. PhD thesis, 2020. Presented 07 May 2020.
- [263] **ATLAS** Collaboration, *Jet mass reconstruction with the ATLAS Detector in early Run 2 data*, *ATLAS-CONF-2016-035* (Jul, 2016).
- [264] B. T. Elder and J. Thaler, *Aspects of Track-Assisted Mass*, *JHEP* **03** (2019) 104 [[1805.11109](#)].
- [265] **ATLAS** Collaboration, *Measurement of large radius jet mass reconstruction performance at $\sqrt{s} = 8$ TeV using the ATLAS detector*, *ATLAS-CONF-2016-008* (Mar, 2016).

- [266] **ATLAS** Collaboration, G. Aad *et. al.*, *Optimisation of large-radius jet reconstruction for the atlas detector in 13 tev proton–proton collisions*, *The European Physical Journal C* **81** (2021), no. 4 334.
- [267] **ATLAS** Collaboration, G. Aad *et. al.*, *Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, *Eur. Phys. J.* **C76** (2016), no. 11 581 [[1510.03823](#)].
- [268] **ATLAS** Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, *ATLAS-CONF-2014-018* (May, 2014).
- [269] **ATLAS** Collaboration, M. Aaboud *et. al.*, *Identification and rejection of pile-up jets at high pseudorapidity with the ATLAS detector*, *Eur. Phys. J.* **C77** (2017), no. 9 580 [[1705.02211](#)]. [Erratum: *Eur. Phys. J.* **C77**, no. 10, 712 (2017)].
- [270] **ATLAS** Collaboration, *Forward jet vertex tagging using the particle flow algorithm*, *ATL-PHYS-PUB-2019-026* (Jul, 2019).
- [271] **ATLAS** Collaboration, M. Aaboud *et. al.*, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **79** (2019), no. 8 639 [[1902.04655](#)].
- [272] **ATLAS** Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*, *ATLAS-CONF-2016-024* (Jun, 2016).
- [273] **ATLAS** Collaboration, *Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run-1 data*, *Eur. Phys. J. C* **76** (2016), no. 12 666 [[1606.01813](#)].
- [274] **ATLAS** Collaboration, M. Aaboud *et. al.*, *Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run 2 data collected in 2015 and 2016*, *Eur. Phys. J. C* **79** (2019), no. 3 205 [[1810.05087](#)].

- [275] **ATLAS** Collaboration, M. Aaboud *et. al.*, *Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton-proton collision data*, *JINST* **14** (2019), no. 03 P03017 [[1812.03848](#)].
- [276] **ATLAS** Collaboration, G. Aad *et. al.*, *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, *Eur. Phys. J. C* **74** (2014), no. 10 3071 [[1407.5063](#)].
- [277] **ATLAS** Collaboration, G. Aad *et. al.*, *Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment in pp collisions at $\sqrt{s}=8$ TeV*, *Eur. Phys. J. C* **75** (2015), no. 7 303 [[1412.7086](#)].
- [278] **ATLAS** Collaboration, *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC*, *ATL-PHYS-PUB-2015-045* (Nov, 2015).
- [279] **ATLAS** Collaboration, G. Aad *et. al.*, *Reconstruction of hadronic decay products of tau leptons with the ATLAS experiment*, *Eur. Phys. J. C* **76** (2016), no. 5 295 [[1512.05955](#)].
- [280] **ATLAS** Collaboration, *Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at $\sqrt{s} = 13$ TeV*, *ATLAS-CONF-2017-029* (May, 2017).
- [281] **ATLAS** Collaboration, G. Aad *et. al.*, *Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton-proton collision data*, *Eur. Phys. J. C* **74** (2014), no. 11 3130 [[1407.3935](#)].
- [282] **ATLAS** Collaboration, G. Aad *et. al.*, *Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **76** (2016), no. 5 292 [[1603.05598](#)].
- [283] **ATLAS** Collaboration, M. Aaboud *et. al.*, *Performance of missing transverse momentum reconstruction with the atlas detector using proton-proton*

