

Measurement of the differential dileptonic $t\bar{t}$ cross section in a BSM phase space with the CMS detector using the full LHC Run 2 data set

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften der RWTH
Aachen University zur Erlangung des akademischen Grades eines Doktors der
Naturwissenschaften genehmigte Dissertation

vorgelegt von

Danilo Meuser, M. Sc. RWTH

aus

Aachen

Berichter: **Univ-Prof. Dr. rer. nat. Lutz Feld**
Univ-Prof. Dr. rer. nat. Michael Krämer

Tag der mündlichen Prüfung: 8. Dezember 2023

Diese Dissertation ist auf den Internetseiten der Universitätsbibliothek verfügbar.



Kurzfassung

Das Top-Quark ist das schwerste bekannte elementare Teilchen, das im Standardmodell der Teilchenphysik (SM) beschrieben wird. In Proton-Proton (pp) Kollisionen wird es überwiegend in Quark-Antiquark-Paaren ($t\bar{t}$) produziert. Messungen des entsprechenden Wirkungsquerschnitts liefern wichtige Tests des SMs, sowie Überprüfungen von Theorien, die Physik jenseits des SMs (BSM) beschreiben. In dieser Arbeit werden differentielle Messungen des Wirkungsquerschnitts für den dileptonischen $t\bar{t}$ -Prozess vorgestellt. Dabei liegt der Fokus auf einem Phasenraum, in dem zusätzliche Beiträge von BSM-Physik gefunden werden könnten. Die Messungen basieren auf pp-Kollisionen bei einer Schwerpunktsenergie von 13 TeV, die mit dem CMS Detektor zwischen 2016 und 2018 während Run 2 des LHCs aufgenommen wurden und einer integrierten Luminosität von 138 fb^{-1} entsprechen. Die Ergebnisse werden durch die Kombination der e^+e^- , $\mu^+\mu^-$ und $e^\pm\mu^\mp$ Zerfallskanäle bestimmt. Die in dieser Arbeit betrachteten BSM-Szenarien umfassen supersymmetrische Modelle und Modelle zur Dunklen Materie, die jeweils Endzustände ähnlich dem des dileptonischen $t\bar{t}$ -Prozesses vorhersagen und durch zwei geladenen Leptonen, zwei b-Jets und einem signifikanten fehlenden Transversalimpuls, der durch nicht-detektierte Teilchen entsteht, gekennzeichnet sind. Die Messung solch komplexer Signaturen basiert auf einer korrekten Rekonstruktion der Teilchen im Endzustand, was unter anderem ein gutes Alignment der Silizium-Sensoren im CMS-Spurdetektor erfordert. Im Rahmen dieser Arbeit werden Performance-Studien des Online-Alignments sowie eine Erweiterung der Alignment-Methode vorgestellt.

Im Gegensatz zu früheren Messungen, bei denen der differentielle Wirkungsquerschnitt hauptsächlich als Funktion kinematischer Variablen der Leptonen oder Top-Quarks bestimmt wurde, konzentriert sich die vorliegende Analyse auf zwei Observablen, die potenzielle BSM-Beiträge vom dileptonischen $t\bar{t}$ -Prozess des SMs trennen. Auf Detektorebene sind diese Observablen durch den fehlenden Transversalimpuls und den minimalen azimuthalen Abstand zwischen dem fehlenden Transversalimpuls und einem Lepton gegeben. Um die Sensitivität der Analyse zu erhöhen, werden zusätzliche Korrekturen des fehlenden Transversalimpulses basierend auf multivariaten Methoden, die im Rahmen dieser Arbeit entwickelt wurden und die entsprechende Auflösung im gesuchten Endzustand deutlich verbessern, angewandt. Im ersten Schritt der Analyse werden auf Detektorebene die gemessenen pp-Kollisionen mit Vorhersagen des Standardmodells basierend auf Monte-Carlo (MC) Simulationen hinsichtlich der eindimensionalen Verteilungen und der zweidimensionalen Kombination beider Observablen verglichen. Für alle drei Verteilungen kann eine gute Übereinstimmung zwischen den Daten und der SM-Vorhersage ohne signifikante Anzeichen für mögliche Beiträge von BSM-Physik beobachtet werden. Im zweiten Schritt werden absolute und normierte differentielle Wirkungsquerschnitte auf Teilchenebene durch die Entfaltung der drei Verteilungen auf Detektorebene bestimmt. Hierbei werden die Wirkungsquerschnitte als Funktion des Transversalimpulses des Dineutrino-Systems und des minimalen azimuthalen Abstands zwischen dem Dineutrino-System und einem Lepton, die für den dileptonischen $t\bar{t}$ -Prozess den Observablen auf Detektorebenen entsprechen, gemessen. Dies ist die erste Messung des $t\bar{t}$ -Wirkungsquerschnitts in Abhängigkeit dieser beiden Observablen am LHC. Um die Auswirkungen verschiedener Beschreibungen des Untergrunds von Single-Top-Produktion mit einem W-Boson (tW) in Simulation zu studieren, werden die Ergebnisse je basierend auf dem Diagram Removal (DR) and Diagram Subtraction (DS) Schema produziert. Die Ergebnisse werden mit zwei MC-Theorievorhersagen, die auf Matricelementen mit NLO-Genauigkeit basieren, sowie zwei Theorie-Rechnungen, die NLO- und NNLO-Genauigkeit entsprechen, verglichen. Alle vier Theorievorhersagen beschreiben die gemessenen absoluten und normierten Wirkungsquerschnitte hinreichend gut, wobei die Messungen des Dineutrino-Impulses die tW -Beschreibung basierend auf dem DS-Schema bevorzugen.

Abstract

The top quark is the heaviest of all known elementary particles described by the standard model of particle physics (SM). In proton-proton (pp) collisions it is mainly produced in quark-antiquark pairs ($t\bar{t}$). Measurements of the corresponding $t\bar{t}$ production cross section yield important precision tests of the SM, while also probing theories describing physics beyond the SM (BSM). In this thesis, differential cross section measurements of the SM dileptonic $t\bar{t}$ process in a phase space, where additional contributions from BSM physics could be found, are presented. The measurements are performed for pp collisions at a center-of-mass energy of 13 TeV, recorded by the CMS experiment between 2016 and 2018 during Run 2 of the LHC, corresponding to an integrated luminosity of 138 fb^{-1} . The results are derived by combining the e^+e^- , $\mu^+\mu^-$ and $e^\pm\mu^\mp$ decay channels. The BSM scenarios considered in this thesis include supersymmetric and dark matter models, which predict final states similar to the dileptonic $t\bar{t}$ process, characterized by two charged leptons, two b jets and significant missing transverse momentum produced by undetected particles. Measurements of such complex signatures require a proper reconstruction of the final state particles, which, among others, relies on a good alignment of the silicon sensors in the CMS tracking system. Within the scope of the thesis performance studies of the online tracker alignment and an improvement of the alignment method are presented.

Unlike previous cross section measurements, where differential results were mainly derived as a function of kinematic variables of the leptons or top quarks, the present analysis focuses on two observables, which separate potential BSM contributions from the SM dileptonic $t\bar{t}$ process. On detector level, these observables are given by the missing transverse momentum and the minimal azimuthal distance between the missing transverse momentum and a lepton. In order to increase the sensitivity of the analysis, additional corrections of the missing transverse momentum based on machine learning techniques are developed within this thesis, which significantly improve the corresponding response and resolution in the final state of interest. In the first step of the measurements, the observed pp-collision data and SM predictions based on Monte Carlo (MC) simulations are compared on detector level based on the one-dimensional distributions as well as the two-dimensional combination of both detector-level observables. For all three distributions, good agreement between the recorded data and the SM predictions with no significant indications for potential BSM contributions is observed. In the second step, absolute and normalized differential cross sections on particle level are derived by unfolding the three detector-level distributions. The cross sections are obtained as a function of the transverse momentum of the dineutrino system and the minimal azimuthal distance between the dineutrino system and a lepton, which match the detector-level observables for the dileptonic $t\bar{t}$ process. So far, differential $t\bar{t}$ cross sections as a function of these two observables have not been measured at the LHC. To study the impact of different descriptions for the background contribution from single top production in association with a W boson (tW) in simulation, two sets of results based on the diagram removal (DR) and the diagram subtraction (DS) scheme are provided. The results are compared to two MC-based theory predictions using NLO matrix element accuracy as well as two fixed-order calculations, which correspond to NLO and NNLO accuracy in QCD. All four theory predictions provide reasonable descriptions of the measured absolute and normalized cross section spectra, while the measurements of the dineutrino momenta seem to prefer the tW description based on the DS scheme.

Contents

1	Introduction	1
2	Theoretical foundation and motivation	5
2.1	Standard model of particle physics	6
2.1.1	Elementary particles	6
2.1.2	Electromagnetic interactions	8
2.1.3	Strong interactions	8
2.1.4	Electroweak unification	9
2.1.5	Electroweak symmetry breaking	11
2.1.6	Shortcomings and open questions	12
2.2	Physics beyond the standard model	14
2.2.1	Supersymmetry	14
2.2.2	Dark Matter	16
2.3	Top quark physics	17
2.3.1	Production and decay	17
2.3.2	Top quarks and BSM physics	18
2.3.3	Cross section measurement in a BSM phase space	19
3	The CMS experiment at the LHC	23
3.1	The Large Hadron Collider	23
3.2	The Compact Muon Solenoid detector	26
3.2.1	Coordinate system	27
3.2.2	Solenoid magnet	28
3.2.3	Tracking system	28
3.2.4	Electromagnetic calorimeter	30
3.2.5	Hadron calorimeter	32
3.2.6	Muon system	34
3.2.7	Trigger system	36
4	Alignment of the CMS tracker	37
4.1	Track-based alignment	37
4.1.1	Alignment algorithm	38
4.1.2	Validation methods	39
4.2	Impact of radiation effects	41
4.3	Automated alignment in the Prompt Calibration Loop	43
4.3.1	Low-granularity alignment for Run 2	43
4.3.2	High-granularity alignment for Run 3	44
5	Data sets and Monte Carlo simulation	51
5.1	Triggers and primary data sets	51
5.2	Monte Carlo simulation	53
5.2.1	Hard process	53
5.2.2	Parton shower, hadronization and underlying event	54
5.2.3	Detector simulation	55
5.2.4	Simulated samples	56
5.2.5	Interference between tW and $t\bar{t}$	58
5.3	Trigger efficiencies	60
5.3.1	Efficiency measurement	60
5.3.2	Scale factors	61

6	Event reconstruction and selection	65
6.1	Particle flow algorithm	65
6.2	Physics object reconstruction	67
6.2.1	Primary vertex	68
6.2.2	Electrons	68
6.2.3	Muons	70
6.2.4	Jets	71
6.2.5	Heavy-flavor jets	72
6.2.6	Missing transverse momentum	73
6.3	Event selection	77
6.3.1	Total yields	78
6.3.2	Control distributions	79
7	Deep neural network regression for p_T^{miss} correction	83
7.1	Phase space dependencies of the p_T^{miss} response	84
7.2	Machine learning and neural networks	85
7.2.1	Mathematical foundation	86
7.2.2	Hyperparameters	88
7.3	Initial architecture and training process	90
7.3.1	Training sample reweighting	91
7.4	Input features	92
7.4.1	Impact ordering	93
7.4.2	Goodness-of-fit tests	94
7.5	Hyperparameter optimization	96
7.6	Performance studies	98
8	Cross section extraction	103
8.1	Fiducial phase space	103
8.2	Binning optimization	104
8.2.1	Binning requirements	104
8.2.2	One-dimensional binning	106
8.2.3	Two-dimensional binning	107
8.3	Unfolding procedure	110
8.3.1	Signal extraction	110
8.3.2	Unfolding with TUnfold	111
8.4	Validation tests	113
9	Systematic uncertainties	119
9.1	Experimental sources	119
9.2	Theoretical sources	123
9.3	Impact on differential measurements	125
9.4	Definition of covariance matrices	127
10	Results	129
10.1	One-dimensional cross sections	130
10.1.1	Measurements in $p_T^{\nu\nu}$	130
10.1.2	Measurements in $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$	133
10.2	Two-dimensional cross sections	137
11	Conclusions and outlook	141
	Appendix 1: Signal and baseline triggers	147
	Appendix 2: Primary data sets	149

Appendix 3: Simulated samples	151
Appendix 4: Trigger efficiencies	153
Appendix 5: Total event yields	159
Appendix 6: Control distributions	163
Appendix 7: DNN input features	167
Appendix 8: DNN performance	169
Appendix 9: Unfolding validation tests	171
Appendix 10: Uncertainties for normalized results	175
Appendix 11: Uncertainties for results using the DS scheme	177
Appendix 12: BSM content in detector-level distributions	181
References	183

Chapter 1

Introduction

Studies concerning the principles and constitution of the matter surrounding us have been pursued for centuries, starting with philosophical approaches made by the ancient Greeks. The modern field of particle physics, however, was not born before the late 19th and early 20th centuries, when J. J. Thomson discovered the electron in 1897 [1] and Ernest Rutherford proved that atoms contain a small and dense nucleus in 1911 [2]. The field gained momentum with the development of the first particle accelerators around the 1930s and particle colliders in the 1950s [3], which enabled the discovery of a plethora of new particles. This evolution was accompanied by the proposal of the electroweak (EWK) theory in the 1960s [4–7] and quantum chromodynamics (QCD) in the 1970s [8, 9], which led to the standard model of particle physics (SM), still providing the best understanding of elementary particles and their interactions, to date. One of the main achievements of the SM has been connected to the prediction of elementary particles before their actual discovery, as shown by the top quark, predicted in 1973 [10] and observed at the Tevatron in 1995 [11, 12], or the Higgs boson, the last missing elementary particle in the SM, which was predicted in 1964 [13, 14] and observed at the LHC in 2012 [15, 16].

Despite the large success of the SM, proven by various experiments over the last half century, the SM fails to provide solutions, among others, for the origin of dark matter [17] or the nature of neutrino masses [18–22]. These shortcomings call for extensions of the SM, also known as BSM physics, which have been numerous proposed over the last decades. Such extensions include the theory of supersymmetry (SUSY), predicting new particles based on an additional symmetry between fermions and bosons, or theories, which expand the spacetime by additional spatial dimensions to provide dark matter candidates.

Searches for such extensions of the SM as well as precision measurements of the SM itself are the primary purpose of the Large Hadron Collider (LHC) [23], which is the largest and most powerful particle collider currently operating. The LHC, situated at the European Center of Nuclear Research (CERN) near Geneva, collides proton beams at center-of-mass energies up to 13.6 TeV. Among other experiments, two multi-purpose detectors, namely the ATLAS (A Toroidal LHC Apparatus) [24] and the CMS (Compact Muon Solenoid) [25] detector, are located at collision points of the LHC aiming to record and study the proton-proton (pp) collisions. Ever since the start of the LHC operation in 2010, both experiments were able to record increasingly larger data sets at increasingly higher collision energies, which led to the discovery of the Higgs boson in 2012 [15, 16] or the measurement of the top-quark mass at an unprecedented precision in 2023 [26].

Most analyses of the pp-collision data, recorded by one of the multi-purpose detectors, can be roughly split into two types of analyses. Precision measurements of the SM study a specific

SM process, such as $t\bar{t}$ production or the Drell-Yan process, covering phase spaces, that offer a large abundance of the process of interest. Searches for BSM physics in contrast usually focus on specific event signatures selecting phase space regions, where potential signals from BSM physics could exceed the contributions from SM processes. In case no significant excess is found, which, so far, holds for all BSM searches conducted at the LHC, limits for specific BSM scenarios are derived. One example of a combination of multiple BSM searches aiming for SUSY signatures, which include photons in the final state, can be found in [27]. This combination is based on the author's Master's thesis [28] and has been published as a CMS publication within the scope of this Ph.D. thesis.

The main analysis performed in the present thesis aims to combine the primary aspects of both types of analyses focusing on signatures with a top and an anti-top quark in the final state, which are mainly produced by SM top-pair ($t\bar{t}$) production but are also well motivated by various extensions of the SM. This combination of aspects is realized by relying on observables, which separate potential BSM contributions from the SM $t\bar{t}$ process, similar to BSM searches. However, rather than focusing only on specific phase space regions, the entire spectrum of the given observables is measured. Thus, if no significant BSM contributions are observed, the analysis yields a precise measurement of the SM $t\bar{t}$ process in terms of a differential cross section rather than improving the limits for different BSM scenarios.

The analysis focuses on the dileptonic decay channel, where both W bosons from the top quark decays decay leptonically. Hence, the final state of interest contains two oppositely-charged leptons, two neutrinos, which lead to significant missing transverse momentum p_T^{miss} , as well as two jets, which predominantly originate from b quarks. The BSM scenarios considered in the analysis lead to similar final states, except for the presence of additional BSM particles, which, similar to the neutrinos, escape the detector without detection. The choice of observables employed in the present analysis is based on the presence of these additional BSM particles. The first observable is given by missing transverse momentum itself, which, compared to the SM $t\bar{t}$ process, receives additional contributions due to the BSM particles. The second observable is the minimal azimuthal distance $\min[\Delta\phi(p_T^{\text{miss}}, \ell)]$ between the missing transverse momentum and a lepton. Given that the leptons and neutrinos originate from the same W bosons, the directions of the missing transverse momentum and the leptons are correlated. This correlation, however, is significantly reduced in the BSM scenarios due to the additional BSM particles.

The analysis is performed on pp collisions recorded by the CMS experiment at a center-of-mass energy of 13 TeV between 2016 and 2018 during the period known as Run 2 of the LHC. The analyzed data set corresponds to a total integrated luminosity of 138 fb^{-1} , containing roughly 100 million $t\bar{t}$ events. To obtain the final results of the present analysis, three lepton decay channels, given by the e^+e^- , $\mu^+\mu^-$ and $e^\pm\mu^\mp$ channel, are combined. Following the strategy sketched above, the detector-level distributions of both observables are compared in the first step of the analysis between the observed data and SM prediction derived from Monte Carlo (MC) simulation. The same comparison is performed for the two-dimensional combination of both observables, increasing the sensitivity to potential BSM contributions. In order to improve the resolution of the missing transverse momentum, which is subject to resolution effects of all reconstructed particles in an event and impacts both detector-level observables, a dedicated correction method based on a deep neural network is developed within the scope of this thesis.

If no significant indications for BSM contributions are found, all three detector-level distributions are unfolded to a fiducial phase space (particle level) after the contributions from other SM processes are removed. Based on these unfolded distributions, normalized and absolute differential cross sections of the dileptonic $t\bar{t}$ process in one and two dimensions as a function of the transverse momentum $p_T^{\nu\nu}$ of the dineutrino system and the minimal azimuthal

distance $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ between the dineutrino system and a lepton, which match the observables on detector-level, are derived. Differential cross sections of the dileptonic $t\bar{t}$ process using these two types of observables have, so far, not been measured at the LHC. The results obtained within the scope of this thesis are labeled as "CMS Work in progress", which marks results not yet published in a journal paper but endorsed by the CMS collaboration for use in a Ph.D. thesis. The differential cross section results are compared to two theory predictions based on MC simulations using next-to-leading order (NLO) matrix element accuracy. In addition, two fixed-order calculations, which are based on the methods described in [29] and correspond to NLO and next-to-next-to-leading order (NNLO) accuracy in QCD, have been provided by theory colleagues [30] and are compared to the obtained results.

This thesis is structured as follows. A short introduction of the standard model of particle physics including shortcomings and potential extensions of the SM as well as a brief review of top quark physics and a discussion about the strategy of this analysis are given in Chapter 2. In Chapter 3, the main aspects of the LHC are discussed, followed by a description of the CMS experiment and its subdetectors. Special attention is paid to the alignment of the CMS tracker, in Chapter 4, where the performance of the automated alignment during Run 2 of the LHC as well as improvements of the alignment procedure for Run 3, developed within the scope of this thesis, are discussed. In Chapter 5, the recorded pp-collision data sets and the corresponding triggers used to collect the data are summarized. In addition, an introduction to MC simulations and the simulated samples used throughout the analysis, as well as a dedicated estimation of the trigger efficiencies are presented. The reconstruction of the physics objects in an event, including charged leptons, (heavy-flavor) jets and different algorithms to obtain the missing transverse momentum, is covered by Chapter 6. This part is followed by the introduction of a dedicated event selection aiming for signatures expected from the dileptonic $t\bar{t}$ process as well as first comparisons between the recorded data and MC simulation. In Chapter 7, an additional correction to the missing transverse momentum based on a deep neural network, developed within the scope of this thesis, is discussed. Here, studies of potential input features with respect to their impact on the correction and their description by simulation as well as a dedicated hyperparameter optimization and the performance of the final corrections are presented. The extraction of differential cross sections based on detector-level distributions is discussed in Chapter 8. This part includes the definition of the fiducial phase space, a dedicated optimization of the binnings used for the final measurements, as well as an introduction of the unfolding procedure and various validation tests of different unfolding approaches. In Chapter 9, an overview of the different experimental and theoretical contributions to the systematic uncertainty of the final cross section results including the definition of the covariance matrix is presented. The results of this analysis including comparisons on detector level between the observed data and the MC-based predictions as well as the differential cross section measurements and comparisons to theory predictions, are shown and discussed in Chapter 10. Finally, the thesis is concluded by a summary and a brief outlook given in Chapter 11.

Chapter 2

Theoretical foundation and motivation

2.1	Standard model of particle physics	6
2.1.1	Elementary particles	6
2.1.2	Electromagnetic interactions	8
2.1.3	Strong interactions	8
2.1.4	Electroweak unification	9
2.1.5	Electroweak symmetry breaking	11
2.1.6	Shortcomings and open questions	12
2.2	Physics beyond the standard model	14
2.2.1	Supersymmetry	14
2.2.2	Dark Matter	16
2.3	Top quark physics	17
2.3.1	Production and decay	17
2.3.2	Top quarks and BSM physics	18
2.3.3	Cross section measurement in a BSM phase space	19

The standard model of particle physics (SM) is the main theory to describe all known elementary particles and their interactions. It has been tested across several energy scales by many different independent experiments studying the properties of the elementary particles as well as the interactions covered by the SM. Some of the most stringent tests of the SM are performed at particle colliders such as the Large Electron-Positron Collider (LEP) [31], the Tevatron [32] and the Large Hadron Collider (LHC) [23], where in 2012 the CMS [16] and the ATLAS collaboration [15] discovered the Higgs boson, the last missing elementary particle predicted by the SM. While the vast majority of the experimental measurements conducted so far are in precise agreement with the predictions of the SM, there are indications, e.g., from cosmological measurements as well as aspects of the SM itself which call for extensions of the standard model.

