



FACULTY
OF SCIENCES



Measurements to Improve the Characterization of Four-Top-Quark Production

David Marckx

Universiteit Gent, Faculty of Sciences



A thesis submitted in fulfillment of the requirements for the degree of

Doctor of Sciences: Physics

2022 – 2026

Copyright - auteursrecht

The author and promoter give the permission to use this thesis for consultation and to copy parts of it for personal use. Every other use is subject to the copyright laws, more specifically the source must be extensively specified when using results from this thesis.

De auteur en promotor geven de toelating deze scriptie voor consultatie beschikbaar te stellen en delen ervan te kopiëren voor persoonlijk gebruik. Elk ander gebruik valt onder de beperkingen van het auteursrecht, in het bijzonder met betrekking tot de verplichting uitdrukkelijk de bron te vermelden bij het aanhalen van resultaten uit deze thesis.

Ghent, July 26, 2026

The promotor,

The author,

.....
Prof. Dr. Didar Dobur

.....
David Marckx

Author: David Marckx

Promotor: Prof. dr. Didar Dobur

Copromotor: Dr. Kirill Skovpen

Chairperson of the examination board:

prof. dr. Toon Verstraelen (UGENT)

Members of the examination board:

prof. dr. Gauthier Durieux (UCLouvain)

prof. dr. Ilse De Looze (UGent)

prof. dr. Pedro Silva (CERN)

prof. dr. Juliana Stachurska (UGent)

prof. dr. Roberto Salerno (IP Paris)

Abstract

This thesis describes the main research contributions of the author throughout his PhD, all performed within the CMS collaboration at the Large Hadron Collider. The work is structured around three main topics.

The first and central contribution is a measurement of the inclusive and differential production cross sections of top quark pair production in association with a W boson ($t\bar{t}W$), using the full CMS RUN 2 dataset collected at $\sqrt{s} = 13$ TeV. The analysis targets final states with two same-sign or three leptons, and employs a multivariate strategy based on boosted decision trees (BDTs). Both an inclusive cross section and, for the first time at CMS, differential cross sections are measured as functions of several kinematic observables. The results are compared against state-of-the-art theoretical predictions, providing new constraints on the modelling of this process, which constitutes a dominant background in searches for four-top-quark production.

The second contribution concerns the development of machine learning tools to improve the discrimination of four-top-quark ($t\bar{t}t\bar{t}$) events from background processes, with particular focus on $t\bar{t}t$ production. This includes studies to incorporate improved $t\bar{t}t$ predictions into the analysis, as well as techniques to handle negative event weights arising in higher-order Monte Carlo generators.

The third contribution, which falls outside the scope suggested by the title of this dissertation, describes the author’s involvement in the LHC light-ion physics program. Luminosity calibrations are presented for the oxygen–oxygen (OO), neon–neon (NeNe), and proton–oxygen (pO) collision datasets, using dedicated Van der Meer programs. This is complemented by a discussion of an ongoing measurement of the inelastic pO cross section and its diffractive components, as well as a study of charged particle production rates through particle identification via energy loss measurements in the CMS tracker. These results will help constrain hadronic interaction models relevant to cosmic ray physics.

Keywords: LHC, CMS, top quark, $t\bar{t}W$, $t\bar{t}t\bar{t}$, differential cross section, machine learning, light ions, luminosity calibration

Samenvatting

Deze thesis beschrijft de belangrijkste onderzoeksbijdragen van de auteur tijdens zijn doctoraat, uitgevoerd binnen de CMS-samenwerking aan de Large Hadron Collider. Het werk is gestructureerd rond drie hoofdonderwerpen.

De eerste en centrale bijdrage is een meting van de inclusieve en differentiële productiedoorsneden van top-quarkpaarproductie in associatie met een W-boson ($t\bar{t}W$), gebruik makend van de volledige CMS RUN 2 dataset opgenomen bij $\sqrt{s} = 13$ TeV. De analyse richt zich op eindtoestanden met twee gelijkgeladen of drie leptonen, en maakt gebruik van een multivariate strategie gebaseerd op boosted decision trees (BDTs). Zowel een inclusieve werkzame doorsnede als, voor de eerste keer bij CMS, differentiële werkzame doorsneden worden gemeten als functies van verschillende kinematische observabelen. De resultaten worden vergeleken met state-of-the-art theoretische voorspellingen, en leveren nieuwe beperkingen op de modellering van dit proces, dat een dominante achtergrond vormt in zoekacties naar vier-top-quarkproductie.

De tweede bijdrage betreft de ontwikkeling van machine-learning methoden om de discriminatie van het vier-top-quark ($t\bar{t}t\bar{t}$) proces van achtergrondprocessen te verbeteren, met bijzondere aandacht voor $t\bar{t}t$ productie. Dit omvat studies om verbeterde $t\bar{t}t$ voorspellingen in de analyse te integreren, alsook technieken om negatieve gebeurtenisgewichten afkomstig van Monte-Carlo-

generatoren op hogere orde te behandelen.

De derde bijdrage, die buiten het bereik valt dat de titel van dit proefschrift doet vermoeden, beschrijft de betrokkenheid van de auteur bij het LHC-programma voor lichte-ionen botsingen. Luminositeitskalibraties worden gepresenteerd voor de zuurstof–zuurstof (OO), neon–neon (NeNe) en proton–zuurstof (pO) botsingsdatasets, verkregen via toegewijde Van der Meer-programma's. Dit wordt aangevuld met een bespreking van een lopende meting van de inelastische pO-werkzame doorsnede en haar diffractieve componenten, alsook een studie van geladen-deeltjesproductiesnelheden via deeltjesidentificatie op basis van energieverliesmeting in de CMS-tracker. Deze resultaten zullen helpen om hadronische wisselwerkingsmodellen relevant voor kosmische-stralennatuurkunde te verbeteren.

Trefwoorden: LHC, CMS, top-quark, $t\bar{t}W$, $t\bar{t}t\bar{t}$, differentiële
werkzame doorsnede, machine learning, lichte ionen,
luminositeitskalibratie

Acknowledgements

I am deeply grateful to my parents for bringing me into this world, and to the doctors and staff at Inselspital in Bern for their skill, care, and commitment in keeping me here.

In all seriousness, I have at times found it difficult to write this acknowledgements section, as I do not necessarily see this thesis itself the milestone worthy of recognition. Rather, I want to acknowledge everyone for the past four joyous years.

First, I would like to thank Didar for believing in my potential as a young master's student. That trust set me on a path that has shaped my life in countless ways. While I may not always have been the easiest student—often absorbed in my own research obsessions and eager to attend every conference—I am sincerely grateful for the opportunities and support you have provided me. I really look forward to working together over the coming years!

I would also like to thank the members of the jury for their time and effort in reading this thesis. For those just beginning: good luck, and my apologies in advance. I am also grateful to everyone who has collaborated with me, with a special thanks to the team in Oviedo and to Michael at CERN.

My thanks also go to my colleagues over these years, especially my office mates Liam, Luka, Niels, and Dati. You guys are the reason I went every day to that asbestos infested, radioactive cesspit of doom, and kept my sanity (partially) intact along the way. I am also grateful to Jules and Ameline for the fun climbing sessions and friendship we developed. A special mention goes to the office Germans Jan and Joscha for keeping the group organised and for all the activities you organised. I'd also like to thank Kirill for his vast knowledge and for reminding me that there is always something exciting to discover in

every corner of physics. There are many others I cannot name individually, but I am truly thankful for all the shared moments.

I would also like to thank my friends, especially Wout and Tippi, for all the memories since our bachelor's years. Jacob and the "Whiskey guys" from Geneva also deserve a mention for the many adventures together. To all other friends not named here: I am equally grateful for your presence in my life.

Finally, I would like to thank my girlfriend, Mercedes. I'm so happy we found each other and thank you for pulling me through the harder times!

I end these acknowledgements as I did in my master's thesis: to conclude this extensive but far from complete list, I wish to thank my parents for their love and encouragement, and for supporting me at every step along the way. Everything I have achieved, and will ever achieve, I owe to you.

List of publications and presentations

Peer reviewed publications

- CMS Collaboration, "Measurements of $t\bar{t}W$ differential cross sections and the leptonic charge asymmetry at $\sqrt{s}=13$ TeV", JHEP 03 (2026) 083 [1], [LINK](#), (primary contribution)

Non-peer-reviewed publication

- ECFA ECR Panel, "Results of the analysis of a survey for young scientists on training quality in HEP instrumentation software and machine learning" [2], [LINK](#), (primary contribution)
- P. Van Mechelen *et al.* , "National input to the ESPPU" [3], [LINK](#), (primary contribution)

Ongoing/Under internal review

- "pO, OO and NeNe luminosity calibrations"¹, stage: "Publication procedure not started", (primary contribution)

¹used by HIN-25-(008, 009, 010, 014, 015, 016, 019, 020, 021, 023), HIN-26-(004, 005, 009, 010, 011, 012, 013)

-
- "Measurement of the inelastic cross section and charged particle production rates in proton-Oxygen collisions at 9.62 TeV", stage: "CADI line" (HIN-26-012), (primary contribution)
 - "3-4 top quark analysis with Run2+Run3 data", stage: "Publication procedure not started", (secondary contribution: Machine learning, dedicated MC sample generation.)
 - "Differential measurement of $t\bar{t}$ c and $t\bar{t}$ b production - Run 2", stage: "CADI line" (TOP-26-003), (minor contribution: dedicated MC sample generation.)

Conference proceedings

- "Towards a differential $t\bar{t}W$ cross section measurement at CMS" [4], Proceedings TOP2024 (Peer reviewed)
- "First differential $t\bar{t}W$ cross section measurement by CMS" [5], Proceedings Rencontres de Moriond 2025 EW
- " $t\bar{t}W$ and $t\bar{t}Z$ measurements from CMS" [6], Proceedings LHCP2025 (Peer reviewed)

Presentations and posters

- "*Inclusive and differential $t\bar{t}W$ cross section measurement*", Young scientist talk at TOP2024, Sep 25, 2024, [LINK](#)

-
- *"First differential ttW cross section measurement by CMS"*, Young scientist talk at Moriond/EW2025, March 29, 2025, [LINK](#)
 - *" $ttW+ttZ$ measurements from CMS"*, parallel talk at LHCP2025, May 8, 2025, [LINK](#)
 - *"First differential ttW cross section measurement by CMS"*, Poster at LHCP2025, May 8, 2025, [LINK](#)
 - *"Impressions from young scientists"*, plenary talk at RECFA visit to Belgium, Sep 12, 2025, [LINK](#)

Further CMS internal documentation

- CMS Collaboration, "Differential cross section measurements of ttW production at $\sqrt{s} = 13$ TeV". CMS-AN-23-077. (CMS internal) [LINK](#) (CMS internal note, restricted access)
- CMS Collaboration, "Charged particle production rates in proton-Oxygen events.". CMS-AN-25-091. (CMS internal) [LINK](#) (CMS internal note, restricted access)

Contents

Copyright	i
Abstract	v
Samenvatting	vii
Acknowledgements	ix
List of publications and presentations	xiii
Lists	xxxvii
List of Figures	xxxvii
List of Tables	xxxix
1. Introduction	1
2. The Standard Model of particle physics	7
2.1. The standard model of particle physics	7
2.1.1. Strong interactions	9
2.1.2. Electroweak interactions	10
2.2. Shortcomings of the standard model	15
2.3. From Feynman diagrams to cross-sections and Monte Carlo predictions	16
2.3.1. Hadronization and parton shower	19
2.3.2. Negative weights in ME generators	20
2.3.3. Interference between processes at NLO	22
2.3.4. Multijet merging and matching	24

2.4. Luminosity: a short overview	25
3. Theoretical and experimental status in top quark physics	31
3.1. Top quark properties	31
3.1.1. Mass	31
3.1.2. Charge asymmetry	33
3.1.3. Spin correlation/polarization	35
3.1.4. CKM matrix	35
3.2. Production of top quarks at the LHC	36
3.3. Cross-sections of top quark processes	37
3.4. The discovery of four top quark production	41
3.4.1. Analysis regions	41
3.4.2. ML approach	42
3.4.3. Results	42
3.4.4. Interplay between top pair production in association with a W boson and four-top production	43
3.4.5. Interplay between three-top and four-top production	45
4. The Large Hadron Collider and CMS experiment	47
4.1. From source to collision	48
4.1.1. Protons	52
4.1.2. Heavy ions	53
4.1.3. Beam property monitoring devices	54
4.1.4. Luminosity	58
4.2. The CMS detector: a short overview	60
4.2.1. CMS coordinate system	60
4.2.2. CMS subdetectors	61
4.2.3. Forward detectors	69
4.3. Upgrades for High-luminosity LHC	71
4.3.1. iRPCs	73
4.3.2. GEMs	76

4.4.	Trigger System	77
4.4.1.	Level-1 Trigger	79
4.4.2.	High-Level Trigger	80
4.4.3.	ZeroBias and Minimum Bias Triggers	80
4.4.4.	Scouting and parking	81
4.5.	Particle flow reconstruction	81
4.5.1.	Track reconstruction	82
4.5.2.	Muon reconstruction	83
4.5.3.	Vertexing	84
4.5.4.	Electron and photon reconstruction	84
4.5.5.	Jet reconstruction	85
4.5.6.	Missing transverse momentum	87
5.	Top quark pair production in association with a W boson	89
5.1.	Introduction	89
5.1.1.	Theoretical difficulties	90
5.1.2.	State-of-the-art	91
5.1.3.	Charge asymmetry	94
5.1.4.	The analysis design	94
5.2.	Data and simulated samples	95
5.2.1.	Data	95
5.2.2.	Simulated samples	95
5.3.	Object selection	98
5.3.1.	Jets and Missing transverse energy	98
5.3.2.	Leptons	99
5.4.	Event selection	101
5.4.1.	Trigger	102
5.4.2.	Signal regions	103
5.5.	Background estimation and control regions	106
5.5.1.	Charge misidentified electrons	106
5.5.2.	Nonprompt leptons	113
5.5.3.	Control regions and validation regions	115

5.6.	Event-level MVA	123
5.6.1.	Input variables	125
5.6.2.	Training setup and hyperparameter optimization	125
5.7.	Calibrations and measurement uncertainties	129
5.7.1.	Experimental calibrations and uncertainties	129
5.7.2.	Theoretical uncertainties	136
5.8.	Maximum likelihood fit and statistical inference	139
5.8.1.	Theoretical background	139
5.8.2.	Inclusive fit distributions	141
5.8.3.	Differential likelihood measurements	142
5.8.4.	Differential fit distributions	147
5.9.	Robustness checks	150
5.9.1.	Injection tests	150
5.9.2.	The Uniform Gradient Boosting with Flatness Loss method	157
5.9.3.	Self-consistency of differential measurements	161
5.10.	Results	163
5.10.1.	Inclusive cross section measurement	164
5.10.2.	Differential cross section measurement	167
5.11.	Conclusions and outlook	176
6.	Machine learning models to discern 4 top events from other processes	179
6.1.	tttj and tttW predictions at NLO QCD + LO EWK	180
6.1.1.	Scale choice	181
6.1.2.	PDF dependence	184
6.1.3.	Interference terms	185
6.2.	Treatment of negative weights	187
6.2.1.	Effects of negative weights on distributions	189
6.2.2.	The neural reweighter and refiner approaches	191
6.2.3.	Results	193

6.3.	Training of the classifier model	196
6.3.1.	Architecture	196
6.3.2.	Training	202
6.3.3.	Performance	206
6.3.4.	Conclusions and outlook	209
6.3.5.	Improved predictions of <i>ttt</i> production	210
7.	Oxygen collisions program at CMS	213
7.1.	Current hadronic models in the field of cosmic ray physics	214
7.2.	Data taking: the pO, OO and NeNe programs	217
7.3.	Luminosity callibrations for the OO and NeNe program	219
7.3.1.	The Van der Meer scan	221
7.3.2.	BSRL: Ghosts and Satellites	221
7.3.3.	The superseparation scan	224
7.3.4.	Beam-Beam effects	226
7.3.5.	arcBPMs and DOROS	228
7.3.6.	Other scans	229
7.3.7.	Consistency, stability and linearity	230
7.4.	pO, NeNe and futher OO calibrations	233
7.4.1.	Ghosts and satellites	234
7.4.2.	Length scale	238
7.5.	Towards a measurement of the inelastic cross section and charged particle production rates in pO collisions.	245
7.5.1.	Introduction	245
7.5.2.	Samples	245
7.5.3.	Detector validation	247
7.5.4.	HF	252
7.6.	Outlook towards planned future measurements	255
7.6.1.	A "dE/dx" measurement	257
7.6.2.	Inelastic cross section	262

7.6.3. Pion exchange	266
7.7. Outlook	268
8. Final outlook	271
8.1. Outlook beyond this thesis	274
Bibliography	277
A. Appendix: ECFA ECR panel	299
A.1. Introduction	299
A.2. Accessibility and quality of schools	303
A.3. Machine Learning	306
B. Appendix: Sample lists $t\bar{t}W$	315
B.1. Data and simulated samples	315
B.1.1. Data	315
B.1.2. Simulation	316
C. Appendix: Triggers	321
C.0.1. triggers used in the $t\bar{t}W$ analysis	321
C.0.2. Orthogonal triggers for the $t\bar{t}W$ analysis	324
C.0.3. triggers used in the fake-rate measurement for the $t\bar{t}W$ analysis	325

List of Figures

2.1.	A full list of the elementary particles described by the standard model of particle physics. Figure from ref. [7].	8
2.2.	Feynman diagrams representing the quantum chromodynamics interactions with quark-gluon couplings (left) and gluon self-couplings (middle and right).	9
2.3.	Feynman diagram representing the coupling of charged particles to photons predicted by QED. . .	10
2.4.	Feynman diagram representing the weak interactions of SM particles.	11
2.5.	Longitudinal gauge-boson scattering cross section as a function of center-of-mass energy. The diagrams added by the introduction of the Higgs cancel the problematic diagrams of the gauge bosons. Figure from Ref. [16].	14
2.6.	Feynman diagram representing the interactions with the Higgs field for massive particles (top) and the Higgs self-interactions (bottom).	15
2.7.	Example (N)(N)LO diagrams for top pair production.	18
2.8.	Schematic representation of a top pair production event at the LHC, as modeled by a PS model. Figure from Ref. [28].	20
2.9.	NLO diagrams of $t\bar{t}$ production include non-resonant modes (left) but also contain resonant $t\bar{t}$ production (right).	23
2.10.	Double counting by ME when multijet merging a $t\bar{t}$ sample.	25

3.1. Origin of charge asymmetry in hadroproduction of top quarks: interference of initial-state (top left) with final-state (top right) real emission diagrams and interference of the box (bottom left) and Born (bottom right) diagrams.	34
3.2. Single top production mode diagrams at LO: t-channel (left), tW (middle) and s-channel (right). .	37
3.3. Illustrative top pair production in association with bosons (top) and single top quark production in association with bosons (bottom) diagrams.	38
3.4. Illustrative $t\bar{t}t\bar{t}$ (left), $tttj$ (middle), $tttW$ (right) production diagrams.	39
3.5. Summary of production cross section measurements by the CMS collaboration involving top quarks. Measurements performed at different collision energies are marked by unique symbols and the coloured bands indicate the combined statistical and systematic uncertainty of the measurement. Grey bands indicate the uncertainty of the theory predictions. Figure from Ref. [59].	40
3.6. Schematic representation of the event selection and categorization in the RUN 2 $t\bar{t}t\bar{t}$ analysis. Figure from Ref. [60].	41
3.7. Measured $t\bar{t}t\bar{t}$ production cross section compared to the Standard Model prediction and results from ATLAS and CMS. Figure from Ref. [61].	44
3.8. Impact plot for the $t\bar{t}t\bar{t}$ cross section fit. The plot shows the 20 uncertainties that have the most impact on the measurement. Figure from Ref. [60]. .	45
3.9. Two-dimensional negative log-likelihood scan of the $t\bar{t}t\bar{t}$ and ttt production cross sections. Figure from Ref. [62].	46
4.1. Cern accelerator complex	49

4.2.	The concept of the three crucial components in a circular accelerator: The dipole magnets, quadrupole magnets and RF cavities. Modified from [65].	51
4.3.	A depiction of charge-exchange injection. Figure taken from [67].	53
4.4.	Schematic drawing of the arcBPM buttons. Figure modified from Ref. [70].	56
4.5.	Cumulative luminosity delivered to CMS during stable beams for pp collisions at nominal center-of-mass energy versus day. Figure from Ref. [71].	58
4.6.	Distribution of the average number of interactions per crossing (pileup) for pp collisions. Figure from Ref. [71].	59
4.7.	The coordinate system applied in CMS. The positive y direction is upwards, the x-axis points to the inside of the LHC while the z-axis is pointed in the direction of the beam towards the Jura. Figure modified from Ref. [73].	61
4.8.	The CMS detector. Figure from Ref. [76].	62
4.9.	The Pixel tracker as it was installed in 2016 (Original) and after the upgrade in 2017 (Phase-1 upgrade). This tracker will be upgraded for the CMS Phase-2 upgrades. Figure from Ref. [74].	63
4.10.	The layout of the strip tracker. The orange lines show the placement of the single-sided silicon strip modules, and the blue lines the stereo modules. The pixel detector layers are shown in green. Figure from Ref. [74].	64
4.11.	Left: sensitivity to additional spacial coordinate. Right: Ambiguity of track locations with multiple hits.	65

4.12. Layout of the muon system inside the CMS experiment at the start of Run 3. Since, several iRPC chambers have been installed to validate the performance of these designs for Phase II. Figure from Ref. [74].	67
4.13. Examples of the gas leak test and spacer test for gaps to be used in RPC chambers. Left: a good gap, with acceptable pressure drop (the rejection threshold is 0.4 mbar / 10 mins) and no apparent broken spacers. Right: a rejected gap, with acceptable pressure drop but two broken spacers.	75
4.14. Dark currents as a function of the applied voltage for a gap showing too high currents before flushing treatment (left) and a good gap (right).	76
4.15. A stack of 4 Modules assembled at B904 in CERN.	77
4.16. Illustrative plots of leak test (QC 3b, top) and electronic test (QC 7, bottom). The peak structures in the noise are due to the routing geometry of the strips to the 6 VFATs that are connected to the Optohybrid/ASIAGO.	78
4.17. b-tagging ROC curves. UParTv2 shows improved performance for both c and light jet rejection.	87
5.1. The LO $t\bar{t}W$ production Feynman diagrams.	89
5.2. Examples of NLO (left), NNLO (middle), and NLO EWK (right) Feynman diagrams contributing to the $t\bar{t}W$ process. All diagrams introduce sizeable corrections to the cross section of this process.	91
5.3. Left: double-loop diagram, for which the amplitude is yet unknown. Right: the convergence of both approximations used to estimate the double-loop contribution. The top panel shows the convergence to the exact calculation at NLO QCD. The bottom panel shows the convergence at NNLO [82], where the exact calculation is of course not available.	92

5.4.	Two real emission $t\bar{t}W$ diagrams where the extra parton is protected from IR divergencies.	93
5.5.	Left: merging scale dependence for $t\bar{t}W$ production with the old FxFx@1j merging scheme [83]. Right: merging scale dependence for the new FxFx scheme [84].	93
5.6.	Leading lepton p_T (upper left), subleading lepton p_T (upper right), scalar p_T sum of selected jets in the event (lower left), and the number of selected jets in the event (lower right), in the two-lepton signal selection for data (points) and predictions (filled histograms) after the fit to the data in the SR and CR. The lower panels show the ratio of the data to the sum of the postfit predictions (points) and the ratio of the data to the prefit predictions (red line). The hatched (filled) band corresponds to the total uncertainty in the postfit (prefit) prediction. Events that exceed the range of the plot are included in the last bin.	105
5.7.	Scalar p_T sum of selected jets in the event (left), number of selected b-jets (right), in the trilepton signal selection for data (points) and predictions (filled histograms) after the fit to the data in the SR and CR. The lower panels show the ratio of the data to the sum of the postfit predictions (points) and the ratio of the data to the prefit predictions (red line). The hatched (filled) band corresponds to the total uncertainty in the postfit (prefit) prediction. Events that exceed the range of the plot are included in the last bin.	106
5.8.	Illustration on how the charge of an electron can be misreconstructed in the CMS tracker.	107
5.9.	Chargeflip maps, as measured in MC events for the data-taking years used in the $t\bar{t}W$ analysis.	109

5.10. Some distributions observed in the closure tests in 2018 simulation samples for the charge misidentification method. Upper row: DY, lower row: $t\bar{t}$	110
5.11. Some distributions observed in the closure tests in data for the charge misidentification method for all data-taking years. Closure is shown before (after) applying the year dependent scale factors in the row above (below).	112
5.12. ROC curves showing the performance of the leptonMVA in the p_T ranges $10 \text{ GeV} < p_T < 25 \text{ GeV}$ (left) and $p_T > 25 \text{ GeV}$ (right). Figure from ref. [60].	113
5.13. Example sketch of a nonprompt lepton.	114
5.14. Schematic representation of the event selection and categorization.	116
5.15. $ \eta $ of the leading lepton (left), p_T of the leading lepton (right), in a CR enriched with charge-misidentified leptons and additional jets after the fit to the data in the SR and CR.	117
5.16. $ \eta $ (top left) and p_T (top right) of the subleading lepton, number of jets (bottom left) and number of b-jets (bottom right) in a CR enriched with nonprompt leptons after the fit to the data in the SR and CR.	118
5.17. Scalar sum of lepton p_T 's (left) and leading jet p_T (right) in a VR enriched with nonprompt leptons using jet requirements for orthogonality.	119
5.18. Example sketch of a photon conversion in the tracker.	119
5.19. η of the leading lepton (left) and number of jets (right) in a VR enriched with photon conversion events.	120
5.20. Number of selected (b-tagged) jets for events passing the trilepton CR requirements after the fit to the data in the SR and CR.	122

5.21. Number of selected Z candidates and b-tagged jets for events passing the fourlepton CR requirements after the fit to the data in the SR and CR.	123
5.22. Left: ROC curve for the BDT model. ‘Other backgrounds’ indicates the background samples that pass the signal region event selection but that are not directly trained on. Right: Training and validation loss values during model training.	127
5.23. Event BDT output score distribution in the dilepton SR after the fit to the data in the SR and CR.	128
5.24. Event BDT output score distribution in the Charge-flip (left) and NP (right) CRs after the fit to the data in the SR and CR.	128
5.25. Correlation of the number of reconstructed primary vertices (color) and the pileup (x -axis).	131
5.26. Response matrix for the H_T observable. The horizontal axis corresponds to the particle-level quantity, while the vertical axis shows the reconstructed observable. An ideal response matrix is fully diagonal. Quality metrics are shown in the lower panels.	148
5.27. Fitted distribution in the 2ℓ SS SR for the $p_T^{\text{jet } 1}$ fit. The binning in the BDT score provides signal sensitivity, while the hyperbins provide the detector-level binning of the response matrix. The $t\bar{t}W$ signal is split in different components that define the particle-level bins.	149

5.28. Results of the pseudodata injection tests for selected differential measurements. Top left: unfolded jet multiplicity distribution using pseudodata in which the signal is reweighted as a function of jet multiplicity. Top right: unfolded H_T distribution using pseudodata reweighted as a function of the leading jet p_T . Bottom left: unfolded ΔR between leptons using pseudodata reweighted as a function of H_T . Bottom right: unfolded subleading jet p_T distribution using pseudo-data reweighted as a function of H_T	153
5.29. Alternative signal number of jets distribution, based on the activation of several SMEFT operators. . . .	155
5.30. Results of the SMEFT pseudodata injection tests for selected differential measurements. Top left: unfolded jet multiplicity distribution using pseudodata with C_{tq}^1 operator. Top right: unfolded H_T distribution using pseudodata with C_{tW} operator. Bottom left: unfolded ΔR between leptons using pseudodata with $C_{Qq}^{(1,8)}$ operator. Bottom right: unfolded subleading jet p_T distribution using pseudo-data with $C_{Qq}^{(3,8)}$ operator.	156
5.31. Optimization of the α parameter. Lower values show that the shape is not well corrected, while $\alpha = 0.8$ shows good decorrelation.	159
5.32. Results of the $C_{Qq}^{(3,8)}$ operator pseudodata injection tests with the uGBFL BDT. Left: leading b-tagged jet pseudorapidity. Right: subleading lepton pseudorapidity.	160
5.33. Fiducial $t\bar{t}W$ cross sections extracted from the differential fits, normalized to the result obtained from the inclusive fit. All measurements, with the exception of the b-jet multiplicity observables, yield consistent fiducial $t\bar{t}W$ cross sections.	161

5.34. Comparison of the pulls for a selection of nuisance parameters across all differential measurements. The fitted statistical models are found to be mutually consistent within uncertainties.	162
5.35. Impact plot for the differential measurement as a function of $\Delta R(\ell_1, \ell_2)$	163
5.36. Inclusive fit results split per year.	165
5.37. Inclusive fit results split per lepton flavor channel. Red datapoints are results from previous analysis.	166
5.38. Inclusive fit results per lepton sign	166
5.39. Two Dimensional scan of the $t\bar{t}W$ signal strengths based on lepton charge. light-colored contours represent results from the previous analysis.	167
5.40. Absolute (upper row) and normalized (lower row) differential cross section measured as a function of the jet multiplicity (left) and H_T (right), using the MVA-based and the counting method.	169
5.41. Absolute (upper row) and normalized (lower row) differential cross section measured as a function of the leading jet p_T (left) and leading jet $ \eta $ (right), using the MVA-based and the counting method.	170
5.42. Absolute (upper row) and normalized (lower row) differential cross section measured as a function of the subleading jet p_T (left), and subleading jet $ \eta $ (right), using the MVA-based and the counting method.	171
5.43. Absolute (upper row) and normalized (lower row) differential cross section measured as a function of the leading (left) and subleading (right) lepton p_T , using the MVA-based and the counting method.	172
5.44. Absolute (upper row) and normalized (lower row) differential cross section measured as a function of the maximum $ \eta $ of the selected leptons (left) and the sum of their p_T , using the MVA-based and the counting method.	173

5.45. Absolute (upper row) and normalized (lower row) differential cross section measured as a function of the invariant mass of the leptons (left) and the $\Delta\eta(\ell_1, \ell_2)$, using the MVA-based and the counting method.	174
5.46. Absolute (upper row) and normalized (lower row) differential cross section measured as a function of the $\Delta R(\ell_1, \ell_2)$ and the $\Delta R(\ell_1, jet)_{\min}$, using the MVA-based and the counting method.	175
6.1. LO terms contributing to the $tttj$ (left) and $tttW$ (right) processes.	179
6.2. Cross sections for the different ttt NLO QCD terms, as predicted by the original setup provided by G. Durieux et al., the CMS implementations at 13 TeV and 13.6 TeV.	181
6.3. Effects of scale variations around the central scale choice $H_T/2$ on the jet spectrum of the inclusive $tttW$ sample. μ_R variations are represented by the color, while μ_F variations are represented by the line style.	183
6.4. Ratio $\sigma_{tttj+}/\sigma_{tttj-}$ for different PDF sets.	184
6.5. Valence PDFs as accessible via LHAPDF [98].	186
6.6. The BCE loss for positive weight events (left) and negative weight events (right).	188
6.7. Initial state parton fractions for weighted events.	190
6.8. Jet and b-jet multiplicity distributions for different treatments of negatively weighted events. Top left: jet multiplicity distribution. Top right: b-jet multiplicity distribution. Bottom left: ratio of $t\bar{t}t\bar{t}$ to ttt as a function of the jet multiplicity. Bottom right: ratio of $t\bar{t}t\bar{t}$ to ttt as a function of the b-jet multiplicity.	191
6.9. Some training performance plots for $t\bar{t}t\bar{t}$ (top), $tttj$ (middle) and $tttW$ (bottom).	194

6.10. Some training performance on physical distributions for $t\bar{t}t\bar{t}$, $t\bar{t}tj$ and $t\bar{t}tW$	195
6.11. Some training performance on number of jets distributions for $t\bar{t}tj$ vs $t\bar{t}t\bar{t}$ (left) and $t\bar{t}tW$ vs $t\bar{t}t\bar{t}$ (right).	196
6.12. A graph representation of an event.	197
6.13. Structure of the Event level neural network architecture.	201
6.14. ROC curves for the GNN trained with batch sizes 64 and batch size 1024. The RUN 2 BDT model is shown as reference.	205
6.15. Loss curve throughout the training of the SR3 ℓ region with RUN 3 samples.	206
6.16. Loss curve throughout the training of the SR2 ℓ region with RUN 3 samples. The best performance model is studied further below.	207
6.17. Output scores of the most performant models in the SR2 ℓ (left) and SR3 ℓ (right). The output nodes target the $t\bar{t}t\bar{t}$ process, $t\bar{t}t$ events and $t\bar{t}(X)$ backgrounds.	208
6.18. Output scores of the most performant models with the $t\bar{t}t$ process split in the $t\bar{t}tj$ and $t\bar{t}tW$ components.	209
6.19. State-of-the-art performance before these studies. Figure from Ref. [107].	210
7.1. The measured mean-logarithmic mass of cosmic rays as a function of their energy. Measurements that use a different mass-sensitive variable in the analysis (variables can roughly be grouped in X_{max} or N_μ) get vastly different results. [111]	215

7.2. Collision systems currently used to constrain hadronic interaction models at the LHC (red), together with the new collision systems collected in July 2025 (blue). The yellow band indicates the collision systems most relevant for cosmic-ray interactions in the atmosphere.	216
7.3. LHC schedule around the light ions program. . . .	217
7.4. Beam positions measured by DOROS during the first Van der Meer scan of Fill 10802. The horizontal (x) and vertical (y) scan coordinates are shown in the left and right panels, respectively.	221
7.5. Pedestal subtracted BSRL data for BCID 1786 in Fill 10802 and BCID 1 in Fill 10813.	222
7.6. Effect of ghost and satellite corrections on the visible cross section of the HFET luminometer for the OO and NeNe fills.	223
7.7. Beam positions measured by DOROS during the superseparation scan of Fill 10802. The horizontal (x) and vertical (y) scan coordinates are shown in the left and right panels, respectively.	224
7.8. Top: observed rates during the superseparation scan. Bottom: observed rates during a headon period in the VdM program.	225
7.9. Corrections to the beam-beam displacement due to beam-beam deflection for Fill 10802 (left) and Fill 10808 (right).	227
7.10. Data as measured by the arcBPMs during the pO program.	228
7.11. Horizontal linear orbit drift values as measured by DOROS and arcBPMs.	229
7.12. Diagonal scan in the VdM program of Fill 10808. .	230
7.13. cLS scan in the VdM program of Fill 10782. . . .	231
7.14. Rate asymmetry observed in pO VdM scans for BCID 172.	234

7.15. Schematic of satellite bunch contributions to the asymmetry.	235
7.16. Bucket populations for the full LHC RF system (top) and buckets surrounding BCID 172 (bottom).	236
7.17. Correlation between satellite contributions and the observed asymmetry (left) and extracted visible cross section (right).	237
7.18. PLT+ and PLT- rates in VdM scan of the pO program.	238
7.19. Beamspot positions for scan steps 6 (left) and 7 (right) of the cLS scan.	240
7.20. Step sizes of the horizontal (left) and vertical (right) pO cLS scans.	240
7.21. Length scale scan results for horizontal (X, left column) and vertical (Y, right column) scans under different orbit correction schemes.	242
7.22. Length scale results for the pO run.	243
7.23. Beam position residuals after multiple corrections. Left: orbit drift, middle: orbit drift+Beam-beam deflection, right: orbit drift+Beam-beam deflection+ length scale.	243
7.24. Two-dimensional distributions of track pseudorapidity versus azimuthal angle in run 393952, for three transverse-momentum intervals: $0.2 < p_T < 0.5$ GeV (left), $0.5 < p_T < 1.0$ GeV (middle) and $p_T > 1.0$ GeV (right).	248
7.25. Removal of the FSCs after the NeNe run.	250
7.26. The FSC stations, as used in RUN 1. The stations #2 and #3 installed for the pO run are highlighted.	250
7.27. ZDC neutral energy deposits for the proton side (left) and oxygen side (right).	251
7.28. HF calorimeter energy deposits measured on the proton-going and oxygen-going sides in data.	252
7.29. Two-dimensional distributions of HF activity in pseudorapidity and azimuth for run 393952.	253

7.30. Average HF energy deposits per lumisection during the superseparation scan. A preceding VdM scan is also visible.	254
7.31. Energy distributions of the leading (blue) and sub-leading (yellow) particle-flow candidates in HF for the two noise-enriched datasets. The 99% and 99.9% background rejection thresholds are indicated in red.	255
7.32. HF calorimeter energy deposits measured on the proton going side for "electromagnetic" (id 7) and "hadronic" particle flow candidates.	256
7.33. dE/dx bands for pions (top left), kaons (top right), protons (bottom left) and deuterons (bottom right), as predicted by EPOS-LHC.R.	258
7.34. Predicted particle production multiplicities for kaons (top left), protons (top right), deuterons (bottom left). Predicted kaon production rate as a function of track multiplicity (bottom right).	259
7.35. Response matrix used for the unfolding of the kaon multiplicity distribution.	261
7.36. Simulated neutral energy distribution in ZDC– (proton side), broken down by the production processes. The π -exchange signal produces a distinct high-energy structure that overlaps with the ND core formation background.	267
A.1. The respondents' career stage distribution (a). Name of the experiment in which the respondents take part (b).	302
A.2. The respondents' fields of research distribution (a). Hot topics indicated by respondents (b).	303
A.3. Knowledge of available training programs (a). Have participants attended such programs? (b)	305

A.4. Coverage of topics in attended programs (a). Participant satisfaction with training programs (percentage) (b).	306
A.5. Usage of ML software (a) and knowledge level (b).	307
A.6. The main use cases for ML (a). The ML libraries that are most often used by respondents (b).	308
A.7. Method of learning ML.	308
A.8. The type of training considered most effective for ML (a). Subjects that are of interest to respondents or were underexplored in previous schools (b).	309
A.9. The average preferred composition of a course on ML.	310
A.10. Usage of detector simulation software tools (a) and knowledge level (b).	311
A.11. Method of learning detector simulation software tools (a). Previous participation to a training program for detector simulation.	311
A.12. The type of training considered most effective for detector simulation (a). Simulation tools that are of interest to respondents or were underexplored in previous schools (b).	312
A.13. The average preferred composition of a course on detector simulation.	313

List of Tables

5.1. Structure of LO and NLO contributions with naive expectations and actual corrections. Green order terms are included in the signal sample, red order terms are excluded.	97
---	----

5.2. Summary of jet identification criteria used in this analysis. Values in color are the requirements in the Jet ID.	99
5.3. Summary of muon identification criteria used in this analysis. Values in parentheses indicate thresholds applied when the leptonMVA requirement is not satisfied.	101
5.4. Summary of electron identification criteria used in this analysis. Values in parentheses indicate thresholds applied when the leptonMVA requirement is not satisfied. Values separated by slashes correspond to 2016PreVFP / 2016PostVFP / 2017 / 2018.	102
5.5. Comparison of dilepton and trilepton signal region selections.	104
5.6. Year dependent scale factors for chargemisID rate.	112
5.7. Sample sizes available for training.	125
5.8. Input variables and corresponding p-values.	126
5.9. Overview of the categorization strategy for the different likelihood fits used in the analysis.	142
5.10. Definition of the fiducial phase space used in the differential measurement.	144
5.11. Observables of interest in the differential cross-section measurements. Highlighted variables are presented in the paper.	146
5.12. Event weight parametrizations used in the analysis.	151
5.13. Used Wilson coefficients.	154
6.1. Fraction of negatively weighted events for the different signal samples.	187
6.2. Input features provided to the graph neural network. Empty jet and lepton nodes are padded with fixed values when fewer than ten jets or three leptons are present in the event.	198

7.1. Target and collected (online calibration) luminosities and main beam parameters.	219
7.2. Overview of the scans for Fills 10802, 10808 and 10813.	220
7.3. Summary of luminosity calibration uncertainties for the OO and NeNe collision programs.	233
7.4. Overview of the pO Fill 10782 scan program.	234
7.5. Summary of luminosity calibration uncertainties for the OO, NeNe and pO collision programs.	244
7.6. Monte Carlo samples and their production cross sections.	247
7.7. Visible cross sections for the baseline preselection, inclusive HF selections, vertex–HF overlap regions, and the main analysis categories in run 393952. No background subtraction or reconstruction-efficiency correction is applied.	264
B.1. Simulated signal and background samples and their effective cross sections	319
B.2. Simulated signal and background samples and their effective cross sections (part 2)	320
C.1. Overview of 2016 triggers used to select good event for this analysis.	321
C.2. Overview of 2017 triggers used to select good event for this analysis.	322
C.3. Overview of 2018 triggers used to select good event for this analysis.	323
C.4. Reference triggers for the orthogonal trigger method.	324
C.5. Overview of triggers used for the fake rate measurement	325

1 Introduction

“Last chapter to be written.” For a long time, that was the only sentence that existed in this introduction. Not because I did not know what to write, but more because I was unsure how to write it. It felt misplaced to begin a thesis by glorifying work that I happened to perform over the past few years. Equally, I saw little value in forcing the reader through a philosophical reflection on science written by someone who is still very much a newcomer to both the field and the wider world.

In reality, I have been only a small cog in a machine far larger than myself. Every result presented in this thesis is the product of thousands of people: physicists, engineers, technicians, accelerator experts, and countless others whose names rarely appear on the author list of a paper. The specific contributions described in these pages will likely be improved upon, superseded, or perhaps even forgotten in the decades to come.

This raises an obvious question: why should anyone read nearly three hundred pages describing it?

I do not have a particularly convincing answer. While writing this section, I found myself drawn to include a section of a book that has strongly resonated with me throughout the writing of this thesis. I add an excerpt below:

""WHAT IS IMPORTANT IN LIFE? It is a question we should all give thought to. The answer is not taught in school, and it is not as easy as it may seem, for a superficial answer is not acceptable. To discover the real truth you have to know

yourself. Then you have to be honest with yourself. Then you have to respect and accept yourself. For me, these were all tough tasks.

*I had gone through college and into academia in a hurry, wanting to rush ahead with my work, to prove to the world that I had been alive, and that it had mattered. That was an external focus to life. That was Murray's way. To accomplish and impress. To be an important person, and a leader. It was the classical path. The traditional one. It seemed to be an obvious and worthy goal. I had accepted it without second thought. **But for me, it was like chasing a rainbow. Even worse, it was like chasing other people's rainbows. Rainbows whose beauty I didn't really see.***

Feynman's example caused me to rethink mine. He didn't seek the leadership role. He didn't gravitate to the sexy "unified" theories. For him satisfaction in discovery was there even if what you discover was already known by others. It was there even if all you are doing is re-deriving someone else's result your own way. And it was there even if your creativity is in playing with your child. It was self-satisfaction. Feynman's focus was internal, and his internal focus gave him freedom.

Our culture is a culture that, by Feynman's characterization, is Greek. It is a culture of logic and proof, rules and order. In our culture people who live their lives like Feynman are considered eccentric, for Feynman was a Babylonian. For Feynman, both physics and life were ruled by intuition and inspiration, and a disdain for rules and customs. He ignored the conventional methods of physics, and invented his own, his sum over paths and his Feynman diagrams. He also ignored academic culture and invented his own. And if his behavior was not approved of, well, what did he care what other people thought?

I chose Feynman's way. A lot of people aren't lucky enough to feel a passion for any particular endeavor like my writing, or else, like my immigrant father, are too occupied with mere survival to have any choices. Especially after my death scare, if I had a choice, I did not want to squander it. I resolved that, as long as I could, I would spend my limited time in life pursuing goals that moved me, whether or not others found them worthy. I resolved to never lose sight of the beauty in physics – and life – whatever that beauty is, personally, for me.

...

Before long, I was asked to move out of my office, to another one on the other side of the building. "Murray wants the office next to him for one of his own people", Helen said. ... I didn't know where my physics or my writing would take me. But I looked forward to the ride. And whether I continued to write as a hobby, or even supported myself with it, I hoped that maybe one day I'd write something that Feynman would admire. And then I thought, no, even better, I hope that someday I will write something that I will admire.""

— Leonard Mlodinow (Some time with Feynman, 2003)

Reading this paragraph for the first time brought tears to my eyes. It speaks to experiences that are, I believe, shared by many young researchers. The idea of following one's convictions, even at personal cost, is, at first glance, deeply inspiring. At the same time, I believe it is important to include the concluding part of the story as well. Leonard Mlodinow's trajectory was not one of unambiguous triumph; his convictions were strong enough to allow him to step away from a predefined path, but not necessarily to reshape the broader culture he was part of.

As for myself, I remain undecided as to whether I will follow the “Babylonian” or the “Greek” path. When starting this endeavour four years ago, I had recently returned from a serious accident in the high mountains that could easily have been fatal. This experience has shaped my perspective on life in a profound way and has led to a greater appreciation of the beauty in my everyday work. Yet I remain uncertain whether I am truly brave enough to fully follow the Babylonian path.

Perhaps, in a few years, this uncertainty will seem naïve, and we will look back and laugh at the ambitions I hold at this stage of my life. Perhaps not. For now, I cannot fully justify to the reader why they should continue through the remaining pages of this thesis, beyond the hope that the process of writing them has been valuable for my own thinking and decisions. I hope that, even if the message has not resonated with the reader as it has with me, they will nevertheless continue to reflect on their own path forward and keep pursuing their own “rainbows”. In writing this work, I have tried to remain close to those rainbows of my own: the ideas and problems in physics that genuinely drew my attention.

There is, however, a recurring structure in nearly every CMS thesis. Certain chapters have become almost ritualistic: one introduces the Standard Model, another the CMS detector. Both topics have been described extensively before, often in excellent detail, and with only modest variation from one thesis to the next. As a result, these sections can become both a repetition for the author and a test of patience for the reader.

To avoid unnecessary repetition, I have aimed to keep the standard material concise, and instead focus on the aspects most directly relevant to the work presented in this thesis, or those I personally find most compelling. The intention is not to provide the most complete or authoritative introduction to particle physics or CMS

ever written; many excellent theses already fulfil that role far better than I could. Rather, my aim is to provide sufficient context for the measurements that follow, while occasionally highlighting the ideas, detectors, and physical concepts that I found particularly striking during this work.

2 The Standard Model of particle physics

A century of work by successive generations of particle physicists has culminated in the Standard Model of particle physics, a quantum field theory that describes all known fundamental particles and their interactions. Although it is often distilled into a compact symbolic expression that fits on a coffee mug, its full structure and underlying mechanisms are remarkably rich and complex. As the Standard Model—and its limitations—have been extensively treated by my predecessors and numerous textbooks, I will not attempt to reproduce those comprehensive accounts here. Instead, I provide only a brief overview, focusing primarily on the aspects most relevant to the work presented in this thesis.

2.1. The standard model of particle physics

This work is based on lecture notes, the book "The Rise of the Standard Model" [8] and Ref. [9] (both are highly recommended reads). The elementary particles described by the Standard Model (see Figure 2.1) can be classified into three categories: fermions, vector bosons, and scalar bosons. Fermions, which have spin $1/2$, constitute the matter content of the theory and comprise six quarks and six leptons. The leptons are further divided into charged leptons and neutrinos. These fermions are organized into three generations, ordered according to the masses of the charged members within each generation. In this thesis, the term "lepton" will be used exclusively to refer to charged leptons, and the discussion will often be restricted to the first two generations.

2. The Standard Model of particle physics

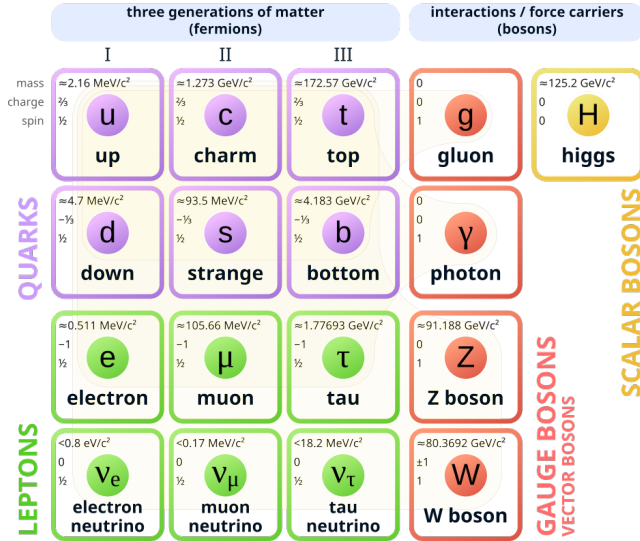


Figure 2.1.: A full list of the elementary particles described by the standard model of particle physics. Figure from ref. [7].

Vector bosons mediate the fundamental interactions: gluons are responsible for the strong interaction, the photon mediates the electromagnetic interaction, and the W and Z bosons mediate the weak interaction. The remaining scalar boson is the Brout-Englert-Higgs (BEH) boson. The BEH field is responsible for giving mass to elementary particles through the mechanism of electroweak symmetry breaking. All particles that couple to this field get a fundamental mass.

The strength of this interaction is given by the Yukawa coupling, defined as $\sqrt{2}m_f/v$, where v denotes the vacuum expectation value of the Higgs field. As the most massive fermion, the top quark has a Yukawa coupling of order unity, whereas the much lighter electron has a correspondingly small coupling, far beyond the reach of current experimental precision.

2.1.1. Strong interactions

Quantum Chromodynamics (QCD) is the quantum field theory that describes the strong interaction. Aptly named by Murray Gell-Mann [10], all particles that have a colour charge are affected by this field. Quarks and antiquarks carry colour and anticolour charges, conventionally labelled red, green, and blue. As a non-Abelian gauge theory based on the $SU(3)$ symmetry group, QCD predicts that the force carriers—gluons—also carry colour charge. This leads not only to quark–gluon interactions but also to gluon self-interactions, as illustrated in Figure 2.2.

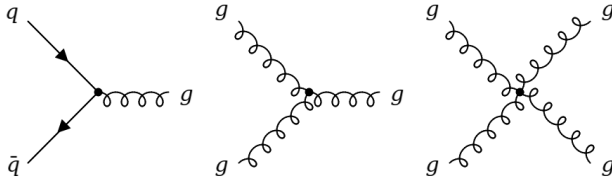


Figure 2.2.: Feynman diagrams representing the quantum chromodynamics interactions with quark–gluon couplings (left) and gluon self-couplings (middle and right).

A direct consequence of this non-Abelian structure is the scale dependence of the strong coupling constant, α_s . The interaction becomes strong at low momentum transfer and weak at high momentum transfer, a behaviour known as the running of α_s . This gives rise to two key phenomena.

Below the QCD scale, coloured particles do not exist in isolation. As two bound coloured particles are separated from each other, a tube of gluons is formed between them. Unlike the electromagnetic force, the corresponding potential between these two particles does not decrease with distance but actually grows approximately linearly as the gluons in this flux tube interact with each other. This means infinite work would be required to fully separate these particles from each other. Rather than allowing

free quarks to emerge, it becomes energetically favourable to produce a quark–antiquark pair from the vacuum, leading to the formation of colour-neutral hadrons. This phenomenon is called colour confinement and explains why quarks and gluons were never seen in isolation, which was seen as a significant weakness of the quark model in its early days [8].

Conversely, due to the small value of α_s at large energies, quarks and gluons in high energy interactions behave as quasi-free particles (“asymptotic freedom”). In this regime, scattering processes can be calculated using perturbative expansions in terms of α_s , an approach that breaks down in the strongly coupled, low-energy domain.

2.1.2. Electroweak interactions

Electromagnetic interactions are described by the quantum field theory of quantum electrodynamics (QED). All electrically charged particles couple to the electromagnetic field, giving rise to interactions mediated by photons, as illustrated in Figure 2.3. The strength of the electromagnetic interaction is set by the fine-structure constant, $\alpha_{\text{em}} \approx 1/137$ [11], which increases logarithmically with momentum transfer. Owing to its relatively small value compared to the strong coupling constant α_s , QED processes can be computed with high precision using perturbative expansions.

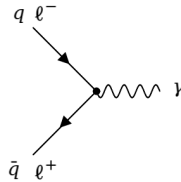


Figure 2.3.: Feynman diagram representing the coupling of charged particles to photons predicted by QED.

The remaining fundamental interaction described within the Standard Model is the weak interaction, to which all known fermions couple. The mediators of this field are represented by the charged W and neutral Z bosons, with masses of approximately 80.4 GeV [12] and 91.2 GeV [13], respectively. The large masses of these gauge bosons imply that the weak interaction is short-ranged at low energies and comparatively suppressed in strength relative to the electromagnetic and strong interactions. The corresponding interaction vertices are shown in Figure 2.4.

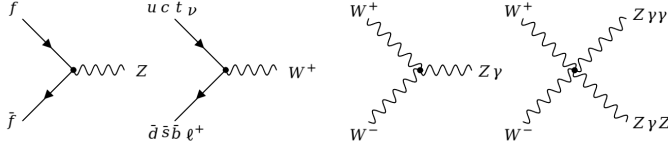


Figure 2.4.: Feynman diagram representing the weak interactions of SM particles.

The charged weak interaction, mediated by the W bosons, couples left-handed fermions within weak isospin doublets. In the lepton sector, this results in generation-diagonal couplings between charged leptons and their corresponding neutrinos. In the quark sector, however, flavour-changing transitions are allowed and are governed by the Cabibbo–Kobayashi–Maskawa (CKM) matrix. The matrix elements $V_{\alpha\beta}$ are multiplied onto the interaction strength, such that $|V_{\alpha\beta}|^2$ quantifies the transition probability between an up-type quark of flavour α and a down-type quark of flavour β . From global fits to experimental results sensitive to the CKM matrix elements, the absolute values of the nine elements are found to be [14]:

$$|V_{\text{CKM}}| = \begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} = \begin{pmatrix} 0.974 & 0.225 & 0.004 \\ 0.225 & 0.974 & 0.042 \\ 0.009 & 0.041 & 0.999 \end{pmatrix}. \quad (2.1)$$

Since neutrinos only interact via the weak force, only left handed neutrinos are observed. While right-handed neutrinos are not included in the Standard Model, they are a well-motivated extension and would correspond to so-called sterile neutrinos. In the SM, there are no mass terms for the neutrinos. Experimental data however show that neutrinos oscillate between flavours when they propagate through space. This observation is explained by small neutrino masses. These mass eigen states, denoted as ν_1, ν_2, ν_3 , are not the same states as the weak interaction eigenstates ν_e, ν_μ, ν_τ . The relation between these bases is described by the Pontecorvo–Maki–Nakagawa–Sakata (PMNS) matrix (values can be found at Ref. [15]):

$$U_{\text{PMNS}} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu1} & U_{\mu2} & U_{\mu3} \\ U_{\tau1} & U_{\tau2} & U_{\tau3} \end{pmatrix} \quad (2.2)$$

$$= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{\text{CP}}} \\ -s_{12}c_{23} - c_{12}s_{13}s_{23}e^{i\delta_{\text{CP}}} & c_{12}c_{23} - s_{12}s_{13}s_{23}e^{i\delta_{\text{CP}}} & c_{13}s_{23} \\ s_{12}s_{23} - c_{12}s_{13}c_{23}e^{i\delta_{\text{CP}}} & -c_{12}s_{23} - s_{12}s_{13}c_{23}e^{i\delta_{\text{CP}}} & c_{13}c_{23} \end{pmatrix}, \quad (2.3)$$

With the observation of radioactive decays, it became clear that the weak interaction plays a fundamental role in nature. However, constructing a consistent theoretical description proved to be a long and challenging process. Fermi's theory, postulated in the 1930's, provided the first quantitative framework but was plagued by serious problems in the UV. It was found to be non-renormalizable and suffered from unitarity violation at the weak scale. The Intermediate Vector Boson (IVB) theory, proposed in 1958, aimed to address some of these shortcomings by introducing a massive charged vector boson, the W, replacing the point-like four-fermion interaction. Although this formulation introduced a dimensionless coupling constant g , the theory remained non-renormalizable and still exhibited unitarity violations in the UV.

A major breakthrough came in 1964, when a gauge theory based on the symmetry group $SU(2)_L \times U(1)_Y$ was developed by Salam and Ward. This framework successfully unified the electromagnetic and weak interactions and predicted a massless photon and the three known massive gauge bosons: W^+ , W^- and Z^0 . While the theory is renormalizable, a fundamental issue arises: explicit mass terms for gauge bosons are forbidden, as they violate local gauge invariance. Similarly, fermion Dirac mass terms are not gauge invariant, since left- and right-handed fermion fields transform differently under the gauge group:

$$m_e \bar{\psi}_e \psi_e = m_e (\bar{\psi}_{eL} + \bar{\psi}_{eR})(\psi_{eL} + \psi_{eR}) \quad (2.4)$$

$$= m_e \bar{\psi}_{eL} \psi_{eR} + m_e \bar{\psi}_{eR} \psi_{eL} \quad (2.5)$$

It was further recognized in the 1960s that introducing masses “by hand” in non-Abelian gauge theories renders them non-renormalizable at energy scales of order 1 TeV, offering no real improvement over Fermi’s original theory. A mechanism for mass generation was therefore essential. Although spontaneous symmetry breaking appeared promising, the apparent absence of massless Goldstone bosons in experiments discouraged its early development. This issue was resolved in 1964 through the work of Brout and Englert; Higgs; and Guralnik, Hagen, and Kibble. They demonstrated that in gauge theories, spontaneous symmetry breaking does not lead to observable massless Goldstone bosons. Instead, these degrees of freedom are absorbed as the longitudinal components of massive gauge bosons, a process now known as the Brout–Englert–Higgs (BEH) mechanism. The BEH mechanism introduces a new potential in the Lagrangian of the SM. This potential takes the shape of a “Mexican hat” or the bottom of a wine bottle, with a vacuum expectation value (VEV) of 246 GeV. The spontaneous breaking of this symmetry results in masses for the weak bosons and extra interaction vertices between the Higgs and the gauge bosons themselves. Building on this framework, the modern electroweak theory was formulated by Weinberg and

others.

In 1971, it was proved by 't Hooft that the theory is indeed renormalizable. A key illustrative process is longitudinal gauge-boson scattering, which, in the absence of the Higgs boson, leads to scattering amplitudes that grow with the centre-of-mass energy, violating unitarity. The inclusion of the Higgs boson restores unitarity through destructive interference, regulating the high-energy behaviour of the cross section, as illustrated in Figure 2.5. This makes longitudinal gauge-boson scattering a particularly important process for probing the electroweak symmetry-breaking mechanism.

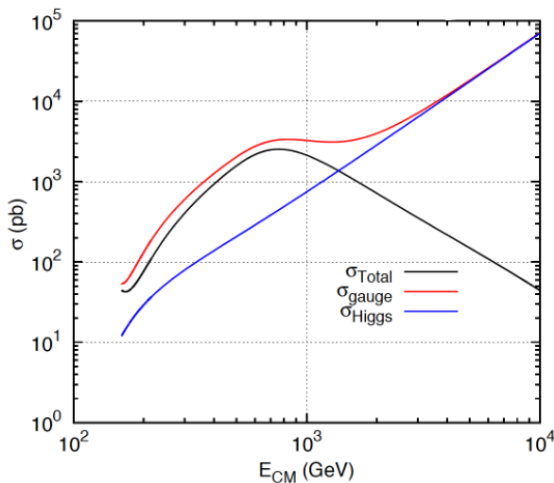


Figure 2.5.: Longitudinal gauge-boson scattering cross section as a function of center-of-mass energy. The diagrams added by the introduction of the Higgs cancel the problematic diagrams of the gauge bosons. Figure from Ref. [16].

The interactions between the Higgs field and fermions are described by Yukawa couplings, which give rise to fermion mass terms proportional to a coupling constant y_f specific to each fermion species. In addition, the Higgs field couples to the massive electroweak gauge bosons and exhibits self-interactions, as illustrated in Figure 2.6.

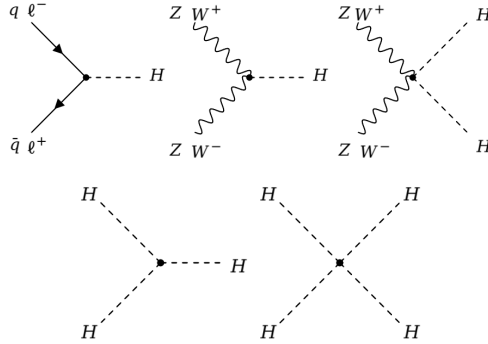


Figure 2.6.: Feynman diagram representing the interactions with the Higgs field for massive particles (top) and the Higgs self-interactions (bottom).

2.2. Shortcomings of the standard model

While the SM has been remarkably successful in describing the data collected by the LHC and other collider experiments, it cannot account for all observations made over the past few decades. This is particularly true when considering phenomena outside collider experiments, where several key questions remain unanswered. These limitations will be outlined in this section. As my work over the past four years has focused primarily on SM physics, I will not explore these topics in exhaustive detail.

A first notable shortcoming is the neutrino mass problem. Neutrino oscillations have been observed in both solar and reactor

neutrino fluxes [17, 18], which require neutrinos to have nonzero masses. However, the SM does not provide a mechanism to generate these masses.

Cosmological observations reveal further limitations of the SM. Ordinary matter accounts for only about 5% of the energy content of the universe, while roughly 26% is dark matter and 69% is dark energy [19]. These components are necessary to explain phenomena ranging from large-scale structure formation to gravitational lensing, as in the Bullet Cluster [20], and the accelerating expansion of the universe. While many Beyond the Standard Model (BSM) theories propose candidates for dark matter, the nature of dark energy remains largely unknown [21, 22].

Additional shortcomings include the absence of gravity in the SM framework. While general relativity successfully describes gravitational phenomena, it is not compatible with quantum field theory, and no fully successful unification with the SM has yet been achieved. Another unresolved issue is the matter–antimatter asymmetry observed in our universe, which the SM cannot fully explain.

2.3. From Feynman diagrams to cross-sections and Monte Carlo predictions

The previous sections provided a qualitative overview of interactions within the Standard Model. This qualitative description however does not describe how quantitative predictions can be made. In the context of proton–proton collisions, the starting point is a description of the proton itself. The proton p is a bound state of two up quarks and one down quark, called “valence quarks”. These quarks carry different colour charges such that the proton as a whole is colour neutral. In addition to the valence quarks, the proton contains a dynamical sea of quark–antiquark pairs and gluons, which arise from quantum fluctuations and

exist over short timescales. The collection of valence quarks and sea particles are referred to as the “partons” of the proton. The internal structure of the proton is described by parton distribution functions (PDFs), $f_\alpha(x, Q)$, which encode the probability of finding a parton of type α carrying a momentum fraction x of the proton at a resolution scale Q . These functions capture the non-perturbative dynamics of the proton and are defined up to the factorisation scale μ_F , which separates long-distance (non-perturbative) from short-distance (perturbative) physics.

Since the dynamics governing parton distribution functions (PDFs) are intrinsically non-perturbative, proton PDFs cannot be calculated from perturbative QCD and must instead be extracted from experimental data. The strongest constraints come from deep-inelastic scattering (DIS) in electron-proton collisions and fixed-target DIS experiments. Proton-proton collision data are also essential inputs in modern global PDF fits [23, 24, 25]. Nuclear PDFs are significantly less constrained and rely on electron-ion scattering and selected heavy-ion and proton-nucleus data [26, 27].

In high-energy collisions, individual partons from each proton undergo hard scattering processes. At sufficiently large momentum transfer, these interactions can be treated perturbatively. The total proton-proton cross section can then be factorised as:

$$\sigma(pp \rightarrow X)(s) = \sum_{\alpha, \beta} \int_0^1 dx_1, dx_2; f_\alpha(x_1, \mu_F), f_\beta(x_2, \mu_F) \cdot \hat{\sigma}(\alpha\beta \rightarrow X)(x_1 x_2 s, \mu_F, \mu_R), \quad (2.6)$$

where $\hat{\sigma}$ denotes the partonic cross section. This can be written as

$$\hat{\sigma}(\alpha\beta \rightarrow X)(\hat{s}, \mu_F, \mu_R) = \frac{1}{2\hat{s}} \int d\Phi_n |\overline{\mathcal{M}(\alpha\beta \rightarrow X)}|^2. \quad (2.7)$$

Here $\hat{s} = s x_1 x_2$ is the partonic centre-of-mass energy, $d\Phi_n$ denotes the n -body phase space, and $\mathcal{M}$ is the scattering amplitude

derived from the relevant Feynman diagrams.

Each partonic cross section is calculated perturbatively as an expansion in the (strong) coupling constant. The calculation is evaluated for the factorisation scale μ_F and the renormalisation scale μ_R that serve as cut-offs to avoid infrared and ultraviolet divergences, respectively. The lowest-order contributions correspond to Leading Order (LO) diagrams, which contain the minimal number of vertices. Higher-order corrections include Next-to-Leading Order (NLO), Next-to-Next-to-Leading Order (NNLO), and beyond, as illustrated in Figure 2.7. At NLO, corrections consist of both virtual contributions, involving loop diagrams, and real emissions, which introduce an additional particle in the final state and thus increase the final-state multiplicity from n to a $n+1$ multiplicity final state. Provided the process is evaluated in the perturbative regime, the expansion in the coupling constant is well behaved, and higher-order corrections systematically converge and improve the accuracy of theoretical predictions.

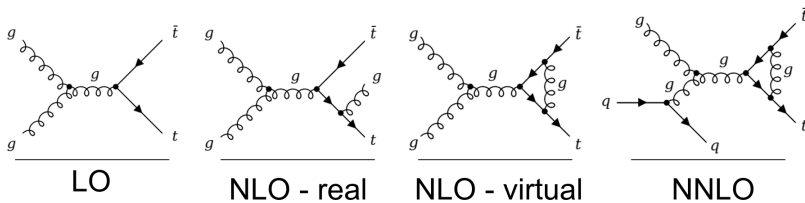


Figure 2.7.: Example (N)(N)LO diagrams for top pair production.

Furthermore, due to the relative strength of the interactions, QCD contributions typically dominate over electroweak ones in most processes. As a result, higher-order corrections in QCD generally provide the largest impact on theoretical predictions. For many of the processes considered in this work, calculations are available at next-to-leading order (NLO) in QCD, while next-to-next-to-leading order (NNLO) QCD results are still unavailable for a significant number of processes. Similarly, next-to-leading order

electroweak (EWK) corrections are not always available.

Focusing on the calculation of partonic cross sections shown in equation 2.7, we see that we need to integrate the squared matrix elements (MEs) over the available phase space. The quantity $|\mathcal{M}(ab \rightarrow X)|^2$ describes the transition probability of an initial state ab to a final state X . Feynman diagrams provide a convenient pictorial representation of these processes at the level of individual particles and interactions. In practice, computing the relevant elements of the S-matrix corresponds to evaluating all Feynman diagrams contributing to a given process up to the desired order in the perturbative expansion. This task, including the integration over phase space, is typically performed using dedicated matrix element (ME) generators.

ME generators are then combined with parton shower (PS) models to predict differential distributions of physical observables. These distributions form the basis for comparing theoretical predictions to collision data in LHC analyses.

ME generators are highly sophisticated tools that incorporate a wide range of physical and computational ingredients. In the following, we restrict the discussion to three aspects that are particularly relevant to the analyses presented in this thesis. These are treated in a qualitative manner, as a detailed mathematical derivation is beyond the scope of this experimental thesis. Parton shower models are also discussed briefly in the next section.

2.3.1. Hadronization and parton shower

While the combination of PDFs and ME calculations accurately describes the hard scatter in DIS interactions, it cannot account for the soft and collinear emissions that accompany these interactions, as these are not perturbatively calculable. Nor can it describe the hadronization of the end products of the hard scatter into colourless particles, as this is also non-perturbative

physics. Similarly, the hadronization of the final-state partons into colourless particles is inherently non-perturbative and cannot be described by the ME alone. To model these effects, PS models are employed. Figure 2.8 illustrates the various phenomena typically handled by the PS.

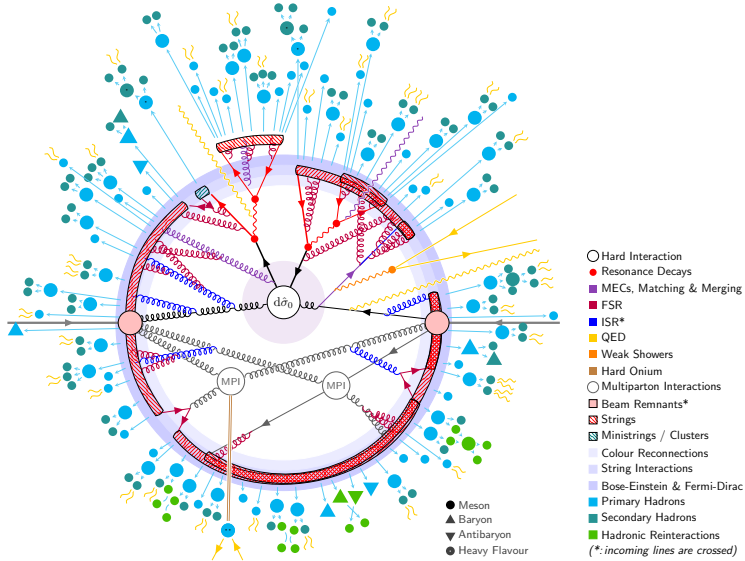


Figure 2.8.: Schematic representation of a top pair production event at the LHC, as modeled by a PS model. Figure from Ref. [28].

2.3.2. Negative weights in ME generators

For a NLO calculation, ME generators must integrate over two distinct phase spaces: the Born (n) phase space and the n+1 phase space with 1 real emission:

$$\begin{aligned}\sigma_{\text{NLO}} &\sim \int d\Phi_n \left[|\mathcal{M}_B|^2 + 2 \operatorname{Re} (\mathcal{M}_B \mathcal{M}_V^*) \right] + \int d\Phi_{n+1} |\mathcal{M}_R|^2 \\ &= \int_n B + V + \int_{n+1} R\end{aligned}\tag{2.8}$$

where B denotes the Born contribution, V the virtual corrections, and R the real emission contribution. While the sum of the $B+V$ contributions and the R contribution yields a finite result for infrared-safe observables, the individual terms are separately divergent due to infrared (IR) and collinear (C) singularities. These singularities cancel only after integration over the full phase space. In practice, this cancellation cannot be performed analytically at the integrand level because the divergent contributions originate from different regions of phase space. This necessitates the use of subtraction schemes that isolate and cancel these singular limits. Common approaches include Catani–Seymour dipole subtraction and transverse-momentum-based methods such as q_T (or related k_T) subtraction [29, 30]. These subtraction methods can be understood intuitively as follows: a counterterm S is introduced to mimic the singular behavior of the real emission:

$$\sigma_{\text{NLO}} = \int_n B + V + \int_1 S + \int_{n+1} (R - S).\tag{2.9}$$

This S term cancels the divergences in both integrals, making them separately integratable. But since S only approximates the singular structure of R , there are regions of phase space where $R-S$ is negative, leading to events with negative weights in the Monte Carlo integration.