- collisions at $\sqrt{s} = 13$ tev*, *The European Physical Journal C* **78** (2018), no. 11 903.
- [284] **ATLAS** Collaboration, E_T^{miss} performance in the ATLAS detector using 2015-2016 LHC p - p collisions, *ATLAS-CONF-2018-023* (Jun, 2018).
- [285] G. Luisoni, “An introduction to powheg.”
https://indico.cern.ch/event/656211/contributions/2673382/attachments/1498653/2333153/luisoni_powheg.pdf, Mar, 2017.
- [286] T. Sjöstrand, S. Mrenna and P. Z. Skands, *A brief introduction to PYTHIA 8.1*, *Comput. Phys. Commun.* **178** (2008) 852–867 [[0710.3820](#)].
- [287] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen and P. Z. Skands, *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159–177 [[1410.3012](#)].
- [288] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert and J. Winter, *Event generation with sherpa 1.1*, *JHEP* **02** (2009) 007 [[0811.4622](#)].
- [289] M. Bahr *et. al.*, *Herwig++ Physics and Manual*, *Eur. Phys. J.* **C58** (2008) 639–707 [[0803.0883](#)].
- [290] J. Bellm *et. al.*, *Herwig 7.0/Herwig++ 3.0 release note*, *Eur. Phys. J.* **C76** (2016), no. 4 196 [[1512.01178](#)].
- [291] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, *JHEP* **11** (2007) 070 [[0709.2092](#)].
- [292] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043 [[1002.2581](#)].

- [293] W. Verkerke, “Profile likelihood.”
<https://indico.in2p3.fr/event/9742/contributions/50419/>, May, 2014.
- [294] **ATLAS** Collaboration, *ATLAS Pythia 8 tunes to 7 TeV datas*,
ATL-PHYS-PUB-2014-021 (Nov, 2014).
- [295] **ATLAS** Collaboration, Z. Marshall, *Simulation of Pile-up in the ATLAS Experiment*, *J. Phys. Conf. Ser.* **513** (2014) 022024.
- [296] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin and C. P. Yuan, *New parton distributions for collider physics*, *Phys. Rev. D* **82** (2010) 074024 [[1007.2241](#)].
- [297] D. J. Lange, *The EvtGen particle decay simulation package*, *Nucl. Instrum. Meth.* **A462** (2001) 152–155.
- [298] **GEANT4** Collaboration, S. Agostinelli *et. al.*, *GEANT4: a simulation toolkit*, *Nucl. Instrum. Meth. A* **506** (2003) 250–303.
- [299] **ATLAS** Collaboration, *A combination of measurements of Higgs boson production and decay using up to 139 fb⁻¹ of proton–proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment*, *ATLAS-CONF-2020-027* (Aug, 2020).
- [300] **ATLAS** Collaboration, M. Aaboud *et. al.*, *Search for Higgs boson decays to beyond-the-Standard-Model light bosons in four-lepton events with the ATLAS detector at $\sqrt{s} = 13$ TeV*, *JHEP* **06** (2018) 166 [[1802.03388](#)].
- [301] **CMS** Collaboration, *Search for a low-mass dilepton resonance in Higgs boson decays to four-lepton final states at $\sqrt{s} = 13$ TeV*, .
- [302] **CMS** Collaboration, A. M. Sirunyan *et. al.*, *A search for pair production of new light bosons decaying into muons in proton-proton collisions at 13 TeV*, *Phys. Lett. B* **796** (2019) 131–154 [[1812.00380](#)].