In the following, an overview of the different ingredients of the standard model is given in Sec. 2.1. In Sec. 2.2, shortcomings and potential extensions of the model are discussed. The chapter is completed by a short review of top quark physics, a discussion about the role of top quarks in relation to potential extensions of the SM as well as a brief introduction of the strategy applied in the present analysis in Sec. 2.3.

2.1 Standard model of particle physics

In the standard model of particle physics [4–7, 33–35], the fundamental forces responsible for the interactions between all known elementary particles are described based on a renormalizable quantum field theory. In this theory, each of the forces results from the local gauge invariance under a given symmetry group. The strong force arises from the local invariance under $SU(3)$ transformation. The electromagnetic force and the weak force are combined in the SM into a consistent gauge theory through the electroweak (EWK) unification, which is based on the local invariance under $SU(2) \times U(1)$. Thus, the full symmetry group of the SM is given by

$$SU(3) \times SU(2) \times U(1) . \quad (2.1)$$

In order to enable massive mediators for the EWK interactions while ensuring local gauge invariance the corresponding $SU(2) \times U(1)$ symmetry is spontaneously broken in the SM by introducing a scalar Higgs field, which also accounts for the masses of the remaining elementary particles via Yukawa couplings.

2.1.1 Elementary particles

The elementary particles of the SM, summarized in Fig. 2.1, can be divided based on their spin into fermions, which correspond to the constituents of matter, and bosons, which mediate interactions between elementary particles. The group of fermions is characterized by half-integer spin and includes a total of twelve particles and their corresponding antiparticles, which differ only by the sign of the additive quantum numbers¹. Fermions can be further divided into quarks, which are strongly interacting particles, and leptons, which do not interact via the strong interaction.

The group of leptons includes a total of six particles, which all interact via the weak force. In addition, three leptons (electron e , muon μ and tau τ) are electrically charged and therefore also participate in electromagnetic interactions. Each of the charged leptons is associated with an electrically neutral lepton (electron neutrino ν_e , muon neutrino ν_μ and tau neutrino ν_τ) forming a generation resulting in a total of three lepton generations. While the charged leptons acquire mass due to their Yukawa coupling to the Higgs fields, which is further discussed in Sec. 2.1.5, neutrinos are predicted to be massless in the SM. Due to the different magnitudes of the Yukawa couplings, which are free parameters in the SM, the charged leptons cover a large mass range from the sub-MeV range for the electron to more than a GeV for the tau lepton.

Unlike leptons, quarks also interact via strong interactions and therefore participate in all three interactions covered by the SM. Similar to leptons, the six quarks can be grouped into three generations, where each generation includes an up-type quark (u , c , t) and a down-type quark (d , s , b). Both quark types differ in their electrical charge with up-type and down-type quarks carrying a charge of $Q = +2/3 e$ and $Q = -1/3 e$, respectively. In addition to the electrical charge, quarks also carry the charge related to the strong interaction known as color charge. Similar to the charged leptons, all quarks acquire their mass through Yukawa interactions with the Higgs field covering a range of around 2 MeV for the up quark to around 173 GeV for the top quark, which is the heaviest known elementary particle.

As mentioned above, leptons and quarks can be grouped into generations given by doublets of the weak isospin $T = 1/2$, which corresponds to the generator of the $SU(2)$ group of the SM. The components of each doublet differ in the third component of the isospin, which equals $+1/2$ for

¹In the following, no distinction between particles and antiparticles is made and different chirality states are neglected, unless otherwise stated.

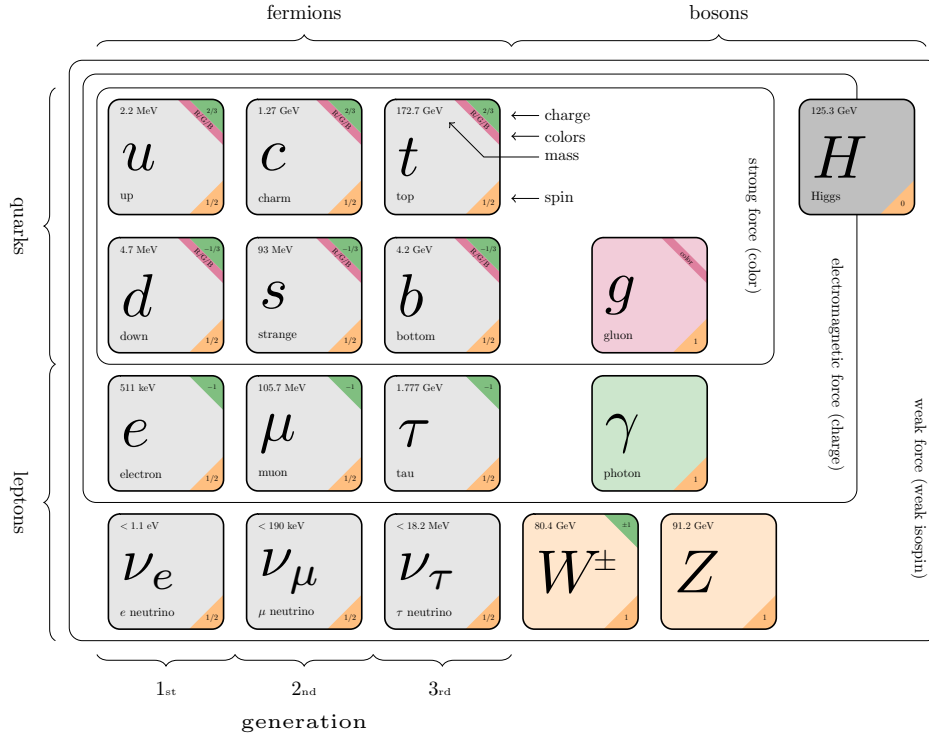


Figure 2.1: The elementary particles of the standard model grouped into fermions and bosons [36]. The three generations of fermions are further divided into quarks and leptons. For each particle, the spin and, if applicable, the electric charge, the color charge and the mass [37] are given.

neutrinos and up-type quarks and $-1/2$ for charged leptons and down-type quarks, neglecting right-handed contributions which are further discussed in Sec. 2.1.4. By introducing the weak hypercharge Y , which is the generator of the SM $U(1)$ group and equals -1 for leptons and $+1/3$ for quarks, again neglecting right-handed contributions, the electric charge of a particle can be expressed by the Gell-Mann-Nishijima relation [38, 39]

$$Q = T_3 + \frac{1}{2}Y, \quad (2.2)$$

where T_3 is the third component of the weak isospin and the electric charge Q is given in units of the elementary charge e .

The bosons carrying integer spin form the second group of elementary particles in the SM. Each of the three fundamental forces described in the SM can be connected to the exchange of a corresponding spin-1 gauge boson. Electromagnetic interactions are mediated by the photon γ , while weak interactions are related to exchanges of the Z boson or the $W^\pm$ bosons, where the latter are electrically charged. The strong force is mediated by eight gluons carrying different color-anticolor combinations. While the photon and the gluons are massless, the three weak gauge bosons acquire mass as a result of the Higgs mechanism. The Higgs boson, which corresponds to the only scalar elementary particle described in the standard model, is an additional consequence of this mechanism and completes the group of bosons.

2.1.2 Electromagnetic interactions

Electromagnetic interactions are one out of the three fundamental interactions described in the SM. The interactions are mediated by massless photons between electrically charged particles. In the standard model, electromagnetic interactions are defined by the theory of quantum electrodynamics (QED). As for all interactions covered by the SM, the interaction itself is a consequence of demanding local gauge invariance under a specific symmetry transformation. In case of QED, these symmetry transformations are given by the abelian group $U(1)$. For a free Dirac particle, which describes massive spin-1/2 particles like the electron, the corresponding Lagrangian

$$\mathcal{L}_\psi = \bar{\psi}(i\gamma_\mu\partial^\mu - m)\psi, \quad (2.3)$$

where ψ is the Dirac spinor, γ_μ are the Pauli matrices and m is the mass of the particle, is not invariant under local $U(1)$ transformation given by $\psi \rightarrow \psi' = \exp[-i\alpha(x)]\psi$. The local gauge invariance can be restored by introducing a new gauge field A^μ , the photon field, which couples to ψ leading to the full Lagrangian of QED

$$\mathcal{L}_{\text{QED}} = \bar{\psi}(i\gamma_\mu\partial^\mu - m)\psi - e\bar{\psi}\gamma_\mu\psi A^\mu - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \quad (2.4)$$

The second term of the Lagrangian describes the interaction between the Dirac particle and the gauge field with a coupling strength equal to the electromagnetic charge e , while the third term corresponds to the free Lagrangian of the free gauge field A^μ including the electromagnetic strength tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. Since a potential mass term for the gauge field proportional to $A_\mu A^\mu$ would break the local invariance of the Lagrangian, the photon has to be massless. As a consequence, electromagnetic interactions, unlike interactions mediated by massive bosons, have an infinite range.

2.1.3 Strong interactions

In the standard model, strong interactions between quarks and gluons are described by the theory of quantum chromodynamics (QCD). The interactions are mediated by massless and electrically neutral gluons, which couple to color charge. While (anti-)quarks carry only one type of (anti-)color given by red, green or blue, gluons must carry a color-anticolor combination to conserve the color-charge at QCD vertices.

Similar to the concept introduced for QED, in QCD, the interactions between color-charged particles and the mediators of the force are a direct consequence of demanding local gauge invariance under specific symmetry transformations. In case of QCD, these transformations are given by the eight generators T^a of the $SU(3)$ group. In the resulting QCD Lagrangian

$$\begin{aligned} \mathcal{L}_{\text{QCD}} &= \bar{q}(i\gamma_\mu\partial^\mu - m)q - g_s(\bar{q}\gamma_\mu T^a q)G_a^\mu - \frac{1}{4}G_{\mu\nu}^a G_a^{\mu\nu} \\ &\text{with } G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a - g_s f_{bc}^a G_\mu^b G_\nu^c, \end{aligned} \quad (2.5)$$

where f_{bc}^a denotes the structure constants of the $SU(3)$ group, the first and the second term can be interpreted analog to the QED Lagrangian describing a free quark q and a coupling of the strength g_s between the quark and the eight gauge fields G_a^μ , representing the gluons, respectively. The main difference to the QED Lagrangian corresponds to the structure of the

field strength tensor $G_{\mu\nu}^a$, which includes a self-interaction between the gluons given by the term proportional to $G_{\mu}^b G_{\nu}^c$. This additional part originates from the non-abelian character of the $SU(3)$ group, where the generators of the group do not commute. As a consequence not only vertices between two quarks and a gluon, but also three-gluon and four-gluon vertices are allowed in QCD. In addition, the strong coupling constant given by

$$\alpha_s = \frac{g_s^2}{4\pi} \sim \left[\log \left(\frac{Q^2}{\Lambda^2} \right) \right]^{-1} \quad (2.6)$$

where Λ is the QCD scale of around 200 MeV, is not constant but a function of the energy scale Q of the interaction. This so-called running of the coupling, which is a consequence of α_s satisfying the renormalization group equation (RGE), has been precisely measured by multiple experiments as shown in Fig. 2.2.

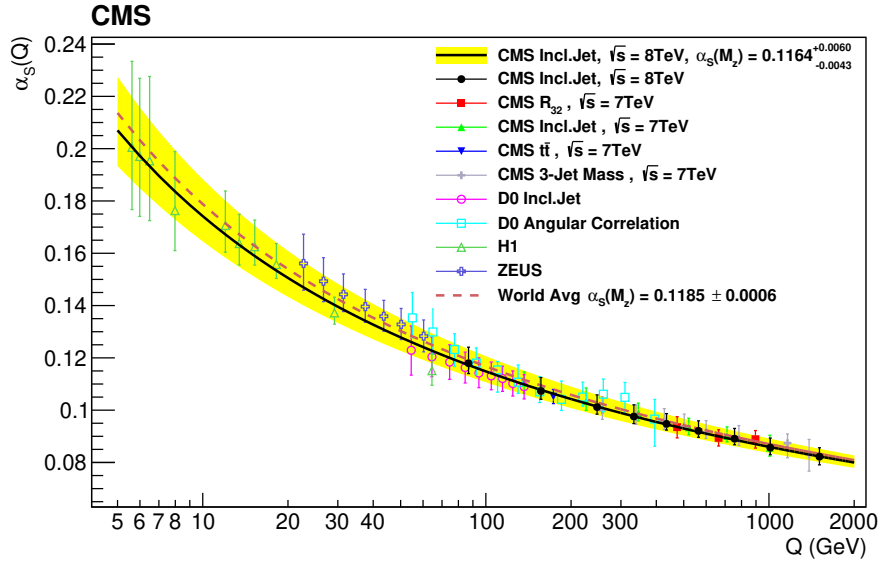


Figure 2.2: Measurements of the strong coupling constant α_s as a function of the momentum transfer Q [40] for several different experiments. The dashed red line shows the world average of the strong coupling for $Q = m_Z$.

At high energies ($Q \gg \Lambda$), corresponding to small distances, the coupling becomes weak enabling perturbative calculations in α_s . This effect is also known as asymptotic freedom of the QCD. At small energies ($Q \approx \Lambda$) or large distances on the other side, the effect of color confinement can be observed. Here, the coupling becomes strong and the quarks form color-neutral bound states also known as hadrons. Hadrons can be grouped into mesons, which are combinations of quark and anti-quark, and baryons consisting of three quarks or anti-quarks. As a consequence of color confinement, free quarks or gluons have not been observed so far, but can be detected based on their hadronization also known as hadronic showers or jets.

2.1.4 Electroweak unification

In the standard model, the electromagnetic force and the weak force are described by a unified theory also known as electroweak (EWK) unification. This unified theory is based on the $SU(2) \times U(1)$ gauge group, which is generated by the weak hypercharge operator Y connected

to $U(1)$ and the three weak isospin operators T_i for the $SU(2)$ part (cf. Eq. (2.2)). The corresponding gauge fields for $U(1)$ and $SU(2)$ are given by B_μ and W_μ^i , respectively. Similar to the two interactions introduced before, the coupling between the gauge fields and the fermions is introduced by demanding local gauge invariance under the given symmetry transformations, which yields the following electroweak Lagrangian (neglecting mass terms):

$$\begin{aligned} \mathcal{L}_{\text{EWK}} = & \bar{L}\gamma^\mu(i\partial_\mu - gT_iW_\mu^i - g'\frac{Y}{2}B_\mu)L + \bar{R}\gamma^\mu(i\partial_\mu - g'\frac{Y}{2}B_\mu)R \\ & - \frac{1}{4}W_{\mu\nu}W^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}. \end{aligned} \quad (2.7)$$

While for QED and QCD there is no distinction between the left-handed spinor component L and the right-handed spinor component R , here both parts couple to the B_μ gauge field but only the left-handed part couples to the W_μ^i gauge fields leading to parity violation in weak interactions. The strength of these couplings is given by g and g' . The field strength tensor $B_{\mu\nu}$ connected to the $U(1)$ part shows the same structure as the strength tensor introduced in QED (cf. Sec. 2.1.2). For the tensor $W_{\mu\nu}$ corresponding to the non-abelian $SU(2)$ group a self-interacting part similar to QCD is introduced. At this point, potential mass terms for the gauge fields proportional to $W_\mu W^\mu$ would break the local gauge invariance of Lagrangian. Only by introducing spontaneous symmetry breaking, which is further discussed in the next section, the gauge bosons can acquire mass, which matches the experimental observations of massive weak gauge bosons [41–44], while keeping the local gauge invariance.

Another consequence of the spontaneous symmetry breaking is the mixing of the gauge fields to mass eigenstates, which can be identified with the mediators of the electromagnetic force (γ) and the weak force ($W^\pm$ and Z):

$$W^\pm = \frac{1}{\sqrt{2}}(W^1 \mp W^2) \quad \text{and} \quad \begin{pmatrix} \gamma \\ Z \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B \\ W^3 \end{pmatrix}. \quad (2.8)$$

The mixing for the neutral mediators is defined by the Weinberg angle θ_W , which is a free parameter of the standard model. Based on this mixing, charged current interactions via $W^\pm$ only occur between left-handed fermions leading to a maximal parity violation. For neutral current interactions via the Z boson, the parity violation is smaller due to contributions from the B_μ gauge field. For electromagnetic interactions there is no parity violation since based on $g\sin\theta_W = g'\cos\theta_W$ the coupling to the photon is the same for left- and right-handed fermions.

While for leptons charged current interactions are in the SM only described between neutrinos and charged leptons of the same generation (cf. Fig. 2.1), for quarks interactions between up- and down-type quarks of different generations can occur. This effect can be explained by the fact that for quarks the weak isospin eigenstates do not correspond to the mass eigenstates. The mixing, also known as quark mixing, is applied to the down-type quarks and is described by the Cabbibo-Kobayashi-Maskawa matrix (CKM matrix) [10, 45], which connects the weak states (d' , s' , b') to the mass eigenstates (d , s , b):

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \quad (2.9)$$

The elements of the matrix are free parameters of the standard model. Due to the large diagonal elements of the mixing matrix charged current interactions within the same generation are still preferred for quarks. This holds especially for the top quark, which almost exclusively decays to a bottom quark due to $|V_{tb}| \approx 1$.

2.1.5 Elektroweak symmetry breaking

Potential mass terms for the weak gauge bosons would break the local gauge invariance of the corresponding Lagrangian, as discussed in the previous section. In order to generate masses for the gauge bosons while still keeping the gauge invariance intact the so-called Brout-Englert-Higgs mechanism [13, 14], which spontaneously breaks the electroweak symmetry, is introduced. For this mechanism, a new $SU(2)$ isospin doublet ϕ of complex scalar fields, also called Higgs field, with hypercharge $Y = 1$ and potential $V(\phi)$ is introduced:

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \quad \text{and} \quad V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2. \quad (2.10)$$

For $\lambda > 0$ and $\mu^2 < 0$, the potential has a "Mexican hat" shape, where the potential itself is symmetric under the $SU(2)$ and $U(1)$ transformation, but the ground state $\phi_0 \neq 0$ is not, which spontaneously breaks the symmetry.

The interactions between the weak gauge fields and the Higgs field are again a consequence of maintaining local gauge invariance under $SU(2)$ and $U(1)$ which yields the following Lagrangian:

$$\mathcal{L} = \mathcal{L}_{\text{EWK}} + \underbrace{\left| \left(i\partial_\mu - gT_i W_\mu^i - g' \frac{Y}{2} B_\mu \right) \phi \right|^2}_{|D_\mu \phi|^2} - V(\phi), \quad (2.11)$$

where $\mathcal{L}_{\text{EWK}}$ corresponds to Eq. (2.7). Choosing a ground state, with $\phi_1 = \phi_2 = \phi_4 = 0$ and $\phi_3^2 = -\mu^2/\lambda = v^2$ spontaneously breaks this invariance and leads to the desired behavior of the Higgs field. Since the third component of the weak isospin T_3 equals $-1/2$, the electric charge of the field is zero based on Eq. (2.2). Thus, there is no coupling between the photon and the Higgs field keeping the photon massless. By inserting the ground state into the new interaction term of Eq. (2.11) and applying Eq. (2.8) the mass terms of the EWK gauge bosons can be identified:

$$|D_\mu \phi_0|^2 = \underbrace{\frac{1}{8} \nu^2 g^2 [(W_\mu^1)^2 + (W_\mu^2)^2]}_{M_W^2 W_\mu^+ W_\mu^-} + \underbrace{\frac{1}{8} \nu^2 [gW_\mu^3 - g'B_\mu]^2}_{\frac{1}{2} M_Z^2 Z_\mu Z^\mu} + \underbrace{0 [gW_\mu^3 + g'B_\mu]^2}_{\frac{1}{2} M_A^2 A_\mu A^\mu}, \quad (2.12)$$

with the mass of the EWK gauge bosons given by $M_W = \nu g/2$ and $M_Z = \nu \sqrt{g^2 + g'^2}/2$ and the mass of the photon M_A equal to zero.

Out of the four additional degrees of freedom introduced by the isospin doublet ϕ , three degrees, also known as Goldstone bosons, can be absorbed into the polarization of the gauge bosons by choosing a specific gauge. The only remnant of the Higgs field is then given by the Higgs boson, which has a mass of $M_H = \sqrt{-2\mu^2}$ as denoted in the mass term of the potential $V(\phi)$.

While the Lagrangian in Eq. (2.11) does only describe the couplings between the Higgs boson and the gauge bosons as well as the Higgs self coupling given by the cubic and quartic terms of the Higgs potential, the same mechanism as described above can be applied to introduce mass terms for the fermions. The corresponding couplings between fermions and the Higgs field are known as Yukawa couplings. Similar to the interaction between the Higgs field and the gauge bosons, the strength of these couplings is proportional to the corresponding particle masses. This dependency has been confirmed by the comparison of different coupling measurements as shown in Fig. 2.3, where the measured strengths of all observed couplings between massive particles and the Higgs field are summarized.

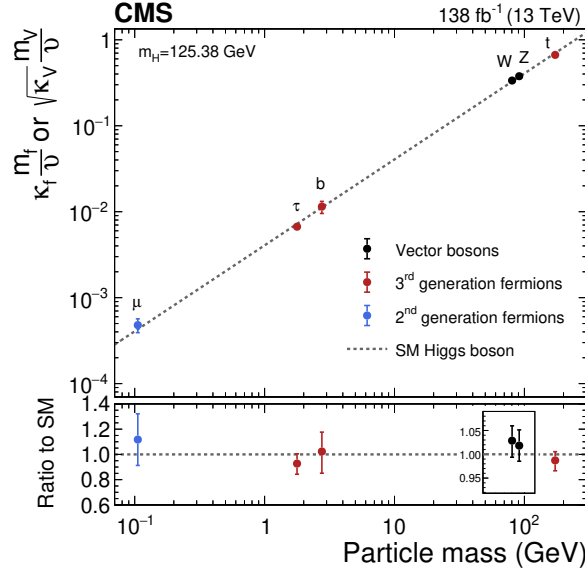


Figure 2.3: Coupling strengths of the Higgs boson to fermions and massive gauge bosons measured by the CMS experiment [46]. The coupling strengths given in terms of the coupling modifiers for fermions κ_f and for bosons κ_V , where ν is the vacuum expectation value of the Higgs field, are shown as a function of the particle mass. The dashed line shows the prediction by the standard model given by $\kappa_f = \kappa_V = 1$. The ratio between the measurements and the SM prediction is shown in the lower part.