This discussion is particularly relevant for the $t\bar{t}t\bar{t}$ and $t\bar{t}t$ processes, that will be covered in chapter 6.

2.3.3. Interference between processes at NLO

While Feynman diagrams provide an intuitive framework for understanding particle interactions in proton collisions, this picture becomes less clear in non-perturbative regions or when calculating NLO cross sections. A concrete example can be seen in the NLO diagrams for tW production, as shown in Figure 2.9. A naive NLO calculation of the tW process would also include LO $t\bar{t}$ diagrams. As some NLO tW diagrams and $t\bar{t}$ share the same final state, interference occurs between the two processes. In terms of amplitudes, this can be expressed as:

$$|\mathcal{A}_{WWbb}|^2 = \underbrace{|\mathcal{A}_{tWb}|^2}_{\text{DR1}} + \overbrace{2\mathcal{R}\{\mathcal{A}_{tWb}\mathcal{A}_{t\bar{t}}^*\}}^{\text{DR2}} + |\mathcal{A}_{t\bar{t}}|^2 \quad (2.10)$$

This interference makes it highly non-trivial to define meaningful and distinct tW and $t\bar{t}$ processes. A complete description of both contributions can be obtained by modeling the full process without assuming an on-shell top quark. Rather than generating $pp \rightarrow t\bar{t}$, one considers the exact final-state process

$$pp \rightarrow \ell^+ \nu_\ell \ell^- \bar{\nu}_\ell b\bar{b},$$

which consistently includes resonant, non-resonant, and interference contributions. This approach is employed, for example, in the `bb4l` method [31].

Generating the full set of resonant and non-resonant diagrams without approximations is, however, computationally demanding, and for some processes remains beyond the reach of current calculations. Moreover, these calculations are specific to the final state under consideration and must therefore be repeated for each top-quark decay channel. To simplify the calculation, the narrow-width approximation (NWA) is often used. Within the NWA, the process

$$pp \rightarrow t\bar{t} \rightarrow (Wb)(Wb)$$

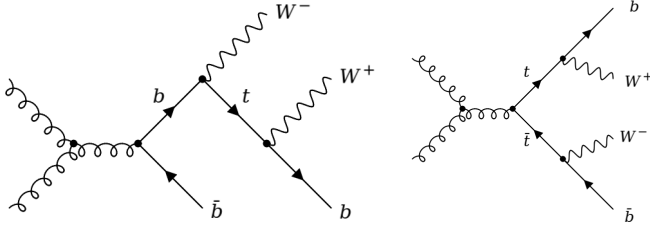


Figure 2.9.: NLO diagrams of tW production include non-resonant modes (left) but also contain resonant $t\bar{t}$ production (right).

factorizes into the gauge-invariant production and subsequent decay of on-shell top quarks, substantially reducing the complexity of the phase-space integration. As a result, resonant $t\bar{t}$ samples can be generated at very high perturbative precision. Non-resonant contributions, including the interference between tW and $t\bar{t}$ production, are then incorporated separately using a variety of dedicated schemes.

DR1: In Diagram removal scheme 1 (DR1) resonant diagrams are removed at the amplitude level. This eliminates the overlap between both processes. While this approach is simple and provides a reasonable description of the data when interference is small, it does not guarantee gauge invariance or unitarity.

DR2: Diagram removal scheme 2 (DR2) keeps the interference term but removes the squared resonant amplitude in equation 2.10. This generally yields a better description of distributions when interference is non-negligible. Like DR1, however, DR2 does not ensure gauge invariance or unitarity.

DS If gauge invariance or unitarity violations are significant in DR1 or DR2, the Diagram Subtraction (DS) scheme can be employed. In DS, the full amplitude is retained, but the resonant contribution

is subtracted in a gauge-invariant way: $|\mathcal{A}_{\text{DS}}|^2 = |\mathcal{A}|^2 - |\mathcal{A}_{\text{Res GI}}|^2$. This ensures that the resonant component is removed while preserving gauge invariance and unitarity.

This discussion is particularly relevant for the $t\bar{t}t$ processes, that will be covered in chapter 6.

2.3.4. Multijet merging and matching

A process generated at NLO, will only describe the hardest emission with the ME generator via the real emission diagrams, while all remaining radiation is described by the PS. Overall the PS can describe well the non-perturbative regions in the emission phase space, the soft region of the p_T spectrum and narrower angular radiation distributions. The PS is however less suited to describe harder emissions, which would be better described by the ME generator as they are perturbative in nature.

Frameworks such as MadGraph5_aMC@NLO allow the inclusion of multiple hard parton emissions at the matrix-element (ME) level through multijet merging and matching. In this approach, samples with increasing partonic multiplicity are combined, denoted as @0 (no additional partons), @1 (one additional parton), and @2 (two additional partons).

The combination of these samples introduces two sources of double counting. First, at the ME level, the real-emission contributions in the NLO calculation of a given multiplicity overlap with the Born-level configurations of the next-higher multiplicity: NLO(@0) overlaps with LO(@1), and NLO(@1) with LO(@2). This is visualized in figure 2.10. Second, the PS acting on lower-multiplicity samples can generate emissions that populate the same phase-space regions as higher-multiplicity ME configurations, leading to an additional overlap between PS emissions and explicit ME partons.

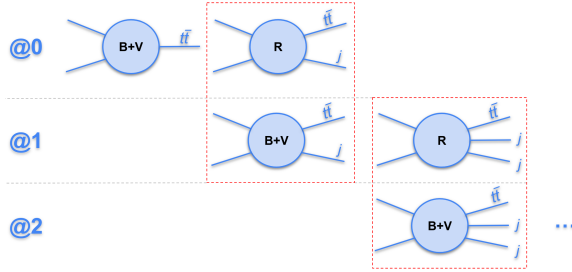


Figure 2.10.: Double counting by ME when multijet merging a $t\bar{t}$ sample.

To resolve this double counting, MADGRAPH5_AMC@NLO relies on two algorithms designed for this purpose: the MLM [32] algorithm for LO merged samples and the FFX [33] algorithm for NLO merged samples. Both approaches introduce a merging scale, Q_{cut} , which defines the boundary in phase space between hard emissions described by matrix elements and softer emissions generated by the parton shower. These techniques are of particular relevance to the modeling of the $t\bar{t}W$ process that will be covered in chapter 5.

2.4. Luminosity: a short overview

It is widely accepted that the SM is only an effective theory at the energies currently accessible. Consequently, the development of accelerators with ever-higher center-of-mass energies (E_{COM}) seems like a natural progression. Historically, such advances have repeatedly yielded significant discoveries, from the detailed cataloguing of the particle “zoo” to the identification of the W and Z bosons. Another crucial parameter, often underappreciated, is the integrated luminosity—a process-independent measure of the total number of hadron–hadron collisions delivered over the lifetime of the machine. The high instantaneous luminosity of the

LHC was instrumental in the discovery of the Higgs boson and was a key reason why accelerators became the dominant tool in high-energy physics compared to cosmic-ray experiments in the 1950s. The expected number of interactions for a specific process is then given by the product of the integrated luminosity and the process cross section:

$$\frac{dN}{dt} = L_{\text{inst}}(t) \sigma \quad (2.11)$$

$$N_{\text{exp}} = \int L_{\text{inst}}(t) \sigma dt = L \sigma \quad (2.12)$$

From this formula, it is clear that in order to translate a number of observed events to a cross section, the luminosity of the experiment needs to be well-known. As the cross section is both calculable and measurable, the comparison of the two values can be used to extract theory parameters from the measurement. The precision of the accordingly measured parameters depends on the precision of the three ingredients: the measurement of the process itself (N_{exp}), the luminosity measurement (L), and the theory calculation. It is hence vital that the luminosity is a well-measured quantity, as it is an inevitable uncertainty for any measurement relying on the observation of a number of events.

For two beams with N_1 and N_2 particles, distributed in the plane transverse to the beam direction according to normalised particle distributions $f_1(x, y)$ and $f_2(x, y)$, and circulating at a revolution frequency ν , the instantaneous luminosity is given by:

$$L_{\text{inst}} = \nu N_1 N_2 \int dx \int dy f_1(x, y) \cdot f_2(x, y) = \nu \frac{N_1 N_2}{A_{\text{overlap}}} \quad (2.13)$$

where A_{overlap} is the effective overlap area of the beams. While ν and N_1 and N_2 are well known quantities, the overlap area of the beams is not directly accessible. It is therefore often advantageous to calibrate the luminosity using equation 2.11 and the observed

number of events for a process with a well-known cross section.

This could easily be done in e^+e^- -colliders, making use of the Bhabha scattering process, which has a clear detector signature and a precisely calculable perturbative cross section. In proton–proton collisions, achieving such precision is more challenging due to the complexity of theoretical predictions. One approach currently under investigation is to use the production rate of Z bosons as a reference process. The fiducial cross section for the Drell–Yan process $pp \rightarrow Z \rightarrow \mu^+\mu^-$ is measured in a low-pileup run using a precise luminosity determination from Van der Meer scans (as discussed later), and then used in regular data-taking for the measurement of the luminosity. For heavy ion collisions we do not have these reference processes available and hence will have to rely on methods that estimate the overlap area of both beams for the foreseeable future. These methods were first developed at the Intersecting Storage Rings (ISR) at CERN by Simon van der Meer [34] and are also still performed for pp collisions [35].

The concept goes as follows. consider a detector ("luminometer") that measures a rate $\mathcal{R}$, proportional to the instantaneous luminosity L_{inst} . The observed rate for two beams separated by $(\Delta x, \Delta y)$ is then described as:

$$\mathcal{R}(\Delta x, \Delta y) = \mathcal{N} \int dx \int dy f_1(x, y) \cdot f_2(x + \Delta x, y + \Delta y) \quad (2.14)$$

where $\mathcal{N}$ is a proportionality constant intrinsic to the detector and the colliding systems. The beam overlap area can then be expressed in terms of the rate for head-on collisions: $A_{\text{overlap}} = \mathcal{N}/R(0, 0)$. To use the rate of this luminometer to get an estimate of the luminosity we however require to find the proportionality constant. This can be done by integrating the rate over the beam separations:

$$\begin{aligned}
 \int d\Delta x \int d\Delta y R(\Delta x, \Delta y) &= \mathcal{N} \left(\int dx \int dy f_1(x, y) \right) \\
 &\quad \times \left(\int d\Delta x \int d\Delta y f_2(x + \Delta x, y + \Delta y) \right) \\
 &= \mathcal{N} \cdot 1 \cdot 1 = \mathcal{N}
 \end{aligned} \tag{2.15}$$

In the following, we assume the particle densities are factorizable: $f_i(x, y) = f_i^x(x) \cdot f_i^y(y)$. This allows the beam overlap area to be expressed as

$$\begin{aligned}
 A_{\text{overlap}} &= \mathcal{N} \left(\int dx \int dy f_1^x(x) \cdot f_1^y(y) \right) \\
 &\quad \times \left(\int d\Delta x \int d\Delta y f_2^x(x + \Delta x) \cdot f_2^y(y + \Delta y) \right) / (R_x(0)R_y(0)) \\
 &:= 2\pi\Sigma_x\Sigma_y.
 \end{aligned} \tag{2.16}$$

where parameters Σ_x and Σ_y describe the effective widths of the beam overlap area along two orthogonal directions. Independent of the functional form of the particle densities, the effective widths can be obtained from so called Van der Meer (VdM) scans, in which the beams are separated along each axis and moved in steps across each other. The integrated rate as a function of beam displacement is measured, as in equation 2.15.

Once the luminosity for head-on collisions under the special beam conditions of the VdM scan is determined, the corresponding luminometer rate can be calibrated to a visible cross section. This cross section is then defined as:

$$\sigma_{vis} = \frac{\mathcal{R}(0, 0)}{L} = \frac{2\pi\Sigma_x\Sigma_y\mathcal{R}(0, 0)}{\nu N_1 N_2}. \tag{2.17}$$

This calibrated fiducial cross section subsequently allows the instantaneous luminosity to be measured during regular data-taking,

independently of the prevailing beam conditions. Several luminometers are used to determine the collected luminosity. Since different luminometers may exhibit varying responses as a function of pileup (“linearity”) and over time (“stability”), the measured luminosities are compared across luminometers and systematic uncertainties are assigned to account for these effects.

This section provided a brief motivation for the methodologies applied in a VdM program at the LHC. A detailed discussion of the luminosity calibrations for three specific datasets is presented in [Section 7](#).

3 Theoretical and experimental status in top quark physics

3.1. Top quark properties

The top quark is the central character in a large part of the work presented here. Hence, I will discuss this particle in a bit more detail than what was discussed in the previous chapter. We will discuss its main properties and the current state of the art in top-quark research at the LHC. Why should you be excited about the top quark? As the heavy partner of the up quark with a mass of approximately 172.5 GeV, the top quark is the heaviest particle in the Standard Model. It was the last fermion to be discovered, which occurred in 1995 at the Tevatron collider, where both the CDF and DØ collaborations reported evidence for top-quark pair production [36, 37].

3.1.1. Mass

While the top quark can be thought of as “just” a very heavy up quark, its large mass makes it one of the most intriguing particles to study at colliders. Its mass corresponds to a Yukawa coupling near unity, far larger than that of any other fermion. This strong coupling makes the top quark a key player in probing the Higgs field and its properties. Radiative corrections involving the top quark significantly modify the effective Higgs potential. In fact, the top quark mass and the Higgs boson mass are the dominant

parameters affecting the shape of this potential.

These corrections can introduce a second minimum at a higher VEV, raising the possibility that the universe could spontaneously tunnel to this alternate vacuum. Several scenarios arise depending on the shape of this potential:

1. absolute stability:

$$V_{\text{eff}}(v) < V_{\text{eff}}(\phi_{\text{min}}^2)$$

2. metastability:

$$V_{\text{eff}}(v) > V_{\text{eff}}(\phi_{\text{min}}^2), \quad \tau > T_U$$

3. instability:

$$V_{\text{eff}}(v) > V_{\text{eff}}(\phi_{\text{min}}^2), \quad \tau < T_U,$$

where ϕ_{min}^2 is the location of the second minimum of the Higgs potential, τ the lifetime of an unstable electroweak vacuum, and T_U the age of the universe. Clearly, an unstable vacuum with $\tau < T_U$ would be inconsistent with the observed age of the universe. Current experimental constraints, however, indicate that we live in a metastable vacuum [38, 39]. Precise knowledge of the Standard Model Higgs potential would allow us to determine the energy scale at which new physics must appear to stabilize the universe and would place strong constraints on models that could render the vacuum unstable.

The top quark's large mass also implies an extremely short lifetime of approximately 10^{-25} s, after which it decays almost exclusively via the weak interaction. Because the third-generation elements of the CKM matrix are nearly diagonal, the top decays to a bottom quark in more than 99% of cases. Due to this short lifetime, the top quark typically does not hadronize into bound states, unlike

any other quark. This makes the top quark unique as the only “bare” quark observable in nature, offering a rare window into fundamental quark properties.

Recent measurements suggest that very soft top quark pairs can form quasi-bound states near production threshold [40, 41, 42], known as toponium, which provide a novel laboratory to study QCD at the shortest timescales. Moreover, the top quark’s lifetime,

$$\tau_t \approx 10^{-25} \text{ s},$$

is significantly shorter than the characteristic hadronization timescale,

$$\tau_{\text{had}} \sim \frac{1}{\Lambda_{\text{QCD}}} \approx 10^{-24} \text{ s},$$

as well as the timescale associated with spin decorrelation due to QCD interactions. As a result, the top quark decays before hadronizing, allowing its spin information to be directly transferred to its decay products.

3.1.2. Charge asymmetry

Charge asymmetry, often also referred to as forward backward asymmetry, in top quark decays is an interesting phenomenon that only starts to appear in calculations beyond LO. The effect appears due to interference between final state and initial state emissions and interference of the box with the lowest-order-diagram [43]. Due to this interference, the top quark is preferentially produced in the direction of the incoming quark, while the top antiquark is preferentially produced in the direction of the incoming antiquark. This effect is only present in quark–antiquark initiated processes, as the gluon–gluon initial state is charge-conjugation symmetric. At the Tevatron, where top-quark production was dominated by quark–antiquark annihilation, the resulting forward–backward asymmetry was relatively large, with an inclusive magnitude of approximately 4–5%. The asymmetry was visible both in the

partonic centre-of-mass frame and in the laboratory frame, owing to the asymmetric proton–antiproton initial state.

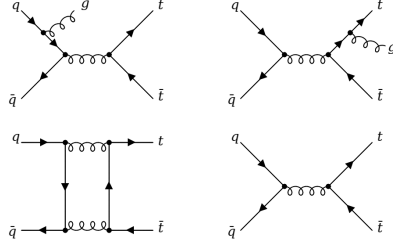


Figure 3.1.: Origin of charge asymmetry in hadroproduction of top quarks: interference of initial-state (top left) with final-state (top right) real emission diagrams and interference of the box (bottom left) and Born (bottom right) diagrams.

At the LHC, the situation is considerably more challenging. Top-quark pair production is dominated by gluon–gluon fusion, which does not contribute to the asymmetry, substantially diluting the observable effect. Furthermore, the symmetric proton–proton initial state removes any preferred direction in the laboratory frame. As a result, the asymmetry manifests itself only through a difference in the absolute rapidity distributions of top quarks and top antiquarks. These effects reduce the inclusive charge asymmetry to below 1%, making its measurement a demanding precision test of the Standard Model [43]. The charge asymmetry is usually defined as:

$$A_c^{t\bar{t}} = \frac{N(\Delta y^{t\bar{t}} > 0) - N(\Delta y^{t\bar{t}} < 0)}{N(\Delta y^{t\bar{t}} > 0) + N(\Delta y^{t\bar{t}} < 0)}. \quad (3.1)$$

As reconstructing top quarks is often not trivial, the variable is often measured in dileptonic $t\bar{t}$ events via the leptonic charge asymmetry proxy variable A_c^ℓ :

$$A_c^\ell = \frac{N(\Delta y^\ell > 0) - N(\Delta y^\ell < 0)}{N(\Delta y^\ell > 0) + N(\Delta y^\ell < 0)}. \quad (3.2)$$

While the inclusive $t\bar{t}$ asymmetry is in general quite low, $t\bar{t}$ production in association with bosons can have higher asymmetries.

3.1.3. Spin correlation/polarization

In $t\bar{t}$ production, the spins of the top quark and antiquark are largely unpolarized, but mutually correlated [44]. The decay products of the W bosons from the top quark decay are however polarized with respect to the top quark as the top decays before the spin decorrelation timescale [45]. This makes $t\bar{t}$ pairs excellent candidates for studying polarization and spin correlation. The study of these correlations has lead, to among other things, the observation of entanglement in $t\bar{t}$ production [46, 47].

3.1.4. CKM matrix

The CKM matrix was introduced in equation 2.1, where the matrix elements $V_{\alpha\beta}$ describe the coupling strength between particles α and β . Interestingly, the third generation couplings are almost diagonal, meaning V_{tb} is close to unity, while V_{td} and V_{ts} are close to zero. This implies that top quark almost exclusively decays into a bottom quark and a W boson, and only very rarely to s or d quarks directly with a W boson.

The most natural measurements to constrain V_{tb} can be found in single top quark production processes. Since both the production and decay of the top quark proceed via the W_{tb} vertex, as shown in Figure 3.2, the cross section depends quadratically on this coupling:

$$\frac{\sigma_{\text{single-}t}}{\sigma_{\text{single-}t}^{\text{SM}}} \propto |V_{tb}|^2. \quad (3.3)$$

As a result, measurements of the single top production cross section can be directly translated into determinations of $|V_{tb}|$, assuming all other parameters entering the prediction are fixed to their Standard Model values. Extensive measurements have been performed by the CMS and ATLAS collaborations, all of which are consistent with $|V_{tb}| = 1$ [48].

3.2. Production of top quarks at the LHC

While the LHC and the CMS detector will be described in detail in chapter 4, it is worth noting the immense scale of top quark production at these experiments. At collision energies of 13–13.6 TeV and with the combined integrated luminosity of Run 2 and Run 3 exceeding 470 fb^{-1} , roughly $4 \cdot 10^8$ top–antitop pairs have been recorded by CMS. In just 30 years, the field has progressed from the discovery of the top quark at the Tevatron in 1995—with only a handful of observed events—to performing high-precision measurements of top quark properties at the LHC.

The most precise determination of the top quark mass using LHC data comes from a combination of Run 1 results by ATLAS and CMS, yielding $172.52 \pm 0.33 \text{ GeV}$ [49]. The single most precise measurement at the LHC was obtained using 2016 CMS data at a center-of-mass energy of 13 TeV [50]. Using a template likelihood fit, this measurement gives

$$m_t = 171.77 \pm 0.37 \text{ GeV}. \quad (3.4)$$

For the inclusive top pair production cross section, the precision is largely limited by the uncertainty on the total integrated luminosity, currently at approximately 2%. Multiple double-differential cross-section measurements have been performed [51, 52], providing rigorous tests of the Standard Model over a wide kinematic range.

At typical LHC resolution scales Q , the probability of finding a gluon parton inside the proton is significantly higher than that of finding a quark. Gluons dominate at low momentum fractions, whereas at large momentum fractions, the valence quarks remain more likely to initiate interactions. As a consequence, roughly 90% of top pair production at the LHC proceeds through gluon–gluon fusion. This is in stark contrast to the Tevatron, where proton–antiproton collisions produced a large fraction of top pairs via $q\bar{q}$ annihilation, reflecting the dominance of valence quark contributions at those energies.

3.3. Cross-sections of top quark processes

As the LHC has collected a vast dataset, the top quark has been observed in a wide variety of production processes. Top pair production has by far the largest cross section at the LHC. However, several other production modes exist and provide complementary information on top quark properties and couplings. The simplest of these is single top production, which proceeds via electroweak interactions. The leading-order production modes are illustrated in Figure 3.2 and include the t-channel, tW, and s-channel processes.

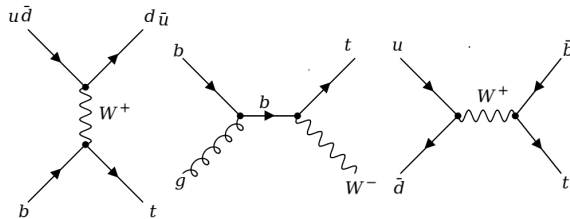


Figure 3.2.: Single top production mode diagrams at LO: t-channel (left), tW (middle) and s-channel (right).

A second important class of processes is top quark associated production, in which a boson is produced alongside a $t\bar{t}$ pair.

The associated boson can be a photon, W, Z, or Higgs boson. Although these processes have significantly smaller cross sections than inclusive $t\bar{t}$ production, they provide direct sensitivity to the couplings between the top quark and electroweak or Higgs bosons, thereby offering powerful probes of physics beyond the Standard Model. Complementary to $t\bar{t}$ associated production are processes involving single top quark production in association with bosons, which probe similar couplings in an electroweak production environment. Key examples include single top production in association with a Higgs boson, photon, or Z boson. Some leading-order production modes are illustrated in Figure 3.3 for both categories.

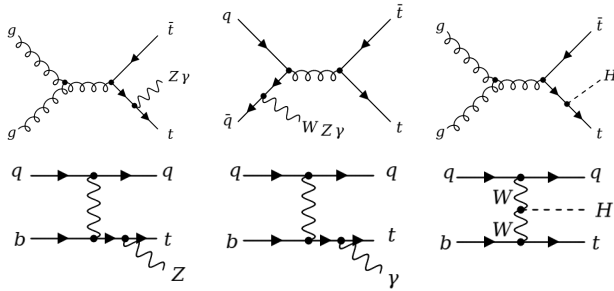


Figure 3.3.: Illustrative top pair production in association with bosons (top) and single top quark production in association with bosons (bottom) diagrams.

Finally, four top quark production ($t\bar{t}t\bar{t}$) represents one of the rarest processes accessible at the LHC, with a predicted cross section of approximately $\mathcal{O}(10)$ fb at $\sqrt{s} = 13$ TeV in the Standard Model [53, 54, 55]. This process is particularly sensitive to the top quark Yukawa coupling, as it receives contributions from diagrams involving Higgs boson exchange. Moreover, due to its small Standard Model cross section, it is highly sensitive to potential contributions from new physics, which could significantly enhance the production rate [56, 57].

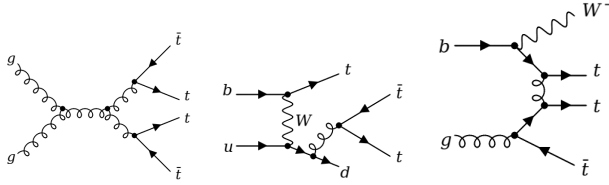


Figure 3.4.: Illustrative $t\bar{t}t\bar{t}$ (left), $tttj$ (middle), $tttW$ (right) production diagrams.

Closely related to four top production are the three top quark processes (ttt). Three top quark production happens through two processes: $tttj$ and $tttW$, with a total cross section of roughly 2 fb [58]. These processes have similar experimental signatures and can contribute as backgrounds in searches for $t\bar{t}t\bar{t}$ production.

Most of these processes have now been observed and measured by both the ATLAS and CMS collaborations. Figure 3.5 summarizes the current status of top quark measurements performed by the CMS Collaboration, with the majority of results showing good agreement with Standard Model predictions within uncertainties. In this thesis, particular focus will be placed on the $t\bar{t}t\bar{t}$, $tttj$, $tttW$ and $t\bar{t}W$ processes. These processes are closely related both theoretically and experimentally: $t\bar{t}W$, $tttj$ and $tttW$ production constitute the most important backgrounds in measurements of four top quark production. As such, they will be discussed in dedicated chapters.

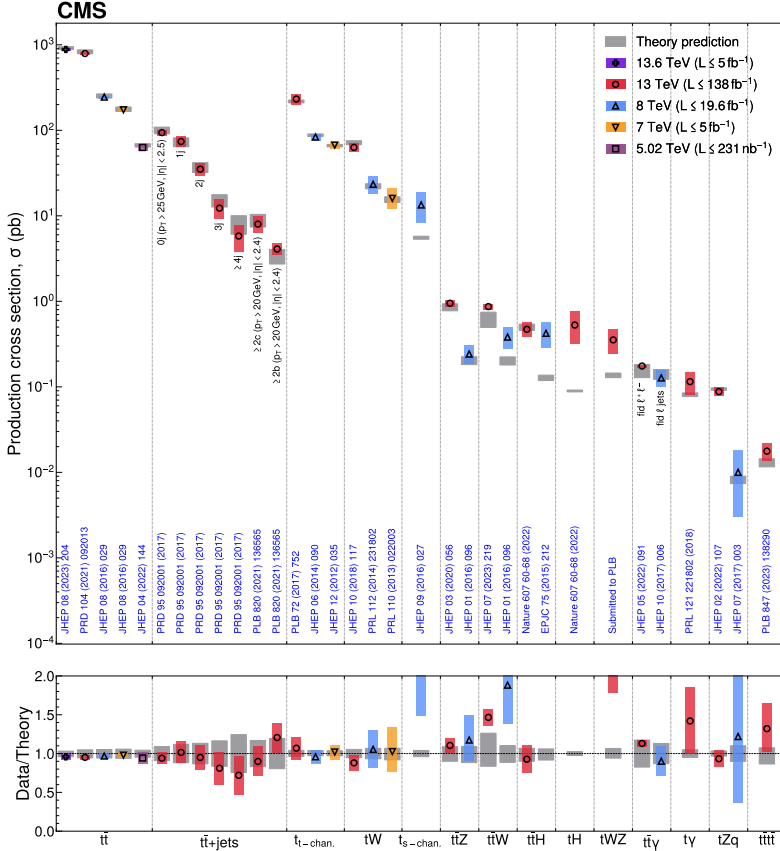


Figure 3.5.: Summary of production cross section measurements by the CMS collaboration involving top quarks. Measurements performed at different collision energies are marked by unique symbols and the coloured bands indicate the combined statistical and systematic uncertainty of the measurement. Grey bands indicate the uncertainty of the theory predictions. Figure from Ref. [59].

3.4. The discovery of four top quark production

Ghent University has taken a leading role in the observation of the $t\bar{t}t\bar{t}$ process using CMS Run 2 data. As this process is central to this thesis, a brief overview of the analysis strategy and its main features is provided. The author did not directly contribute to this work. This chapter is based on the discovery paper [60].

3.4.1. Analysis regions

The $t\bar{t}t\bar{t}$ process is characterized by final states with high jet and b -jet multiplicities, as well as multiple isolated leptons. To maximize sensitivity, the analysis focuses on multilepton final states, in particular the same-sign dilepton, trilepton, and four-lepton channels, which provide a favorable signal-to-background ratio. The same-sign dilepton and trilepton regions provide similar sensitivity, while the four-lepton signal region contains very few events, providing much less to the total result.

The dominant backgrounds arise from $t\bar{t}$ production in association with vector bosons ($t\bar{t}W$, $t\bar{t}Z$), as well as from processes with non-prompt or charge-misidentified leptons. Dedicated control regions (CRs) are defined to constrain these backgrounds directly

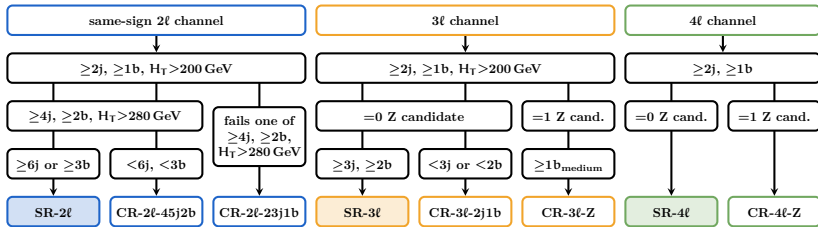


Figure 3.6.: Schematic representation of the event selection and categorization in the RUN 2 $t\bar{t}t\bar{t}$ analysis. Figure from Ref. [60].

from data. The $t\bar{t}Z$ control regions are constructed by requiring three leptons with an opposite-sign same-flavour (OSSF) pair with an invariant mass compatible with the Z boson mass, as well as a corresponding four-lepton region. In contrast, other background processes do not exhibit such a clean signature and are therefore constrained using a combination of control regions and signal regions, with additional discrimination provided by machine learning techniques. A summary of the analysis regions is shown in Figure 3.6.

3.4.2. ML approach

Multiclass Boosted Decision Trees (BDTs) are trained separately in the dilepton and trilepton+four-lepton analysis regions, using input variables that capture both event topology and kinematic properties. The models are trained to discriminate the $t\bar{t}t\bar{t}$ signal from the dominant backgrounds: $t\bar{t}$ production and $t\bar{t}X$ processes (including $t\bar{t}H$, $t\bar{t}Z$, and $t\bar{t}W$). The $t\bar{t}$ contribution is particularly important, as it constitutes the main source of non-prompt and charge-misidentified leptons. The training is performed using Monte Carlo events that satisfy the selection criteria of the corresponding signal regions.

Key discriminating features include jet and b -jet multiplicities, lepton kinematics, global event variables such as the scalar sum of transverse momenta, and observables sensitive to the presence of multiple top quarks. The multivariate discriminants are subsequently used as final observables in a simultaneous maximum likelihood fit across all signal and control regions.

3.4.3. Results

The measurement is performed using a template-based maximum-likelihood fit to the BDT discriminant distributions in all signal and control regions. The details of the fitting procedure are presented in the following section; for the present discussion, it is

sufficient to regard the fit as a framework that simultaneously incorporates all relevant sources of uncertainty and extracts the signal cross section together with its associated uncertainties. The measured cross section is found to be consistent with the Standard Model prediction within uncertainties. In addition, a significant excess relative to the background-only hypothesis is observed, corresponding to the first observation of four-top-quark production with a significance greater than five standard deviations.

This result represents an important milestone, confirming the Standard Model prediction in one of its rarest accessible processes and providing a new probe of the top quark Yukawa coupling and potential contributions from physics beyond the Standard Model. While the discovery was an impressive feat, significant questions remain. These questions lie at the heart of this thesis and are discussed in the following subsections.

3.4.4. Interplay between top pair production in association with a W boson and four-top production

The likelihood fit incorporates a wide range of systematic uncertainties. Their relative effect on the measurement is commonly summarized using impact plots, which quantify how individual uncertainties influence the fitted cross section. A detailed discussion of these plots is also deferred to later sections. Figure 3.8 shows the impact plot for the $t\bar{t}t\bar{t}$ cross section measurement, displaying the 20 uncertainties with the largest influence on the measurement result. The dominant uncertainties are related to the jet energy scale, which describes how accurately the energy of reconstructed jets is measured, and flavour tagging performance, which quantifies how reliably jets originating from (b) quarks can be identified. This is expected, as the BDT discriminants used in the analysis rely heavily on jet and b-jet multiplicities to separate $t\bar{t}t\bar{t}$ signal events from the background. Uncertainties affecting the reconstruction of jets and the identification of jets originating

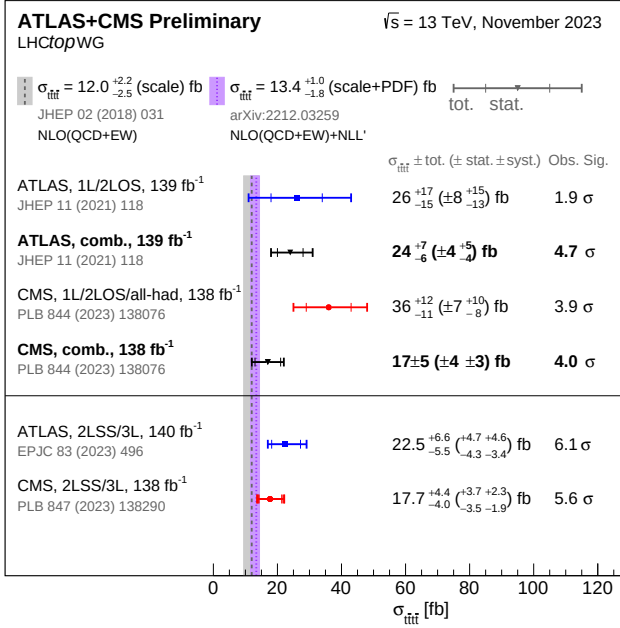


Figure 3.7.: Measured $t\bar{t}t$ production cross section compared to the Standard Model prediction and results from ATLAS and CMS. Figure from Ref. [61].

from b quarks therefore have a particularly strong impact on the measurement sensitivity.

The next most significant uncertainties are those associated with the modelling of additional radiation in the $t\bar{t}W$ process. Due to its very similar final-state topology—particularly in events with additional (b-)jet activity— $t\bar{t}W$ production constitutes a major irreducible background to $t\bar{t}t$. Consequently, uncertainties in its modelling have a substantial impact on the $t\bar{t}t$ cross section measurement. A more precise understanding of $t\bar{t}W$ production is therefore essential for improving future $t\bar{t}t$ analyses.

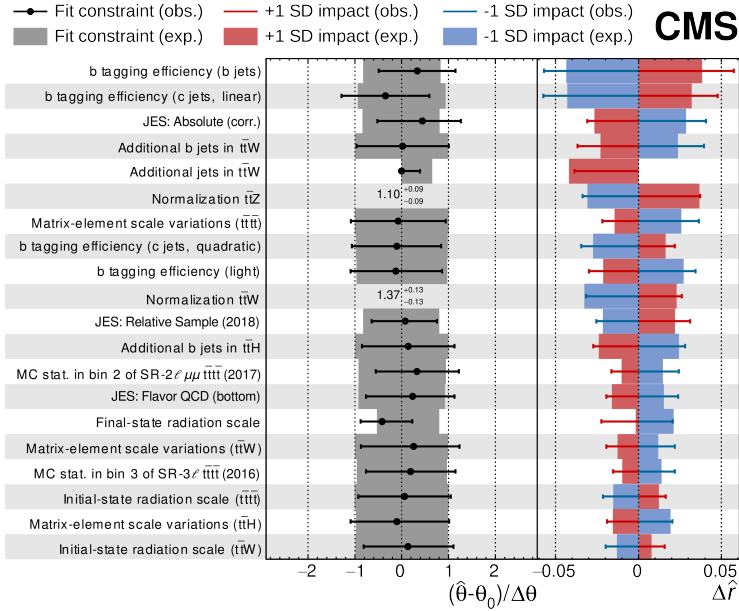


Figure 3.8.: Impact plot for the $t\bar{t}t$ cross section fit. The plot shows the 20 uncertainties that have the most impact on the measurement. Figure from Ref. [60].

3.4.5. Interplay between three-top and four-top production

While the $t\bar{t}W$ process shares many experimental features with $t\bar{t}t$ production, even closer signatures arise from three-top quark production processes, such as $tttj$ and $tttW$. Although these processes have very small predicted cross sections of order $\mathcal{O}(2)$ fb at NLO QCD [58], their final states closely resemble those of $t\bar{t}t$ events, making them difficult to distinguish experimentally.

To assess the sensitivity of the analysis to three-top production, a two-dimensional likelihood scan is performed, simultaneously measuring the ttt and $t\bar{t}t$ cross sections. The result of this scan is shown in Figure 3.9. A strong anticorrelation between the two

processes is observed, indicating that the fit is unable to disentangle their individual contributions. The likelihood contours are relatively broad, allowing for sizable variations in both cross sections, while only the sum of the contributions remains well constrained.

This behaviour reflects the limited discriminating power of the analysis between $t\bar{t}t$ and $t\bar{t}t\bar{t}$ final states. As a consequence, any observed excess in the measured cross section cannot be unambiguously attributed to $t\bar{t}t\bar{t}$ production alone. It may instead originate from an enhanced $t\bar{t}t$ contribution or potential effects from physics beyond the Standard Model. Improving the modelling and understanding of $t\bar{t}t$ production is therefore crucial for interpreting such deviations in future analyses!

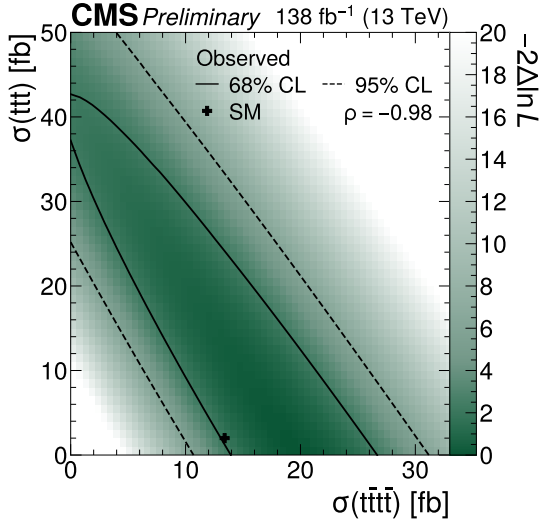


Figure 3.9.: Two-dimensional negative log-likelihood scan of the $t\bar{t}t\bar{t}$ and $t\bar{t}t$ production cross sections. Figure from Ref. [62].

4 The Large Hadron Collider and CMS experiment

The field of particle physics has been driven by technological innovation ever since its conception in the 1930s. The quest to study particles at ever higher energies has pushed experiments from many early, small-scale cosmic-ray observations to a few increasingly complex and powerful machines. Over the decades, electrostatic accelerators such as Van de Graaff generators gave way to fixed-target experiments using circular accelerators, leading to the Proton Synchrotron at CERN and the Alternating Gradient Synchrotron at Brookhaven National Laboratory, which made use of the strong focussing principle. Breakthrough discoveries in the 1970s, made possible by the SPEAR collider at SLAC, firmly established collider-based experiments as the leading tools of the field. Around the same time, the successful operation of CERN's Intersecting Storage Rings (ISR) demonstrated the viability of hadron colliders, paving the way for future generations of machines [8]. This continuous cycle of innovation ultimately culminated in the Large Hadron Collider, the largest and most powerful particle collider ever constructed.

The Large Hadron Collider (LHC) is truly a machine built on extremes: a 27-kilometre-long ring buried roughly 100 metres underground, designed to probe physical processes occurring on the attometre scale. Its accelerator complex consumes power comparable to that of a small city, yet all of this effort delivers to a single proton an energy no greater than that of a falling needle or a mosquito in flight. This seemingly modest energy is enough to accelerate protons to within a fraction of the speed of light and

to recreate conditions similar to those that existed moments after the Big Bang.

4.1. From source to collision

Before hadrons are collided in the LHC, their journey first passes through a complex sequence of pre-accelerators. One could describe this as a quick journey through the history of CERN accelerators before traveling back to the Big Bang. In this section, the paths of both protons and heavy ions towards the LHC will be briefly described. This is the first section in which I go into somewhat more detail on aspects that are often only briefly discussed or omitted altogether. The operation of this machine is far from trivial, and I find it worthwhile to highlight some of the ingenuity that has made its performance possible. It is here already that some of my personal “rainbows” I alluded to in the introductory chapter reside. The CERN accelerator complex is depicted in Figure 4.1.

Accelerating hadrons efficiently in an accelerator complex of this scale is a challenging task, where the discussion of even one small subsystem could exceed the scope of one thesis. As the end user, we are often not aware of each intricacy and small detail of the machine. I believe a foundational familiarity with the machine providing our data is however crucial. Therefore, one must first understand the main concepts and limits of the machine, after which we can better understand the sequence of acceleration for protons and heavier ions.

In the accelerators discussed below, each beam consists of multiple bunches of charged particles. In circular machines, dipole magnets provide the guiding field that forces the particles to follow a closed orbit. Superimposed on this reference trajectory, the particles perform transverse betatron oscillations, which describe their motion around the ideal circular path and is described by

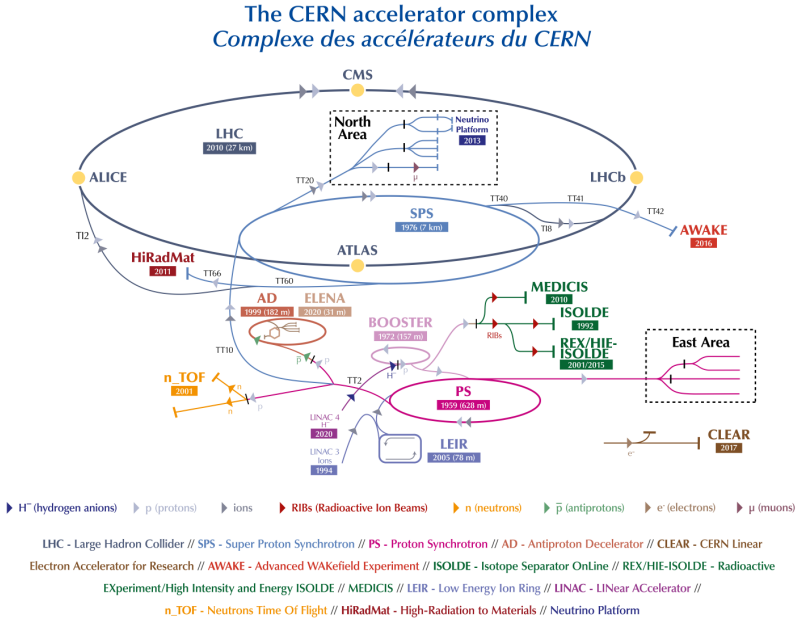


Figure 4.1.: The CERN accelerator complex. Figure from [63]

the machine tune. As the beams are stored and accelerated, the beam will defocus over time as all the particles in the beam have the same charge, leading to phase-space dilution and potential loss of beam quality. In lepton machines, synchrotron radiation provides a natural damping mechanism: the emission of photons reduces transverse oscillation amplitudes and stabilises the beam.

In hadron machines, however, synchrotron radiation is much smaller due to the much larger particle mass. As a consequence, no comparable radiation damping is available, and transverse oscillations are not naturally suppressed. Without additional focusing mechanisms, these oscillations can grow under the influence of imperfections, leading to particle loss. Lost particles eventually interact with the beam pipe, leading to charge deposits and

local heating. In extreme cases, this can trigger superconducting magnet quenches, making beam stability a central challenge in hadron collider design.

A major conceptual breakthrough in overcoming these limitations was the introduction of strong focusing using quadrupole magnets, first developed and implemented at the CERN Intersecting Storage Rings (ISR) [64]. Quadrupole magnets provide focusing in one transverse plane while defocusing in the orthogonal plane; by alternating the orientation of successive quadrupoles, net focusing is achieved in both transverse directions. This principle allows for much stronger overall focusing than earlier weak-focusing designs, significantly reducing beam size and improving stability.

Longitudinal beam stability is provided by radio-frequency (RF) cavities. These cavities generate oscillating electric fields that act on the longitudinal momentum of the particles, effectively grouping them into stable bunches and compensating for energy losses. The RF system therefore plays a dual role: it maintains the bunch structure of the beam and provides the accelerating field for the bunches. Together, the strong-focusing magnet system and RF systems form the backbone of modern circular accelerator operation. A visual representation of all these components is shown in Figure 4.2.

With this setup, 4 important machine limits still need to be understood:

- **Space charge limit:** As a bunch becomes more charged, either due to more particles in the bunch or a higher level of ionization, its electric self-repulsion starts to grow the beamsizes in both the longitudinal and transverse directions. This can result in a beam with growing emittance or bunches that are not well contained in the RF-buckets. The high electric charge furthermore results in a space-charge tune shift, denoted by ΔQ_{sc} . This ΔQ_{sc} can result in the effective

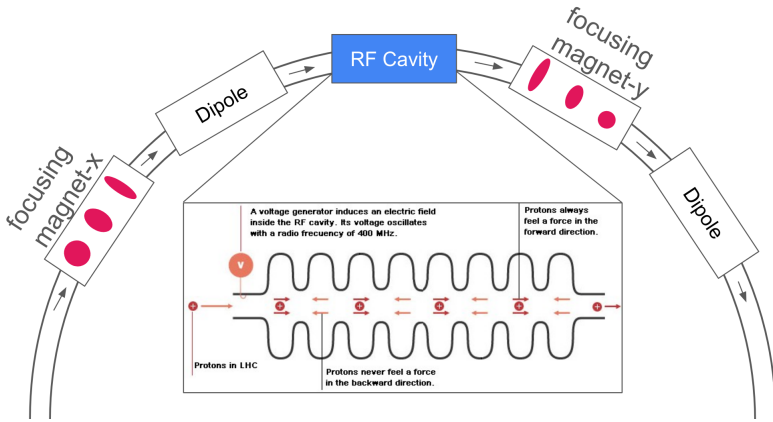


Figure 4.2.: The concept of the three crucial components in a circular accelerator: The dipole magnets, quadrupole magnets and RF cavities. Modified from [65].

beam tune to pass a resonant frequency, making the beam unstable. This effect is mainly present in highly charged beams at lower energies.

- **Beam cooling efficiency:** Particle beams have to be cooled when traversing the accelerator complex. This allows us to have uniform compact beams that can stay longer in the LHC storage rings. This in turn results in higher achievable luminosities during data-taking. Electron cooling is an often applied technique to have a cooler low energy beam. In this process, a cold electron beam is combined with the hot ion beam, transferring heat to the electron beam, which is afterwards removed with a strong magnetic field. The cooling time is proportional to A/q^2 , indicating that this cooling process is more challenging for lightly ionized beams.
- **Intra beam scattering:** In dense beams, particles in the beam can scatter with each other. This results in an exchange of momentum, also leading to wider, longer beams. This effect

is closely linked with the space charge limit, which is the collective effect of the same origin.

- **Interactions with rest gas:** While the beamlines in the accelerator complex are made to be near vacuum, residual rest gasses can still be present in the beamlines. Interactions with these rest gasses can lead to electron exchange (both capture and loss), after which the differently ionized particles will be lost in the magnetic field.

4.1.1. Protons

Protons start their journey from a simple bottle of hydrogen gas. In RUN 3, the 86 metre long Linac4 ionizes the gas as H^- ions and then accelerates it to 160 MeV. The Linac4 system replaced the previous Linac2 system, still in use in RUN 2, which instead ionized the gas to H^+ ions, or bare protons, and accelerated these protons to only 50 MeV. The benefits of this new pre-accelerator are twofold. The higher energy beam increases the beam brilliance (the ratio of the beam intensity over its emittance) but also allows the use of the charge-exchange injection into the PSB. This injection scheme is depicted in Figure 4.3. This technique overcomes space-charge limits by merging the beams of opposite sign charge in one magnetic field before stripping the excess electrons, allowing for much higher intensity beams as this system is essentially loss-free apart from a few percent owing to the stripping efficiency [66].

The beam intensity is increased in a multi-turn injection scheme in the PSB, after which the beams are accelerated to 2 GeV and injected into the Proton Synchrotron (PS), where they are in turn accelerated to 26 GeV. Finally, the Super Proton Synchrotron (SPS) is used to increase their energy further to 450 GeV before they are at last injected into the main LHC ring. Here, the proton bunches are accumulated, accelerated to their peak energy, and finally circulated for up to 50 hours while collisions occur at the four experimental interaction points. During RUN 2 data taking,

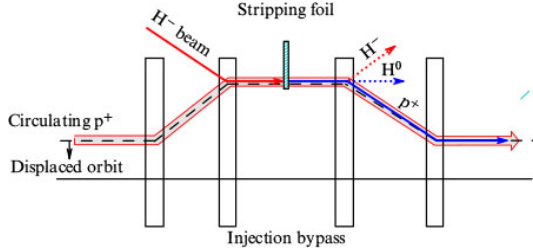


Figure 4.3.: A depiction of charge-exchange injection. Figure taken from [67].

the peak beam energy was 6.5 TeV, resulting in collisions with a center of mass energy (E_{COM}) of 13 TeV. In RUN 3, this beam energy was increased to 6.8 TeV, resulting in a E_{COM} of 13.6 TeV. The beam is divided into 3600 bunches, out of which at most 2808 are filled with protons. These bunches are 25 ns long. Each bunch slot is divided into ten subintervals of 2.5 ns, called RF buckets. Ideally only one RF bucket in each bunch contains protons. RF cavity fields are used to accelerate the particle bunches. Dipole magnets are used to bend the bunch trajectories, while magnetic elements of increasing multipole order—such as quadrupoles, sextupoles, and higher-order magnets—are employed to focus the beams and mitigate instabilities.

4.1.2. Heavy ions

While heavy ions follow a very similar acceleration procedure, the initial stages are very different as creating a highly ionized heavy ion beam is highly non-trivial. Depending on the choice of ion species, they can not always be found in gas form. The most often used heavy ion species for instance, lead, is not available in gas form. These species are heated in a micro oven, where evaporated gas is collected into the Electron Cyclotron Resonance Ion Source.

In this source, RF waves are supplied into the ECRIS cavity, creating a hot electron plasma, while a magnetic trap captures the ion gas in the center of the cavity. An oxygen gas is often also present as a support gas to help create the electron plasma. The hot electron plasma collides with the gas atoms, stripping many electrons and creating highly charged positive ions. These highly ionized ions are then extracted in the afterglow mode. In the case of lead, Pb^{29+} are collected and accelerated in Linac3, after which the beams are collected in multiple bursts into the Low Energy Ion Ring (LEIR). Before entering LEIR, they pass a stripper foil, only collecting Pb^{54+} ions. This ionization level is chosen as it has preferable electron capture and loss rates compared to e.g. Pb^{53+} ions. In LEIR, electron cooling is applied and the luminosity of the bunches are increased. Afterwards, the beams are injected into the PS, following a similar path as the protons. Ions are fully stripped in between the PS and SPS. In a heavy ion run, the number of filled bunches is much lower than in normal pp runs. This might change in LHC Phase II if the PS RF system receives upgrades [68].

4.1.3. Beam property monitoring devices

Controlling the properties of hadron beams in the LHC is very challenging. To understand the development of beam parameters during the run, several monitoring devices are installed over the accelerator complex. In this section, only a few are highlighted that have been most relevant to the work in this thesis.

DCCTs and FBCTs

The beam current is a key quantity that characterizes the total charge circulating in the accelerator. It is measured using two Direct Current Current Transformers (DCCTs) and two Wall Current Transformers (WCTs)—often referred to as the older system: Fast Beam Current Transformers (FBCTs)—per ring. These devices operate based on the fluxgate magnetometer principle. In a simplified picture, the beam and the sensor form a transformer system, where the beam itself acts as the primary

winding with a single turn. The beam current is then proportional to the measured voltage:

$$I_{\text{beam}} = \frac{N_{\text{torus}} \cdot V_{\text{measured}}}{R_{\text{transformer}}}. \quad (4.1)$$

The DCCTs provide a precise measurement of the total beam current integrated over all bunches. In contrast, the FBCTs deliver time-resolved measurements proportional to the charge in each 25 ns bunch slot. The FBCT cannot pick up ghosts because their charge falls below the FBCT threshold current. However, its measurement will include satellites since it measures the whole charge per bunch slot rather than per RF bucket. The DCCT adds the integrated charge of both ghosts and satellites to its measurement.

BSRL

Non-colliding protons contribute to the measured beam current but not to the luminosity, and must therefore be subtracted when determining the effective beam populations for VdM calibrations. Three categories of such non-colliding contributions can be identified. First, there are bunches that are filled but do not participate in collisions according to the filling scheme. The remaining two categories arise from unintended beam populations: so-called “ghost” particles, which occupy nominally empty bunch slots, and “satellite” particles, which are located in colliding bunch slots but outside the nominal RF bucket.

These contributions are measured using Longitudinal Density Monitors (LDMs), which probe the longitudinal distribution of particles in the beam. The main LDM system is the BSRL and is based on single-photon counting of synchrotron radiation. The detector consists of a silicon avalanche photodiode operated in Geiger mode, enabling measurements of the longitudinal beam structure with a time resolution of about 90 ps [69]. This is more than fast enough to have the longitudinal resolution to resolve 2.5 ns RF buckets.

The total bunch charge, corrected for ghost and satellite fractions, f_g and f_{sat}^j , is then given by:

$$I_{\text{bunch}}^j = n_{\text{FBCT}}^j (1 - f_{sat}^j) \frac{N_{\text{DCCT}}(1 - f_g)}{\sum_k n_{\text{FBCT}}^k}. \quad (4.2)$$

DOROS and arcBPMs

Another crucial beam parameter to be measured is the exact transverse beam position in the LHC tube or inside of the experiment. This information is provided by Beam Position Monitors (BPMs), of which two important types in the LHC are the DOROS BPMs inside the CMS experiment and the arcBPMs inside the LHC rings. Both systems use button electrodes to find the transverse beam position. The basic principle of these button electrodes is image current, which is the current induced by the passing bunches on the buttons. The closer the bunch passes by the electrode, the stronger the resulting current. With two oppositely facing buttons, one can use the difference of the currents in both plates as a position measurement.

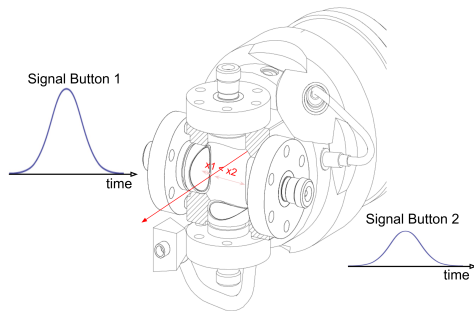


Figure 4.4.: Schematic drawing of the arcBPM buttons. Figure modified from Ref. [70].

While the LHC arcBPMs use this simple setup, the DOROS system process the signal differently from standard BPMs, using diodes to achieve a higher resolution and be sensitive to small oscillations.

BSRT and BBQ

Three remaining crucial parameters in beam optics are the emittance, the beta function, and the machine tune. These parameters play a central role in determining the beam size, stability, and ultimately the achievable luminosity.

The emittance quantifies the spread of particles in transverse phase space, i.e. in position–momentum space (x, p_x) . It can be interpreted as the effective “area” occupied by the beam in phase space and is therefore a measure of the beam quality. A smaller emittance corresponds to a more collimated beam with particles following similar trajectories. The beta function β describes the evolution of the beam envelope along the accelerator. Its value at the interaction point (IP), denoted as β^* , is of particular importance as the transverse beam size at the IP is given by:

$$\sigma = \sqrt{\epsilon\beta^*}. \quad (4.3)$$

A smaller β^* hence results in a smaller beam size at the collision point, which directly enhances the instantaneous luminosity. The emittance can be measured in multiple (non-)destructive ways. A non-destructive way of measuring the emittance is via the Beam Synchrotron Radiation Telescope (BSRT). This detector measures the width of the beam by imaging the synchrotron radiation. The emittance is then extracted from the measured beam width by using formula 4.3 in combination with the β at the BSRT position.

The tune Q of the accelerator is defined as the number of betatron oscillations a particle performs per revolution around the ring. It is defined separately for the x and y plane and is usually the same for both rings. The choice of tune is crucial for beam stability. If the tune approaches a rational value, resonances can occur, leading to a coherent buildup of oscillations and potential beam loss. The beam tunes at the LHC are usually $Q_x = 64.3165$ and $Q_y = 59.3177$. The tune of the beam is measured using the Base-Band Tune (BBQ) system.

4.1.4. Luminosity

The main work in this thesis was performed with the RUN 2 dataset. This dataset was collected in 2016–2018 and corresponds to an integrated luminosity of 138 fb^{-1} of proton–proton collisions. As of the writing of this thesis, the proton–proton program of RUN 3 has concluded, bringing the total integrated luminosity recorded by CMS in this run to approximately 326 fb^{-1} . Several heavy-ion datasets have also been collected during this period, some of which will be discussed in more detail in Chapter 7. The LHC will conclude its current operations with high-intensity beam studies before entering a well-deserved period of rest in July.

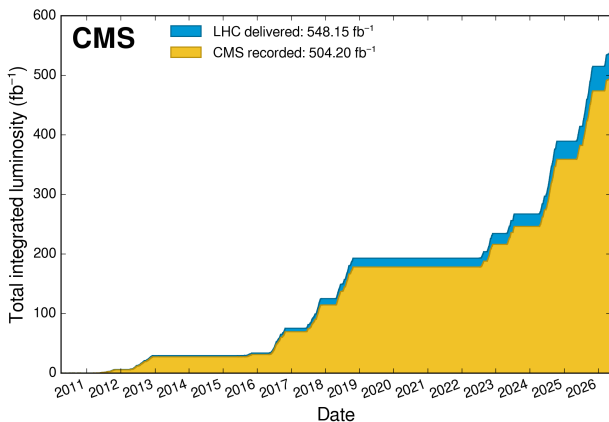


Figure 4.5.: Cumulative luminosity delivered to CMS during stable beams for pp collisions at nominal center-of-mass energy versus day. Figure from Ref. [71].

Over the following three years, the accelerator complex and experiments will undergo extensive upgrades and refurbishment in preparation for the final major phase of the LHC program: the High-Luminosity LHC (HL-LHC). The HL-LHC is expected to deliver approximately 3 ab^{-1} of proton–proton collision data to CMS. Such an increase in luminosity will not merely improve the

precision of existing measurements. It will fundamentally change the range of physics that can be explored. Processes that currently remain statistically limited will become precision measurements, while rare signatures that today produce only a handful of events may become accessible for detailed study!

These gains, however, do not come without challenges. To achieve the required luminosities, the number of simultaneous proton–proton interactions per bunch crossing is expected to increase from roughly 30–60 during RUN 2 and RUN 3, as shown in figure 4.6, to around 200 during HL-LHC operation. This extreme pileup environment requires substantial upgrades to both the accelerator and the detectors. CMS is therefore undergoing a major transformation, including a completely new tracking detector, enhanced trigger capabilities, high-granularity calorimetry in the endcap region, and the introduction of precision timing detectors capable of separating interactions in both space and time.

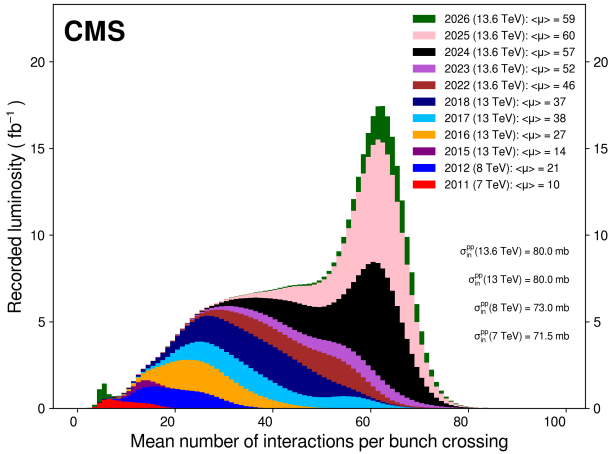


Figure 4.6.: Distribution of the average number of interactions per crossing (pileup) for pp collisions. Figure from Ref. [71].

4.2. The CMS detector: a short overview

During the inception of the CMS detector, one of the primary physics goals was the discovery of the Higgs boson. This objective played a central role in shaping the detector’s design. The most promising signatures for Higgs boson discovery were expected in the so-called “golden channels”: the diphoton final state and the four-lepton final state. To allow for the best sensitivity in these channels, the performance of the electromagnetic calorimeter and muon system were central to the technical design [72].

4.2.1. CMS coordinate system

Before describing the CMS detector, it is necessary to introduce the coordinate system used at CMS, illustrated in Figure 4.7. This system is based on a Cartesian coordinate system in which the z-axis follows the beamline, with the positive direction pointing to the Jura. The x-axis is perpendicular to the beamline in the horizontal plane and points towards the inside of the LHC, while the y-axis points vertically upwards. As this coordinate system is ill suited for describing particle trajectories from the interaction point, two more angles are defined. The angle in the transverse (x-y) plane from the x-axis, ϕ , and the angle to the z-axis, θ . The angle θ is most often transformed into the pseudorapidity, defined as

$$\eta = -\ln \left[\tan \left(\frac{\theta}{2} \right) \right]. \quad (4.4)$$

The pseudorapidity is preferred over θ because differences in η are invariant under Lorentz boosts along the beam direction for highly relativistic particles. In particular, the pseudorapidity difference between two massless particles remains unchanged under longitudinal boosts. This property is especially important in hadron collider physics, where the colliding partons carry different momentum fractions x of their parent hadrons. As a result, the rest frames of the parton-parton collisions will have a

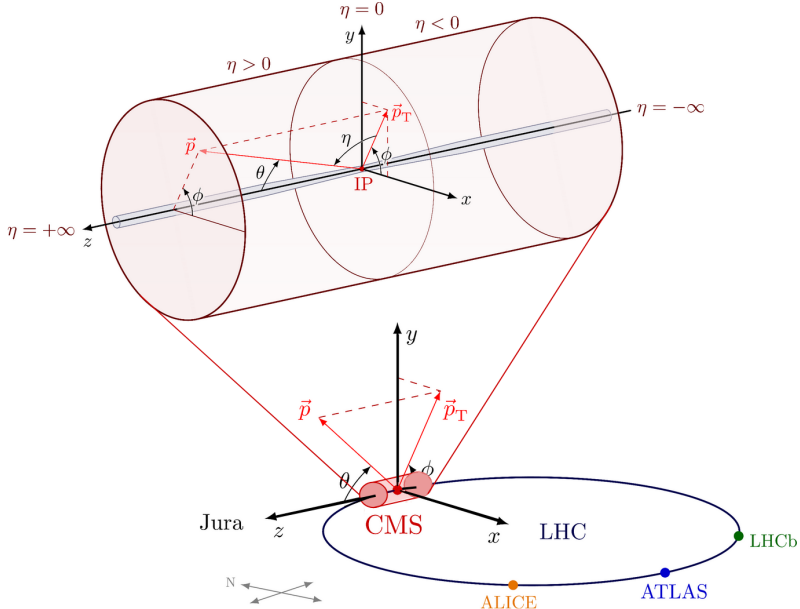


Figure 4.7.: The coordinate system applied in CMS. The positive y direction is upwards, the x -axis points to the inside of the LHC while the z -axis is pointed in the direction of the beam towards the Jura. Figure modified from Ref. [73].

non-zero longitudinal boost with respect to the laboratory frame of reference.

4.2.2. CMS subdetectors

The CMS detector, often colloquially referred to as a “cylindrical onion,” is a multilayered general-purpose detector in which each subdetector is optimized to measure specific classes of particles. A superconducting solenoid magnet, embedded within the detector, provides a uniform magnetic field of approximately 3.8 T, enabling the bending of charged particle trajectories in the transverse plane

and thereby allowing precise momentum measurements.

The detector is divided into a central barrel region and two endcap regions, ensuring full solid-angle coverage around the interaction point. The description of the subdetector systems is therefore ordered according to the typical path particles follow outward from the interaction point. This chapter is based on references [74, 75].

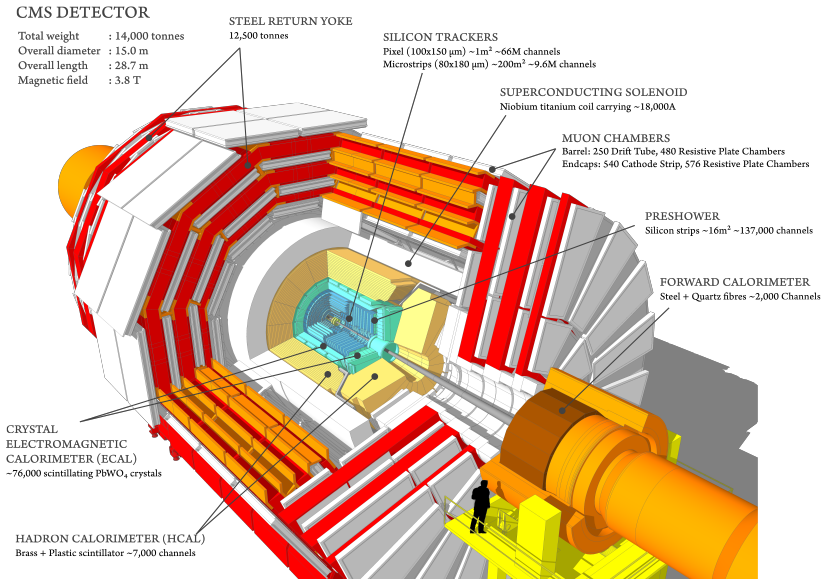


Figure 4.8.: The CMS detector. Figure from Ref. [76].

Tracker

The first detector surrounding the interaction point is the silicon tracker. The tracker is designed to reconstruct the tracks of charged particles going through it and reconstructing track vertices. The track of a charged particle will be curved by the magnetic field, with a radius R of:

$$R = \frac{p_T}{qB}, \quad (4.5)$$

where p_T is the transverse momentum of the particle, q its electric charge, and B the magnetic field strength. The measurement of the track curvature therefore provides an estimate of both the transverse momentum and the charge sign of the particle. Accurate reconstruction of primary and secondary vertices is essential for distinguishing pileup interactions and is also a key ingredient in heavy-flavour jet identification (flavour tagging).

The tracker must be highly radiation tolerant and feature fast readout capabilities, as it is exposed to the highest particle fluxes close to the interaction point. In addition, it must maintain a low material budget, since multiple scattering and energy loss in detector material significantly degrade vertex resolution and impact parameter precision. To meet these requirements, the CMS tracker is composed of two complementary silicon-based systems: the pixel detector, located closest to the interaction point, and the silicon strip detector, which provides tracking at larger radii. The pixel detector was upgraded during the 2016–2017 technical stop, resulting in improved spatial resolution and enhanced reconstruction of both primary and secondary vertices. A comparison of the two designs is shown in Figure 4.9.

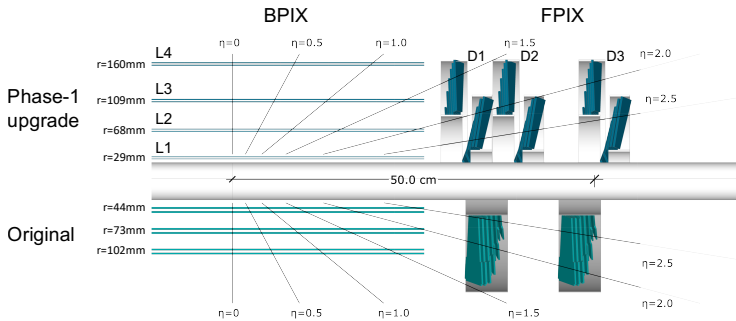


Figure 4.9.: The Pixel tracker as it was installed in 2016 (Original) and after the upgrade in 2017 (Phase-1 upgrade). This tracker will be upgraded for the CMS Phase-2 upgrades. Figure from Ref. [74].

The silicon strip detector surrounds the pixel detector and extends the tracking coverage up to a pseudorapidity of $|\eta| < 2.5$. Its layout is shown in Figure 4.10. Two types of strip modules are employed: single-sided and double-sided strip detectors, the latter commonly referred to as stereo modules. In stereo modules, the second silicon strip sensor is mounted at a small relative angle of approximately 100 mrad with respect to the first. This configuration provides an additional spatial coordinate for each hit: the z -coordinate in the barrel region and the radial coordinate r in the endcap disks.

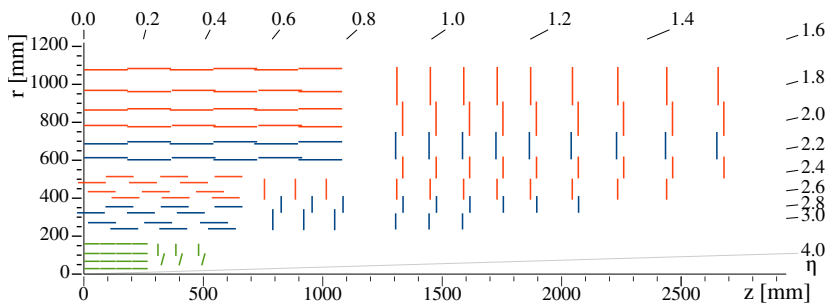


Figure 4.10.: The layout of the strip tracker. The orange lines show the placement of the single-sided silicon strip modules, and the blue lines the stereo modules. The pixel detector layers are shown in green. Figure from Ref. [74].

The choice of stereo angle represents a compromise between the achievable z resolution, the complexity of detector construction, and the rate of fake tracks arising from hit ambiguities. The impact of the stereo configuration on the z resolution and on the occurrence of fake-track ambiguities is illustrated in Figure 4.11.

Electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) surrounds the tracker and is designed to measure the energy of electrons and photons, as well as a fraction of the energy of hadronic particles. It is a

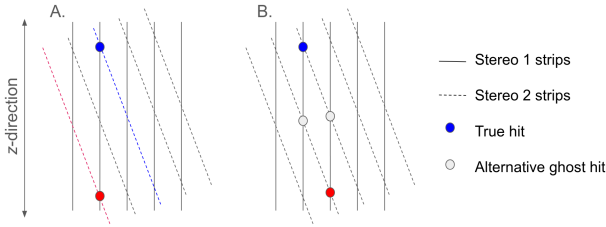


Figure 4.11.: Left: sensitivity to additional spacial coordinate. Right: Ambiguity of track locations with multiple hits.

homogeneous, fine-grained calorimeter composed of lead tungstate (PbWO_4) crystals.