- [303] **CMS** Collaboration, A. M. Sirunyan *et. al.*, *Search for a light pseudoscalar Higgs boson in the boosted $\mu\mu\tau\tau$ final state in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **08** (2020) 139 [[2005.08694](#)].
- [304] **CMS** Collaboration, A. M. Sirunyan *et. al.*, *Search for light pseudoscalar boson pairs produced from decays of the 125 GeV Higgs boson in final states with two muons and two nearby tracks in pp collisions at $\sqrt{s} = 13$ TeV*, *Phys. Lett. B* **800** (2020) 135087 [[1907.07235](#)].
- [305] **CMS** Collaboration, A. M. Sirunyan *et. al.*, *Search for an exotic decay of the Higgs boson to a pair of light pseudoscalars in the final state of two muons and two τ leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **11** (2018) 018 [[1805.04865](#)].
- [306] **CMS** Collaboration, A. M. Sirunyan *et. al.*, *Search for an exotic decay of the Higgs boson to a pair of light pseudoscalars in the final state with two muons and two b quarks in pp collisions at 13 TeV*, *Phys. Lett. B* **795** (2019) 398–423 [[1812.06359](#)].
- [307] **ATLAS** Collaboration, *Search for Higgs boson decays into two spin-0 particles in the $b\bar{b}\mu\mu$ final state with the ATLAS detector in pp collisions at $\sqrt{s} = 13$ TeV*, *ATLAS-CONF-2021-009* (Mar, 2021).
- [308] **ATLAS** Collaboration, M. Aaboud *et. al.*, *Search for the Higgs boson produced in association with a vector boson and decaying into two spin-zero particles in the $H \rightarrow a\bar{a} \rightarrow 4b$ channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *JHEP* **10** (2018) 031 [[1806.07355](#)].
- [309] **ATLAS** Collaboration, G. Aad *et. al.*, *Search for Higgs boson decays into two new low-mass spin-0 particles in the $4b$ channel with the ATLAS detector using pp collisions at $\sqrt{s} = 13$ TeV*, *Phys. Rev. D* **102** (2020) 112006 [[2005.12236](#)].
- [310] **ATLAS** Collaboration, *Summary of Exotic Higgs Boson Decays from the ATLAS Experiment*, *ATL-PHYS-PUB-2021-008* (Mar, 2021).

- [311] R. Garisto, *Editorial: Theorists react to the cern 750 gev diphoton data*, *Phys. Rev. Lett.* **116** (Apr, 2016) 150001.
- [312] K. Agashe, M. Ekhterachian, D. Kim and D. Sathyan, *LHC Signals for KK Graviton from an Extended Warped Extra Dimension*, *JHEP* **11** (2020) 109 [[2008.06480](#)].
- [313] P. T. Komiske, E. M. Metodiev and J. Thaler, *Energy Flow Networks: Deep Sets for Particle Jets*, *JHEP* **01** (2019) 121 [[1810.05165](#)].
- [314] M. Zaheer, S. Kottur, S. Ravanbakhsh, B. Póczos, R. Salakhutdinov and A. Smola, *Deep sets*, 2018.
- [315] G. Kasieczka and D. Shih, *Robust Jet Classifiers through Distance Correlation*, *Phys. Rev. Lett.* **125** (2020), no. 12 122001 [[2001.05310](#)].
- [316] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040 [[hep-ph/0409146](#)].
- [317] T. Sjöstrand, S. Mrenna and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, *JHEP* **05** (2006) 026 [[hep-ph/0603175](#)].
- [318] **ATLAS** Collaboration, *Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, *JHEP* **09** (2014) 145 [[1406.3660](#)].
- [319] **NNPDF** Collaboration, R. D. Ball *et. al.*, *Parton distributions for the LHC Run II*, *JHEP* **04** (2015) 040 [[1410.8849](#)].
- [320] G. Cullen, N. Greiner, G. Heinrich, G. Luisoni, P. Mastrolia, G. Ossola, T. Reiter and F. Tramontano, *Automated One-Loop Calculations with GoSam*, *Eur. Phys. J. C* **72** (2012) 1889 [[1111.2034](#)].
- [321] K. Hamilton, P. Nason and G. Zanderighi, *MINLO: Multi-Scale Improved NLO*, *JHEP* **10** (2012) 155 [[1206.3572](#)].