So far the couplings of the Higgs field with the EWK gauge bosons as well as with the two heaviest leptons and quarks have been observed or strong evidence has been found [46]. All measured couplings are well in agreement with predictions of the standard model showing the connection between particle mass and coupling strength and confirming the last missing piece of the standard model.

2.1.6 Shortcomings and open questions

So far multiple experimental tests and precision measurements have confirmed the standard model. Over the last decades, various particles have been observed only after their prediction by the standard model. Examples of such particles are the Higgs boson, discussed in the previous section, or the c quark, which was proposed in 1970 [47] and measured at SLAC and BNL in 1974 [48, 49]. Despite this large success of the model, there are some striking experimental results which can not be explained by the SM as well as flaws in the model itself. To motivate the need for extensions of the standard model, some of these shortcomings are briefly discussed in the following.

Dark Matter

The existence of non-baryonic matter, so-called dark matter (DM), in our universe has been confirmed by various astrophysical results [17]. One of the strongest evidence for the existence of dark matter is based on the measurement of rotation curves of galaxies [50, 51]. An example of such a measurement is shown in Fig. 2.4. From Newtonian dynamics, one would expect the rotational velocities inside the visible mass distribution first to increase with increasing distance R and then to decrease outside of the visible mass as shown by the contribution of the stellar disk. However, the measured rotational velocities show an almost flat behavior outside the visible mass, which matches an additional contribution expected from a dark matter halo inside the galaxy.

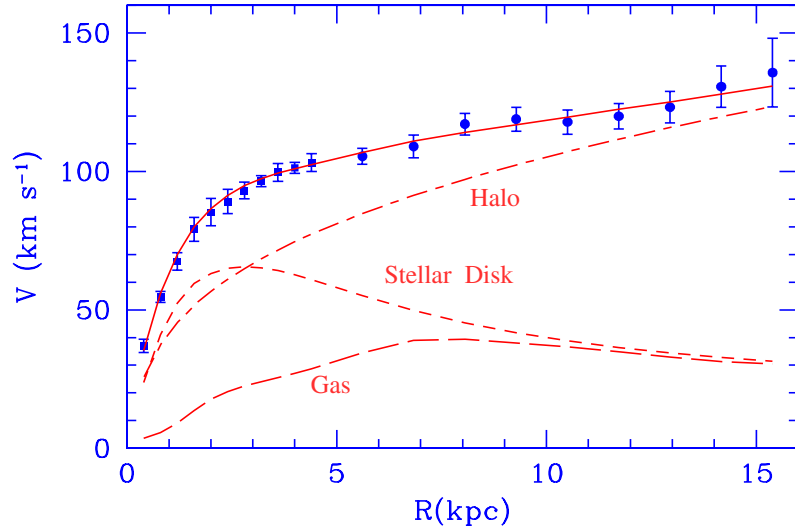


Figure 2.4: Measured rotational velocities as a function of the distance to the galactic center given in kpc $\approx 3 \cdot 10^{19}$ m of the M33 galaxy [51]. In addition, the best fit (continuous line) as well as contributions from the DM halo, the stellar disk and gas are shown.

Further evidence for the existence of dark matter can be found in observations of gravitational lensing [52, 53] and measurements of bullet clusters [54], which consist of two colliding galaxy clusters. From measurements of the Cosmic Microwave Background (CMB) [55] and analyses of the large-scale structures of the universe [56], it can be concluded that around 85% of the total matter of the universe is not made of ordinary baryonic matter but consists of dark matter. Since dark matter does not interact via the electromagnetic interaction, the only potential particle of the standard model to account for DM is the neutrino. However, neutrinos can only account for a small part of DM. The neutrino mass is too small to be the only source of the entire dark matter in the universe. In addition, neutrinos can only contribute to so-called hot dark matter consisting of light relativistic particles, which due to constraints from early universe structure formation can only account for a fraction of the DM [57].

Hierarchy problem

Due to the coupling between the Higgs field and massive particles, discussed in Sec. 2.1.5, the mass of the Higgs boson m_H receives quantum corrections [58]. For a Dirac fermion, whose

coupling to the Higgs field H is given by $-\lambda_f H \bar{f} f$ in the Lagrangian, the Higgs boson mass obtains a correction of

$$\Delta m_H^2 = -\frac{|\lambda_f|^2}{8\pi^2} \Lambda_{UV}^2 + \dots, \quad (2.13)$$

where λ_f corresponds to the strength of the Yukawa coupling and Λ_{UV} is an energy cutoff scale up to which the SM is assumed to be valid. The largest corrections are introduced by the coupling between the top quark and the Higgs field, where the corresponding Yukawa coupling is almost unity. Assuming that the SM is valid up to the Planck scale of around 10^{19} GeV, the Higgs mass would receive corrections, which are tens of orders of magnitude larger than the Higgs mass itself. In order to still end up with a Higgs mass of around 125 GeV the corrections to the mass would need to almost perfectly cancel each other. This cancellation, which is considered to be unnatural, is also known as the hierarchy problem of the standard model.

Neutrino masses

During the last two decades, several experiments have observed neutrino oscillations for solar, atmospheric, reactor as well as accelerator neutrinos [18–22]. Since these oscillations can only occur if there is a mass difference between the neutrino mass eigenstates, the neutrinos themselves cannot be massless as assumed by the standard model. Although the neutrino masses are small compared to the remaining massive SM particles, with current neutrino mass limits below 1 eV [59], an extension of the SM is required to account for the mass of the neutrinos.

An example of such an extension is the addition of right-handed sterile neutrinos to the SM. Due to their sterile nature, these additional particles do not interact via any of the three fundamental interactions of the SM. However, including right-handed neutrinos would allow for Yukawa couplings between the neutrinos and the Higgs field introducing a Dirac mass for the neutrinos similar to the mechanism assumed for charged leptons. In addition, neutrinos could also be their own antiparticle (Majorana particle), which would introduce an additional mass term in the Lagrangian [60]. This assumption is used in the so-called seesaw mechanism, in which the left-handed neutrinos can acquire small masses of the order of eV, while the right-handed neutrino masses get large. Based on this mechanism the large mass difference between neutrinos and the remaining massive SM particles could be explained.

2.2 Physics beyond the standard model

As discussed in the previous section, there are multiple experimental results as well as shortcomings of the standard model itself, which call for extensions of the SM also known as physics beyond the SM or BSM physics. Over the last decades, numerous extensions have been proposed by theoretical physicists [61]. In the following, supersymmetry as one of the most famous extensions as well as different models explaining dark matter are briefly introduced.

2.2.1 Supersymmetry

One of the most elegant extensions of the standard model, which accounts for some of the shortcomings introduced in Sec. 2.1.6, is the theory of supersymmetry (SUSY) [58, 61, 62]. In this theory, the particle content of the SM is extended by assigning an additional supersymmetric partner to each elementary particle of the SM, together forming a supermultiplet. The particles inside such a supermultiplet share the same quantum numbers except for their spin, which differs

by half a unit. Thus, the supersymmetric extension effectively introduces a symmetry between bosonic and fermionic states by adding an operator $\mathbf{Q}$ which transforms bosonic into fermionic states and vice versa. This operator commutes with the generators of the remaining SM gauge transformations keeping all quantum numbers except for the spin unchanged.

The symmetry is also represented in the naming scheme of the additional SUSY particles, which are usually marked with a tilde. The names of the superpartners of the fermions are built with a prepending "s", e.g., smuon and stop. For the bosonic superpartners, an "ino" is appended to the corresponding SM particle name, e.g., higgsino.

The extension introduced above adding only one superpartner for each SM particle corresponds to the Minimal Supersymmetric Standard Model (MSSM). The complete particle content of the MSSM is shown in Fig. 2.5. In order to account for gravitation an additional graviton can be added in the MSSM. Furthermore, the Higgs sector has to be extended to two Higgs doublets to avoid gauge anomalies and to introduce the correct Yukawa couplings. These doublets correspond to H_u , which is responsible for the up-type quark masses, and H_d , which introduces masses to the down-type quarks as well as to the charged leptons. Due to EWK symmetry breaking these doublets mix into five Higgs scalar mass eigenstates including the SM Higgs boson. In a similar fashion the superpartners of the SM gauge eigenstates, namely bino $\tilde{B}$, winos $\tilde{W}$, and higgsinos $\tilde{H}$, mix to mass eigenstates. These eigenstates are grouped by their electrical charge into charginos $\tilde{\chi}_{1,2}^\pm$ and neutralinos $\tilde{\chi}_{1-4}^0$, where higher indices correspond to larger masses.

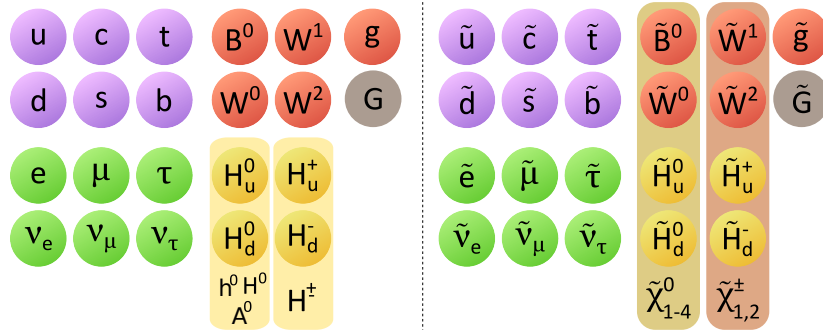


Figure 2.5: Particle content of the MSSM. The left side shows the standard model content. The SM particles are grouped into quarks (purple), leptons (green), gauge eigenstates (red), and two Higgs doublets (yellow), which mix to mass eigenstates. The corresponding supersymmetric partners are shown on the right. Gauginos (red) and higgsinos (yellow) mix to the mass eigenstates $\tilde{\chi}_{1-4}^0$ and $\tilde{\chi}_{1,2}^\pm$.

While all fermions introduce a correction to the Higgs boson mass given by Eq. (2.13), the coupling between the Higgs field H and a scalar boson S with a term $-\lambda_S |H|^2 |S|^2$ in the Lagrangian introduces a correction of

$$\Delta m_H^2 = \frac{\lambda_S}{16\pi^2} \Lambda_{UV}^2 + \dots \quad (2.14)$$

This correction exactly cancels the contribution from the fermions assuming $|\lambda_f|^2 = \lambda_S$ holds for a given supermultiplet and taking into account that each fermion is accompanied by two scalars due to the two helicity states of the fermion. Thus, the supersymmetric extension of the SM introduced above would solve the hierarchy problem.

In general, a supersymmetric extension of the SM does not forbid baryon number B and lepton number L violating terms in the Lagrangian, which could, for example, allow for proton decays [58]. Since there are strong constraints on the stability of the proton [63] as well as on other processes violating the baryon or lepton number [37], the effects of these violating terms have to be either negligible or B and L have to be conserved in SUSY. Postulating the conservation of the so-called R -parity

$$R = (-1)^{3B+L+2S}, \quad (2.15)$$

would be one way to enforce the B and L conservation in a supersymmetric extension. Taking the different spin S of the superpartners into account, the R -parity yields $+1$ for SM particles and -1 for SUSY particles. As a consequence of R -parity conservation SUSY particles would be only produced in pairs and the lightest supersymmetric particle (LSP) would be stable. If the LSP is electrically neutral and interacts only via EWK interactions, it could serve as a dark matter candidate and thus solve the dark matter problem of the SM introduced earlier.

Since so far no supersymmetric particles have been experimentally observed up to particle masses of a few TeV [37], the masses of the SUSY particles have to be significantly larger than the masses of the corresponding SM partners. Such a mass difference would follow if SUSY is not an exact symmetry but is broken due to some additional mechanism. In order to avoid reintroducing large corrections to the Higgs boson mass caused by the symmetry breaking, usually a breaking mechanism similar to the spontaneous EWK symmetry breaking is introduced also known as soft symmetry breaking. However, for broken SUSY theories to be able to account for the hierarchy problem, especially mass differences between superpartners, where the SM particle yields large contributions to the Higgs mass corrections, can not be too large [64].

2.2.2 Dark Matter

As discussed in Sec. 2.1.6 the lack of an adequate DM candidate is one of the most striking shortcomings of the standard model. While R -parity conserving supersymmetric extensions can offer a DM candidate given by the LSP, there are many other theories predicting weakly interacting massive particles (WIMPs), which could account for the DM in our universe [17, 65, 66]. WIMPs are among the most popular DM candidates since based on their mass of the order of the EWK scale and their weak coupling they can quite precisely account for the measured DM relic density.

Along with SUSY, extensions based on extra spatial dimensions are another group of theories predicting WIMPs. In these theories, the 3+1 dimensional spacetime is expanded by adding additional dimensions which appear at higher energy scales. An example of such theories are Kaluza-Klein theories [17], which predict compactified extra dimensions. SM particles propagating in these extra dimensions have their momenta quantized due to the limited size of the extra dimension. This quantization corresponds to excited states similar to a particle in an infinite potential wall. In certain cases, the lightest of these states, also known as Lightest Kaluza-Klein particle (LKP), can be stable and serve as a DM candidate.

Other theories implementing DM candidates different from WIMPs are often based on so-called hidden sectors [65]. These additional hidden sectors could contain new particles, which have no SM gauge interactions, as well as new non-SM-like gauge interactions. In general, the new particles of the hidden sector could be only detected based on their gravitational interaction. However, there are certain models predicting particles, so-called mediators, connecting the SM

and the hidden sector. Mediators would have to interact via at least one of the SM interactions making models including mediators accessible for particle collider experiments.

2.3 Top quark physics

Being the heaviest of all known elementary particles with a mass of 172.69 ± 0.30 GeV [37], the top quark plays an important role in many aspects of the standard model as well as in BSM theories. It was predicted in 1973 by Kobayashi and Maskawa [10] together with the bottom quark forming the third generation of quarks. Its existence was experimentally confirmed in 1995 by the CDF and the D0 experiments at the Tevatron proton-antiproton collider [11, 12].

In the following, the production of top quarks at hadron colliders, the different decay modes of the top quark as well as its role concerning BSM physics is discussed and a short introduction of the strategy applied in the present analysis is given.

2.3.1 Production and decay

At hadron colliders, the top quark is predominantly produced in quark-antiquark pairs ($t\bar{t}$). The two dominant production mechanisms of top pairs are shown in Fig. 2.6. While at Tevatron, where proton-antiproton collisions were studied, the largest contribution to the top-pair production was given by quark-antiquark annihilation, at the LHC, which collides protons, the production is dominated by gluon-gluon fusion, accounting for around 90% of the total $t\bar{t}$ production [37].

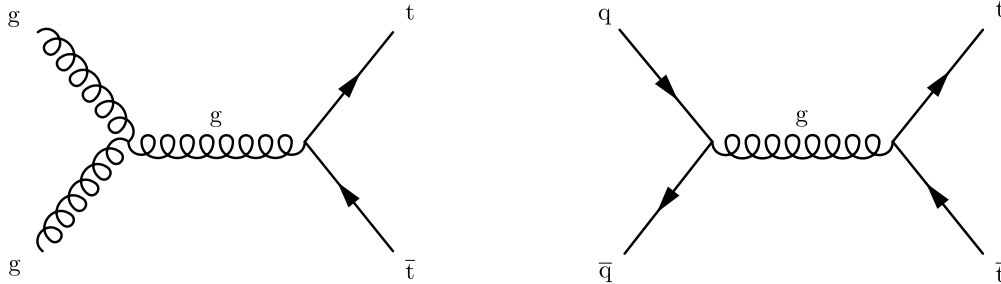


Figure 2.6: Example diagrams for two top-pair production mechanisms at hadron colliders. The diagram on the left corresponds to production via gluon-gluon fusion while the diagram on the right shows quark-antiquark annihilation.

The difference between the dominant production mechanism at Tevatron and LHC originates from the parton density functions (PDFs) of the colliding particles, which include antiquarks as valence quarks in case of proton-antiproton collisions. Using the factorization theorem, the total top-pair production cross section can be derived as a convolution of the PDFs f_i and the partonic cross section $\hat{\sigma}$ [67]:

$$\sigma_{pp \rightarrow t\bar{t}}(s, m_t) = \sum_{i,j=q,\bar{q},g} \int dx_i dx_j f_i(x_i, \mu_F^2) f_j(x_j, \mu_F^2) \hat{\sigma}_{ij \rightarrow t\bar{t}}(\hat{s}, m_t, \mu_F, \mu_R, \alpha_s), \quad (2.16)$$

where s is the squared center-of-mass energy of the collider, m_t is the top-quark mass and x_i are the momentum fractions of the colliding partons. The partonic cross section is derived as a function of the partonic center-of-mass energy $\hat{s}$, the top-quark mass, factorization and renormalization scales $\mu_{F,R}$ and the strong coupling constant α_s . The total $t\bar{t}$ production cross section

has been measured at multiple center-of-mass energies as shown in Fig. 2.7. The measurements across the entire studied energy range up to a center-of-mass energy of 13.6 TeV show good agreement with the corresponding theoretical predictions.

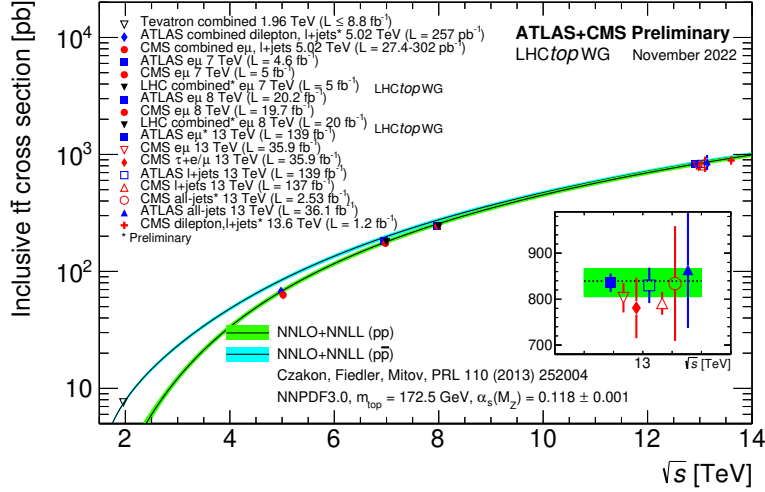


Figure 2.7: Measurements of the inclusive top-pair production cross section as a function of the center-of-mass energy [68] conducted at the LHC and Tevatron. The results are compared to theory predictions derived with Top++ [69].

Due to its short lifetime of the order of 10^{-25} s, the top quark is the only quark which decays before hadronization [37, 67]. Thus, at particle colliders, the properties of the bare quark rather than its hadrons can be tested. Caused by the large diagonal element $|V_{tb}|$ of the CKM matrix, shown in Eq. (2.9), top quarks decay almost exclusively to a W boson and a bottom quark. Focusing on $t\bar{t}$ production the signatures produced at particle colliders can be grouped into three different decay channels depending on the decay modes of the two W bosons. The all-hadronic decay channel includes signatures where both W bosons decay hadronically. Signatures where only one of the two W bosons decays into quarks while the other boson decays into a charged lepton and a neutrino are denoted as semileptonic decays. The channel, which is of most interest in this thesis, is given by the dileptonic channel, where both W bosons decay leptonically. The different contributions as well as the branching ratios of the three different decay channels taking into account different lepton flavors are shown in Fig. 2.8.

While the all-hadronic decay channel accounts for almost half of the $t\bar{t}$ decays, the reconstruction of its final states in a particle detector is rather challenging due to the hadronization of its decay products and large background contributions from QCD events, which can mimic the signatures of the all-hadronic channel. The dileptonic channel on the other hand, ignoring final states with τ leptons, amounts to only 4% of the total branching fraction. However, the rather small contributions from other SM processes to this final state as well as the more precise detection of charged leptons in a particle detector compared to quarks yield much cleaner signatures in the dileptonic decay channel.

2.3.2 Top quarks and BSM physics

There are several ways in which the top quark could be connected to physics beyond the standard model. The main argument why the top quark might play an important role in potential extensions of the SM is usually based on its large mass and therefore the large Yukawa coupling

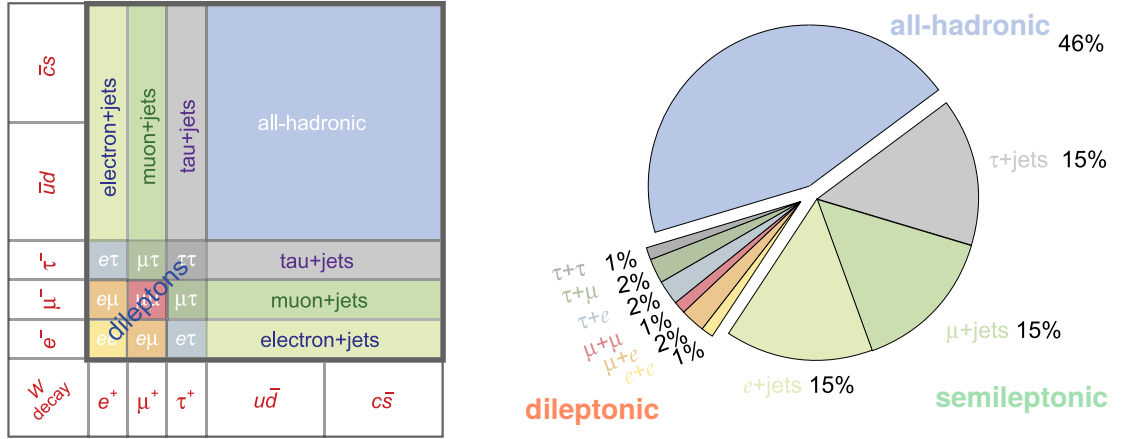


Figure 2.8: The sketch on the left shows the compositions of the three $t\bar{t}$ decay channels grouped into the all-hadronic, the lepton+jets (semileptonic) and the dilepton channel depending on the decay of the two W bosons [70]. In the diagram on the right, the corresponding branching fractions of the different decay channels are displayed based on the same color coding as shown in the sketch on the left [71].

between the Higgs field and the top quark [58, 67]. As a consequence, the top quark contributes to the largest corrections of the Higgs boson mass, which give rise to the hierarchy problem of the SM as introduced in Sec. 2.1.6. Following the same argumentation, also the stability of the electroweak vacuum is strongly correlated to the mass of the top quark. Depending on the Higgs mass a top-quark mass which is too large leads to a meta-stability of the vacuum [72].