High-energy photons entering the ECAL interact with the material and convert into electron–positron pairs via pair production. These electrons and positrons subsequently radiate bremsstrahlung photons, which in turn produce additional pairs. This cascade process repeats, resulting in an electromagnetic shower that develops within the crystals, analogous to electromagnetic showers in the atmosphere. The deposited energy is converted into scintillation light, which is collected by photodetectors. This destructive measurement allows the energy of the incoming particle to be reconstructed with high precision. The ECAL barrel covers the pseudorapidity range up to $|\eta| < 1.479$, while the endcap regions extend the coverage to $|\eta| < 3.0$. The total thickness of the ECAL corresponds to approximately 26 radiation lengths, ensuring full containment of electromagnetic showers up to high energies.

Hadronic calorimeter

The ECAL is followed by the hadronic calorimeter (HCAL). The barrel and endcap HCALs are situated in a similar pseudorapidity range of $|\eta| < 1.3$ and $1.3 < |\eta| = 3$. The HCAL is a sampling calorimeter, meaning it is a combination of alternating layers

of absorbing material and scintillating material. The absorbing material is a combination of layers of steel and brass, while the active material is plastic with wavelength shifting fibers. The total thickness of the HCAL is roughly 5.8 - 10.6 interaction lengths in the barrel region.

While the ECAL is designed to measure the energy of electrons and photons and the HCAL is designed for hadronic particles, large energy electrons/photons can also deposit energy in the HCAL further away from the interaction point and hadrons also deposit a small part of their energy in the ECAL, which has a thickness of roughly 1.1 interaction lengths for hadrons.

Muon system

The outermost layers of the CMS detector are instrumented with the muon system. This system employs four different detector technologies: Drift Tubes (DT), Cathode Strip Chambers (CSC), Resistive Plate Chambers (RPC), and the more recent Gas Electron Multiplier (GEM) detectors. These detectors are installed outside the solenoid, embedded in the gaps of the steel return yoke, and are designed to measure muons, which are the only charged particles capable of traversing the entire detector. The layout of the muon system is shown in Figure 4.12.

All subsystems are based on gaseous detector technologies. As charged particles traverse the detector, they ionize the gas mixture, producing electron-ion pairs. The liberated electrons drift towards the anode under an applied electric field, inducing a measurable signal. The choice of gas mixture is crucial and is optimized to ensure efficient ionization, high electron drift velocity, and effective photon quenching. Many of the currently used gas mixtures provide excellent detector performance but have a high Global Warming Potential (GWP). To mitigate the environmental impact, significant efforts are underway to develop more sustainable alternatives. In this context, the ECOgas@GIF++

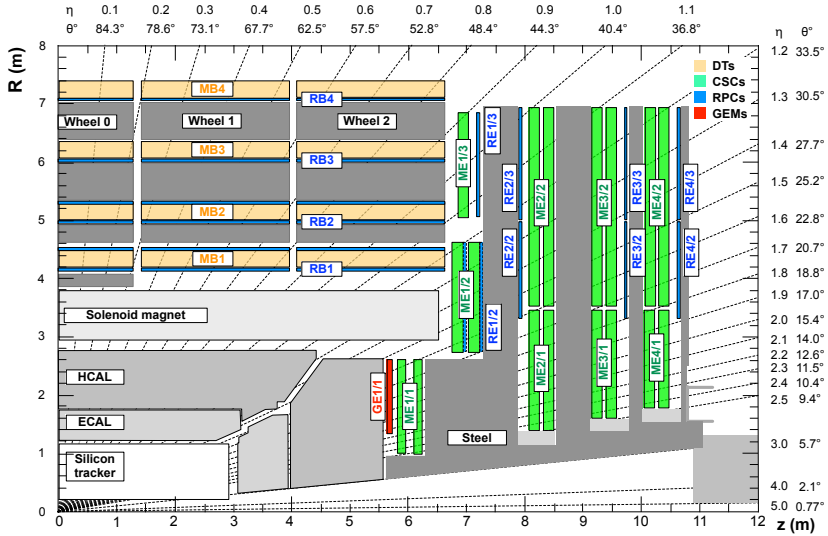


Figure 4.12.: Layout of the muon system inside the CMS experiment at the start of Run 3. Since, several iRPC chambers have been installed to validate the performance of these designs for Phase II. Figure from Ref. [74].

collaboration, involving several CERN experiments and detector groups, is pursuing an extensive R&D program to identify eco-friendly gas mixtures.

The DT system is employed in the barrel region, where the particle rates are moderate and the magnetic field rather homogeneous. DTs provide precise position measurements by recording the drift time of ionization electrons in gas-filled tubes, allowing for accurate reconstruction of muon trajectories.

In the endcap regions, where particle fluxes and radiation levels are significantly higher and the magnetic field is less uniform, Cathode Strip Chambers (CSCs) are employed. These detectors

combine fast response times with robust performance in strong and non-uniform magnetic fields, making them well suited for operation in the harsher endcap environment. CSCs consist of layers of positively charged anode wires interleaved with negatively charged cathode strips, oriented perpendicular to each other. As a muon traverses the gas volume, it ionizes the gas, producing electron-ion pairs. The electrons drift towards the anode wires, while the ions are collected on the cathode strips. The combination of wire and strip readout provides two spatial coordinates for each hit, enabling precise reconstruction of muon trajectories.

Resistive Plate Chambers (RPCs) are installed in both the barrel and endcap regions and are primarily used for triggering due to their excellent time resolution. RPCs are gaseous detectors composed of two parallel resistive plates, made of bakelite, separated by a thin gas gap. A high voltage is applied across the plates, creating a strong and uniform electric field. When a charged particle passes through the gas, it induces a fast signal on external readout strips. These readout strips are typically arranged in one coordinate per chamber, providing coarse spatial information but very precise timing on the order of a few nanoseconds.

The RPC system is arranged in multiple layers interleaved with the DT chambers in the barrel and the CSC chambers in the endcaps. This layout provides redundancy in muon detection and improves the robustness of the trigger system. While RPCs have a lower spatial resolution compared to DTs and CSCs, their fast timing capability makes them essential for bunch crossing identification and for assigning muons to the correct LHC bunch crossing.

More recently, GEM detectors have been introduced in the forward regions to cope with the increasing particle rates expected at higher luminosities. GEMs offer high rate capability, good spatial resolution, and excellent radiation tolerance. In addition,

upgraded RPC systems (iRPCs) are being developed to further improve performance under high-luminosity conditions. These newer technologies, GEMs and iRPCs, will be discussed in more detail in Chapter 4.3.

4.2.3. Forward detectors

HF:

The forward HCAL (HF) is located 11.2 m from the interaction point on either side of the detector and provides the most forward calorimetry for standard pp-collision data-taking at a pseudorapidity range of $3 < |\eta| < 5.2$. The detector is a Cherenkov calorimeter made of steel absorber material and embedded radiation-hard quartz fibers. The calorimeter is installed with 2 types of fibers: long fibers that run over the full depth and short fibers that start at a depth of 22 cm. The signal in each HF cell (or tower) is combined from the long and short fibers. Some sensitivity to EM and hadronic showers can be reached by estimating these components as:

$$\begin{aligned} E_{\text{tower}} &= E_{\text{long}} + E_{\text{short}} \\ E_{\text{EM}} &= E_{\text{long}} - E_{\text{short}} \\ E_{\text{Had}} &= 2 \cdot E_{\text{short}} \end{aligned} \tag{4.6}$$

While this artificial separation into quasi-EM and quasi-hadronic calorimetry seems attractive, this results in a challenging particle flow algorithm which is hard to interpret at the physics level. Some of these challenges will be discussed in chapter 7.

PPS:

The Total, elastic and diffractive cross-section measurement (TOTEM) experiment has several detectors scattered at distances of more than 200m around the CMS cavern. In 2015, the TOTEM and CMS collaborations started coordinating the use of their detectors to perform combined measurements. This allows for CMS

measurements using the CMS-TOTEM Proton Precision Spectrometer (CT-PPS, or simply PPS). The PPS system consists of tracking and timing stations located inside the LHC tunnel on both sides of the CMS detector, approximately 200 m from the interaction point. Tracking is provided by silicon pixel detectors, while timing measurements are achieved using diamond detectors. The detectors are housed in movable beamline insertions, known as “Roman pots,” which allow them to be remotely positioned to within a few millimeters of the beam after stable beam conditions are reached.

The PPS can detect very forward produced protons, their momentum can be described by their fractional momentum loss, defined as

$$\xi = \frac{|\vec{p}_i| - |\vec{p}_f|}{|\vec{p}_i|}. \quad (4.7)$$

As the Roman pots are roughly 200 m away from the IP, the displacement of these protons is dependent on the beam optics, so the PPS requires calibration before measurements of displacement can be translated into the fractional momentum loss. Tagging these protons is crucial as this means the proton has undergone minimal interaction, indicating an event where there was no color exchange between the protons.

ZDC:

Another forward detector is the Zero Degree Calorimeter (ZDC). The ZDCs are installed symmetrically on both sides of the CMS interaction point at a distance of approximately 140 m, embedded in the neutral particle absorber (TAN). They are designed to detect neutral particles, primarily photons and neutrons, emitted at very small angles relative to the beam direction ($|\eta| \geq 8.3$). As these neutral particles are unaffected by the LHC beam optics, the ZDC acceptance is more easily understood. These detectors are only installed during low pileup data-taking periods (mainly

heavy-ion run periods).

The ZDCs are sampling calorimeters composed of alternating layers of tungsten and quartz fibers. Charged particles produced in showers traversing the detector generate Cherenkov light in the quartz fibers, which is then collected and read out by photodetectors. Separate electromagnetic and hadronic sections allow for a coarse discrimination between photons and neutrons, similar to the HF. In between these separate sections, the Beam RATE for Neutrals (BRAN) detector is installed. Unlike the ZDC, the BRAN is always present and is used, among other purposes, as a luminometer.

FSC: The Forward Shower Counters (FSCs) are a set of scintillation detectors installed along the beamline at several locations on both sides of the CMS interaction point, at distances of 59 m, 85 m and 114 m. Their primary purpose is to detect showers induced by very forward particles interacting with beamline elements. The FSCs consist of plastic scintillator tiles coupled to photodetectors. While they do not provide precise energy or position measurements, they are highly effective in tagging the presence of forward activity. The FSC system is particularly useful for luminosity monitoring and the identification of diffractive and exclusive processes, where large rapidity gaps are expected. By detecting or vetoing forward activity, the FSCs help to distinguish between different classes of events and to suppress background contributions in analyses sensitive to forward particle production.

4.3. Upgrades for High-luminosity LHC

In July 2026, the last collisions of the Run 3 campaign were recorded. In the following years, the LHC will enter the third Long Shutdown (LS3), during which it will undergo significant upgrades over approximately four years in preparation for the next operational phase. The primary goal of these upgrades is

to increase the instantaneous luminosity of the machine through stronger focusing magnets and various improvements to the injector chain and collider optics. This is expected to increase the yearly integrated luminosity to approximately 250 fb^{-1} , with a total target of around 3 ab^{-1} over the full HL-LHC program. No substantial increase in the centre-of-mass energy is foreseen; however, improved magnet performance may allow operation closer to the nominal design energy of 14 TeV.

To cope with the increased instantaneous luminosity and the corresponding rise in pileup, the CMS detector will also undergo a major upgrade program during LS3. As previously explained, several new or upgraded subdetectors will be installed. A completely new silicon tracker will replace the existing system, which has accumulated significant radiation damage over Run 2 and Run 3, resulting in substantially improved tracking performance and flavour tagging capabilities. The MIP Timing Detector (MTD) will provide precision timing measurements for charged particles with a projected resolution of 30–60 ps. This will enable four-dimensional vertex reconstruction (4D tracking), significantly improving track-to-vertex association and object reconstruction/identification.

In the forward regions, the CMS endcaps will be equipped with a High Granularity Calorimeter (HGCal), enabling high-precision particle-flow reconstruction in a high-radiation environment. The Beam Radiation, Instrumentation and Luminosity (BRIL) system will also be upgraded, including a new dedicated bunch-by-bunch luminometer, the Fast Beam Condition Monitor (FBCM). While these upgrades provide a wide range of new physics opportunities, this section focuses on the iRPC and GEM upgrades of the muon system, for which the author contributed to production and validation activities in collaboration with other members of the CMS group at Ghent University. These activities represent a minor

component of the overall work presented in this thesis.

4.3.1. iRPCs

In this section, the production and validation of two new improved Resistive Plate Chamber (iRPC) stations in the CMS endcap region are discussed: RE3/1 and RE4/1. These stations will be installed in the third and fourth endcap disks, extending the RPC coverage up to a pseudorapidity of $|\eta| < 2.4$. This extension was already foreseen in the original CMS Technical Design Report (TDR), but the previous RPC design lacked sufficient rate capability to operate reliably in the high particle flux environment of the forward regions. The additional instrumentation improves muon reconstruction in a region that is currently instrumented only with Cathode Strip Chambers (CSCs). Furthermore, the improved timing performance enhances trigger efficiency and background rejection.

The new iRPC detectors are based on an upgraded version of the standard RPC design and include several key improvements:

1. readout of the RPC strips is done on both sides. This allows to extract the position of the signal on the strip based on the time difference between both readouts.
2. The gaps of the new RPC modules are narrower than the previous version. Furthermore, the operating voltage, and hence the avalanche charge, are reduced. This modification increases the rate capability of the detector.
3. Improved electronics in the Front-End Board (FEB) retain better efficiency, even though the avalanche charge will be reduced.

During the production of the new iRPC chambers, several Quality Control (QC) tests were performed to assess the quality of

individual components and assembled chambers, and to identify and repair or reject defective elements. As an illustrative example, the first batch of RE3/1 gas gaps delivered from KODEL (Korea) exhibited significant mechanical bending, with deformations of the order of 1–3 cm. To ensure proper chamber assembly and uniform detector performance across the full active area, it was essential to minimise this bending. In addition, as these detectors are expected to be operational for decades without intervention, prolonged mechanical stress could lead to damage of the peripheral spacers, potentially resulting in gas leaks after installation in the cavern.

To mitigate these effects and prevent further deformation, all gas gaps were stored in a climate-controlled environment, maintaining a relative humidity between 30% and 40% using a portable humidifier. Following the implementation of these measures, subsequent batches of gas gaps received at both Ghent and CERN did not exhibit any bending issues.

QC2: gap leak test, spacer test, and HV test

In the first QC test performed at Ghent, incoming gas gaps undergo a visual inspection to identify any obvious defects. Subsequently, the gaps are filled with argon gas to a pressure of approximately 15 mbar. The gas inlets are then sealed, and the system is left to stabilize for 10 minutes, followed by a 10 minute measurement period during which the internal pressure is monitored. A significant pressure drop, exceeding 0.4 mbar over 10 minutes, is indicative of a gas leak. In such cases, the gap is either repaired or rejected for further use.

In a second stage of the test, mechanical integrity is assessed by applying pressure at the positions of the internal spacers that maintain the electrode separation. Defective or broken spacers are identified through sudden pressure spikes, which arise from the changes in gas volume caused by the applied force.

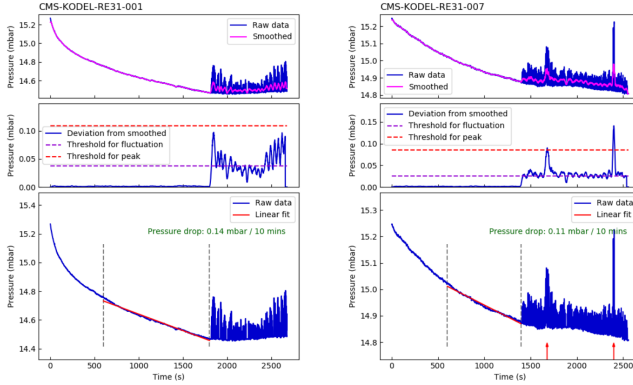


Figure 4.13.: Examples of the gas leak test and spacer test for gaps to be used in RPC chambers. Left: a good gap, with acceptable pressure drop (the rejection threshold is 0.4 mbar / 10 mins) and no apparent broken spacers. Right: a rejected gap, with acceptable pressure drop but two broken spacers.

Following the initial QC tests, the dark current (DC) of the gas gaps is measured. Prior to these measurements, all gaps are placed in the AirCo room and flushed with a standard RPC gas mixture for at least 48 hours. This procedure stabilizes the detector response and helps remove residual impurities or dust particles inside the gaps. The iRPC gaps are expected to exhibit ohmic behaviour, and the measured currents are required to remain below $3 \mu\text{A}$ at an applied voltage of 7.3 kV. Gaps that do not satisfy this criterion are rejected or subjected to further inspection.

After passing these tests, the assembly of the full chamber is performed. Once assembly is completed, a similar quality control step (QC3) is carried out. In this stage, both a gas leak test and a dark current measurement are repeated on the fully assembled iRPC module to ensure proper detector integrity and performance.

After this, the chambers are transported to CERN for further QC tests before installation inside the CMS cavern.

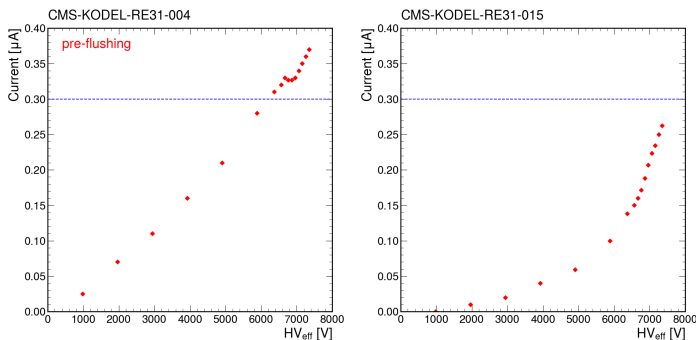


Figure 4.14.: Dark currents as a function of the applied voltage for a gap showing too high currents before flushing treatment (left) and a good gap (right).

4.3.2. GEMs

Next to the iRPCs, Gas Electron Multiplier (GEM) detectors will also be installed in the muon system. Similar to RPCs, GEMs are gaseous avalanche detectors in which primary ionization electrons initiate a cascade of secondary ionizations in the strong electric field.

In contrast to RPCs, the gas volume in GEM detectors is segmented into multiple regions by one or more thin polymer foils, typically made of Kapton and coated with copper on both sides. These foils are perforated with a high density of microscopic holes with diameters of approximately $70\text{ }\mu\text{m}$. The region between the cathode and the first GEM foil, referred to as the drift region, is operated at a moderate electric field, sufficient to transport primary electrons towards the amplification stages but not strong enough to initiate avalanche multiplication. When electrons en-

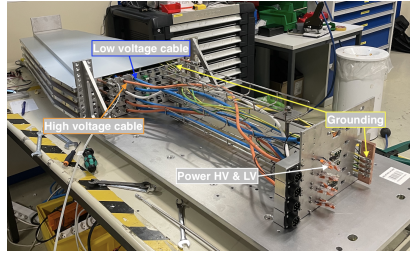


Figure 4.15.: A stack of 4 Modules assembled at B904 in CERN.

ter the GEM holes, they encounter a very strong electric field, leading to localized avalanche multiplication within the holes. This confinement of the amplification process results in high rate capability, reduced discharge probability, and excellent spatial resolution.

My contribution to GEM production at CERN was helping the assembly of GEM stacks from modules produced at other institutes such as Ghent. An illustrative stack is shown in Figure 4.15. During the construction of GEM stacks and modules, similar QC tests were performed as for the iRPC QC tests. This includes gas leak tests and noise readings of the GEM strips in so-called S-curves, as shown in Figure 4.16.

4.4. Trigger System

In early fixed-target and bubble chamber experiments, event rates were sufficiently low that all interactions could be recorded, often even as photographic images for offline analysis. As accelerator technology progressed, interaction rates increased to the point where selective readout became necessary, leading to the development of the first trigger systems.

At the LHC, this challenge has reached an extreme scale. The LHC produces events in the CMS detector at a rate of 40 MHz.

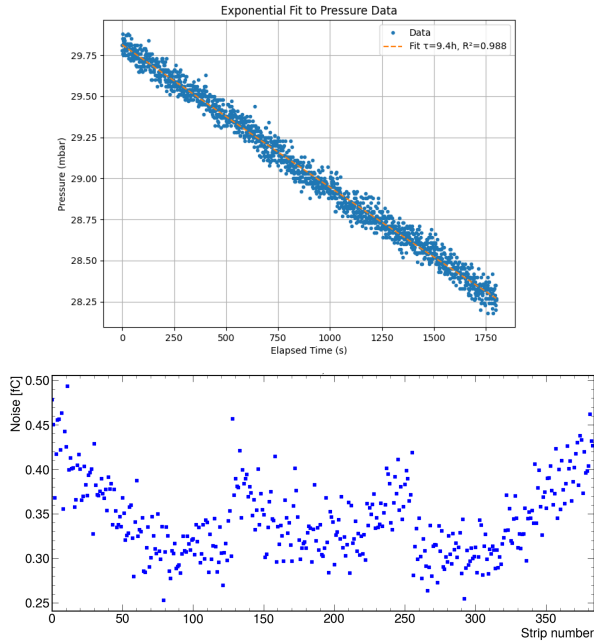


Figure 4.16.: Illustrative plots of leak test (QC 3b, top) and electronic test (QC 7, bottom). The peak structures in the noise are due to the routing geometry of the strips to the 6 VFATs that are connected to the Optohybrid/ASIAGO.

Recording all events at this rate is far beyond the capabilities of any data acquisition or storage system. Moreover, only a tiny fraction of these collisions contain signatures of interest, such as rare Standard Model processes or potential signals of new physics. Most events contain QCD interactions that are not of much interest.

To address this, the CMS experiment employs a real-time event selection system, known as the trigger, to reduce the event rate to a manageable level while retaining high efficiency for relevant physics processes. The CMS trigger system reduces the input rate from 40 MHz to approximately $\mathcal{O}(\text{kHz})$ for permanent storage. This is achieved through a two-level architecture, consisting of a hardware-based Level-1 (L1) trigger and a software-based High-Level Trigger (HLT), which are described in the following sections.

4.4.1. Level-1 Trigger

The Level-1 (L1) trigger performs the first stage of event selection, reducing the collision rate from 40 MHz to approximately 100 kHz. As this decision must be taken within a fixed latency of a few microseconds, detector signals are temporarily stored in pipeline buffers while the trigger decision is formed.

To meet these stringent timing constraints, the L1 trigger is implemented in custom hardware located close to the detector, minimizing signal transmission delays. The system is based on a combination of Field-Programmable Gate Arrays (FPGAs) and Application-Specific Integrated Circuits (ASICs). The L1 trigger employs fast, simplified reconstruction algorithms to identify high-level physics objects such as energetic muons, electrons, photons, and jets. Events satisfying predefined selection criteria generate a Level-1 Accept (L1A) signal and are forwarded to the next stage of processing.

If the instantaneous L1 accept rate exceeds the capacity of the downstream readout and data acquisition systems, the Trigger Control System (TCS) temporarily inhibits further L1A signals, introducing so-called “dead time”. This mechanism prevents buffer overflows at the cost of a small loss in data-taking efficiency.

4.4.2. High-Level Trigger

Events accepted by the L1 trigger are transferred to the High-Level Trigger (HLT), which is implemented as a large computing farm located on the surface above the experimental cavern. The HLT further reduces the event rate from $\mathcal{O}(100\text{ kHz})$ to $\mathcal{O}(\text{kHz})$ for permanent storage.

In contrast to the L1 trigger, the HLT has access to the full detector readout and employs software-based event reconstruction algorithms. While these algorithms are conceptually similar to the offline reconstruction, they are optimized for speed to allow processing times of order 100 ms per event. The HLT performs refined object reconstruction and applies more sophisticated selection criteria.

4.4.3. ZeroBias and Minimum Bias Triggers

In addition to physics-driven trigger paths, dedicated triggers are used to record unbiased or minimally biased samples of proton-proton collisions.

ZeroBias triggers select events based solely on the presence of colliding bunches, without applying any physics-based selection criteria. These events provide an unbiased snapshot of the detector conditions and are primarily used for detector performance studies and luminosity measurements.

Minimum Bias (MinBias) triggers apply very loose selection requirements, typically based on activity in forward detectors or

minimal energy deposits in the calorimeters. These triggers aim to sample generic inelastic proton-proton interactions while maintaining a high acceptance efficiency. Such datasets are essential for studying soft QCD processes and for tuning Monte Carlo simulations.

4.4.4. Scouting and parking

To extend the physics reach beyond the limitations imposed by the standard trigger bandwidth, CMS employs dedicated data-taking strategies known as scouting [77] and parking [78].

In data scouting, a reduced event content is recorded directly at the HLT level, storing only high-level reconstructed objects instead of the full detector information. This significantly decreases the event size, allowing data to be recorded at much higher rates. Scouting is particularly useful for searches targeting low-mass resonances or rare signatures that would otherwise be rejected by standard trigger thresholds.

In contrast, data parking involves recording events at a higher rate than can be processed promptly, storing them for deferred offline reconstruction during periods without data taking. This strategy allows CMS to collect larger datasets for analyses that are not time-critical, effectively increasing the usable luminosity for certain physics studies. Neither of these datastreams are used in the work presented in this thesis.

4.5. Particle flow reconstruction

In section 4.2, we have discussed the different subdetectors in CMS. Each subdetector system provides low-level data such as a charge excess of a single pixel in the tracker or a raw PMT signal in a HF tower. This data is however not very interpretable without relating this information to all other observed signals. To bridge

the gap from rough detector-level observables to the physics-level interpretation of them (e.g. a muon with reconstructed p_T, η and ϕ coordinates), CMS relies on the Particle Flow (PF) algorithm. This step is called event reconstruction.

4.5.1. Track reconstruction

In the high-pileup environment of the LHC, tracking and vertex reconstruction are among the most critical and computationally demanding steps in event reconstruction. A single proton-proton interaction typically produces on the order of 30 charged particles within the acceptance of the tracking system. With average pileup conditions in CMS reaching values of up to 60 interactions per bunch crossing, this results in thousands of charged particles traversing the tracker simultaneously. In the presence of the strong magnetic field, charged particles follow helical trajectories. These trajectories can be parameterized by five quantities: the signed transverse momentum p_T (where the sign reflects the particle charge), the azimuthal angle ϕ in the transverse plane, the pseudorapidity η , the transverse impact parameter d_{xy} (defined as the distance of closest approach of the track to the beamline or primary vertex in the transverse plane), and the longitudinal impact parameter d_z (measured along the beam axis).

Track reconstruction begins with the formation of seeds, which provide initial estimates of the track parameters. These seeds are first constructed from a small number of hits in the inner detector, where the hit density is highest and the spatial resolution is optimal. Starting from these seeds, track candidates are built using a combinatorial track finding algorithm based on the Kalman filter formalism. At each step, the track parameters are propagated to the next detector layer, taking into account the magnetic field and material effects. Compatible hits are added to the track candidate, and the track parameters are updated iteratively. The Kalman filter provides estimates of the five track

parameters but also provides their associated uncertainties.

To reduce the computational cost and improve the reconstruction efficiency, the tracking is performed in multiple iterations. Early iterations focus on reconstructing high-quality tracks originating from the primary vertex, using tight selection criteria. Hits associated with successfully reconstructed tracks are then removed from consideration, significantly reducing the combinatorial complexity for subsequent iterations. Later iterations employ progressively looser criteria to recover more challenging tracks, such as low- p_T particles, displaced tracks, or tracks with missing hits. After the track finding stage, track candidates are refitted two times using an inward and an outward going Kalman filter, with more complicated propagators than in the track finding stage, and the final track parameters are determined from a combination of both fits.

4.5.2. Muon reconstruction

Muon reconstruction in CMS combines information from both the inner tracker and the muon system. “Tracker muons” are identified by matching tracks reconstructed in the silicon tracker with compatible signals in the muon detectors. Tracks in the muon system follow a similar, but simplified algorithm as is employed in the tracker. Track seeds are here considered from clusters of nearby hits in the DTs or CSCs. Reconstructed tracks here are referred to as “standalone” muons. These muon tracks are then combined with tracker tracks to define the global muon and its estimated parameters. The extremely long lever-arm of reconstructed muons results in a very precise reconstruction of its parameters such as its p_T , η and ϕ , compared to other objects such as jets or electrons. After reconstructing the muons, the PF components used in their reconstruction are masked for the further steps of electron, photon and hadron reconstruction.

4.5.3. Vertexing

After the tracks are reconstructed, vertices are reconstructed from these tracks. A vertex is described as the point in space where an interaction took place. Primary vertex (PV) reconstruction is a crucial component of event reconstruction, especially in high pileup conditions. Vertices are reconstructed by clustering tracks based on their compatibility in the longitudinal direction and fitting them to a common point. The primary vertex is typically defined as the vertex with the largest sum of squared transverse momenta of its associated tracks.

Accurate vertex reconstruction allows for the separation of multiple proton-proton interactions within the same event and plays a key role in object reconstruction, such as lepton isolation and jet flavour tagging.

4.5.4. Electron and photon reconstruction

Bremsstrahlung emission significantly degrades both the energy measurement and the track reconstruction of electrons compared to muons. To mitigate this effect and retain accurate track reconstruction, electron candidates are refitted using a Gaussian Sum Filter (GSF) [79], instead of the standard Kalman filter [80].

The GSF models the electron energy loss as a weighted sum of multiple Gaussian components, rather than a single Gaussian approximation. This allows for a more accurate description of the highly non-Gaussian energy loss due to bremsstrahlung and improves the reconstruction of electron trajectories, which can exhibit substantial kinks as photons are emitted. In addition, to recover the energy radiated via bremsstrahlung, the electron track is extrapolated tangentially from successive tracker layers toward the ECAL. Energy deposits consistent with this extrapolation are associated with the electron candidate and interpreted as originating from emitted photons. As a result, an electron is

reconstructed as a combination of a charged-particle track and multiple energy deposits in the ECAL.

Simultaneously with the reconstruction of electrons, isolated photons are identified, as both rely on similar signatures in the ECAL. Since photons are not used in the physics analyses presented in this thesis, their reconstruction will not be discussed further. Just like for muons, the corresponding PF candidates of electrons and photons are removed from the pool of inputs used for hadronic reconstruction. Non-isolated photons and hadrons are then reconstructed with the leftover PF objects.

4.5.5. Jet reconstruction

As previously explained, partons produced in the hard interaction undergo parton showering and hadronization, resulting in collimated sprays of particles known as jets. The reconstruction of jets as a pseudoparticle provides an effective way to infer the kinematic properties of the original quarks and gluons produced in the primary interaction. To allow for meaningful comparisons between theoretical predictions and experimental measurements, jet definitions must be robust against soft radiation and collinear splittings. This property, known as infrared and collinear (IRC) safety, ensures that the addition of low-energy particles or the splitting of a particle into collinear fragments does not significantly alter the reconstructed jets.

In CMS, jets are reconstructed using sequential clustering algorithms that satisfy IRC safety. The most widely used algorithm is the anti- k_T algorithm, which clusters particles based on their relative transverse momenta and angular separation. This algorithm produces nearly conical jets and is both experimentally robust and theoretically well-defined. A distance parameter R controls the size of the reconstructed jets and is typically chosen as $R = 0.4$ or 0.8 depending on the analysis. The algorithm applies a distance metric between two PF candidates as:

$$d_{ij} = \min(p_{Ti}^{-2}, p_{Tj}^{-2}) \frac{\Delta_{ij}^2}{R^2} \quad (4.8)$$

$$d_{iB} = k_{Ti}^{-2}, \quad (4.9)$$

where p_T^i is the transverse momentum of entity i and $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, with y and ϕ respectively the rapidity and azimuth angle. R is the radius parameter and gives a measure of how wide jets are allowed to be. The clustering proceeds by identifying the smallest of the distances. If this is a d_{ij} , entities i and j are recombined into a new jet. While if it is d_{iB} , i is called a jet and is removed from the list of entities.

The flavour of the parton initiating a jet is an important piece of information that can be inferred using flavour tagging techniques. In the case of b quarks, this identification is facilitated by the relatively long lifetime of B hadrons produced during hadronization. These hadrons travel a measurable distance, typically of the order of a few millimetres, before decaying. As a result, their decay products form tracks that originate from a displaced secondary vertex rather than the primary interaction vertex. This characteristic displacement, along with additional information such as track impact parameters and the presence of soft leptons from leptonic hadron decays, provides strong discrimination between b -jets and jets originating from lighter quarks or gluons. Early flavour tagging algorithms exploited these features using simple, cut-based or likelihood-based approaches.

Modern techniques employ machine learning methods to optimally combine a large number of observables. In the analysis presented in Chapter 5, the DEEPJET algorithm is used for b -jet identification. This algorithm relies on Long Short-Term Memory (LSTM) layers. Since, even more advanced architectures are available that show significant improvements in performance. The latest models are based on transformer architectures that make use of

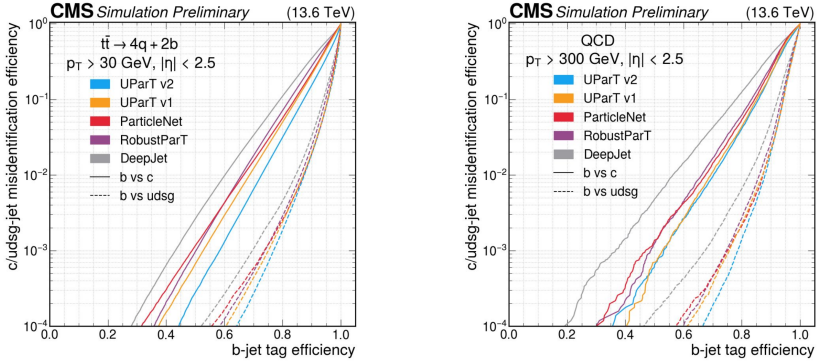


Figure 4.17.: b-tagging ROC curves. UParTv2 shows improved performance for both c and light jet rejection.

the attention-mechanism. The current state-of-the-art model is the UParTv2 model [81]. This improved tagger is used in the studies presented in chapter 6. The increased performance of this model is shown in receiver operating characteristic (ROC) curves in figure 4.17.

4.5.6. Missing transverse momentum

In the previous subsections, we shortly explained the reconstruction of electrons, muons, photons and jets. The only SM particles that escape detection are neutrinos. There could also be weakly interacting particles beyond the Standard Model that evade detection. The missing transverse momentum (p_T^{miss}) is a key observable in proton-proton collisions, sensitive to these particles that do not interact with the detector. It is defined as the negative vector sum of the transverse momenta of all reconstructed particles in the event:

$$p_T^{\text{miss}} = - \sum_i \vec{p}_{T,i}, \quad (4.10)$$

where the sum runs over all reconstructed PF candidates.

5 Top quark pair production in association with a W boson

5.1. Introduction

In section 3.4.4, the importance of the $t\bar{t}W$ process to our understanding of $t\bar{t}t\bar{t}$ production was already explained. The process is however also a crucial background in $t\bar{t}H$ analyses and is an interesting process in its own right.

At leading order, only two diagrams contribute to the $t\bar{t}W$ process (one diagram for $t\bar{t}W^+$ and one for $t\bar{t}W^-$). In these diagrams, the process proceeds via annihilation of an up-type quark with a down-type antiquark or vice versa, where the W boson is radiated from one of the initial state quarks. These facts make the process looks deceptively simple. Several interesting phenomena however make this process both challenging and interesting.

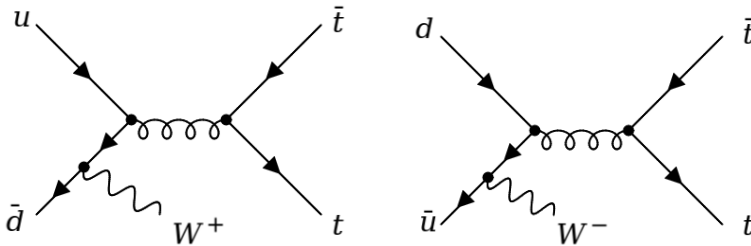


Figure 5.1.: The LO $t\bar{t}W$ production Feynman diagrams.

5.1.1. Theoretical difficulties

The simplicity at LO however also means that quark-gluon induced channels are only present at NLO and gluon-gluon induced channels only become available at NNLO (or at NLO@1). The significant presence of gluons in the proton PDFs at LHC energy scales however means that these channels usually dominate production processes at the energies of the LHC. This leads to significant higher-order and electroweak corrections and modifies significantly the predicted event kinematics. For instance, the NLO QCD term allows for quark-gluon induced channels. With a much larger gluon parton distribution function contribution, this term modifies the overall cross section with a k -factor of roughly 1.40, while one would expect an effect of around 10% from naive power counting of the coupling strengths [55].

As an even more extreme example, one can consider the sub-leading EWK NLO contributions. From naive power counting assumptions, one can expect modifications of the total cross section to be of the order 0.1%. But the t - W scattering diagrams, highlighted on the right side of figure 5.2, actually result in a k -factor of 1.12 [55]. These larger than expected contributions show the importance of including higher-order contributions in our theoretical calculations. Many of these higher-order contributions are however extremely difficult to compute. In order to compute the full NNLO QCD contributions to $t\bar{t}W$ production, the tree-level amplitudes and one-loop scattering amplitudes are already available but the two-loop virtual corrections are currently only known up to significant approximations.

A consistent tension has been observed between observed cross sections and the latest theory predictions. While improvements can still be made on the theory side, as we will see in the next section, the excess in data makes this process worthy of detailed investigation.

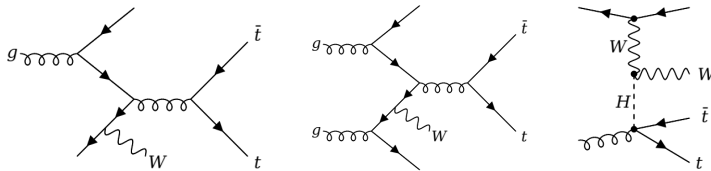


Figure 5.2.: Examples of NLO (left), NNLO (middle), and NLO EWK (right) Feynman diagrams contributing to the $t\bar{t}W$ process. All diagrams introduce sizeable corrections to the cross section of this process.

5.1.2. State-of-the-art

Currently the state-of-the-art calculations can only approximately calculate two-loop virtual contributions to the full NNLO QCD cross section of $t\bar{t}W$ (this picture has changed recently, see the outlook section). The first approximation calculates the amplitude in the ultra-relativistic limit with a massification method. On a simplistic level, with $Q \gg m_t$, we can get an amplitude with "massless" top quarks, which is known from NNLO W +light jets calculations. The other approximation assumes the radiation of a soft W boson. In this limit the amplitude can be factorised into the virtual amplitude of $t\bar{t}$ production and a soft W emission. As we do know that these corrections are significant, we can not yet trust these approximations for kinematic distributions at NNLO. Therefore, only the inclusive cross-section of $745.3 \text{ fb} \pm 6.7\%$ (scale) $\pm 1.8\%$ (approx.) is provided by the authors [82]. The two approximations at NNLO and their convergence to the true value for NLO calculations are shown in figure 5.3.

Apart from higher-order calculations, there have been numerous other improvements in the modelling of the $t\bar{t}W$ process. As the approximate NNLO predictions are not available at the level of MC events, the current state-of-the-art at this level is NLO calculations with up to two extra partons in the matrix element.

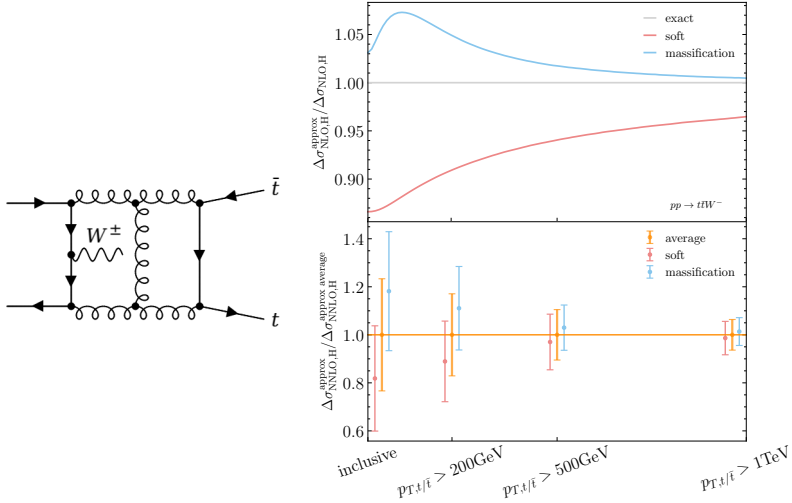


Figure 5.3.: Left: double-loop diagram, for which the amplitude is yet unknown. Right: the convergence of both approximations used to estimate the double-loop contribution. The top panel shows the convergence to the exact calculation at NLO QCD. The bottom panel shows the convergence at NNLO [82], where the exact calculation is of course not available.

While these calculations do not include higher order loop diagrams, they do include important gluon-gluon induced diagrams. Related to these state-of-the-art samples, an improved parton-jet merging scheme was developed by the original FxFx authors [83]. For this improved scheme, it was realised that the emission of a W from final-state extra partons protects these partons from infrared (IR) divergences. Two such diagrams are displayed in figure 5.4. Kinematically identifying these "electroweak" jets allows one to exclude them from the merging procedure and hence keep these jets from the matrix element even below the merging scale Q_{cut} . Having jets described by the matrix element below the merging

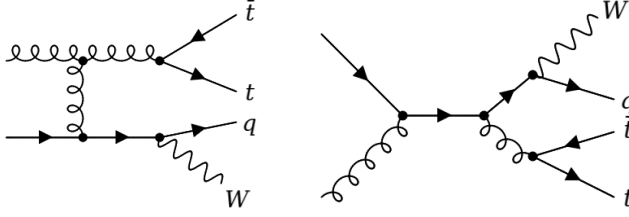


Figure 5.4.: Two real emission $t\bar{t}W$ diagrams where the extra parton is protected from IR divergencies.

scale could improve the modelling of these extra emissions [84] and lowers the merging scale dependence, as shown in Fig. 5.5.

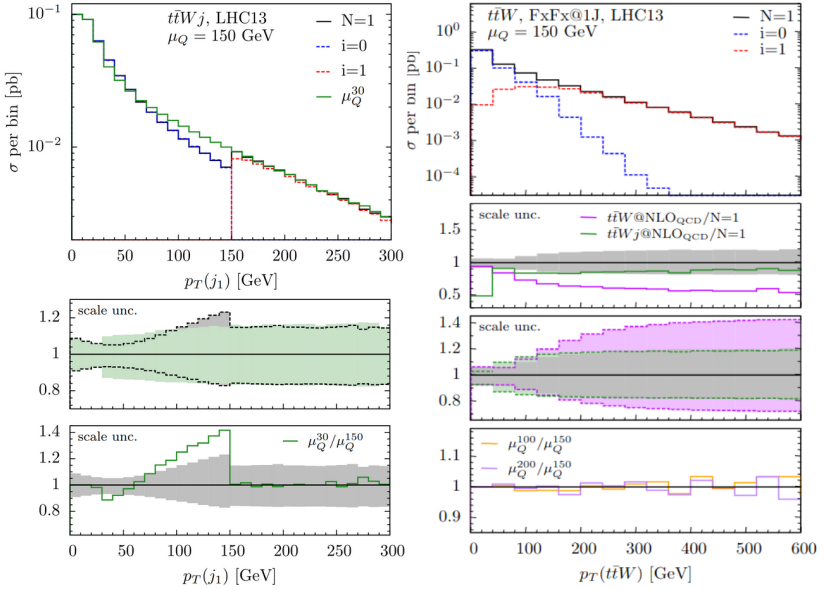


Figure 5.5.: Left: merging scale dependence for $t\bar{t}W$ production with the old FxFx@1j merging scheme [83]. Right: merging scale dependence for the new FxFx scheme [84].

5.1.3. Charge asymmetry

The presence of charge asymmetry in $t\bar{t}$ events was already explained in section 3.1.2. This effect is however strongly enhanced in $t\bar{t}W$ production as the process is quark induced at LO. The top quarks furthermore tend to inherit the polarization of the initial-state quarks due to the W boson emission, further enhancing the charge asymmetry [85]. This enhancement makes the process an enticing candidate to probe this interesting QCD phenomenon.

5.1.4. The analysis design

This study, performed with CMS data collected during RUN 2, is focused on these two interesting aspects of the $t\bar{t}W$ process. The analysis targets mainly the same-sign dilepton final state, where the W boson and one of the top quarks decay leptonically (to electrons or muons) and the other top quark hadronically. We also study the trilepton final state, but this channel suffers from a lower number of events and a worse signal purity. In the analysis we perform differential measurements as a function of a selection of interesting observables, collected from both theory-oriented and experiment-oriented sources.

This research project was carried out in close collaboration with the CMS group at Oviedo. Both teams pursued two conceptually different approaches for the differential measurement as a mutual cross-check (unfolding with and without a BDT discriminator), and the Oviedo group also performed a measurement of the charge asymmetry in $t\bar{t}W$. The focus in this dissertation will be only on the work I performed in collaboration with Dr. Luka Lambrecht: a colleague PhD student at Ghent university who has since graduated. Presented work I have not directly contributed to is always clearly highlighted.

5.2. Data and simulated samples

5.2.1. Data

The data used in this analysis consist of proton–proton collisions delivered by the LHC at a center-of-mass energy of 13 TeV and recorded by the CMS detector between 2016 and 2018. The full RUN 2 dataset is used, corresponding to an integrated luminosity of 138 fb^{-1} . Only data certified for physics analyses are considered, requiring stable beam conditions and the proper operation of all relevant detector subsystems.

To maximize the signal acceptance, multiple datasets collected with different high-level trigger (HLT) paths are combined. Any overlap between these datasets is removed by requiring unique run, luminosity block, and event identifiers.

The analysis is performed using the Ultra Legacy (UL) reprocessing of the RUN 2 dataset, which incorporates the most complete understanding of the detector performance during this period. This reprocessing includes improved detector calibrations, corrections for detector effects identified after data taking, and refined reconstruction algorithms. As a result, systematic biases are reduced and the corresponding experimental uncertainties are generally improved relative to earlier data reconstructions.

5.2.2. Simulated samples

We employ Monte Carlo (MC) simulated samples for all SM processes relevant to the phase space of this analysis. A complete list of the simulated processes is provided in Appendix B. Most samples are generated at next-to-leading-order (NLO) accuracy in QCD using MADGRAPH5_AMC@NLO or POWHEG. In cases where NLO samples are unavailable or suffer from limited statistics, leading-order (LO) samples generated with MADGRAPH5 are

used instead. This is only the case for processes that contribute very little to the phase space of this analysis.

The modeling of parton showering, hadronization, and the underlying event is performed with the PYTHIA event generator using the CP5 tune. In Monte Carlo event generators, a “tune” denotes a set of phenomenological parameters that cannot be predicted directly from first principles and must instead be constrained using experimental data. The CP5 tune was developed by CMS using a wide range of LHC measurements and provides an improved description of soft-QCD effects, multiple parton interactions, colour reconnection, and parton-shower evolution. To avoid double counting of partons generated at the matrix-element and parton-shower levels, the FxFx scheme for NLO samples and the MLM scheme for LO samples are applied, as explained in previous chapters. All simulated events are processed through a detailed GEANT4-based simulation of the CMS detector and subsequently reconstructed using the same algorithms as applied to collision data. The effect of additional proton–proton interactions (pileup) is included by overlaying simulated minimum-bias events generated with NeutrinoGun. The resulting pileup distribution is then reweighted to match that observed in data. This reweighting is done by tuning the inelastic cross section parameter to optimize the data - MC agreement for some pileup related observables.

The signal process is simulated using MADGRAPH5_AMC@NLO. One sample is generated at NLO in QCD, including up to two additional parton emissions. A second, dedicated sample includes electroweak contributions of order α_{EW}^3 and mixed QCD–electroweak terms of order $\alpha_s \alpha_{\text{EW}}^3$. An overview of the full set of NLO components and their corresponding corrections is shown in Table 5.1.

Since the contribution at order $\alpha_s \alpha_{\text{EW}}^2$ is zero, the LO₃ and NLO₃ terms can be consistently matched to the parton shower as an

independent sample. Both signal samples are generated with the W boson decaying to leptons.

	$\begin{array}{c} \text{LO}_1 \\ \swarrow \quad \searrow \\ \text{NLO}_1 \quad \text{NLO}_2 \end{array}$		$\begin{array}{c} \text{LO}_2 = 0 \\ \swarrow \quad \searrow \\ \text{NLO}_3 \quad \text{NLO}_4 \end{array}$	
Naive expectation	$\mathcal{O}(10\%)$	$\mathcal{O}(1\%)$	$\mathcal{O}(0.1\%)$	$\mathcal{O}(0.01\%)$
Actual correction	30–60%	–(4.0–4.4)%	12–13%	0.02–0.05%

Table 5.1.: Structure of LO and NLO contributions with naive expectations and actual corrections. Green order terms are included in the signal sample, red order terms are excluded.

Next to these samples, another signal sample was generated following the improved FxFx scheme discussed in section 5.1.2. This sample was generated without detector simulations and is only used to compare our results to these state-of-the-art theory predictions. All the samples used in this analysis were not produced by the author.

The reference cross-section of this process is chosen to be 745.3 fb, based on the approximate NNLO calculation described in section 5.1.2 and as agreed with the ATLAS collaboration. The QCD-only cross-section is 710.7, following the same reference. The cross sections for the two samples are then:

$$QCD : \sigma_{t\bar{t}W}^{\text{QCD}} Br(W \rightarrow l\nu) = 231.6 \text{ fb} \quad (5.1)$$

$$EW : (\sigma_{t\bar{t}W}^{\text{total}} - \sigma_{t\bar{t}W}^{\text{QCD}}) Br(W \rightarrow l\nu) = 11.3 \text{ fb} \quad (5.2)$$

5.3. Object selection

The reconstruction of physics objects using the PF algorithm was described in Section 4.5. While this algorithm provides a detailed and efficient reconstruction of individual particles, additional quality criteria must be applied to ensure the reliability of the reconstructed objects. These object selection requirements are discussed in this section. No special objects are introduced in this analysis, all reconstructed objects and their selection criteria follow standard definitions commonly used in CMS analyses.

5.3.1. Jets and Missing transverse energy

Jets are required to pass the requirements summarized in table 5.2. Jets have to fall inside of tracker and calorimeter acceptance and have a reasonable transverse momentum. Jet objects are removed from leptons within a ΔR radius of 0.4 to orthogonalise the interpretation of these objects. Other selections are summarized in the "Jet ID". The tight jet ID imposes conditions on the jet substructure (e.g. number of constituents and relative amount of charged and neutral components), aimed at removing anomalous energy deposits.

To identify jets originating from B-hadrons, we make use of the DEEPJET algorithm. The architecture of this model is based on LSTM nodes. In this analysis, jets are tagged as b-jets using the loose working point. This working point has an efficiency of about 95% (for jets with $p_T > 30$ GeV). The light jet misidentification rate (i.e. the probability that a light jet is identified as a b-jet) is about 10%, and the c jet misidentification rate is about 50%. Scale factors to correct for known differences in b-tagging efficiency between simulation and data are taken from central measurements by the b-tagging Physics Object Group (POG), following their recommendations. In this analysis, we consider jets with p_T greater than 25 GeV.

Table 5.2.: Summary of jet identification criteria used in this analysis. Values in color are the requirements in the Jet ID.

	Requirement
p_T	$> 25 \text{ GeV}$
$ \eta $	< 2.4
ΔR closest lepton	> 0.4
Neutral Hadron Fraction	< 0.9
Neutral EM Fraction	< 0.9
Number of Constituents	> 1
Muon Fraction	< 0.8
Charged Hadron Fraction	> 0
Charged Multiplicity	> 0
Charged EM Fraction	< 0.8

The missing transverse momentum is constructed as the negative vector sum of all "jet" objects reconstructed by the jet finding algorithm.

5.3.2. Leptons

Next to b-tagging performance, lepton identification performance is vital to this analysis. In this analysis, we make use of an MVA based approach with a gradient Boosted Decision Tree (gBDT) model, referred to as the leptonMVA. Firstly, baseline selections are applied to both muons and electrons that make sure the objects are within the limits of what can be accurately reconstructed by the CMS detector and that they are within the phase space of the leptonMVA training. The ‘fakeable object’ (FO) ID is used to estimate the non-prompt background in a data-driven way. This method is explained in detail in section 5.5.2.

The definitions of these selections are described in detail in the coming sections and are shown in table 5.3 for muons and table 5.4

for electrons.

Muons

For muons we also have p_T and $|\eta|$ requirements to fall within the acceptance of the muon system. Additional loose cuts are applied on impact parameter, relative isolation and a POG-defined ID. These are exactly the same as the cuts used to delimit the phase space for the leptonMVA training and testing. Relative isolation is defined as the scalar p_T sum of physics objects in a cone around the lepton, divided by the p_T of the lepton itself. The physics objects include charged hadrons, neutral hadrons and photons, and a correction for pileup contributions. The cone size is variable, scaling with $1/p_T$ of the lepton:

$$p_T(l) < 50 \text{ GeV} : \Delta R = 0.2 \quad (5.3)$$

$$50 \text{ GeV} < p_T(l) < 200 \text{ GeV} : \Delta R = \frac{10 \text{ GeV}}{p_T(l)} \quad (5.4)$$

$$p_T(l) > 200 \text{ GeV} : \Delta R = 0.05 \quad (5.5)$$

Thanks to the isolation requirement, many leptons originating from hadron decays that fell outside of the cone of the corresponding jet can still be removed. The tight muons are then required to pass the leptonMVA WP with a prompt muon efficiency of 93% and a nonprompt mistagging rate of 3%. A few other variables are introduced for the FO and tight selections, which will be discussed later in the fakerate method section: section 5.5.2.

Electrons Electrons have very similar quality requirements as muons. There are however two more fake electron sources that can be reduced with dedicated physics object selections. The first source of "fake" electrons are electrons coming from photon conversions (i.e. real photons converting to an e^-e^+ pair in the detector material), which can be reduced with a dedicated "conversion veto". The second source comes from electrons with a misidentified charge. These are reduced by requiring that three

Table 5.3.: Summary of muon identification criteria used in this analysis. Values in parentheses indicate thresholds applied when the leptonMVA requirement is not satisfied.

	Loose	Fakeable	Tight
p_T	$> 10 \text{ GeV}$	$> 10 \text{ GeV}$	$> 10 \text{ GeV}$
$ \eta $	< 2.4	< 2.4	< 2.4
$ d_{xy} $	$< 0.05 \text{ cm}$	$< 0.05 \text{ cm}$	$< 0.05 \text{ cm}$
$ d_z $	$< 0.1 \text{ cm}$	$< 0.1 \text{ cm}$	$< 0.1 \text{ cm}$
d/σ_d	< 8	< 8	< 8
Rel. isolation	< 0.4	< 0.4	< 0.4
PF muon ID	$\geq \text{M WP}$	$\geq \text{M WP}$	$\geq \text{M WP}$
DeepJet nearest jet	–	(< 0.025)	–
Jet p_T ratio	–	(> 0.45)	–
LeptonMVA	–	> 0.64 (< 0.64)	> 0.64

independent estimations of their charge agree. These two additional selections were observed to significantly increase the signal to background ratio. The charge misidentified electrons will however be described in more detail in section 5.5.1. Tight electrons are then required to pass the leptonMVA working point with a prompt efficiency of 89% for electrons (with a nonprompt mistagging rate of 2%). A few other variables are introduced for the FO selection, which will be discussed later in the fakerate method section: section 5.5.2. Electrons that fall in the region between the barrel and endcap of the ECAL ($1.44 < |\eta| < 1.57$) are not considered in this analysis as their reconstruction performance is suboptimal.

5.4. Event selection

The leptonMVA working points described in the last chapter are quite loose with a large signal efficiency and correspondingly

Table 5.4.: Summary of electron identification criteria used in this analysis. Values in parentheses indicate thresholds applied when the leptonMVA requirement is not satisfied. Values separated by slashes correspond to 2016PreVFP / 2016PostVFP / 2017 / 2018.

	Loose	Fakeable	Tight
p_T	> 10 GeV	> 10 GeV	> 10 GeV
$ \eta $	< 2.5	< 2.5	< 2.5
$ d_{xy} $	< 0.05 cm	< 0.05 cm	< 0.05 cm
$ d_z $	< 0.1 cm	< 0.1 cm	< 0.1 cm
d/σ_d	< 8	< 8	< 8
Rel. isolation	< 0.4	< 0.4	< 0.4
Missing hits	< 2	< 2	< 2
Conversion rejected	–	✓	✓
3 Charge measurements	–	agree	agree
EGamma MVA	–	($> \text{WP-loose}$)	–
DeepJet nearest jet	–	(< 0.1)	–
Jet p_T ratio	–	($> 0.5/0.5/0/4/0.4$)	–
Lepton MVA	–	> 0.81 (< 0.81)	> 0.81

sizable nonprompt contribution passing the leptonMVA. The larger event yield in the nonprompt contribution is however taken care of by the event-level MVA, which benefits from a larger training set in this strategy. Moreover, the nonprompt prediction can more accurately be checked and constrained, both in the signal region (at low event-level MVA scores) and in dedicated control regions. This strategy gives the analysis more statistical power than an analysis requiring a tight lepton working point.

5.4.1. Trigger

The baseline lepton phase space relevant for this analysis consists of events with two same-sign leptons, with transverse momenta

exceeding 25 GeV (15 GeV) for the (sub)leading lepton. In addition, a trilepton signal region (of secondary importance) and control regions with three or four leptons are defined. A common trigger strategy is employed across all these regions, based on a combination of single-lepton, dilepton, and trilepton triggers, listed in Tables C.1–C.3 (Appendix C).

The trigger efficiency is estimated using the orthogonal trigger method. In this approach, events are selected from data sets triggered independently of leptons (e.g. MET- or H_T -based triggers), and a lepton selection similar to that used in the analysis is applied. The fraction of these events that also satisfy the lepton triggers provides an unbiased estimate of the trigger efficiency. This procedure was not performed by the author.

5.4.2. Signal regions

To get a pure $t\bar{t}W$ sample, we focus on two signal regions: the 2 lepton (2ℓ) same-sign (SS) signal region and the trilepton (3ℓ) signal region. Most sensitivity can be reached in the 2ℓ SS SR. A high number of (b-)jets allows us to exclude many non-top-related background processes. Furthermore requiring two SS leptons excludes a large amount of other top-related processes such as top pair production and top pair production in association with a Z boson or γ . Events with a lepton pair whose invariant mass is below 30 GeV are rejected to reduce contributions from low-mass resonances. The detailed selections are presented in table 5.5.

Events passing the 2ℓ SS selections are shown in Figure 5.6, while the 3ℓ region is shown in Figure 5.7. The 2ℓ SS signal region features a relatively high signal selection efficiency, but remains dominated by several background contributions. The largest background arises from events with nonprompt leptons, predominantly originating from $t\bar{t}$ production. In such events, one top quark typically decays leptonically, while the second lepton is non-prompt, i.e. originates from a hadron decay. After applying the electron

Table 5.5.: Comparison of dilepton and trilepton signal region selections.

Selection	Dilepton SR	Trilepton SR
# leptons	2	3
Lepton charge	Same-sign	OS pair
p_T thresholds	25, 15 GeV	25, 15, 15 GeV
Inv. mass cut	$m(\ell\ell) > 30$ GeV	$m(\ell\ell) > 30$ GeV
$ m(ii) - m(Z) > 10$ GeV	ee	$\ell\ell_{\text{OSSF}}$
E_T^{miss}	> 30 GeV	—
# jets	≥ 3	≥ 2
# b-jets	≥ 2 (loose WP)	≥ 2 (loose WP)

charge consistency requirement, a non-negligible contribution from charge-misidentified electrons remains, again primarily from $t\bar{t}$ events. Additional contributions arise from photon conversions. These backgrounds are often ill simulated and are the topic of the following section. These backgrounds are referred to as *reducible*, as they originate from misidentified or nonprompt objects and can therefore be suppressed through improved object reconstruction, identification, and selection criteria. Each of these background types will be discussed in the next section.

In addition to these backgrounds, there exists an *irreducible* background component. These processes produce a very similar final-state signature as the signal, with all reconstructed objects being genuine prompt particles. As a result, they cannot be removed through tighter object selections and must be constrained through simulation and dedicated control regions. The dominant contributions originate from top quark pair production in association with bosons, in particular $t\bar{t}H$ and $t\bar{t}Z$. Further contributions arise from diboson processes such as WZ and ZZ , multiboson production, and single top production in association with bosons, such as tqZ and tWZ . These irreducible backgrounds are estimated using simulated samples, as listed in appendix B.

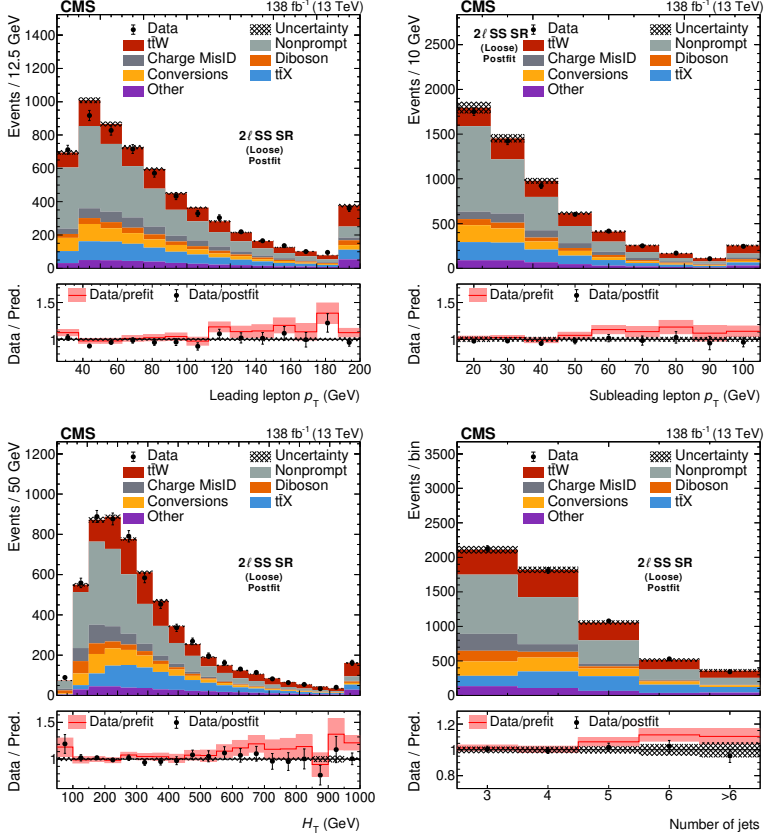


Figure 5.6.: Leading lepton p_T (upper left), subleading lepton p_T (upper right), scalar p_T sum of selected jets in the event (lower left), and the number of selected jets in the event (lower right), in the two-lepton signal selection for data (points) and predictions (filled histograms) after the fit to the data in the SR and CR. The lower panels show the ratio of the data to the sum of the postfit predictions (points) and the ratio of the data to the prefit predictions (red line). The hatched (filled) band corresponds to the total uncertainty in the postfit (prefit) prediction. Events that exceed the range of the plot are included in the last bin.

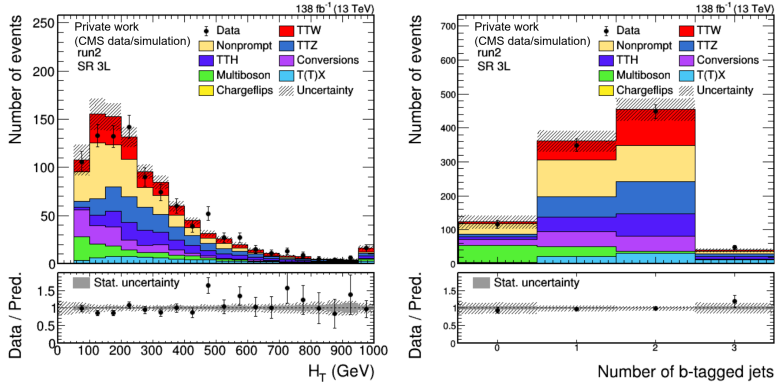


Figure 5.7.: Scalar p_T sum of selected jets in the event (left), number of selected b-jets (right), in the trilepton signal selection for data (points) and predictions (filled histograms) after the fit to the data in the SR and CR. The lower panels show the ratio of the data to the sum of the postfit predictions (points) and the ratio of the data to the prefit predictions (red line). The hatched (filled) band corresponds to the total uncertainty in the postfit (prefit) prediction. Events that exceed the range of the plot are included in the last bin.

5.5. Background estimation and control regions

5.5.1. Charge misidentified electrons

In CMS, the electric charge of a reconstructed particle is primarily determined from the curvature of its track in the magnetic field. For electrons, additional information can be obtained from the position of the associated ECAL supercluster relative to the track. Despite requiring consistency between these charge determination methods, a small but non-negligible probability of charge misiden-

tification remains.

This effect is predominantly caused by bremsstrahlung radiation. Electrons can emit energetic photons while traversing the tracker material, which may subsequently convert into electron–positron pairs. The resulting tracks can complicate the reconstruction and lead to an incorrect charge assignment. This mechanism explains why charge misidentification is significantly more pronounced for electrons than for muons. This also shows that the misidentification rate is strongly dependent on the material budget that has to be traversed in the tracker.

At high transverse momentum, the curvature of tracks becomes very small, making the charge determination more susceptible to detector effects such as noise or missing hits. In such cases, the reconstructed track may be assigned the wrong curvature and hence the wrong charge. Both effects are illustrated in Figure 5.8.

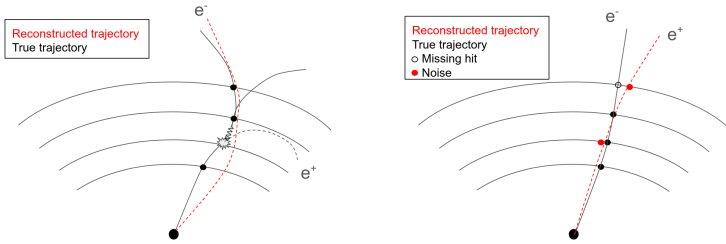


Figure 5.8.: Illustration on how the charge of an electron can be misreconstructed in the CMS tracker.

Charge misidentification is particularly relevant in same-sign dilepton signatures. Processes such as DY and $t\bar{t}$, which have large production cross sections, predominantly produce opposite-sign dilepton pairs. Even though the probability for charge misidentification is small, the large event yields of these processes can result in a non-negligible contribution to the same-sign signal region. In this way, events originating from opposite-sign dilepton

production can migrate into a phase space where the expected rates of genuine same-sign processes are substantially lower. As the charge misidentification rate is not reliably modeled in simulation, a semi-data-driven method is employed to estimate this contribution. The central assumption of this approach is that the charge misidentification probability is largely independent of the event topology and can be parameterized as a function of the lepton p_T and pseudorapidity η .

First, the charge misidentification rate is measured in simulated DY and $t\bar{t}$ samples. Electrons are required to pass the tight selection criteria and to be prompt, based on generator-level information, in order to avoid overlap with the nonprompt background estimate described in Section 5.5.2. The misidentification rate is defined as the fraction of electrons for which the reconstructed charge differs from the true charge over the total number of electrons. These rates are stored in two-dimensional maps as a function of p_T and η of the electron.

Due to the intrinsically low probability of charge misidentification, the statistical precision of this measurement is limited. To mitigate this effect, a binning scheme is chosen that is sufficiently coarse to ensure adequate statistical precision in each bin, while still capturing the dependence on p_T and $|\eta|$.

The charge misidentification rate may vary over time and is therefore measured separately for each data-taking period: 2016PreVFP, 2016PostVFP, 2017, and 2018 (and even finer during the development of this method). The resulting maps are shown in Figure 5.9. They are derived using both DY and $t\bar{t}$ samples, which are found to yield statistically compatible results. The final maps are obtained by combining both processes to maximize statistical precision. The measured charge misidentification rates are small, typically ranging from $\mathcal{O}(10^{-4})$ to $\mathcal{O}(10^{-2})$. The charge

misidentification rates for muons were estimated in the same way and were found to be negligible (maximally 10^{-5}).

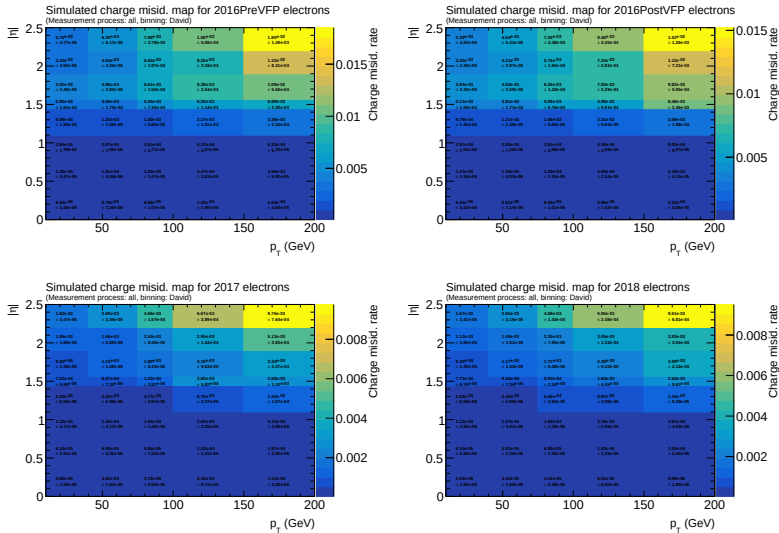


Figure 5.9.: Chargeflip maps, as measured in MC events for the data-taking years used in the $t\bar{t}W$ analysis.

5.5.1.1. Closure test in MC

The charge misidentification maps are validated through a closure test using simulated DY and $t\bar{t}$ samples. This test serves as a self-consistency check of the method and of the chosen binning. The following event selection is applied:

1. Exactly two tight, prompt electrons (25 GeV, 15 GeV)
2. $|m(ee) - m(Z)| < 15$ GeV (for DY events only).

An event is considered “observed” if at least one of the electrons has a misidentified charge, as determined using generator-level information. The corresponding “prediction” is obtained by assigning a weight to each event with opposite sign electrons based

on the per-electron charge misidentification probabilities. Neglecting the small probability that both electrons are simultaneously misidentified, the event weight is given by

$$w_{\text{event}} = w(e_1) + w(e_2) = \frac{f(e_1)}{1 - f(e_1)} + \frac{f(e_2)}{1 - f(e_2)}. \quad (5.6)$$

Selected distributions from this closure test are shown in Figure 5.10. Good agreement between the observed and predicted yields as a function of lepton kinematic variables demonstrates that the charge misidentification maps accurately reproduce the relevant kinematic dependencies. There are however two biases observed in this parametrisation.