- [322] G. Luisoni, P. Nason, C. Oleari and F. Tramontano, $HW^\pm/HZ + 0$ and 1 jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO, *JHEP* **10** (2013) 083 [[1306.2542](#)].
- [323] **Sherpa** Collaboration, E. Bothmann *et. al.*, *Event Generation with Sherpa 2.2*, *SciPost Phys.* **7** (2019), no. 3 034 [[1905.09127](#)].
- [324] T. Gleisberg and S. Hoeche, *Comix, a new matrix element generator*, *JHEP* **12** (2008) 039 [[0808.3674](#)].
- [325] F. Buccioni, J.-N. Lang, J. M. Lindert, P. Maierhöfer, S. Pozzorini, H. Zhang and M. F. Zoller, *OpenLoops 2*, *Eur. Phys. J. C* **79** (2019), no. 10 866 [[1907.13071](#)].
- [326] F. Buccioni, S. Pozzorini and M. Zoller, *On-the-fly reduction of open loops*, *Eur. Phys. J. C* **78** (2018), no. 1 70 [[1710.11452](#)].
- [327] A. Denner, S. Dittmaier and L. Hofer, *Collier: a fortran-based Complex One-Loop Library in Extended Regularizations*, *Comput. Phys. Commun.* **212** (2017) 220–238 [[1604.06792](#)].
- [328] F. Cascioli, P. Maierhofer and S. Pozzorini, *Scattering Amplitudes with Open Loops*, *Phys. Rev. Lett.* **108** (2012) 111601 [[1111.5206](#)].
- [329] S. Carrazza, S. Forte, Z. Kassabov, J. I. Latorre and J. Rojo, *An Unbiased Hessian Representation for Monte Carlo PDFs*, *Eur. Phys. J. C* **75** (2015), no. 8 369 [[1505.06736](#)].
- [330] **NNPDF** Collaboration, R. D. Ball *et. al.*, *Parton distributions for the LHC Run II*, *JHEP* **04** (2015) 040 [[1410.8849](#)].
- [331] **ATLAS** Collaboration, G. Aad *et. al.*, *Modelling and computational improvements to the simulation of single vector-boson plus jet processes for the ATLAS experiment*, [2112.09588](#).

- [332] **ATLAS** Collaboration, *The ATLAS Simulation Infrastructure*, *Eur. Phys. J. C* **70** (2010) 823 [[1005.4568](#)].
- [333] **ATLAS** Collaboration, “The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie–Landshoff diffractive model.” ATL-PHYS-PUB-2016-017, 2016.
- [334] M. Cacciari, G. P. Salam and G. Soyez, *The anti- k_t jet clustering algorithm*, *JHEP* **2008** (apr, 2008) 063–063.
- [335] M. Aaboud, G. Aad, B. Abbott *et. al.*, *Jet reconstruction and performance using particle flow with the atlas detector*, *The European Physical Journal C* **77** (2017), no. 7 466.
- [336] **ATLAS** Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, *Eur. Phys. J. C* **76** (2016) 581 [[1510.03823](#)].
- [337] P. Gras, S. Höche, D. Kar, A. Larkoski, L. Lönnblad, S. Plätzer, A. Siódmok, P. Skands, G. Soyez and J. Thaler, *Systematics of quark/gluon tagging*, *JHEP* **07** (2017) 091 [[1704.03878](#)].
- [338] S. Bright-Thonney and B. Nachman, *Investigating the Topology Dependence of Quark and Gluon Jets*, *JHEP* **03** (2019) 098 [[1810.05653](#)].
- [339] B. Nachman, L. S. Miller, D. Gillberg, S. Kaur, A. P. Suresh and M. Pettee, *A simultaneous measurement of all charged particle properties in Z+jets events in $\sqrt{s} = 13$ TeV pp collisions using the ATLAS detector*, *ATL-COM-PHYS-2021-121* (Mar, 2021).
- [340] A. J. Larkoski, G. P. Salam and J. Thaler, *Energy Correlation Functions for Jet Substructure*, *JHEP* **06** (2013) 108 [[1305.0007](#)].
- [341] **CMS** Collaboration, *Search for a Higgs boson in the decay channel $H \rightarrow ZZ^{(*)} \rightarrow q\bar{q}\ell\ell^+$ in pp collisions at $\sqrt{s} = 7$ TeV*, *JHEP* **04** (2012) 036 [[1202.1416](#)].