Contributions from top quarks are also well motivated in terms of experimental signatures covered by searches for BSM physics. As discussed in Sec. 2.2.1, the mass difference between superpartners, where the SM particle yields large corrections to the Higgs mass, can not be too large in order to solve the hierarchy problem. Therefore, the mass of the stop, the superpartner of the top quark, should not be too large motivating searches for signatures which include stops. Since the stop predominantly decays to top quarks and LSPs [67], such searches naturally include top quarks in their targeted final states, making SM $t\bar{t}$ production one of the main SM backgrounds in searches for supersymmetry as well as dark matter.

Two examples of such signatures including two top quarks and two BSM particles in the final state are shown in Fig. 2.9 and compared to the signature of SM $t\bar{t}$ production [73–75]. All three signatures include the production of two top quarks, while the final states of the two BSM signatures hold additional BSM particles, which are assumed to interact only via the weak force.

2.3.3 Cross section measurement in a BSM phase space

This thesis focuses on the difference between the BSM signatures discussed above and the signatures produced in SM $t\bar{t}$ production using the dileptonic decay channel of the top pair. The top-quark decays in both types of processes produce two charged leptons originating from the W-boson decays and two jets produced by the bottom quarks, which both can be detected in a particle detector. The remaining final state particles given by the neutrinos from the W-boson decays as well as the potential BSM particles interact only via weak interactions and therefore escape a particle detector, like the CMS or ATLAS detector, undetected. The only way to still measure these particles is to make use of the momentum imbalance introduced by the non-detected particles quantified by the so-called missing transverse momentum $\vec{p}_T^{\text{miss}}$, which

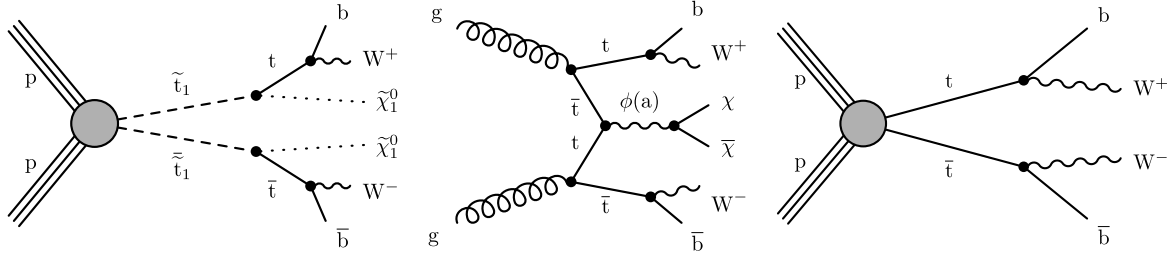


Figure 2.9: The left diagram shows a pair production of the lightest stop quark $\tilde{t}_1$, where both stops decay to a top quark and the lightest neutralino $\tilde{\chi}_1^0$. The diagram in the middle sketches $t\bar{t}$ production in association with a (pseudo-)scalar DM mediator $\phi(a)$ decaying to a pair of DM Dirac fermions. The diagram for SM $t\bar{t}$ production is shown on the right. The blobs in the diagrams represent different production mechanisms of the corresponding pair production.

is introduced in Sec. 6.2.6. While the missing transverse momentum yields a measure of the transverse momentum of undetected particles in an event, there is no way to distinguish between the different contributions to the transverse momentum imbalance.

However, comparing the BSM and SM process shown in Fig. 2.9, different particles contribute to the transverse momentum imbalance. While for the SM process only the neutrinos from the leptonic W-boson decays account for the missing transverse momentum, additional contributions from the neutralinos or the DM fermions are present in the BSM signatures. These additional contributions will on one hand affect the magnitude² of $\vec{p}_T^{\text{miss}}$, also denoted as p_T^{miss} , and on the other hand also affect the correlation between the directions of $\vec{p}_T^{\text{miss}}$ and the remaining detected particles. The left plot in Fig. 2.10 shows a comparison between the normalized p_T^{miss} spectra of the SM $t\bar{t}$ process and the two BSM processes introduced above. Due to additional contributions to the momentum imbalance caused by the BSM particles, both BSM processes show a significantly harder spectrum compared to the SM process. The impact on the correlation between the directions of $\vec{p}_T^{\text{miss}}$ and the leptons, quantified by the minimal azimuthal distance between $\vec{p}_T^{\text{miss}}$ and a lepton $\min[\Delta\phi(\vec{p}_T^{\text{miss}}, \ell)]$, can be seen in the right plot of Fig. 2.10. For events with a significant missing transverse momentum, the BSM contributions yield a much weaker correlation due to the additional undetected particles.

In the present analysis, both observables shown in Fig. 2.10 will be employed to separate SM $t\bar{t}$ events from potential BSM contributions. Unlike conventional searches for new physics, as performed by CMS [76] and ATLAS [77] based on the signatures introduced before, this analysis will not focus only on the phase space where large contributions from BSM processes could be present but measure the entire spectrum to also characterize the SM $t\bar{t}$ process itself. This translates into a method where in case no significant BSM contributions are found in the measurement on detector level, differential cross section results for the dileptonic $t\bar{t}$ process are derived, rather than improving the limits for different BSM scenarios. The corresponding studies on detector level are performed following a strict blinding strategy, where the sensitive detector-level distributions are only studied in recorded data once all analysis ingredients are fixed. To derive the differential cross section results the measured spectra are unfolded to a fiducial phase space by correcting for detector and acceptance effects, as further discussed in Sec. 8. The cross sections are measured in one and two dimensions as a function of the quantities matching the two observables introduced above for the dileptonic $t\bar{t}$ process. Assuming the

²In the following, the distinction between the vectorial missing transverse momentum $\vec{p}_T^{\text{miss}}$ and its magnitude p_T^{miss} is usually neglected.

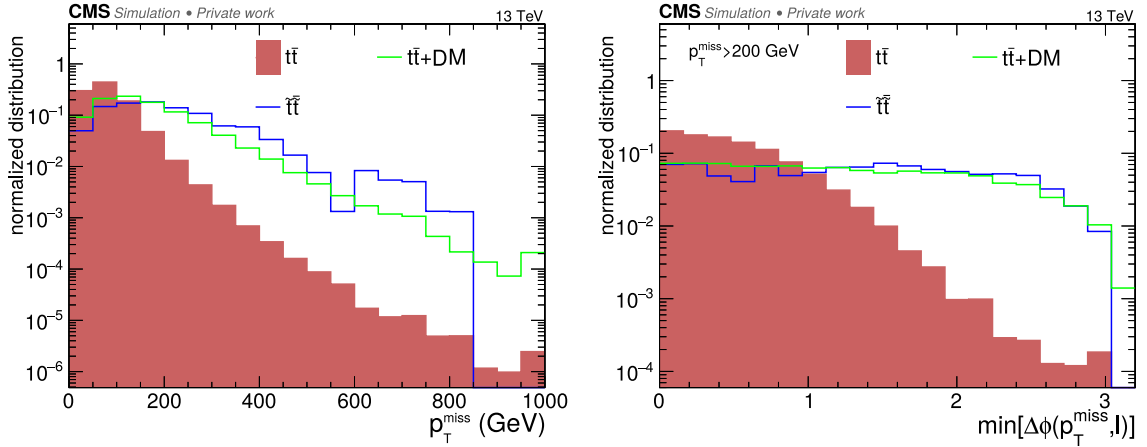


Figure 2.10: Normalized distributions of p_T^{miss} (left) and minimal azimuthal distance between p_T^{miss} and a lepton (right) comparing SM $t\bar{t}$ production (red), stop-pair production (blue) and $t\bar{t}$ production in association with a scalar DM mediator (green). For the comparison on the right only events with p_T^{miss} above 200 GeV are taken into account. The SUSY process is based on stop and neutralino masses of 650 GeV and 350 GeV, respectively. The DM process is simulated with DM fermion and scalar masses of 200 GeV and 1 GeV, respectively.

dileptonic $t\bar{t}$ process and neglecting any mismeasurements and neutrinos produced in b-meson decays present in the hadronization of b quarks, the missing transverse momentum is caused by the momentum sum of the two neutrinos denoted as $p_T^{\nu\nu}$. Thus, if no significant BSM contributions are observed, the analysis would yield differential cross sections as a function of $p_T^{\nu\nu}$ and $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ as well as the 2D combination of both observables.

With this approach, the dileptonic $t\bar{t}$ cross section can be measured up to a phase where potential BSM contributions become relevant. This enables potential re-interpretations of the results in a BSM context, e.g., by using effective field theory methods [78, 79]. In addition, taking into account that the SM $t\bar{t}$ process is often the main background in BSM searches, where missing transverse momentum plays a major role, the results allow to validate existing background estimation methods. Furthermore, the analysis provides a precise SM $t\bar{t}$ measurement on its own in a phase space not yet measured at the LHC, which can be compared to theory predictions.

Chapter 3

The CMS experiment at the LHC

3.1	The Large Hadron Collider	23
3.2	The Compact Muon Solenoid detector	26
3.2.1	Coordinate system	27
3.2.2	Solenoid magnet	28
3.2.3	Tracking system	28
3.2.4	Electromagnetic calorimeter	30
3.2.5	Hadron calorimeter	32
3.2.6	Muon system	34
3.2.7	Trigger system	36

Particle colliders are among the most powerful tools to test the standard model and to search for BSM physics. The collider currently providing particle collisions with the highest center-of-mass energies of up to 13.6 TeV is the Large Hadron Collider (LHC), located at the European Center of Nuclear Research (CERN). The proton-proton (pp) collisions provided by the LHC, which are analyzed in this thesis, were recorded by the Compact Muon Solenoid (CMS), which is a multi-purpose detector located at one of the four interaction points of the LHC. In this chapter, the main aspects of the LHC are briefly introduced in Sec. 3.1, while the main part of the chapter is dedicated to the CMS experiment, where the coordinate system, the different subdetectors and the trigger system are described in Sec. 3.2.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [23] based at CERN is the largest and most powerful particle accelerator and collider ever built. The LHC has a circumference of 26.7 km and is located in the tunnel already used for the Large Electron-Positron Collider (LEP) [80] at a depth between 45 and 170 m underground. It is designed to collide proton beams at a center-of-mass energy of $\sqrt{s} = 14$ TeV and heavy ions with an energy of up to 2.8 TeV per nucleon using two separate rings, which allow to simultaneously accelerate same-charged particles in opposite directions.

At the LHC, there are four interaction points, where the two separate rings cross and the beams can collide. The particle collisions at the interaction points are recorded by four different particle detectors, as sketched in Fig. 3.1. The CMS detector (Compact Muon Solenoid) [25] and ATLAS detector (A Toroidal LHC Apparatus) [24] are the most complex out of the four detectors and

located at the northernmost and southernmost interaction point, respectively. Both detectors are multi-purpose detectors designed to perform precision measurements of the standard model as well as searches for physics beyond the standard model. The remaining two detectors, the ALICE detector (A Large Ion Collider Experiment) [81] and the LHCb detector (LHC beauty) [82], cover more specific topics of particle physics research. The ALICE detector is mainly designed to record heavy ion collisions enabling studies of the quark-gluon plasma. The physics program of the LHCb detector, which unlike the three other detectors has an asymmetric design, is dedicated towards heavy flavor physics focusing on CP violation and rare decays of beauty and charm hadrons.

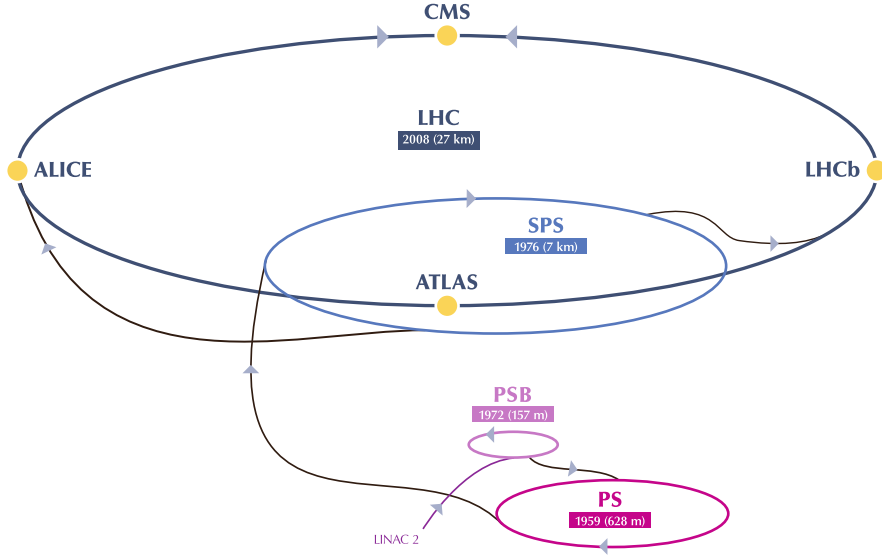


Figure 3.1: Sketch of the LHC accelerator scheme used for proton acceleration during Run 2 of the LHC [83]. The four large detectors located at the LHC are shown as yellow dots. Before injection into the LHC, the protons run through a linear accelerator (LINAC 2) and three synchrotrons (PSB, PS and SPS). For the circular accelerators, the beginning of operation as well as the circumference of the accelerator are given.

Following the scheme shown in Fig. 3.1, before injection into the LHC the protons run through a set of different acceleration steps. In the first step, the protons are accelerated by a linear accelerator (LINAC 2)¹ up to an energy of 50 MeV. In the following steps, the protons are further accelerated by three synchrotrons with increasing size. The Proton Synchrotron Booster (PSB) accelerates the protons to an energy of 1.4 GeV, while the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS) increase the energy of the protons to 25 GeV and 450 GeV, respectively. The protons are then injected every 25 ns into the LHC filling up to 2808 bunches, which each hold around 10^{11} protons. In the last step, the protons are further accelerated by the LHC up to their collision energy.

¹In 2020, LINAC 2 was replaced by LINAC 4, which accelerates negative hydrogen ions up to 160 MeV [84]. Both electrons are then stripped during the injection into the PSB.

Apart from the center-of-mass energy, which affects the cross section for a given process, as shown in Eq. (2.16), the performance of a particle collider is quantified by its instantaneous luminosity $\mathcal{L}$ and integrated luminosity L_{int} given by

$$\mathcal{L} = \frac{n_b f N^2 F}{4\pi\sigma_x\sigma_y}, \quad L_{\text{int}} = \int \mathcal{L} \cdot dt. \quad (3.1)$$

The instantaneous luminosity is proportional to the number of bunches n_b , the number of particles per bunch N and the revolution frequency of the bunches f . The additional factor F takes losses due to the crossing angle at the interaction point into account. The denominator accounts for the beam geometry, where σ_x and σ_y denote the Gaussian width of the beam in the plane perpendicular to the beam direction [85]. Based on the luminosity, the rate $\dot{N}$ as well as the total number of events N_{tot} of a process with a given cross section σ can be expressed as

$$\dot{N} = \mathcal{L} \cdot \sigma, \quad N_{\text{tot}} = L_{\text{int}} \cdot \sigma. \quad (3.2)$$

Vice versa, as performed in this thesis, the cross section σ of a specific process can be derived by measuring the total number of events at a known integrated luminosity.

Since the first collisions at the LHC in 2010, the center-of-mass energy and the instantaneous luminosity of the pp collisions have been constantly increased across several different data-taking periods [86, 87]. Starting with collisions at a center-of-mass energy of 7 TeV (2010-2011) and 8 TeV (2012), the first period of data-taking, denoted as Run 1, reached an instantaneous luminosity of up to $0.8 \cdot 10^{34} \text{ cm}^{-2}\text{s}^{-1}$. The second data-taking period started in 2015 with collisions at $\sqrt{s} = 13 \text{ TeV}$ and provided a total integrated luminosity of around 160 fb^{-1} across the four years of data taking (2015-2018), as shown in Fig. 3.2, while reaching an instantaneous luminosity of up to $2.1 \cdot 10^{34} \text{ cm}^{-2}\text{s}^{-1}$. In 2022, the latest data-taking period (Run 3) started with an even higher center-of-mass energy of 13.6 TeV already improving on the peak instantaneous luminosity reached during Run 2.

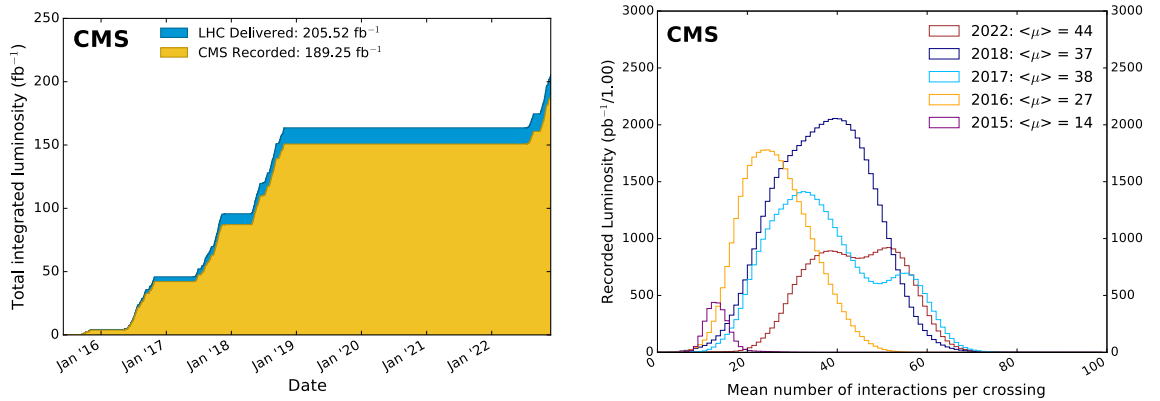


Figure 3.2: Cumulative luminosity delivered by the LHC (blue) and recorded by CMS (yellow) as a function of time for Run 2 and the first year of Run 3 (left) [87]. The integrals of both measures are shown in the legend. Recorded luminosity as a function of the mean number of interactions per bunch crossing (pileup) for the years of data taking between 2015 and 2022 (right). The mean number of interactions for each year is denoted in the legend.

One way to achieve higher luminosities is to squeeze the proton beam, effectively decreasing the σ_i in Eq. (3.1). As a consequence, the number of simultaneous interactions during a bunch crossing, known as pileup, increases, which is shown in Fig. 3.2. Therefore, while larger luminosities allow collecting more collision data on shorter time scales, the reconstruction of individual events becomes more involved due to the increasing number of simultaneous interactions, which will be one of the main challenges when entering the era of the High-Luminosity-LHC [88].

3.2 The Compact Muon Solenoid detector

The Compact Muon Solenoid (CMS) detector [25, 89], sketched in Fig. 3.3, is a multi-purpose detector located at the northernmost interaction point of the LHC. With a length of 21.6 m and a diameter of 14.6 m the detector is rather compact compared to the ATLAS detector [24], the second multi-purpose detector located at the LHC, but has a significantly larger total weight of around 12500 t. The CMS detector is designed to cover a wide physics program ranging from precision measurements of the standard model including measurements of the Higgs boson and the top quark to heavy-flavor physics and searches for various BSM signatures.

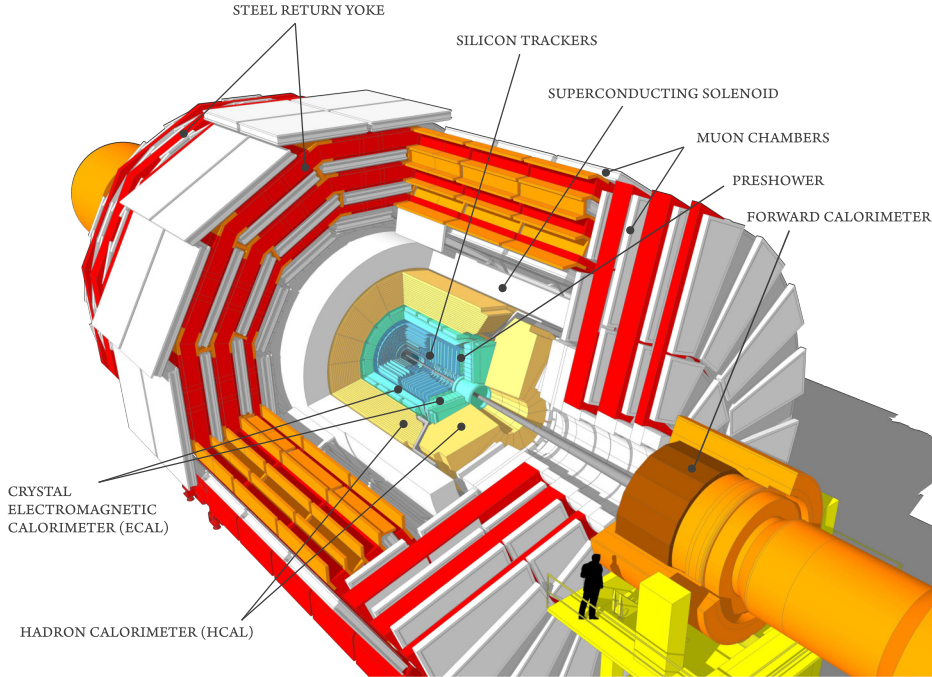


Figure 3.3: Sketch of the CMS detector showing the main components of the different subdetectors and the solenoid magnet [90].

The detector is composed of four main subdetectors, which are cylindrically placed around the interaction point in the center of the detector. The innermost part of the detector corresponds to the tracking system recording the trajectories of charged particles. The electromagnetic calorimeter (ECAL) and hadron calorimeter (HCAL) are built around the tracking system allowing for precise energy measurements of various types of particles. The tracking system and both calorimeters are surrounded by a superconducting solenoid, which provides an almost homogeneous magnetic field of 3.8 T inside the solenoid enabling a precise momentum measurement for charged particles based on their trajectories. The muon system, which is partly embedded into the iron return yoke of the solenoid, marks the outermost part of the detector and allows for

efficient detection of muons escaping the calorimeters. Each of the four subdetectors consists of a barrel part, which is built around the beam pipe and two so-called endcaps, which close both open ends of the barrel part. Besides high energy and momentum resolution, another key aspect in the design of the subdetectors is given by the radiation tolerance of detector components close to the beam pipe and the high occupancy caused by the large flux of particles produced in the pp collisions. In the following, an overview of the main features including the solenoid magnet, the different subdetector systems as well as the triggering system, which is crucial for a reasonable data acquisition, is given.

3.2.1 Coordinate system

The coordinate system adopted at the CMS detector, which is sketched in Fig. 3.4, is centered around the nominal collision point. The x -axis of the system points radially inward towards the center of the LHC, while the y -axis is chosen to point vertically upward. In order to create a right-handed coordinate system, the z -axis, which runs parallel to the beam axis, is selected to point towards the direction of the beam running counter-clockwise in the collider.