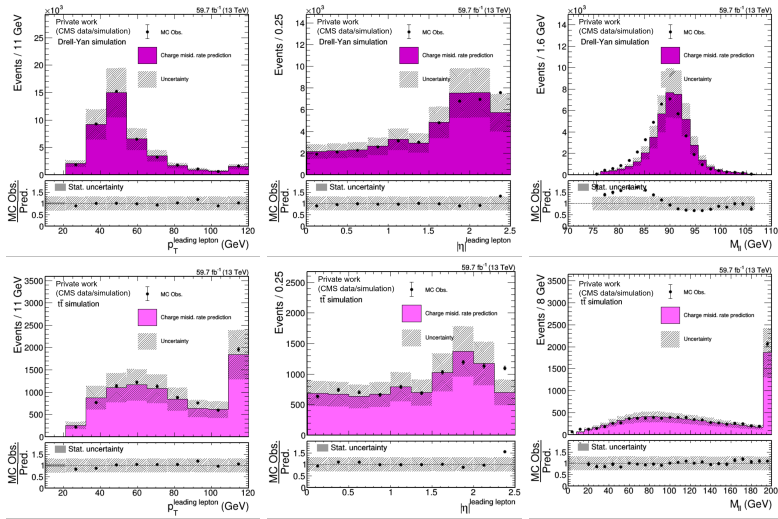


Figure 5.10.: Some distributions observed in the closure tests in 2018 simulation samples for the charge misidentification method. Upper row: DY, lower row: $t\bar{t}$.

The first systematic deviation is observed around the Z boson mass peak in the “observed” distributions. This effect arises because charge-misidentified electrons often undergo significant hard bremsstrahlung, leading to energy loss via the photons that are

not (fully) captured in the electron supercluster and a distortion of the reconstructed invariant mass. Since the charge misidentification method does not account for this energy loss, the effect is not reproduced in the “predicted” distributions.

As charge misidentified leptons in our analysis phase space are predominantly originating from $t\bar{t}$ decays, they do not exhibit a resonant structure and are hence not affected by this bias of the method.

The second systematic deviation is observed in the highest- η bin. The effect exhibits a clear year dependence and is most pronounced in the 2018 dataset. Although the exact origin of this discrepancy is not fully understood, studies indicate that it arises from the simplified parameterization of the fake rate as a function of only p_T and η . In particular, correlations between the kinematic properties of the two leptons are neglected, leading to residual biases that become most apparent in the forward region.

As the effect is considerably smaller in the other data-taking years, the granularity of the differential measurement is much coarser than the observed bias, and a sizeable systematic uncertainty is assigned to the method, these residual discrepancies have a negligible impact on the final results of the analysis.

5.5.1.2. Calibration of charge MisID rates in data

Following this test, we perform a similar closure test in data around the Z peak to extract year-dependent scale factors. The same selection is performed at the reconstructed level, of course excluding the prompt requirement as this is based on generator-level information. The small amount of SS events with correctly charge-identified leptons (both prompt and non-prompt) are taken from MC. The charge-misidentified component is selected from OS events following the same procedure as in equation 5.6.

The results are shown in Figure 5.11 and the scale factors are shown in table 5.6. An underprediction of the order of 40-50% is observed in both 2017 and 2018, while the agreement in 2016 is better (with a slight overprediction in 2016PreVFP however). These differences and discrepancies have also been observed in previous analyses [86, 60] and are caused by the strong dependence of the charge misidentification on the modeling of the material budget in the tracker. The substantial difference between 2016 and 2017/2018 is most likely caused by the upgrade of the pixel tracker in between those two data-taking periods.

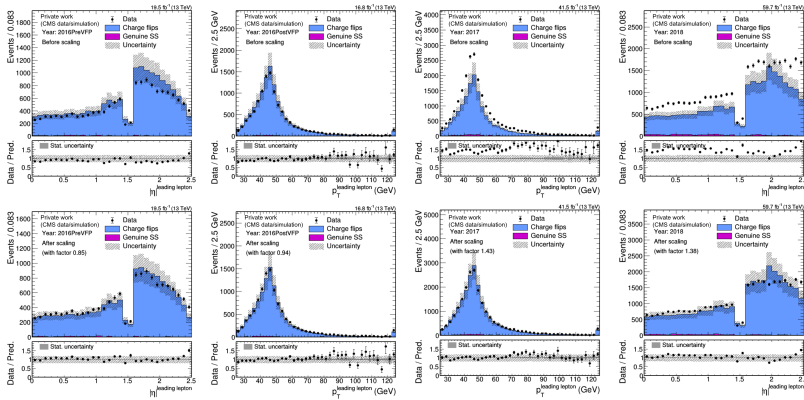


Figure 5.11.: Some distributions observed in the closure tests in data for the charge misidentification method for all data-taking years. Closure is shown before (after) applying the year dependent scale factors in the row above (below).

Table 5.6.: Year dependent scale factors for chargemisID rate.

Year	2016PreVFP	2016PostVFP	2017	2018
Scalefactor	0.86	0.94	1.42	1.39

5.5.2. Nonprompt leptons

We have already introduced the selection requirements for muons and electrons used in this analysis. A key component of this selection is the leptonMVA discriminator. The lepton MVA is a gradient boosted decision tree (gBDT) trained to distinguish prompt leptons from nonprompt leptons. The training sample consists of prompt leptons from simulated $t\bar{t}Z$, $t\bar{t}W$, tqZ , $t\bar{t}t\bar{t}$, and $t\bar{t}H$ events, while nonprompt leptons are taken from simulated $t\bar{t}$ samples.

The hyperparameters of the model are optimized using cross-validation. The performance of the leptonMVA has been calibrated in data, resulting in scale factors and associated uncertainties that have been validated and approved by the Muon and EGamma POGs [87]. This calibration work was carried out within the Ghent University CMS group, but was not performed by the author. The performance of the leptonMVA is illustrated by the ROC curves shown in Figure 5.12.

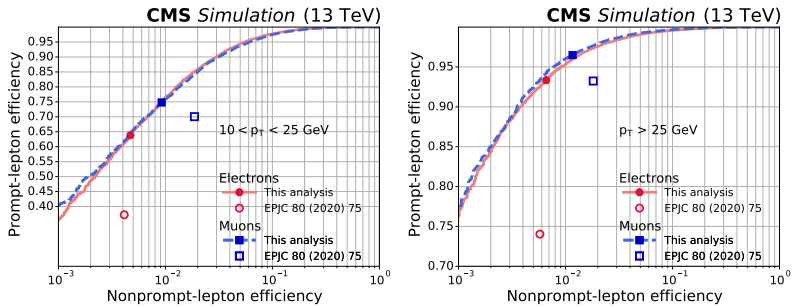


Figure 5.12.: ROC curves showing the performance of the leptonMVA in the p_T ranges $10 \text{ GeV} < p_T < 25 \text{ GeV}$ (left) and $p_T > 25 \text{ GeV}$ (right). Figure from ref. [60].

The dominant source of nonprompt leptons in this analysis originates from semileptonic decays of b hadrons produced in $t\bar{t}$ events.

Smaller contributions arise from charm-hadron decays, electrons from photon conversions and Dalitz decays of neutral mesons, and leptons produced in charged pion and kaon decays. The model exploits several observables to discriminate between prompt and nonprompt leptons. Some of the most crucial observables are listed below:

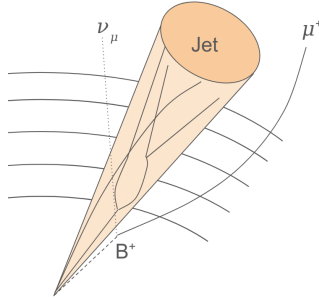


Figure 5.13.: Example sketch of a nonprompt lepton.

1. **Isolation:** Nonprompt leptons, originating from hadron decays, are typically embedded in hadronic activity. The isolation variable, defined as the scalar sum of transverse momenta of particles within a cone around the lepton, quantifies how isolated the lepton is from surrounding activity.
2. **Impact parameter:** Leptons from hadron decays can have a measurable displacement with respect to the primary vertex. A large impact parameter therefore indicates a higher probability for the lepton to be nonprompt.
3. **Closest jet observables:** Nonprompt leptons are often associated with nearby jets. Variables such as p_T^{rel} (the transverse momentum of the lepton relative to the direction of the closest jet) and p_T^{ratio} (the ratio of the lepton p_T to that of the closest jet) provide additional discrimination power.

4. **Lepton kinematics and identification variables:** Additional inputs include basic kinematic properties and quality criteria from the standard lepton reconstruction, which further enhance the separation power of the classifier.

5.5.2.1. Background estimation

To estimate the number of background events with nonprompt leptons that still pass the leptonMVA requirement, we rely on the fakerate method. In this method, we want to measure the probability, called the fake rate, that a nonprompt lepton passing the FO requirement also passes the Tight requirement. The fake rate is hence defined as:

$$f = P(\text{lepton} = \text{Tight} | (\text{lepton} = \text{FO}) \ \& \ (\text{lepton} = \text{nonprompt})) \quad (5.7)$$

The method then follows a similar approach to the chargeflip method, where a sideband region is now constructed using the fakeable object (FO) definition, which is made exclusive from the tight selection requirement. This results in an orthogonal application region that is rich in nonprompt leptons. Just like we used the OS sideband to predict the SS chargeflip background, we will use this FO sideband to predict the nonprompt background in our analysis regions. The correct weighting factor for an N-lepton region is:

$$w = -1 \prod_{i=1}^N \frac{-f}{1-f}. \quad (5.8)$$

As the fake rate was not measured by the author, I refer to the thesis of Luka Lambrecht [88] for more details on this subject.

5.5.3. Control regions and validation regions

Several control regions (CRs) and validation regions (VRs) are developed to validate our description of crucial backgrounds in this

analysis. CRs will be included in the fit discussed in section 5.8, while VRs are only used to validate the background predictions but are not included in the fit. An overview of all CRs and the already discussed SRs is shown in Figure 5.14.

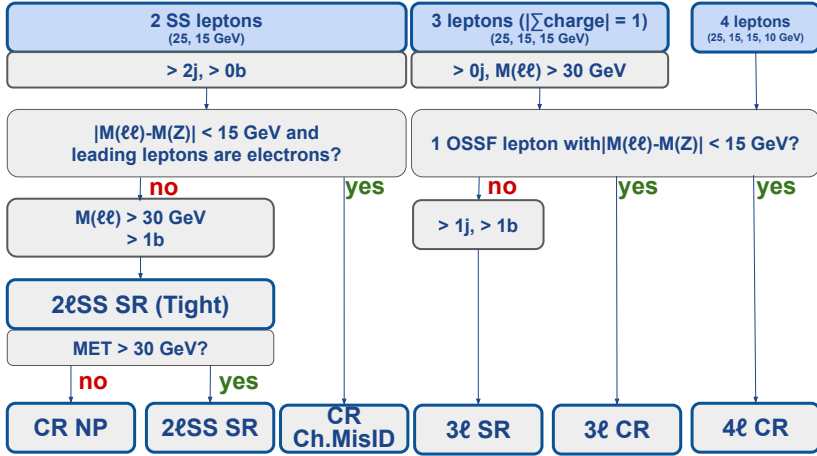


Figure 5.14.: Schematic representation of the event selection and categorization.

5.5.3.1. Reducible backgrounds

Charge misidentified electrons:

The chargeflip control region makes use of the large cross section of the DY process. As this process is so abundant around the Z pole, we expect to see the SS electrons resulting from charginemisidentification dominate any true SS process in this window. The validation region requires exactly two tight SS leptons with $|m(\ell\ell) - m(Z)| < 15$ GeV and p_T 's of 25 and 15 GeV. A tighter CR was developed which also requires more than 2 jets and at least 1 b-jet, bringing the region closer to the phase space of this analysis. Some distributions are shown in Figure 5.15.

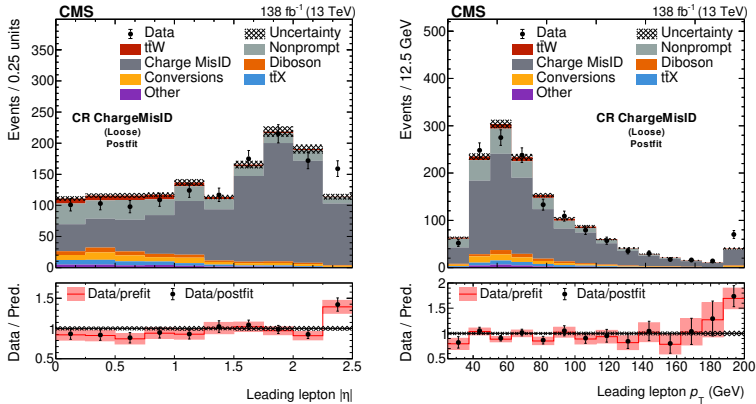


Figure 5.15.: $|\eta|$ of the leading lepton (left), p_T of the leading lepton (right), in a CR enriched with charge-misidentified leptons and additional jets after the fit to the data in the SR and CR.

Nonprompt leptons:

A good understanding of nonprompt leptons and a good separation of our signal from this background is crucial to this analysis. To this end, multiple control and validation regions were developed for this background.

The first control region is a subsection of the 2ℓ SS SR definition employed in the Oviedo analysis (modulo the other leptonMVA). As our analysis strategy uses a loose working point, we have a much larger acceptance in this region and can have a dedicated nonprompt CR (NP CR) with $p_T^{\text{miss}} < 30$ GeV. Some distributions are shown in Figure 5.16.

In addition to this control region, a second region is defined in which, instead of applying a p_T^{miss} requirement, the jet and b -jet multiplicities are used to ensure orthogonality with respect to the signal region. Events are selected into this nonprompt control region if they satisfy the same requirements as the dilepton signal

5. Top quark pair production in association with a W boson

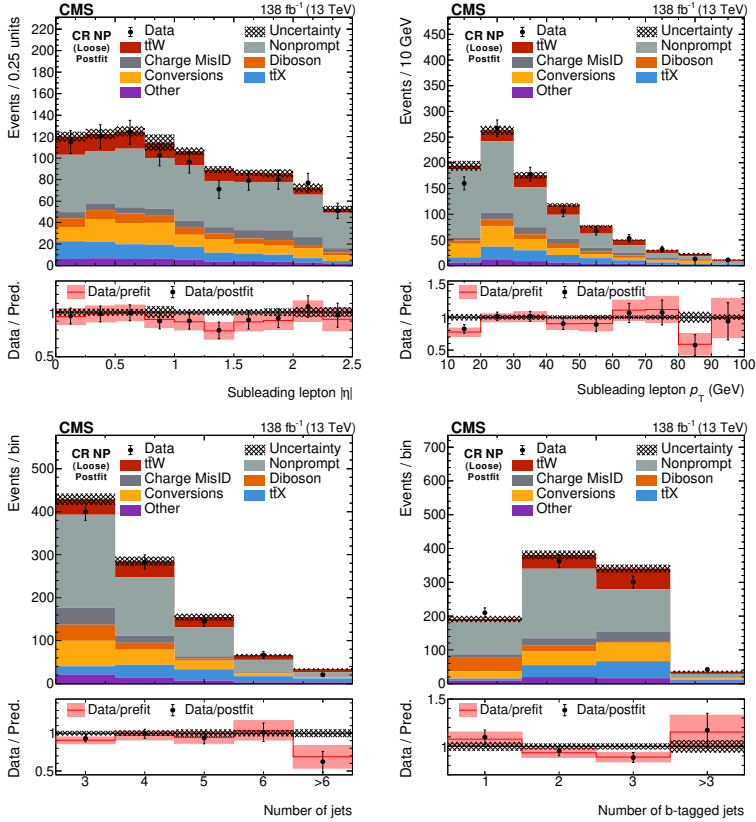


Figure 5.16.: $|\eta|$ (top left) and p_T (top right) of the subleading lepton, number of jets (bottom left) and number of b -jets (bottom right) in a CR enriched with nonprompt leptons after the fit to the data in the SR and CR.

region, with the exception that the number of jets is restricted to one or two (instead of at least three), and at least one b -tagged jet is required. Unlike the NP CR, this region is not included in the fit. Some distributions are shown in Figure 5.17.

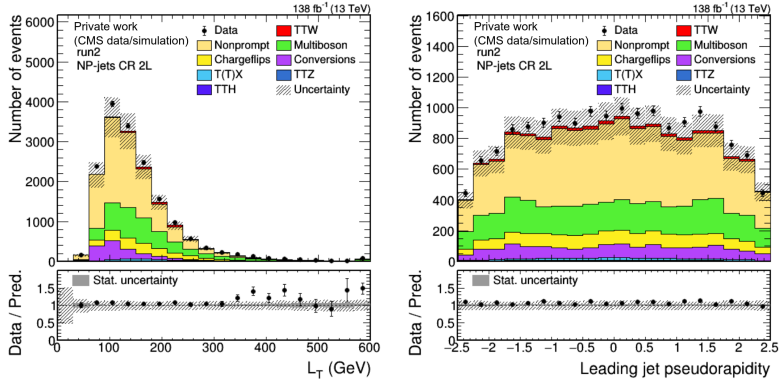


Figure 5.17.: Scalar sum of lepton p_T 's (left) and leading jet p_T (right) in a VR enriched with nonprompt leptons using jet requirements for orthogonality.

Photon conversions:

Photon conversions occur when a photon interacts with the detector material and converts into an electron–positron pair, as illustrated in figure 5.18. The simulation of photon conversions is validated by selecting collision events where three leptons have an invariant mass consistent with a Z boson, with a OSSF lepton pair in the combination. This selection targets DY events where

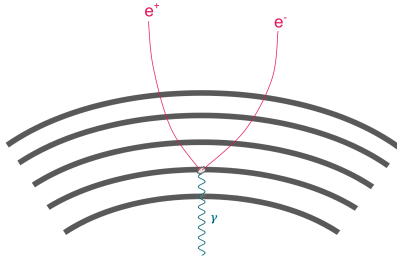


Figure 5.18.: Example sketch of a photon conversion in the tracker.

one lepton undergoes a hard bremsstrahlung, which in turn forms an OSSF lepton pair. The detailed selections are:

1. Exactly three tight leptons (25 GeV, 20 GeV, 10 GeV)
2. one OSSF lepton pair
3. $|m(\ell\ell) - m(Z)| < 15$ GeV and $|m(\ell\ell)_{\text{OSSF}} - m(Z)| > 15$ GeV
4. $m(\ell\ell) > 12$ GeV (for all combinations)

Some distributions are shown in Figure 5.19. Good agreement between data and MC is observed, showing that the photon conversions are well modeled by MC.

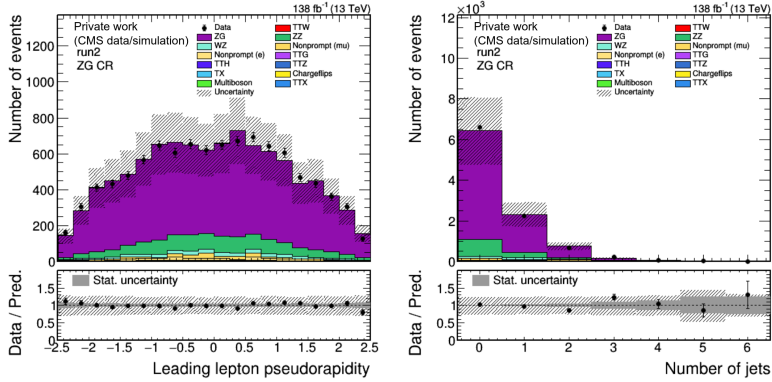


Figure 5.19.: η of the leading lepton (left) and number of jets (right) in a VR enriched with photon conversion events.

5.5.3.2. Irreducible backgrounds

Apart from the backgrounds discussed in the last section, which are more related to the modeling of the detector response, there are other backgrounds that enter the analysis phase space, without nonprompt or charge misidentified leptons. The control regions that focus on these processes are the focus of this section.

Trilepton control region:

The trilepton control region is used to infer the modeling of mainly the WZ and $t\bar{t}Z$ backgrounds. The event selection details are shown below:

1. Exactly three tight, prompt electrons (25 GeV, 15 GeV, 15 GeV)
2. one OSSF lepton pair with $|m(\ell\ell)_{\text{OSSF}} - m(Z)| < 15 \text{ GeV}$
3. $m(\ell\ell) > 30 \text{ GeV}$ (for all combinations)
4. > 0 jets

We show the CR as a function of number of jets and number of b-jets. At low (b-)jet multiplicities we get more sensitivity for the WZ process, while $t\bar{t}Z$ is dominant at higher multiplicities. This distribution, shown in Figure 5.20, is included in the fit.

Fourlepton control region:

The fourlepton control region is constructed in a similar way and is used to infer the modeling of mainly the ZZ and $t\bar{t}Z$ backgrounds. The selections are as follows:

1. Exactly four tight, prompt electrons (25 GeV, 15 GeV, 15 GeV and 10 GeV)
2. at least one OSSF lepton pair with $|m(\ell\ell)_{\text{OSSF}} - m(Z)| < 15 \text{ GeV}$

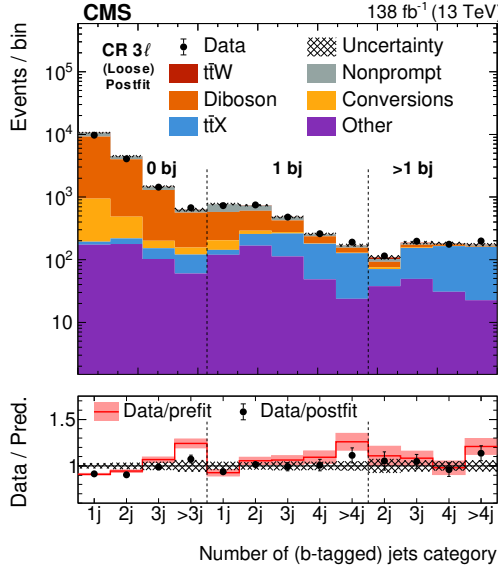


Figure 5.20.: Number of selected (b-tagged) jets for events passing the triplepton CR requirements after the fit to the data in the SR and CR.

This CR is displayed as a function of the number of Z candidates (1 or 2) and number of b-tagged jets. The split in number of Z candidates gives strong separating power between the ZZ and $t\bar{t}Z$ processes. As the statistical power of this selection is much lower than in the triplepton CR, $t\bar{t}Z$ uncertainties are mostly constrained in the triplepton CR. This distribution, shown in Figure 5.21, is included in the fit.

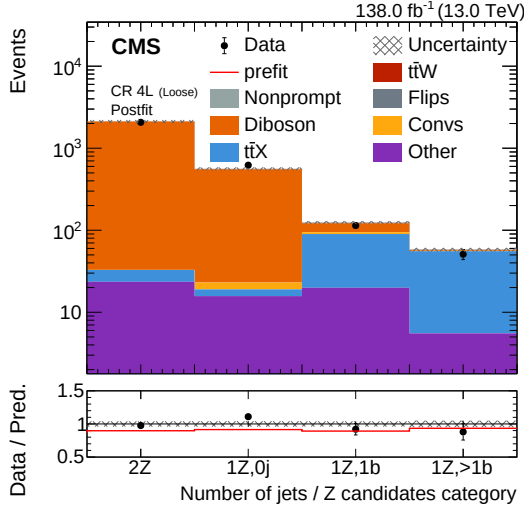


Figure 5.21.: Number of selected Z candidates and b-tagged jets for events passing the fourlepton CR requirements after the fit to the data in the SR and CR.

5.6. Event-level MVA

Now that the analysis regions and the main fitted observables have been introduced, a crucial remaining ingredient is the event-level multivariate classifier used to distinguish $t\bar{t}W$ signal events from the dominant background contributions. For this purpose, a gradient Boosted Decision Tree (gBDT) implemented in the XGBoost library [89] is employed.

The model is trained to separate the $t\bar{t}W$ signal from the most relevant backgrounds in the analysis phase space, namely events with nonprompt leptons and other $t\bar{t}X$ processes such as $t\bar{t}H$, $t\bar{t}Z$ and $t\bar{t}\gamma$ production. For the training and validation procedure, 20% of the available simulated signal and $t\bar{t}X$ samples are reserved, while the remaining 80% are used in the rest of the analysis. The background contribution from nonprompt leptons

is modeled using 20% of the inclusive $t\bar{t}$ simulated sample. The resulting dataset is further split into independent training (80%) and validation (20%) subsets. These events are excluded from the final statistical analysis in order to avoid potential biases arising from overtraining or statistical correlations between the training and evaluation samples. All event selections of the dilepton signal region, described in Section 5.4.2, are applied prior to the training procedure.

The possibility of training separate classifiers for each individual data-taking year was investigated. However, due to the reduced amount of available training statistics per year, this approach resulted in a degradation of the overall classifier performance. A single combined model is therefore trained using all years simultaneously, with the data-taking period included as an additional input feature.¹ Checks were performed to see if one-hot-encoding the years and lepton categories improved the performance, which was not the case.

During the optimization of the model, it was observed that the BDT output exhibited a strong sensitivity to jet energy correction (JEC) variations. In particular, variations of the jet energy scale produced large shifts in the output score distributions, indicating that the classifier was learning features that were insufficiently robust against detector-related systematic effects. To mitigate this issue, the training dataset was augmented with samples corresponding to the JEC up and down variations. This procedure effectively exposes the model to a broader range of jet energy configurations during training, thereby improving its robustness against JEC uncertainties while simultaneously increasing the effective training sample size. The available sample sizes for all processes entering the training are summarized in Table 5.7.

¹The years are numerically encoded as follows: 2018: 3, 2017: 2, 2016PostVFP: 1, and 2016PreVFP: 0.

Table 5.7.: Sample sizes available for training.

Process	Total	Training + validation	Training
$t\bar{t}W$	1 783 021	356 604	285 283
$t\bar{t}$	178 491	35 698	28 558
$t\bar{t}Z$	807 389	161 478	129 182
$t\bar{t}H$	685 764	137 153	109 722
$t\bar{t}\gamma$	57 987	11 597	9 278

5.6.1. Input variables

In total, 37 variables are selected as input features to the BDT, all of them listed below in table 5.8. Good data vs MC agreement is observed for each of the input features. To quantify this, the p-value (resulting from a goodness-of-fit test of the prediction to the data) is given in the list below for each of the input features. Furthermore, we checked the correlation matrix between the input features both in simulation and data, and good agreement is observed between both.

5.6.2. Training setup and hyperparameter optimization

All training samples are given their MC generator weights and are reweighted to their expected cross sections and luminosity, so that the model trains against a realistic background composition. Events with negative MC weights due to NLO corrections are excluded from the training, as negative weights can destabilize the training. It was also tested to include these samples with positive weights but no significant difference in performance was observed. After these weights are implemented, the $t\bar{t}W$ sample weights are rescaled to correct for the class imbalance present in the training sample. This reweighting factor is calculated on the training sample, excluding the validation sample.

A hyperparameter optimization was performed by running a grid search with 5-fold cross-validation over the following grid (optimal values highlighted in bold):

Table 5.8.: Input variables and corresponding p-values.

Leptons	
leading lepton p_T ($p_T(l1)$)	0.17
subleading lepton p_T ($p_T(l2)$)	0.33
leading lepton η ($\eta(l1)$)	0.85
subleading lepton η ($\eta(l2)$)	0.29
leading lepton energy ($E(l1)$)	0.65
subleading lepton energy ($E(l2)$)	0.82
Jets	
leading jet p_T ($p_T(j1)$)	0.61
subleading jet p_T ($p_T(j2)$)	0.29
leading jet η ($\eta(j1)$)	0.37
subleading jet η ($\eta(j2)$)	0.98
leading jet mass	0.47
subleading jet mass	0.87
leading b-tagged jet p_T	0.47
leading b-tagged jet η	0.75
highest deepFlavor score among all jets	0.78
deepFlavor score of leading jet	0.26
deepFlavor score of subleading jet	0.98
number of jets	0.72
number of loose, medium and tight b-tagged jets	0.69
$ \eta(j) _{max}$	0.59
Correlations	
$\Delta R(l,b)_{min}$	0.40
$M(j,j)_{max}$	0.43
$p_T(j,j)_{max}$	0.81
$\Delta R(l,j)_{min}$	0.40
$\Delta R(l1,b)_{min}$	0.15
$\Delta R(l1,non-b)_{min}$	0.96
$M(l1,l2)$	0.31
$\Delta R(l1,l2)$	0.23
$\Delta \eta(l1,l2)$	0.31
Global variables	
H_T (scalar sum of p_T of the selected jets)	0.34
l_T (scalar sum of lepton p_T and MET p_T)	0.16
lepton flavour category (mm=3, me=2, em=1, ee=0)	0.21
lepton charge (-1 or +1)	0.24
p_T MET	0.18
data taking year (2018=3, ..., 2016PreVFP=0)	—

- Learning rate: 0.25, **0.1**, 0.05, 0.01, 0.005
- Number of trees: 1000, 2000, **3000**, 4000, 5000
- Maximum depth of each tree: 2, **3**, 4, 5

Balanced accuracy was used as a scoring metric to identify the best-performing hyperparameters. This test resulted in a model with a learning rate of 0.1, 3000 trees and an individual tree depth of 3. Other parameters are kept to their default values. The resulting ROC and loss curves are shown in Figure 5.22

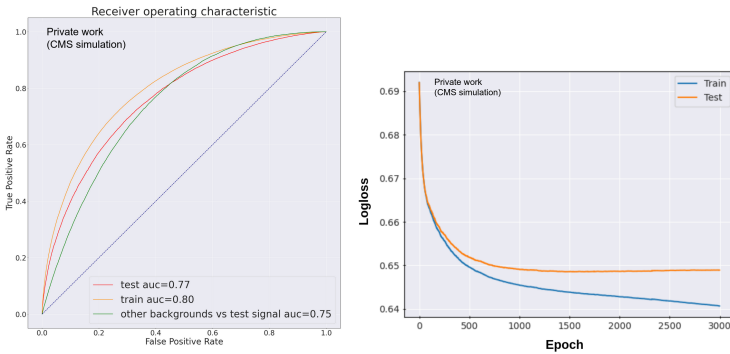


Figure 5.22.: Left: ROC curve for the BDT model. ‘Other backgrounds’ indicates the background samples that pass the signal region event selection but that are not directly trained on. Right: Training and validation loss values during model training.

In order to improve the binning of the BDT, a sigmoid function was applied on the output score of the model. The chosen sigmoid function is of the form:

$$\frac{1}{1 + \exp(-(2x - 1)/0.3)} \quad (5.9)$$

The BDT output distribution in the 2ℓ SS signal region is shown in figure 5.23. The model is additionally evaluated in the charge-flip and nonprompt control regions, where it is not explicitly

trained. The corresponding BDT output distributions are shown in Figure 5.24. The agreement between data and prediction across all three phase spaces validates the robustness of the model.

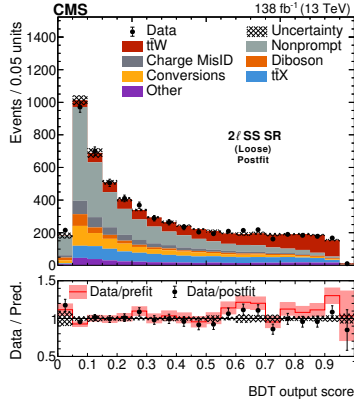


Figure 5.23.: Event BDT output score distribution in the dilepton SR after the fit to the data in the SR and CR.

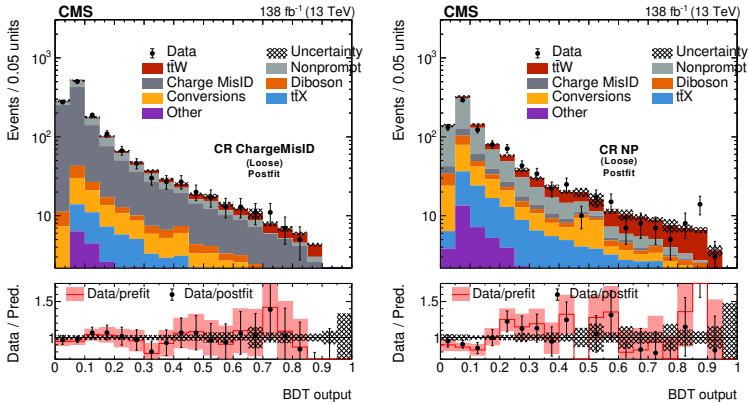


Figure 5.24.: Event BDT output score distribution in the Charge-flip (left) and NP (right) CRs after the fit to the data in the SR and CR.

5.7. Calibrations and measurement uncertainties

5.7.1. Experimental calibrations and uncertainties

Much of the work required to derive and validate these calibrations is performed centrally within the CMS collaboration, with standardized recommendations provided to individual analyses. In addition, several technical implementations were already available from previous analyses within the Ghent group. As a result, the discussion presented here focuses primarily on the associated uncertainties and their treatment in the analysis rather than on the technical details of the implementations themselves.

Luminosity: The measurement of the integrated luminosity collected during a given data-taking period is a complex procedure, which will be discussed in more detail for a dedicated dataset in Section 7.3. The luminosity determination introduces several systematic uncertainties, some of which are correlated across different years. For the dataset used in this analysis, the integrated luminosity was determined by the CMS luminosity POG. The corresponding uncertainties amount to 1.2%, 2.3%, and 2.5% for the 2016, 2017, and 2018 data-taking periods, respectively.

Trigger efficiency: In addition to the integrated luminosity, an accurate determination of the trigger efficiency is required to translate the observed event yield into a production cross section. As discussed previously, the trigger efficiency is measured using the orthogonal trigger method.

For the trilepton and four-lepton regions, the trigger efficiency is found to be compatible with 100% in both data and simulation within the statistical precision of the measurement. A flat uncertainty of 2% is therefore assigned to these regions. In the dilepton region, residual differences between data and simulation are corrected using a two-dimensional reweighting procedure as a

function of the transverse momenta of the leptons. The associated systematic uncertainty is evaluated by varying the reweighting factors within their uncertainties and propagating the effect to the final result. The trigger efficiency uncertainties are treated as uncorrelated between different data-taking years, as they are dominated by statistical uncertainties. The trigger uncertainty was measured by Luka Lambrecht [88].

Pileup: The number of proton–proton interactions occurring in a single bunch crossing cannot be controlled on an event-by-event basis by the LHC. Only the average number of pileup interactions can be adjusted through machine parameters such as the beam separation, crossing angle and the value of β^* . As a result, the pileup conditions observed in data can differ from those simulated in Monte Carlo samples. In addition to differences in the average number of interactions, the modeling of the kinematic properties of pileup interactions may also be imperfect. Detector-related effects, such as an inaccurate collision center description, can further contribute to discrepancies between data and simulation in pileup-sensitive observables.

Since the true number of pileup interactions is not directly observable and several effects are intertwined, the number of pileup interactions in simulation is treated as a tuning parameter rather than a directly measurable physics quantity. The simulated pileup distribution is therefore reweighted using the total inelastic proton–proton cross section in order to reproduce the distributions observed in data. The quality of this tuning is validated using several pileup-sensitive observables, including:

1. **Number of reconstructed primary vertices:** The number of reconstructed interaction vertices provides a direct estimate of the pileup activity in an event. The correlation between both parameters is shown in figure 5.25.

2. **Energy density observables (ρ):** The quantity ρ estimates the average transverse momentum density per unit area in the event (split in PF components) and is particularly sensitive to diffuse soft activity originating from pileup interactions.

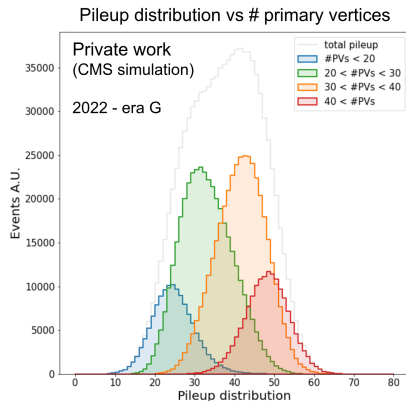


Figure 5.25.: Correlation of the number of reconstructed primary vertices (color) and the pileup (x -axis).

After the nominal pileup reweighting is applied, the total inelastic cross section is varied by $\pm 4.6\%$ to estimate the associated systematic uncertainty. This variation is chosen to cover the residual disagreements observed between data and simulation in the pileup-sensitive observables, which are not all simultaneously described equally well. The resulting uncertainty affects both the normalization and the shape of the simulated distributions and is applied to all simulated samples.

The pileup uncertainty is treated as fully correlated across all processes and data-taking years. While the correlation across processes is well motivated, the assumption of full correlation between years is less rigorous, as the inelastic cross section is effectively used as a tuning parameter rather than as a directly

measured physical observable. The pileup uncertainty recipe was centrally provided for RUN 2. The author only contributed to central studies for early RUN 3 data-taking years.

Prefiring issue: During RUN 2 data-taking, a data acquisition issue affected a small fraction of events in the ECAL trigger system. In these cases, ECAL energy deposits were incorrectly assigned to the previous bunch crossing. This effect, referred to as *prefiring*, can cause the Level-1 trigger to accept the previous bunch crossing instead of the correct one. Due to bandwidth limitations and trigger dead time, this may subsequently lead to the rejection of the actual event of interest in the current bunch crossing. In addition, the incorrect assignment of energy deposits can distort the reconstructed missing transverse momentum in both bunch crossings. The effect was first identified in 2016. In subsequent years, mitigation measures were implemented that strongly reduced its impact.

The probability for prefiring is measured for all classes of physics objects, as a function of their kinematic properties. These measurements are used to derive event weights that model the probability for an event not to prefire. In simulation, each event is assigned a weight corresponding to the product of the non-prefiring probabilities of the individual physics objects in the event. These weights should be very close to 1 as this prefiring is a very small effect.

The prefiring probabilities are varied within their uncertainties and propagated through the full analysis chain. Similar to the pileup reweighting procedure, the prefiring correction is applied to all simulated samples and affects both the normalization and shape of the predicted distributions. The prefiring issue was studied centrally within the CMS collaboration, and the corresponding correction was derived and validated by dedicated groups. In this analysis, the author only implemented the recommended

prescription.

Lepton identification: Small discrepancies are observed between data and simulation in the efficiency of selecting tight leptons using the leptonMVA discriminator. These differences are corrected by applying per-lepton scale factors to simulated events. The scale factors are measured using the tag-and-probe method in DY events as a function of the relevant lepton kinematic variables. Details of the measurement procedure and the resulting scale factors used in this analysis can be found in ref. [88].

The uncertainties associated with these scale factors are separated into statistical and systematic components. The systematic uncertainties are treated as fully correlated across the different data-taking years, while the statistical uncertainties are assumed to be uncorrelated. In addition, the electron and muon scale factors are treated as independent sources of uncertainty. Although the corresponding efficiencies are measured using similar methodologies, their dominant uncertainty contributions originate from detector- and reconstruction-specific effects that are largely distinct for electrons and muons.

Lepton energy calibrations: The simulated momentum and energy measurements of electrons [90] and muons [91] are calibrated centrally to match the detector response observed in data.

For electrons, the calibration is derived using the resonant DY process, where the reconstructed invariant mass distribution of the e^+e^- pair is compared between data and simulation. Residual differences in the position and width of the Z boson mass peak are used to extract correction factors and associated uncertainties. For muons, the calibration is obtained using several well-known resonances, including the Z boson, the J/Ψ , and the $\Upsilon(1S)$ states. These measurements constrain both the momentum scale and

resolution of reconstructed muons.

The resulting uncertainties are treated independently for electrons and muons, while being considered fully correlated across all simulated processes and data-taking years.

Jet energy corrections and resolution: In contrast to leptons, jets originate from the hadronization of quarks and gluons and are therefore governed by non-perturbative QCD effects. Their reconstruction is consequently much more challenging than that of isolated leptons. Individual particles inside jets can produce overlapping calorimeter deposits, pileup interactions contaminate the detector response, and the calorimeter response itself is neither perfectly uniform nor fully linear with respect to the true jet energy.

To mitigate these effects, a sequence of jet energy corrections (JEC) is applied. These corrections account for contributions from pileup, detector response effects derived from simulation, and residual discrepancies between data and simulation measured in dedicated calibration studies. In addition, the jet energy resolution (JER) in simulation is smeared to reproduce the broader resolution observed in data.

The uncertainties associated with the jet energy corrections are evaluated by varying the jet energy scale up and down by one standard deviation for each independent JEC uncertainty source. In this analysis, a slightly simplified scheme of combined sources is considered. These variations are propagated through the full event reconstruction and selection chain, thereby consistently accounting for their impact on derived quantities such as p_T^{miss} , jet multiplicities, and multivariate discriminants. Depending on the underlying source, the uncertainties are treated as either correlated or uncorrelated between data-taking years, while remaining

correlated across all simulated processes.

The jet energy resolution uncertainty is evaluated using an analogous procedure, where the JER smearing factors are varied within their uncertainties. Similarly, uncertainties associated with unclustered energy contributions to $p_{\text{T}}^{\text{miss}}$ are propagated through the full analysis chain.

b-tagging: Identifying the flavour of the parton initiating a jet is a crucial aspect of top quark physics analyses. In this work, the DEEPJET algorithm is used to identify jets originating from bottom quarks. No other tagging information (i.e. charm- or gluon-tagging output scores) is employed in the analysis.

As flavour-tagging algorithms rely on subtle signatures in the internal structure of jets, small differences in tagging performance are observed between data and simulation. These differences are corrected using data-to-simulation scale factors derived from dedicated calibration measurements in a variety of control regions and using several complementary methods. The different approaches yield mutually consistent correction factors within their uncertainties.

Per-event correction factors are computed as the product of the per-jet scale factors for all jets in the event. These factors account for the measured difference in the probability of correctly identifying a jet of a given flavour and kinematic configuration in data relative to simulation.

The associated systematic uncertainties are evaluated by varying the scale factors up and down within their uncertainties. Several independent uncertainty sources are considered, corresponding to both statistical and systematic components of the calibration procedure. These uncertainties are treated as fully correlated

across all simulated processes, while their correlation across data-taking years depends on the origin of the uncertainty source.

5.7.2. Theoretical uncertainties

Renormalization and factorization scales:

As discussed in Section 2.3, matrix-element calculations are performed at a finite order in perturbative QCD and therefore remain affected by uncertainties from missing higher-order contributions. These theoretical uncertainties are estimated by varying the renormalization and factorization scales up and down by a factor of two with respect to their nominal values. Variations are performed both independently and simultaneously, excluding the extreme combination where one scale is varied up and the other down. The renormalization scale determines the energy scale at which the strong coupling constant α_s is evaluated, while the factorization scale separates the perturbative hard-scattering process from the non-perturbative parton distribution functions. In an all-orders calculation, physical predictions would be independent of these scale choices. The residual dependence observed at finite order therefore provides an estimate of the size of missing higher-order corrections and serves as a proxy for the convergence of the perturbative series.

These uncertainties are treated as correlated across data-taking years, but uncorrelated between different simulated processes, since the impact of missing higher-order corrections depends on the specific process topology and order of calculation.

PDF uncertainties: The description of the proton partons is separated by μ_F and is described using the PDFs. These distributions cannot be calculated perturbatively and are instead determined from global fits to a wide range of experimental measurements. As a consequence, the PDFs carry intrinsic uncertainties that propagate to the predicted cross sections and kinematic distributions used in the analysis. The PDF uncertainty is evaluated using the

eigenvector variations provided with the PDF set employed in the simulation. In this analysis, the full set of 100 PDF eigenvector variations is considered. For each eigenvector variation, the event weights provided by the generator are propagated through the full analysis chain, allowing both normalization and shape effects to be evaluated. These uncertainties are treated as fully correlated between the data-taking years and between the processes to which they are applicable.

Initial state and final state radiation:

The remaining non-perturbative physics is described by the parton shower model. The modeling of additional QCD radiation emitted before or after the hard scattering process constitutes an important source of uncertainty in simulated event samples. Initial-state radiation (ISR) corresponds to radiation emitted by the incoming partons prior to the hard interaction, while final-state radiation (FSR) describes radiation emitted by the outgoing particles produced in the scattering process.

The uncertainty associated with ISR and FSR modeling is estimated by varying the corresponding parameters controlling the parton shower evolution in the simulation. Dedicated variations of the shower scales are applied independently for ISR and FSR, resulting in alternative simulated event weights that probe the sensitivity of the analysis to the amount of additional radiation. The corresponding uncertainties are treated as correlated across data-taking years and different simulated processes.

Process yield and other dedicated uncertainties: While renormalization and factorization scale variations are often used to estimate normalization uncertainties of simulated processes, this approach is not always well justified in a restricted analysis phase space. The selected signal and control regions may probe only a limited kinematic regime of a given process, where the behavior of scale variations is not necessarily representative of the true theoretical

uncertainty. Additional dedicated normalization uncertainties are therefore assigned to several background processes in order to provide a more realistic estimate of their overall modeling uncertainty.

An a priori normalization uncertainty of 30% is assigned to the $Z\gamma$ and $t\bar{t}\gamma$ processes. For the WZ , ZZ , and $t\bar{t}Z$ backgrounds, the overall normalizations are allowed to float freely in the likelihood fit and are constrained directly from data through the inclusion of dedicated control regions.

The $t\bar{t}H$ process is assigned a normalization uncertainty of 10%, corresponding approximately to the theoretical uncertainty associated with PDF and renormalization/factorization scale variations. Other rare background categories, including the ‘TX’ and ‘TTX’ processes (dominated by contributions such as $t\bar{t}VV$) as well as multiboson production, are assigned a conservative normalization uncertainty of 50% to account for the limited theoretical precision and the reduced experimental constraints available in the analysis phase space.

Next to these uncertainties, the WZ and ZZ processes are known to exhibit significant mismodeling in their jet multiplicity spectra, especially in regions with high jet activity relevant to this analysis. To account for this effect, dedicated shape nuisance parameters are introduced that modify the jet spectrum of these processes. These nuisance parameters are constrained *in situ* through the inclusion of the trilepton and four-lepton control regions in the likelihood fit. A similar dedicated shape uncertainty is introduced for the charge-misidentification background, allowing the jet activity spectrum of events containing charge-misidentified electrons to vary within the fit.

5.8. Maximum likelihood fit and statistical inference

One of the most important components of a CMS physics analysis is the statistical interpretation of the observed data. For this analysis, this is performed using the **Combine** framework [92], a generic maximum-likelihood fitting framework developed within the CMS collaboration and originally designed for Higgs boson measurements. The framework is based on template likelihood fits and has since become the standard statistical tool used throughout CMS for measurements, searches, and interpretations. Much of the theoretical discussion below is based on the same combine reference (ref. [92]).

5.8.1. Theoretical background

At its core, the method compares the observed data with a prediction consisting of signal and background contributions. To illustrate the principle, consider a simple counting experiment with a single analysis bin. The predicted number of events is given by

$$n_{\text{exp}} = s + b, \quad (5.10)$$

where s and b denote the expected signal and background yields respectively. If the observed number of events is n_{obs} , the probability of observing this outcome is described by a Poisson distribution:

$$\mathcal{L} = \text{Poiss}(n_{\text{obs}} \mid n_{\text{exp}}). \quad (5.11)$$

To measure the signal contribution, a signal strength parameter μ is introduced, scaling the signal prediction relative to its nominal expectation:

$$n_{\text{exp}}(\mu) = \mu s + b. \quad (5.12)$$

The parameter μ acts as the parameter of interest (POI) in the fit. A value of $\mu = 1$ corresponds to the nominal prediction, while $\mu = 0$ represents the background-only hypothesis.

In realistic analyses, the predicted signal and background yields are affected by many systematic uncertainties, as discussed in the last section. These effects are incorporated through nuisance parameters $\vec{\theta}$ that modify the expected event yields and distributions. The likelihood is therefore extended to

$$\mathcal{L}(\mu, \vec{\theta}) = \prod_j \text{Pois} \left(n_j^{\text{obs}} \mid n_j^{\text{exp}}(\mu, \vec{\theta}) \right) \times \prod_i p(\theta_i), \quad (5.13)$$

where the first product runs over all analysis bins and the second product contains constraint terms for the nuisance parameters (typically Gaussian or log-normal). These constraints encode prior knowledge on the systematic uncertainties before fitting to the observed data.

The best-fit values of the signal strength and nuisance parameters are obtained by maximizing the likelihood, or equivalently by minimizing the negative log-likelihood. In practice, the quantity of interest is the profile likelihood ratio:

$$q(\mu) = -2 \ln \left(\frac{\mathcal{L}(\mu, \hat{\hat{\theta}}_\mu)}{\mathcal{L}(\hat{\mu}, \hat{\theta})} \right), \quad (5.14)$$

where $(\hat{\mu}, \hat{\theta})$ denote the global maximum likelihood estimators, while $\hat{\hat{\theta}}_\mu$ represents the conditional best-fit values of the nuisance parameters for a fixed value of μ . This profiling procedure naturally propagates systematic uncertainties into the final uncertainty on the measured signal strength.

The same framework is also used to evaluate the significance of an observed signal. For this purpose, the likelihood of the

background-only hypothesis ($\mu = 0$) is compared to the nominal best-fit hypothesis using the test statistic

$$q_0 = -2 \ln \left(\frac{\mathcal{L}(0, \hat{\theta}_0)}{\mathcal{L}(\hat{\mu}, \hat{\theta})} \right). \quad (5.15)$$

The evaluation of the significance can be performed using toys generated with the background-only hypothesis, for which the test statistic distribution is compared to the observed test statistic. This is however computationally very expensive. Under the conditions of Wilks' theorem [93] and the asymptotic approximation derived by Cowan *et al.* [94], this test statistic approximately follows a χ^2 distribution in the large-sample limit. The statistical significance can therefore be approximated directly as

$$Z \approx \sqrt{q_0}, \quad (5.16)$$

with Z expressed in units of Gaussian standard deviations. In high-energy physics, an observed significance larger than 5σ is conventionally required to claim the observation of a process.

5.8.2. Inclusive fit distributions

Throughout Sections 5.4.2 and 5.5.3, the various signal and control regions used in the analysis were introduced and motivated individually. In this section, we provide a concise overview of the distributions entering the final inclusive likelihood fits, together with the categorization strategy applied in each region. Three complementary fitting strategies are performed for the inclusive measurement. These fits primarily serve as robustness checks by partitioning the dataset into independent subsets.

The first fit splits all distributions by data-taking year. Since the measured cross section is a fundamental physical quantity, it should be independent of the year in which the data were recorded. Consistent results across the three years therefore

provide an important validation of the detector calibrations, background estimations, and analysis methodology.

A second fit divides the events according to the lepton-flavor composition of the final state. The relative yields in the different channels are determined by the well-known branching fractions of the W boson, and therefore all channels are expected to yield compatible measurements of the fiducial cross section.

The final inclusive fit categorizes events according to the lepton charges. Unlike the previous two partitions, the expected yields in these categories are not required to be identical, as the $t\bar{t}W$ process exhibits a genuine asymmetry between positively and negatively charged leptons in the final state arising from the proton–proton initial state. The distributions entering these fits are summarized in Table 5.9.

Table 5.9.: Overview of the categorization strategy for the different likelihood fits used in the analysis.

Analysis region	Distribution	Per-year fit	Per-flavor fit	Per-charge fit
Chargeflip CR	# jets	per year	per year	per year
Trilepton CR	# jets, b-jets	per year	per year	per year
Four-lepton CR	# Z , b-jets	per year	per year	per year
Nonprompt CR	event BDT	per year	per flavor	per charge
2ℓ SR	event BDT	per year	per flavor	per charge
3ℓ SR	event BDT	per year	per year	per year

5.8.3. Differential likelihood measurements

In a differential measurement, the cross section of the process of interest is measured as a function of a given observable, such as the transverse momentum of a particle or the jet multiplicity in the event. Instead of a single parameter of interest, the fit therefore determines multiple signal strengths simultaneously,

each corresponding to a specific bin of the observable under study. In this analysis, differential measurements are performed using a maximum-likelihood-based unfolding approach, which can be viewed as a natural extension of the inclusive fitting strategy discussed previously. Other traditional unfolding strategies are discussed in chapter 7.

The distributions measured by the CMS detector are not direct representations of the underlying particle-level distributions. Detector inefficiencies, finite resolutions, acceptance limitations, and reconstruction effects distort the observed event yields and cause migrations of events between bins of the measured observable. While these effects can often be treated globally in an inclusive measurement through efficiency corrections, differential measurements require a more detailed treatment in order to recover the underlying shapes of the physical distributions.

To account for these detector effects, an unfolding procedure is applied. In the likelihood-based unfolding approach, the differential cross section bins are introduced directly as parameters of interest in the fit, while the detector response connecting particle-level and reconstructed-level observables is modeled using simulated events. The likelihood fit is then performed simultaneously across all reconstructed bins and analysis regions, allowing detector effects, backgrounds, and systematic uncertainties to be treated consistently within a single statistical framework. Compared to traditional unfolding techniques, this approach naturally incorporates correlations between bins and nuisance parameters, while avoiding the need for explicit matrix inversion procedures that can become numerically unstable.

5.8.3.1. Fiducial volume

This analysis is performed at particle level. The truth observables are therefore defined using stable final-state particles after parton

showering and hadronization, but before detector simulation. To reduce the extrapolation between detector level and particle level, a fiducial phase space is defined through a set of particle-level object and event selections closely matching the detector-level analysis selections.

Particle-level leptons are constructed by combining prompt electrons or muons with nearby prompt photons using the anti- k_T algorithm with a cone size of $R = 0.1$. Selected leptons are required to satisfy $p_T > 10$ GeV and $|\eta| < 2.5$ for electrons or $|\eta| < 2.4$ for muons.

Jets are clustered using the anti- k_T algorithm with radius parameter $R = 0.4$, starting from all stable particles except prompt dressed leptons and neutrinos. Selected jets are required to satisfy $p_T > 25$ GeV and $|\eta| < 2.4$. Jets originating from b quarks are identified using ghost association of b hadrons to the reconstructed jets.

The missing transverse momentum at particle level is defined as the negative vectorial sum of all neutrinos originating from W , Z , or prompt τ decays projected onto the transverse plane.

The fiducial phase space is defined through the event selections summarized in Table 5.10.

Table 5.10.: Definition of the fiducial phase space used in the differential measurement.

Selection	Requirement
Two light leptons	Same-sign
Leading lepton p_T	> 25 GeV
Subleading lepton p_T	> 15 GeV
Number of jets	≥ 3
Number of b -jets	≥ 1

5.8.3.2. Differential variables

The variables selected for the differential measurements are summarized in Table 5.11. The choice of observables is motivated both by their sensitivity to specific $t\bar{t}W$ modeling effects [84] and by their relevance for understanding $t\bar{t}W$ production as a background in other analyses [95].

5.8.3.3. Response matrices and choice of binning

As explained, observed distributions will differ from the true distributions due to detector response. The detector response connecting reconstructed and particle-level observables is often encoded in response matrices derived from simulation. These matrices describe the probability for events generated in a given particle-level bin to be reconstructed in another detector-level bin. Response matrices are used to define an appropriate binning for the differential measurement. In particular, the binning is chosen such that the matrices remain sufficiently diagonal, limiting migrations between neighboring bins and ensuring a stable unfolding procedure.

A response matrix is typically considered well-conditioned when the off-diagonal contributions remain small compared to the diagonal entries. Several quantities are commonly used to quantify the quality of a response matrix and the associated binning choice:

- **Stability:** fraction of events from a generator level bin y_i that are observed in the corresponding detector level bin x_i .
- **Purity:** fraction of events from a detector level bin x_i that are originating from the corresponding generator level bin y_i .
- **Efficiency:** fraction of events from a generator level bin y_i that are reconstructed and selected at detector level (in any detector level bin). This is 1 if there is no fiducial selection.

Table 5.11.: Observables of interest in the differential cross-section measurements. Highlighted variables are presented in the paper.

Variable	Description
b-jet kinematics	
$\eta^{\text{b-jet } 1}$	Pseudorapidity of highest- p_T b-jet
$ \eta^{\text{b-jet } 1} $	Absolute pseudorapidity of highest- p_T b-jet
$p_T^{\text{b-jet } 1}$	Transverse momentum of highest- p_T b-jet
Angular observables	
$\Delta\eta(\ell_1, \ell_2)$	Difference in pseudorapidity between leptons
$\Delta R(\ell_1, \ell_2)$	Angular separation between leptons
$\Delta R(\ell_1, \text{b-jet})$	Min. angular separation between leading lepton, any b-jet
$\Delta R(\ell_1, \text{jet})$	Min. angular separation between leading lepton, any jet
Jet kinematics	
H_T	Scalar sum of jet transverse momenta
$\eta^{\text{jet } 1}$	Pseudorapidity of highest- p_T jet
$ \eta^{\text{jet } 1} $	Absolute pseudorapidity of highest- p_T jet
$p_T^{\text{jet } 1}$	Transverse momentum of highest- p_T jet
$\eta^{\text{jet } 2}$	Pseudorapidity of second-highest- p_T jet
$ \eta^{\text{jet } 2} $	Absolute pseudorapidity of second-highest- p_T jet
$p_T^{\text{jet } 2}$	Transverse momentum of second-highest- p_T jet
Lepton kinematics	
η^{ℓ_1}	Pseudorapidity of highest- p_T lepton
$ \eta^{\ell_1} $	Absolute pseudorapidity of highest- p_T lepton
$p_T^{\ell_1}$	Transverse momentum of highest- p_T lepton
η^{ℓ_2}	Pseudorapidity of second-highest- p_T lepton
$ \eta^{\ell_2} $	Absolute pseudorapidity of second-highest- p_T lepton
$p_T^{\ell_2}$	Transverse momentum of second-highest- p_T lepton
$\text{Max}(\eta^{\ell_i})$	Maximum absolute pseudorapidity among leptons
$\sum p_T^\ell$	Scalar sum of lepton transverse momenta
$m(\ell\ell)$	Invariant mass of the lepton pair
Object multiplicities	
$N_{\text{b-jets}}$	Number of selected b-jets
N_{jets}	Number of selected jets
$N_{\text{electrons}}$	Number of selected electrons
N_{muons}	Number of selected muons

Events failing the particle-level fiducial selection are considered outside the acceptance of the measurement and are therefore treated separately as $t\bar{t}W$ out-of-acceptance (OOA) contributions. These events do not enter the response matrix and can be considered as a background process.

For events inside the fiducial phase space that fail the detector-level selection, different conventions are possible when constructing the response matrix and defining quantities such as purity and stability. In this analysis, such events are excluded from the response matrix definition, allowing reconstruction efficiencies and migration effects to remain factorized.

The resulting response matrices are used to optimize the binning of the differential observables. An example response matrix is shown in Figure 5.26. More visually, purity of a given reconstructed bin is defined as the fraction of events reconstructed and generated in the same bin relative to all events in that row, while the stability quantifies the fraction of generated events in a column that are reconstructed in the correct bin. Although the response matrix itself is not directly inverted during the unfolding, its information is fully encoded in the likelihood model used in the fit.

5.8.4. Differential fit distributions

The same analysis regions included in the inclusive fits, presented in Section 5.8.2, are also used for the differential measurements, with the $t\bar{t}W$ signal contribution now split into the different particle-level bins of the observable under study following the binning procedure discussed in the previous section. In practice, the measurement sensitivity is dominated by the dilepton signal region, which provides the highest purity and event yield for the $t\bar{t}W$ process. Although the trilepton signal region is also defined, its fiducial phase space differs substantially from that of the dilepton selection. Rather than combining both regions into a single measurement, the trilepton region is therefore used to

define a dedicated 3ℓ fiducial measurement. Owing to its more limited sensitivity in this analysis strategy, this measurement is not discussed further in this thesis.

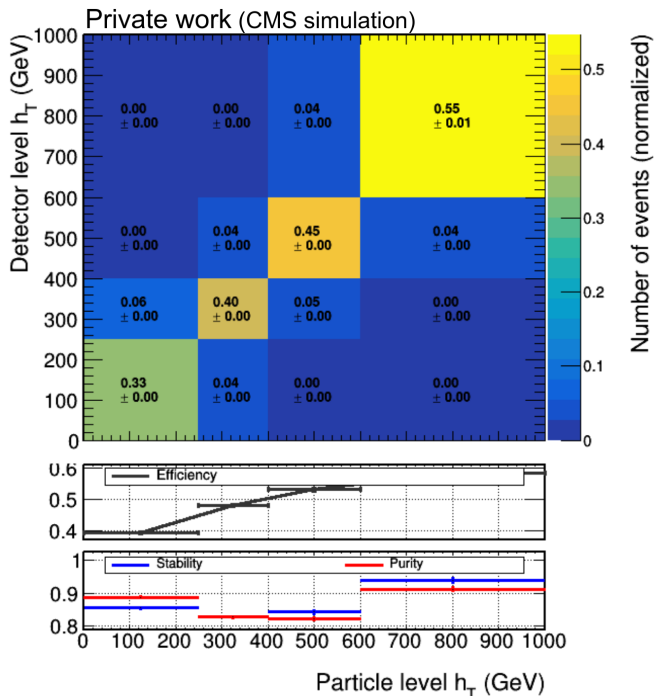


Figure 5.26.: Response matrix for the H_T observable. The horizontal axis corresponds to the particle-level quantity, while the vertical axis shows the reconstructed observable. An ideal response matrix is fully diagonal. Quality metrics are shown in the lower panels.

For the inclusive measurement, the event-level BDT output provides strong discrimination between the $t\bar{t}W$ signal and the dominant backgrounds. However, the inclusive BDT score alone has limited separation power between the different differential particle-level bins. To retain both signal discrimination and sensitivity to the unfolded observable, a two-dimensional template

is constructed using the event BDT score together with the reconstructed value of the observable of interest. The simultaneous likelihood fit can then constrain the signal strength independently in each differential bin while consistently accounting for detector migrations and systematic uncertainties.

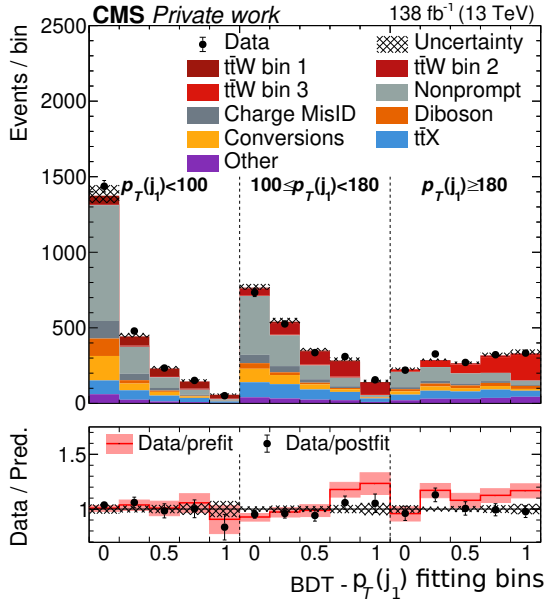


Figure 5.27.: Fitted distribution in the 2ℓ SS SR for the $p_T^{\text{jet } 1}$ fit. The binning in the BDT score provides signal sensitivity, while the hyperbins provide the detector-level binning of the response matrix. The $t\bar{t}W$ signal is split in different components that define the particle-level bins.

5.9. Robustness checks

Before applying this analysis strategy to data, numerous studies were performed to validate the robustness of the method against potential biases. Since the $t\bar{t}W$ process is known to be poorly modelled by MC simulations, particular care must be taken to ensure that the use of these samples in the unfolding procedure, and in the measurement more generally, does not introduce biases in the results. This issue is especially relevant because the behaviour of the BDT could, at least naively, induce biases in the measurement.

Several validation studies were performed and are described in this section. First, a series of injection tests is presented, in which pseudo-data samples are generated using different signal hypotheses. In the second subsection, the consistency of all differential measurements is validated.

5.9.1. Injection tests

As a first signal injection test, all differential measurements were performed on pseudo-data, in which the observed data were replaced by the sum of all nominal background contributions and a signal sample reweighted at the event level according to the PL value of a single unfolding variable at a time. This test was carried out for a selection of 21 unfolding variables. Agreement between the measured differential cross sections and the modified signal cross sections injected into the pseudo-data indicates that the analysis is unbiased with respect to these simple signal modifications.

It was chosen to reweight all events based on a linear function of the unfolding variable, with a shape difference between the first and last bin of up to 40%. The selected variables of study and their reweighting functions are shown in the table below.

Table 5.12.: Event weight parametrizations used in the analysis.

Observable	Weight
# jets	$1 + 0.1 \times (\# \text{jets} - 3)$
# b-jets	$1 + 0.1 \times (\# \text{b-jets})$
# loose b-jets	$1 + 0.2 \times (\# \text{loose b-jets})$
# muons	$1 + 0.2 \times (\# \text{muons})$
H_T	$1 + 0.0004 \times H_T$
p_T leading jet	$1 + 0.0015 \times p_T \text{ leading jet}$
p_T subleading jet	$1 + 0.0015 \times p_T \text{ subleading jet}$
p_T leading b-jet	$1 + 0.002 \times p_T \text{ leading b-jet}$
$ \eta $ leading jet	$1 + 0.2 \times \eta_{\text{leading jet}} $
η leading jet	$1 + 0.2 \times \eta_{\text{leading jet}}$
$ \eta $ subleading jet	$1 + 0.2 \times \eta_{\text{subleading jet}} $
η subleading jet	$1 + 0.2 \times \eta_{\text{subleading jet}}$
$ \eta $ leading b-jet	$1 + 0.2 \times \eta_{\text{leading b-jet}} $
η leading b-jet	$1 + 0.1 \times \eta_{\text{leading b-jet}}$
$\Delta R^{\min}(l_1, \text{jet})$	$1 + 0.15 \times \Delta R^{\min}(l_1, \text{jet})$
$\Delta R^{\min}(l_1, l_2)$	$1 + 0.15 \times \Delta R(l_1, l_2)$
$ \eta_{l1} $	$1 + 0.2 \times \eta_{l1} $
η_{l1}	$1 + 0.2 \times \eta_{l1}$
$ \eta_{l2} $	$1 + 0.2 \times \eta_{l2} $
η_{l2}	$1 + 0.2 \times \eta_{l2}$
$\max(\eta_{l1} , \eta_{l2})$	$1 + 0.2 \times \max(\eta_{l1} , \eta_{l2})$

All differential distributions of the reweighted pseudodata samples are measured in the same strategy as for data. A few of these results are shown in figure 5.28. Following this method, the unfolded jet multiplicity distribution using pseudodata in which the signal is reweighted as a function of jet multiplicity should close by construction, as observed. This is however not true for when a different distribution is measured than the reweighted one. While no bias was found to be problematic, a small bias was found

for variables that are not strongly correlated with H_T when H_T or correlated variables, such as jet p_T are modified. The reason for this bias is a significant correlation of the BDT output with the measured H_T . A vastly different H_T distribution will hence lead to a significantly different overall BDT shape of the signal, leading to a bias for variables that are not strongly correlated with H_T and hence can not recover this shape modification. This small bias and ways of mitigating such biases will be discussed in more detail in section 5.9.2.

A more sophisticated set of signal injection studies was also performed. In these studies, all differential measurements were repeated on pseudo-data samples in which the observed data were replaced by the sum of all nominal background contributions and an alternative LO signal sample containing variations of several SMEFT operators. These operators are activated by assigning non-zero values to their corresponding Wilson coefficients (WCs). Each operator was varied individually, and an overview of the chosen Wilson coefficients and their magnitudes is provided in Table 5.13. Event weights were then modified according to

$$\text{weight}_{\text{EFT}} = \text{weight}_{\text{SM}} \cdot (1 \pm WC \cdot w_1 + WC^2 \cdot w_2), \quad (5.17)$$

where the linear and quadratic EFT contributions are included through the coefficients w_1 and w_2 , respectively.

The resulting modifications to the reconstructed jet multiplicity distribution are shown in Figure 5.29. This set of SMEFT operators produces a broad range of distorted distributions with different correlation structures, thereby providing a stringent test of the robustness of the analysis method. Compared to the simpler reweighting studies described in the previous section, these tests probe sensitivity to significantly more complex signal modifications that are still physical in nature.

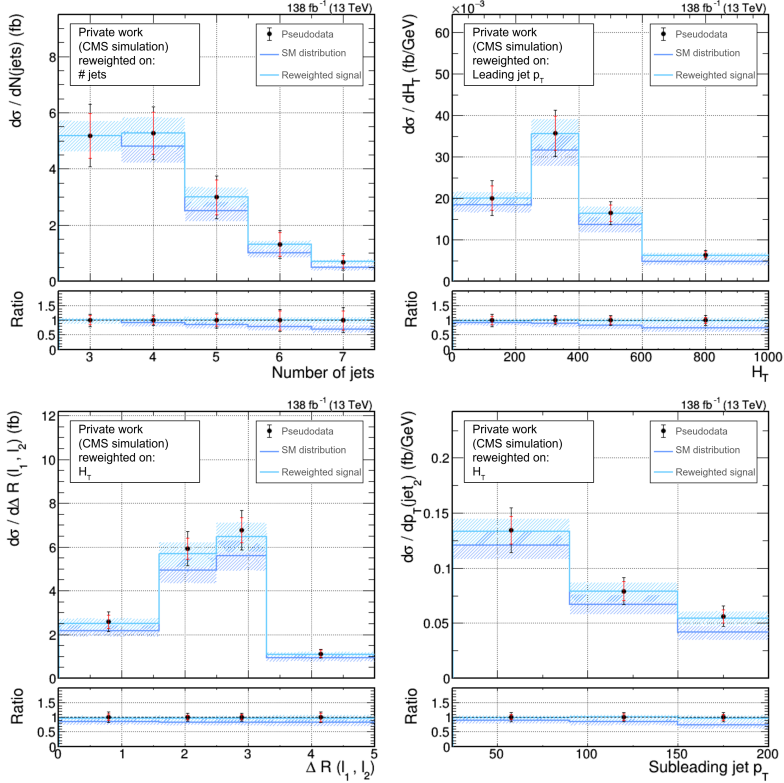


Figure 5.28.: Results of the pseudodata injection tests for selected differential measurements. Top left: unfolded jet multiplicity distribution using pseudodata in which the signal is reweighted as a function of jet multiplicity. Top right: unfolded H_T distribution using pseudodata reweighted as a function of the leading jet p_T . Bottom left: unfolded ΔR between leptons using pseudodata reweighted as a function of H_T . Bottom right: unfolded subleading jet p_T distribution using pseudo-data reweighted as a function of H_T .

Among the considered operators, $C_{Qq}^{(1,8)}$ and $C_{Qq}^{(3,8)}$, both with Wilson coefficients of magnitude 1, produce the largest deviations from the SM expectations. Although these operators generate highly pathological distributions that are not expected to occur in real data and the underlying physics model is not relevant to these tests, they nevertheless provide a useful benchmark for evaluating potential biases in the analysis procedure.

Table 5.13.: Used Wilson coefficients.