- [342] A. Andreassen and B. Nachman, *Neural Networks for Full Phase-space Reweighting and Parameter Tuning*, *Phys. Rev. D* **101** (2020) 091901(R) [[1907.08209](#)].
- [343] K. Cranmer, J. Pavez and G. Louppe, *Approximating Likelihood Ratios with Calibrated Discriminative Classifiers*, [1506.02169](#).
- [344] P. Baldi, K. Cranmer, T. Faucett, P. Sadowski and D. Whiteson, *Parameterized neural networks for high-energy physics*, *Eur. Phys. J. C* **76** (2016), no. 5 235 [[1601.07913](#)].
- [345] T. Hastie, R. Tibshirani and J. Friedman, *The Elements of Statistical Learning*. Springer Series in Statistics. Springer New York Inc., New York, NY, USA, 2001.
- [346] M. Sugiyama, T. Suzuki and T. Kanamori, *Density Ratio Estimation in Machine Learning*. Cambridge University Press, 2012.
- [347] **ATLAS** Collaboration, *Deep Sets based Neural Networks for Impact Parameter Flavour Tagging in ATLAS*, *ATL-PHYS-PUB-2020-014* (2020).
- [348] F. Chollet, “keras.” <https://github.com/fchollet/keras>, 2015.
- [349] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard *et. al.*, *Tensorflow: A system for large-scale machine learning.*, in *OSDI*, vol. 16, pp. 265–283, 2016.
- [350] D. Kingma and J. Ba, *Adam: A method for stochastic optimization*, [1412.6980](#).
- [351] A. Andreassen, S.-C. Hsu, B. Nachman, N. Suaysom and A. Suresh, *Parameter estimation using neural networks in the presence of detector effects*, *Phys. Rev. D* **103** (2021), no. 3 036001 [[2010.03569](#)].
- [352] M. Stoye, J. Brehmer, G. Louppe, J. Pavez and K. Cranmer, *Likelihood-free inference with an improved cross-entropy estimator*, [1808.00973](#).

- [353] J. Hollingsworth and D. Whiteson, *Resonance Searches with Machine Learned Likelihood Ratios*, [2002.04699](#).
- [354] J. Brehmer, K. Cranmer, G. Louppe and J. Pavez, *Constraining Effective Field Theories with Machine Learning*, *Phys. Rev. Lett.* **121** (2018), no. 11 111801 [[1805.00013](#)].
- [355] J. Brehmer, K. Cranmer, G. Louppe and J. Pavez, *A Guide to Constraining Effective Field Theories with Machine Learning*, *Phys. Rev. D* **98** (2018), no. 5 052004 [[1805.00020](#)].
- [356] J. Brehmer, F. Kling, I. Espejo and K. Cranmer, *MadMiner: Machine learning-based inference for particle physics*, *Comput. Softw. Big Sci.* **4** (2020), no. 1 3 [[1907.10621](#)].
- [357] J. Brehmer, G. Louppe, J. Pavez and K. Cranmer, *Mining gold from implicit models to improve likelihood-free inference*, *Proc. Nat. Acad. Sci.* (2020) 201915980 [[1805.12244](#)].
- [358] C. Badiali, F. Di Bello, G. Frattari, E. Gross, V. Ippolito, M. Kado and J. Shlomi, *Efficiency Parameterization with Neural Networks*, [2004.02665](#).
- [359] A. Andreassen, B. Nachman and D. Shih, *Simulation Assisted Likelihood-free Anomaly Detection*, *Phys. Rev. D* **101** (2020), no. 9 095004 [[2001.05001](#)].
- [360] A. Andreassen, P. T. Komiske, E. M. Metodiev, B. Nachman and J. Thaler, *OmniFold: A Method to Simultaneously Unfold All Observables*, *Phys. Rev. Lett.* **124** (2020), no. 18 182001 [[1911.09107](#)].
- [361] M. Erdmann, B. Fischer, D. Noll, Y. Rath, M. Rieger and D. Schmidt, *Adversarial Neural Network-based data-simulation corrections for jet-tagging at CMS*, in *Proc. 19th Int. Workshop on Adv. Comp., Anal. Techn. in Phys. Research, ACAT2019*, 2019.
- [362] **ATLAS** Collaboration, *Search for resonances decaying into a weak vector boson and a Higgs boson in the fully hadronic final state produced in*

- proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Rev. D* **102** (2020), no. 11 112008 [[2007.05293](#)].
- [363] CMS Collaboration, *Search for Higgs boson pair production in the four b quark final state*, *CMS-PAS-HIG-20-005* (2021).
- [364] A. Andreassen and B. Nachman, *Neural Networks for Full Phase-space Reweighting and Parameter Tuning*, *Phys. Rev. D* **101** (2020), no. 9 091901 [[1907.08209](#)].
- [365] G. J. Székely, M. L. Rizzo and N. K. Bakirov, *Measuring and testing dependence by correlation of distances*, *Ann. Statist.* **35** (2007), no. 6 2769–2794.
- [366] G. J. Székely and M. L. Rizzo, *Brownian distance covariance*, *Ann. Appl. Stat.* **3** (2009), no. 4 1236–1265.
- [367] G. J. Székely and M. L. Rizzo, *The distance correlation t -test of independence in high dimension*, *J. Multivar. Anal.* **117** (2013) 193–213.
- [368] G. J. Székely and M. L. Rizzo, *Partial distance correlation with methods for dissimilarities*, *Ann. Statist.* **42** (2014), no. 6 2382–2412.
- [369] ATLAS Collaboration, *Quark versus Gluon Jet Tagging Using Charged-Particle Constituent Multiplicity with the ATLAS Detector*, *ATL-PHYS-PUB-2017-009* (2017).
- [370] ATLAS Collaboration, *Quark versus Gluon Jet Tagging Using Jet Images with the ATLAS Detector*, *ATL-PHYS-PUB-2017-017* (2017).
- [371] ATLAS Collaboration, G. Aad *et. al.*, *Charged-particle multiplicities in pp interactions measured with the ATLAS detector at the LHC*, *New Journal of Physics* **13** (may, 2011) 053033.
- [372] L.-G. G. and, *Development and performance of track reconstruction algorithms at the energy frontier with the ATLAS detector*, *Journal of Physics: Conference Series* **898** (oct, 2017) 042007.

- [373] J. D. Mansour, *Tracking in dense environments and its inefficiency measurement using pixel $d\epsilon/dx$* , *Journal of Instrumentation* **11** (dec, 2016) C12041–C12041.
- [374] M. Cepeda *et. al.*, *Report from Working Group 2: Higgs Physics at the HL-LHC and HE-LHC*, *CERN Yellow Rep. Monogr.* **7** (Dec, 2018) 221–584. 364 p [[1902.00134](#)].
- [375] **ATLAS** Collaboration, G. Aad *et. al.*, *Search for invisible higgs-boson decays in events with vector-boson fusion signatures using 139 fb^{-1} of proton-proton data recorded by the atlas experiment*, *Journal of High Energy Physics* **2022** (2022), no. 8 104.
- [376] **ATLAS** Collaboration, *Expected Tracking Performance of the ATLAS Inner Tracker at the HL-LHC*, *ATL-PHYS-PUB-2019-014* (Mar, 2019).
- [377] **ATLAS** Collaboration, *Tagging and suppression of pileup jets*, *ATL-PHYS-PUB-2014-001* (Jan, 2014).
- [378] **ATLAS** Collaboration, G. Aad *et. al.*, *Optimisation of large-radius jet reconstruction for the ATLAS detector in 13 TeV proton–proton collisions*, *Eur. Phys. J. C* **81** (Sep, 2020) 334. 47 p [[2009.04986](#)].
- [379] **ATLAS** Collaboration, *Generalized Numerical Inversion: A Neural Network Approach to Jet Calibration*, *ATL-PHYS-PUB-2018-013* (Jul, 2018).
- [380] P. J. Huber, *Robust estimation of a location parameter*, *Ann. Math. Statist.* **35** (03, 1964) 73–101.
- [381] **CMS** Collaboration, *A deep neural network for simultaneous estimation of b quark energy and resolution*, *CMS-PAS-HIG-18-027* (2019).
- [382] R. Koenker and G. Bassett, *Regression Quantiles*, *Econometrica* **46** (1978) 33.
- [383] J. Taylor, *A Quantile Regression Neural Network Approach to Estimating the Conditional Density of Multiperiod Returns*, *Journal of Forecasting* **19** (2000) 299.