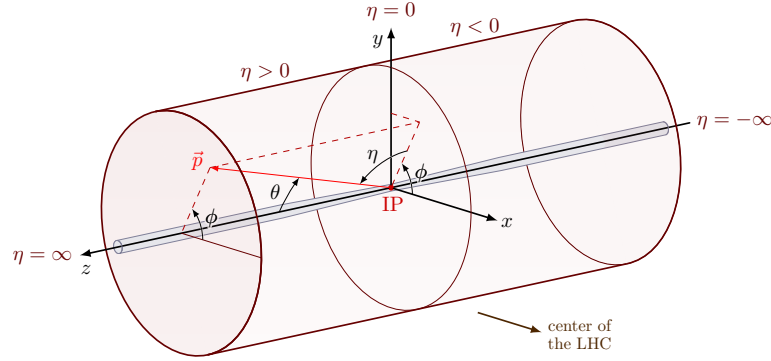


Figure 3.4: Sketch of the coordinate system used at the CMS detector [91]. The system is centered around the interaction point (IP). The red arrow visualizes a potential particle momentum.

At the CMS detector, particle trajectories are usually described using spherical coordinates as well as transverse properties. The spherical coordinates are defined such that the azimuthal angle ϕ describes the direction inside the xy -plane starting from the x -axis in a range from $-\pi$ to π . The polar angle θ , ranging between 0 and π , yields the direction in the yz -plane starting from the z -axis. To simplify the discussion of tracks from relativistic particles, the polar angle θ is usually replaced by the pseudorapidity

$$\eta = -\ln \left(\tan \frac{\theta}{2} \right). \quad (3.3)$$

Unlike distances with respect to the beam plane described by $\Delta\theta$, distances in $\Delta\eta$ are invariant against Lorentz boosts along the z -axis in the limit of massless particles. Distances combining the azimuthal angle and the pseudorapidity are usually given in ΔR :

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}, \quad (3.4)$$

which is also invariant against the Lorentz boost described above. The transverse properties of particle trajectories usually include the transverse momentum p_T and transverse energy E_T , which are projections onto the xy -plane:

$$p_T = \sqrt{p_x^2 + p_y^2}, \quad E_T = \sqrt{E_x^2 + E_y^2}, \quad (3.5)$$

where the x - and the y -component correspond to the particle momenta (energy) times the cosine and sine of the azimuthal angle, respectively. Similar to differences in the pseudorapidity, both transverse quantities are invariant against Lorentz boosts along the z -axis.

3.2.2 Solenoid magnet

One of the key features of the CMS detector is its superconducting solenoid, which has a length of 12.5 m and an inner diameter of 6 m. Inside the coil, which consists of four layers of Niobium-titanium (NbTi) conductors, an almost homogeneous magnetic field of 3.8 T is provided. The strong magnetic field leads to a large stress on the components of the magnet, which is partially compensated by using the conductors themselves as part of the structural material of the coil. The flux of the solenoid is returned by an iron return yoke, which marks the outermost part of the detector and makes up for roughly 80% of the total detector weight.

The magnetic field inside the coil is employed to bend the trajectories of charged particles, recorded by the tracking system allowing for a precise momentum measurement of charged particles. In addition, the magnetic field inside the return yoke is used by the muon system, which is partly embedded in the yoke, to derive an independent measurement of the muon momenta.

3.2.3 Tracking system

The tracking system of the CMS detector built around the beam pipe is the innermost part of the detector. It has a total length of 5.8 m and extends to a radius of 1.1 m with respect to the beam pipe. The tracking system is designed to provide efficient and precise measurements of charged particle trajectories used to derive their momenta and to reconstruct secondary vertices crucial for the identification of b hadrons. In addition, the high granularity and the fast response time of the subdetector allow for a precise assignment of the trajectories to the correct bunch crossing as well as the correct vertex. The different parts of the subdetector are sketched in Fig. 3.5. Based on the sensor type used for the detection of charged particles and the distance to the interaction point, the tracking system can be divided into two parts given by the inner pixel detector and the outer strip detector.

The inner pixel detector provides tracking measurements in the region closest to the interaction point enabling a precise reconstruction of primary and secondary vertices. Due to the small distance between the interaction point and the pixel detector, the original pixel detector, which was only designed for luminosities up to $10^{34} \text{ cm}^{-2}\text{s}^{-1}$, had to be replaced after the data-taking period in 2016 by the new CMS Phase-1 pixel detector [92]. In the following, the details of the new CMS Phase-1 pixel detector used for the majority of pp collisions analyzed in this thesis are briefly discussed. The features of the old pixel detector can be found in [25].

The Phase-1 pixel detector includes a total of 1856 individual silicon sensor modules, where 1184 modules are located in the barrel part (BPIX) and the remaining 672 modules are used in the two endcap parts (FPIX). The main part of each module is a sensor consisting of 160×400 pixels, where each pixel has a size of $100 \times 150 \mu\text{m}^2$, enabling a precise track res-

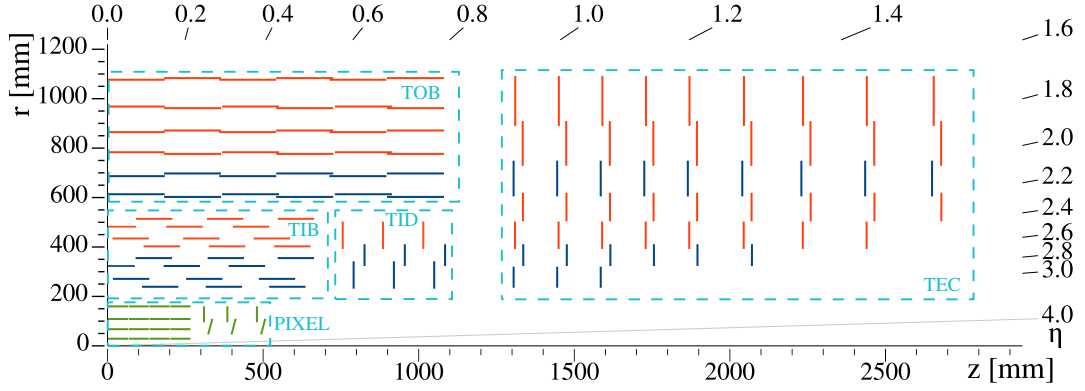


Figure 3.5: Sketch of one quarter of the CMS tracking system shown in the rz -plane as used for the data taking in 2017 and 2018 [93]. The dashed lines in light blue mark the different parts of the subdetector. The pixel detector is shown in green, while the single-sided and double-sided components of the strip detector are visualized in red and blue, respectively.

olution of the order of $\mathcal{O}(10\mu\text{m})$ in the $r\phi$ - and z -direction and yielding a small occupancy despite the high particle flux. The BPIX is built in four layers placed at radii of 29, 68, 109 and 160 mm around the beam pipe, as shown in Fig. 3.5. Each layer is formed by a different number of ladders, where each ladder holds eight modules. Each of the two parts of the FPIX contains three disks placed at distances of 291, 396 and 516 mm from the center of the detector, where each disk consists of 112 modules. Combining BPIX and FPIX, the Phase-1 pixel detector employs a total of 124 million readout channels and has a total silicon area of 1.9 m^2 covering a range up to $|\eta| < 2.5$.

The outer strip detector is built around the pixel detector extending the radius of the tracking system from 0.2 m up to 1.1 m. Here, silicon microstrip sensors with different pitch sizes between 80 and $180\mu\text{m}$ are used, matching the decreasing particle flux with increasing distance to the interaction point. Similar to the pixel detector, the strip detector is divided into a barrel and an endcap part, as shown in Fig. 3.5. The barrel part includes the Tracker Inner Barrel (TIB) built from four layers covering a range up to $|z| < 65\text{ cm}$ and the Tracker Outer Barrel (TOB), which consist of six layers extending up to $|z| < 110\text{ cm}$. The main structure of the endcap part corresponds to the Tracker End Cap (TEC), which consists of nine disks. The second structure of the endcap part, the Tracker Inner Disk (TID), fills the gap between the TIB and the TEC and includes three disks. To enable track measurements in the $r\phi$ - as well as in the rz -direction, doubled-sided modules, also known as stereo modules, with a stereo angle of 100 mrad are used in parts of the strip detector (cf. Fig. 3.5). In total, the strip detector is built from more than 15000 individual modules covering a combined area of around 200 m^2 .

With the combination of both parts of the tracking system, a precise momentum measurement of charged particles based on their bent trajectories as well as a high resolution in the vertex reconstruction can be achieved. In Fig. 3.6, examples for momentum and vertex resolution are shown comparing the tracking system used during the data-taking periods in 2016 and 2017, which were based on the old and new inner pixel detector, respectively. In the central part of the detector, a relative transverse momentum resolution of up to 1% can be realized, while in the region of the endcaps, the resolution is reduced to around 2-5%. The improvement in the transition region between the barrel and endcap part ($1.2 < |\eta| < 1.6$) is caused by the additional BPIX layer of the Phase-1 pixel detector, which improves the track extrapolation from the pixel

region to the strip detector. An even larger effect of the upgraded pixel detector can be found in the vertex resolution shown in Fig. 3.6. While at large track multiplicities, a vertex resolution of the order of $10\,\mu\text{m}$ can be achieved with both pixel detectors, at smaller track multiplicities the Phase-1 pixel detector can improve the resolution by up to 50%. One key requirement in realizing such precise momentum measurements and vertex resolutions of $\mathcal{O}(10\,\mu\text{m})$ is a correct alignment of the different components of the tracking system, which are only mounted with a precision of $\mathcal{O}(100\,\mu\text{m})$. For this, CMS employs a complex track-based alignment [94], which is further discussed in Sec. 4.

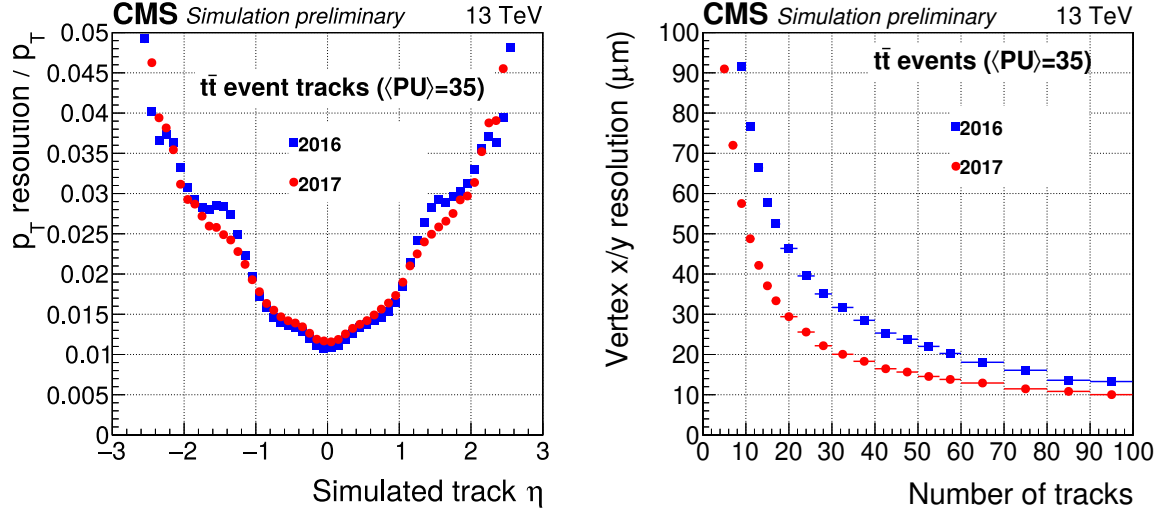


Figure 3.6: Relative resolution of the transverse momentum as a function of the track η (left) [95]. Transverse vertex resolution as a function of the number of tracks used in the fit of the vertex (right). Both plots are based on simulated top-pair events using a mean pileup of 35. The results for the data-taking periods of 2016 (blue) and 2017 (red), which use the old and new pixel detector, respectively, are compared.

3.2.4 Electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) is one of the two calorimeter systems of the CMS detector. Its main purpose is the precise energy measurement of electrons and photons based on their electromagnetic shower induced in the calorimeter. While electrons and photons lead to similar signatures in the ECAL, the combination of signals in the ECAL and the tracking system is used to accurately identify both particle types.

The ECAL system of the CMS detector is a homogeneous calorimeter built from scintillating lead tungstate (PbWO_4) crystals, which act as absorbing and scintillating material at the same time. The crystals have a high density of $8.27\,\text{g}/\text{cm}^3$ as well as a short radiation length of $X_0 = 0.89\,\text{cm}$ and a small Molière radius of $2.2\,\text{cm}$, which allows for a compact calorimeter design with fine granularity. Due to the short scintillating decay time of the crystals, around 80% of the total light yield, which is proportional to the energy deposit in a crystal, is emitted within the nominal LHC bunch crossing time of 25 ns. The light yield produced inside the crystals is detected by avalanche photodiodes and vacuum phototriodes in the barrel (EB) and endcap (EE) part of the ECAL, respectively, which are placed at the back side of each crystal. Due to the close placement of the ECAL to the interaction point, the crystals, especially in the EE, are exposed to high radiation levels and particle fluxes, which lead to a wavelength-dependent loss of light

transmission inside the crystals. This damage is tracked and corrected using a laser monitoring system, which measures the optical transparency of the crystals in different regions of the ECAL.

The ECAL is divided into three different parts as shown in Fig. 3.7. The EB consists of a total of 61200 crystals, covers a range up to $|\eta| < 1.479$ and extends up to a radial distance of 1.77 m from the beam pipe. The crystals installed in the EB have a length of 230 mm, corresponding to $25.8 X_0$, and a front face cross section of $22 \times 22 \text{ mm}^2$ matching the Molière radius of the crystals. Each of the two endcap parts of the ECAL cover a pseudorapidity range of $1.479 < |\eta| < 3.0$ and include a total of 7324 crystals with a length of 220 mm ($24.7 X_0$) and a front face cross section of $28.6 \times 28.6 \text{ mm}^2$. An additional two-layer sampling calorimeter called preshower (ES), which consists of alternating lead and silicon strip sensor layers, is installed in front of the endcap parts covering a range of $1.653 < |\eta| < 2.6$. The preshower system allows for a better rejection of photons originating from π^0 decays as well as an improved position determination for electrons and photons in the endcap due to the high-granularity design of the preshower.

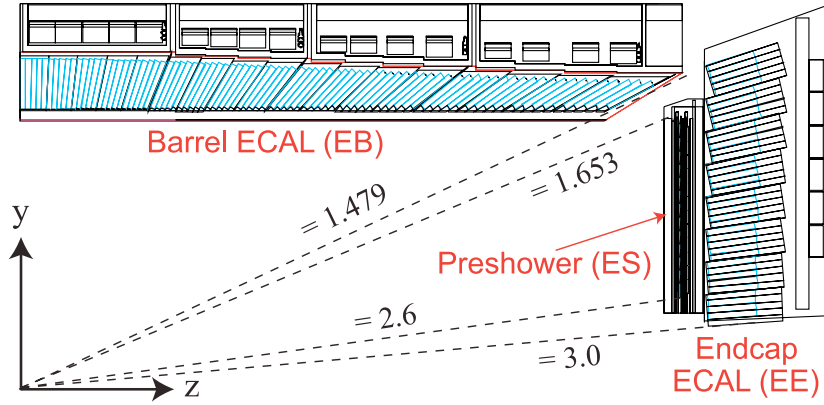


Figure 3.7: Sketch of one quarter of the CMS electromagnetic calorimeter shown in the longitudinal plane [89]. The three different parts of the calorimeter are denoted in red.

The electron reconstruction performance of the ECAL is shown in Fig. 3.8. For electrons with a transverse momentum between 20 and 500 GeV, a flat reconstruction efficiency of about 96% can be achieved up to a pseudorapidity of $|\eta| < 2.5$, which corresponds to the coverage of the tracking system. Only the transition region between EB and EE at $|\eta| \approx 1.5$ shows a slight reduction in efficiency and is therefore usually excluded when analyzing electrons or photons. The energy resolution of the ECAL can be parameterized as a function of the energy taking three different contributions into account:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2, \quad (3.6)$$

where S describes a stochastic contribution including, for example, fluctuations in the number of photons, N represents a noise term originating from electronic noise and pileup contributions and C covers constant contributions like non-uniformities of the light collection [96]. The measured electron energy resolution based on simulation can be found in Fig. 3.8. At low energies, the tracking system yields a better resolution compared to the ECAL. At high energies, the energy measurement is dominated by the ECAL resolution, where a relative resolution of around 1% can be achieved for electrons with a transverse momentum above 100 GeV.

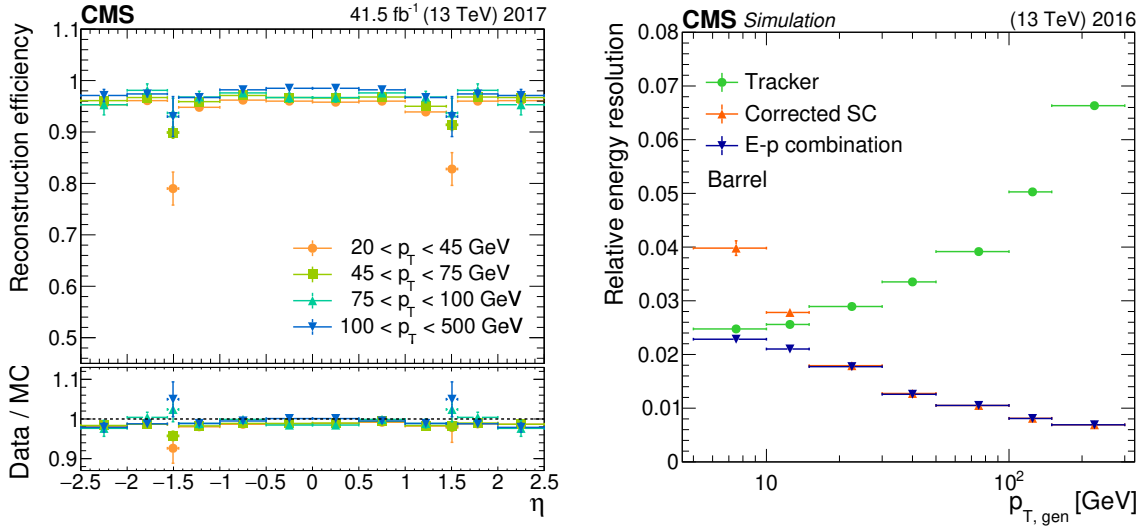


Figure 3.8: Left plot: Electron reconstruction efficiency as a function of the pseudorapidity for different electron p_T ranges measured in data recorded in 2017 [97]. The lower panel shows the ratio to a measurement based on simulation. Right plot: Relative energy resolution versus electron p_T measured by the tracking system (green), the ECAL (orange) and the combination of both measurements (blue) in simulation for the 2016 data-taking period. The vertical error bars in both plots correspond to the total uncertainty.

3.2.5 Hadron calorimeter

The second of the two CMS calorimeter systems is the hadron calorimeter (HCAL), which, unlike the ECAL, is a sampling calorimeter with individual absorber and scintillator layers. Its main purpose is the energy measurement of charged and neutral hadrons based on their hadronic showers induced in the HCAL. It is also designed to provide good containment and hermeticity in order to improve the resolution of the missing transverse momentum used to detect neutrinos or potential BSM particles. As shown in Fig. 3.9, the HCAL contains four individual parts, out of which the barrel part (HB) and the endcap part (HE) are contained within the magnetic coil. The HCAL is completed by an outer part (HO) placed just outside the magnetic coil and a forward part (HF) covering the high pseudorapidity range.

The HE and HB are designed to maximize the absorber material within the magnetic coil. They are built from alternating layers of brass absorbers and plastic scintillators, which for most of the layers have a thickness of 3.7 mm and are connected via wavelength-shifting fibers and clear fibers to photodetectors. To increase the structural stability of the HB, its innermost and outermost absorber layers are made from steel. The HB extends from a radial distance of 1.77 m to 2.95 m with respect to the beam pipe and covers a range up to $|\eta| < 1.3$. It is built from a total of 32 sectors in η , also called towers, and 36 wedges in ϕ , which are each divided into four sections, resulting in a scintillator granularity of $\Delta\phi \times \Delta\eta = 0.087 \times 0.087$. The total absorber thickness of the HB ranges from 5.8 interaction lengths (λ_I) at $\eta = 0$ up to $10.6 \lambda_I$ at $|\eta| = 1.3$. The HE consists of 13 towers in η extending the range of the HCAL to $|\eta| < 3.0$. While for the first five towers, the granularity matches the one used in the HB, for the remaining towers it is increased up to 0.17×0.17 . Including the ECAL crystals, the total length of the calorimeter in the forward directions sums up to around $10 \lambda_I$.

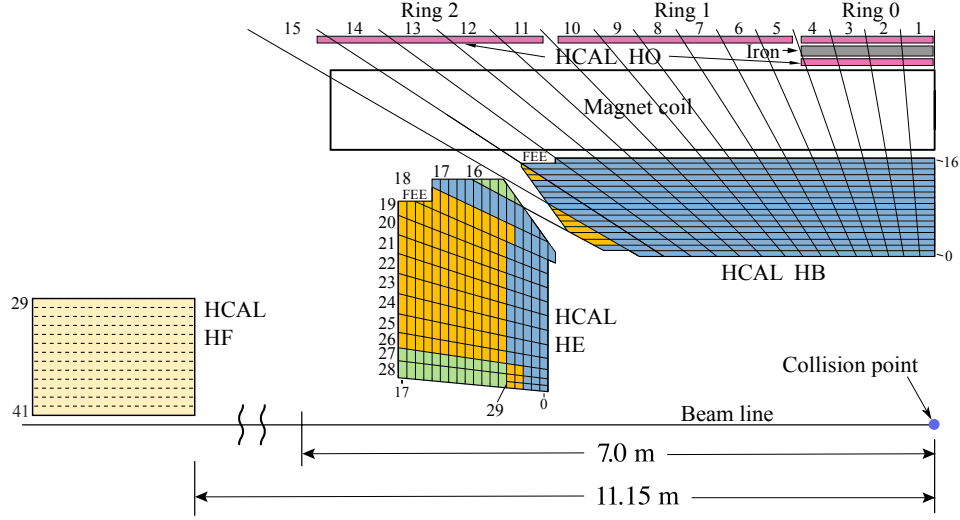


Figure 3.9: Sketch of one quarter of the CMS hadron calorimeter as used during the 2016 data-taking period shown in the longitudinal plane [98]. The calorimeter consists of four individual parts given by the barrel (HB), endcap (HE), outer (HO) and forward part (HF). The coloring scheme marks individual readout channels, while the numbers denote the layers and towers.