WC	magnitude	WC	magnitude	WC	magnitude
C_{Qei}	5	$C_{Qq}^{(1,8)}$	1	C_{pt}	5
C_{Ql3i}	5	$C_{Qq}^{(3,8)}$	1	C_{ptb}	5
C_{QlMi}	5	$C_{Qq}^{(1,1)}$	0.1	C_{tG}	0.3
$C_{Qq}^{(1,3)}$	0.1	C_{bW}	5	C_{tW}	1
C_{pQ3}	5	C_{pQM}	5	C_{tZ}	5
C_{tei}	5	C_{tlSi}	5	C_{tp}	5
C_{tTi}	5	C_{tli}	1	C_{tq}^1	0.3
C_{tq}^8	0.3				

Some representative results of this study are shown in Fig. 5.30. Good closure is observed for all considered Wilson coefficients, with the exception of the two highly pathological cases: $C_{Qq}^{(3,8)}$ and $C_{Qq}^{(1,8)}$. For these operators, the same bias discussed in the previous section is observed. The strong correlation between the BDT output and H_T leads to significant modifications of the overall BDT shape. For observables that are only weakly correlated with H_T , this shape distortion cannot be fully recovered, resulting in an overcompensation of the background contribution during the fit.

A harder H_T spectrum, as observed for $C_{Qq}^{(3,8)}$, causes the fit to suppress the background contribution while enhancing the signal contribution. Conversely, a softer H_T spectrum would induce the opposite behaviour. Although such extreme distortions are not

expected to occur in real data, these studies nevertheless revealed a small potential bias in the analysis strategy.

To provide a tool when pathological distributions would have been observed during unblinding, a more complex alternative strategy was developed. This strategy is discussed in detail in the following section. This strategy was however not needed for the final results.

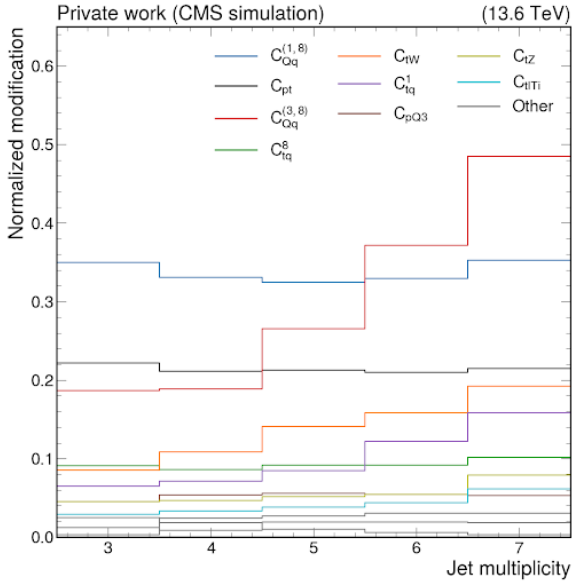


Figure 5.29.: Alternative signal number of jets distribution, based on the activation of several SMEFT operators.

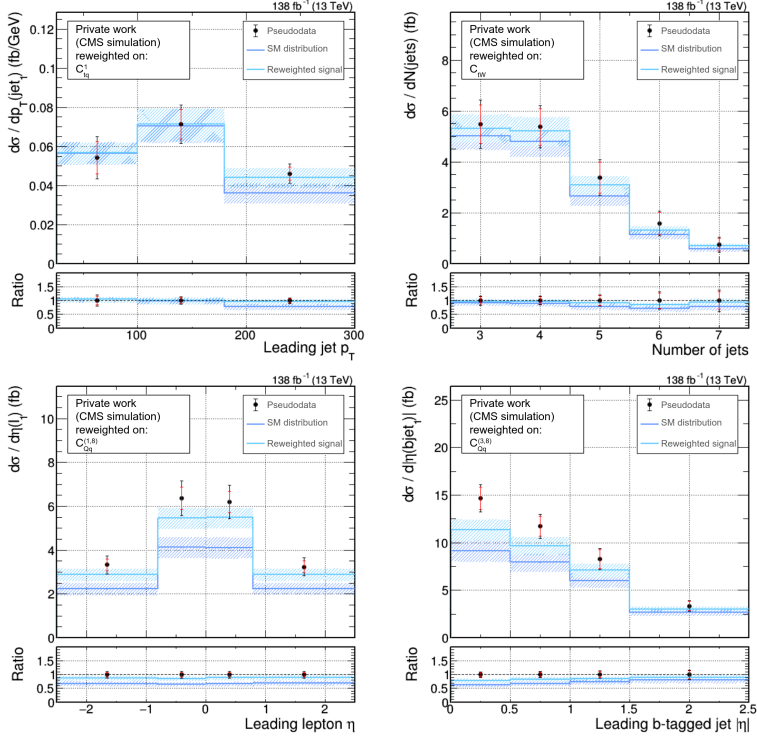


Figure 5.30.: Results of the SMEFT pseudodata injection tests for selected differential measurements. Top left: unfolded jet multiplicity distribution using pseudodata with C_{tq}^1 operator. Top right: unfolded H_T distribution using pseudodata with C_{tW} operator. Bottom left: unfolded ΔR between leptons using pseudodata with $C_{Qq}^{(1,8)}$ operator. Bottom right: unfolded sub-leading jet p_T distribution using pseudo-data with $C_{Qq}^{(3,8)}$ operator.

5.9.2. The Uniform Gradient Boosting with Flatness Loss method

Given the small potential bias observed in the previous section, a possible modification of the BDT training procedure is presented here to mitigate these effects. The observed bias becomes visible when the signal BDT shape is strongly correlated with a measured observable whose distribution is poorly modelled. The most straightforward way to address this issue is therefore to decorrelate the signal BDT output from the problematic observable. After such a decorrelation, modifications of this observable no longer alter the overall shape of the signal BDT output distribution.

In the specific case of H_T , simply removing the variable from the training is insufficient, since a large fraction of the information contained in H_T can still be reconstructed indirectly from other input variables. Furthermore, the method should ideally allow simultaneous decorrelation with respect to multiple observables if required. Finally, removing variables from the training generally results in a significant degradation in classification performance. It is important to note, however, that only the signal BDT response must be decorrelated from the problematic variable, unless the background modelling of that variable is also unreliable. Consequently, part of the lost discrimination power can still be recovered through modifications of the background response shape.

Motivated by these considerations, a loss function was developed and integrated into the XGBoost framework, that later turned out to be very similar to the Uniform Gradient Boosting with Flatness Loss (uGBFL) method. For a detailed description of the method, the reader is referred to the uGBFL publication [96]. The central idea is to introduce an additional term in the loss function that quantifies the difference between the signal BDT shape in each bin of the decorrelated observable and the inclusive signal BDT shape. This flatness term is combined with the classification loss through a weighting factor α , which controls the trade-off

between classification performance and decorrelation strength. The flatness loss is based on the Standard Deviation of Efficiency on Bins (SDE) metric. The binning definitions were chosen to match those used in the analysis.

$$\begin{aligned}\epsilon_{\text{BDT bin } i}^{\text{sig}} &= \frac{N_{\text{BDT bin } i}^{\text{sig}}}{N_{\text{tot}}^{\text{sig}}}, \\ \epsilon_{\text{BDT bin } i}^{(\text{var bin } j)} &= \frac{N_{\text{BDT bin } i, \text{var bin } j}^{\text{sig}}}{N_{\text{var bin } j}^{\text{sig}}}, \\ SDE(\epsilon_{\text{BDT bin } i}) &= \sqrt{\sum_j w_j \left(\epsilon_{\text{BDT bin } i}^{(\text{var bin } j)} - \epsilon_{\text{BDT bin } i} \right)^2}, \\ \mathcal{L} &= \mathcal{L}_{\text{classification}} + \alpha \cdot \mathcal{L}_{\text{flatness}} \\ &= \mathcal{L}_{\text{logloss}} + \alpha \sum_i SDE(\epsilon_i)^2.\end{aligned}$$

Minimizing this additional term effectively introduces a transport mechanism that redistributes BDT bin contents in order to make the signal BDT shape more uniform across all bins of the decorrelated observable. Within the XGBoost framework, however, the loss function itself is not explicitly implemented; instead, only its gradient and Hessian are defined. For simplicity, the flatness contribution was propagated only to the gradient, while the Hessian was kept identical to that of the classification loss. Since the Hessian only affects the Newton boosting step, this approximation has a minimal impact on the training behaviour.

The proposed method introduces two important hyperparameters: α and w_j . The parameter α must be sufficiently large to achieve meaningful decorrelation, while excessively large values dominate the classification gradient and impair the learning process. It

was observed that assigning enhanced weights only to the highest H_T bin, which exhibits the strongest problematic correlations, allowed effective decorrelation to be achieved with relatively small α values. In addition, improved stability for larger α values was obtained by first pretraining the classifier and only introducing the flatness loss at a later stage of the training.

Based on these observations, the model was first pretrained for 1500 iterations using only the classification loss, after which the uGBFL loss term was activated for an additional 1500 iterations. The w_j values were chosen to be 2 for the highest H_T bin and 1 for all other bins. After manual optimisation of the remaining hyperparameters, the value $\alpha = 0.8$ was selected as the smallest value that still provided satisfactory shape decorrelation, as can be seen in figure 5.31.

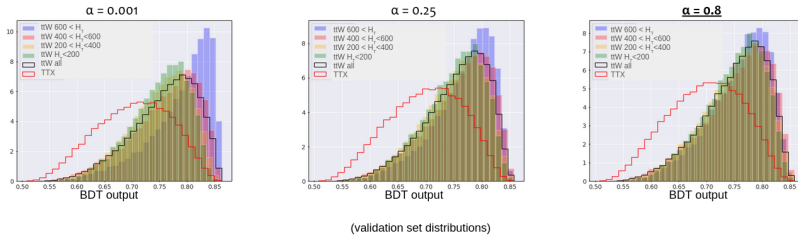


Figure 5.31.: Optimization of the α parameter. Lower values show that the shape is not well corrected, while $\alpha = 0.8$ shows good decorrelation.

The improvement in the robustness of fits for observables weakly correlated with H_T is illustrated in Fig. 5.32. However, for observables strongly correlated with H_T , a reduced robustness is observed. In these cases, the BDT becomes increasingly sensitive to secondary correlations with H_T in order to recover classification performance.

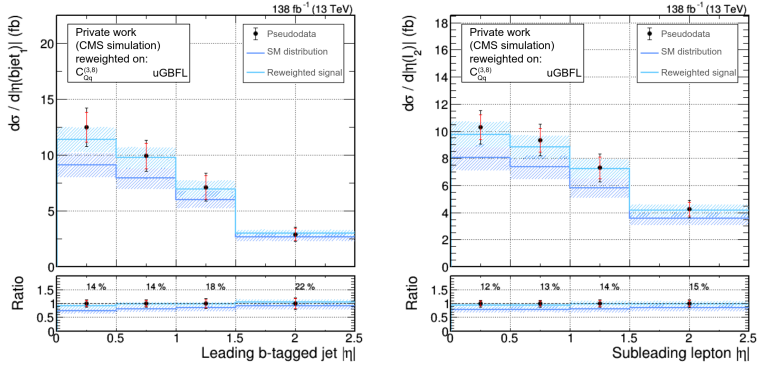


Figure 5.32.: Results of the $C_{Qq}^{(3,8)}$ operator pseudodata injection tests with the uGBFL BDT. Left: leading b-tagged jet pseudorapidity. Right: subleading lepton pseudorapidity.

A multidimensional extension of the uGBFL method was also investigated. However, no satisfactory closure behaviour was obtained, indicating that the current architecture is not well suited for such complex decorrelation tasks. Alternative approaches, such as adversarially trained networks, may offer greater flexibility in handling multidimensional decorrelations. Nevertheless, these methods generally suffer from significantly larger extrapolation instabilities and are therefore considerably less attractive for analyses that require high robustness.

Since the observed bias is small and data-MC differences are not as extreme, no further studies of these alternative approaches were pursued. The recommendation to future analyses is therefore to apply the uGBFL method only for unfolding observables that are weakly correlated with the poorly modelled variable against which the classifier was decorrelated.

5.9.3. Self-consistency of differential measurements

Several checks were performed to validate the consistency of all differential measurements after the unblinding. Since all differential measurements correspond to the same fiducial phase-space definition, the sum of the differential cross sections is expected to reproduce the inclusive fiducial cross section. To verify this, the fiducial cross section obtained from the inclusive fit was compared to the total fiducial cross sections extracted from each differential measurement. The results of this comparison are shown in Figure 5.33.

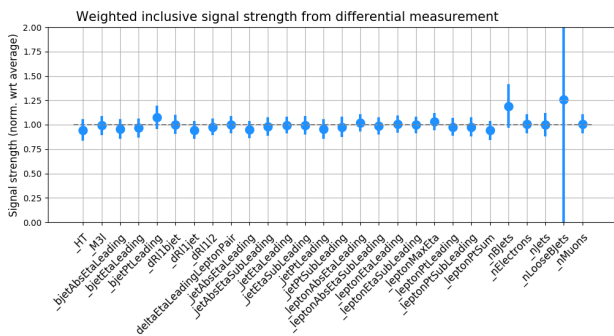


Figure 5.33.: Fiducial $t\bar{t}W$ cross sections extracted from the differential fits, normalized to the result obtained from the inclusive fit. All measurements, with the exception of the b-jet multiplicity observables, yield consistent fiducial $t\bar{t}W$ cross sections.

The differential measurements based on the b-jet multiplicity exhibited poor fit performance, primarily due to limited statistical precision in the higher multiplicity bins. Furthermore, these observables mainly probe the b-tagging acceptance and provide limited sensitivity to additional heavy-flavour radiation. For these reasons, the b-jet multiplicity observables were excluded from the

final results.

All remaining differential measurements were found to be mutually consistent. Nevertheless, only the subset of observables listed in Table 5.11 was retained for the final publication. The remaining observables were omitted either because they mainly served as validation distributions for reconstructed physics objects (e.g. the number of electrons or muons), contained less informative quantities (such as η instead of $|\eta|$), or were highly correlated with other selected observables (for example the leading b-jet and leading jet observables).

Next to the measured cross sections, the systematic uncertainty models of the different differential fits are also expected to converge towards a consistent underlying physics description within uncertainties. To validate this behaviour, the impact plots of all differential measurements were studied in detail and the consistency of the nuisance parameter pulls was examined. Representative examples are shown in Figure 5.34. Overall, the fitted nuisance parameter pulls were found to be highly consistent across the different differential measurements, indicating good agreement between the statistical models within their associated uncertainties.

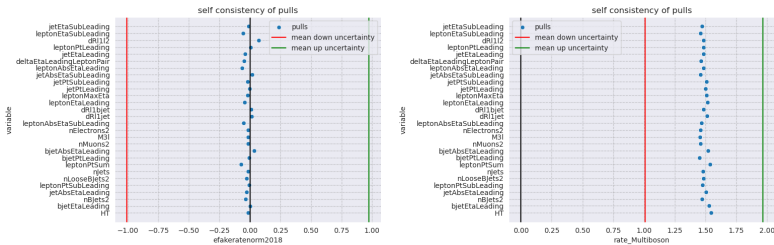


Figure 5.34.: Comparison of the pulls for a selection of nuisance parameters across all differential measurements. The fitted statistical models are found to be mutually consistent within uncertainties.

For completeness, an impact plot is shown for the differential measurement as a function of N_{jets} in figure 5.35. The plot serves as a representative example of the uncertainty breakdown and will not be discussed further.

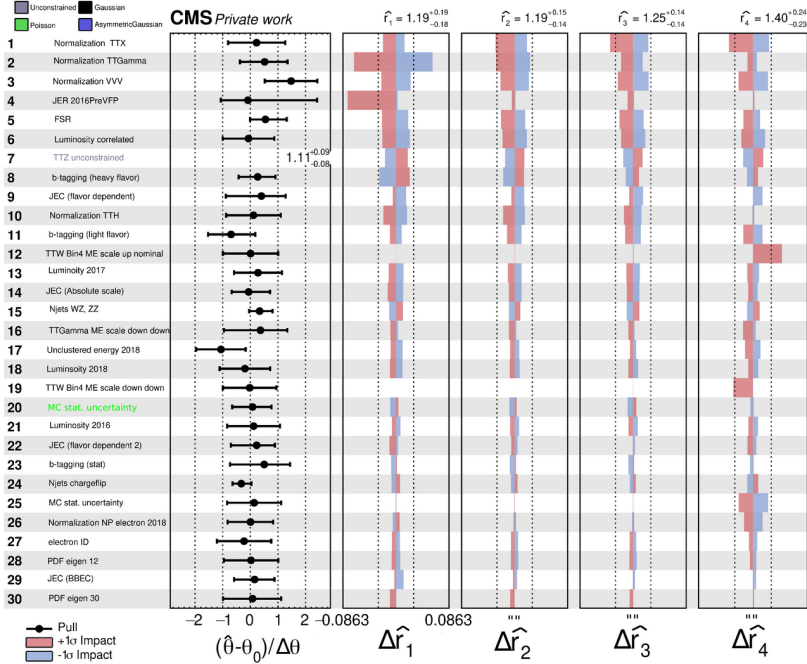


Figure 5.35.: Impact plot for the differential measurement as a function of $\Delta R(\ell_1, \ell_2)$.

5.10. Results

In this section, the results of the analysis are presented. First, the inclusive $t\bar{t}W$ cross-section measurements are discussed. These measurements primarily serve as validation and cross-checks with respect to the earlier CMS analyses presented in Refs. [86] and [97], rather than as new or improved measurements. Subsequently, the

main results of this work are presented: the differential $t\bar{t}W$ cross section measurements.

5.10.1. Inclusive cross section measurement

The inclusive $t\bar{t}W$ production cross section is measured using several fit configurations and event categorizations. All considered strategies yield mutually consistent results within their uncertainties. The most precise result (by a small margin) is obtained using the fit split by lepton flavor channels, leading to a measured cross section of

$$\sigma_{t\bar{t}W} = 938.4 \pm 44.6 \text{ (stat.)} \pm 52.1 \text{ (syst.) fb.} \quad (5.18)$$

This measurement is consistent with previous CMS results and with the Standard Model prediction within uncertainties (with a slight but persistent excess). The measured cross sections obtained with the different fit strategies are summarized below.

While the result is consistent with the previous CMS measurement within uncertainties, the statistical overlap between both analyses must first be understood before this comparison can be interpreted rigorously. Several dedicated studies were therefore performed to evaluate the overlap between the event selections used in both analyses. Since the previous measurement relied on pre-UL detector calibrations and reconstruction, the overlap between the selected event samples is found to be relatively small at only 56% in the 3ℓ signal region.

As such, both statistical and systematic uncertainties can differ substantially between both measurements. Small deviations with respect to the previous measurement are hence expected.

5.10.1.1. Fit per data-taking year

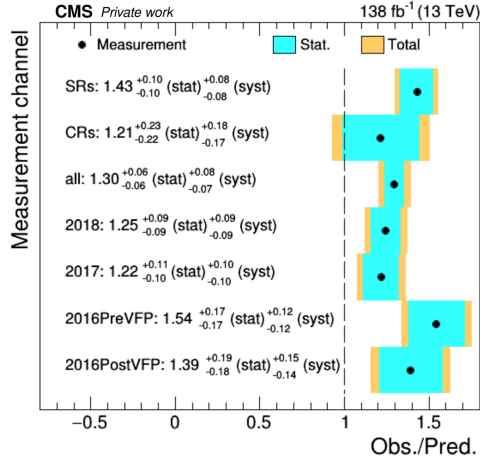


Figure 5.36.: Inclusive fit results split per year.

5.10.1.2. Fit per channel

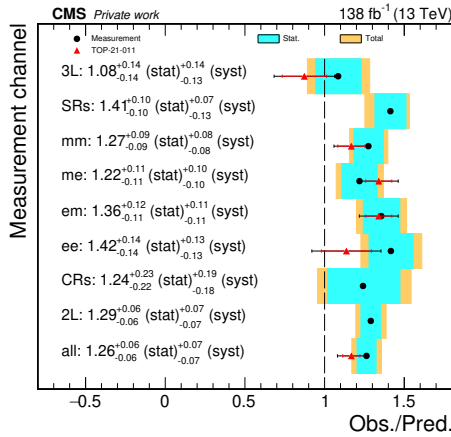


Figure 5.37.: Inclusive fit results split per lepton flavor channel. Red datapoints are results from previous analysis.

5.10.1.3. Fit per charge

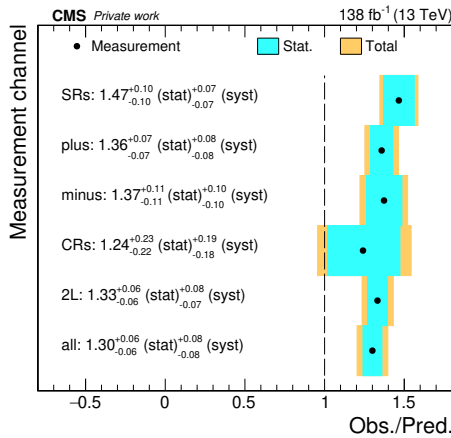


Figure 5.38.: Inclusive fit results per lepton sign

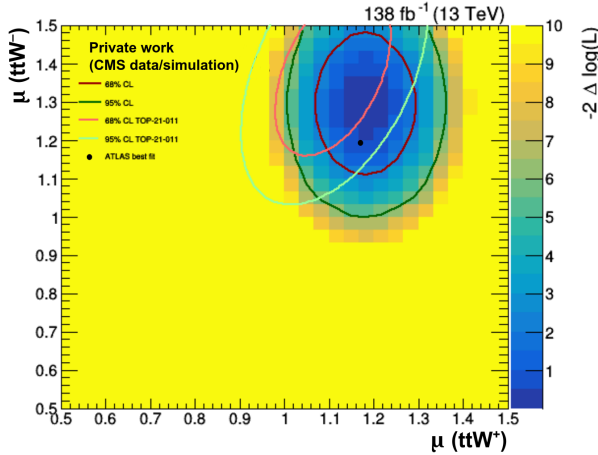


Figure 5.39.: Two Dimensional scan of the $t\bar{t}W$ signal strengths based on lepton charge. light-colored contours represent results from the previous analysis.

5.10.2. Differential cross section measurement

The analysis strategy presented in this thesis was developed in parallel with a complementary approach by the team at the University of Oviedo. The method described here relies on a relatively loose lepton selection in order to maximize the signal acceptance, while an event-level MVA is used to recover discrimination power against the dominant backgrounds. Throughout this chapter, this approach is referred to as the *MVA-based* method.

The complementary strategy developed in Oviedo instead employs a different lepton identification strategy with a substantially tighter lepton selection. This leads to signal regions with a significantly higher intrinsic $t\bar{t}W$ purity, reducing the need for an additional event-level discriminator. The tighter object selection, however, comes at the cost of reduced signal acceptance and therefore somewhat lower statistical precision in most dif-

ferential bins. This method is referred to as the *Counting* method.

The unfolded differential cross sections obtained with both analysis strategies are presented below. The MVA-based approach constitutes the primary result of this analysis. Nevertheless, the good agreement observed with the complementary strategy, which relies on a substantially different treatment of the nonprompt background, provides an important cross-check and significantly strengthens the robustness of the measurement.

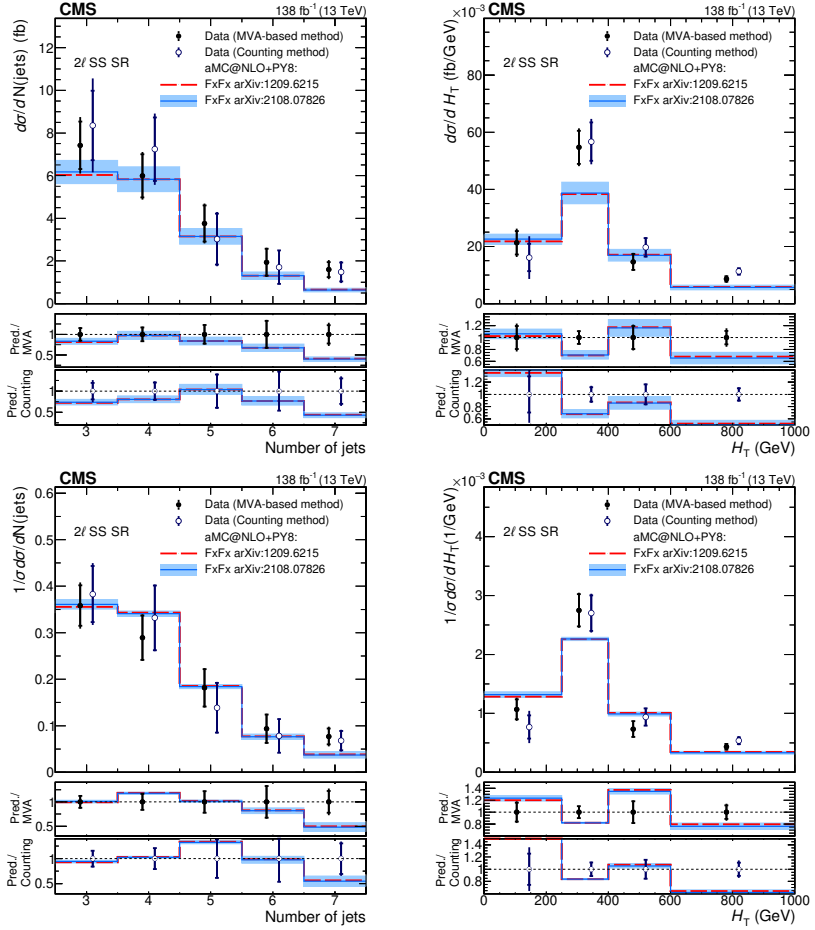


Figure 5.40.: Absolute (upper row) and normalized (lower row) differential cross section measured as a function of the jet multiplicity (left) and H_T (right), using the MVA-based and the counting method.

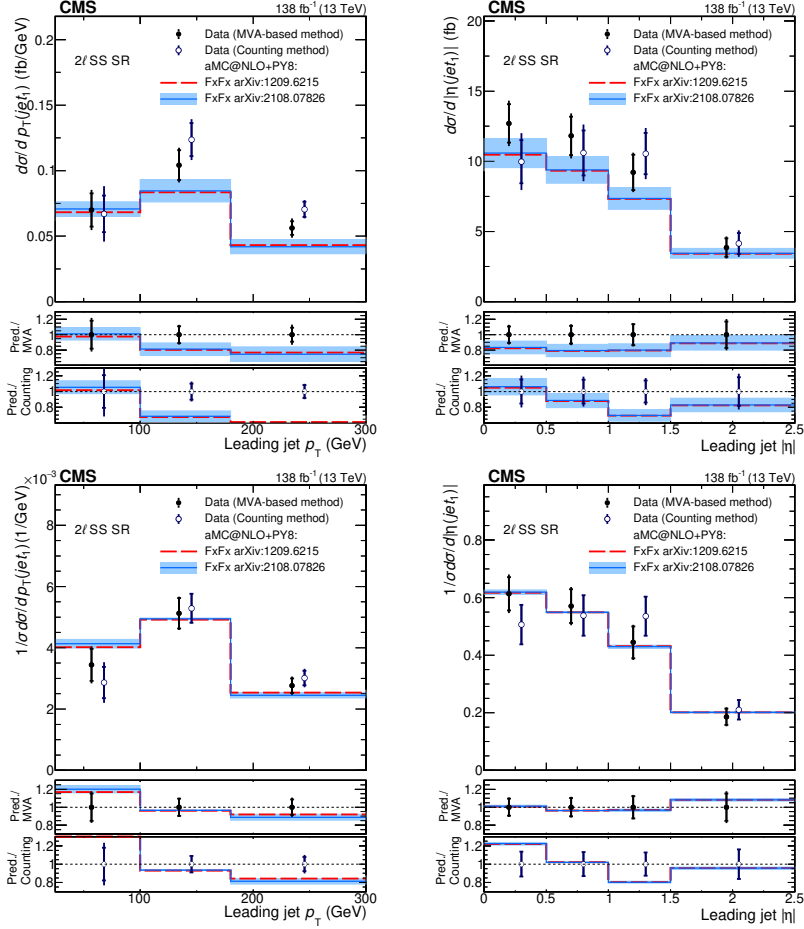


Figure 5.41.: Absolute (upper row) and normalized (lower row) differential cross section measured as a function of the leading jet p_T (left) and leading jet $|\eta|$ (right), using the MVA-based and the counting method.

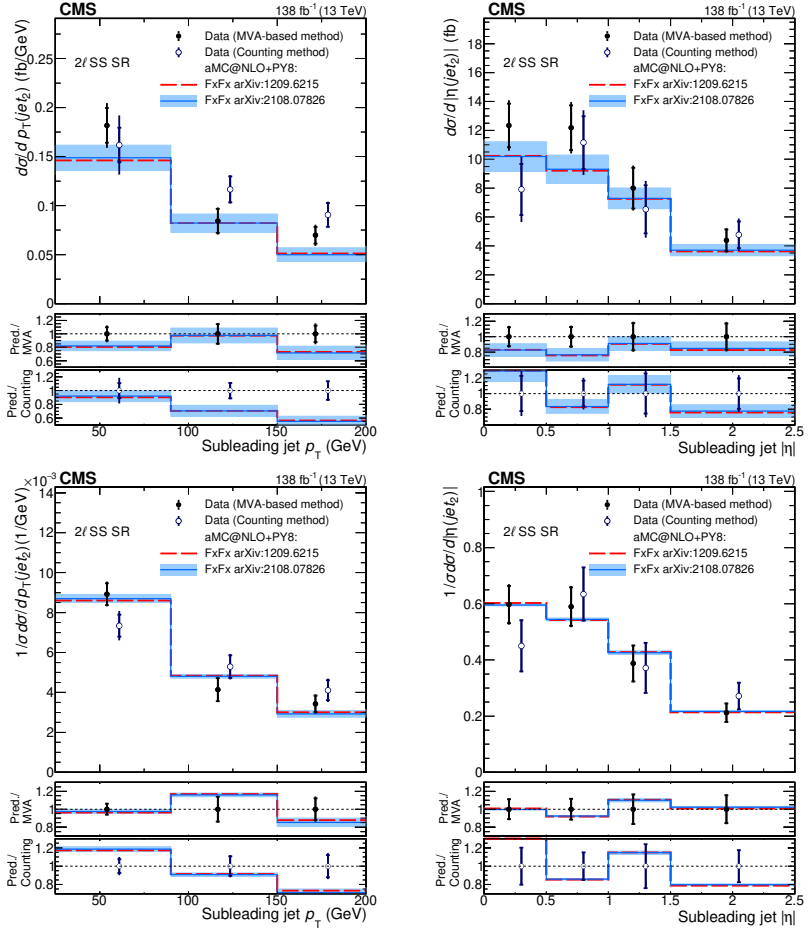


Figure 5.42.: Absolute (upper row) and normalized (lower row) differential cross section measured as a function of the subleading jet p_T (left), and subleading jet $|\eta|$ (right), using the MVA-based and the counting method.

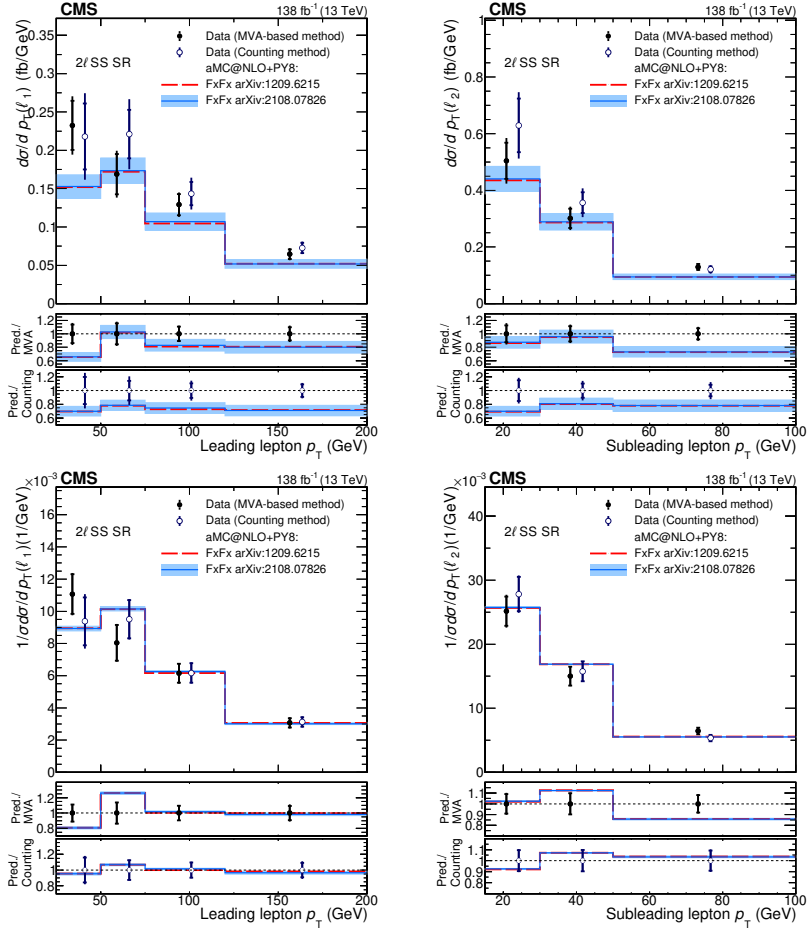


Figure 5.43.: Absolute (upper row) and normalized (lower row) differential cross section measured as a function of the leading (left) and subleading (right) lepton p_T , using the MVA-based and the counting method.

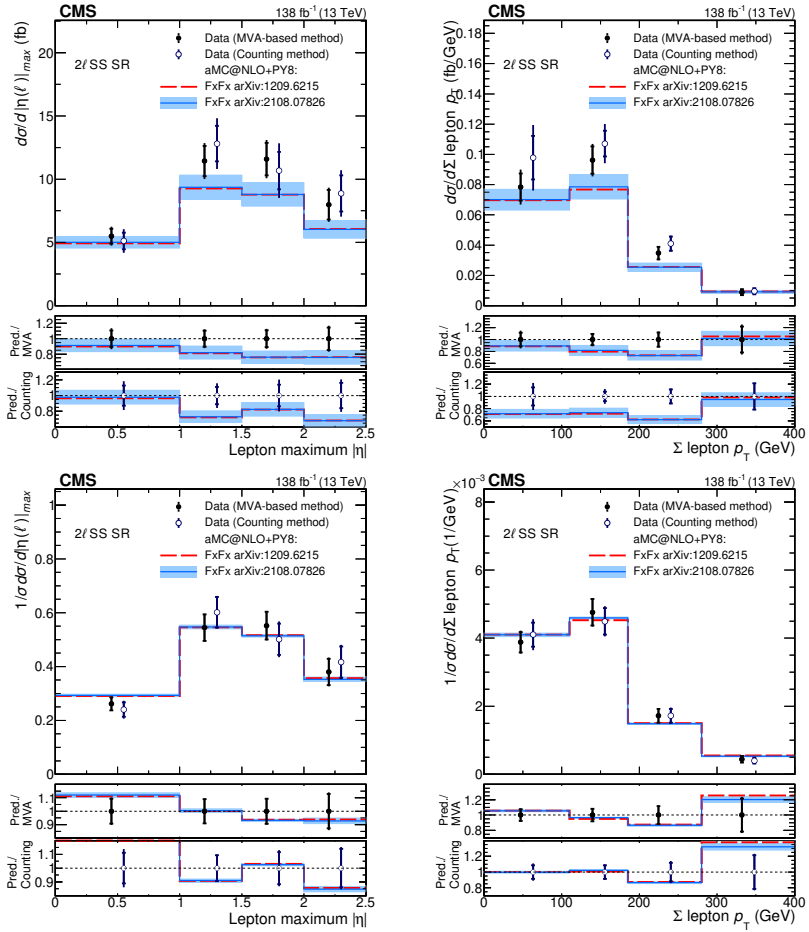


Figure 5.44.: Absolute (upper row) and normalized (lower row) differential cross section measured as a function of the maximum $|\eta|$ of the selected leptons (left) and the sum of their p_T , using the MVA-based and the counting method.

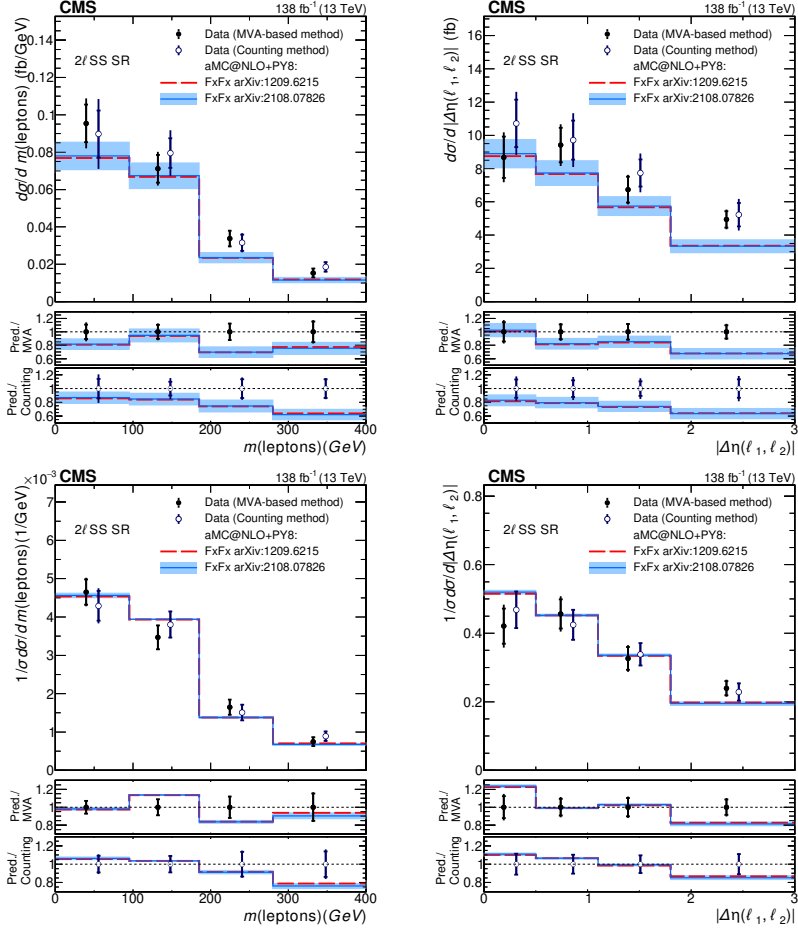


Figure 5.45.: Absolute (upper row) and normalized (lower row) differential cross section measured as a function of the invariant mass of the leptons (left) and the $\Delta\eta(\ell_1, \ell_2)$, using the MVA-based and the counting method.

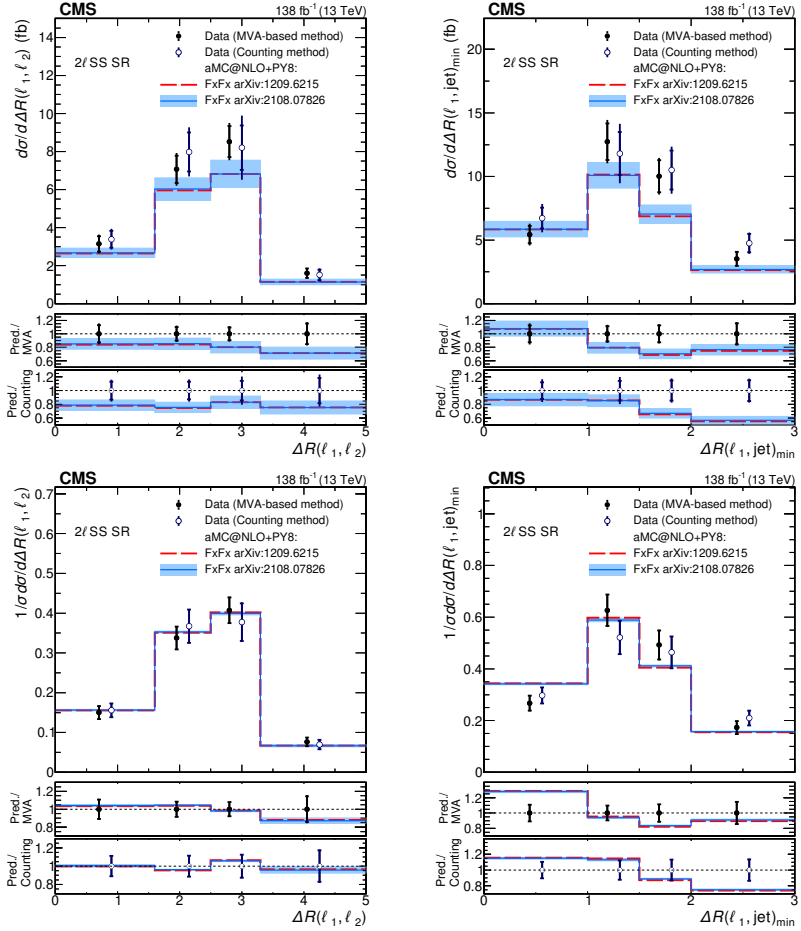


Figure 5.46.: Absolute (upper row) and normalized (lower row) differential cross section measured as a function of the $\Delta R(\ell_1, \ell_2)$ and the $\Delta R(\ell_1, jet)_{\min}$, using the MVA-based and the counting method.

5.11. Conclusions and outlook

The $t\bar{t}W$ process has proven to be one of the more intriguing challenges in the last years within top quark physics. Across multiple measurements performed by the ATLAS and CMS collaborations, the measured cross section has consistently been found to be somewhat higher than the latest state-of-the-art Standard Model predictions. Although these excesses are well established experimentally, the sizeable theoretical uncertainties and known modeling challenges have long prevented a clear interpretation in terms of physics beyond the Standard Model. A more precise understanding of the $t\bar{t}W$ process is therefore essential to determine whether these tensions originate from shortcomings in the theoretical description or from genuinely new physical effects.

The work presented in this thesis constitutes an important step in this direction. A broad set of differential measurements has been performed in order to probe possible kinematic dependencies of the observed excess. Most observables show a rather uniform scaling with respect to the inclusive cross section and do not exhibit statistically significant differential deviations from the Standard Model predictions. One of the more notable trends is observed in the jet activity associated with the $t\bar{t}W$ process, where the data tend to prefer a harder jet spectrum than predicted. Although the current statistical precision remains limited, a similar tendency has also been reported by the ATLAS collaboration.

A better understanding of the jet spectrum in $t\bar{t}W$ production will be particularly important for future measurements of processes such as ttt and $t\bar{t}t\bar{t}$ production, for which $t\bar{t}W$ constitutes one of the most challenging backgrounds. Improved modeling of $t\bar{t}W$ events in regions of high jet and b-tagged jet multiplicity will therefore directly benefit a broad range of top-quark measurements and searches.

The measurements performed by both collaborations are, however, still statistically limited. Since the completion of this analysis, the RUN 3 data-taking period has concluded, providing a dataset that is more than twice as large as that available for the measurement presented in this thesis. This substantial increase in statistics, combined with improvements in prompt-lepton identification and b-tagging algorithms, will enable significantly more precise measurements in the near future.

The larger dataset will also open new opportunities to improve our understanding of multiboson and $t\bar{t}$ + multiboson production, which constitute important irreducible backgrounds in this analysis but remain poorly constrained experimentally. As statistical uncertainties decrease, these processes will become increasingly well understood.

At the same time, considerable effort is being invested by the theory community to improve the description of $t\bar{t}W$ production. Most notably, the two-loop amplitudes required for NNLO predictions have recently been computed for the first time in the leading-colour approximation. These theoretical advances, together with the substantially larger RUN 3 dataset, will play an important role in the next generation of precision top-quark measurements and may help clarify the persistent differences between measured and predicted $t\bar{t}W$ production rates.

6 Machine learning models to discern 4 top events from other processes

In section 3.4.5, it was shown that the discovery papers of $t\bar{t}t\bar{t}$ production by both ATLAS and CMS still keep a few crucial questions unanswered. Especially the contribution of $tttj$ and $tttW$ processes is yet to be well understood. To gain a better understanding of the interplay of 4top process with this background, new techniques need to be developed. The development of these techniques is the focus of this chapter. Firstly, the development of better ttt predictions are discussed in section 6.1. The development of better ML architectures are discussed in the remaining sections. This work is part of an ongoing analysis effort. In this dissertation I will however only discuss my direct contributions.

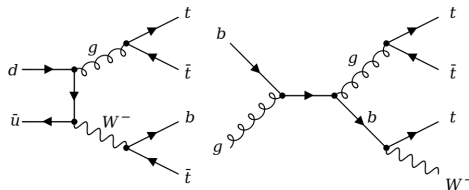


Figure 6.1.: LO terms contributing to the $tttj$ (left) and $tttW$ (right) processes.

6.1. $t\bar{t}j$ and $t\bar{t}W$ predictions at NLO QCD + LO EWK

During the period in which the ATLAS and CMS analyses leading to the observation of $t\bar{t}t\bar{t}$ production were developed, theoretical predictions for $t\bar{t}t$ processes were limited. The best available estimates for $t\bar{t}t$ production were based on leading-order (LO) samples generated with MadGraph, while next-to-leading-order (NLO) predictions for $t\bar{t}t\bar{t}$ production were already available within the same generator framework. Since then, improved theoretical calculations have become available, providing NLO QCD accuracy and leading-order electroweak (EWK) precision. These developments include a consistent treatment of interference effects between $t\bar{t}j$ – $t\bar{t}W$ and $t\bar{t}W$ – $t\bar{t}t\bar{t}$ processes, handled using a DR1 subtraction scheme. The corresponding implementation was provided by the authors as a public GitHub repository.¹ Issues with this model are discussed later on.

The work presented in this section is primarily technical in nature and focuses on integrating this model into the CMS Monte Carlo (MC) production chain. This required adapting the private implementation to the CMS simulation framework and ensuring compatibility with standard production workflows. Once integrated, the model was used to generate event samples intended for use in CMS analyses, after completing a series of validations.

The implementation was first validated by comparing the predicted cross sections of individual contributions with the reference results provided by the authors. This comparison was performed at a center-of-mass energy of 13 TeV, corresponding to RUN 2 LHC conditions. Agreement between the reproduced and reference cross sections was used as the primary consistency check of the implementation. The results of this validation are shown in Figure 6.2 for the NLO QCD term. The LO EWK terms showed

¹<https://github.com/gdurieux/triple-top-nlo>

similar agreement.

A small disagreement was observed with the implementation using the CMS nominal PDF choices. This discrepancy was resolved by adopting the NNPDF23_nlo_as_0118 parton distribution function set, as used by the authors. An increase in the cross section of approximately 20% is observed for all processes when going from 13 TeV to 13.6 TeV proton–proton collisions.

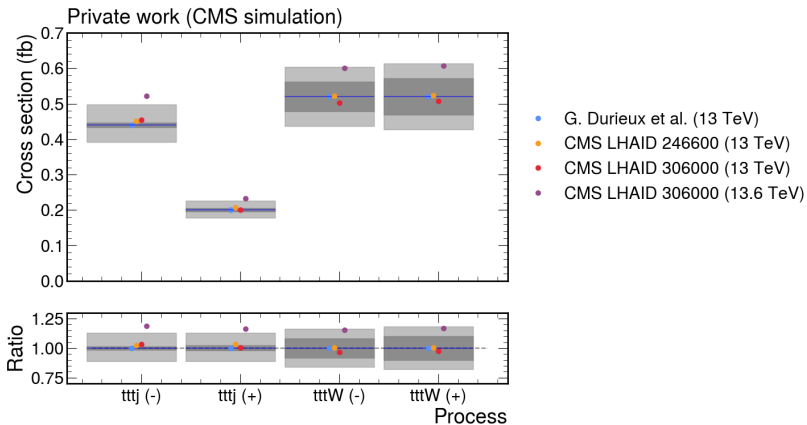


Figure 6.2.: Cross sections for the different $t\bar{t}t$ NLO QCD terms, as predicted by the original setup provided by G. Durieux et al., the CMS implementations at 13 TeV and 13.6 TeV.

With the validation of the software implementation completed, several studies were performed to select an appropriate choice of renormalisation and factorisation scales, μ_R and μ_F , and to assess the PDF dependence of the processes.

6.1.1. Scale choice

While the choice of scales is an unphysical artefact of perturbative calculations, it is nevertheless important for producing physically

meaningful Monte Carlo samples. Since the process has not yet been observed experimentally, no data-driven constraints are available to guide this choice. Instead, the primary criterion is the perturbative convergence of the calculations.

First, the convergence from leading order (LO) to next-to-leading order (NLO) can be used as an indicator of a robust scale choice. A small K -factor ($\sigma_{\text{NLO}}/\sigma_{\text{LO}}$) suggests better perturbative convergence. Second, stability under variations of the scale choice provides an additional criterion: reduced sensitivity to scale variations indicates improved convergence of the perturbative expansion.

Both approaches were investigated by the original authors, and a preliminary recommendation of $H_T/8$ was identified as optimising these criteria. However, minimising scale variation bands alone is not necessarily optimal, since these variations are intended to estimate the theoretical uncertainty on the cross section. Certain scale choices can also induce significant shape distortions, leading to crossings of the uncertainty bands and artificially reduced uncertainty estimates. For these reasons, it is essential to validate the chosen scales and assess their impact on the final results.

The author investigated multiple renormalization and factorization scale choices, including H_T , $H_T/2$, $H_T/4$, and $H_T/8$, thereby covering both the standard CMS prescription and the preliminary recommendations proposed by the original authors of the model. In addition, a fixed scale choice of $3m_t$ ($+ m_W$) was studied. Most observables exhibit only limited shape dependence on the scale choice, although moderate effects are observed in the tails of distributions such as the jet multiplicity (Figure 6.3), the Born-system mass, and the Born-system p_T . Based on these studies, the conservative choice of $H_T/2$ was adopted as the nominal scale. This choice leads to a broader and therefore less constrained estimate of the associated scale uncertainties. The normalization differences induced by the various scale choices are not propagated directly,

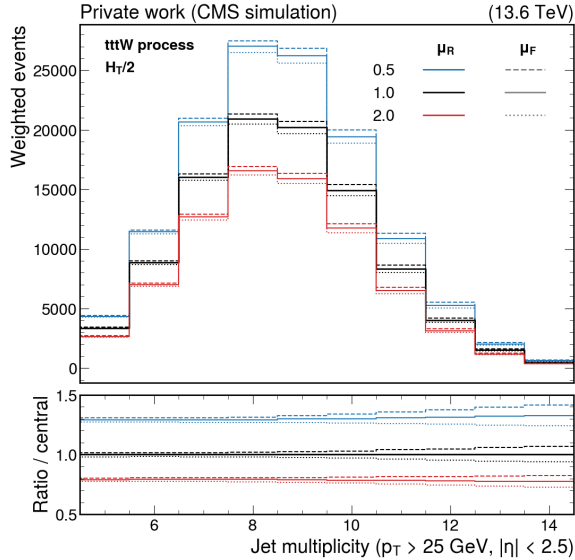


Figure 6.3.: Effects of scale variations around the central scale choice $H_T/2$ on the jet spectrum of the inclusive $tttW$ sample. μ_R variations are represented by the color, while μ_F variations are represented by the line style.

since the reference cross sections are defined independently within the context of the LHC Top Working Group. The author serves as the CMS contact for these efforts. At the time of writing, no common central prescription has yet been agreed upon within the working group; therefore, preliminary values provided by the original authors of the model are used in this work:

Process	cross section (fb)
$tttj^+$	0.21
$tttj^-$	0.49
$tttW^+$	0.66
$tttW^-$	0.66

6.1.2. PDF dependence

Both the $tttj$ and $tttW$ processes exhibit a non-trivial dependence on the proton PDFs. At leading order (LO), the $tttW$ process couples predominantly to bottom- and gluon-initiated partons and therefore closely resembles tW and t-channel single top quark production. On the other hand, the $tttj$ process, which is similar to $t\bar{t}W$ production, is initiated only by quark partons at LO. This leads to a charge ratio of $tttj^+/tttj^- \approx 0.5$. These characteristics suggest that both processes may exhibit significant sensitivity to the proton PDFs, making this an interesting field for further study. The dependence of the charge asymmetry in $tttj$ production ($tttj^+/tttj^-$) on the chosen PDF set was investigated. An overview is shown in Figure 6.4.

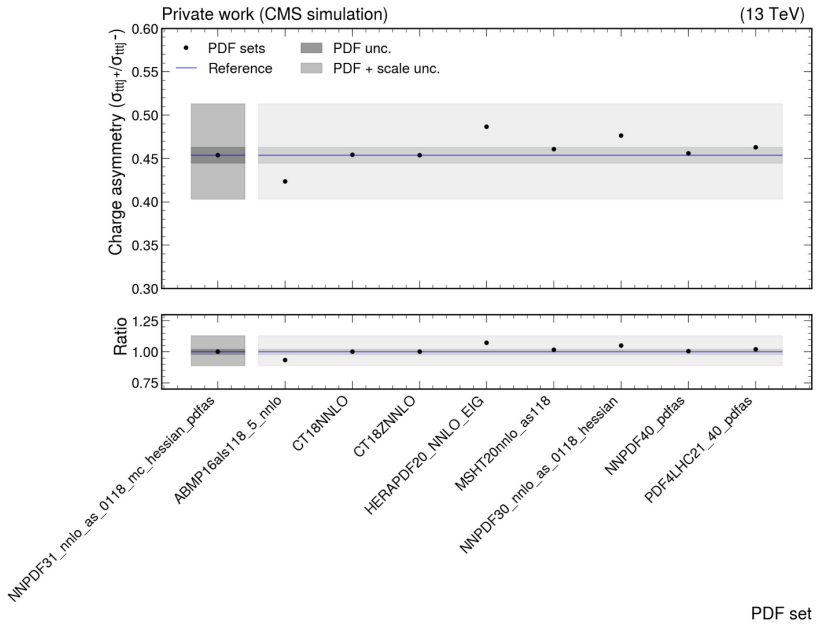


Figure 6.4.: Ratio $\sigma_{tttj^+}/\sigma_{tttj^-}$ for different PDF sets.

The charge asymmetry variations are not covered by the dedicated Hessian uncertainties of the reference PDF set. To understand these differences, the two most pronounced differences were studied in more detail. The valence quark content for the reference PDF set (NNPDF31_nnlo_as_0118_mc_hessian_pdfas) and HERAPDF20_NNLO_EIG is shown in figure 6.5. The valence down quark content is noticeably higher while the valence up quark content is lower in the HERAPDF PDF set compared to the reference PDF set, explaining the observed disagreement. Similar behavior is also observed in $t\bar{t}W$ and justifies the use of multiple PDF sets to estimate the PDF uncertainty.

In $tttW$, strong dependence on the mass of the b quark used in the PDF fits was observed. This behavior in turn strongly resembles the observations made in t -channel single top analyses. This also justifies the need of evaluating multiple PDF sets to really estimate the PDF uncertainty.

6.1.3. Interference terms

The implementation used in these samples follows a DR1 approach, in which the interference terms are neglected in the NLO calculation. Subsequent studies by the original authors, which are still ongoing, have shown that these interference contributions are in fact non-negligible and can lead to sizeable K-factors. They further observed that the DR1 amplitudes do not preserve gauge invariance and violate unitarity at high energies.

These issues clearly demonstrate the limitations of the current modeling and motivate the development of an improved treatment that consistently incorporates the interference effects. The approach currently being developed by the original authors defines a resonant window in $m(bW)$ around the top-quark mass. Nonresonant diagrams are included over the full phase space, while resonant diagrams are excluded inside this window to avoid overlap with the on-shell $tt\bar{t}$ prediction. This procedure provides

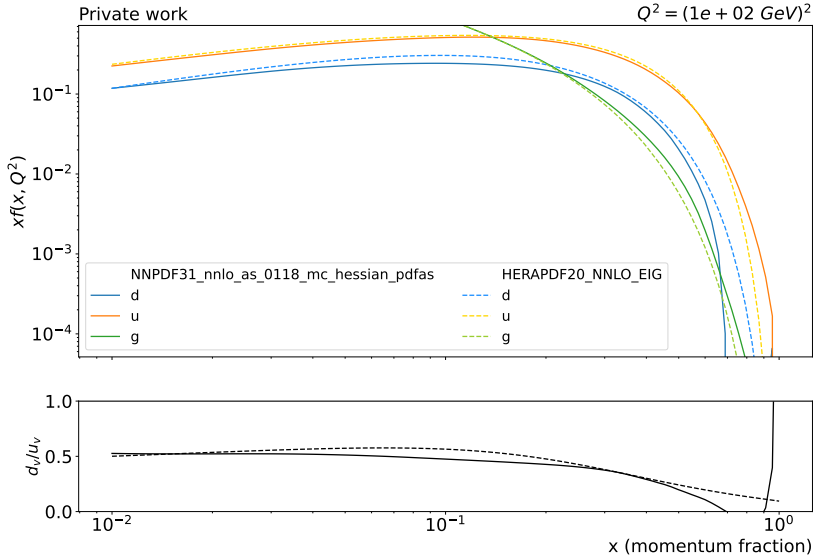


Figure 6.5.: Valence PDFs as accessible via LHAPDF [98].

a theoretically consistent combined description of the $t\bar{t}t + t\bar{t}t\bar{t}$ final state.

The implementation of this improved model, once available, will be important for future analyses. However, all further work presented in this section has been performed using the currently available model, despite its limitations, as no better alternative was accessible at the time.

Throughout this PhD, the author has generated additional samples for other analyses, including a $t\bar{t}b\bar{b}$ sample produced with Sherpa [99], a 3FS $t\bar{t}cc$ sample generated with MadGraph, and SMEFT $t\bar{t}VV$ samples generated with MadGraph. The author has also produced DR1 and DR2 $tW\gamma$ samples. As these samples are not relevant to the present work, they are not described in further detail.

6.2. Treatment of negative weights

Now that the signal samples and their limitations are well understood, another topic of relevance to these new NLO samples can be introduced: negative weights. The origin of negative weights in NLO MC samples was discussed in more detail in section 2.3.2.

Both in the ttt samples and the $t\bar{t}t\bar{t}$ sample, large fractions of negative weights can be observed, as shown in table 6.1.

Table 6.1.: Fraction of negatively weighted events for the different signal samples.

Process	Fraction of negatively weighted events
$tttj^+$	23.36%
$tttj^-$	23.89%
$tttW^+$	27.74%
$tttW^-$	28.13%
$t\bar{t}t\bar{t}$	28.39%

Negative event weights are required to reproduce the correct NLO distributions. However, their inclusion in the training of machine learning models can introduce several complications. To illustrate this, consider the weighted binary cross-entropy (BCE) loss function commonly used in classification tasks:

$$\mathcal{L}_{\text{BCE}} = - \sum_i w_i [y_i \log(p_i) + (1 - y_i) \log(1 - p_i)]. \quad (6.1)$$

For positively weighted events, both contributions are bounded from below by zero. Correctly classified events therefore yield losses approaching zero, but never negative values. Now consider a negatively weighted event belonging to class 0, for which the model prediction approaches 1. The corresponding loss contribution becomes

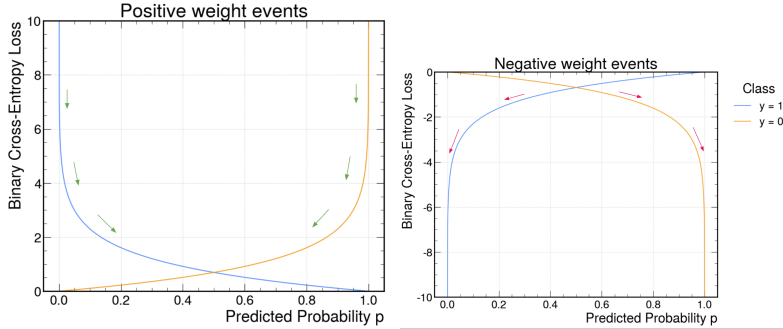


Figure 6.6.: The BCE loss for positive weight events (left) and negative weight events (right).

$$\begin{aligned}\mathcal{L}_{\text{BCE}}(i) &= -w_i \log(1 - p_i), \\ &= -w_i(-\infty) = -\infty.\end{aligned}$$

The strong penalisation of highly confident wrong predictions is therefore transformed by the negative event weight into a pathological contribution to the loss. As a consequence, the optimizer can reduce the total loss indefinitely by intentionally assigning catastrophically wrong predictions with arbitrarily high confidence to certain negatively weighted events. Instead of learning a meaningful decision boundary, the model is effectively incentivised to exploit these pathological configurations. Larger batch sizes can partially mitigate this behaviour by averaging over many events and steering the optimisation away from such extreme regions. Nevertheless, if the model still enters one of these pathological configurations, the loss remains fundamentally unbounded from below.

This issue can be alleviated to some extent by instead using a mean squared error (MSE) loss function:

$$\mathcal{L}_{\text{MSE}} = \sum_i w_i (p_i - y_i)^2. \quad (6.2)$$

Since $(p_i - y_i)^2$ is bounded from both sides, the loss no longer exhibits pathological divergences. However, the gradient contribution of negatively weighted events still points in the opposite direction, effectively encouraging the network to misclassify these samples. Although this behaviour is less problematic than in the BCE case, it can still introduce training instabilities and degrade performance when large fractions of negatively weighted events are present or when complex models are employed.

These observations highlight the importance of properly addressing negative event weights in the present use case. Historically, such issues were often mitigated either by excluding negatively weighted events from the training altogether or by replacing their weights with their absolute values. However, both approaches imply that the model is trained on an effective probability density that no longer corresponds to the physical NLO prediction. In the case where absolute event weights are used, the corrections introduced by negatively weighted events are effectively applied in the wrong direction, thereby worsening the disagreement in the distribution shapes. Excluding negatively weighted events from the training does not introduce such additional distortions, but it reduces the effective training sample size.

6.2.1. Effects of negative weights on distributions

To evaluate the bias introduced by these commonly used treatments, the impact of negatively weighted events on physical distributions was investigated in more detail. If negatively weighted events populated the phase space uniformly, replacing their weights by their absolute values would still preserve the correct probability density while retaining the full training statistics. However, since negative weights originate from the technical implementation of the virtual and real corrections at NLO, this assumption is highly

non-trivial. The non-physical structure of negatively weighted events can already be observed in several distributions. One illustrative example is the distribution of the initiating partons, shown in Figure 6.7.

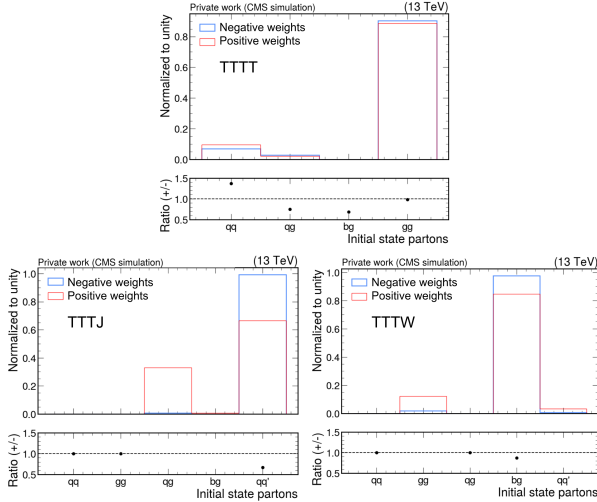


Figure 6.7.: Initial state parton fractions for weighted events.

This is particularly evident in the case of $tttj$, where no negatively weighted events are observed for configurations with real initial-state emissions at matrix-element level (qg), while all negative-weight contributions originate exclusively from quark-induced (qq) processes. This demonstrates that care must be taken to ensure that observed differences in distributions are not purely technical artefacts, but instead correspond to genuine effects.

Two physical distributions that are crucial in this analysis, are the number of jets and the number of b-jets in events after applying the PS. In figure 6.8, one can see that these distributions are significantly biased following historic treatment of negative weights. The modifications of these spectra also differ between

processes, changing the local contribution fractions of processes. These are strong indications that a better treatment is needed.

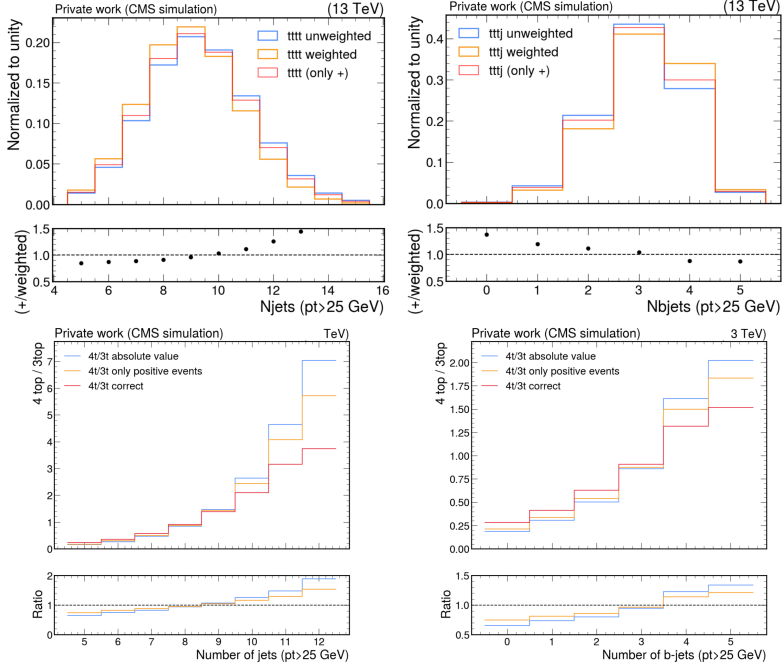


Figure 6.8.: Jet and b-jet multiplicity distributions for different treatments of negatively weighted events. Top left: jet multiplicity distribution. Top right: b-jet multiplicity distribution. Bottom left: ratio of $t\bar{t}t\bar{t}$ to $t\bar{t}t$ as a function of the jet multiplicity. Bottom right: ratio of $t\bar{t}t\bar{t}$ to $t\bar{t}t$ as a function of the b-jet multiplicity.

6.2.2. The neural reweighter and refiner approaches

Some algorithmic approaches were developed by the author to address these issues, albeit with limited success. Owing to the large fraction of negatively weighted events, simple self-developed algorithms lacked sufficient phase-space awareness to correctly

reweight the events. Consequently, implementations of the Neural Reweighter and Refiner methods proposed by B. Nachman and D. Noll [100] were adopted instead.

While the reader is referred to the original paper for a detailed description of the methods, the basic concepts are briefly summarized here. Most reweighting algorithms aim to replace individual event weights by the average weight associated with a given phase-space point x_i . This can be achieved using a modified binary cross-entropy loss function:

$$\mathcal{L}_{\text{Reweighter}} = - \int dx dz p(x, z) w(x, z) \log(g(x)) - \int dx p(x) \log(1 - g(x)). \quad (6.3)$$

Minimizing this loss is equivalent to learning the difference between the weighted and unweighted probability density distributions. Consequently, after optimisation, the ratio of the network outputs can be interpreted as a scale factor relating the unweighted distribution to the weighted one.

Following the arguments presented in the previous section, this loss function remains well behaved only if negatively weighted events do not dominate over positively weighted events throughout the phase space. The authors however note that this loss function reweights even fully positive samples towards their local absolute weight, which is not required for the present application. To address this issue, the following alternative loss function, referred to as the Refiner loss, was proposed:

$$\begin{aligned} \mathcal{L}_{\text{Refiner}} = & - \int dx dz p(x, z) \frac{|w(x, z)| + w(x, z)}{2} \log(g(x)) \\ & - \int dx dz p(x, z) \frac{|w(x, z)| - w(x, z)}{2} \log(1 - g(x)). \end{aligned} \quad (6.4)$$

Since the MC generators considered here produce event weights of the form $w_i = \pm a$, this expression can be rewritten as

$$\begin{aligned} \mathcal{L}_{\text{Refiner}} = & - \int dx \, p(x) a \, \delta(\text{positive}) \log(g(x)) \\ & - \int dx \, p(x) a \, \delta(\text{negative}) \log(1 - g(x)). \end{aligned} \quad (6.5)$$

This form makes the interpretation of the loss particularly transparent: the network effectively learns to distinguish negatively weighted events from positively weighted ones. The output score $g(x)$ can therefore be interpreted as the probability that an event at phase-space point x carries a positive weight. The average weight at a given phase-space point is then given by

$$\mathcal{P}_{\text{positive}}(x) - \mathcal{P}_{\text{negative}}(x) = g(x) - (1 - g(x)) = 2g(x) - 1. \quad (6.6)$$

Both methods were tested for the present use case.

6.2.3. Results

Each process exhibits significantly different kinematic distributions. In addition, observables that are strongly correlated with the occurrence of negatively weighted events, such as the jet multiplicity, are further modified by the applied lepton selection criteria. For these reasons, separate models are trained for each process and for each of the two relevant lepton selection regions ($2\ell\text{SS}$ and 3ℓ SRs).

The models are trained at generator level using the following input variables:

1. the p_T and η of the 10 leading jets,
2. the jet and b-jet multiplicities,
3. the scalar sum of transverse momenta, H_T ,

4. the transverse momentum of the Born system.

Some performance metrics are shown in figure 6.9.

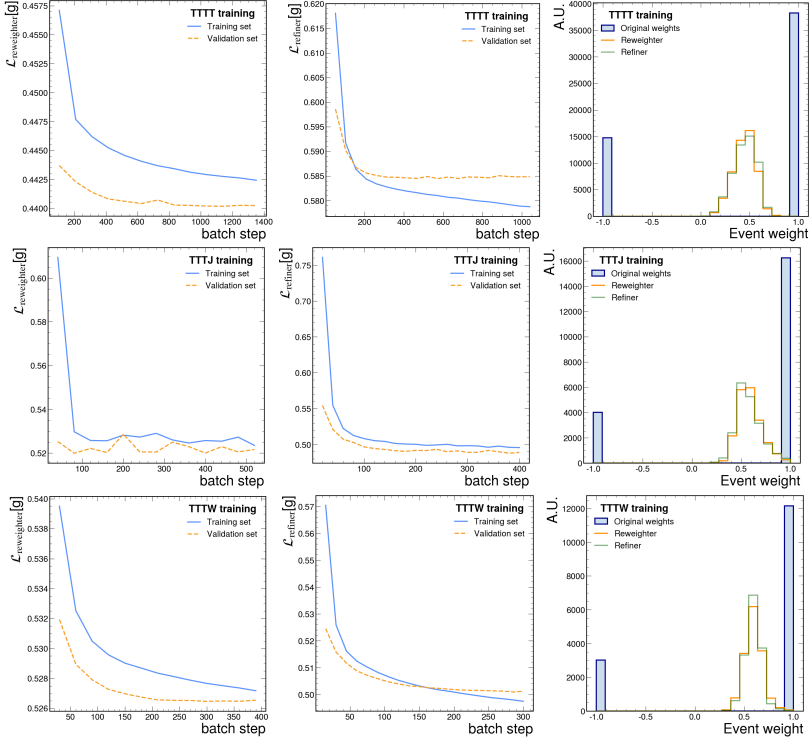


Figure 6.9.: Some training performance plots for $tt\bar{t}\bar{t}$ (top), $tttj$ (middle) and $tttW$ (bottom).