- [384] S. Zheng, *Gradient descent algorithms for quantile regression with smooth approximation*, *International Journal of Machine Learning and Cybernetics* **2** (2011) 191.
- [385] K. Hatalis *et. al.*, *Smooth Pinball Neural Network for Probabilistic Forecasting of Wind Power*, *NREL/CP-5D00-71247* (2017).
- [386] N. Tagasovska and D. Lopez-Paz, *Frequentist uncertainty estimates for deep learning*, [1811.00908](#).
- [387] M. Lee, *Mode Regression*, *Journal of Econometrics* **42** (1986) 337.
- [388] W. Yao and L. Li, *A New Regression Model: Modal Linear Regression*, *Scandinavian Journal of Statistics* **41** (2013) 656.
- [389] R. Hogg and A. Craig, *Introduction to Mathematical Statistics*. Macmillan, New York, 1978.
- [390] J. E. Gaiser, *Charmonium Spectroscopy From Radiative Decays of the J/ψ and ψ'* , *SLAC-0255*, *UMI-83-14449-MC*, *SLAC-R-0255*, *SLAC-R-255* (8, 1982).
- [391] R. D. Ball *et. al.*, *Parton distributions with LHC data*, *Nucl. Phys.* **B867** (2013) 244–289 [[1207.1303](#)].
- [392] E. Bothmann, G. Singh Chahal, S. Höche, J. Krause, F. Krauss, S. Kuttimalai, S. Liebschner, D. Napoletano, M. Schönherr, H. Schulz and *et al.*, *Event generation with sherpa 2.2*, *SciPost Physics* **7** (Sep, 2019).
- [393] S. Dulat, T.-J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt, D. Stump and C. P. Yuan, *New parton distribution functions from a global analysis of quantum chromodynamics*, *Phys. Rev. D* **93** (2016), no. 3 033006 [[1506.07443](#)].
- [394] J. Bellm *et. al.*, *Herwig 7.1 Release Note*, [1705.06919](#).

- [395] L. A. Harland-Lang, A. D. Martin, P. Motylinski and R. S. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*, *Eur. Phys. J. C* **75** (2015), no. 5 204 [[1412.3989](#)].
- [396] B. R. Webber, *A QCD model for jet fragmentation including soft gluon interference*, *Nucl. Phys. B* **238** (1984) 492–528.
- [397] **ATLAS** Collaboration, *Multijet simulation for 13 TeV ATLAS Analyses*, *ATL-PHYS-PUB-2019-017* (2019).
- [398] F. Chollet, “Keras.” <https://github.com/fchollet/keras>, 2017.
- [399] A. L. Maas, A. Y. Hannun and A. Y. Ng, *Rectifier nonlinearities improve neural network acoustic models*, in *in ICML Workshop on Deep Learning for Audio, Speech and Language Processing*, 2013.
- [400] D. Carney, T. Cecil, J. Ellis, R. F. G. Ruiz, A. A. Geraci, D. Hanneke, J. Hogan, N. R. Hutzler, A. Jayich, S. Kolkowitz, G. W. Morley, H. Muller, Z. Pagel, C. Panda and M. S. Safronova, “Snowmass 2021: Quantum sensors for hep science – interferometers, mechanics, traps, and clocks.” <https://arxiv.org/abs/2203.07250>, 2022.
- [401] S. Chattopadhyay, R. Falcone and R. Walsworth, *Quantum Sensors at the Intersections of Fundamental Science, Quantum Information Science & Computing*, *DOE Technical Report* (2, 2016).
- [402] E. Farhi, S. Jordan, P. Hayden, M. Lukin, J. Maldacena, J. Preskill, P. Shor, J. Taylor and C. William, *Grand Challenges at the Interface of Quantum Information Science, Particle Physics, and Computing*, *DOE Study Group Report* (1, 2015).
- [403] M. Battaglieri *et. al.*, *US Cosmic Visions: New Ideas in Dark Matter 2017: Community Report*, [1707.04591](#).
- [404] A. M. Smith and K. Burnett, *Effect of diffusion on laser cooling of atoms with long-range collisions*, *J. Opt. Soc. Am. B* **9** (Aug, 1992) 1240–1255.