The energy resolution of the barrel and endcap part of the calorimeter has been measured in pion test beams and can be parametrized as follows:

$$\frac{\sigma_E}{E} = \frac{A}{\sqrt{E/\text{GeV}}} \oplus B \quad (3.7)$$

For the HB, using pions with energies between 20-300 GeV a resolution with $A = 115\%$ and $B = 5.5\%$ was measured [99], while for the HE using pions in an energy range of 5-300 GeV, the parameters were measured to $A = 106\%$ and $B = 4.0\%$ [100].

In order to increase the effective thickness provided by the HB, an additional part of the HCAL, also called tail catcher or HO, is installed just outside the magnetic coil as shown in Fig. 3.9. The HO is designed to employ the coil as an additional absorber layer and to sample the energy of hadronic showers leaking through the end of the HB. For the sampling, two scintillation layers with an iron absorber layer in between are used in the most central part, while for the more forward parts, only one scintillation layer is installed. Due to the HO, the minimum length of the total calorimeter system is increased to around $11.8 \lambda_I$, neglecting the boundary region between HE and HO.

The most forward part of the HCAL corresponds to the hadron forward calorimeter (HF), which is placed at a distance of 11.2 m with respect to the interaction point and covers a range up to $|\eta| < 5.0$. The HF is built from steel absorbers and embedded radiation-hard quartz fibers, to withstand the extreme particle fluxes present in the region close to the beam line. The signal in the HF is generated by Cherenkov light produced by charged particles inside the hadronic shower, which is channeled by the quartz fibers to photomultipliers. To enable the detection of electromagnetic showers originating from electrons and photons, which in this pseudorapidity range are not measured by the ECAL due to its smaller coverage, the HF is divided into two longitudinal parts. In the first 22 cm long part, only one set of quartz fibers is embedded,

while for the remaining part, a second set of fibers is installed with both sets being read out separately. Since electrons and photons deposit the main part of their energy in the first part, while hadrons produce similar signals in both parts, the showers originating from hadrons and electrons or photons can be distinguished.

3.2.6 Muon system

The muon system marks the outermost part of the CMS detector and is partially embedded into the return yoke of the solenoid magnet. Its main purposes are detection, identification and triggering of muons, which are usually not stopped by the inner subdetectors. The system also provides a precise measurement of the muon momentum which is independent of the one derived in the tracking system. To cope with the large detection surface to be covered as well as the different radiation environments and muon rates in the barrel and endcap part of the system, three different types of gaseous detectors are used (cf. Fig. 3.10).

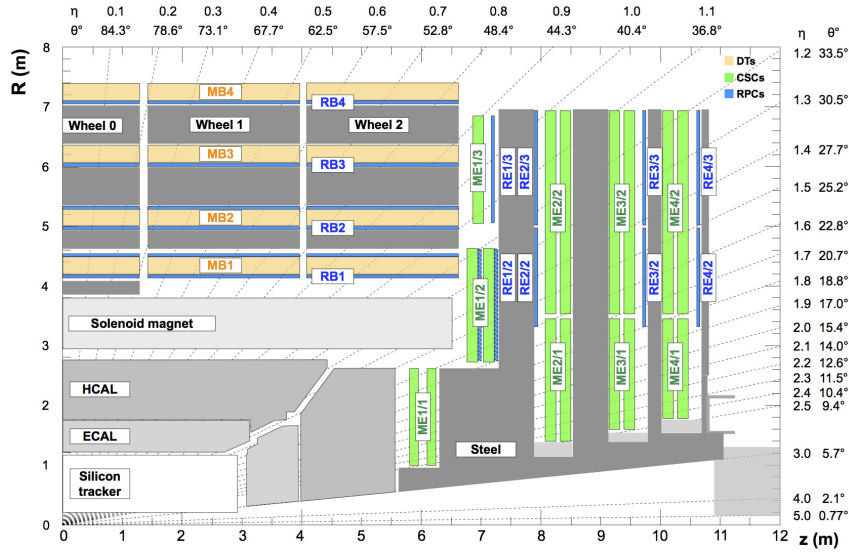


Figure 3.10: Sketch of one quarter of the CMS muon system, which is partially embedded in the return yoke (dark grey), as used for the data-taking periods between 2016 and 2018 shown in the longitudinal plane [101]. The drift tubes (DTs) are shown in orange, the cathode strip chambers (CSCs) in green and the resistive plate chambers (RPCs) in blue. The remaining CMS subdetectors are shown in light gray and white.

In the barrel part of the muon system, which covers a range up to $|\eta| < 1.2$, a total of 250 drift tube chambers are installed in four layers, also called stations (MB1-MB4), at radial distances of approximately 4 to 7 m from the beam axis. Each chamber in the first three layers contains a total of twelve drift tubes (DTs), where the four inner- and outermost DTs are built from $r\phi$ -measuring planes, while the four DTs in the middle correspond to z -measuring planes. In the chambers of MB4, no z -measuring planes are installed. Inside the DTs, which are filled with a gas mixture of 85% Ar and 15% CO₂, muons have a maximal drift length of 2.0 cm. With the design of the chambers, a spatial resolution of about 100 μm and an angular resolution of about 1 mrad can be achieved. The DT chambers are complemented by a total of 480 resistive plate chambers (RPCs), which due to their precise time resolution yield an important input to the CMS hardware trigger (cf. Sec. 3.2.7). As shown in Fig. 3.10, the first two muon stations in the barrel are sandwiched between two layers of RPCs, while for MB3/4 only one layer of RPCs is installed in front of the station.

The endcap part of the muon system, covering a range up to $|\eta| < 2.4$, consists of a total of 468 cathode strips chambers (CSCs), which can cope with the higher muon rates and the non-uniform magnetic field present in the endcap region. The CSCs are grouped into four muon endcap stations (ME1-ME4), installed at longitudinal distances between about 5.7 to 10.6 m from the interaction point. The chambers consist of six gas gaps, which are each filled with a plane of radial cathode strips and anode wires running perpendicular to the strips. While the wire planes enable fast signals, which are also used in the CMS hardware trigger, the position measurement is improved by including the slower signals from the cathode strips. This results in a spatial resolution of about $200\ \mu\text{m}$ ($100\ \mu\text{m}$ for ME1/1) and an angular resolution of about 10 mrad . Similar to the barrel part, the CSCs in the endcaps are complemented by RPCs covering a range up to $|\eta| < 1.9$ [101]. The distribution of the RPC layers in the endcap part is shown in Fig. 3.10.

An estimation of the muon momentum resolution achieved with the muon system is shown in Fig. 3.11. The estimation is based on a method introduced in [101] and employs the difference of the momentum measurements in the upper and lower half of the detector using cosmic ray muons, which cross the detector. For low- p_T muons, the resolution is dominated by the momentum measurement derived in the tracking system, while at transverse momenta above $\sim 200\text{ GeV}$ the measurement from the muon system significantly improves the resolution. This effect is caused by multiple scattering, which occurs when muons transverse the material before the muon system and has a much larger impact on low- p_T muons.

As pointed out above, parts of the muon system contribute to the muon hardware trigger of the CMS detector, introduced in Sec. 3.2.7. An example of an efficiency measurement for this trigger is shown in Fig. 3.11. Across all data-taking periods between 2016 and 2018, the trigger shows a sharp turn-on at the p_T threshold and a constant efficiency of about 96% across the measured muon- p_T range.

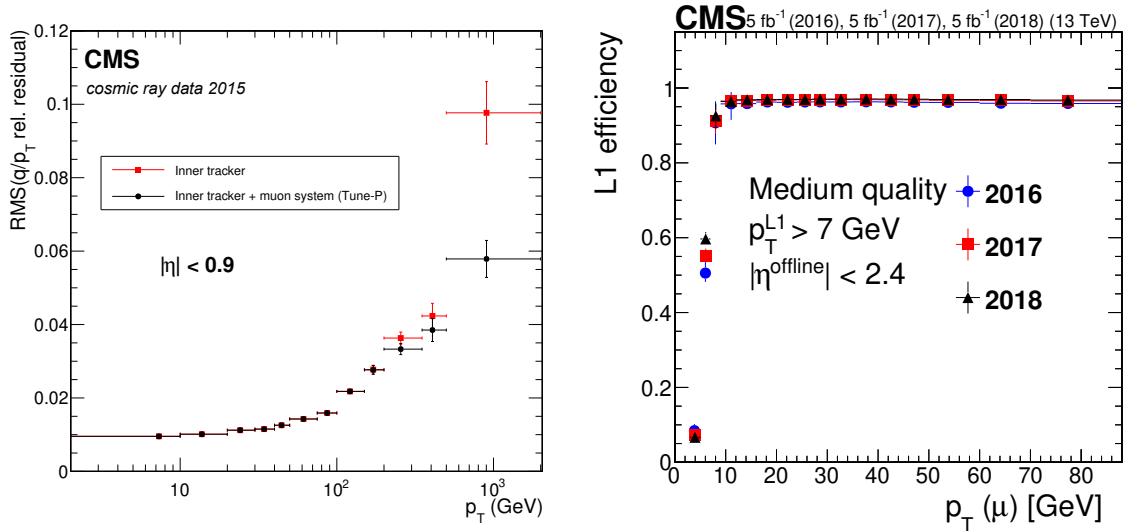


Figure 3.11: Left plot: Estimation of the muon momentum resolution as a function of the muon p_T based on the method introduced in [101] for muons with $|\eta| < 0.9$ comparing tracking system measurement (red) and the combined measurement (black) for cosmic muons. Right plot: Trigger efficiency for the CMS hardware muon trigger as a function of muon p_T for the data-taking periods of 2016 (blue), 2017 (red) and 2018 (black), based on medium quality requirements [102]. The vertical bars in both plots correspond to statistical uncertainties.

3.2.7 Trigger system

A dedicated trigger system [103, 104] is used to reduce the enormous event rates of the order of 10^9 Hz, which are produced at the LHC running with 40 MHz bunch crossings, by six orders of magnitude. To cope with the high rates and to ensure that predominantly events of potential interest for physics analyses are stored to permanent storage by the data acquisition system (DAQ), the trigger system is divided into two main levels. The first level of the system (L1 trigger), which is entirely hardware-based, is dedicated to a fast decision process based on simple detector information. The second level, also called high-level trigger (HLT), corresponds to a software-based trigger and processes more complex detector information and event reconstruction.

At the L1 trigger, information from both calorimeter systems as well as the muon system are processed by programmable electronics towards a trigger decision. For this, data with a reduced resolution and granularity from the three subdetectors are used to build so-called trigger primitive objects, which correspond to photons, electrons, muons or jets. The trigger decision is then based on the presence of such objects with a given energy or momentum threshold or the combination to global sums like the total transverse energy or the missing transverse energy. The L1 trigger processes an event within a fixed trigger latency of $4\,\mu\text{s}$, during which the high-resolution data is held in pipelined memory. The event rate is reduced by the L1 trigger from $\mathcal{O}(10^9\text{ Hz})$ to a maximal design rate of 100 kHz, which is usually not fully employed to keep a safety margin. An example of the efficiency of a particular L1 trigger can be found in Fig. 3.11.

After an event has been accepted by the L1 trigger, the high-resolution event data are read from multiple front-end readout buffers. Combining the information from several readout buffers, the complete data of one single event are transferred to a processor, which runs the high-level trigger software. The HLT software contains similar reconstruction and selection algorithms as used in the CMS offline reconstruction (cf. Sec. 6), which in the case of the HLT are usually tuned towards shorter processing times. These algorithms, which unlike the L1 trigger also have access to information from the tracking system, are usually employed in different steps. First, only partial reconstructions of the event are produced aiming to discard an event as fast as possible. To maximize the filter efficiency, only during later steps the full event information including a full reconstruction of the tracks is processed and used for the trigger decision. At the HLT, the maximal incoming event rate of 100 kHz is reduced by another two orders of magnitude yielding an event rate of $\mathcal{O}(1\text{ kHz})$, which is stored to permanent storage. During the data-taking period of 2018, a computing farm of up to 30000 cores was employed at the HLT to achieve this kind of rate reduction [105].

Chapter 4

Alignment of the CMS tracker

4.1	Track-based alignment	37
4.1.1	Alignment algorithm	38
4.1.2	Validation methods	39
4.2	Impact of radiation effects	41
4.3	Automated alignment in the Prompt Calibration Loop	43
4.3.1	Low-granularity alignment for Run 2	43
4.3.2	High-granularity alignment for Run 3	44

The tracking system of the CMS detector is the world’s largest silicon tracker and is crucial for the vertex reconstruction and momentum measurement of charged particles. To capitalize on the excellent hit resolution of its silicon sensors, their positions and orientations in the tracking system have to be precisely known. At the CMS detector, a procedure based on reconstructed tracks, also called track-based alignment, is used to derive these sensor-specific parameters. To mitigate the impact of magnet cycles and temperature variations and to cope with the aging of the sensors, the geometry of the tracking system has to be updated regularly. For this, an automated online alignment procedure for the inner pixel detector was implemented for Run 2 of the LHC, which supplies geometries on short-time scales supporting the manually derived offline alignments. During Run 2, this automated alignment was only based on movements of larger-scale structures of the pixel detector. Within the scope of this thesis, a detailed discussion of the method and the performance of the automated alignment used during Run 2 has been published [94]. For Run 3, the automated alignment procedure has been updated as part of this thesis towards an alignment with a significantly increased granularity. In this chapter, the general concept of track-based alignment including the alignment algorithm and different validation methods are briefly explained in Sec. 4.1. In Sec. 4.2, the interplay between sensor aging due to radiation and the tracker alignment is discussed, while in Sec. 4.3, the automated alignment procedure including its high-granularity update for Run 3 is reviewed.

4.1 Track-based alignment

The main goal of the track-based alignment [94] used for the CMS tracking system is to improve the precision of the position measurement of the individual components from $\mathcal{O}(100\mu\text{m})$ achieved by the mechanical alignment down to $\mathcal{O}(10\mu\text{m})$ matching the hit resolution of the silicon sensors. Without this precise alignment, biases in the track and vertex reconstruction could

be introduced, leading, e.g., to degradation in the identification of jets from b quarks or inefficiencies already at trigger level when tracker information is used at the HLT. The parameters of the correction, also called alignment parameters, are derived by performing track fits based on the hits in the tracking system and minimizing the distance (residuals) between fitted tracks and measured hits, which is sketched in Fig. 4.1 and further discussed in the next section.



Figure 4.1: Sketch of a comparison between a fitted trajectory (solid blue) with misaligned (left) and aligned modules (right) [106]. The true trajectory of the charged particle is shown as a dashed blue line, while the predicted and measured hits are shown as blue-red and orange-red dots, respectively. The residuals between both hit types are visualized as solid red lines.

The alignment parameters usually include three translational and three rotational degrees of freedom¹ per aligned structure, leading to a total of several hundreds of thousands of parameters, if each module of the pixel and strip detector is individually aligned. Such complex alignments with high granularity, also called offline alignments in the following, are usually performed manually and only a few times during each data-taking period. Movements of larger-scale structures of the pixel detector, like the two half barrels of the BPIX and the two half cylinders of each FPIX endcap, are derived on much shorter time scales by an automated online alignment. Besides the difference in granularity, also the type of events differ between the offline and online alignment. The online alignment is only based on collision data, while for the offline alignment also specialized data sets including events containing dimuon resonances or cosmic muons, which cross the entire detector, are used to correlate groups of modules in different detector regions.

4.1.1 Alignment algorithm

As pointed out in the previous section, the alignment parameters are derived based on minimizing the residuals between the hits predicted by a fitted trajectory and the hits measured by the modules (cf. Fig. 4.1). This track-based approach is formalized in the following χ^2 function:

$$\chi^2(\mathbf{p}, \mathbf{q}) = \sum_j^{\text{tracks}} \sum_i^{\text{hits}} \left(\frac{m_{ij} - f_{ij}(\mathbf{p}, \mathbf{q}_j)}{\sigma_{ij}} \right)^2, \quad (4.1)$$

where the χ^2 is a function of $\mathbf{p}$ corresponding to the alignment parameters and $\mathbf{q}$ representing the track parameters. The measured hits are given by m with a measurement uncertainty of σ , while the hits predicted by the fitted track are denoted by f . The χ^2 function is derived by summing over all hits of all tracks in a given data set of events.

¹Here, surface deformations of modules, which are usually only used for offline alignments, are neglected.

To solve the minimization problem, which, depending on the number of alignment parameters and tracks, can have millions of parameters, the prediction term f is linearized. This is achieved by assuming only small corrections $\Delta\mathbf{p}$ and $\Delta\mathbf{q}$ from an initial set of parameters $(\mathbf{p}_0, \mathbf{q}_0)$, which are usually derived from a starting geometry. The resulting set of linear equations is further reduced using block matrix algebra effectively removing blocks of equations that determine only the track parameters. This method yields a reduced set of linear equations

$$\mathbf{C}' \times \Delta\mathbf{p} = \mathbf{b}', \quad (4.2)$$

which includes the corrections $\Delta\mathbf{p}$ of the alignment parameters, whose correlation with the track parameters is encoded in the global matrix $\mathbf{C}'$ and the source term $\mathbf{b}'$. To derive the corrections $\Delta\mathbf{p}$, the matrix $\mathbf{C}'$ has to be inverted, which can be performed exactly in case of a small number of alignment parameters, but is usually achieved with approximative methods.

At the CMS detector, the task of collecting all necessary inputs to construct the ingredients of Eq. (4.2) and solving the set of linear equations by inverting the global matrix is mainly performed by the MILLEPEDE-II algorithm [107–109]. The algorithm is structured in two steps. The first step, also called mille step, is integrated into the CMS track fitting software and produces the residuals for each hit of a track including uncertainties as well as the derivatives required for the linearization of Eq. (4.1). In the second step, called pede step, the linear equation system shown in Eq. (4.2) is constructed and solved by inverting the global matrix either with an exact or approximative method. While the alignment results discussed in this work are derived using the MILLEPEDE-II algorithm, results from an independent algorithm, called HIPPY algorithm [110, 111], have been mainly used in the commissioning of the detector and to cross-check the results derived by the MILLEPEDE-II algorithm.

4.1.2 Validation methods

The performance of geometries derived in the track-based alignment is evaluated based on several different validation methods. While all of these methods in some way test the accuracy achieved by the alignment process, a part of the validation methods has been developed to also detect specific systematic misalignments of the detector. These systematic misalignments include the so-called weak modes, which are systematic distortions that transform a set of valid tracks into another set of valid tracks and have a negligible impact on the χ^2 minimization. A full discussion of systematic distortions and the validation methods employed to detect such distortions can be found in [94]. In the following, two validation methods, which are mainly studied when evaluating the performance of the automated online alignment, are further discussed.

Track-hit residuals

One of the main validation methods to evaluate the accuracy of the local alignment of individual modules is based on track-hit residuals. The residuals are derived by refitting a track without the hit of interest and measuring the distance between the measured hit and the hit predicted by the refitted track with respect to the local module coordinates². Based on these residuals, distributions of median residuals (DMRs) per module are derived for individual substructures of the tracking system. For a perfect alignment, the distributions are expected to be centered

²The local coordinates are defined for each module, with the x' -direction corresponding to the most precisely measured coordinate. The y' -direction is orthogonal to x' in the module plane and z' is orthogonal to the module plane.

around zero, while shifted distributions would indicate a systematic misalignment. The width of the distributions is a measure of the remaining random misalignments, but also includes a statistical component based on the number of tracks used for the validation. An example of a DMR validation for the BPIX comparing three different geometries is shown in Fig. 4.2. The alignment derived during data taking, which is mainly based on the automated online alignment correcting the high-level structures of the pixel detector, shows a rather broad distribution. The two remaining alignments, which utilize a much finer granularity as well as additional specialized data sets, show a reduced width indicating an improved accuracy derived in the alignment.

In addition to the nominal DMR validation, where the median for each module of a substructure is taken into account, so-called split DMR validations are derived by separating different types of modules. The separation is performed based on the orientation of the electric field distinguishing between modules with an electric field pointing towards the center of the detector (inward-pointing) and towards the muon system (outward-pointing). The difference between the means of the DMRs derived for the two different module types can be used as a measure of the residual bias caused by radiation in the silicon sensors, further discussed in Sec. 4.2.

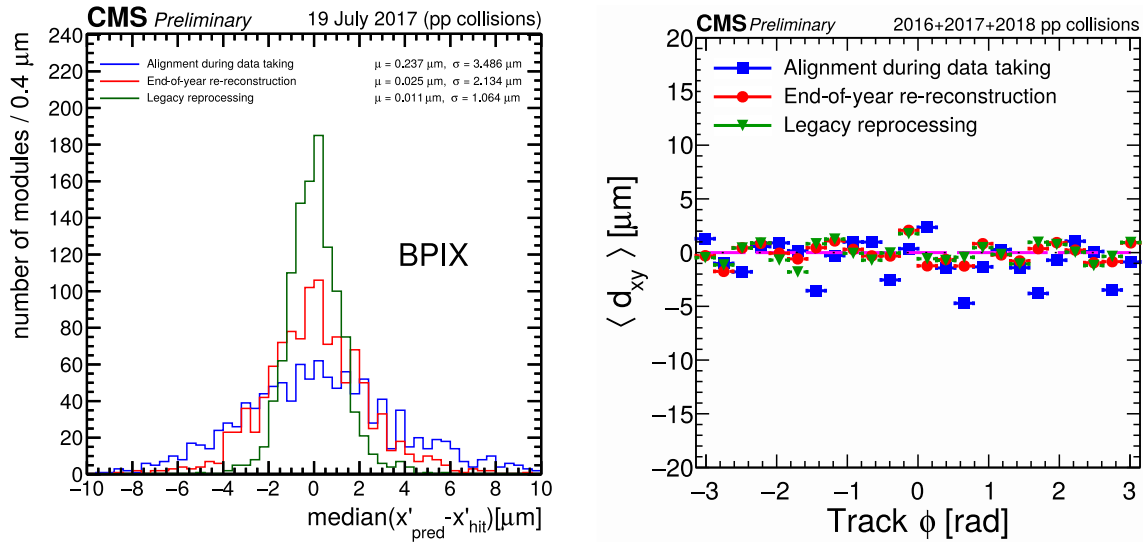


Figure 4.2: Left plot: DMRs for each module of the BPIX for the local x' -direction measured for an example run in 2017 [112]. The mean and the width of the distributions are shown in the legend. Right plot: Mean track-vertex impact parameter in the transverse plane as a function of track azimuthal angle averaged for the data-taking periods between 2016 and 2018. For both plots, the results from an alignment during data taking (blue), an improved alignment derived during the end-of-year re-reconstruction (red) and the final legacy alignment (green) are compared.