The models are trained for only three epochs. This limited training duration reduces the risk of overtraining, thereby keeping the resulting event weights close to unbiased and preserving realistic probability density distributions. Any residual overtraining would not introduce a physical bias, but would instead merely result in suboptimal event weights.

Overall, all networks exhibit excellent stability and convergence behaviour. In addition, the Refiner and Reweigher approaches produce very similar weight distributions. More importantly, their performance was validated using several physics observables, as shown in Figure 6.10. Excellent agreement is observed for all considered variables in all regions for both methods. Similar ratios are shown in figure 6.11. Both methods were used to provide a positively weighted sample with the same probability densities as the true sample. The performance of a network with these weights is discussed in the next section.

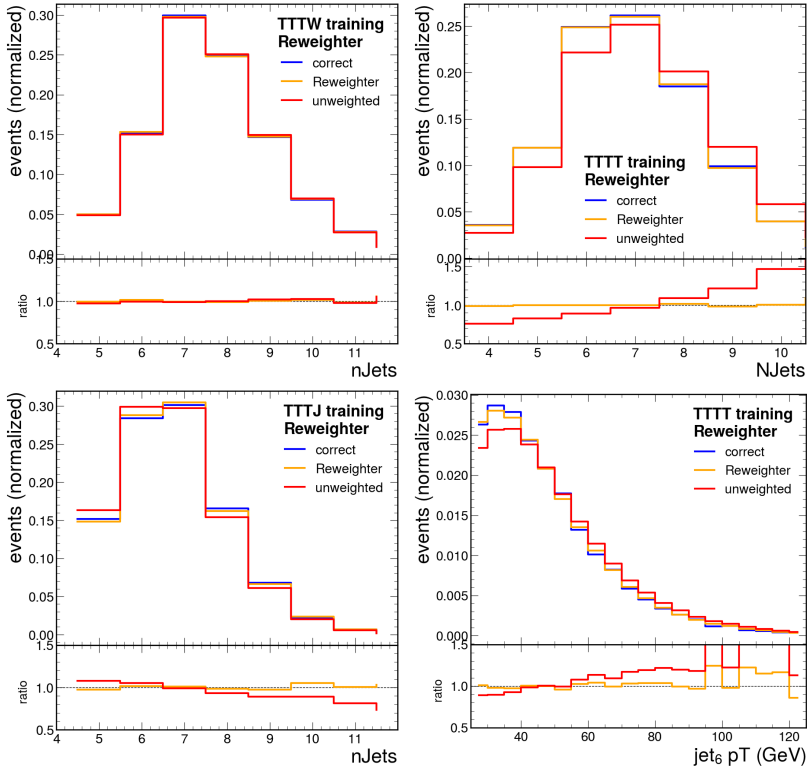


Figure 6.10.: Some training performance on physical distributions for $tt\bar{t}t$, $tttj$ and $tttW$.

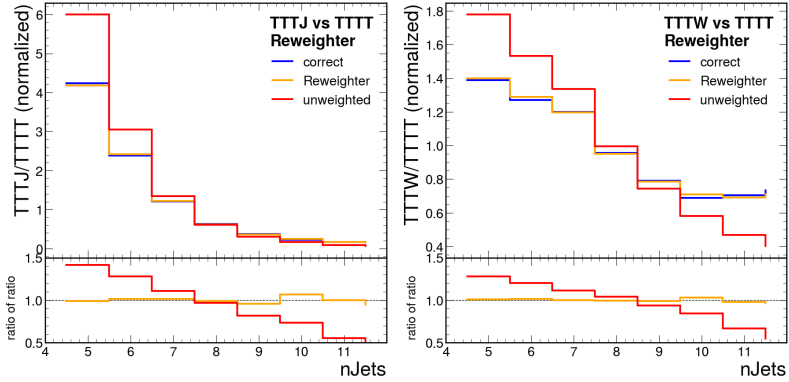


Figure 6.11.: Some training performance on number of jets distributions for $tttj$ vs $t\bar{t}t\bar{t}$ (left) and $tttW$ vs $t\bar{t}t\bar{t}$ (right).

6.3. Training of the classifier model

Now that the state-of-the-art samples are available and their weights are optimized for training, the training of the actual classifier can be discussed. The classifier model is a graph neural network that incorporates attention blocks. These blocks are the crucial component in transformer networks. The model is developed in Pytorch Geometric [101, 102].

6.3.1. Architecture

This discussion will inevitably involve a considerable amount of jargon. I therefore apologise to the reader. Non-experts are encouraged to refer to Figure 6.12, and may take away from this section the central idea that the model provides event-level predictions for whether a given event corresponds to $t\bar{t}t\bar{t}$, ttt , or $t\bar{t}$ (X). The precise structure of the network is not crucial for the discussion that follows.

The architecture begins by representing each event as a graph, an example of which is shown in Figure 6.12. In this representation,

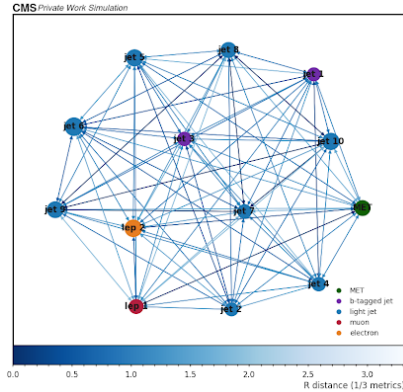


Figure 6.12.: A graph representation of an event.

each reconstructed particle corresponds to a node. Nodes are connected to all other nodes through edges, resulting in a fully connected graph. Each node is described by a set of features containing information about the corresponding particle, including its object type (jet, lepton, or MET), kinematic properties, and e.g. b-tagging information. The complete list of input features is given in Table 6.2.

For each event, the graph contains ten jet nodes, three lepton nodes, and one MET node. If fewer than ten jets or three leptons are present, the missing nodes are padded with fixed values. Conversely, when more than ten jets are reconstructed, only the ten highest- p_T jets are retained. This fixed graph size ensures a consistent input dimensionality across all events, enabling efficient batch processing and matrix operations during training.

In addition to the graph representation, the high-level variables used in the RUN 2 BDT analysis are also provided as inputs to the network [60]. With these inputs defined, the architecture of the network can be discussed. The full network layout is shown in Figure 6.13. Its central building block is the GATv2Conv layer,

Table 6.2.: Input features provided to the graph neural network. Empty jet and lepton nodes are padded with fixed values when fewer than ten jets or three leptons are present in the event.

Graph component	Features
Jet nodes	
Jet	$(1, p_T, \phi, \eta, m, \text{jet tagger node(s)})$
Empty jet	$(1, 0, 0, 0, 0, 0(, 0...))$
Lepton nodes	
Lepton	$(2, p_T, \phi, \eta, \text{charge, isMuon}(, 0...))$
Empty lepton	$(2, 0, 0, 0, 0, 0(, 0...))$
MET node	
MET node	$(3, p_T^{\text{miss}}, \phi^{\text{miss}}, N_\mu, N_{\text{jets}}, H_T(, 0...))$
Edge features	
Edge	$(\Delta R(\text{obj}_1, \text{obj}_2), k(\text{obj}_1, \text{obj}_2), M(\text{obj}_1, \text{obj}_2))$

originally introduced in Ref. [103].

Any graph neural network (GNN) layer updates every node representation by aggregating its neighbors' representations. This general operation can be written as

$$h'_i = f_\theta(h_i, \text{AGGREGATE}(\{h_j \mid j \in \mathcal{N}_i\})), \quad (6.7)$$

where h_i is the representation of particle i (the particle features), AGGREGATE is some aggregation function that aggregates the representations of all neighboring nodes (the set $\mathcal{N}_i$) and f_θ is some operation on the concatenation of h_i and this aggregate.

While most AGGREGATE operations give each neighbor equal importance, GATConv performs this aggregation based on the principle of the attention mechanism. The block constructs a scoring function e_{ij} for every edge (j, i) , which indicates the importance of the features of the neighbor j to the particle i :

$$e_{ij} = \text{LeakyReLU}\left(\mathbf{a}^\top \cdot [Wh_i \parallel Wh_j]\right), \quad (6.8)$$

where $[\parallel]$ denotes vector concatenation, the W functions are learned to map the representations to a different latent space and $\mathbf{a}^\top$ is a learnable function that describes the importance of Wh_j to Wh_i . These attention scores are normalized across all neighbors

$$\alpha_{ij} = \text{softmax}_j(e(h_i, h_j)) = \frac{\exp(e(h_i, h_j))}{\sum_{j' \in \mathcal{N}_i} \exp(e(h_i, h_{j'}))} \quad (6.9)$$

and used to get the weighted aggregation function

$$h'_i = \sigma\left(\sum_{j \in \mathcal{N}_i} \alpha_{ij} \cdot Wh_j\right). \quad (6.10)$$

In the GATv2Conv paper, the authors note that the ranking of attention coefficients in the original GAT layer is effectively independent of the query node. This can be seen by decomposing the learned attention vector $\mathbf{a} = [\mathbf{a}_k \parallel \mathbf{a}_q]$ and rewriting the attention logits as

$$e_{ij} = \text{LeakyReLU}\left(\mathbf{a}_k^\top Wh_i + \mathbf{a}_q^\top Wh_j\right). \quad (6.11)$$

For a fixed source node i , the term $\mathbf{a}_k^\top Wh_i$ is constant with respect to all neighbors j . As a result, it does not affect the relative ordering of the logits e_{ij} across different j , but only contributes a uniform shift before the softmax normalization. Consequently, the attention weights α_{ij} are determined solely by the compatibility of each neighbor embedding h_j with a global scoring function, rather than a true pairwise interaction between h_i and h_j .

This implies that the original GAT formulation cannot express query-dependent ranking of neighbors: the identity of the source node only affects the normalization, not the ordering of attention scores. In this sense, the attention mechanism is limited to a

form of context-independent scoring of candidate neighbors rather than fully conditional attention. To address this limitation, the authors propose a simple modification:

$$e_{ij} = \mathbf{a}^\top \text{LeakyReLU}(W [h_i \parallel h_j]). \quad (6.12)$$

This is further improved by concatenating the edge features into the latent space transformation

$$e_{ij} = \mathbf{a}^\top \text{LeakyReLU}\left(W \left[h_i \parallel h_j \parallel h_{ij}^{\text{edge}}\right]\right). \quad (6.13)$$

This scoring function yields an attention distribution that is fully dependent on h_i , h_j , and h_{ij}^{edge} . We therefore adopt this improved graph attention block in our network. Each block uses four attention heads, meaning that the model learns four independent sets of parameters $(\mathbf{a}, W)_k$. These attention heads can capture different correlations between nodes when computing attention scores. Intuitively, one head may become sensitive to hadronic W -boson decays, while another may respond to signatures compatible with Z or Higgs decays. In practice, however, interpreting individual attention heads in terms of physically meaningful features is highly non-trivial.

After two stacked GATv2Conv layers, the graph still consists of 14 nodes; however, these nodes no longer correspond directly to the original physical objects. Instead, each node now represents a learned latent embedding of its neighbourhood and interaction history. The feature dimension of the node representations is also no longer tied to the original transformation $W [h_i \parallel h_j \parallel h_{ij}^{\text{edge}}]$, but is determined by the chosen latent space of the network. In this implementation, we use a 30-dimensional latent space per attention head, resulting in a 120-dimensional feature vector per node after concatenation over all heads.

After these transformations, the graph representation is reduced to a 120-dimensional vector using a global mean pooling operation,

which averages the learned node features over all nodes in the event. This latent representation is then concatenated with the high-level input variables from the BDT of the RUN 2 analysis and passed to a final multilayer perceptron (MLP). The MLP outputs three scores corresponding to the probabilities of the event being classified as $t\bar{t}\bar{t}\bar{t}$, $t\bar{t}t$, or $t\bar{t}$ (X).

Normalization layers are applied throughout the network to stabilise training and improve gradient flow. In addition, residual connections are introduced around the GATv2Conv layers to facilitate information propagation and mitigate the risk of vanishing gradients. Dropout layers are applied after each GATv2Conv block. During training, dropout randomly sets 10% of the node features to zero, preventing over-reliance on individual inputs and improving the generalisation performance of the model.

A schematic overview of the full architecture is shown in Figure 6.13.

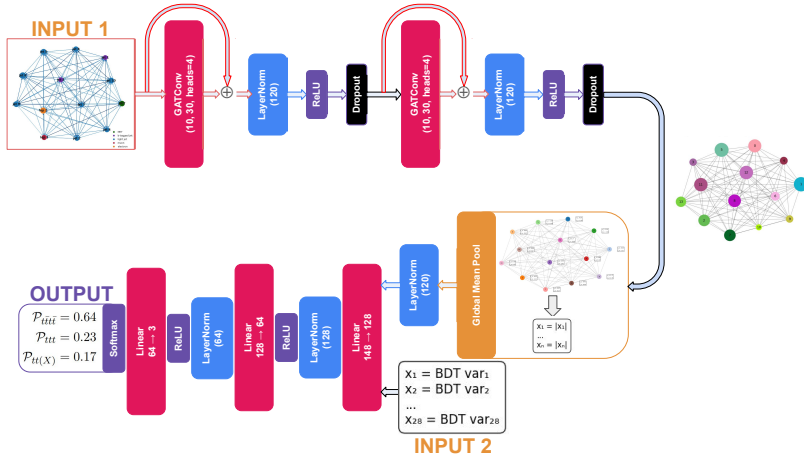


Figure 6.13.: Structure of the Event level neural network architecture.

6.3.2. Training

Currently, the model is trained separately in the SR 2ℓ and SR 3ℓ signal regions, and independently for the RUN 2 and RUN 3 data-taking campaigns. The network is trained to distinguish the following three classes:

1. $t\bar{t}t\bar{t}$
2. $t\bar{t}t$ ($t\bar{t}tj$ and $t\bar{t}tW$)
3. $t\bar{t}(X)$ ($t\bar{t}$, $t\bar{t}W$, $t\bar{t}Z$ and $t\bar{t}H$)

In early iterations, the model was found to suffer from limited training statistics. To mitigate this, the lepton selection and jet multiplicity requirements were slightly relaxed. Similar considerations are planned for future analysis selections, as the acceptance for $t\bar{t}t$ events is otherwise too low in the current fiducial phase space. These adjustments ensure that the model is trained on a dataset that better reflects the kinematic phase space relevant to the analysis.

The available statistics can be further increased by training the model simultaneously across both signal regions and all data-taking campaigns. While the classification tasks are not fully identical between regions, similar intermediate representations are expected to be learned, in particular in the attention layers. A combined training approach is therefore expected to provide a modest improvement in performance through increased effective statistics. These trainings are currently not being performed as they are computationally too expensive at this stage of the analysis.

The employed loss function is a weighted cross-entropy loss

$$\mathcal{L}(p) = -\left(w_{t\bar{t}t\bar{t}}y_{t\bar{t}t\bar{t}}\log(p_{t\bar{t}t\bar{t}})+w_{t\bar{t}t}y_{t\bar{t}t}\log(p_{t\bar{t}t})+w_{\text{Bgd}}y_{\text{Bgd}}\log(p_{\text{Bgd}})\right). \quad (6.14)$$

Each class is assigned a weight reflecting its physical composition, while an additional reweighting is applied such that all classes contribute equally to the total loss. This procedure mitigates biases arising from class imbalances.

At the start of training, all network weights are initialized with random values. During each training step, the loss function is evaluated on a batch of events. The network then computes the gradient of the loss with respect to every trainable parameter through the backpropagation algorithm. Using this information, the optimizer updates the weights in the direction that reduces the loss. Repeating this procedure over many iterations gradually improves the network performance. This optimization procedure is generally referred to as gradient descent.

During the early stages of training, relatively large weight updates are beneficial, as the network is still far from an optimal solution and rapid progress can be made. As training progresses, however, the network begins to learn increasingly subtle correlations in the data. Large parameter updates at this stage can become detrimental, causing the optimizer to overshoot minima in the loss landscape or erase previously learned features. It is therefore advantageous to gradually reduce the learning rate during training, allowing the optimization to transition from a coarse exploration of the parameter space to a more precise refinement of the learned representations. After a sufficient number of epochs, the model reaches the limit of the information that can be learned from the training dataset. Further minimization of the loss then begins to fit statistical fluctuations and peculiarities of the training sample rather than genuine underlying features of the data. As a result, the training performance continues to improve while the performance on an independent validation dataset stagnates or even deteriorates. This phenomenon is known as overtraining (or overfitting). If left unchecked, overtraining can lead to degraded performance when the model is applied to unseen data and may

introduce biases in the analysis. Training is therefore typically stopped once the validation performance ceases to improve.

It is important to note that a difference between the performance on the training and validation datasets does not, by itself, indicate overtraining. A model is expected to perform somewhat better on the data it was trained on. Overtraining occurs only when continued improvements on the training sample are accompanied by a deterioration of the performance on independent datasets. As long as the model retains its predictive power on unseen data and the training sample is not used in the final analysis, a moderate gap between training and validation performance is entirely normal.

Together with the normalization, residual, and dropout layers discussed previously, this improves both the stability of the training procedure and the generalisation performance of the final model.

During training, the generalisation performance of the network was found to depend strongly on the batch size. Smaller batch sizes yielded noticeably improved performance, as illustrated in Figure 6.14. To accommodate this behaviour, the original batch normalisation layers were replaced by Layer Normalisation layers, which do not rely on batch-level statistics and instead normalise activations within each event in latent space.

In addition, the batch size is initially chosen to be relatively small and is gradually increased during training as a form of implicit regularisation. Motivated by the observations of Ref. [104], a coupled schedule of batch size and learning rate is employed. The implementation follows Refs. [105, 106], where the batch size is doubled after a fixed number of training steps N_{steps} is reached. After this optimization, the regularization grid that was optimized, with the grid search consisting of the following parameters:

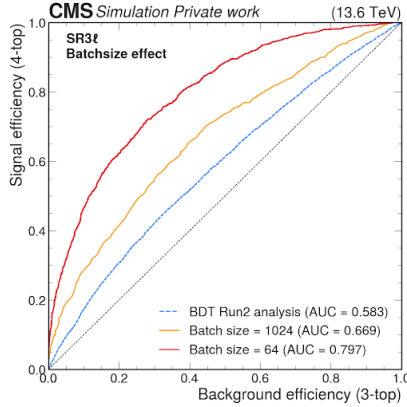


Figure 6.14.: ROC curves for the GNN trained with batch sizes 64 and batch size 1024. The RUN 2 BDT model is shown as reference.

- Training weights = [Absolute value of MC weight, Refiner weights, Reweighter weights]
- learning rate = [0.00005, 0.0005, 0.005]
- epochs = [200 (with early stopping)]
- batchsizes = [8, 16, 32, 64, 128]
- $N_{\text{steps}} = [5, 10, 30, 50, 201]$

Best results were found for very small starting batch values (8), with intermediate increase throughout the training ($N_{\text{steps}} = 30$). A too low increase would lead to fast overfitting after only $\mathcal{O}(30)$ epochs with a slight performance decrease. The hyperparameter search leads to similar optimal hyperparameters for the trainings in the dilepton, trilepton and RUN 2 and RUN 3 trainings. Dilepton trainings however reach their minimum at earlier epochs ($\mathcal{O}(50)$) and are found to be more volatile, even at low learning rates. An optimal learning curve is shown for the SR3 ℓ training with RUN 3 samples in Figure 6.16. The corresponding curve

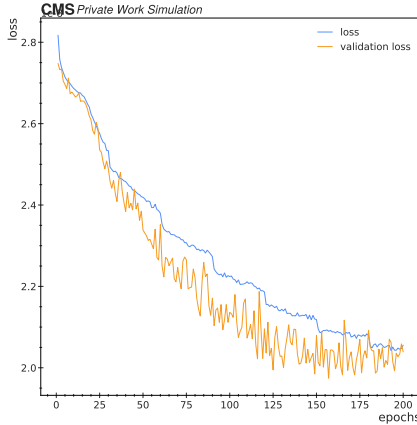


Figure 6.15.: Loss curve throughout the training of the SR3 ℓ region with RUN 3 samples.

of the SR2 ℓ training is shown in Figure 6.15. The model with the best performance on the validation set is used in the studies below.

6.3.3. Performance

The output scores of both networks are shown in Figure 6.17. The obtained results exceed initial expectations. The model achieves strong discrimination between the $t\bar{t}t\bar{t}$ signal and the $t\bar{t}(X)$ backgrounds, which is slightly better than what was obtained by the previous model. At the same time, the model is providing excellent separation between the $t\bar{t}t\bar{t}$ and ttt processes despite their very similar final-state topologies. The SR2 ℓ model achieves an area-under-the-curve (AUC) value of 0.86, while the SR3 ℓ model reaches an AUC of 0.83.

Interestingly, the network exhibits a comparable ability to distinguish both $tttj$ and $tttW$ events from the $t\bar{t}t\bar{t}$ signal, as shown in

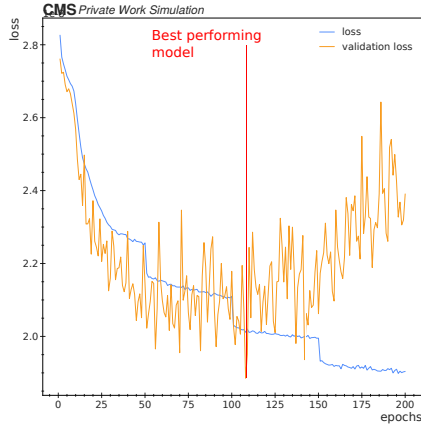


Figure 6.16.: Loss curve throughout the training of the $\text{SR}2\ell$ region with RUN 3 samples. The best performance model is studied further below.

Figure 6.18. Given the current theoretical uncertainties associated with the separation and interference of the $t\bar{t}j$ and $t\bar{t}W$ processes, no attempt was made to further distinguish these contributions from one another. Instead, they are treated as a single $t\bar{t}t$ category throughout this study.

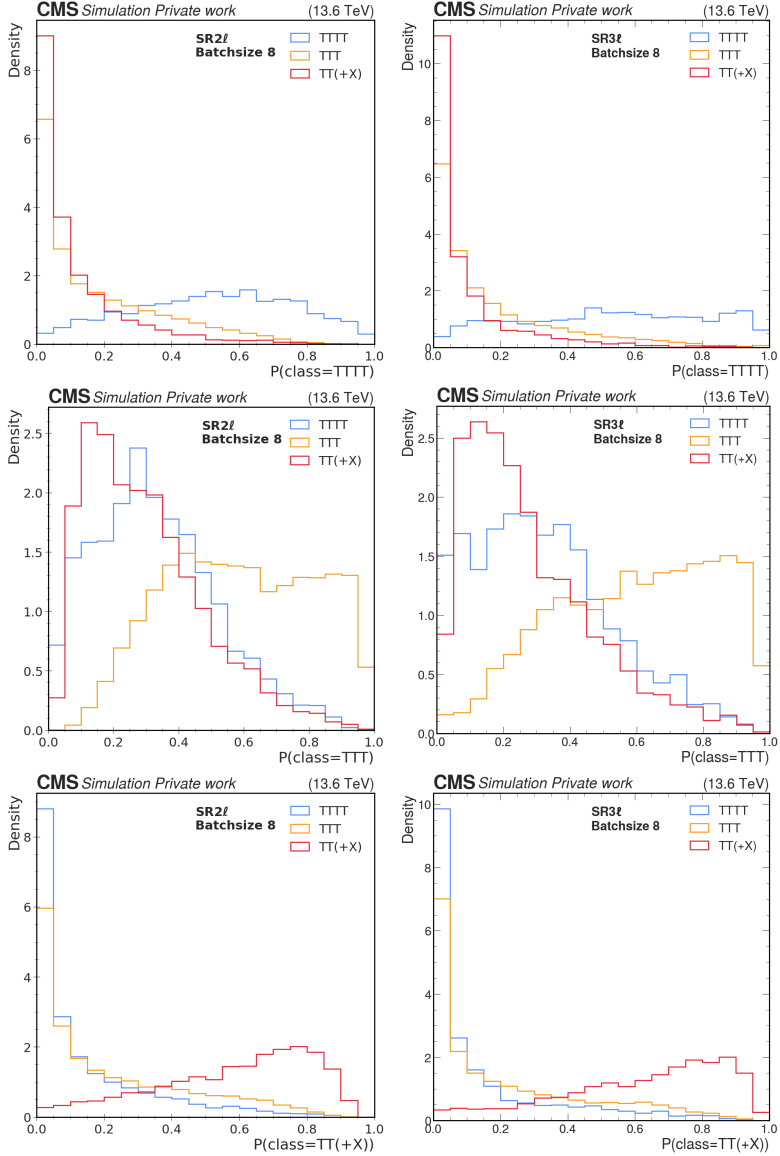


Figure 6.17.: Output scores of the most performant models in the SR2 ℓ (left) and SR3 ℓ (right). The output nodes target the $t\bar{t}t\bar{t}$ process, $t\bar{t}t$ events and $t\bar{t}(X)$ back-grounds.

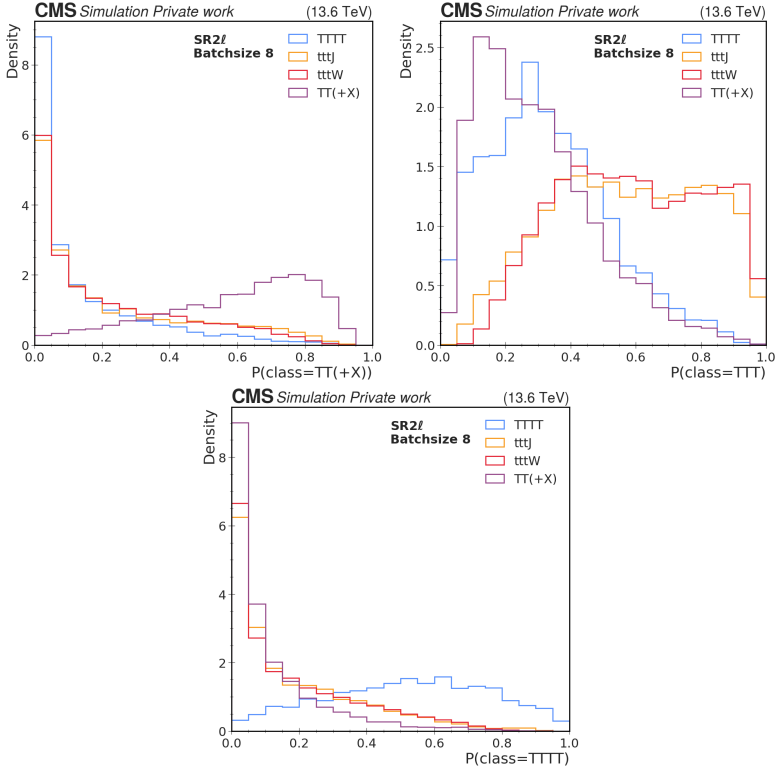


Figure 6.18.: Output scores of the most performant models with the ttt process split in the $ttjt$ and ttW components.

6.3.4. Conclusions and outlook

The successful demonstration that substantial separation can be achieved between the DR1-defined ttt processes and the $t\bar{t}t$ process indicates that a simultaneous experimental measurement of both contributions may be feasible, provided that theoretically robust and well-defined process definitions can be established.

The performance achieved in this work represents a significant improvement over previous approaches. For example, the first

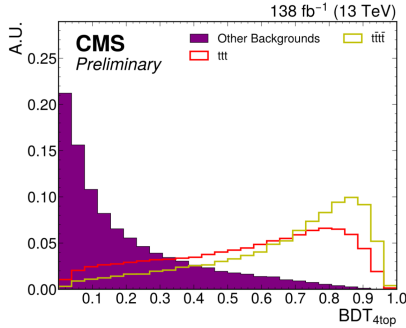


Figure 6.19.: State-of-the-art performance before these studies.
Figure from Ref. [107].

CMS search for ttt production was unable to achieve comparable discrimination with a single classifier and instead relied on two separate models: one trained to distinguish ttt from background processes and another trained to separate $tt\bar{t}\bar{t}$ from backgrounds [107].

The combined score space nevertheless remains significantly correlated between the ttt and $tt\bar{t}\bar{t}$ processes, as illustrated in Figure 6.19. The model developed in this thesis substantially improves upon previous strategies and opens promising opportunities for future measurements targeting these rare multi-top production processes.

6.3.5. Improved predictions of ttt production

While these results show incredible promise for what can be achieved in future analyses, the final hurdle has yet to be reached. Recent theoretical work, performed by G. Durieux et al., has shown that the interference term between $tttW$ and $tt\bar{t}\bar{t}$ is very large. These observations made the use of a DR1 sample ill-motivated and necessitate the use of DR2 samples unless the interference terms were found to be flat over the full phase space.

Further studies by G. Durieux et al., presented in the LHCTop working group [108], have however shown that both DR1 and DR2 implementations violate gauge invariance and are non-unitary. This means that ttt and $t\bar{t}t\bar{t}$ are ill-defined processes. Only the sum of both parts is a gauge invariant, unitary quantity that is well described.

These observations require a totally new approach to dealing with these processes. A new model will soon be made available by G. Durieux et al. that will be able to describe the full $ttt + t\bar{t}t\bar{t}$ phase space. These models will be implemented within the analysis. At that point, particle level definitions of both processes will have to be defined. This could for instance be based on the requirement of all top systems to be in a window around the top mass resonance, or could be based on the emission of an extra b-quark from the $tttW$ process.

Once these processes are well defined, the architecture developed in this thesis can be applied to these new process definitions. This would allow for a measurement that is able to shed light on the important questions that were raised in section 3.4.5.

7 Oxygen collisions program at CMS

As top quark physicists might sometimes forget, the LHC has a much richer physics program than only pp collisions. Already at the end of the first year of operation, the first lead-lead ($PbPb$) collisions were recorded in the LHC experiments. These were the first collisions of lead ions at the TeV scale and opened a brand new research program at CERN. In 2017, the LHC performed a brief run colliding Xenon nuclei, allowing scientists to study heavy-ion collisions with a different system size than $PbPb$ collisions. In July 2025, an exceptionally ambitious ten-day program was conducted at the LHC, featuring two days of unprecedented proton–oxygen collisions, followed by additional firsts: two days of oxygen–oxygen collisions at energies roughly 55 times higher than previously achieved at the RHIC facility and one day of first-ever recorded neon–neon collisions, with several days of machine setup in between.

The primary motivation for this campaign was to improve hadronic interaction models relevant for cosmic-ray physics—an area explicitly included in the CERN Convention, which mandates research on high-energy particles “including work in the field of cosmic rays.” [109] In section 7.1, the shortcomings of current hadronic models are highlighted. Then, the 10 day program is discussed in section 7.2 and the luminosity calibrations performed by the author are discussed in section 7.3. A final section, section 7.5 is dedicated to the ongoing and planned future works of the author on this subject.

7.1. Current hadronic models in the field of cosmic ray physics

Cosmic rays (CRs) span an enormous energy range, extending up to approximately 10^{21} eV. Interactions of such particles with the Earth's atmosphere correspond to center-of-mass energies far beyond those accessible at current or foreseeable human-made accelerators. The origin and composition of ultra-high-energy cosmic rays (UHECRs), typically defined as CRs with energies above 10^{18} eV, remain among the central open questions in astroparticle physics.

The energy and composition of these particles can be inferred from the extensive air showers (EAS) produced when they interact with nuclei in the upper atmosphere. A precise determination of these properties would provide crucial information about the sources and acceleration mechanisms of the most energetic particles in the universe. However, the field is currently limited by significant uncertainties in the theoretical modelling of EAS development, preventing cosmic ray physics from reaching the level of precision achieved in other areas of high-energy physics.

This limitation is illustrated in Figure 7.1, which shows the average logarithmic mass of cosmic rays as a function of their energy, as measured by several experiments. Different experiments rely on different EAS observables to estimate the primary cosmic-ray mass. For example, the Pierre Auger collaboration uses the atmospheric depth at which the shower reaches its maximum development for fluorescence light (X_{max}) as a primary mass-sensitive observable. Experiments such as IceTop instead rely more strongly on measurements of the muonic component of the air shower at ground level.

In addition, the predicted observables depend significantly on the choice of hadronic interaction model [110]. This dependence

is visible, for example, in the differently coloured data points corresponding to the same Auger measurements on the right side of Figure 7.1.

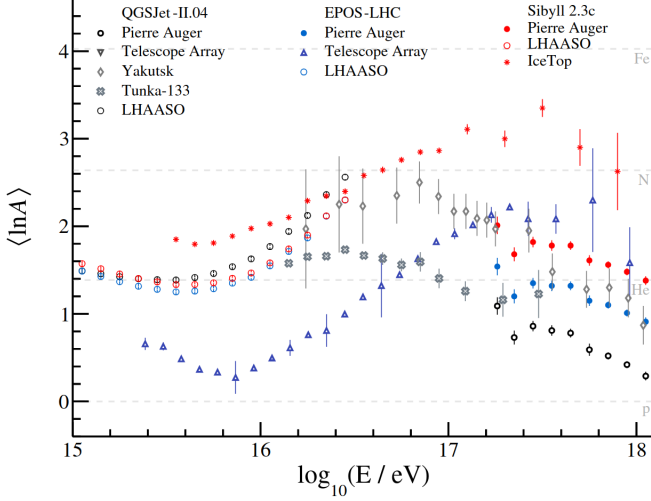


Figure 7.1.: The measured mean-logarithmic mass of cosmic rays as a function of their energy. Measurements that use a different mass-sensitive variable in the analysis (variables can roughly be grouped in X_{max} or N_μ) get vastly different results. [111]

Currently, these models are constrained primarily using proton–proton collision data collected at the LHC, supplemented in some cases by heavy-ion collision data, such as lead collisions used in the tuning of EPOS.LHC-R [112]. However, applying these constraints to atmospheric interactions requires substantial extrapolations in both collision energy and nuclear mass. As a consequence, the resulting hadronic interaction models carry large systematic uncertainties, which still fail to fully cover the large discrepancies observed between experimental measurements.

These observations clearly demonstrate the need for additional LHC data to constrain hadronic interaction models in a more robust and physically motivated manner. Improved experimental constraints would reduce the associated model uncertainties and help resolve the tensions currently observed between different cosmic-ray experiments. As can be seen in figure 7.2, proton–oxygen and oxygen–oxygen collisions would be perfectly situated to describe cosmic ray–air interactions. Proton–proton and lead collisions are however clearly far from realistic cosmic ray–air interactions.

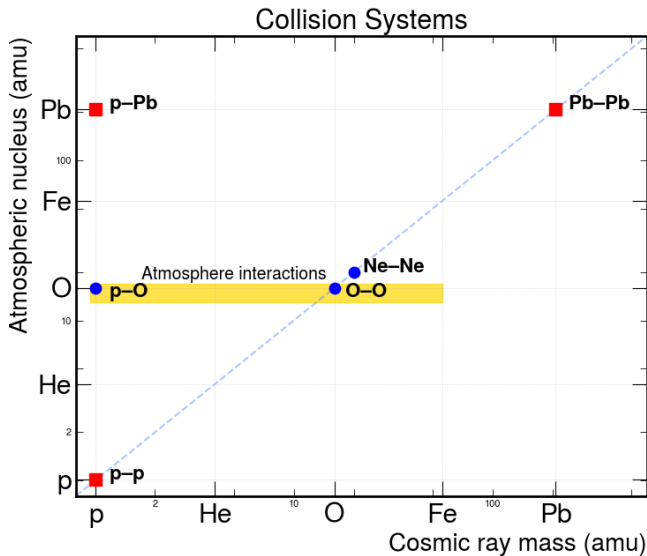


Figure 7.2.: Collision systems currently used to constrain hadronic interaction models at the LHC (red), together with the new collision systems collected in July 2025 (blue). The yellow band indicates the collision systems most relevant for cosmic-ray interactions in the atmosphere.

7.2. Data taking: the pO, OO and NeNe programs

Motivated by this need for improved experimental constraints on hadronic interaction models, an exceptionally ambitious ten-day data-taking campaign was conducted at the LHC in July 2025. The program featured two days of unprecedented proton–oxygen collisions, followed by two days of oxygen–oxygen collisions at center-of-mass energies approximately 55 times higher than those previously achieved at the RHIC facility, and concluded with the first-ever recorded neon–neon collisions. Several dedicated machine-development and setup periods were required between the different collision systems.

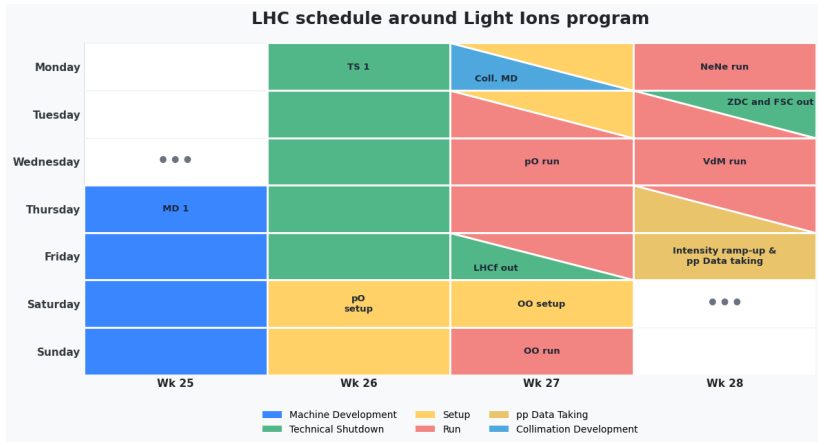


Figure 7.3.: LHC schedule around the light ions program.

The idea for this special run was first proposed in 2018–2019. After several postponements, detailed feasibility studies and preparations began in April 2025. Beyond its relevance for cosmic-ray physics, the program also opened new opportunities for studies of QCD in small and intermediate collision systems, as well as investigations of collective effects and particle production in previously unexplored nuclear environments. While numerous people within

the collaboration are interested in exploring these intermediate systems, few are exploring the crucial physics for the cosmic ray community. The contributions of the author in this program are focussed in this direction.

To enable these cosmic-ray measurements, the installation and operation of several specialised detector systems were essential. The combination of the CMS–TOTEM Precision Proton Spectrometer (PPS), the Zero Degree Calorimeter (ZDC), and the Forward Shower Counters (FSC), together with the excellent particle-identification capabilities of CMS in the central detector region, places CMS in a uniquely strong position to deliver a broad range of complementary and high-impact measurements from these data.

The author contributed to the installation of the FSC detectors and participated in the preparations for this data-taking campaign. In particular, the luminosity requirements for the ultra-low-pileup ZeroBias dataset were estimated by the author. This dataset will play a crucial role in the studies discussed in Section 7.6, which constitute the primary physics use case for these data. Once these luminosity targets were reached, pileup interactions were raised from the ultra-low 0.1 interactions per crossing to a handful of interactions per crossing. This increase in instantaneous luminosity was accompanied by a much harder ZeroBias prescale and more MinBias datastreams. Due to the excellent work by collaborators at the LHC and CMS, we were able to collect much larger datasets than were originally targeted. The target and online calibrated collected luminosity estimates and other crucial machine parameters are shown in table 7.1.

Table 7.1.: Target and collected (online calibration) luminosities and main beam parameters.

System	Target (nb ⁻¹)	Delivered (nb ⁻¹)	Ratio
pO	3	48.4	16
OO	0.8	9.4	11.8
NeNe	0.1	0.91	9.1

System	Parameter	Injection	Flat top	Physics
OO/NeNe	Energy (Z TeV)	0.45	5.36	5.36
	β^* (m)	11	1.0	0.5
	Half ext. crossing (μ rad)	170	100	100
	Net half crossing (μ rad)	100	100	100
pO	Energy (Z TeV)	0.45	6.8	6.8
	β^* (m)	11	1.0	1.0
	Half ext. crossing (μ rad)	170	60	60
	Net half crossing (μ rad)	60	60	60

7.3. Luminosity callibrations for the OO and NeNe program

The work presented in this chapter was performed in the two months following the successful data-taking, allowing CMS to present results at IS2025. For this timeline, only analyses using OO and NeNe data were targeted. The luminosity callibrations for the pO dataset (in combination with the final OO and NeNe callibration) are being performed on a longer timeline. Some ongoing work for these datasets is discussed in the next section. The OO program included two VdM programs in Fill 10802 and Fill 10808, while the NeNe program included a short VdM program in Fill 10813. The scans included in these programs are shown in table 7.2. CMS is the only LHC collaboration that performed a VdM program in the NeNe program, giving us the opportunity to provide very precise measurements that would otherwise be constrained by the luminosity uncertainty. Each scan included in these programs will be discussed in the following subsections.

Table 7.2.: Overview of the scans for Fills 10802, 10808 and 10813.

Activity	Start time	End time
Fill 10802		
Optimisation	07:40	07:48
vdm xy - standard	07:48	08:20
Optimisation	08:20	08:23
diag $\pm 45^\circ$	08:23	08:55
Optimisation	08:55	08:58
vdm xy - standard	08:58	09:29
super separation	09:30	09:36
head-on cLS (20s version)	09:37	09:53
Optimisation	09:53	09:56
Fast scan pair	09:56	10:09
Fill 10808		
vdm xy - standard	11:49	12:21
Optimisation	12:21	12:25
diag $\pm 45^\circ$	12:25	12:50
Optimisation	12:50	12:54
vdm xy - standard	12:54	13:20
Optimisation	13:20	13:23
diag $\pm 45^\circ$	13:23	13:48
Optimisation	13:48	13:51
vdm xy - standard	13:51	14:17
super separation	14:17	14:23
Optimisation	14:23	14:25
head-on cLS (20s version)	14:25	14:42
Fill 10813		
vdm xy - Fast for NeNe	18:00	18:25
head-on CLS (20s version)	18:29	18:45
Optimisation	21:36	21:38
vdm xy	21:38	22:04
Optimisation	22:04	22:07
diag $\pm 45^\circ$	22:07	22:32
Optimisation	22:32	22:35
vdm xy	22:35	23:01
Optimisation	23:01	23:05
diag $\pm 45^\circ$	23:05	23:29
super separation	23:32	23:37

7.3.1. The Van der Meer scan

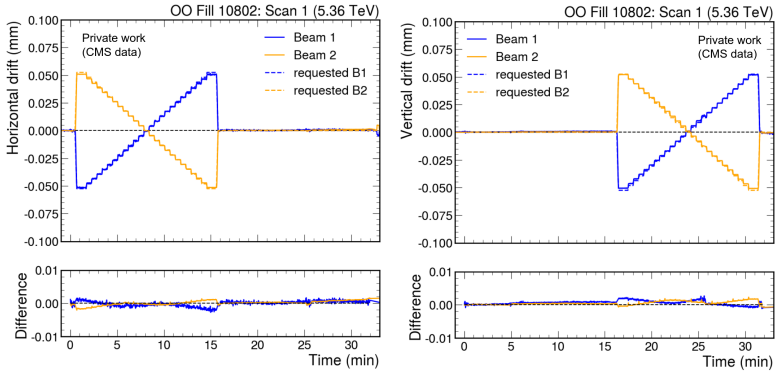


Figure 7.4.: Beam positions measured by DOROS during the first Van der Meer scan of Fill 10802. The horizontal (x) and vertical (y) scan coordinates are shown in the left and right panels, respectively.

The most important scan of any VdM program are the VdM scans. As described in section 2.4, these scans are used to integrate the rate of the luminometers over the beam separations. This is used to find the effective widths of the beam overlap area along the x and y directions. After this, we can use:

$$\sigma_{vis} = \frac{2\pi\Sigma_x\Sigma_y\mathcal{R}(0,0)}{\nu N_1 N_2}, \quad (7.1)$$

which was derived in section 2.4, to find the fiducial cross section of the luminometer. Before this integration can be performed, detailed studies need to be performed to understand all systematic uncertainties of this procedure.

7.3.2. BSRL: Ghosts and Satellites

The BSRL, introduced in section 4.1.3, is used to measure the ghost and satellite fractions f_g, f_{sat}^j in the LHC beams. These

components of the beam do not collide at the IP and are subtracted from N_1 and N_2 following equation 4.2.

To measure these fractions, the raw BSRL data is binned in 50 ps beam slots and integrated over 5 minutes. As electronic noise is not synchronized with the LHC RF timing, this forms a flat background under the signal. The signals are calibrated by removing this pedestal and are then grouped in 2.5 ns RF buckets. The pedestal subtracted BSRL data for one BCID in Fill 10802 and Fill 10813 are shown in figure 7.5.

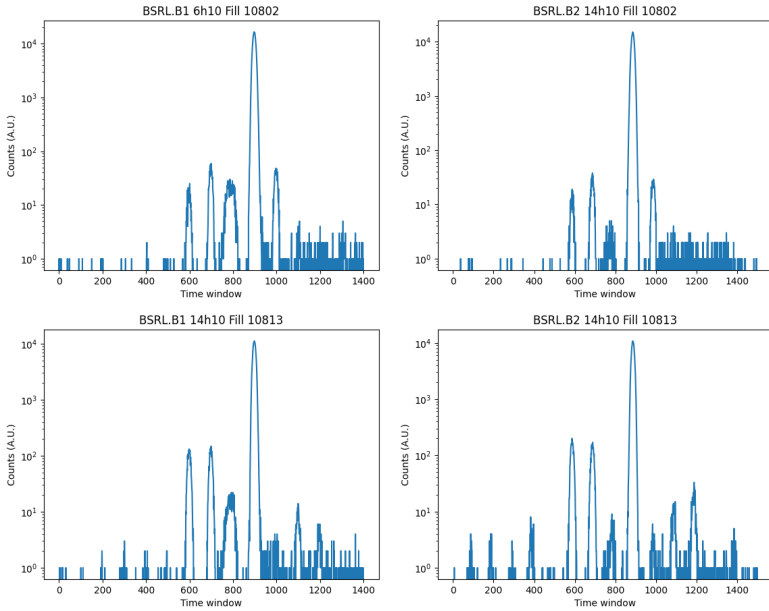


Figure 7.5.: Pedestal subtracted BSRL data for BCID 1786 in Fill 10802 and BCID 1 in Fill 10813.

The satellite fractions are then estimated in each BCID, while the ghost fraction is a global value. The ghost and satellite contributions without pedestal removal are calculated as a systematic uncertainty, which results in an uncertainty of roughly 0.3% on

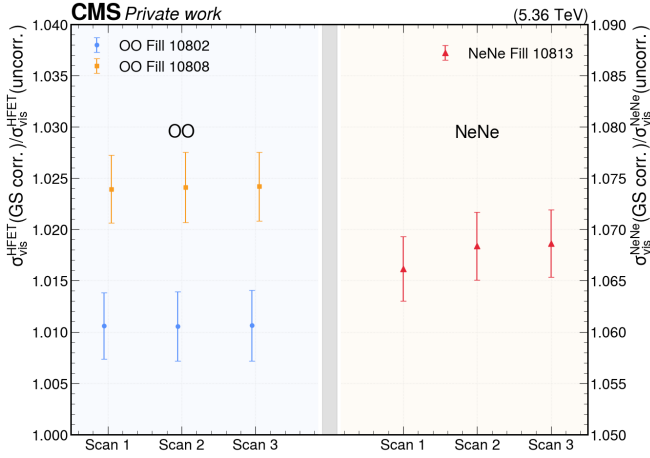


Figure 7.6.: Effect of ghost and satellite corrections on the visible cross section of the HFET luminometer for the OO and NeNe fills.

the measured visible cross section. The ghost and satellite fractions evolve over time and can differ significantly between fills. As summarized in figure 7.6, ghost and satellite corrections have a 1-2.5% effect on the cross section for OO, while the NeNe fill had much larger ghost and satellite fractions with an effect of 6.7% on the visible cross section. These corrections are not dependent on the luminometer as they just change the N terms in equation 7.1. These satellite fractions are sizeable compared to what can be achieved in better-understood collision systems such as Pb-Pb or p-p collisions. To mitigate the effect of these satellite fractions, a non-zero crossing angle was adopted at all detector IPs. A more detailed study of satellite bunches is being performed for the pO VdM program and final calibrations of OO and NeNe. This is discussed in a bit more detail in section 7.4.1.

7.3.3. The superseparation scan

Based on the BSRL measurements, the beam currents are well constrained. However, several additional systematic effects still impact the determination of the beam widths and the head-on interaction rate. Firstly, the luminometers used to measure the interaction rates are themselves subject to electronic and detector noise. This noise can exhibit a non-trivial dependence on the presence of beams and therefore does not cancel trivially in standard subtraction procedures.

To study these noise components, dedicated superseparation scans are performed, in which the two beams are separated by distances well beyond the typical beam width. In this regime, any remaining coincidence rate is dominated by uncorrelated background contributions and luminometer noise rather than true beam–beam interactions. These measurements allow for a more robust subtraction of non-collision backgrounds in the determination of the visible rate.

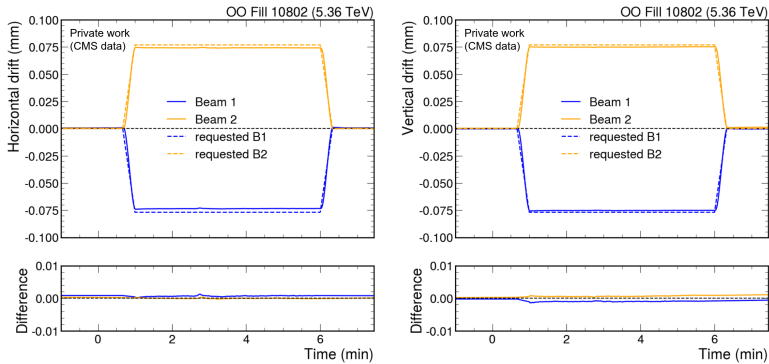


Figure 7.7.: Beam positions measured by DOROS during the superseparation scan of Fill 10802. The horizontal (x) and vertical (y) scan coordinates are shown in the left and right panels, respectively.

Superseparation scan noise measurements are shown in figure 7.8 for the BCM1F luminometer. These backgrounds are further studied in the tails of the VdM scans, empty BCIDs, single filled BCIDs and BCIDs just after a colliding BCID. These studies can help factorize background components to individual components such as afterglow or backgrounds related to one beam. Overall, the online noise reduction of HF was found to preform well, while sometimes slightly overestimated. The PLT and BCM1F luminometers were found to have extremely low backgrounds, in agreement with previous observations. The application of superseparation background corrections showed an slightly improved agreement between BCIDs for the $\Sigma_x \Sigma_y$ product and resulted in a visible cross section that is 0.6% lower. This modification is added as a systematic uncertainty.

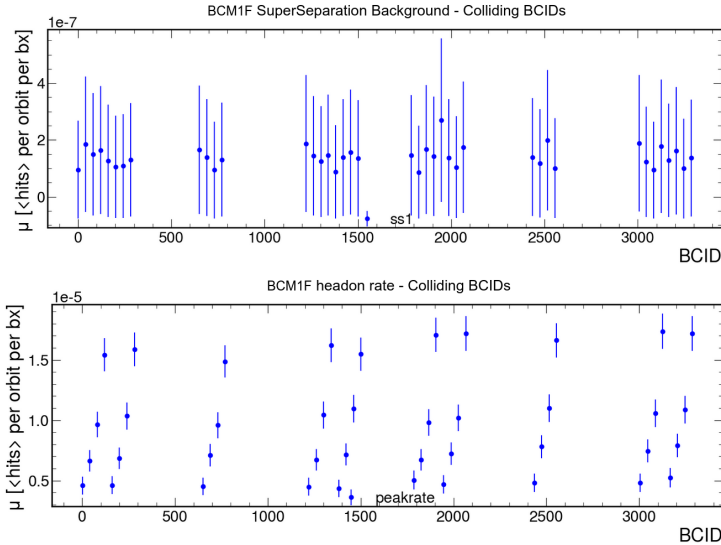


Figure 7.8.: Top: observed rates during the superseparation scan.
Bottom: observed rates during a headon period in the VdM program.

7.3.4. Beam–Beam effects

The oxygen beams have beam charge numbers of the order of 13×10^{11} . This corresponds to approximately 2.7×10^9 fully ionized oxygen atoms per bunch. When two bunches pass by each other at an interaction point, the electromagnetic fields generated by the large bunch charges induce sizeable beam–beam forces. These forces lead to two important effects that must be considered in precision luminosity calibrations: beam–beam deflection and the dynamic β effect.

Beam–beam deflection arises from the electromagnetic repulsion between the two colliding bunches. During a Van der Meer scan, the bunches are intentionally separated in the transverse plane. At non-zero separation, the electromagnetic force introduces an additional transverse kick, causing the closed orbit of each beam to shift. As a result, the true beam separation differs from the nominal separation inferred from the magnet settings. If left uncorrected, this effect biases the determination of the beam overlap and therefore the extracted visible cross section.

For Gaussian beams, the transverse kick can be calculated using the Bassetti–Erskine formalism [113]. The kick experienced by a particle at transverse position (x, y) is given by:

$$\Delta x' + i\Delta y' = \frac{2Nr_0}{\gamma} \frac{i\sqrt{\pi}}{\sqrt{2(\sigma_x^2 - \sigma_y^2)}} \left[w\left(\frac{x + iy}{\sqrt{2(\sigma_x^2 - \sigma_y^2)}}\right) - \exp\left(-\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2}\right) w\left(\frac{x\sigma_y/\sigma_x + iy\sigma_x/\sigma_y}{\sqrt{2(\sigma_x^2 - \sigma_y^2)}}\right) \right]. \quad (7.2)$$

where N is the opposing bunch population, r_0 is the classical particle radius, γ is the energy divided by mass, σ_x and σ_y are the transverse beam sizes, and $w(z)$ denotes the complex error

function. When adapting this expression to oxygen beams, the formal structure of the equation remains unchanged, but the beam parameters entering the prefactor must be modified to account for the ion charge and mass. In particular, γ is still defined as $E_{\text{ion}}/\text{mass}$ and $r_{\text{ion}} = \frac{(Ze)^2}{4\pi\epsilon_0 m_{\text{ion}} c^2} = \frac{Z^2 m_p}{m_{\text{ion}}} r_p$.

The resulting orbit distortion modifies the effective beam separation during the scan. While most parameters are known well, the transverse beam sizes are extrapolated from the measured $\Sigma_{x/y}$ values, assuming identical beam profiles. Following this assumption, the correction is evaluated and shown in figure 7.9. The calculated beam-beam deflection effects are extremely low, even lower than what was calculated in [114] for proton – lead data taking.

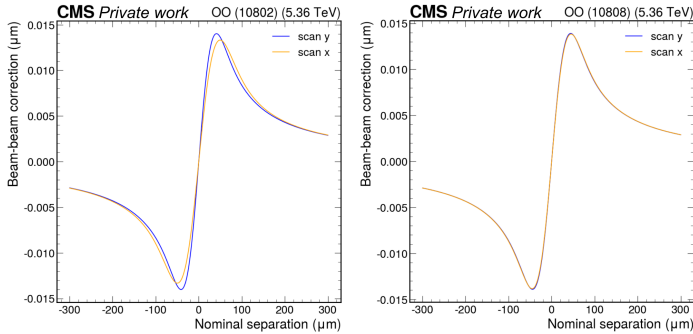


Figure 7.9.: Corrections to the beam-beam displacement due to beam-beam deflection for Fill 10802 (left) and Fill 10808 (right).

A second consequence of the beam–beam interaction is the dynamic β effect. The electromagnetic field generated by one bunch acts as an additional lens on the particles of the opposing bunch. This modifies the optical functions of the machine and leads to a shift of the effective β^* at the interaction point.

A precise evaluation of the dynamic β effect requires detailed knowledge of the transverse beam distributions. Given the limited set of available scan configurations and the absence of dedicated destructive beam measurements, no explicit calculation of the dynamic β correction was performed for the fast analyses presented in this thesis. Since the dynamic β effect is further expected to be anticorrelated with the beam–beam deflection and of comparable or smaller magnitude, a conservative combined uncertainty of 0.5% is assigned to cover both effects.

7.3.5. arcBPMs and DOROS

Over the course of a scan, beams might undergo small movements. If these movements are not correctly taken into account, the true beam separation would differ from the nominal separation requested via the magnet settings. If left uncorrected, this effect biases the determination of the beam overlap and therefore the extracted visible cross section. To partially correct for these movements, orbit drift corrections are applied. The positions of the beams are inferred using two systems: the arcBPMs and DOROS.

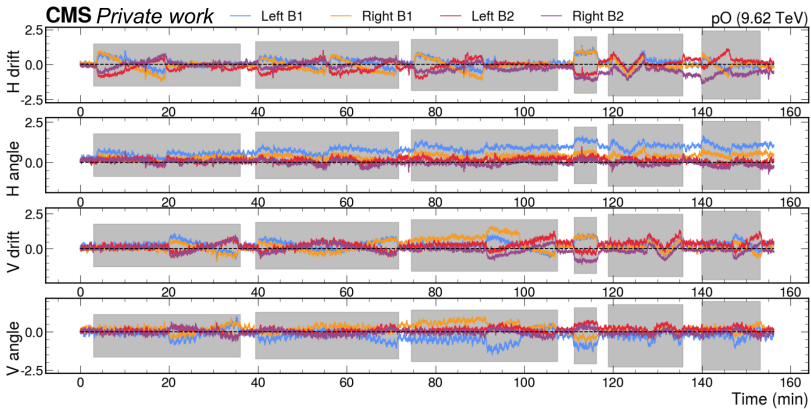


Figure 7.10.: Data as measured by the arcBPMs during the pO program.

For each system, the position of the beam is measured before and after the scan. The orbit drift during the scan is undone with a linear correction of the nominal positions. Horizontal measured orbit drift values are shown in figure 7.11 for all colliding systems.

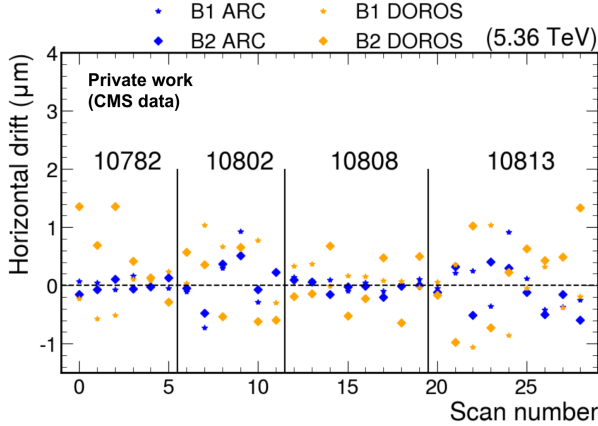


Figure 7.11.: Horizontal linear orbit drift values as measured by DOROS and arcBPMs.

As these values are small, no further investigations were performed for the fast callibrations and a conservative uncertainty of 1% is assigned for orbit drifts. For the final callibration, more detailed studies will be performed, including residual orbit drift corrections. These studies are however not performed by the author.

7.3.6. Other scans

As the timeline was extremely short, not all callibrations could be performed. Hence, two scan types are yet to be discussed: the diagonal scans and the constant length scale scans.

In section 2.4, it was assumed that the particle densities are factorizable, allowing the beam overlap to be parametrized as $2\pi\Sigma_x\Sigma_y$. This assumption is however not always well justified. Several

methods exist to measure the nonfactorization bias. However, the beam optics during the light ion runs did not allow for some of these methods as the event rates would be too low. In these fills, diagonal scans, as shown in figure 7.12, are used to estimate the nonfactorization bias. By scanning in the x+y and x-y directions, one can detect correlations between the horizontal and vertical beam distributions.

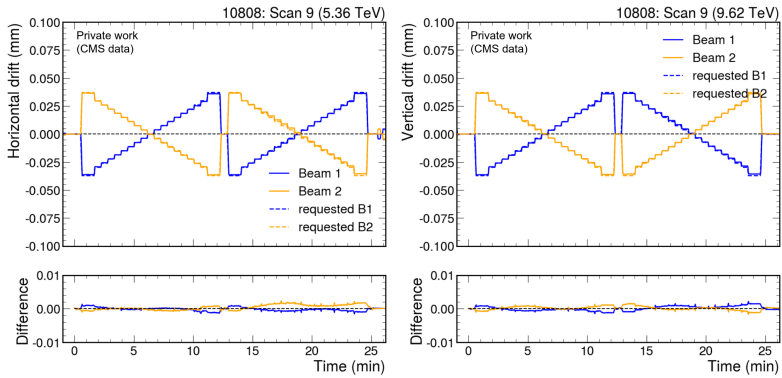


Figure 7.12.: Diagonal scan in the VdM program of Fill 10808.

The last scans to be discussed are the constant length scale scans. The beam movements in all scans are controlled by the magnet system. It is however uncertain if the magnet system moves the beams exactly as intended. To test the length scale of beam movements, a length scale analysis is performed. An example of a constant length scale scan is shown in figure 7.13. This analysis could not yet be performed for the fast calibration as the needed data was yet to be processed. The length scale analysis is discussed in section 7.4.

7.3.7. Consistency, stability and linearity

After the corrections described in the previous sections have been applied, several additional studies are performed to validate the robustness of the luminosity calibration. A first validation is

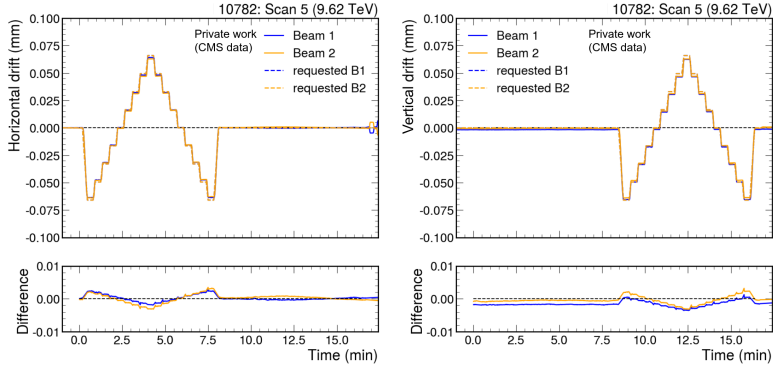


Figure 7.13.: cLS scan in the VdM program of Fill 10782.

performed by comparing the visible cross sections extracted from individual colliding bunch pairs. In the absence of unaccounted systematic effects, all bunches are expected to yield consistent calibration constants within their statistical uncertainties. Any residual spread beyond the expected statistical variation is interpreted as a measure of the bunch-to-bunch reproducibility of the calibration procedure. Systematic differences can originate from bunch characteristics that are not well understood. These systematic effects often show correlation with the injection scheme.

A second consistency check is obtained by comparing the visible cross sections extracted from the different Van der Meer scans performed during the calibration program. Since the same detector and beam conditions are probed multiple times, the resulting calibrations should be mutually consistent. Inconsistencies are an indication of changing beam conditions throughout the program.

The consistency between independent luminometers is also investigated. Several detector systems with different acceptances, technologies, and reconstruction techniques are used to measure the luminosity. Since these detectors are subject to different instrumental effects, a comparison of the instantaneous luminosities

obtained with each luminometer provides an important validation of the calibration procedure. In particular, the agreement between the HFET, PLT and BCM1F luminometers is studied in detail for this fast analysis.

Beyond the calibration scans themselves, the long-term stability of the luminosity measurement must be verified for these luminometers. The visible cross sections extracted during the Van der Meer program are propagated to the full data-taking period and the ratios between different luminometers are monitored as a function of time. Significant deviations from a constant response would indicate changes in detector performance, calibration drifts, or other time-dependent effects. The observed variations are used to derive the stability uncertainty associated with the luminosity measurement. In Fill 10805, PLT was found to be misbehaving and is excluded from this stability uncertainty estimation.

Finally, the linearity of the luminometer response is investigated. While the Van der Meer scans are performed with physics beam optics, later fills ran with higher average pileup interactions. The detector response must therefore remain proportional to the interaction rate across the full range of operating conditions. This is evaluated by comparing luminosity measurements from independent luminometers as a function of interaction rate and pileup. Any deviations from a linear response are quantified and propagated as a systematic uncertainty in the final luminosity calibration.

These systematic studies are summarized in Table 7.3 and form the basis of the luminosity uncertainty used in all fast analyses presented at IS2025. The author was solely responsible for the studies described in this section and for the resulting calibration.

The final luminosity calibrations for the full OO/NeNe program, as well as the corresponding calibration for the pO dataset, are

still ongoing and involve contributions from multiple collaborators. Consequently, the remainder of this chapter focuses exclusively on the author's contributions to these efforts.

Table 7.3.: Summary of luminosity calibration uncertainties for the OO and NeNe collision programs.

Source	OO (%)	NeNe (%)
Normalization		
Nonfactorization bias	3.0	3.0
Beam current calibration	0.3	0.3
Ghosts and satellites	0.3	0.3
Length scale	3.0	3.0
Orbit drift and residual beam position	1.0	1.0
Beam-beam and dynamic- β	0.5	0.5
Bunch-to-bunch variation	0.1	0.2
Scan-to-scan variation	1.7	0.7
Cross-detector consistency	0.4	0.3
Background correction	0.6	0.3
Integration		
Stability	0.4	0.0
Linearity	0.3	0.5
Afterglow	3.0	3.0
Total	5.7	5.4

7.4. *pO*, NeNe and futher OO calibrations

The *pO* VdM program is listed in Table 7.4. For this dataset, several studies analogous to those performed for the OO and NeNe program have already been carried out, including ghost and satellite corrections, beam-beam deflection calculations, and linear orbit drift studies. However, since the final calibrations are still being refined, this section focuses on more detailed investigations of ghost and satellite charge contributions for the *pO* run, as well as the constant length scale analysis applied across all datasets.

Table 7.4.: Overview of the pO Fill 10782 scan program.

Activity	Start time	End time
vdm xy - standard	16:40	17:12
Optimisation	17:12	17:16
diag $\pm 45^\circ$	17:16	17:47
Optimisation	17:47	17:51
vdm xy - standard	17:51	18:21
super separation	18:28	18:34
head-on cLS (20s version)	18:36	18:53
Optimisation	18:53	18:56
Fast scan pair	18:56	19:10

7.4.1. Ghosts and satellites

In the pO scans, an asymmetry in the measured rates between the two scan directions was observed. An example of this behaviour is shown in Figure 7.14, where a clear difference between the forward and backward scan profiles can be seen.

This asymmetry arises from differences in ghost and satellite charge populations between the proton and oxygen beams, combined with the fact that these contributions are not symmetrically distributed with respect to the filled bunch pattern and the non-zero crossing angle. As a result, the effective

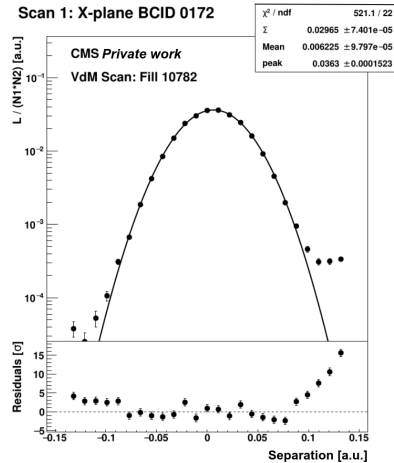


Figure 7.14.: Rate asymmetry observed in pO VdM scans for BCID 172.

overlap luminosity depends on the scan direction. While the non-zero crossing angle reduces the impact of ghosts and satellites on the head-on rate, it introduces a residual background asymmetry that cannot be fully addressed using conventional methods such as superseparation-based background estimates or scan-tail background subtraction techniques.

This effect is illustrated schematically in Figure 7.15. Early satellite bunches of beam 1 and late satellite bunches of beam 2 contribute to the asymmetry where $\text{position}_{B2} > \text{position}_{B1}$. The other satellite bunches contribute to the other asymmetry where $\text{position}_{B2} < \text{position}_{B1}$.

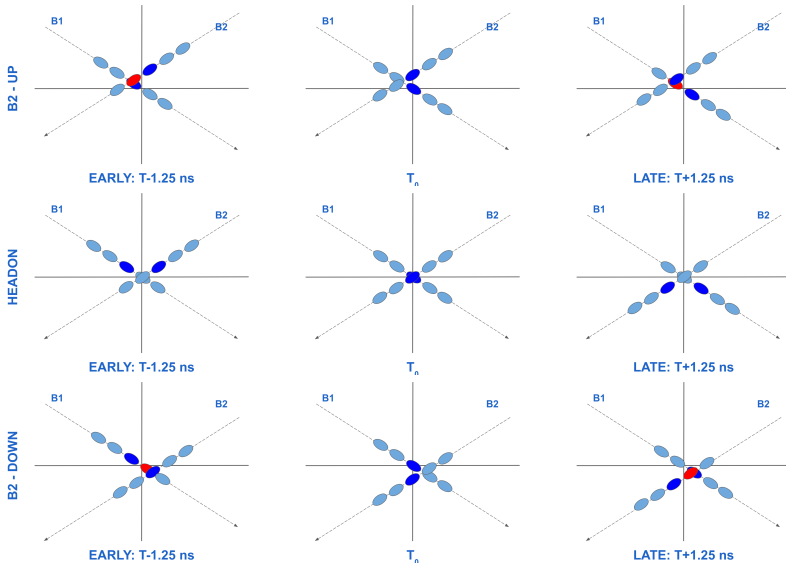


Figure 7.15.: Schematic of satellite bunch contributions to the asymmetry.

Looking at BSRL data, large satellites (ghosts) were observed in the oxygen beam at $n=N-4$ ($N-6$). This can be observed in

figure 7.16. Satellite contributions in the proton beam were not as sizeable.

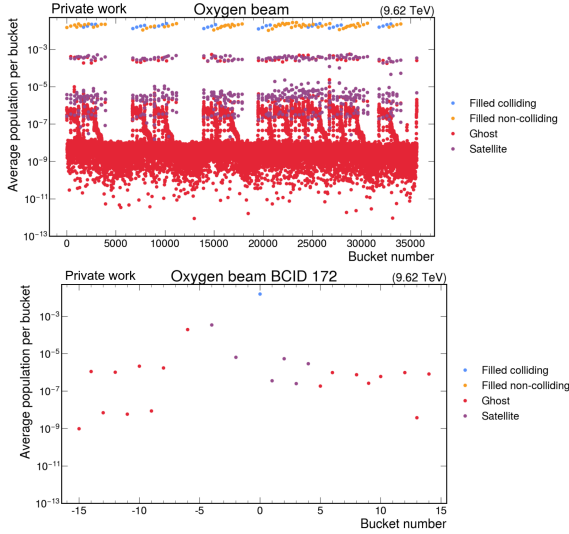


Figure 7.16.: Bucket populations for the full LHC RF system (top) and buckets surrounding BCID 172 (bottom).

With these findings, the correlation of the extracted visible cross section with this asymmetry is studied in more detail. In figure 7.17, a clear correlation can be observed between the satellite fractions and the observed asymmetry. This correlation can also be observed with the visible cross section.