- [405] E. Hecht, A. Zajac and A. Zając, *Optics*. Addison-Wesley world student series. Addison-Wesley Publishing Company, 1987.
- [406] G. Wu, B. Masia, A. Jarabo, Y. Zhang, L. Wang, Q. Dai, T. Chai and Y. Liu, *Light field image processing: An overview, IEEE Journal of Selected Topics in Signal Processing* **11** (2017), no. 7 926–954.
- [407] S. Zhou, T. Zhu, K. Shi, Y. Li, W. Zheng and J. Yong, *Review of light field technologies, Visual Computing for Industry, Biomedicine, and Art* **4** (Dec, 2021) 29.
- [408] M. Reiss, *The $\cos^4$ law of illumination**, *J. Opt. Soc. Am.* **35** (Apr, 1945) 283–288.
- [409] “BMF HTL resin data sheet.”
https://2oec4j3twl711d5klt43pn63-wpengine.netdna-ssl.com/wp-content/uploads/2022/06/BMFHTLResinDatasheet_060922.pdf.
Accessed: 2022-04-14.
- [410] M. Vandegar, S. Gasiorowski, M. Safdari, S. Cheong, A. Schwartzman and M. Kagan, “Gradoptics: Differentiable optics via ray tracing.”
<https://github.com/magis-slac/GradOptics>, 2022.
- [411] M. Pharr, W. Jakob and G. Humphreys, *Physically Based Rendering: From Theory to Implementation*. Elsevier Science, 2016.
- [412] V. Sitzmann, M. Zollhöfer and G. Wetzstein, *Scene representation networks: Continuous 3d-structure-aware neural scene representations, CoRR* **abs/1906.01618** (2019) [[1906.01618](#)].
- [413] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi and R. Ng, *Nerf: Representing scenes as neural radiance fields for view synthesis, CoRR* **abs/2003.08934** (2020) [[2003.08934](#)].

- [414] V. Sitzmann, J. N. P. Martel, A. W. Bergman, D. B. Lindell and G. Wetzstein, *Implicit neural representations with periodic activation functions*, *CoRR* **abs/2006.09661** (2020) [[2006.09661](https://arxiv.org/abs/2006.09661)].
- [415] Y. Xie, T. Takikawa, S. Saito, O. Litany, S. Yan, N. Khan, F. Tombari, J. Tompkin, V. Sitzmann and S. Sridhar, “Neural fields in visual computing and beyond.” <https://arxiv.org/abs/2111.11426>, 2021.
- [416] A. C. Kak and M. Slaney, *Principles of computerized tomographic imaging*. SIAM, 2001.
- [417] P. Wang, L. Liu, Y. Liu, C. Theobalt, T. Komura and W. Wang, “Neus: Learning neural implicit surfaces by volume rendering for multi-view reconstruction.” <https://arxiv.org/abs/2106.10689>, 2021.
- [418] D. B. Lindell, J. N. P. Martel and G. Wetzstein, *AutoInt: Automatic integration for fast neural volume rendering*, in *Proc. CVPR*, 2021.
- [419] G. Bradski, *The OpenCV Library*, *Dr. Dobbs’s Journal of Software Tools* (2000).
- [420] K. Zhang, G. Riegler, N. Snavely and V. Koltun, “Nerf++: Analyzing and improving neural radiance fields.” <https://arxiv.org/abs/2010.07492>, 2020.
- [421] Z. Wang, A. Bovik, H. Sheikh and E. Simoncelli, *Image quality assessment: from error visibility to structural similarity*, *IEEE Transactions on Image Processing* **13** (2004), no. 4 600–612.
- [422] Dawson-Haggerty et al., “trimsh.” <https://trimsh.org/>.
- [423] M. van Heel and M. Schatz, *Fourier shell correlation threshold criteria*, *Journal of Structural Biology* **151** (2005), no. 3 250–262.
- [424] P. B. Rosenthal and R. Henderson, *Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy*, *Journal of Molecular Biology* **333** (2003), no. 4 721–745.

- [425] O. Behnke, K. Kroninger, G. Schott and S.-S. Thomas, *Data analysis in high energy physics: a practical guide to statistical methods*. Wiley-VCH, 2013.
- [426] L. Heinrich, M. Feickert and G. Stark, “pyhf: v0.6.3.”
<https://github.com/scikit-hep/pyhf/releases/tag/v0.6.3>.
- [427] L. Heinrich, M. Feickert, G. Stark and K. Cranmer, *pyhf: pure-python implementation of histfactory statistical models*, *Journal of Open Source Software* **6** (2021), no. 58 2823.
- [428] **ROOT** Collaboration, K. Cranmer, G. Lewis, L. Moneta, A. Shibata and W. Verkerke, *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, *CERN-OPEN-2012-016* (Jan, 2012).
- [429] A. L. Read, *Presentation of search results: the CL_s technique*, *Journal of Physics G: Nuclear and Particle Physics* **28** (sep, 2002) 2693–2704.
- [430] G. Cowan, K. Cranmer, E. Gross and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, *The European Physical Journal C* **71** (feb, 2011).