Track-vertex residuals

Since the reconstruction of the primary vertex (PV) is mainly driven by the inner pixel detector, one of the main validation methods to evaluate the alignment of the pixel detector is based on the residuals between reconstructed vertices and tracks. In this method, the PV is reconstructed for each associated track without the track of interest. The track-vertex residual corresponds to the distance between the reconstructed PV and the given track measured by the impact parameter in the longitudinal (d_z) and transverse plane (d_{xy}). The mean of the impact parameters as a function of the track azimuth and pseudorapidity yields a measure of the alignment quality.

A perfect alignment would lead to flat distributions centered around zero. The same holds for a geometry including random misalignments, which would only affect the width but not the mean of the track-vertex residuals. Thus, significant deviations from a flat distribution would indicate a systematic misalignment. An example of a validation based on the transverse track-vertex residuals derived as a function of track azimuth is shown in Fig. 4.2, which indicates that the alignment derived during data taking might not be able to correct for all systematic misalignments.

4.2 Impact of radiation effects

The main goal of the track-based alignment is to correct for physical movements of the detector modules caused by, e.g., cycles of the magnet or temperature variations. However, it has been observed that an alignment based on a sufficiently high granularity can also account for shifts of the measured hit position on an individual module. Such shifts occur based on changes in the characteristics of a module being operated in a high-radiation environment, including changes in the electric field, the thickness of the active layer and the mobility of charge carriers.

The Lorentz drift of the charge carriers inside a module, which is sketched in Fig. 4.3, is sensitive to the impact of radiation effects on an individual detector module. A charged particle crossing a module creates electron-hole pairs, which are collected by an electric field resulting in a cluster charge, from which the hit position can be derived. Due to the external magnetic field, the charge carriers experience a lateral shift orthogonal to the electric field of the module, also called Lorentz drift. While changes in the external magnetic field are negligible, the size of the drift strongly depends on the strength of the electric field as well as on the mobility of the charge carriers, which is affected by the irradiation of a module and therefore highly time-dependent. The impact of the shift on the hit position is quantified by the Lorentz angle (LA), which is defined as the angle between the electric field and the drift direction of the charge carriers. The Lorentz angles of individual modules, which can be used to correct the shifted hit position, are derived based on dedicated calibration methods [92, 113].

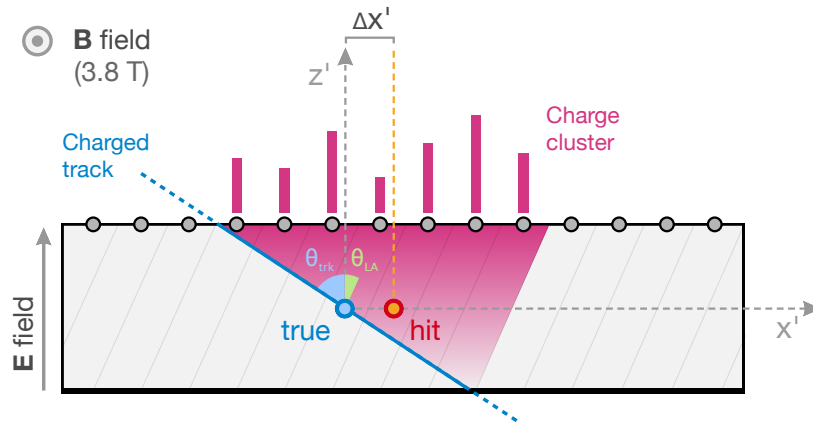


Figure 4.3: Sketch of a transverse slice of a silicon module, which operates in an external magnetic field and is traversed by a charged particle (blue line) with an incident angle of θ_{trk} [94]. The true and shifted hit positions, which have a residual of $\Delta x'$, are shown as blue-cyan and red-orange dots, respectively. The grey lines indicate the direction of the Lorentz drift, whose angle to the electric field is given by θ_{LA} . The volume where charge carriers are released is visualized by the magenta-shaded area, while the corresponding cluster charge is shown as magenta bars.

While the calibrations of the LA are usually only performed on larger time scales during data-taking, the alignment of the detector can be used to absorb residual effects caused by radiation-induced changes of the LA. Since the sign of the hit shift caused by the Lorentz drift depends on the orientation of the electric field with respect to the external magnetic field, the hit position is shifted in opposite directions for modules with inward- and outward-pointing electric fields. An alignment, which distinguishes between these two module types, can therefore introduce artificial movements of the modules to correct for the shifted hit positions. This correction can be validated based on the split DMR validation by monitoring the difference $\Delta\mu$ between the means of the DMRs for modules with inward- and outward-pointing electric fields. Changes in this parameter as a function of the delivered integrated luminosity, which can be used as a proxy for the radiation dose, would indicate residual effects caused by radiation, which are not absorbed in the alignment or dedicated calibrations.

In Fig. 4.4, the trend of the $\Delta\mu$ parameter for the data-taking periods between 2016 and 2018 comparing different alignments is shown. For the alignment during data taking, a continuous slope in $\Delta\mu$ as a function of the delivered integrated luminosity can be observed. This continuous change in $\Delta\mu$ is caused by cumulative radiation effects, which can not be absorbed by the alignment during data taking. The reason is the granularity used in the automated online alignment employed during Run 2 (cf. Sec. 4.3.1), which is too coarse and does not distinguish between the two module types discussed above. The slope is only absorbed by updates of the LA calibration or manually derived offline alignments with a higher granularity usually provided with the pixel calibration updates. For the other two sets of alignments, the $\Delta\mu$ parameter is significantly more stable due to the higher granularity used in these approaches. Here, the largest differences between the two offline alignments can be observed at the beginning of the 2017 data-taking period, when the new CMS pixel detector was put into operation. In this period, only the legacy reprocessing, which includes much more frequent geometry updates, is able to absorb the radiation effects, which are especially large at the beginning of a detector's lifetime.

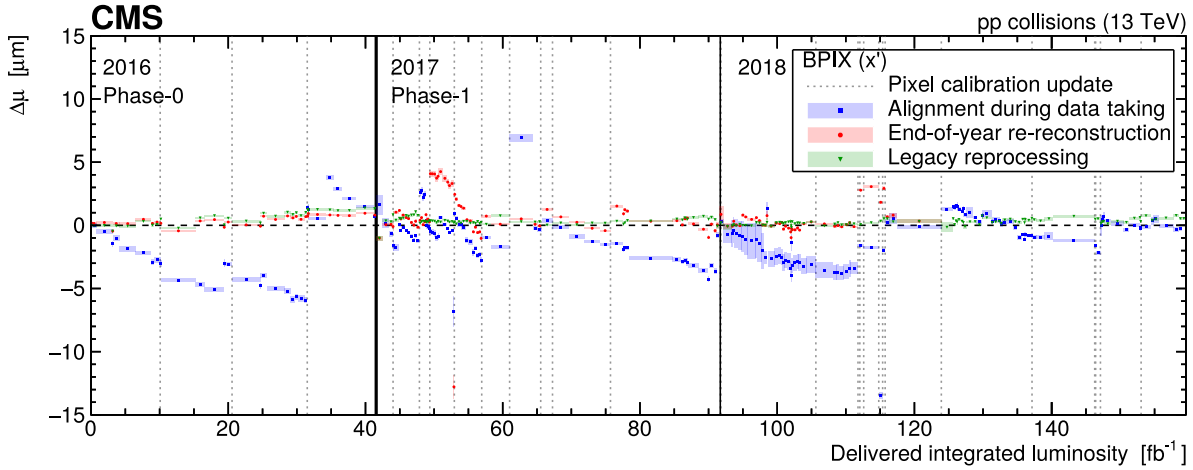


Figure 4.4: Difference $\Delta\mu$ between the means of the DMRs with respect to the local x' -direction for modules of the BPIX with inward- and outward-pointing electric fields as a function of the delivered integrated luminosity [94]. The results for the data-taking periods between 2016 and 2018 from an alignment during data taking (blue), an alignment derived during the end-of-year re-reconstruction (red) and the final legacy alignment (green) are compared. The vertical bands correspond to the widths of the DMRs. The horizontal bands indicate the interval of validity for a given alignment. Updates of the pixel calibration are shown as vertical dashed lines.

4.3 Automated alignment in the Prompt Calibration Loop

For Run 2 of the LHC, an automated online alignment, which aims to correct movements of the pixel detector components on a short time scale, has been put in place. Based on this online alignment, the alignment parameters of the inner pixel detector can be updated within 48 hours after the end of an individual collision run to be used for the first main reconstruction available for physics analysis (prompt reconstruction). For this, the online alignment runs as part of the prompt calibration loop (PCL) [114], which processes different calibration and alignment routines and automatically updates the corresponding detector-related parameters. This enables a much faster response to potential movements compared to the nominal offline alignments, which require significant preparation and computing time and therefore usually lead to a delay of the order of one week.

4.3.1 Low-granularity alignment for Run 2

The online alignment procedure employed during Run 2 of the LHC has been designed to correct the movements of the high-level structures of the pixel detector, which include the two half barrels of the BPIX and two half cylinders in each of the two endcaps of the FPIX. For each of the substructures, changes in the position (x, y, z) and the rotation $(\theta_x, \theta_y, \theta_z)$ are derived based on the MILLEPEDE-II algorithm leading to a total of 36 alignment parameters. The online alignment is performed based on events selected by an inclusive L1 trigger selection and does not include any specialized data sets. A new set of parameters is triggered if at least one movement with respect to the previous geometry is observed, which has a significance of at least 2.5σ and lies above a given threshold. The thresholds for the positions and rotations are given in Tab. 4.1. In addition, the uncertainty of each movement derived by the alignment procedure is checked as well as the compliance with a maximal movement, which prevents unphysical shifts from being deployed. If one out of the 36 parameters does not fit the requirements regarding maximal uncertainty or maximal movement, which are shown in Tab. 4.1, the update of the geometry is vetoed.

Table 4.1: Threshold values, maximal movements and maximal uncertainties used for the update logic of the automated online alignment employed during Run 2. The values are shown for the six different parameter types and hold for all high-level structures of the pixel detector.

Parameter	Threshold	Max. Movement	Max. Uncertainty
Δx	$5 \mu\text{m}$	$200 \mu\text{m}$	$10 \mu\text{m}$
Δy	$10 \mu\text{m}$	$200 \mu\text{m}$	$10 \mu\text{m}$
Δz	$15 \mu\text{m}$	$200 \mu\text{m}$	$10 \mu\text{m}$
$\Delta\theta_x$	$30 \mu\text{rad}$	$200 \mu\text{rad}$	$10 \mu\text{rad}$
$\Delta\theta_y$	$30 \mu\text{rad}$	$200 \mu\text{rad}$	$10 \mu\text{rad}$
$\Delta\theta_z$	$30 \mu\text{rad}$	$200 \mu\text{rad}$	$10 \mu\text{rad}$

The success of the automated online alignment can be seen in Fig. 4.5, which shows the movement in the x -direction of all six high-level structures for a specific data-taking period in 2016. During this period, three individual magnet cycles were performed, which can cause rather large movements of the high-level structures. After each of these cycles, the online alignment observed significant movements in parts of the pixel detector, which triggered the deployment of a new

geometry. In all three cases, the movements observed in the following alignment are significantly reduced indicating that the update was able to correct for the movements caused by the magnet cycles.

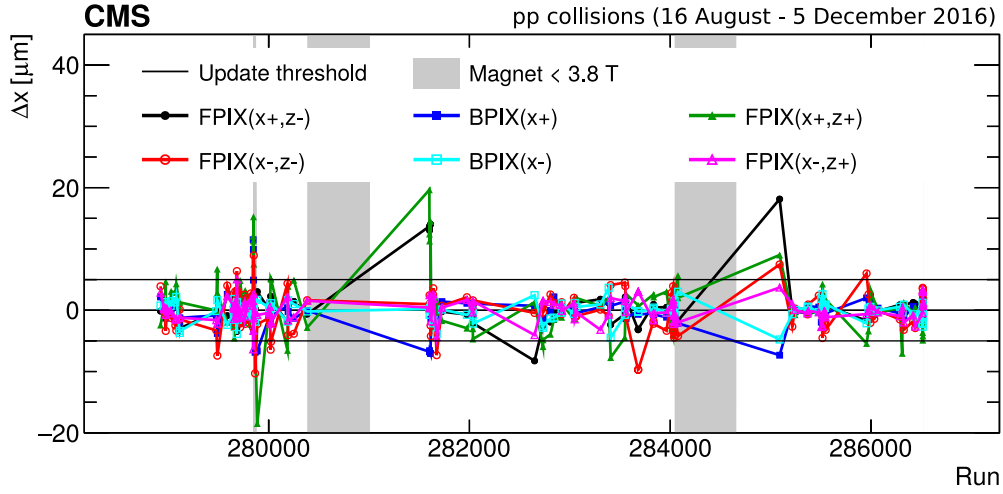


Figure 4.5: Observed movements in the x -direction for the six high-level structures of the inner pixel detector as a function of run number for a specific data-taking period in 2016 [94]. Each point corresponds to a run, where the deployment of a new geometry was triggered by at least one of the 36 alignment parameters derived in the online alignment. The thresholds for movements in the x -direction are indicated as horizontal lines, while the periods during which the magnet was not at 3.8 T are shown as grey-shaded bands.

In summary, the online alignment employed during Run 2 was able to successfully cope with movements of the larger structures of the pixel detector as induced by magnet cycles or temperature variations. However, due to the limited granularity used in this alignment procedure, it did not succeed in absorbing the radiation effects as shown in Fig. 4.4 and discussed in Sec. 4.2.

4.3.2 High-granularity alignment for Run 3

Motivated by the shortcomings of the automated online alignment regarding the absorption of residual radiation effects during Run 2, an improved version of the online alignment for Run 3 has been developed within the scope of this thesis. The main change applied to the online alignment concerns the granularity used in the alignment procedure, which is increased from the high-level structures of the pixel detector, denoted as low granularity (LG), to a granularity incorporating ladders in the BPIX and individual modules in the FPIX, denoted as high granularity (HG). This increase in granularity is sufficient to separate modules with inward- and outward-pointing electric fields (cf. Sec. 4.2), since all modules of an individual BPIX ladder have the same electric-field orientation. Another motivation for this update of the online alignment is the potential reduction of the number of manual offline alignments required during data-taking periods.

All main challenges introduced by the higher granularity concern the increased complexity of the alignment problem, where the number of alignment parameters grows from 36 parameters considered during Run 2 to more than five thousand parameters. Due to the increased complexity, the absolute computing time required to derive the alignment parameters with the MILLEPEDE-II algorithm grows significantly. Nevertheless, the online alignment has to run with-

out any manual intervention and is required to supply the updated alignment parameters within 48 hours to be used in the prompt reconstruction process. Another challenge concerns the increased number of tracks and hits required for the alignment, which still have to be supplied by individual collision runs. Finally, also the update logic employed during Run 2 has to be revised, to cope with the increased number of parameters.

Hits per substructure

In the LG online alignment, the movement of a structure is only derived if at least 500 hits for the given structure are recorded. While this requirement is usually not a limitation when aligning only the high-level structure, for the increased granularity its impact has to be studied. For this, the HG online alignment was performed with 50, 300 and 500 required hits per structure using dedicated runs with low pileup, which include significantly fewer tracks per bunch crossing. The track-vertex residuals for one of these runs are shown in Fig. 4.6. The alignment based on 50 required hits shows a significantly reduced performance. The remaining two alignments perform on a similar level, with small improvements from the alignment with 300 and 500 required hits in the azimuth- and pseudorapidity-dependent validation, respectively. Based on this study and on the observation that also for the HG alignment the requirement of 500 hits per structure is usually fulfilled in nominal pp-collision runs, it was decided to keep the conservative option and align only structures with at least 500 hits.

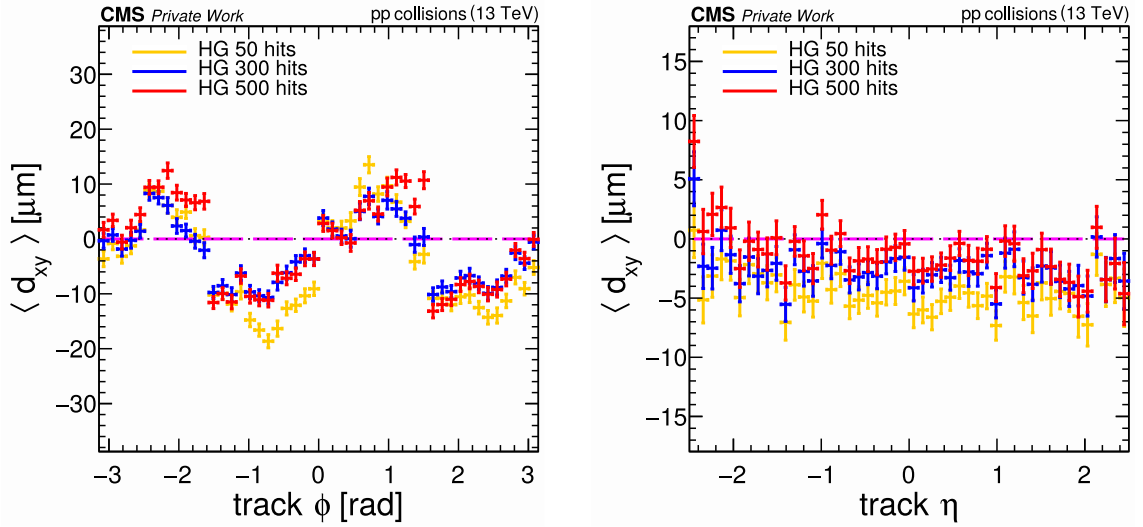


Figure 4.6: Mean track-vertex impact parameter in the transverse plane as a function of track azimuthal angle (left) and track pseudorapidity (right) for a dedicated low-pileup run in 2018. The results for the HG online alignment based on 50, 300 and 500 required hits per structure are shown in yellow, blue and red, respectively. The vertical error bars correspond to the statistical uncertainty.

New update and veto logic

The HG online alignment requires a revised update and veto logic to cope with the increased number of parameters. The logic used during Run 2 is not feasible anymore, since individual outliers, which are expected when aligning more than five thousand parameters, would essentially trigger the deployment of a new geometry after each run. This would lead to a large frequency

of geometry updates, which are only based on statistical fluctuations or on movements of small individual structures, which have no significant impact on the track reconstruction. This problem is solved by implementing a collective threshold logic per substructure of the pixel detector. In this context, the substructures of the pixel detector distinguish between the different layers of the BPIX and the different disks of the FPIX. The collective logic triggers an update of the geometry if a given fraction of aligned ladders (modules) in at least one layer (disk) shows a significant change above a given threshold in the same parameter type. The values for the thresholds as well as the fractions used for the collective threshold for the six different parameter types are shown in Tab. 4.2. The threshold values are kept unchanged compared to the set used in the online alignment during Run 2 (cf. Tab. 4.1). The fraction parameters, however, have been optimized based on dedicated long-term studies reanalyzing pp-collision runs of the 2018 data-taking period using fraction values between 0 and 0.5. It has been observed, that the choice of the fraction parameters has no significant impact on the performance of the alignment. To avoid updates of the geometry triggered by each run while still keeping a reasonably large update frequency, a fraction parameter of 0.25 is selected for all parameter types.

Table 4.2: Threshold values, threshold fractions and maximal movements for the BPIX and FPIX structures used for the update logic of the automated online alignment employed during Run 3. The values are shown for the six different parameter types.

Parameter	Threshold	Fraction	Max. Movement BPIX	Max. Movement FPIX
Δx	5 μm	0.25	150 μm	200 μm
Δy	10 μm	0.25	150 μm	200 μm
Δz	15 μm	0.25	150 μm	300 μm
$\Delta\theta_x$	30 μrad	0.25	150 μrad	2000 μrad
$\Delta\theta_y$	30 μrad	0.25	2000 μrad	5000 μrad
$\Delta\theta_z$	30 μrad	0.25	150 μrad	1000 μrad

In the same long-term studies, the values for the maximal allowed movements were optimized by analyzing the typical size of the movements derived in the HG online alignment. In contrast to the setup used for Run 2, the thresholds for the maximal movements distinguish between BPIX and FPIX, since larger and more frequent outliers have been observed in the alignment of the individual FPIX modules compared to the BPIX ladders. While this difference is especially present in the rotational parameters, for both types of structures, only small sensitivities for rotations around the y -axis ($\Delta\theta_y$) have been observed motivating the large values for the maximal movements regarding this parameter. A check of movement uncertainties with respect to a maximal threshold is currently not used in the veto logic.

Performance during Run 3

While all the components of the HG online alignment were in place for the beginning of Run 3 in mid-2022, the first set of pp-collision runs of Run 3 was used to validate the performance of the new alignment setup without enabling the automated deployment of geometry updates. During this period, the automated deployments were still produced by the LG online alignment, allowing for comparisons between the performance of the HG and LG online alignments. An example of such a comparison is shown in Fig. 4.7, where the track-vertex residuals in the transverse plane

are compared for the two alignment setups. In both validations, the geometries derived by the HG online alignment show a significant improvement compared to the LG online alignment.