Further studies aimed at improving the understanding of the systematic effects associated with satellite charge contributions are currently ongoing within the collaboration. One of the main challenges is the determination of the luminometer acceptance for satellite interactions. Since these collisions occur at time offsets of $n \times 1.25$ ns with respect to the nominal bunch crossing and are displaced by $\pm n \times 37.5$ cm from the nominal interaction point, their acceptance depends in a non-trivial manner on both the

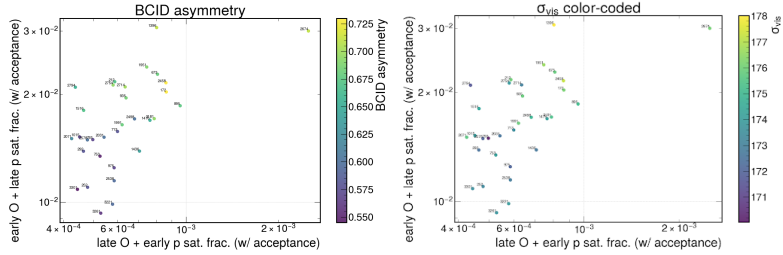


Figure 7.17.: Correlation between satellite contributions and the observed asymmetry (left) and extracted visible cross section (right).

detector geometry and timing response. The work discussed below was performed by Dr. P. Major.

As the interactions of these satellite interactions are significantly displaced, different luminometers exhibit different sensitivities to satellite collisions. Furthermore, for luminometers installed on both sides of the interaction point, the acceptance can differ significantly between the two detector arms. During these studies, it was observed that the PLT+ detector is largely insensitive to the satellite interactions, as the dominant satellite contributions originate from the oxygen beam and occur predominantly on the opposite side of the interaction point, as illustrated in Figure 7.18.

This observation allows the separate luminosity measurements from PLT+ and PLT- to be used as a probe of the residual impact of satellite charge contributions. The observed difference between the two measurements is used to estimate the associated systematic uncertainty. Until more detailed acceptance studies become available, a conservative uncertainty of 0.4% is assigned to cover this effect for PLT. For the other luminometers, such a study is not possible, as no equivalent ghost insensitive measurements are available. Their sensitivity to ghost-charge asymmetries must therefore be estimated indirectly through comparisons to PLT.

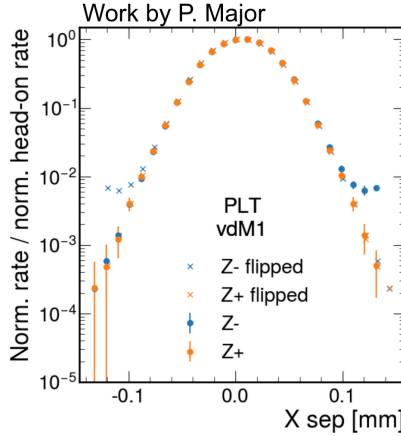


Figure 7.18.: PLT+ and PLT- rates in VdM scan of the pO program.

7.4.2. Length scale

As the length scale had not yet been evaluated for the fast calibrations of the OO and NeNe datasets, this study was performed simultaneously for all three collision systems. The length scale is defined through

$$x_{\text{true}} = \alpha_x x_{\text{requested}}, \quad (7.3)$$

with an analogous definition for the vertical direction. The goal of the analysis is therefore to determine the calibration factors α_x and α_y .

To measure these factors, the excellent track and vertex reconstruction capabilities of the CMS tracker are exploited. While the tracker cannot determine the positions of the two beams individually, it allows a precise reconstruction of the beamspot position, i.e. the centre of the luminous region, from the distribution of reconstructed primary vertices.

The calibration is performed using dedicated constant-separation length scale (cLS) scans, as illustrated in Figure 7.13. During these scans, the two beams are kept at a fixed transverse separation while being moved together in a series of steps across the detector center. In the scan configuration used for the OO, NeNe, and pO programs, all cLS scans were performed under head-on conditions with 17 steps.

As the beams are displaced away from the nominal interaction point, the measured beamspot position is expected to follow the requested beam displacement. The length scale calibration factors can therefore be extracted from the slope of the linear relationship between the reconstructed beamspot position and the nominal beam position derived from the machine settings.

The first step of the analysis is the reconstruction of the beamspot position for each scan point. To ensure that only high-quality vertices with a precise position measurement are used, a set of selection criteria is applied. Vertices are required to have at least 25 associated tracks and must be reconstructed within the time interval corresponding to the scan step under consideration. To avoid potential transient effects during beam movements, the first and last two seconds of each scan step are excluded from the analysis.

For each scan point, the distributions of the reconstructed vertex positions in the transverse plane are fitted with Gaussian functions. The mean of the fitted distribution provides an estimate of the beamspot position for that scan step, while the fitted width reflects the size of the luminous region and the vertex resolution. Two example fits for consecutive scan steps in the horizontal direction are shown in Figure 7.19.

Once the vertex beamspots are extracted, the length scales could be extracted as the slopes between the nominal beamspot positions

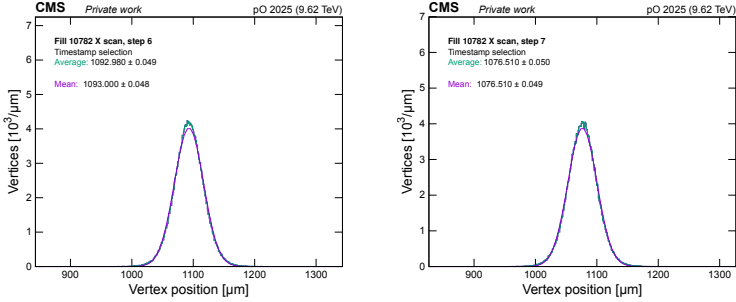


Figure 7.19.: Beamspot positions for scan steps 6 (left) and 7 (right) of the cLS scan.

and the mean vertex positions in the X and Y scans. The presence of orbit drift can however result in extra changes of the beam positions that can not be attributed to the length scale. To study the presence of orbit drift, the sizes of all the scan steps are plotted. If no orbit drift were present, all scan steps would show the same size. Figure 7.20 shows the step sizes for the cLS scan, as requested and as measured by the vertex beamspot, arcBPMs and DOROS. While some stochastic movement of the measured positions can be observed in all measurement methods, no noticeable linear orbit drift is present in these scan steps.

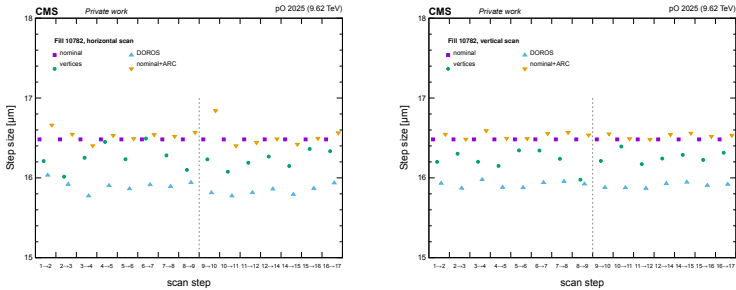


Figure 7.20.: Step sizes of the horizontal (left) and vertical (right) pO cLS scans.

As we know that the steps are supposed to be the same size, the orbit drift can be estimated for each subsystem as

$$\text{size}_{\text{step } i} - \text{size}_{\text{average}} = \text{size}_{\text{orbit drift step } i}. \quad (7.4)$$

The orbit drifts measured by each system are incorporated as corrections to the nominal beam positions. This procedure ensures that the corrected beam positions correspond to a constant step size throughout the scan, even in the presence of non-linear orbit movements.

Figure 7.21 shows the difference between the reconstructed beamspot position and the (orbit-drift-corrected) nominal position as a function of the latter, separately for both scan directions in the horizontal and vertical planes. For each scan direction, the distributions are shifted such that the corresponding linear fit has a zero intercept. This removes any dependence on the arbitrary choice of coordinate origin and allows the analysis to focus solely on the relative displacement scale.

The slopes of the resulting linear fits quantify the correction that must be applied to the nominal beam displacement to obtain the true beam position. Separate measurements are performed for the horizontal and vertical scan directions, yielding the calibration factors α_x and α_y , respectively. The final result of the length scale measurement for the pO program is shown in Figure 7.22. Similar results are found for the OO runs and NeNe run, which overall result in slightly lower length scales of ≈ 0.98 .

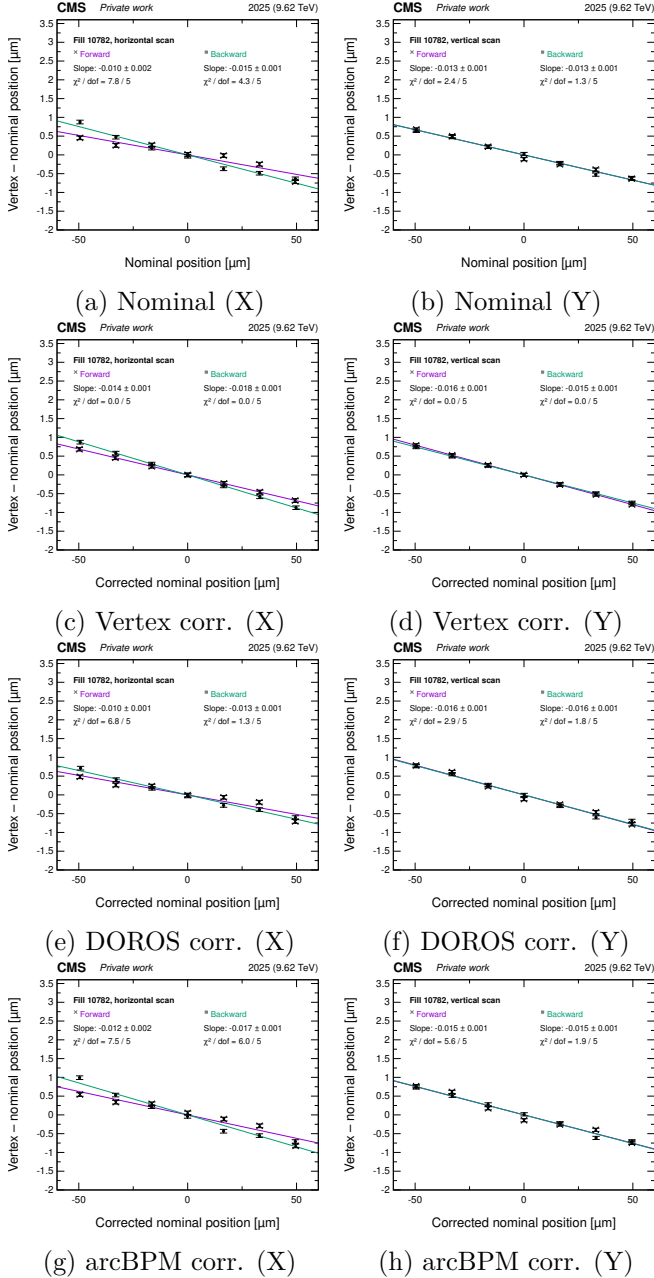


Figure 7.21.: Length scale scan results for horizontal (X, left column) and vertical (Y, right column) scans under different orbit correction schemes.

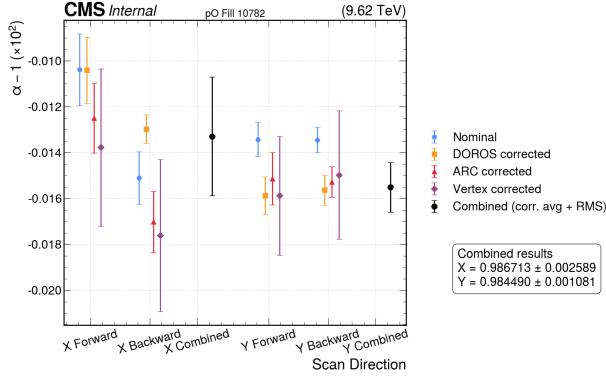


Figure 7.22.: Length scale results for the pO run.

Figure 7.23 shows residuals of the beam position as measured by DOROS and the position requested by the magnet system. The length scale effect and beam-beam corrections are extrapolated to the DOROS position from the beam optics settings. The remaining residuals are anticorrelated to the beam-beam deflection and can largely be attributed to the dynamic β effect. These corrections show a good understanding of these corrections and validates the assumption of small dynamic beta and residual orbit drift effects.

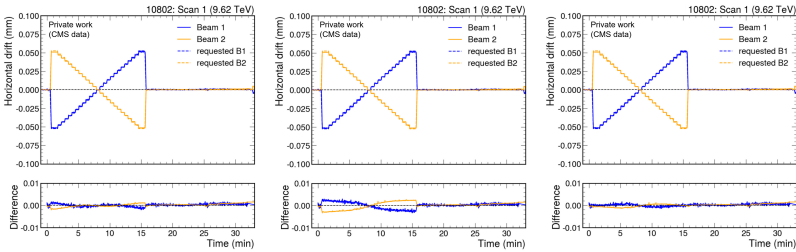


Figure 7.23.: Beam position residuals after multiple corrections. Left: orbit drift, middle: orbit drift+Beam-beam deflection, right: orbit drift+Beam-beam deflection+length scale.

Also for the pO program, a summary table of systematic uncertainties is constructed and shown in table 7.5. Updated values for the OO and NeNe program are highlighted in green, while my contributions to the final pO calibration are highlighted in orange.

The studies presented in this section provide the first luminosity calibrations for these unique datasets. These calibrations are essential for many analyses where luminosity is the dominant systematic uncertainty and therefore enable precision measurements that can have a significant impact on the field. As CMS is the only experiment that performed a dedicated VdM calibration program in the NeNe dataset, CMS is expected to play a leading role in analyses that rely on an accurate luminosity determination.

Table 7.5.: Summary of luminosity calibration uncertainties for the OO, NeNe and pO collision programs.

Source	OO (%)	NeNe (%)	pO (%)
Normalization			
Nonfactorization bias	3.0	3.0	3.0
Beam current calibration	0.3	0.3	0.3
Ghosts & satellites	0.3	0.3	0.7
Length scale	-3.31 corr.±0.3	-4.38 corr.±0.4	-2.86 corr.±0.3
OD & residual beam pos.	1.0	1.0	1.0
Beam-beam & dynamic- β	0.5	0.5	0.5
Bunch-to-bunch variation	0.1	0.2	0.83
Scan-to-scan variation	1.7	0.7	1.0
Cross-detector consistency	0.4	0.3	0.25
Background correction	0.6	0.3	0.3
Integration			
Stability	0.4	0.0	0.27
Linearity	0.3	0.5	0.23
Afterglow	3.0	3.0	0.0
Total	5.7	5.4	3.6

7.5. Towards a measurement of the inelastic cross section and charged particle production rates in pO collisions.

7.5.1. Introduction

A central objective of the LHC light-ion program is to provide measurements that constrain hadronic interaction models used in the interpretation of ultra-high-energy cosmic-ray air showers. Within this program, the author's work has focused on ensuring access to the required data and detector capabilities for key measurements that provide direct experimental input to these models. To this end, the author initiated, in collaboration with Dr. Michael Pitt, an analysis of the inelastic proton–oxygen (pO) cross section and charged-particle production rates in proton–oxygen collisions at $\sqrt{s_{\text{NN}}} = 9.62$ TeV.

The analysis exploits the dE/dx information measured in the CMS silicon tracker (explained in the following section) for charged-hadron identification and the determination of production rates. Forward detectors are also used to enhance sensitivity to the diffractive components of the inelastic cross section. These measurements provide important constraints on particle production in a collision system directly relevant to cosmic-ray interactions in the atmosphere.

At the time of writing, the analysis is in the early stages of CMS internal review. The following sections summarise the measurement strategy and the current status of the analysis.

7.5.2. Samples

The analysis uses proton–oxygen collision data collected by CMS during the 2025 LHC oxygen run at $\sqrt{s_{\text{NN}}} = 9.62$ TeV. Events are selected using a ZeroBias trigger, providing an unbiased sample of minimum-bias interactions. The pO run is composed of 3 fills,

each of a few runs. The pileup rate during the pO run varied between runs, as well as with the trigger prescales. This analysis is restricted to the ultra-low pileup sample, taken at the start of the program, corresponding to fill 10779 and run 393952. During this period, the average number of interactions per bunch crossing was very low ($\mu \simeq 0.1$), and beam conditions remained stable under luminosity leveling. Only data samples that were certified to be of excellent quality for physics analysis are considered in this analysis. After accounting for the trigger prescale, the effective integrated luminosity of the dataset is $L_{\text{eff}} = 110 \pm 4 \mu\text{b}^{-1}$.

Several Monte Carlo event generators are used to model the collision system and evaluate detector effects. These include PYTHIA 8 Angantyr, EPOS LHC-R and DPMJET-III. The base EPOS LHC-R model includes the capability to switch the Ultra-Relativistic Quantum Molecular Dynamics (UrQMD) ON or OFF, allowing the study of final-state hadron interactions. Hence, a second EPOS sample is generated that does not contain this hadron rescattering modeling. While the Pythia sample is generated independently, the other samples are generated using the crmc (Cosmic Ray Monte Carlo) program [115].

For these strong-interaction hadronic models, an estimate of the total inelastic pO cross section, $\sigma_{p\text{O}}^{\text{inel}}$, is derived from the corresponding nucleon–nucleon cross section, σ_{NN} , using a simple geometric scaling with the oxygen atomic mass number A :

$$\sigma_{p\text{O}}^{\text{inel}} = \sigma_{\text{NN}} A^{2/3}. \quad (7.5)$$

This scaling is used only to provide a common normalization reference for the hadronic MC samples.

In addition to the hadronic component, electromagnetic ultra-peripheral and photon-induced processes can contribute to low-multiplicity final states in $p\text{O}$ collisions. These processes are particularly relevant for measurements based primarily on reconstructed tracks, since exclusive or semi-exclusive photon-induced

events may pass track-based selections while having very different forward activity and event topology from genuine hadronic inelastic collisions. Dedicated photon-induced samples are therefore generated to estimate possible biases in the track-only selection used for the inclusive cross-section measurement.

An inclusive photoproduction sample, corresponding to γp interactions induced by the electromagnetic field of the oxygen ion, is generated with PYTHIA 8.317. In addition, exclusive vector-meson and lepton-pair production samples are generated with SUPERCHIC. These include exclusive ρ and ϕ photoproduction, as well as exclusive di-muon production through $\gamma\gamma \rightarrow \mu^+\mu^-$. The cross sections used for the respective MC samples are summarized in Table 7.6. The author has not contributed to the production of these samples but has contributed to the validation of all these samples.

Table 7.6.: Monte Carlo samples and their production cross sections.

Sample name	σ [mb]	σ_{NN} [mb]
Minimum-bias samples		
MinBias_Angantyr_Py8317_ffTune_Op	459.7	72.40
MinBias_crmcv2_2_1_dpmjetIII_Op	481.1	75.77
MinBias_crmcv2_2_1_eposlhc_Op	465.6	73.32
MinBias_crmcv2_2_1_eposnhs_Op	465.6	73.32
Photoproduction samples		
MinBias_PhotoInc_pythia_Op (γp inclusive)	6.5	—
Exclusive_PhotoRho_Superchic_Op ($\gamma p \rightarrow \rho p$)	0.25	—
Exclusive_PhotoPhi_Superchic_Op ($\gamma p \rightarrow \phi p$)	7.6×10^{-3}	—
Exclusive_Dimuon_Superchic_Op ($\gamma\gamma \rightarrow \mu^+\mu^-$)	1.9×10^{-3}	—

7.5.3. Detector validation

As the forward instrumentation at CMS is crucial to this analysis, significant work was performed into the validation of all crucial

subdetectors before the analysis grew in maturity. In this section, the performance of several detectors is discussed.

7.5.3.1. Tracker

The CMS tracker is the detector that is placed closest to the beamline. As discussed in section 4.2.2, the tracker is exposed to significant amounts of radiation. As the inner tracker was installed in 2017, significant amounts of radiation damage has been collected over data-taking years. As the tracker is crucial for this analysis, a good description of the tracker performance is required. To be able to correct small differences in efficiency for different p_T and η regions in the unfolding procedure, described in Sec. 7.6.1.1, we have to make sure that the p_T and η distributions are well described in MC.

The first check consists of validating the track densities in the tracker volume. The maps are produced separately for low- and high- p_T tracks, as shown in Fig. 7.24.

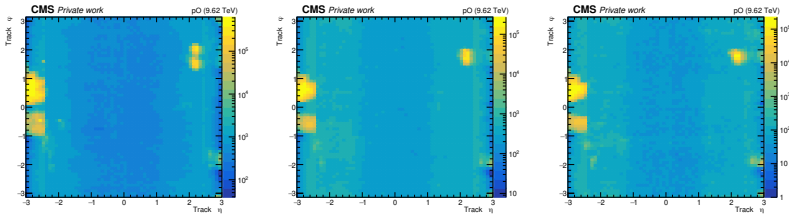


Figure 7.24.: Two-dimensional distributions of track pseudorapidity versus azimuthal angle in run 393952, for three transverse-momentum intervals: $0.2 < p_T < 0.5$ GeV (left), $0.5 < p_T < 1.0$ GeV (middle) and $p_T > 1.0$ GeV (right).

Large discrepancies between data and MC are noted in the end-caps. This effect was also observed during datataking of the Oxygen-Oxygen and Neon-Neon collisions. This effect likely orig-

inates from heavy damage to the pixel detector by radiation over the years. These issues can not be mitigated in a short timescale and requires us to perform our measurement in the central region with $|\eta| < 2$ region. Decent agreement is observed between Data and MC in the central region. Residual discrepancies are still observed, mainly for low p_T tracks. These will however be studied again once updated detector calibration and alignment is available. It is expected that such reprocessed data will improve the agreement further.

7.5.3.2. PPS and FSC

The calibration of the PPS and FSC is currently ongoing and is not being performed by the author. Both detectors are however crucial to the analysis and program as a whole. The FSCs were installed for the first time in these runs since RUN 1.

At $6 < |\eta| < 8$, the FSC detectors cover an important rapidity region that is otherwise uninstrumented between the HF and ZDC detectors. Their inclusion therefore has the potential to significantly improve the identification and characterization of rapidity gaps. The detector acceptance, however, remains only partially understood, as the scintillators detect secondary particles produced through interactions of forward particles with the beam pipe and surrounding material.



Figure 7.25.: Removal of the FSCs after the NeNe run.

Additional refurbished FSC stations were installed during the 2026 PbPb running period, further extending the available forward

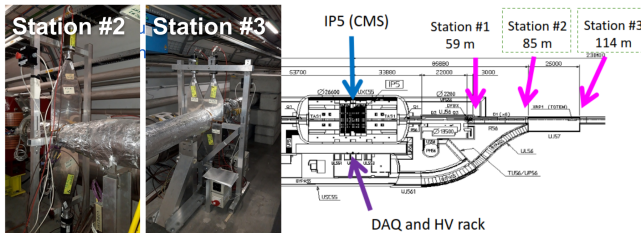


Figure 7.26.: The FSC stations, as used in RUN 1. The stations #2 and #3 installed for the pO run are highlighted.

instrumentation. The full capabilities and physics potential of these detectors have yet to be established and remain the subject of ongoing studies.

7.5.3.3. ZDC

In the asymmetric pO collision system at $\sqrt{s_{NN}} = 9.62$ TeV, the two ZDC sides probe different beam directions. The ZDC− detector corresponds to the proton-going side and is sensitive to forward neutral fragments from the proton beam. The ZDC+ detector corresponds to the oxygen-going side and is sensitive to neutral fragments and spectator neutrons from the oxygen ion. Since spectator nucleons retain the per-nucleon beam energy of approximately 3.4 TeV, multiple neutral spectators can produce very large energy deposits on the oxygen-going side.

The ZDC showed excellent performance, as shown in figure 7.27. The ZDC+ data shows the clear 1 neutron and 2 neutron peaks. These peaks are much more pronounced in MC as no Fermi smearing is applied on these neutrons in the hadronic models.

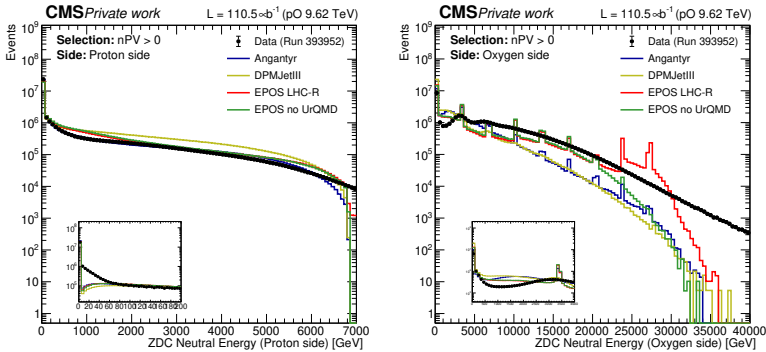


Figure 7.27.: ZDC neutral energy deposits for the proton side (left) and oxygen side (right).

7.5.4. HF

The HF calorimeter, covering the pseudorapidity range $3.0 < |\eta| < 5.2$, is one of the most important subdetectors for this analysis. Its information is used to characterize event activity and identify the presence of large rapidity gaps. As observed for the ZDC, a significant asymmetry is present between the two forward directions. The HF+ detector (oxygen-going side) typically records larger energy deposits and higher particle multiplicities than the HF− detector (proton-going side), as illustrated in Figure 7.28.

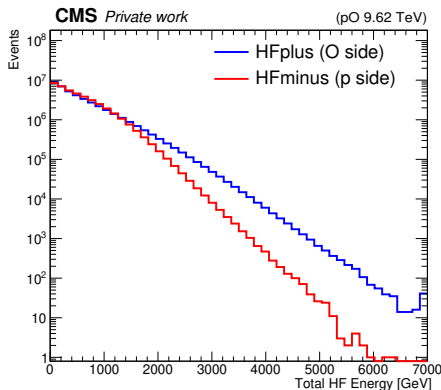


Figure 7.28.: HF calorimeter energy deposits measured on the proton-going and oxygen-going sides in data.

To verify the operational status of the detector, occupancy distributions were studied in a manner similar to that used for the tracker. All HF channels were found to be active, as demonstrated by the occupancy maps shown in Figure 7.29. The apparent absence of signals in half of the bins of the two most forward rings is a consequence of the detector geometry: these cells span twice the azimuthal acceptance of the neighbouring cells at larger radii, resulting in only every second azimuthal bin being recorded. Since the HF calorimeter is used to tag rapidity gaps, it must be operated close to its noise threshold, requiring a detailed under-

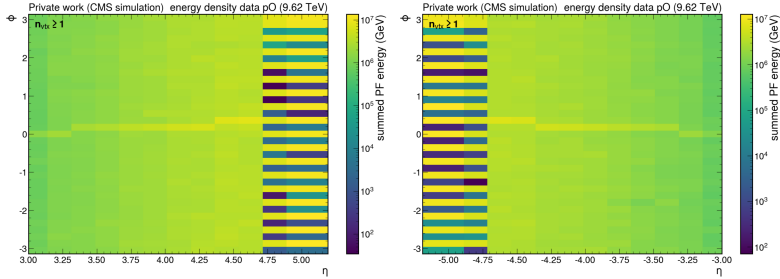


Figure 7.29.: Two-dimensional distributions of HF activity in pseudorapidity and azimuth for run 393952.

standing of detector backgrounds. Located in one of the most highly irradiated regions of CMS, HF is affected by several sources of noise. One important contribution is afterglow, whereby energy deposited in a given bunch crossing induces a residual response in subsequent crossings. In addition, satellite bunches can produce collisions at large longitudinal displacements from the nominal interaction point. Such events may leave little or no activity in the tracker while producing substantial energy deposits in one of the HF calorimeters.

To study these backgrounds, several complementary approaches were employed. HF energy deposits were investigated in events without reconstructed tracker activity, in ZeroBias events corresponding to empty bunch crossings, and using a method proposed by the author based on the superseparation scan performed during the VdM program, shown in Figure 7.30.

The superseparation scan provides the cleanest measurement of electronic noise and beam–gas backgrounds, while remaining largely insensitive to ghost collisions, satellite collisions, and afterglow effects. Empty bunch-crossing events retain sensitivity to afterglow backgrounds, whereas events without tracker activity are sensitive to satellite collisions. The latter method is consid-

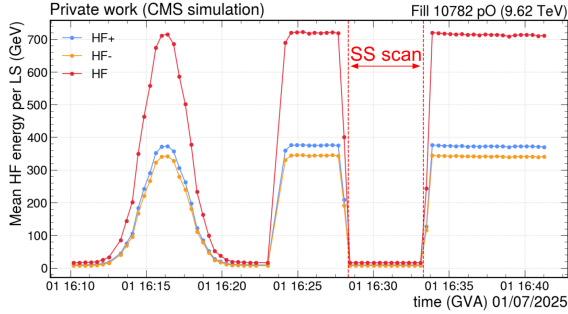


Figure 7.30.: Average HF energy deposits per lumisection during the superseparation scan. A preceding VdM scan is also visible.

ered less robust because genuine diffractive events can produce rapidity gaps extending into the central detector region. Consequently, only the first two methods are retained for the final noise determination.

For each dataset, the energy distribution of the leading particle-flow candidate reconstructed in HF was studied. The resulting spectra are in good agreement, with the empty bunch-crossing sample exhibiting slightly larger backgrounds at high energy due to afterglow effects. Noise thresholds were then determined from the empty bunch-crossing dataset by requiring a rejection of 99% of noise-induced activity. The resulting thresholds are shown in Figure 7.31. Based on these studies, a threshold of 8 GeV was selected for the analysis.

In Section 4.2.3, a coarse particle-identification method in HF based on the relative energy deposits in the long and short fibers was introduced. The corresponding energy distributions observed in the collision dataset are shown in Figure 7.32 for both categories. The electromagnetic component is primarily measured using the long fibers, while the hadronic component is estimated

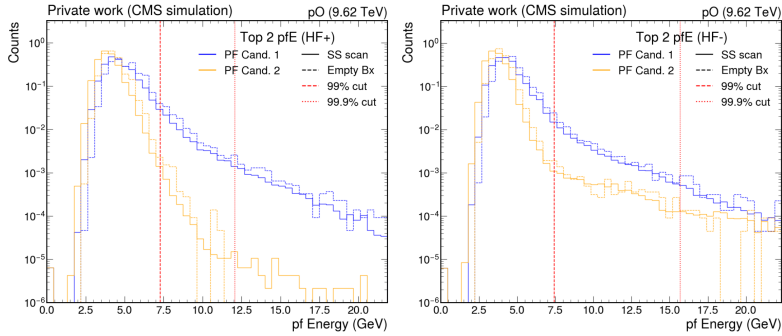


Figure 7.31.: Energy distributions of the leading (blue) and sub-leading (yellow) particle-flow candidates in HF for the two noise-enriched datasets. The 99% and 99.9% background rejection thresholds are indicated in red.

from the short fibers.

It has been observed that the current calibration of the long fibers is not yet satisfactory and is the subject of ongoing studies. In addition, the contributions from ion excitations and photon-induced ultra-peripheral interactions are still under investigation. An accurate description of the HF response will therefore play a crucial role in the next stages of the analysis.

7.6. Outlook towards planned future measurements

The work presented in this chapter remains at an early stage and continues to evolve rapidly. The LHC light-ion program has opened a new frontier in high-energy physics, with strong relevance to both collider and astroparticle physics. These new collision systems introduce challenges at every stage of the experimental program, from beam preparation and operation—including effects

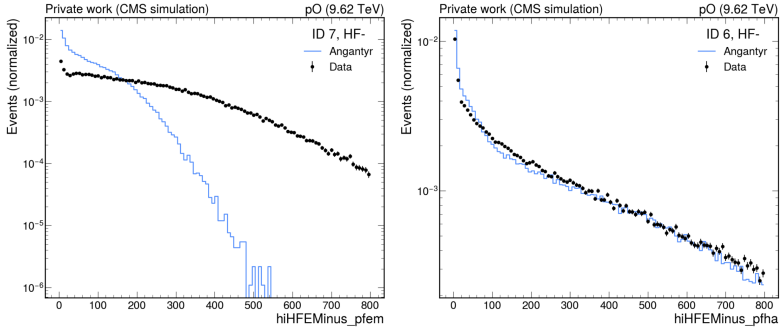


Figure 7.32.: HF calorimeter energy deposits measured on the proton going side for "electromagnetic" (id 7) and "hadronic" particle flow candidates.

such as ghost and satellite bunches, beam transmutation, and electron cooling challenges—to the calibration and performance validation of both new detectors, such as the FSC, and established systems such as HF.

Progress in this field requires the combined expertise of accelerator physicists, detector experts, reconstruction specialists, and phenomenologists. Addressing these challenges demands both detailed technical understanding and the development of novel analysis techniques. I hope this chapter has conveyed that the light-ion program operates in a significantly less explored regime than the work discussed previously, and that this combination of new opportunities and new challenges makes it a particularly exciting area of research.

The remainder of this chapter focuses on the planned analysis strategy, which is still under development. While this aspect is more closely aligned with my previous experience in data analysis, it nevertheless presents a number of unique challenges and has the potential to provide input of incredible valuable to both the collider and cosmic-ray communities.

7.6.1. A "dE/dx" measurement

Charged particles traversing the CMS silicon tracker lose energy through ionization. The measured specific energy loss per traversed length, referred to as dE/dx , can be estimated based on the discrete charge deposits in the silicon strips and pixels. The dependence of the energy loss on the particle momentum is given by the well-known Bethe-Bloch formula

$$-\left\langle \frac{dE}{dx} \right\rangle = K z^2 \frac{Z}{A} \frac{1}{\beta^2} \left[\frac{1}{2} \ln \left(\frac{2m_e c^2 \beta^2 \gamma^2 T_{\max}}{I^2} \right) - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right] \quad (7.6)$$

Here, $\langle dE/dx \rangle$ is the mean energy loss per unit length. K is a constant, z is the charge of the incident particle, Z is the atomic number of the material, and A is the atomic mass of the material. β is the particle velocity in units of the speed of light, and γ is the corresponding Lorentz factor $\frac{1}{\sqrt{1-\beta^2}}$. $T_{\max}$ is the maximum energy that can be transferred in a single collision, I is the mean excitation energy of the material, and $\delta(\beta\gamma)$ is the density effect correction.

This equation can be approximated by $K' \frac{m^2}{p^2} + C$ for reasonably high momenta. By combining the reconstructed track momentum with the measured dE/dx , distinct bands corresponding to pions, kaons, protons and deuterons can hence be identified as we get sensitivity to the mass of the particle.

It is exactly this separation that will be exploited by this analysis to determine particle yields within a fiducial phase space where both tracking performance and particle identification are well understood. The extracted particle multiplicities form the basis of the charged-particle production measurement. The dE/dx bands are shown in figure 7.33.

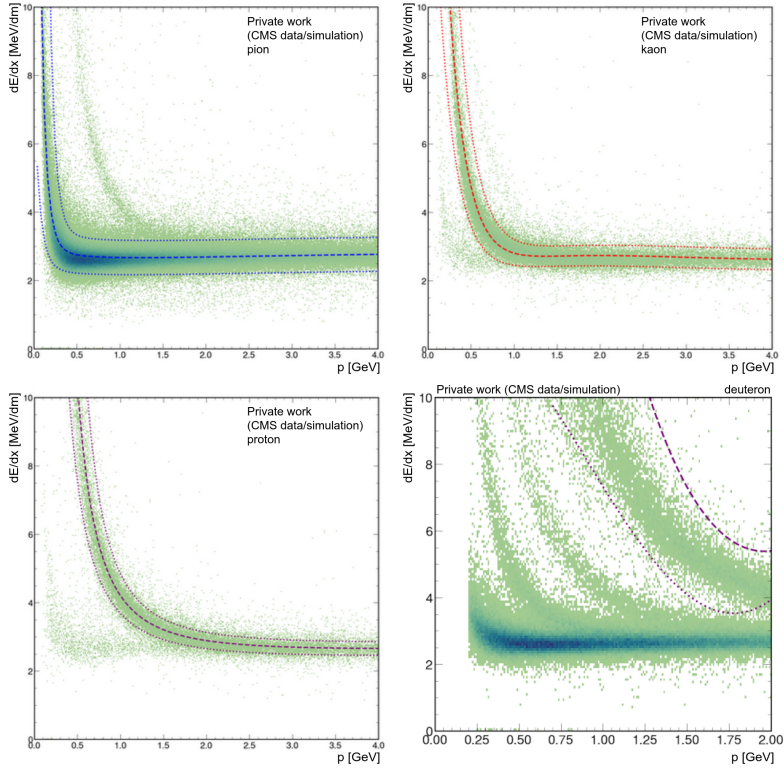


Figure 7.33.: dE/dx bands for pions (top left), kaons (top right), protons (bottom left) and deuterons (bottom right), as predicted by EPOS-LHC.R.

These multiplicities provide important input for constraining strangeness and baryon production in hadronic interaction models. As shown in Figure 7.34, the predicted production rates exhibit substantial variation between models, indicating that they are currently only weakly constrained by existing data. The production of deuterons is furthermore predicted only by the EPOS model. However, deuterons have already been identified in the data during the validation of the dE/dx performance, making

them a particularly interesting probe for further study.

Particle production rates will also be measured as a function of event multiplicity. This differential measurement strategy significantly reduces systematic uncertainties associated with tracking efficiencies, as many of these effects cancel in the ratio of identified-particle yields to the total charged-particle yield.

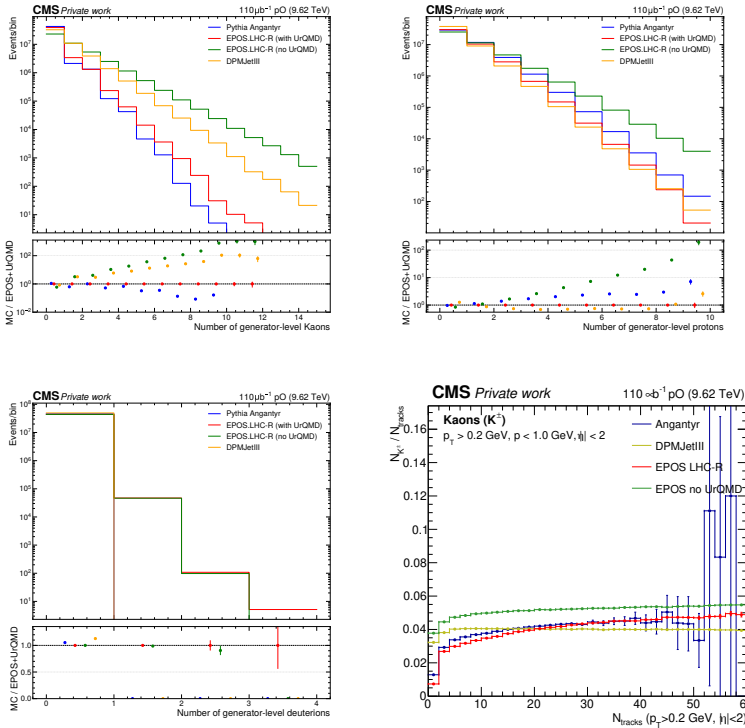


Figure 7.34.: Predicted particle production multiplicities for kaons (top left), protons (top right), deuterons (bottom left). Predicted kaon production rate as a function of track multiplicity (bottom right).

7.6.1.1. Unfolding

As discussed in the $t\bar{t}W$ chapter, the measured detector-level observables are affected by detector inefficiencies and resolution effects that must be corrected before a meaningful comparison with theoretical predictions can be made. These detector effects are corrected through an unfolding procedure that maps reconstructed observables to a particle-level fiducial phase space. Given the minimum-bias nature of this analysis, the particle-level fiducial phase space is defined using the following loose track selections:

- $p_T > 200$ MeV,
- $|\eta| < 2$,
- $p < p_{\max}$, where $p_{\max} = 750$ MeV for kaons, 1.0 GeV for protons, and 1.5 GeV for deuterons.

These requirements reflect the experimental acceptance and particle-identification capabilities of the detector. The lower momentum threshold avoids the region of rapidly decreasing tracking efficiency, while the pseudorapidity selection corresponds to the tracker acceptance. The upper momentum requirements arise from the dE/dx response, where the different particle species remain sufficiently separated for reliable identification below these momenta.

At the event level, the fiducial selection requires the presence of a reconstructed interaction vertex. Response matrices are constructed from simulated events and account for tracking inefficiencies, momentum migration, and particle-identification effects. An example response matrix used for the unfolding of the kaon multiplicity distribution is shown in Figure 7.35.

The primary unfolding strategy employed in this analysis is based on TUnfold with Tikhonov regularization. The technical implementation is built on the RooUnfold framework [116]. Two alternative unfolding techniques were also considered: iterative

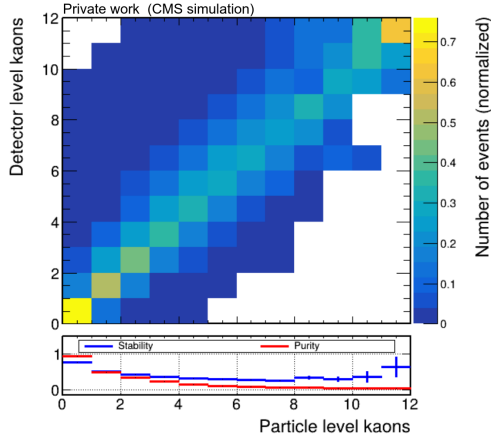


Figure 7.35.: Response matrix used for the unfolding of the kaon multiplicity distribution.

Bayesian unfolding (D’Agostini unfolding) and singular value decomposition (SVD) unfolding.

Tikhonov regularization is particularly well suited to this analysis, as the observables under consideration are expected to exhibit smooth underlying distributions. In contrast, iterative Bayesian unfolding relies on a regularization driven by the prior distribution. Given the sizeable differences between the available event generators and the expectation that none of them provides a fully accurate description of the data, such a prior-dependent regularization is less desirable.

The unfolding strategy was validated through a series of stress tests in which each of the available Monte Carlo predictions was treated as pseudo-data and unfolded using response matrices derived from the other generators. These studies demonstrated that the TUnfold approach provides the most robust and stable performance for the observables considered in this analysis.

7.6.2. Inelastic cross section

In addition to the measurement of charged-particle production rates, the determination of the total inelastic p–O cross section is of significant interest to the cosmic-ray community. A further decomposition of this cross section into its diffractive components would provide additional constraints for hadronic interaction models.

In this analysis, events are selected using the ZeroBias trigger (*HLT_OxyZeroBias*), which records bunch crossings independently of detector activity and therefore provides an unbiased sample of inelastic proton–oxygen interactions. Since the trigger is independent of detector response, the recorded dataset includes both empty bunch crossings and non-collision backgrounds. For $\mu = 0.1$, the probability of an empty bunch crossing is approximately $\exp(-\mu) \simeq 90.5\%$. An offline selection is therefore required to define the visible minimum-bias sample.

The inelastic p–O cross section is extracted from the visible event yields obtained using the ZeroBias trigger and vertex selection. In addition, the HF detector is used to define an alternative visible minimum-bias sample requiring at least one reconstructed calorimeter tower above 8 GeV, which reduces noise contributions. Beyond the inclusive selection, distinct HF- and PPS-based event categories are defined to construct orthogonal measurement regions with enhanced sensitivity to non-diffractive, diffractive, photon-induced, and background contributions. These components will be extracted through a template-based fit to the event yields in the different categories.

For the inelastic cross section measurement, the visible cross section of each selection is defined as the cross section corresponding to inelastic collisions satisfying the respective event criteria. Since multiple independent inelastic interactions may occur within a single recorded bunch crossing, the raw event counts must be cor-

rected for pileup. The visible cross section is therefore computed as:

$$\sigma_{\text{vis}}^i = \frac{(N_{\text{obs}}^i - N_{\text{bkg}}^i) \cdot f_{\text{PU}}}{\mathcal{L}_{\text{eff}}}, \quad (7.7)$$

where N_{obs}^i denotes the number of observed bunch crossings passing the selection (e.g. requiring at least one reconstructed primary vertex, $n_{\text{PV}} > 0$), N_{bkg}^i is the estimated background contribution from sources such as instrumental noise, beam–gas interactions, and ultra-peripheral photon-induced processes, and $\mathcal{L}_{\text{eff}}$ is the effective integrated luminosity accounting for the ZeroBias trigger prescale. For the dataset under study, $\mathcal{L}_{\text{eff}} = 110 \mu\text{b}^{-1}$ with a relative uncertainty of 4%.

The pileup correction factor f_{PU} accounts for the probability of multiple inelastic interactions occurring within a single triggered bunch crossing. Assuming Poisson statistics with mean interaction rate μ , this factor is defined as the ratio of the true mean number of interactions to the probability of observing at least one interaction:

$$f_{\text{PU}} = \frac{\mu}{1 - e^{-\mu}}. \quad (7.8)$$

For the nominal dataset with $\mu = 0.10$, this corresponds to a correction of approximately +5.1%, ensuring that σ_{vis}^i reflects the true interaction rate rather than the raw event rate.

The number of recorded events passing the baseline vertex selection, inclusive HF selections, vertex–HF overlap regions, and the main HF/PPS categories are summarized in Table 7.7. The cross section values in Table 7.7 represent pileup-corrected detector-level visible cross sections for the corresponding event selections and do not yet include any background subtraction or efficiency corrections.

Table 7.7.: Visible cross sections for the baseline preselection, inclusive HF selections, vertex–HF overlap regions, and the main analysis categories in run 393952. No background subtraction or reconstruction-efficiency correction is applied.

Selection	Recorded events	σ_{vis} [mb]
Inclusive selections		
Vertex ($n_{\text{PV}} > 0$)	42,204,330	403.2
Single-arm _{HF} , $E_{\text{HF}} > 8$ GeV	49,969,496	477.4
Double-arm _{HF} , $E_{\text{HF}+} > 8$ & $E_{\text{HF}-} > 8$ GeV	35,266,632	336.9
$n_{\text{PV}} > 0$ OR single-arm _{HF}	71,591,329	683.9
$n_{\text{PV}} > 0$ OR double-arm _{HF}	62,436,138	596.4
Vertex–HF overlap regions		
$n_{\text{PV}} > 0$ & double-arm HF	15,034,824	143.6
$n_{\text{PV}} > 0$ & no double-arm HF	27,169,506	259.5
Double-arm HF & no $n_{\text{PV}} > 0$	20,231,808	193.3
Mutually exclusive categories after preselection		
Vtx & no PPS & no HF	21,499,572	205.4
Vtx & no PPS & HF+ only	2,439,884	23.3
Vtx & no PPS & HF- only	2,837,927	27.1
Vtx & no PPS & HF double-arm	14,597,634	139.5
Vtx & PPS & no HF	122,261	1.17
Vtx & PPS & HF+ only	147,061	1.40
Vtx & PPS & HF- only	122,801	1.17
Vtx & PPS & HF double-arm	437,190	4.18

7.6.2.1. Phase-space overlap and topological decomposition

The baseline vertex selection ($n_{\text{PV}} > 0$) and the inclusive double-arm HF selection probe only partially overlapping regions of the $p\text{O}$ collision phase space.

As shown in Table 7.7, the visible cross section for the $n_{\text{PV}} > 0$ selection is 403.2 mb, while the inclusive double-arm HF selection gives 336.9 mb. The overlap between the two selections, corresponding to events satisfying both $n_{\text{PV}} > 0$ and the double-arm

HF requirement, is only 143.6 mb. Thus, only about 24% of the events selected by the logical OR of the vertex and double-arm HF requirements satisfy both selections simultaneously. The data are therefore dominated by two largely disjoint populations.

This limited overlap is not reproduced by the baseline hadronic Monte Carlo predictions studied so far. In simulation, standard non-diffractive hadronic interactions typically produce particle activity over a broad pseudorapidity range, leading to a much larger overlap between the central vertex and double-arm HF selections. The observed topological disjointness may point to sizable contributions from processes that are not fully described by the baseline hadronic event-generator mixtures:

1. The substantial *vertex-only* population is compatible with contributions from low-multiplicity central processes, including possible photon-induced topologies such as vector-meson photoproduction. However, this region can also receive contributions from detector effects, fake vertices, or other non-collision backgrounds, and therefore requires dedicated validation.
2. The *HF-only* population is sensitive to very-forward activity with little reconstructed central tracking activity. Possible contributions include diffractive dissociation, electromagnetic dissociation, beam-related backgrounds, detector noise, and other ultra-peripheral topologies.

Consequently, a simple logical OR of the vertex and double-arm HF selections, which yields a pileup-corrected visible cross section of 596.4 mb, should not be interpreted directly as the hadronic inelastic cross section. It can include substantial contributions from photon-induced topologies, electromagnetic dissociation, beam-related backgrounds, and detector noise. Studies to better understand these components are currently ongoing.

Once these contributions are constrained, the fiducial cross section can be obtained by correcting the visible cross section for the combined trigger and vertex reconstruction efficiencies:

$$\sigma_{\text{fid}} = \frac{\sigma_{\text{vis}}}{\epsilon_{\text{reco}}}. \quad (7.9)$$

If the fiducial phase space provides sufficient coverage of the total inelastic phase space, an extrapolation to the full inelastic cross section can be performed using simulation:

$$\sigma_{\text{inel}} = \frac{\sigma_{\text{fid}}}{A_{\text{fid}}}, \quad (7.10)$$

where A_{fid} is defined as the fraction of generated inelastic $p\text{O}$ events that fall within the particle-level fiducial phase space.

7.6.3. Pion exchange

A notable mechanism contributing to forward neutral-particle production is pion exchange, in which a beam proton emits a pion and transitions into a leading neutron, $p \rightarrow n + \pi^+$. In EPOS LHC-R, this process contributes at the level of about 9% of the generated inclusive sample. The exact cross section of this process is however not well known and is directly related to the pion-air cross section in cosmic ray hadronic models.

The initial proton-air interaction is a crucial parameter to describe the atmospheric depth at which the shower reaches its maximum development for (fluorescence/Cherenkov) light and is directly probed by measuring the inelastic cross section, as discussed before. Most muons are however produced at the end of the hadronic cascade after many generations of mainly pion-air interactions. This means that in order to describe the muonic content of the air shower, the pion-air cross section and its elasticity are crucial inputs to the models. A measurement of pion exchange would hence be of great benefit to the modeling community.

Events governed by π -exchange are characterized by a pronounced high-energy deposition in the ZDC— detector. The energy distribution broken down by the different physics processes is shown in Figure 7.36. It should be noted that non diffractive interactions involving high-density core formation can mimic the forward single-neutron signal produced by π -exchange. To effectively resolve and separate these overlapping spatial profiles, a finer transverse segmentation of the ZDC would be required, which is not currently available in the CMS forward configuration. Currently, a selection of $E > 3.5$ TeV in ZDC— results in a data sample that is augmented with pion exchange events by roughly a factor of 3. Dedicated studies to create a more pure region are however still under investigation.

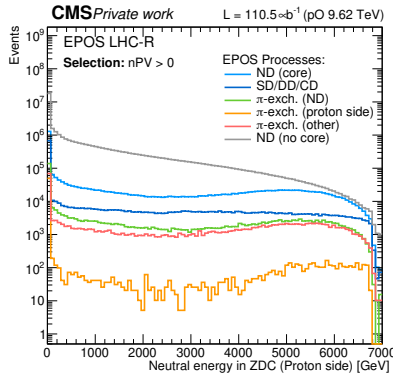


Figure 7.36.: Simulated neutral energy distribution in ZDC— (proton side), broken down by the production processes. The π -exchange signal produces a distinct high-energy structure that overlaps with the ND core formation background.

7.7. Outlook

In this section, I have discussed the LHC light-ion program and my involvement in its early stages. As the project remains under active development, no final physics results could yet be presented. I hope, however, that these pages have conveyed some of the excitement that comes with working in a genuinely new area of research, where many of the most important questions remain unanswered.

This program has confronted its contributors with challenges at every level, from accelerator operations and detector commissioning to reconstruction techniques and physics interpretation. Some of these challenges have already been overcome, while many others still lie ahead. Yet this is precisely what makes the effort so compelling. We are not merely refining established measurements; we are helping to build the foundations of a field that is only beginning to reveal its potential.

If successful, this program may play an important role in transforming cosmic-ray physics into a precision science. For decades, nature has provided us with particle collisions at energies far beyond those achievable in any laboratory. The limiting factor has not been the availability of data, but rather our understanding of the hadronic interactions that govern the development of air showers. The prospect of starting to unlock the full scientific power of these natural accelerators should be exciting to every particle physicist. At the same time, I have observed a tendency within our community to remain closer to familiar territory, with few other people interested in exploring these datasets in new directions.

I strongly believe it is important that CERN, and the particle-physics community more broadly, remains willing to venture beyond well-trodden paths. Many of the most significant discov-

eries in our field emerged not from asking familiar questions with greater precision, but from exploring regions where expectations were weakest. Young scientists have a particularly important role to play in this process. Yet the realities of modern academic careers often encourage caution, steering researchers toward safer and more predictable directions.

The coming decades will be a defining period for hadron physics. Beyond the final runs of the HL-LHC, the long-term future of energy-frontier hadron colliders remains uncertain. Whatever machine eventually succeeds the LHC, it is likely that few members of the current generation will witness another hadron collider operating at the highest energies during their careers. With the closure of RHIC and the finite lifetime of the LHC program, we are approaching the end of an era that has shaped particle physics for many decades.

This reality raises an important question: what legacy should the LHC leave behind? My own view is that we should use these remaining opportunities not only to complete established programs, but also to explore the unknown. The pp and lead-lead programs are undeniably important, but the unexplored regions of phase space and the unconventional collision systems may ultimately prove to be among the most valuable datasets we can leave for future generations of physicists.

For this reason, I am particularly interested in the evolution of the LHC Heavy-Ion Working Group initiative “New Ion Collisions @ LHC Run 4”. The decisions taken in the coming years will determine which datasets are recorded and which scientific questions remain accessible long after data taking has ended. In many ways, the future legacy of hadron physics is being shaped today. It is my hope that we choose to leave behind not only precision, but also curiosity, ambition, and the courage to explore

what lies beyond the horizon.

8 Final outlook

This thesis has presented the work of the author over the past four years. As each section already contains its own conclusions and outlook, only a brief summary is given here. During this PhD, I have worked across multiple subfields that are significantly different in nature.

$t\bar{t}W$: The first project concerned the first CMS measurement of differential cross sections of the $t\bar{t}W$ process. This analysis required high precision and extensive cross-checks throughout. The presence of complicated background contributions necessitated data-driven estimation techniques, while potential mismodelling of the signal required a robust machine-learning approach that balanced sensitivity and stability. This project provided a solid foundation for performing complex collider analyses.

The measurement provided valuable input to the modelling community. While no statistically significant deviations from Standard Model predictions were observed in the differential distributions, a mild excess at high jet multiplicities was noted. A similar trend has also been reported in the corresponding ATLAS measurement [117].

In parallel, theoretical developments have continued to improve the description of this process. At the time of writing, new NNLO calculations, including two-loop amplitudes at leading colour accuracy, have become available [118].

These advances, however, do not yet resolve the observed tension between theoretical predictions and measured cross sections. As theoretical uncertainties continue to decrease and the current trends persist, these effects may attract increased attention in

future studies.

With the RUN 3 dataset approximately doubling the available integrated luminosity and the HL-LHC program approaching, this process is expected to remain an important topic of study in the coming years.

$t\bar{t}t\bar{t}$ machine learning: The next project discussed in this thesis is the $t\bar{t}t\bar{t}$ ML analysis. Ghent was the institute that led the CMS measurement that resulted in the first observation of the $t\bar{t}t\bar{t}$ process [119]. This work provided an important benchmark for understanding rare top-quark production processes, but also left several open questions. In particular, the role of $t\bar{t}t$ production as a background and its potential overlap with the $t\bar{t}t\bar{t}$ signal definition remains an interesting topic of study.

Unlike the previous topic, this project does not rely solely on established analysis techniques and is instead positioned at the frontier of current analysis methodologies. The experimental separation of $t\bar{t}t$ and $t\bar{t}t\bar{t}$ processes has remained an open question since the original observation. While previous studies using more conventional multivariate techniques have shown limited separation power, the present work investigates whether more expressive machine-learning architectures can improve this discrimination. These first results seem extremely promising, far exceeding expectations.

Although the results presented here represent an important first step, the problem is not yet fully resolved. Theoretical studies have also highlighted that the separation between $t\bar{t}t$ and $t\bar{t}t\bar{t}$ is not uniquely defined beyond leading-order interpretations, motivating a more careful treatment of the underlying process definitions. Future work will therefore include the incorporation of improved $t\bar{t}t$ simulations and the development of a consistent particle-level definition of both processes that is both theoretically

well defined and meaningful. The architecture developed in this thesis can then be applied within this refined framework.

This work thus constitutes an important initial step towards a more detailed understanding of rare multi-top production processes.

proton–oxygen:

The final project differs substantially in scope and methodology from the previous two. While it may be considered less central within the traditional LHC program, its relevance to both collider and astroparticle physics is considerable. This work is situated on a different frontier of particle physics, where many fundamental measurements have yet to be performed and where the experimental challenges are often as novel as the physics questions themselves. Consequently, much of the work presented in this thesis is operational and preparatory in nature. The questions encountered in this program are rarely solved using established tools alone; instead, progress often requires a combination of expertise, ingenuity, and creative problem-solving.

Such preparatory work is, however, the foundation upon which a successful physics program is built. The progress achieved by the relatively small group of physicists working on these datasets has been remarkable. New collision systems, detector configurations, reconstruction techniques, and analysis strategies all had to be developed and validated before precision measurements could even be contemplated. The contributions described in this thesis form part of that effort and help establish the infrastructure and understanding required for future analyses.

The long-term ambition of this program extends well beyond the commissioning and understanding of a new collision system. Proton–oxygen collisions provide a unique opportunity to improve the modelling of hadronic interactions relevant to ultra-

high-energy cosmic rays, connecting collider physics directly to questions traditionally studied in astroparticle physics. While the results presented here represent only the first steps, they lay part of the groundwork for future measurements that may significantly improve our understanding of cosmic-ray air showers. I hope to continue contributing to this effort and to help realise the full scientific potential of this emerging field. the proposed analysis presented in this thesis constitutes the first step in this direction.

8.1. Outlook beyond this thesis

The author has worked on three topics that are all positioned at the forefront of the field, yet differ greatly in both methodology and scientific motivation. While each of these research directions is expected to remain relevant in the coming years, the author's primary focus will shift towards the emerging light-ion program.

As the first physics results from these datasets begin to mature, new opportunities may arise through the “New Ion Collisions @ LHC Run 4” initiative. At the same time, many open questions remain in top-quark physics. For this reason, the author expects to remain closely connected to both research areas and hopes to continue contributing to their development in the years ahead.

Looking further into the future, the question naturally arises: what will succeed the LHC? This question takes a central roll for the author. The decisions made in the coming years will define the future of this field. These decisions will determine which measurements can be performed, which questions can be addressed, and ultimately how the field evolves beyond the HL-LHC era.

The author hopes not only to witness these developments, but also to contribute to shaping them where possible. To this end, he has been actively involved in activities within ECFA and broader discussions on the future of European particle physics. Appendix [A](#)

presents one small contribution to these efforts. While such activities represent only a modest part of the work described in this thesis, they reflect a broader interest in helping define the long-term future of the field alongside contributing to its present scientific program. At this stage, my contributions remain small, but I hope they will mature alongside my own experience and allow me to play a more meaningful role in shaping the future of particle physics.

In a decade from now, I am curious to look back at these writings, at my position within the field, and at the state of the field as a whole!

Bibliography

- [1] CMS Collaboration. “Measurements of ttW differential cross sections and the leptonic charge asymmetry at $\sqrt{s} = 13$ TeV”. In: *Journal of High Energy Physics* 2026.3 (Mar. 2026). ISSN: 1029-8479. DOI: [10.1007/jhep03\(2026\)083](https://doi.org/10.1007/jhep03(2026)083). URL: [http://dx.doi.org/10.1007/JHEP03\(2026\)083](http://dx.doi.org/10.1007/JHEP03(2026)083).
- [2] Cecilia Borca et al. *Results of the analysis of a survey for young scientists on training quality in HEP instrumentation software and machine learning*. 2026. arXiv: [2603.16293](https://arxiv.org/abs/2603.16293) [hep-ex]. URL: <https://arxiv.org/abs/2603.16293>.
- [3] Pierre Van Mechelen et al. *National Input from Belgium to the 2026 Update to the European Strategy for Particle Physics*. 2026. URL: <https://indico.cern.ch/event/1439855/contributions/6461676/>.
- [4] David Marckx on behalf of the CMS collaboration. “Towards a differential ttW cross section measurement at CMS”. In: *SciPost Phys. Proc.* (2026), p. 009. DOI: [10.21468/SciPostPhysProc.18.009](https://doi.org/10.21468/SciPostPhysProc.18.009). URL: <https://scipost.org/10.21468/SciPostPhysProc.18.009>.
- [5] David Marckx on behalf of the CMS collaboration. “First differential ttW cross section measurement by CMS”. In: *Proceedings of the 59th Rencontres de Moriond, 2025-Electroweak Interactions and Unified Theories* (2025), p. 183. DOI: <https://doi.org/10.58027/m18g-zj54>. URL: <https://moriond.in2p3.fr/download/2025-ew-book.pdf>.

- [6] David Marckx on behalf of the CMS collaboration. “ttW and ttZ measurements from CMS”. In: *Proceedings of LHCP2025* (2025). DOI: <https://doi.org/10.22323/1.499.0251>. URL: <https://pos.sissa.it/499/251>.
- [7] Andrej B. Arbuzov. *Quantum Field Theory and the Electroweak Standard Model*. 2018. arXiv: [1801.05670](https://arxiv.org/abs/1801.05670) [hep-ph]. URL: <https://arxiv.org/abs/1801.05670>.
- [8] Lillian Hoddeson et al., eds. *The Rise of the Standard Model: Particle Physics in the 1960’s and 1970’s*. Cambridge University Press, 1997. ISBN: 978-0521578165.
- [9] S. Weinberg. “The making of the Standard Model”. In: *The European Physical Journal C* 34.1 (May 2004), pp. 5–13. ISSN: 1434-6052. DOI: [10.1140/epjc/s2004-01761-1](https://doi.org/10.1140/epjc/s2004-01761-1). URL: <http://dx.doi.org/10.1140/epjc/s2004-01761-1>.
- [10] H. Fritzsch, M. Gell-Mann, and H. Leutwyler. “Advantages of the color octet gluon picture”. In: *Physics Letters B* 47.4 (1973), pp. 365–368. ISSN: 0370-2693. DOI: [https://doi.org/10.1016/0370-2693\(73\)90625-4](https://doi.org/10.1016/0370-2693(73)90625-4). URL: <https://www.sciencedirect.com/science/article/pii/0370269373906254>.
- [11] Morel L et al. “Determination of the fine-structure constant with an accuracy of 81 parts per trillion.” In: *Nature* 558.7836 (Dec. 2020), pp. 61–65. DOI: [10.1038/s41586-020-2964-7](https://doi.org/10.1038/s41586-020-2964-7). URL: <https://pubmed.ncbi.nlm.nih.gov/33268866/>.
- [12] CMS Collaboration. “High-precision measurement of the W boson mass with the CMS experiment.” In: *Nature* 652 (2026), pp. 321–327. DOI: <https://doi.org/10.1038/>

-
- s41586-026-10168-5. URL: <https://doi.org/10.1038/s41586-026-10168-5>.
- [13] ALEPH DELPHI L3 OPAL SLD Collaborations LEP EW, SLD EW, and HF Groups. “Precision electroweak measurements on the Z resonance”. In: *Physics Reports* 427.5-6 (May 2006), pp. 257–454. DOI: [10.1016/j.physrep.2005.12.006](https://doi.org/10.1016/j.physrep.2005.12.006). URL: <http://dx.doi.org/10.1016/j.physrep.2005.12.006>.
- [14] S. Navas et al. (Particle Data Group). “CKM Quark-Mixing Matrix.” In: *Phys. Rev. D* 110.030001 (Dec. 2024 and 2025 update), p. 12. DOI: [PDG](https://doi.org/10.1103/PhysRevD.110.030001). URL: <https://ccwww.kek.jp/pdg/2025/reviews/rpp2025-rev-ckm-matrix.pdf>.
- [15] S. Navas et al. (Particle Data Group). “Neutrino Masses, Mixing, and Oscillations.” In: *Phys. Rev. D* 110.030001 (Dec. 2024), p. 36. DOI: [PDG](https://doi.org/10.1103/PhysRevD.110.030001). URL: <https://pdg.lbl.gov/2024/reviews/rpp2024-rev-neutrino-mixing.pdf>.
- [16] A. Denner and T. Hahn. “Radiative corrections to $W^+W^- \rightarrow W^+W^-$ in the electroweak standard model”. In: *Nuclear Physics B* 525.1-2 (Aug. 1998), pp. 27–50. ISSN: 0550-3213. DOI: [10.1016/S0550-3213\(98\)00287-9](https://doi.org/10.1016/S0550-3213(98)00287-9). URL: [http://dx.doi.org/10.1016/S0550-3213\(98\)00287-9](http://dx.doi.org/10.1016/S0550-3213(98)00287-9).
- [17] K. Eguchi et al. “First Results from KamLAND: Evidence for Reactor Antineutrino Disappearance”. In: *Phys. Rev. Lett.* 90 (2 Jan. 2003), p. 021802. DOI: [10.1103/PhysRevLett.90.021802](https://doi.org/10.1103/PhysRevLett.90.021802). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.90.021802>.
- [18] Q. R. Ahmad et al. “Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the

- Sudbury Neutrino Observatory”. In: *Phys. Rev. Lett.* 89 (1 June 2002), p. 011301. DOI: [10.1103/PhysRevLett.89.011301](https://doi.org/10.1103/PhysRevLett.89.011301). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.89.011301>.
- [19] Planck Collaboration. “Planck 2018 results. VI. Cosmological parameters”. In: *Astron. Astrophys.* 641 (2020), A6. DOI: [10.1051/0004-6361/201833910](https://doi.org/10.1051/0004-6361/201833910).
- [20] D. Clowe, M. Bradac, A. H. Gonzalez, et al. “A Direct Empirical Proof of the Existence of Dark Matter”. In: *Astrophys. J. Lett.* 648 (2006), pp. L109–L113. DOI: [10.1086/508162](https://doi.org/10.1086/508162).
- [21] A. G. Riess et al. “Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant”. In: *Astron. J.* 116 (1998), pp. 1009–1038. DOI: [10.1086/300499](https://doi.org/10.1086/300499).
- [22] S. Perlmutter et al. “Measurements of Omega and Lambda from 42 High-Redshift Supernovae”. In: *Astrophys. J.* 517 (1999), pp. 565–586. DOI: [10.1086/307221](https://doi.org/10.1086/307221).
- [23] Tie-Jiun Hou et al. “New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC”. In: *Phys. Rev. D* 103 (1 Jan. 2021), p. 014013. DOI: [10.1103/PhysRevD.103.014013](https://doi.org/10.1103/PhysRevD.103.014013). URL: <https://link.aps.org/doi/10.1103/PhysRevD.103.014013>.
- [24] L. A. Harland-Lang et al. “Parton distributions in the LHC era: MMHT 2014 PDFs”. In: *Eur. Phys. J. C* 75 (2015), p. 204. DOI: [10.1140/epjc/s10052-015-3397-6](https://doi.org/10.1140/epjc/s10052-015-3397-6).
- [25] NNPDF Collaboration. “Parton distributions from high-precision collider data”. In: *Eur. Phys. J. C* 77 (2017), p. 663. DOI: [10.1140/epjc/s10052-017-5199-5](https://doi.org/10.1140/epjc/s10052-017-5199-5).

-
- [26] K. J. Eskola et al. “EPPS16: nuclear parton distributions with LHC data”. In: *Eur. Phys. J. C* 77 (2017), p. 163. DOI: [10.1140/epjc/s10052-017-4725-9](https://doi.org/10.1140/epjc/s10052-017-4725-9).
- [27] K. Kovařík et al. “nCTEQ15: Global analysis of nuclear parton distributions with uncertainties in the CTEQ framework”. In: *Phys. Rev. D* 93 (8 Apr. 2016), p. 085037. DOI: [10.1103/PhysRevD.93.085037](https://doi.org/10.1103/PhysRevD.93.085037). URL: <https://link.aps.org/doi/10.1103/PhysRevD.93.085037>.
- [28] Christian Bierlich et al. “A comprehensive guide to the physics and usage of PYTHIA 8.3”. 2022. arXiv: [2203.11601](https://arxiv.org/abs/2203.11601) [hep-ph]. URL: <https://arxiv.org/abs/2203.11601>.
- [29] S. Catani and M. H. Seymour. “A general algorithm for calculating jet cross sections in NLO QCD”. In: *Nucl. Phys. B* 485 (1997), pp. 291–419. DOI: [10.1016/S0550-3213\(96\)00589-5](https://doi.org/10.1016/S0550-3213(96)00589-5).
- [30] S. Catani et al. “ q_T subtraction formalism for NNLO QCD calculations”. In: *Phys. Rev. Lett.* 103 (2009), p. 082001. DOI: [10.1103/PhysRevLett.103.082001](https://doi.org/10.1103/PhysRevLett.103.082001).
- [31] Tomáš Ježo et al. “An NLO+PS generator for $t\bar{t}$ and tW production and decay including non-resonant and interference effects”. In: *The European Physical Journal C* 76.12 (Dec. 2016). ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-016-4538-2](https://doi.org/10.1140/epjc/s10052-016-4538-2). URL: <http://dx.doi.org/10.1140/epjc/s10052-016-4538-2>.
- [32] Michelangelo L Mangano et al. “Matching matrix elements and shower evolution for top-pair production in hadronic collisions”. In: *Journal of High Energy Physics* 2007.01 (Jan. 2007), pp. 013–013. ISSN: 1029-8479. DOI: [10.1088/](https://doi.org/10.1088/)

- 1126-6708/2007/01/013. URL: <http://dx.doi.org/10.1088/1126-6708/2007/01/013>.
- [33] Rikkert Frederix and Stefano Frixione. “Merging meets matching in MC@NLO”. In: *Journal of High Energy Physics* 2012.12 (Dec. 2012). ISSN: 1029-8479. DOI: [10.1007/jhep12\(2012\)061](https://doi.org/10.1007/jhep12(2012)061). URL: [http://dx.doi.org/10.1007/JHEP12\(2012\)061](http://dx.doi.org/10.1007/JHEP12(2012)061).
 - [34] S. van der Meer. “Calibration of the effective beam height in the ISR”. Tech. rep. CERN-ISR-PO-68-31. CERN, 1968. URL: <https://cds.cern.ch/record/296752>.
 - [35] CMS Collaboration. “Precision luminosity measurement in proton-proton collisions at $\sqrt{s} = 13$ TeV in Run 2 of the LHC”. In: *Eur. Phys. J. C* 81 (2021), p. 800. DOI: [10.1140/epjc/s10052-021-09538-2](https://doi.org/10.1140/epjc/s10052-021-09538-2).
 - [36] CDF Collaboration. “Observation of Top Quark Production in $p\bar{p}$ Collisions”. In: *Phys. Rev. Lett.* 74 (1995), pp. 2626–2631. DOI: [10.1103/PhysRevLett.74.2626](https://doi.org/10.1103/PhysRevLett.74.2626).
 - [37] DØ Collaboration. “Search for High Mass Top Quark Production in $p\bar{p}$ Collisions”. In: *Phys. Rev. Lett.* 74 (1995), pp. 2422–2426. DOI: [10.1103/PhysRevLett.74.2422](https://doi.org/10.1103/PhysRevLett.74.2422).
 - [38] J. de Blas et al. “Impact of the Recent Measurements of the Top-Quark and W-Boson Masses on Electroweak Precision Fits”. In: *Physical Review Letters* 129.27 (Dec. 2022). ISSN: 1079-7114. DOI: [10.1103/physrevlett.129.271801](https://doi.org/10.1103/physrevlett.129.271801). URL: <http://dx.doi.org/10.1103/PhysRevLett.129.271801>.
 - [39] Dario Buttazzo et al. “Investigating the near-criticality of the Higgs boson”. In: *Journal of High Energy Physics* 2013.12 (Dec. 2013). ISSN: 1029-8479. DOI: [10.1007/](https://doi.org/10.1007/)

-
- jhep12(2013)089. URL: [http://dx.doi.org/10.1007/JHEP12\(2013\)089](http://dx.doi.org/10.1007/JHEP12(2013)089).
- [40] CMS Collaboration. “Observation of a pseudoscalar excess at the top quark pair production threshold”. In: *Reports on Progress in Physics* 88.8 (Aug. 2025), p. 087801. ISSN: 1361-6633. DOI: [10.1088/1361-6633/adf7d3](https://doi.org/10.1088/1361-6633/adf7d3). URL: <http://dx.doi.org/10.1088/1361-6633/adf7d3>.
- [41] ATLAS Collaboration. “Observation of a cross-section enhancement near the $t\bar{t}$ production threshold in $\sqrt{s} = 13\text{TeV}$ pp collisions with the ATLAS detector”. In: *Reports on Progress in Physics* 89.5 (May 2026), p. 057801. DOI: [10.1088/1361-6633/ae60a0](https://doi.org/10.1088/1361-6633/ae60a0). URL: <https://doi.org/10.1088/1361-6633/ae60a0>.
- [42] CMS Collaboration. “Observation of a pseudoscalar excess at the top quark pair production threshold in the single lepton channel”. Tech. rep. Geneva: CERN, 2026. URL: <https://cds.cern.ch/record/2957425>.
- [43] J. H. Kühn and G. Rodrigo. “Charge Asymmetry in Hadroproduction of Heavy Quarks”. In: *Physical Review Letters* 81.1 (July 1998), pp. 49–52. ISSN: 1079-7114. DOI: [10.1103/PhysRevLett.81.49](https://doi.org/10.1103/PhysRevLett.81.49). URL: <http://dx.doi.org/10.1103/PhysRevLett.81.49>.
- [44] CMS Collaboration. “Measurement of the top quark polarization and $t\bar{t}$ spin correlations using dilepton final states in proton-proton collisions at $\sqrt{s} = 13\text{ TeV}$ ”. In: *Phys. Rev. D* 100 (7 Oct. 2019), p. 072002. DOI: [10.1103/PhysRevD.100.072002](https://link.aps.org/doi/10.1103/PhysRevD.100.072002). URL: <https://link.aps.org/doi/10.1103/PhysRevD.100.072002>.

- [45] Gregory Mahlon and Stephen J. Parke. “Spin correlation effects in top quark pair production at the LHC”. In: *Phys. Rev. D* 81 (7 Apr. 2010), p. 074024. DOI: [10.1103/PhysRevD.81.074024](https://doi.org/10.1103/PhysRevD.81.074024). URL: <https://link.aps.org/doi/10.1103/PhysRevD.81.074024>.
- [46] CMS Collaboration. “Observation of quantum entanglement in top quark pair production in proton–proton collisions at 13 TeV”. In: *Reports on Progress in Physics* 87.11 (Oct. 2024), p. 117801. ISSN: 1361-6633. DOI: [10.1088/1361-6633/ad7e4d](https://doi.org/10.1088/1361-6633/ad7e4d). URL: <http://dx.doi.org/10.1088/1361-6633/ad7e4d>.
- [47] CMS Collaboration. “Observation of magic states of top quark pairs produced in proton-proton collisions at $\sqrt{s} = 13$ TeV”. Tech. rep. Geneva: CERN, 2025. URL: <https://cds.cern.ch/record/2926751>.
- [48] ATLAS and CMS Collaborations. “Combinations of single-top-quark production cross-section measurements and $|f_{\text{LV}}V_{tb}|$ determinations at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS experiments”. In: *Journal of High Energy Physics* 2019.5 (May 2019). ISSN: 1029-8479. DOI: [10.1007/jhep05\(2019\)088](https://doi.org/10.1007/jhep05(2019)088). URL: [http://dx.doi.org/10.1007/JHEP05\(2019\)088](http://dx.doi.org/10.1007/JHEP05(2019)088).
- [49] ATLAS and CMS Collaborations. “Combination of Measurements of the Top Quark Mass from Data Collected by the ATLAS and CMS Experiments at 7 and 8 TeV”. In: *Physical Review Letters* 132.26 (June 2024). ISSN: 1079-7114. DOI: [10.1103/physrevlett.132.261902](https://doi.org/10.1103/physrevlett.132.261902). URL: <http://dx.doi.org/10.1103/PhysRevLett.132.261902>.

-
- [50] CMS Collaboration. “Measurement of the top quark mass using a profile likelihood approach with the lepton+jets final states in proton–proton collisions at $\sqrt{s} = 13$ TeV”. In: *The European Physical Journal C* 83.10 (Oct. 2023). ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-023-12050-4](https://doi.org/10.1140/epjc/s10052-023-12050-4). URL: <http://dx.doi.org/10.1140/epjc/s10052-023-12050-4>.
- [51] CMS collaboration. “Measurement of differential tt production cross sections in the full kinematic range using lepton+jets events from proton-proton collisions at $\sqrt{s}=13$ TeV”. In: *Physical Review D* 104.9 (Nov. 2021). ISSN: 2470-0029. DOI: [10.1103/physrevd.104.092013](https://doi.org/10.1103/physrevd.104.092013). URL: <http://dx.doi.org/10.1103/PhysRevD.104.092013>.
- [52] CMS collaboration. “Measurement of tt normalised multi-differential cross sections in pp collisions at sqrt s=13 TeV, and simultaneous determination of the strong coupling strength, top quark pole mass, and parton distribution functions”. In: *The European Physical Journal C* 80.7 (July 2020). ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-020-7917-7](https://doi.org/10.1140/epjc/s10052-020-7917-7). URL: <http://dx.doi.org/10.1140/epjc/s10052-020-7917-7>.
- [53] Melissa van Beekveld et al. *Invariant-mass threshold resummation for the production of four top quarks at the LHC*. 2025. arXiv: [2505.10381 \[hep-ph\]](https://arxiv.org/abs/2505.10381). URL: <https://arxiv.org/abs/2505.10381>.
- [54] Tomáš Ježo and Manfred Kraus. “Hadroproduction of four top quarks in the powheg box”. In: *Physical Review D* 105.11 (June 2022). ISSN: 2470-0029. DOI: [10.1103/physrevd.105.114024](https://doi.org/10.1103/physrevd.105.114024). URL: <http://dx.doi.org/10.1103/PhysRevD.105.114024>.

- [55] Rikkert Frederix, Davide Pagani, and Marco Zaro. “Large NLO corrections in $t\bar{t}W^\pm$ and $t\bar{t}t\bar{t}$ hadroproduction from supposedly subleading EW contributions”. In: *JHEP* 02 (2018), p. 031. DOI: [10.1007/JHEP02\(2018\)031](https://doi.org/10.1007/JHEP02(2018)031). arXiv: [1711.02116](https://arxiv.org/abs/1711.02116) [hep-ph].
- [56] E. Abasov et al. “Search for Dark Matter Mediator in the Production of Three and Four Top Quarks”. In: *Physics of Particles and Nuclei* 56.2 (Apr. 2025), pp. 440–446. ISSN: 1531-8559. DOI: [10.1134/s1063779624701764](https://doi.org/10.1134/s1063779624701764). URL: <http://dx.doi.org/10.1134/S1063779624701764>.
- [57] Luc Darmé, Benjamin Fuks, and Mark Goodsell. “Cornering sgluons with four-top-quark events”. In: *Physics Letters B* 784 (Sept. 2018), pp. 223–228. ISSN: 0370-2693. DOI: [10.1016/j.physletb.2018.08.001](https://doi.org/10.1016/j.physletb.2018.08.001). URL: <http://dx.doi.org/10.1016/j.physletb.2018.08.001>.
- [58] Gauthier Durieux. *Triple top-quark production at NLO in QCD*. 2023. Zenodo: (v1.0) (hep-ph). URL: <https://doi.org/10.5281/zenodo.7682519>.
- [59] CMS Collaboration. “Stairway to discovery: A report on the CMS programme of cross section measurements from millibarns to femtobarns”. In: *Physics Reports* 1115 (Apr. 2025), pp. 3–115. ISSN: 0370-1573. DOI: [10.1016/j.physrep.2024.11.005](https://doi.org/10.1016/j.physrep.2024.11.005). URL: <http://dx.doi.org/10.1016/j.physrep.2024.11.005>.
- [60] CMS Collaboration. “Observation of four top quark production in proton-proton collisions at $\sqrt{s} = 13$ TeV”. In: *Physics Letters B* 847 (Dec. 2023), p. 138290. ISSN: 0370-2693. DOI: [10.1016/j.physletb.2023.138290](https://doi.org/10.1016/j.physletb.2023.138290). URL: <http://dx.doi.org/10.1016/j.physletb.2023.138290>.

-
- [61] LHC Top Working Group. *LHCTopWG Summary Plots: Four Top Quark Production Cross Section*. Accessed 8 June 2026. 2025. URL: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCTopWGSummaryPlots#Four_Top_Quark_Production.
- [62] CMS Collaboration. *Search for physics beyond the standard model in four and three top quark production events using proton-proton collisions at $\sqrt{s} = 13$ TeV*. Tech. rep. Geneva: CERN, 2025. URL: <https://cds.cern.ch/record/2944177>.
- [63] Ewa Lopienska. *The CERN accelerator complex, layout in 2022*. Tech. rep. General Photo. CERN, 2022. URL: <https://cds.cern.ch/record/2800984>.
- [64] Ernest D. Courant and Hartland S. Snyder. “Theory of the alternating-gradient synchrotron”. In: *Annals of Physics* 3 (1958), pp. 1–48. DOI: [10.1016/0003-4916\(58\)90012-5](https://doi.org/10.1016/0003-4916(58)90012-5).
- [65] S Baird. *Accelerators for pedestrians*. Tech. rep. AB-Note-2007-014. Geneva: CERN, 2007. URL: <https://cds.cern.ch/record/1017689>.
- [66] E Renner et al. “Tailoring transverse beam characteristics with the new CERN PS Booster charge-exchange injection system”. In: *Journal of Physics: Conference Series* 2687.5 (Jan. 2024), p. 052031. DOI: [10.1088/1742-6596/2687/5/052031](https://doi.org/10.1088/1742-6596/2687/5/052031). URL: <https://doi.org/10.1088/1742-6596/2687/5/052031>.
- [67] V. G. Dudnikov. “Charge exchange injection into accelerators and storage rings”. In: *Phys. Usp.* 62.4 (2019), pp. 405–412. DOI: [10.3367/UFNe.2019.02.038528](https://doi.org/10.3367/UFNe.2019.02.038528). URL: <https://ufn.ru/en/articles/2019/4/f/>.

- [68] E. Waagaard et al. *Charting the Luminosity Capabilities of the CERN Large Hadron Collider with Various Nuclear Species*. 2025. DOI: <https://doi.org/10.1016/j.nima.2025.171118>. arXiv: 2508.19653 [physics.acc-ph]. URL: <https://arxiv.org/abs/2508.19653>.
- [69] A. Jeff et al. “Longitudinal density monitor for the LHC”. In: *Phys. Rev. ST Accel. Beams* 15 (3 Mar. 2012), p. 032803. DOI: [10.1103/PhysRevSTAB.15.032803](https://doi.org/10.1103/PhysRevSTAB.15.032803). URL: <https://link.aps.org/doi/10.1103/PhysRevSTAB.15.032803>.
- [70] D Brandt, L Vos, and B Spataro. *Impedance of the LHC Arc Beam position Monitors*. Tech. rep. Geneva: CERN, 2002. URL: <https://cds.cern.ch/record/691832>.
- [71] CMS Collaboration. *Public CMS Luminosity Information*. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>. Accessed: 2026-06-08. 2026.
- [72] CMS Collaboration. “CMS Physics: Technical Design Report Volume 1: Detector Performance and Software”. In: (2006). DOI: [10.2172/2510878](https://doi.org/10.2172/2510878).
- [73] Izaak Neutelings. *CMS Coordinate System*. https://tikz.net/axis3d_cms/. TikZ.net, accessed 2026-06-08. 2024.
- [74] CMS Collaboration. “Development of the CMS detector for the CERN LHC Run3”. In: *Journal of Instrumentation* 19.05 (May 2024), P05064. ISSN: 1748-0221. DOI: [10.1088/1748-0221/19/05/p05064](https://doi.org/10.1088/1748-0221/19/05/p05064). URL: <http://dx.doi.org/10.1088/1748-0221/19/05/P05064>.
- [75] CMS Collaboration. “The CMS experiment at the CERN LHC”. In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08004. DOI: [10.1088/1748-0221/3/08/S08004](https://doi.org/10.1088/1748-0221/3/08/S08004). URL: <https://doi.org/10.1088/1748-0221/3/08/S08004>.

-
- [76] CMS Collaboration. *The CMS Detector*. <https://cms.cern/detector>. Accessed: 2026-06-08. 2026.
- [77] CMS Collaboration. “Data Scouting in the CMS experiment at the LHC”. In: *Journal of Instrumentation* 12 (2017), P04016. DOI: [10.1088/1748-0221/12/04/P04016](https://doi.org/10.1088/1748-0221/12/04/P04016). URL: <https://doi.org/10.1088/1748-0221/12/04/P04016>.
- [78] CMS collaboration. “Enriching the physics program of the CMS experiment via data scouting and data parking”. In: *Physics Reports* 1115 (2025). CMS physics results from the first decade of LHC data, pp. 678–772. ISSN: 0370-1573. DOI: <https://doi.org/10.1016/j.physrep.2024.09.006>. URL: <https://www.sciencedirect.com/science/article/pii/S0370157324003247>.
- [79] W Adam et al. “Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC”. In: *Journal of Physics G: Nuclear and Particle Physics* 31.9 (2005), N9–N20. ISSN: 1361-6471. DOI: [10.1088/0954-3899/31/9/n01](https://doi.org/10.1088/0954-3899/31/9/n01). URL: <http://dx.doi.org/10.1088/0954-3899/31/9/N01>.
- [80] R. Frühwirth. “Application of Kalman filtering to track and vertex fitting”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 262.2 (1987), pp. 444–450. ISSN: 0168-9002. DOI: [https://doi.org/10.1016/0168-9002\(87\)90887-4](https://doi.org/10.1016/0168-9002(87)90887-4). URL: <https://www.sciencedirect.com/science/article/pii/0168900287908874>.
- [81] CMS collaboration. “Flavour tagging performance of the updated Unified Particle Transformer algorithm with the

- CMS experiment at $\sqrt{s}=13.6$ TeV”. In: (2025). URL: <https://cds.cern.ch/record/2948917>.
- [82] L. Buonocore et al. “Precise predictions for the associated production of a W boson with a top-antitop quark pair at the LHC”. In: *Phys. Rev. Lett.* 131 (2023), p. 231901. DOI: [10.1103/PhysRevLett.131.231901](https://doi.org/10.1103/PhysRevLett.131.231901). arXiv: [2306.16311](https://arxiv.org/abs/2306.16311) [hep-ph].
- [83] R. Frederix and S. Frixione. “Merging meets matching in MC@NLO”. In: *JHEP* 12 (2012), p. 061. DOI: [10.1007/JHEP12\(2012\)061](https://doi.org/10.1007/JHEP12(2012)061). arXiv: [1209.6215](https://arxiv.org/abs/1209.6215) [hep-ph].
- [84] R. Frederix and I. Tsinikos. “On improving NLO merging for $t\bar{t}W$ production”. In: *JHEP* 11 (2021), p. 029. DOI: [10.1007/JHEP11\(2021\)029](https://doi.org/10.1007/JHEP11(2021)029). arXiv: [2108.07826](https://arxiv.org/abs/2108.07826) [hep-ph].
- [85] F. Maltoni et al. “Top-quark charge asymmetry and polarization in $t\bar{t}W$ production at the LHC”. In: *Physics Letters B* 736 (Sept. 2014), pp. 252–260. ISSN: 0370-2693. DOI: [10.1016/j.physletb.2014.07.033](https://doi.org/10.1016/j.physletb.2014.07.033). URL: <http://dx.doi.org/10.1016/j.physletb.2014.07.033>.
- [86] CMS Collaboration. “Measurement of the cross section of top quark-antiquark pair production in association with a W boson in proton-proton collisions at $\sqrt{s} = 13$ TeV”. In: *Journal of High Energy Physics* 2023.7 (July 2023). ISSN: 1029-8479. DOI: [10.1007/jhep07\(2023\)219](https://doi.org/10.1007/jhep07(2023)219). URL: [http://dx.doi.org/10.1007/JHEP07\(2023\)219](http://dx.doi.org/10.1007/JHEP07(2023)219).
- [87] CMS Collaboration. “Muon identification using multivariate techniques in the CMS experiment in proton-proton collisions at $\sqrt{s}=13$ TeV”. In: *Journal of Instrumentation* 19.02 (Feb. 2024), P02031. ISSN: 1748-0221. DOI: [10.1088/1748-0221/19/02/p02031](https://doi.org/10.1088/1748-0221/19/02/p02031). URL: <http://dx.doi.org/10.1088/1748-0221/19/02/P02031>.

-
- [88] Luka Lambrecht. *Measurements of rare processes with top quarks and electroweak bosons in proton-proton collisions at the CMS detector*. May 2024. URL: <https://repository.cern/records/9758f-5jd52>.
- [89] Tianqi Chen and Carlos Guestrin. “XGBoost: A Scalable Tree Boosting System”. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. San Francisco, California, USA: Association for Computing Machinery, 2016, pp. 785–794. ISBN: 9781450342322. DOI: [10.1145/2939672.2939785](https://doi.org/10.1145/2939672.2939785). URL: <https://doi.org/10.1145/2939672.2939785>.
- [90] CMS Collaboration. “Energy calibration and resolution of the CMS electromagnetic calorimeter in pp collisions at $\sqrt{s} = 7$ TeV”. In: *Journal of Instrumentation* 8.09 (Sept. 2013), P09009. DOI: [10.1088/1748-0221/8/09/P09009](https://doi.org/10.1088/1748-0221/8/09/P09009). URL: <https://doi.org/10.1088/1748-0221/8/09/P09009>.
- [91] CMS Collaboration. “Muon identification using multivariate techniques in the CMS experiment in proton-proton collisions at $\sqrt{s}=13$ TeV”. In: *Journal of Instrumentation* 19.02 (Feb. 2024), P02031. ISSN: 1748-0221. DOI: [10.1088/1748-0221/19/02/p02031](https://doi.org/10.1088/1748-0221/19/02/p02031). URL: <http://dx.doi.org/10.1088/1748-0221/19/02/P02031>.
- [92] CMS Collaboration. “The CMS Statistical Analysis and Combination Tool: Combine”. In: *Computing and Software for Big Science* 8.1 (Nov. 2024). ISSN: 2510-2044. DOI: [10.1007/s41781-024-00121-4](https://doi.org/10.1007/s41781-024-00121-4). URL: <http://dx.doi.org/10.1007/s41781-024-00121-4>.
- [93] S. S. Wilks. “The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses”. In: *Annals*

- Math. Statist.* 9.1 (1938), pp. 60–62. DOI: [10.1214/aoms/1177732360](https://doi.org/10.1214/aoms/1177732360).
- [94] Glen Cowan et al. “Asymptotic formulae for likelihood-based tests of new physics”. In: *The European Physical Journal C* 71.2 (Feb. 2011). ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-011-1554-0](https://doi.org/10.1140/epjc/s10052-011-1554-0). URL: <http://dx.doi.org/10.1140/epjc/s10052-011-1554-0>.
- [95] Lars Ferencz et al. *Study of ttbb and ttW background modelling for ttH analyses*. 2023. arXiv: [2301.11670](https://arxiv.org/abs/2301.11670) [hep-ex].
- [96] A. Rogozhnikov et al. “New approaches for boosting to uniformity”. In: *Journal of Instrumentation* 10.03 (Mar. 2015), T03002. DOI: [10.1088/1748-0221/10/03/T03002](https://doi.org/10.1088/1748-0221/10/03/T03002). URL: <https://doi.org/10.1088/1748-0221/10/03/T03002>.
- [97] CMS Collaboration. “Measurement of the Higgs boson production rate in association with top quarks in final states with electrons, muons, and hadronically decaying tau leptons at $\sqrt{s} = 13$ TeV”. In: *The European Physical Journal C* 81.4 (Apr. 2021). ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-021-09014-x](https://doi.org/10.1140/epjc/s10052-021-09014-x). URL: <http://dx.doi.org/10.1140/epjc/s10052-021-09014-x>.
- [98] Andy Buckley et al. “LHAPDF6: parton density access in the LHC precision era”. In: *The European Physical Journal C* 75.3 (Mar. 2015). ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-015-3318-8](https://doi.org/10.1140/epjc/s10052-015-3318-8). URL: <http://dx.doi.org/10.1140/epjc/s10052-015-3318-8>.
- [99] Fabio Cascioli et al. “NLO matching for ttbb production with massive b -quarks”. In: *Phys. Lett. B* 734 (2014), pp. 210–214. DOI: [10.1016/j.physletb.2014.05.040](https://doi.org/10.1016/j.physletb.2014.05.040). arXiv: [1309.5912](https://arxiv.org/abs/1309.5912) [hep-ph].

-
- [100] Benjamin Nachman and Dennis Noll. “Neural refinement of sample weights”. In: *Physical Review D* 112.9 (Nov. 2025). ISSN: 2470-0029. DOI: [10.1103/ynx16-h7n9](https://doi.org/10.1103/ynx16-h7n9). URL: <http://dx.doi.org/10.1103/ynx16-h7n9>.
- [101] Matthias Fey and Jan E. Lenssen. “Fast Graph Representation Learning with PyTorch Geometric”. In: *ICLR Workshop on Representation Learning on Graphs and Manifolds*. 2019.
- [102] Matthias Fey et al. “PyG 2.0: Scalable Learning on Real World Graphs”. In: *Temporal Graph Learning Workshop @ KDD*. 2025.
- [103] Shaked Brody, Uri Alon, and Eran Yahav. *How Attentive are Graph Attention Networks?* 2022. arXiv: [2105.14491](https://arxiv.org/abs/2105.14491) [cs.LG]. URL: <https://arxiv.org/abs/2105.14491>.
- [104] Samuel L. Smith et al. *Don’t Decay the Learning Rate, Increase the Batch Size*. 2018. arXiv: [1711.00489](https://arxiv.org/abs/1711.00489) [cs.LG]. URL: <https://arxiv.org/abs/1711.00489>.
- [105] George Stoica and Mihaela Elena Breabăn. “bs-scheduler: A Batch Size Scheduler library compatible with PyTorch DataLoaders”. In: *SoftwareX* 30 (2025), p. 102162. ISSN: 2352-7110. DOI: <https://doi.org/10.1016/j.softx.2025.102162>. URL: <https://www.sciencedirect.com/science/article/pii/S2352711025001293>.
- [106] Stoica George. “Increasing the Training Speed with Batch Size Schedulers”. In: *2025 27th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC)*. 2025, pp. 540–547. DOI: [10.1109/SYNASC69064.2025.00078](https://doi.org/10.1109/SYNASC69064.2025.00078).

- [107] CMS collaboration. *Measurement of 3 Top Quark Production*. Tech. rep. Geneva: CERN, 2026. URL: <https://cds.cern.ch/record/2957967>.
- [108] Gauthier Durieux. *The inseparable three and four tops*. Contribution to LHC TOP WG meeting. Accessed: 10 Jun 2026. 2026. URL: <https://indico.cern.ch/event/1674087/contributions/7074746/>.
- [109] CERN. *Convention for the Establishment of a European Organization for Nuclear Research*. <https://council.web.cern.ch/en/content/convention-establishment-european-organization-nuclear-research>. Accessed: 2025-07-16.
- [110] Tanguy Pierog, Benjamin Guiot, and Klaus Werner. “Air Shower Development, pion interactions and modified EPOS Model”. In: Aug. 2016, p. 337. DOI: [10.22323/1.236.0337](https://doi.org/10.22323/1.236.0337).
- [111] Karl-Heinz Kampert and Michael Unger. “Measurements of the cosmic ray composition with air shower experiments”. In: *Astroparticle Physics* 35.10 (2012), pp. 660–678. ISSN: 0927-6505. DOI: <https://doi.org/10.1016/j.astropartphys.2012.02.004>. URL: <https://www.sciencedirect.com/science/article/pii/S0927650512000382>.
- [112] Tanguy Pierog and Klaus Werner. *EPOS.LHC-R : a global approach to solve the muon puzzle*. 2025. arXiv: [2508.07105 \[astro-ph.HE\]](https://arxiv.org/abs/2508.07105). URL: <https://arxiv.org/abs/2508.07105>.
- [113] M Bassetti and George A Erskine. *Closed expression for the electrical field of a two-dimensional Gaussian charge*. Tech. rep. Geneva: CERN, 1980. URL: <https://cds.cern.ch/record/122227>.

-
- [114] “CMS luminosity measurement using 2016 proton-nucleus collisions at nucleon-nucleon center-of-mass energy of 8.16 TeV”. In: (2018).
- [115] Colin Baus Ralf Ulrich Tanguy Pierog. “The Cosmic Ray Monte Carlo Package, CRMC (v2.0.1)”. In: *Zenodo* (2021). DOI: <https://doi.org/10.5281/zenodo.4558705>. URL: <https://doi.org/10.5281/zenodo.4558705>.
- [116] Tim Adye. *Unfolding algorithms and tests using RooUnfold*. 2011. arXiv: [1105.1160](https://arxiv.org/abs/1105.1160) [[physics.data-an](https://arxiv.org/abs/1105.1160)]. URL: <https://arxiv.org/abs/1105.1160>.
- [117] ATLAS Collaboration. “Measurement of the total and differential cross-sections of $t\bar{t}W$ production in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *Journal of High Energy Physics* 2024.5 (May 2024). ISSN: 1029-8479. DOI: [10.1007/jhep05\(2024\)131](https://doi.org/10.1007/jhep05(2024)131). URL: [http://dx.doi.org/10.1007/JHEP05\(2024\)131](http://dx.doi.org/10.1007/JHEP05(2024)131).
- [118] Matteo Becchetti et al. *NNLO QCD predictions for $t\bar{t}W$ production at hadron colliders*. 2026. arXiv: [2606.09503](https://arxiv.org/abs/2606.09503) [[hep-ph](https://arxiv.org/abs/2606.09503)]. URL: <https://arxiv.org/abs/2606.09503>.
- [119] CMS Collaboration. “Observation of four top quark production in proton-proton collisions at 13 TeV”. In: *Physics Letters B* 847 (Dec. 2023), p. 138290. ISSN: 0370-2693. DOI: [10.1016/j.physletb.2023.138290](https://doi.org/10.1016/j.physletb.2023.138290). URL: <http://dx.doi.org/10.1016/j.physletb.2023.138290>.
- [120] J. Alwall et al. “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”. In: *JHEP* 07 (2014), p. 079. DOI: [10.1007/JHEP07\(2014\)079](https://doi.org/10.1007/JHEP07(2014)079). arXiv: [1405.0301](https://arxiv.org/abs/1405.0301) [[hep-ph](https://arxiv.org/abs/1405.0301)].

- [121] C. Oleari. “The POWHEG BOX”. In: *Nuclear Physics B - Proceedings Supplements* 205-206 (2010), pp. 36–41. ISSN: 0920-5632. DOI: <https://doi.org/10.1016/j.nuclphysbps.2010.08.016>. URL: <http://www.sciencedirect.com/science/article/pii/S0920563210001994>.
- [122] Torbjörn Sjöstrand et al. “An introduction to PYTHIA 8.2”. In: *Computer Physics Communications* 191 (2015), pp. 159–177. ISSN: 0010-4655. DOI: <https://doi.org/10.1016/j.cpc.2015.01.024>. URL: <http://www.sciencedirect.com/science/article/pii/S0010465515000442>.
- [123] Albert M Sirunyan et al. “Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements”. In: *Eur. Phys. J. C* 80.arXiv:1903.12179. CMS-GEN-17-001-003. 1 (Mar. 2019), 4. 61 p. DOI: [10.1140/epjc/s10052-019-7499-4](https://cds.cern.ch/record/2669320). URL: <https://cds.cern.ch/record/2669320>.
- [124] Rikkert Frederix and Stefano Frixione. “Merging meets matching in MC@NLO”. In: *Journal of High Energy Physics* (2012). DOI: [10.1007/JHEP12\(2012\)061](https://doi.org/10.1007/JHEP12(2012)061).
- [125] S. Agostinelli et al. “Geant4—a simulation toolkit”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303. ISSN: 0168-9002. DOI: [https://doi.org/10.1016/S0168-9002\(03\)01368-8](https://doi.org/10.1016/S0168-9002(03)01368-8). URL: <http://www.sciencedirect.com/science/article/pii/S0168900203013688>.
- [126] CMS Collaboration. *Measurements of the pp to $t\bar{t}W$ process at $\sqrt{s}=13$ TeV using the full Run-2 data sample*. CMS AN-2019-127 (CMS internal note).

-
- [127] CMS collaboration. *Measurement of four top quark production in the same-sign dilepton and multilepton channels*. CMS AN-2021-182 (CMS internal note).
- [128] Luca Buonocore et al. *Associated production of a W boson with a top-antitop quark pair: next-to-next-to-leading order QCD predictions for the LHC*. 2023. arXiv: [2306.16311 \[hep-ph\]](#).
- [129] Rikkert Frederix, Davide Pagani, and Marco Zaro. “Large NLO corrections in ttW and $tttt$ hadroproduction from supposedly subleading EW contributions”. In: *Journal of High Energy Physics* 2018.2 (Feb. 2018). DOI: [10.1007/jhep02\(2018\)031](#). URL: <https://doi.org/10.1007/2Fjhep02%282018%29031>.
- [130] CMS Collaboration. *Measurement of the cross section of top quark-antiquark pair production in association with a W boson in proton-proton collisions at $\sqrt{s} = 13$ TeV*. 2022. arXiv: [2208.06485 \[hep-ex\]](#).
- [131] Rikkert Frederix and Ioannis Tsinikos. “On improving NLO merging for ttW production”. In: *Journal of High Energy Physics* 2021.11 (Nov. 2021). DOI: [10.1007/jhep11\(2021\)029](#). URL: <https://doi.org/10.1007/2Fjhep11%282021%29029>.

A Appendix: ECFA ECR panel

This section discusses a survey that was conducted to assess the quality of training in machine learning and instrumentation software within the particle physics community. The work was carried out as part of the author's responsibilities as the Belgian representative on the ECFA Early Career Researcher (ECR) Panel and as a member of the ECFA Training Panel. The section is largely taken over from the corresponding paper [2].

The section is included to illustrate a small part of these activities and to provide an introduction to the role, objectives, and ongoing initiatives of the ECFA panels. While distinct from the scientific work presented elsewhere in this thesis, these efforts contribute to the development of the skills, training opportunities, and research environment of the next generation of particle physicists.

A.1. Introduction

A 2021 study by the ECFA Early-Career Researchers Panel ([arXiv:2107.05739](https://arxiv.org/abs/2107.05739)) revealed that 71% of 334 respondents used open-source software tools in their instrumentation work, yet 70% reported receiving no training for these tools. In response, the Software and Machine Learning for Instrumentation group was formed to assess the accessibility and quality of training programs in machine learning and software for early-career researchers in

experimental and applied physics.

A new survey was published by the Software and Machine Learning for Instrumentation group in 2025 to gain new insights and guide future decision-making around training programs. This survey is the focus of this document and was structured as follows. At the outset, we aimed to gain a general understanding of the respondents; therefore, the first section consisted of broad, general questions designed to characterize the surveyed community. The subsequent section, also general in nature, focused on whether participants were aware of the educational offerings provided by the schools and whether they made use of them. The remaining sections were completed only by participants who expressed interest in specific topics, such as machine learning, detector simulation, DAQ and DCS systems, and detector electronics.

The survey included 174 participants, around 86% of them work in Europe, the remaining 14% came from numerous other nations across Asia, the Americas, and Africa. Regarding career stage, participants included postdoctoral researchers, master's degree holders, senior researchers, university professors, and technical or engineering staff. The percentage distribution of the respondents' scientific career stages can be seen in Figure [A.1a](#). Since the survey was primarily targeted at Early Career Researchers (ECRs), postdoctoral researchers and students make up the majority of the sample. The majority of the surveyed community is involved in large collaborations at the Large Hadron Collider (LHC), each comprising more than 1,000 members, as shown in Figure [A.1b](#). The next largest group, representing 6.5% of the responses, consists of participants in fixed-target experiments at CERN's SPS accelerator, such as NA61/SHINE, COMPASS,

AMBER, and NA62. Respondents affiliated with research groups or experiments outside CERN constitute 8.9% of the sample. These primarily include neutrino experiments such as DUNE, T2K, ENUBET, and Hyper-Kamiokande (HK). A smaller category includes participants involved in Software and Computing initiatives (e.g. GEANT4, ICSC Spoke 3). Responses labeled as *Other* (single entries) correspond to diverse projects such as Belle II, PADME, CTA, HESS, ePIC, and ALTO/CoMET. Finally, 4.1% of the respondents reported participation in projects or experiments currently under preparation, including FCC, HI-BEAM, CODEX-b, QUART&T, LUXE, and ILD.

The respondents' fields of research are shown in Figure A.2a. The majority indicated working in data analysis. Other frequently mentioned areas included software and detector design and development. This distribution reflects the strong computational and analytical focus of the surveyed research community. In the last part of the general questions, we aimed to identify which skills are currently of greatest interest to ECRs; the results can be found in the Figure A.2b. Taking into account the growing popularity of machine learning techniques and the multitude of possibilities of their application, as well as the fact that the respondents mainly work in data analysis, the most popular topics to learn are machine learning and advanced statistical techniques used in data analysis.

Machine learning (ML) techniques have become increasingly important in research across numerous scientific domains, even more so in High Energy Physics (HEP). As the complexity and volume of data continue to grow, the HEP community is turning to ML to enhance data analysis, simulation, and interpretation. However, the pathways through which these scientists acquire

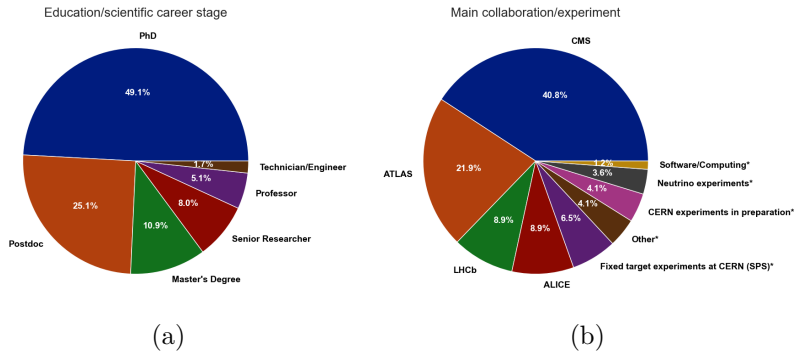
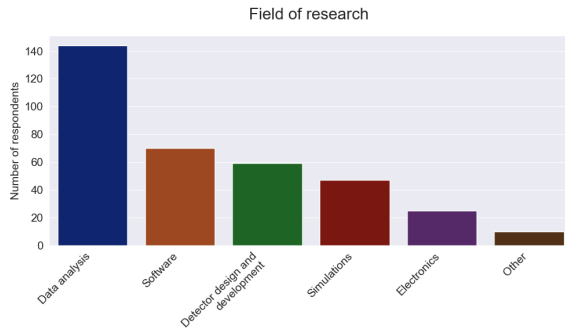
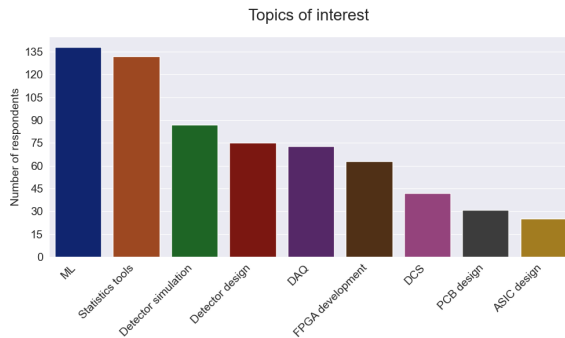


Figure A.1.: The respondents' career stage distribution (a). Name of the experiment in which the respondents take part (b).

ML skills—and their perspectives on the most effective modes of learning—remain underexplored. Understanding how young researchers currently engage with ML, including the resources they use and the challenges they face, is crucial for tailoring educational programs that meet their evolving needs. The insights gained are intended to inform the organizers of future schools and workshops, helping to design targeted, effective training programs that better support the next generation of HEP scientists in mastering machine learning. This section is organized as follows: first, we review the current use cases of machine learning within the community, along with the software tools commonly employed. The second subsection explores the preferences and learning approaches of young researchers regarding ML education. Finally, we discuss how young researchers would organize a well-balanced curriculum for future machine learning schools.



(a)



(b)

Figure A.2.: The respondents' fields of research distribution (a). Hot topics indicated by respondents (b).

A.2. Accessibility and quality of schools

In the first part of the survey, we examined participants' awareness of the available training courses and their accessibility, as well as the quality of the existing schools. We also sought to

identify gaps in the programs and other obstacles that might hinder participation in the schools or workshops. In Figure A.3a we can see that access to information about training and schools remains a significant challenge. More than half of the respondents are unaware of the available training opportunities. Additionally, 7% of respondents indicated that, although they are aware of school offerings, the programs do not meet their expectations. The lack of knowledge may stem from the fact that there are no accessible websites that allow users to easily browse schools and filter them by category or required level of knowledge.

Only 28.7% of respondents have participated in such a school, as shown in Figure A.3b. Among those who attended a school, more than half participated in a machine learning school (56.9%). The second largest group attended Detector Simulation workshops (20.8%), followed by participants in Data Acquisition (DAQ) workshops (11.1%). Satisfaction with the schools is shown in Figure A.4b, with over 72% of respondents reporting a satisfaction level above 70%.

We also provided respondents with an open-ended question, giving them the opportunity to describe what they felt was missing in the school they attended, as well as what they considered to be the main challenges in participating in such training. Although we received only 21 responses out of 174 participants, several recurring themes emerged. Respondents noted that some courses could feel too advanced or fast-paced, with material often assuming prior knowledge that beginners did not always have. A need for more practical, hands-on experience was also frequently mentioned. Several participants felt that the schools were sometimes too theoretical, offering limited examples directly relevant to HEP

analyses or providing short and demanding hands-on sessions.

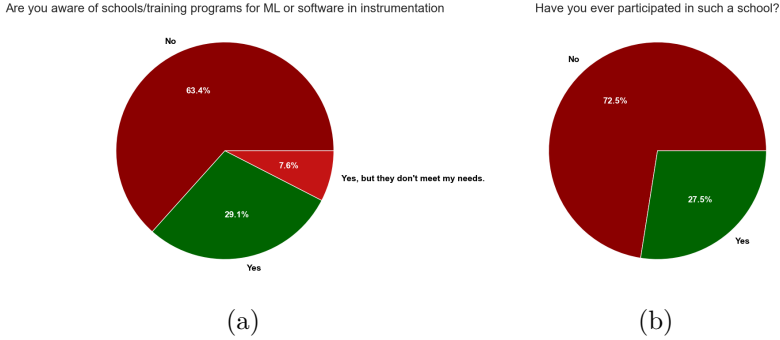


Figure A.3.: Knowledge of available training programs (a). Have participants attended such programs? (b)

Participants additionally expressed interest in more structured learning paths that gradually progress from basic to advanced topics, supported by shorter but more frequent lessons, homework assignments, and more HEP-focused examples. Finally, several participants pointed to time-zone problems and restricted access to computing resources for practice after the course as factors that further limited their ability to fully benefit from the training. While the programs were generally appreciated for the expertise of the instructors and the usefulness of the content, respondents indicated that slower pacing, more complete tutorials, and content tailored to the audience would further enhance the learning experience.

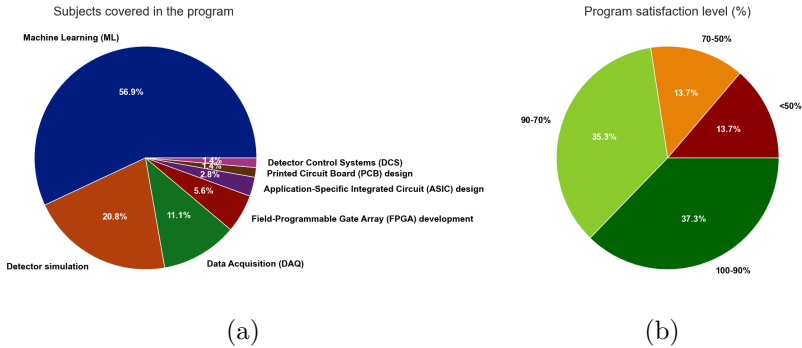


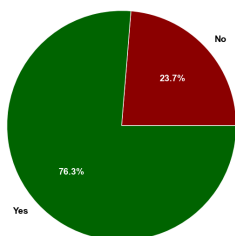
Figure A.4.: Coverage of topics in attended programs (a). Participant satisfaction with training programs (percentage) (b).

A.3. Machine Learning

As can be seen in Figure A.5, a large majority of survey respondents are actively using ML or show interest in the field. Only 2% of people claim to not need any ML tools in their work, highlighting the importance of ML in the field of HEP. More than 90% of respondents would however like to learn more of the subject, while only 6% are satisfied with their current knowledge of the subject. Figure A.6a highlights the main use cases for ML in our field. We see the vast majority of use cases are focused on Classification tasks, both on the event level as well as on the object level (e.g. jet tagging). ML is often also applied on reconstruction tasks such as clustering or tracking but also including other regression tasks. Few respondents use machine learning for other goals. The survey also asked what ML tools we employ in our work. Figure A.6b shows that mainly Python-based ML libraries are

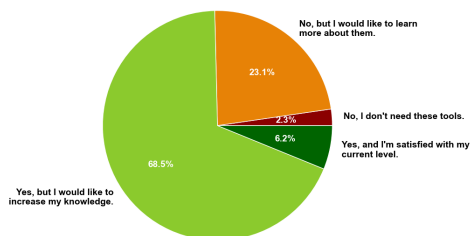
used. While most respondents use PyTorch, the survey shows that the two other well-known frameworks Keras and its base library, TensorFlow, are also often employed. Training for all three libraries will hence remain of use to the community in the near future. We see that ROOT-TMVA, the machine learning library in the ROOT software framework developed at CERN is also used. ML packages that do not include neural network functionalities are less employed. Scikit-Learn and XGBoost are the libraries that are most often used for those cases.

Are you currently involved in projects that utilize ML?



(a)

Have you ever applied ML techniques?



(b)

Figure A.5.: Usage of ML software (a) and knowledge level (b).

Roughly two thirds of respondents learned their ML techniques either independently or by getting help from more experienced colleagues. We see that roughly 13% take courses at their university, while another 13% refined their ML knowledge via online courses. The remaining 9% of respondents have learned via participation to a ML school. While the number of people learning via participation to dedicated schools is quite low, it is higher

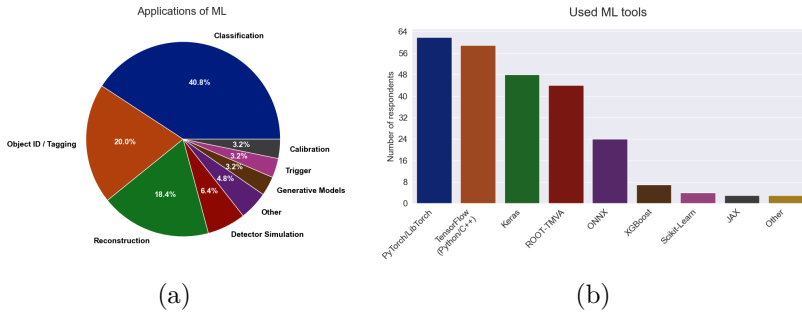


Figure A.6.: The main use cases for ML (a). The ML libraries that are most often used by respondents (b).

than for any of the other tools investigated in the survey. These numbers are highlighted in Figure A.7.

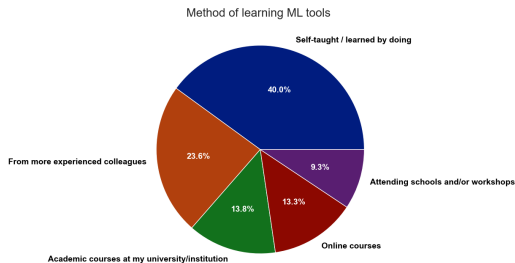


Figure A.7.: Method of learning ML.

When asked what type of training would be more effective for ML, the answers appear very evenly distributed. The most prominent answer, given by 25% of respondents, is having a comprehensive documentation available with instructions and examples. Both organizing schools and organizing shorter workshops receive roughly

21% of the votes each. The remaining 32% of respondents would prefer to have online courses, where 40% of those respondents want to have these courses live. When asked what aspects of the training they would want to receive or what part was missing, most responses were application and hands-on oriented, while the mathematical background and specific integration onto hardware scored lower. These numbers are shown in Figure A.8.

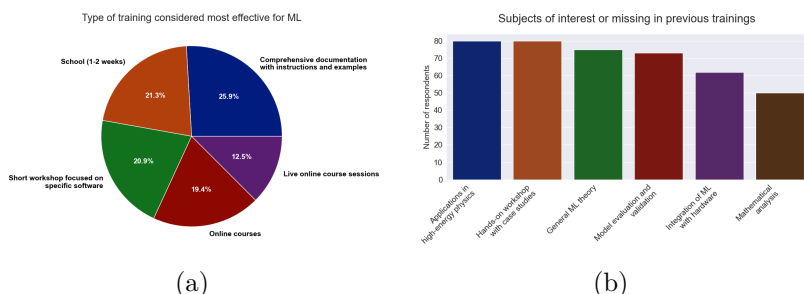


Figure A.8.: The type of training considered most effective for ML (a). Subjects that are of interest to respondents or were underexplored in previous schools (b).

Survey participants were asked to compose their preferred structure for a ML school. They could select 6 topics and attribute a percentage of time to attribute to one subject. All responses were gathered and the average score per topic is shown in Figure A.9. Respondents attributed less time to the theoretical foundations of ML and time to work on team-based projects. Instead, hands-on sessions with experts and workshops on best practices received more priority.

Moving to the second HEP tools discussed in the survey, the experience of the respondents in using detector simulation tools was

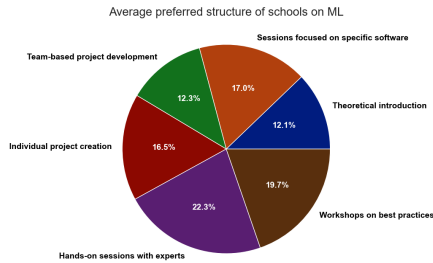


Figure A.9.: The average preferred composition of a course on ML.

investigated. Respondents with detector simulation experience was notably lower at 56%, as shown in Figure A.10a. Notable is however that people with prior experience in detector simulation have a higher probability of being satisfied with their current level of knowledge on the tools, as shown in Figure A.10b. Roughly 6% of respondents report that they do not need these tools.

When asked how people acquired their knowledge on detector simulation tools, we see that roughly 80%, notably more than is the case for ML, is trained by themselves or by experienced coworkers. Less people take courses from their institute, from organised schools or online, as shown in Figure A.11a. Figure A.11b shows that when asked if people have attended a training program or workshop focused on detector simulation before, a large majority have not done so, while 20% has. Another 20% is still planning to, while a large fraction of 42% is not aware of any schools on these subjects.

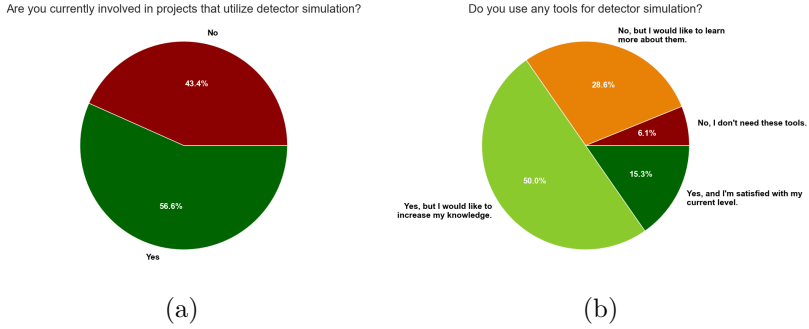


Figure A.10.: Usage of detector simulation software tools (a) and knowledge level (b).

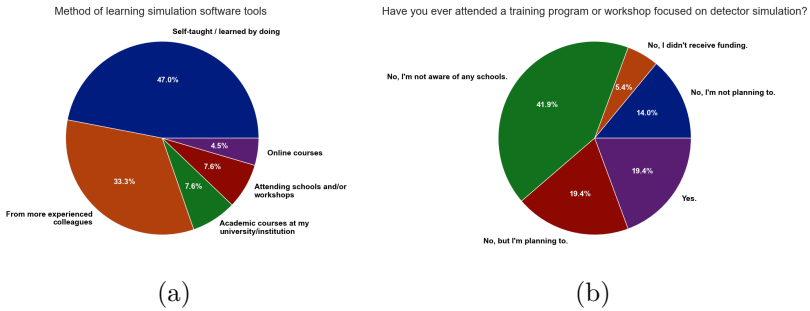


Figure A.11.: Method of learning detector simulation software tools (a). Previous participation to a training program for detector simulation.

When asked what type of training would be beneficial for learning to work with detector simulation tools, 30% of respondents highlight comprehensive documentation with instructions and examples as the most effective tool. This is notably more than was the case for ML tools. Short workshops also get identified

as an effective method 24% of the time, while longer schools get selected as most effective 15% of the time. This seems to indicate that teaching tools for detector simulation software are often preferred to be shorter compared to ML teaching methods. A similar amount (32%) of respondents select online schools as an effective tool to teach detector simulation tools. Less respondents however see the need for these courses to be given live. These numbers are shown in Figure A.12a. When asked which tools respondents would like to learn more about, Geant4 clearly stands out as the most requested software, while other less requested detector simulation tools are also reported in Figure A.12b.

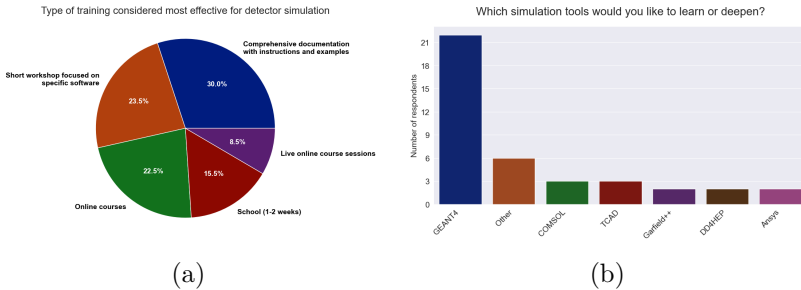


Figure A.12.: The type of training considered most effective for detector simulation (a). Simulation tools that are of interest to respondents or were underexplored in previous schools (b).

Survey participants were asked to compose their preferred structure for a detector simulation software training school. They could select 6 topics and attribute a percentage of time to attribute to one subject. All responses were gathered and the average score per topic is shown in Figure A.13. Respondents attributed less

time to the theoretical foundations of detector simulation and time to work on team-based projects. Instead, hands-on sessions with experts, workshops on best practices and sessions focused on specific software received more priority.

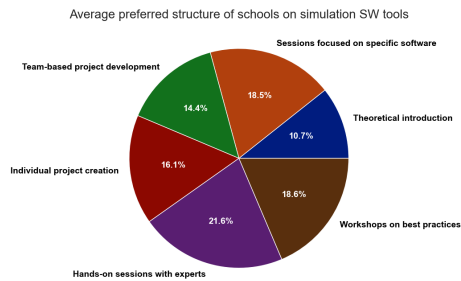


Figure A.13.: The average preferred composition of a course on detector simulation.

B Appendix: Sample lists

ttW

B.1. Data and simulated samples

B.1.1. Data

This analysis uses a set of proton-proton collisions delivered by the LHC at a center-of-mass energy of 13 TeV and collected by the CMS detector. We use the full Run II data set, consisting of data collected in 2016, 2017 and 2018. The so-called “golden-json” files listing the data sections certified for physics analyses are taken from here:

- https://cms-service-dqmdc.web.cern.ch/CAF/certification/Collisions16/13TeV/Legacy_2016/Cert_271036-284044_13TeV_Legacy2016_Collisions16_JSON.txt
- https://cms-service-dqmdc.web.cern.ch/CAF/certification/Collisions17/13TeV/Legacy_2017/Cert_294927-306462_13TeV_UL2017_Collisions17_GoldenJSON.txt
- https://cms-service-dqmdc.web.cern.ch/CAF/certification/Collisions18/13TeV/Legacy_2018/Cert_314472-325175_13TeV_Legacy2018_Collisions18_JSON.txt

In order to collect as many relevant data events as possible, we combine several primary datasets, where overlap between them (i.e. duplicate events) is removed based on run / lumiblock / event number. The relevant datasets are:

- 2016/2017: SingleElectron, SingleMuon, DoubleEG, DoubleMuon and MuonEG
- 2018: EGamma, SingleMuon, DoubleMuon and MuonEG

We use the “ultra-legacy” (UL) reprocessing of all data sets. The specific versions of each dataset are:

- 2016 PreVFP: 21Feb2020_UL2016_HIPM
- 2016 PostVFP: 21Feb2020_UL2016
- 2017: 09Aug2019_UL2017
- 2018: 12Nov2019_UL2018

B.1.2. Simulation

B.1.2.1. Overview

In this analysis, we include Monte-Carlo (MC) simulated samples corresponding to all standard model processes relevant in the phase space of this analysis. Most samples are generated at next-to-leading-order (NLO) precision in QCD, using the MGvATNLO [120] or POWHEG [121] programs. For some cases, in which NLO samples are not available or suffer from low statistics, leading order (LO) samples, generated with MADGRAPH5 [120] are used. The modeling of parton showering, hadronization and the underlying event is performed using PYTHIA [122], with the CP5 tune [123]. Two separate matching schemes are used to avoid double counting of partons generated in the different simulation steps: the FFX [124] scheme for NLO samples, and the MLM [32] scheme for LO samples. The specific samples used to model each process are listed in tables B.1–B.2.

All simulated events are processed by a GEANT4-based [125] simulation of the CMS detector, and in a next step reconstructed using the same algorithms applied to data. For the modeling of additional proton-proton collisions present in data events (pileup), the simulated events are extended with simulated pileup events, and the distribution of the number of additional interactions in the simulated samples is reweighted to match the distribution observed in data.

We use the “ultra-legacy” (UL) production of all simulated samples. The specific versions of each sample are:

- 2016 PreVFP: RunIISummer20UL16MiniAODAPV
- 2016 PostVFP: RunIISummer20UL16MiniAOD
- 2017: RunIISummer20UL17MiniAOD
- 2018: RunIISummer20UL18MiniAOD

The samples in tables B.1–B.2 are grouped per background category. Notice that the samples listed under ‘nonprompt background’ are only used for cross-checking intermediate results; in the final analysis, the nonprompt background is estimated directly from data.

The cross-sections used to scale all simulated samples have been taken from both [126] and [127]. The $t\bar{t}W$ signal samples have however been rescaled to the latest cross-section calculation [128], that will be used as the reference in both CMS and ATLAS, as agreed upon in an LHCtopWG meeting ¹. More details on the $t\bar{t}W$ cross-section are given in the paragraph below.

¹<https://indico.cern.ch/event/1254906/>

B.1.2.2. Signal sample

The default $t\bar{t}W$ sample (TTWJETSToLNu) in table B.1 contains the LO $t\bar{t}W$ production diagrams of order $\alpha_S^2\alpha_{EW}$, as well as NLO contributions in QCD of order $\alpha_S^3\alpha_{EW}$, where α_S and α_{EW} are the strong and electroweak coupling constant respectively. It does however not include electroweak corrections, i.e. terms of the order α_{EW}^3 at ‘LO’ in EWK or $\alpha_S\alpha_{EW}^3$ at NLO [129]. These electroweak correction term are produced and included as an additional sample

(TTWJETSToLNu_5F_EWK). Note that in the previous $t\bar{t}W$ measurement [126, 130], the same procedure was applied, but with a privately produced electroweak corrections sample instead of a centrally produced one (as it was not available at the time). Additionally, we use a privately produced, state-of-the art $t\bar{t}W$ sample with an improved matching procedure tailored specifically to this process [131]. This sample has undergone multiple checks before production. We produced the sample only up to NanoGen level and will use it to compare to our differential predictions. We refer to the Oviedo group (AN-2023-079)² for more details on the private sample production settings and testing.

²<https://icms.cern.ch/tools/publications/notes/entries/AN/2023/079>.

Table B.1.: Simulated signal and background samples and their effective cross sections

Sample name	Cross section [pb]
<i>ttW</i> signal	
TTWJetsToLNu_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8	0.2316
ttWJetsToLNu_5f_EWK_TuneCP5_13TeV_amcatnlo-pythia8	0.01127
TTH	
ttHJetTobb_M125_TuneCP5_13TeV_amcatnloFXFX_madspin_pythia8	0.4040
ttHJetToNonbb_M125_TuneCP5_13TeV_amcatnloFXFX_madspin_pythia8	0.2914974
TTZ	
TTZToLLNuNu_M-10_TuneCP5_13TeV-amcatnlo-pythia8	0.28872
TTZToQQ_TuneCP5_13TeV-amcatnlo-pythia8	0.662
TTZToLL_M-1to10_TuneCP5_13TeV-amcatnlo-pythia8	0.08416
TTG	
TTGamma_Dilept_TuneCP5_13TeV-madgraph-pythia8	2.22
TTGamma_Hadronic_TuneCP5_13TeV-madgraph-pythia8	6.162
TTGamma_SingleLept_TuneCP5_13TeV-madgraph-pythia8	7.509
TGJets_leptonDecays_TuneCP5_13TeV-amcatnlo-pythia8	1.02
TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8	88.40
TTToHadronic_TuneCP5_13TeV-powheg-pythia8	377.96
TTToSemiLeptonic_TuneCP5_13TeV-powheg-pythia8	365.34
TTX	
TTTT_TuneCP5_13TeV-amcatnlo-pythia8	0.01197
TTHH_TuneCP5_13TeV-madgraph-pythia8	0.0006655
TTWH_TuneCP5_13TeV-madgraph-pythia8	0.001141
TTWW_TuneCP5_13TeV-madgraph-pythia8	0.007003
TTWZ_TuneCP5_13TeV-madgraph-pythia8	0.002453
TTZH_TuneCP5_13TeV-madgraph-pythia8	0.00113
TTZZ_TuneCP5_13TeV-madgraph-pythia8	0.001386
TTTJ_TuneCP5_13TeV-madgraph-pythia8	0.0003974
TTTW_TuneCP5_13TeV-madgraph-pythia8	0.0007314
TX	
ST_s-channel_4f_leptonDecays_TuneCP5_13TeV-amcatnlo-pythia8	3.74
ST_t-channel_antitop_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8	136.02
ST_t-channel_top_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8	80.95
ST_tW_antitop_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8	21.585204
ST_tW_top_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8	21.585204
THQ_ctcvcv_4f_Hincl_TuneCP5_13TeV_madgraph_pythia8	0.07096
tZq_l_4f_ckm_NLO_TuneCP5_13TeV-amcatnlo-pythia8	0.07358
TWZToLL_thad_Wlept_5f_DR_TuneCP5_13TeV-amcatnlo-pythia8	0.003
TWZToLL_tlept_Whad_5f_DR_TuneCP5_13TeV-amcatnlo-pythia8	0.003
TWZToLL_tlept_Wlept_5f_DR_TuneCP5_13TeV-amcatnlo-pythia8	0.0015
THW_ctcvcv_5f_Hincl_TuneCP5_13TeV_madgraph_pythia8	0.01561
WZ	
WZTo1L1Nu2Q_TuneCP5_13TeV-amcatnloFXFX-pythia8	11.66
WZTo3LNu_TuneCP5_13TeV-amcatnloFXFX-pythia8	4.9173
WZJJ_EWK_InclusivePolarization_TuneCP5_13TeV_madgraph-madspin-pythia8	0.01629

B. Appendix: Sample lists ttW

Table B.2.: Simulated signal and background samples and their effective cross sections (part 2)

Sample name	Cross section [pb]
ZZ	
ZZTo2L2Nu_TuneCP5_13TeV_powheg_pythia8	0.6204
ZZTo4L_TuneCP5_13TeV_powheg_pythia8	1.256
GluGluHToZZTo4L_M125_TuneCP5_13TeV_powheg2_JHUGenV7011_pythia8	0.01212
GluGluToContinToZZTo2e2mu_TuneCP5_13TeV-mcfm701-pythia8	0.005423
GluGluToContinToZZTo2e2nu_TuneCP5_13TeV-mcfm701-pythia8	0.005423
GluGluToContinToZZTo2e2tau_TuneCP5_13TeV-mcfm701-pythia8	0.005423
GluGluToContinToZZTo2mu2nu_TuneCP5_13TeV-mcfm701-pythia8	0.005423
GluGluToContinToZZTo2mu2tau_TuneCP5_13TeV-mcfm701-pythia8	0.005423
GluGluToContinToZZTo4e_TuneCP5_13TeV-mcfm701-pythia8	0.0027
GluGluToContinToZZTo4mu_TuneCP5_13TeV-mcfm701-pythia8	0.0027
GluGluToContinToZZTo4tau_TuneCP5_13TeV-mcfm701-pythia8	0.0027
VBF_HToZZTo4L_M125_TuneCP5_13TeV_powheg2_JHUGenV7011_pythia8	0.001034
VHToNonbb_M125_TuneCP5_13TeV-amcatnloFXFX_madspin_pythia8	0.9561
ZG	
ZGToLLG_01J_5f_TuneCP5_13TeV-amcatnloFXFX-pythia8	55.48
DYJetsToLL_M-10to50_TuneCP5_13TeV-madgraphMLM-pythia8	15810.0
DYJetsToLL_M-50_TuneCP5_13TeV-amcatnloFXFX-pythia8	6077.22
Multiboson	
WWTo1L1Nu2Q_TuneCP5_13TeV-amcatnloFXFX-pythia8	45.68
WWTo2L2Nu_TuneCP5_13TeV-powheg-pythia8	11.09
WWTo4Q_4f_TuneCP5_13TeV-amcatnloFXFX-pythia8	45.2
WWTo2L2Nu_TuneCP5_DoubleScattering_13TeV-pythia8	0.2232
WWW_4F_TuneCP5_13TeV-amcatnlo-pythia8	0.2086
WWZ_4F_TuneCP5_13TeV-amcatnlo-pythia8	0.1651
WZZ_TuneCP5_13TeV-amcatnlo-pythia8	0.05565
ZZZ_TuneCP5_13TeV-amcatnlo-pythia8	0.01476

C Appendix: Triggers

C.0.1. triggers used in the $t\bar{t}W$ analysis

Table C.1.: Overview of 2016 triggers used to select good event for this analysis.

Category	Trigger
m e	HLT_IsoMu24, HLT_IsoTkMu24, HLT_Mu50, HLT_TkMu50 HLT_Mu45_eta2p1
	HLT_Ele27_WPTight_Gsf
	HLT_Ele105_CaloIdVT_GsfTrkIdT
	HLT_Ele115_CaloIdVT_GsfTrkIdT HLT_Photon175
mm	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL(_DZ)
	HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL(_DZ)
	HLT_TkMu17_TrkIsoVVL_TkMu8_TrkIsoVVL(_DZ)
	HLT_Mu30_TkMu11
em	HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL(_DZ)
	HLT_Mu23_TrkIsoVVL_Ele8_CaloIdL_TrackIdL_IsoVL(_DZ)
	HLT_Mu30_Ele30_CaloIdL_GsfTrkIdVL
	HLT_Mu33_Ele33_CaloIdL_GsfTrkIdVL
ee	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ
	HLT_DoubleEle33_CaloIdL_GsfTrkIdVL
	HLT_DoubleEle33_CaloIdL_GsfTrkIdVL_MW
mmm	HLT_TripleMu_12_10_5
emm	HLT_DiMu9_Ele9_CaloIdL_TrackIdL
eem	HLT_Mu8_DiEle12_CaloIdL_TrackIdL
eee	HLT_Ele16_Ele12_Ele8_CaloIdL_TrackIdL

Table C.2.: Overview of 2017 triggers used to select good event for this analysis.

Category	Trigger
m	HLT_IsoMu24
	HLT_IsoMu24_eta2p1
	HLT_IsoMu27
	HLT_Mu50
	HLT_OldMu100
	HLT_TkMu100
e	HLT_Ele32_WPTight_Gsf
	HLT_Ele35_WPTight_Gsf
	HLT_Ele115_CaloIdVT_GsfTrkIdT
	HLT_Photon200
mm	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL
	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ
	HLT_Mu37_TkMu27
	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass8
	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8
em	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ
	HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ
	HLT_Mu12_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ
	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL
	HLT_Mu27_Ele37_CaloIdL_MW
	HLT_Mu37_Ele27_CaloIdL_MW
ee	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL
	HLT_DoubleEle33_CaloIdL_MW
mmm	HLT_TripleMu_10_5_5_DZ
	HLT_TripleMu_5_3_3_Mass3p8to60_DZ
	TripleMu_12_10_5
emm	HLT_DiMu9_Ele9_CaloIdL_TrackIdL_DZ
eem	HLT_Mu8_DiEle12_CaloIdL_TrackIdL
eee	HLT_Ele16_Ele12_Ele8_CaloIdL_TrackIdL

Table C.3.: Overview of 2018 triggers used to select good event for this analysis.

Category	Trigger
m	HLT_IsoMu24
	HLT_IsoMu27
	HLT_Mu50
	HLT_OldMu100
	HLT_TkMu100
e	HLT_Ele32_WPTight_Gsf
	HLT_Ele115_CaloIdVT_GsfTrkIdT
	HLT_Photon200
mm	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8
	HLT_Mu37_TkMu27
em	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ
	HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ
	HLT_Mu12_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ
	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL
	HLT_Mu27_Ele37_CaloIdL_MW
	HLT_Mu37_Ele27_CaloIdL_MW
ee	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL
	HLT_DoubleEle25_CaloIdL_MW
mmm	HLT_TripleMu_10_5_5_DZ
	HLT_TripleMu_5_3_3_Mass3p8to60_DZ
	TripleMu_12_10_5
	HLT_TripleMu_5_3_3_Mass3p8to60_DZ
emm	HLT_DiMu9_Ele9_CaloIdL_TrackIdL_DZ
eem	HLT_Mu8_DiEle12_CaloIdL_TrackIdL
eee	HLT_Ele16_Ele12_Ele8_CaloIdL_TrackIdL

C.0.2. Orthogonal triggers for the $t\bar{t}W$ analysis

Table C.4.: Reference triggers for the orthogonal trigger method.

Year	Trigger
2016	HLT_MET200, HLT_PFMET300
	HLT_PFMET170_HBHECleaned
	HLT_PFMET120_PFMHT120_IDTight
	HLT_PFHT300_PFMET110, HLT_PFHT900
	HLT_PFHT350_DiPFJetAve90_PFAIalphaT0p53
	HLT_PFHT400_DiPFJetAve90_PFAIalphaT0p52
	HLT_PFHT400_SixJet30_DoubleBTagCSV_p056
	HLT_PFHT650_WideJetMJJ900DEtaJJ1p5
2017	HLT_CaloJet500_NoJetID
	HLT_PFJet500, HLT_AK8PFJet500
	HLT_PFMET140_PFMHT140_IDTight
	HLT_PFHT500_PFMET100_PFMHT100_IDTight
	HLT_PFHT700_PFMET85_PFMHT85_IDTight
	HLT_PFHT800_PFMET75_PFMHT75_IDTight
2018	HLT_CaloJet500_NoJetID
	HLT_CaloMET350_HBHECleaned
	HLT_AK8PFJet500, HLT_AK8PFJet400_TrimMass30
	HLT_DiJet110_35_Mjj650_PFMET110
	HLT_PFHT800_PFMET75_PFMHT75_IDTight
	HLT_PFHT700_PFMET85_PFMHT85_IDTight
	HLT_PFHT500_PFMET100_PFMHT100_IDTight
	HLT_PFHT1050, HLT_PFJet500
	HLT_PFMET120_PFMHT120_IDTight
	HLT_PFMET250_HBHECleaned
	HLT_PFMET200_HBHE_BeamHaloCleaned
	HLT_PFMETTypeOne140_PFMHT140_IDTight
	HLT_PFMETTypeOne200_HBHE_BeamHaloCleaned
	HLT_TripleJet110_35_35_Mjj650_PFMET110

C.0.3. triggers used in the fake-rate measurement for the $t\bar{t}W$ analysis

Table C.5.: Overview of triggers used for the fake rate measurement

Year	Trigger
2016	HLT_Mu3_PFJet40
	HLT_Mu8
	HLT_Mu17
	HLT_Mu20
	HLT_Mu27
	HLT_Ele8_CaloIdM_TrackIdM_PFJet30
	HLT_Ele12_CaloIdM_TrackIdM_PFJet30
	HLT_Ele17_CaloIdM_TrackIdM_PFJet30
2017, 2018	HLT_Ele23_CaloIdM_TrackIdM_PFJet30
	HLT_Mu3_PFJet40
	HLT_Mu8
	HLT_Mu17
	HLT_Mu20
	HLT_Mu27
	HLT_Ele8_CaloIdM_TrackIdM_PFJet30
	HLT_Ele17_CaloIdM_TrackIdM_PFJet30
	HLT_Ele23_CaloIdM_TrackIdM_PFJet30