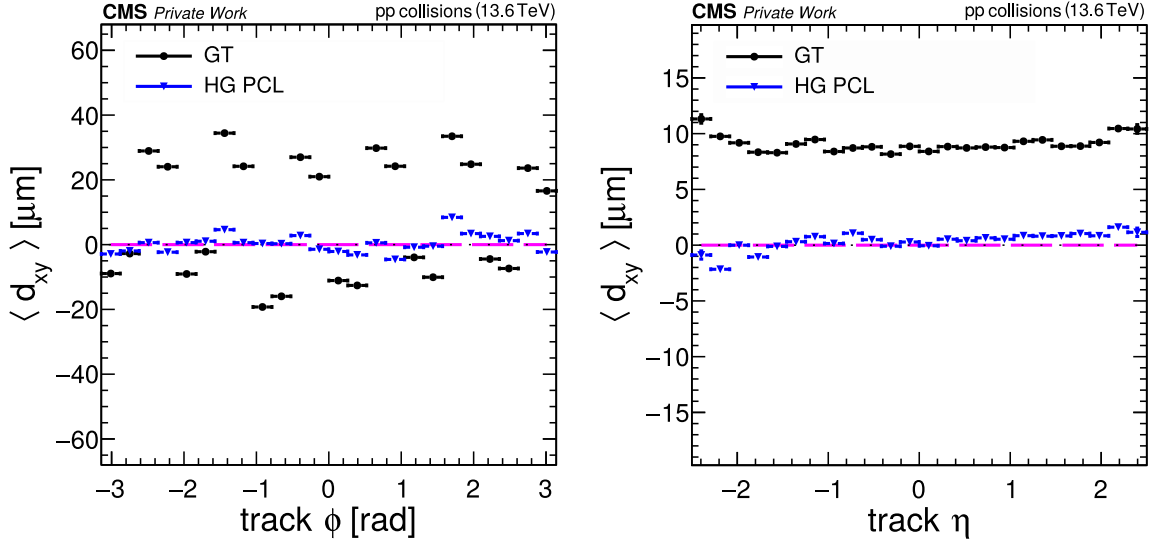


Figure 4.7: Mean track-vertex impact parameter in the transverse plane as a function of track azimuthal angle (left) and track pseudorapidity (right) for a random pp-collision run in 2022. The result for the HG online alignment is shown in blue and the result of the LG online alignment, which was used for the automated deployment in this run, is shown in black.

Similar improvements can be observed for the split DMR validation, which is shown in Fig. 4.8. As expected, the $\Delta\mu$ parameter for the BPIX is significantly reduced in the HG online alignment due to the separation between modules with inward- and outward-pointing electric fields. In the FPIX, the main improvement is observed with respect to the width of the combined DMR, which indicates a smaller residual misalignment in the HG online alignment. The difference between the widths of the DMRs in the BPIX and FPIX for the HG online alignment can be explained by the different granularities used for the alignments. While in the BPIX, ladders of eight modules are aligned, in the FPIX, the alignment is performed for individual modules allowing for a more precise correction of the geometry.

Based on these comparisons, it was decided to activate the new HG online alignment including the automated deployment of the geometry updates derived by the new HG setup. The performance of the HG online alignment for the runs of the 2022 data-taking period after its activation, which coincides with a technical stop of the LHC, can be seen in Fig. 4.9. For the period before the technical stop, the alignment during data taking was derived based on the LG online alignment and an additional manual offline alignment was required to cope with the remaining misalignments. For the period after the technical stop, the performance of the HG online alignment was sufficient to be used for the end-of-year re-reconstruction and no additional manual offline alignment was performed. Using the revised update and veto logic introduced above, new geometries have been triggered by about 45% of the nominal pp-collision runs³ in 2022, while the old setup employed during Run 2 only triggered updates in about 20% of the runs.

The trend of the mean of the DMR in the FPIX shows a significant improvement introduced by the activation of the HG online alignment. The validation for the LG online alignment yields a

³Here, only runs after the activation of the new HG online alignment are considered.

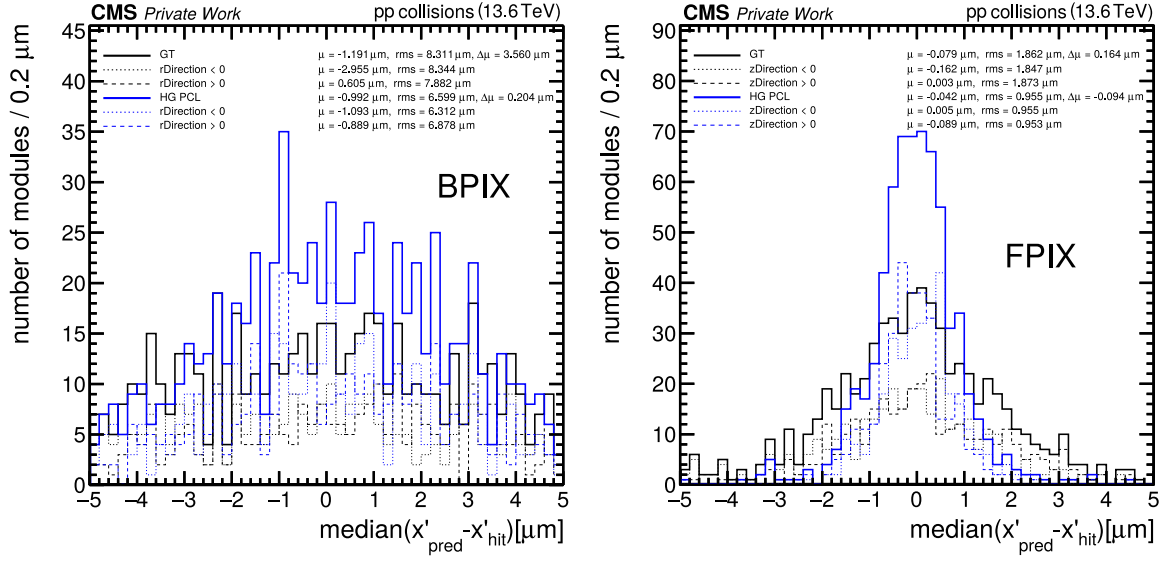


Figure 4.8: Distribution of median residuals for each module of the BPIX (left) and the FPIX (right) for the local x' -direction measured for a random pp-collision run in 2022. The results based on distinguishing between modules with different orientations of the electric field are shown as dashed and dotted lines, while the solid lines are based on the combined result. The mean and the width of the individual distributions as well as the difference between the mean of the two module types are shown in the legend. The result for the HG online alignment is shown in blue and the result of the LG online alignment, which was used for the automated deployment in this run, is shown in black.

constant increase in the mean, while with the HG online alignment, a flat mean close to zero can be achieved after the technical stop. This improvement in performance is mainly based on increasing the granularity from half cylinders to individual modules in the FPIX. The largest improvement can be found in the trend of the $\Delta\mu$ parameter only taking the first layer of the BPIX into account. Since this layer was replaced at the beginning of Run 3, a large impact of the radiation effects is expected for this layer. This is confirmed by the significant decrease of $\Delta\mu$ as a function of the delivered integrated luminosity visible in the LG online alignment, which is not able to account for the radiation effects. While the additional manual offline alignment is able to recover the flat behavior before the technical stop, for the remaining runs, the HG online alignment is able to cope with the radiation effects without any additional offline alignment. This clearly indicates the improvement introduced by aligning the individual ladders of the BPIX in contrast to considering only the two half barrels. In addition, also the trend of the RMS for the vertex-track residuals shows a significant improvement introduced by the HG online alignment, which supplies a quite stable performance across the luminosity range after the technical stop.

In summary, the performance during Run 3 shows that the new HG online alignment can absorb the residual radiation effects, as intended, by using a higher granularity and providing more frequent updates compared to the online alignment employed during Run 2. In addition, the necessity of manual offline alignments can be significantly reduced by the HG online alignment. However, the new online alignment is not able to completely replace the offline alignments, since the outer strip detector is not aligned and so far there is no method implemented to reduce the impact of weak modes by including specialized data sets (cf. Sec. 4.3). These two points could be potential goals for further developments of the online alignment.

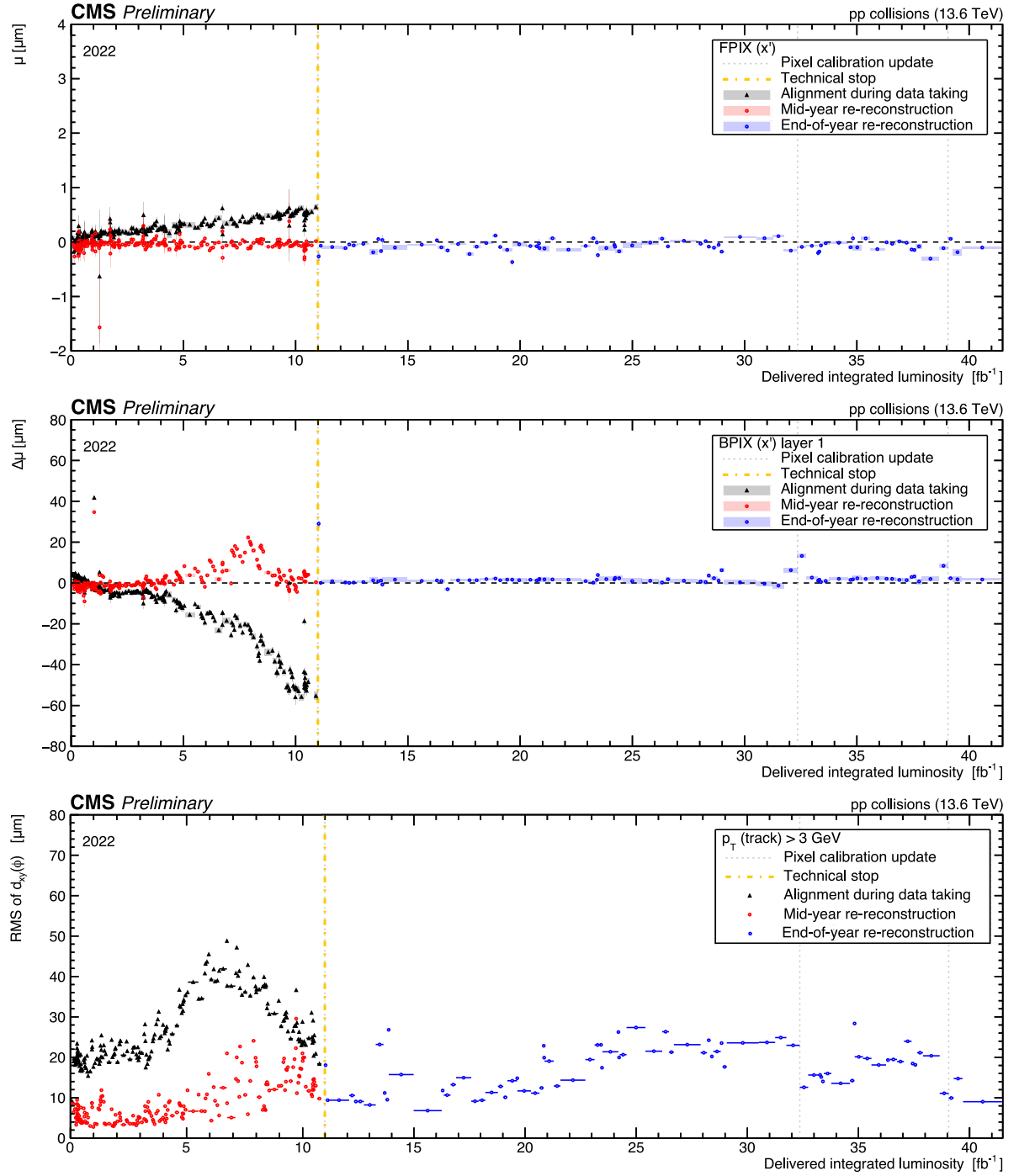


Figure 4.9: Different trends of validation parameters as a function of the delivered integrated luminosity [115]: Mean of the DMR derived for the FPIX (top). Difference between the means of the DMRs for modules with inward- and outward-pointing electric fields derived for layer 1 of the BPIX (center). RMS of the vertex-track residual in the transverse plane as a function of the track azimuthal angle (bottom). In all trends, the results from the LG online alignment (black), the HG online alignment (blue) and an offline alignment (red) are shown. The activation of the HG online alignment and the updates of the pixel calibration are indicated as yellow- and grey-dashed lines, respectively.

Chapter 5

Data sets and Monte Carlo simulation

5.1	Triggers and primary data sets	51
5.2	Monte Carlo simulation	53
5.2.1	Hard process	53
5.2.2	Parton shower, hadronization and underlying event	54
5.2.3	Detector simulation	55
5.2.4	Simulated samples	56
5.2.5	Interference between tW and $t\bar{t}$	58
5.3	Trigger efficiencies	60
5.3.1	Efficiency measurement	60
5.3.2	Scale factors	61

In this thesis, proton-proton (pp) collision data recorded by the CMS detector are used to derive differential cross section measurements for the $t\bar{t}$ process in the dileptonic channel. The collision data have been collected over a period of three years using different sets of high-level triggers matching the targeted event signatures. In addition to the recorded data, Monte Carlo (MC) simulation is employed in many parts of this analysis, e.g., to predict contributions from non- $t\bar{t}$ SM processes, to correct for detector effects such as resolution or acceptance or to interpret the final cross section results. The recorded pp-collision data sets and the corresponding high-level triggers used to collect these data are summarized in Sec. 5.1. The different steps towards a reliable Monte Carlo simulation and the simulated data sets used in this analysis are briefly introduced in Sec. 5.2. In the last part of this chapter (Sec. 5.3), the estimation of trigger efficiencies and the difference between the efficiencies in recorded data and simulation, which is one of the main corrections applied to the simulated samples, are discussed.

5.1 Triggers and primary data sets

The pp-collision data analyzed in this thesis were recorded by the CMS detector in the period between 2016 and 2018, during which the LHC provided pp collisions at a center-of-mass energy of $\sqrt{s} = 13$ TeV. The data set combined across the three years corresponds to a total integrated luminosity of 138 fb^{-1} [116–118]. To consider different detector setups and data-taking conditions, the recorded data set is split into four separate data-taking periods, whose partial integrated luminosities are listed in Tab. 5.1. The collision data recorded in 2015, which correspond to about 2 fb^{-1} , are not used in this analysis. The data set recorded in 2016 is split into

two separate data-taking periods motivated by a change of detector settings related to the strip tracker readout [119, 120]. At the beginning of 2016, inefficiencies in the hit detection by the strip tracker were observed. These inefficiencies were caused by the higher instantaneous luminosity, which resulted in a saturation of a readout chip used in the strip tracker. The efficiency was fully recovered by adapting the settings of the so-called Preamplifier Feedback Voltage Bias (VFP). To take this change into account, the recorded data set as well as the simulated samples are split into a period before the change (2016preVFP) corresponding to about 19.5 fb^{-1} and a period after the change (2016postVFP) corresponding to about 16.8 fb^{-1} .

Table 5.1: Integrated luminosity and corresponding uncertainty for each of the four data-taking periods [116–118].

Period	Int. Luminosity [fb^{-1}]	Uncertainty [%]
2016preVFP	19.52	1.2
2016postVFP	16.81	1.2
2017	41.48	2.3
2018	59.83	2.5
Run 2	137.64	1.6

While all four data-taking periods are based on the same collision energy, the instantaneous luminosity differs significantly between the periods [87], leading to different pileup distributions, as shown in Fig. 3.2. Here, the largest differences can be observed between both 2016 periods, which yield a mean pileup of 27 collisions per bunch crossing, and the 2017/2018 data-taking periods, which show on average 37 to 38 collisions per crossing.

The collision data analyzed in this thesis are selected according to the latest list of runs and lumisections¹ certified as “good for physics analysis” [121]. This certification is only granted if all subdetectors and triggers as well as physics objects, introduced in Sec. 6.2, show a reasonable performance. Furthermore, the latest and most precise set of calibrations, alignments and corrections, also known as legacy reprocessing, is applied to the data analyzed in this thesis.

A variety of different high-level triggers are used to collect the collision data. The triggers are mainly selected to match the expected event signature of two charged leptons (electrons or muons) produced by $t\bar{t}$ processes in the dileptonic channel. The transverse momentum thresholds employed in these dilepton triggers range from 17 to 23 GeV for the leading lepton and from 8 to 12 GeV for the subleading lepton depending on the lepton flavor and data-taking period. For all dilepton triggers, different sets of additional requirements are imposed. These requirements are, for example, based on the lepton isolation, identification criteria related to the track or calorimeter entries, constraints on the maximal longitudinal impact parameter or lower bounds on the invariant dimuon mass to reduce contributions from low-mass resonances like J/ψ . To account for potential inefficiencies of the dilepton triggers, a selection of single-lepton triggers is employed. Here, the requirements on the transverse momentum range between 24 and 27 GeV for muons and between 27 and 32 GeV for electrons. In addition to the lepton triggers, also denoted as signal triggers, triggers based on the missing transverse momentum are used for the estimation of the trigger efficiencies, which is further discussed in Sec. 5.3. The complete set of high-level triggers employed in this analysis can be found in Appx. 1.

Depending on the sets of triggers, which require similar high-level objects, the recorded data sets are divided into individual primary data sets (PDs), also known as data streams. In this analysis,

¹One lumisection corresponds to approximately 23 seconds of data-taking [114].

the DOUBLEMUON, the DOUBLEEG² and the MUONEG PDs containing dimuon, dielectron and electron-muon triggers, respectively, are analyzed. To increase the overall trigger efficiency, events from the SINGLEMUON and SINGLEELECTRON PD collected by the corresponding single-lepton triggers, but not selected by the employed dilepton triggers, are taken into account. In addition, the MET PD is used to provide event samples collected with independent p_T^{miss} -triggers, which are used in the measurement of the trigger efficiencies discussed in Sec. 5.3. The full list of PDs for all four data-taking periods including the paths used to access the data sets via the CMS data set bookkeeping service (DBS) [122] can be found in Appx. 2.

5.2 Monte Carlo simulation

The simulation of the SM $t\bar{t}$ process as well as other SM processes [123], which contribute to the background considered in this analysis, plays a large role in the measurement of the $t\bar{t}$ cross section. The applications of simulated event samples include the estimation of SM background contributions, the correction of detector effects and the interpretation of the final cross section results. All these parts of the analysis rely on an accurate simulation of the given processes, which usually contain hundreds of particles in the final state, whose momenta range across many orders of magnitude. The simulation of such events is split into different steps according to the scale of momentum transfer. The highest scales are covered by the hard process. At smaller scales, the parton shower followed by a potential hadronization of produced particles has to be taken into account. In addition, contributions to the final state not originating from the hard process, also known as underlying event, have to be generated. Finally, the interactions between the particles and the detector material resulting in the detector response are produced by a dedicated detector simulation. In all of these steps, which are briefly introduced in the following, several Monte Carlo techniques are employed to generate the different ingredients of a simulated event.

5.2.1 Hard process

The core of the event generation is given by the simulation of the hard process, which corresponds to the subprocess with the largest momentum transfer. The hard process describes the interaction between partons, which are the constituents of the colliding protons. The cross section of the hard process can be derived based on the factorization theorem, as given by Eq. (2.16), using the parton distribution functions (PDFs) for the incoming partons as well as a parton-level cross section $\hat{\sigma}$ of the process of interest. The parton-level cross section for producing a final state n from the incoming partons a and b is given by an integral over the phase space of the final state:

$$\hat{\sigma}_{ab \rightarrow n}(\hat{s}, \mu_F, \mu_R) \sim \int \frac{1}{2\hat{s}} |\mathcal{M}_{ab \rightarrow n}|^2(\Phi_n, \mu_F, \mu_R) d\Phi_n, \quad (5.1)$$

where Φ_n denotes the final-state phase space, $\hat{s}$ corresponds to the partonic center-of-mass energy and $\mathcal{M}$ is the matrix element (ME) of the hard process, also called amplitude, which depends on the factorization scale μ_F and the renormalization scale μ_R . Since the integral can get rather involved due to the complexity of the final state, it is usually solved based on Monte Carlo sampling methods. The matrix element itself can be described by perturbation theory in the strong coupling constant α_s . The accuracy of the matrix element is given by the order in α_s . In this analysis, simulated samples based on matrix elements, where the leading

²For the 2018 data-taking period, the DOUBLEEG PD and the SINGLEELECTRON PD are replaced by the EGAMMA PD containing electron- and photon-related triggers.

order (LO) and next-to-leading order (NLO) in α_s are considered, are used. While in general, the factorization theorem holds to all orders in perturbation theory, simulating the hard process up to NLO becomes much more involved, since additional virtual- and real-emission contributions have to be taken into account. The additional contributions introduce ultraviolet and infrared divergences, which have to be absorbed, and increase the complexity of the combination between the hard process and the parton shower due to the additional emissions considered in the matrix elements. While simulations based on NLO matrix elements have been made available for some of the main SM processes, such as Drell-Yan or $t\bar{t}$ production, more complex SM processes are still described based on LO ME simulation. To improve the accuracy of the LO simulation, additional k -factors, which correspond to the ratio between the NLO and LO total cross section of a given process, are applied to the overall normalization of the simulated samples.

The hard process in the simulated samples used in this analysis is provided by different ME generators. For the majority of the simulated samples the POWHEGV2 [124–129] generator, which supplies NLO matrix element computations, and the MADGRAPH5_AMC@NLO [130–132] generator, where the matrix element computation is performed either at LO or NLO, are used. Both generators are not self-contained and have to be interfaced with a simulation of the parton shower, hadronization and underlying event. In a few cases, the hard process is also supplied by the PYTHIA8 [133] generator, which is self-contained and therefore contains said additional steps. The parton distribution functions are provided by the NNPDF3.1 set [134] computed at next-to-next-to-leading-order (NNLO) accuracy, which, based on the factorization theorem (cf. Eq. (2.16)), can be also applied to LO and NLO ME simulation.

5.2.2 Parton shower, hadronization and underlying event

The simulation of the hard process includes only effects up to a fixed order in α_s and is therefore not sufficient to describe the internal structure of jets or additional gluon emission. These higher-order effects are simulated by dedicated parton shower algorithms. The algorithms are formulated in an evolution of the momentum transfer from a scale Q associated with the hard process down to a cutoff scale Q_0 , where perturbation theory breaks down and hadronization effects of the partons have to be taken into account. Within this range, successive gluon emissions based on the corresponding probability distributions starting from the incoming and outgoing partons, also known as initial-state radiation (ISR) and final-state radiation (FSR), are generated iteratively. The ordering of the emissions differs between different algorithms, which usually employ an ordering based on the momentum transfer or the opening angle of the emission. For the simulated samples used in this analysis, the parton shower is generated by the PYTHIA8 [133] generator. Here, an ordering based on the momentum transfer is applied, but the gluon emissions are generated with respect to color dipoles, which are given by parton pairs, instead of individual partons.

The parton shower algorithms are based on approximations, which only cover the soft and collinear emissions of gluons. Thus, hard wide-angle emissions, which are usually of interest, have to be still included in the matrix element of the hard process. Combining both parts to get an accurate description of different partonic states is a complex procedure, especially for NLO matrix element generators, and problems like potential double counting or undercounting in specific phase spaces have to be taken into account. For this task, different matching and merging methods have been developed including the MLM [132] and FxFx [131] algorithms, which are applied for the LO and NLO ME simulation, respectively, using the MADGRAPH5_AMC@NLO generator, or the POWHEG method [124], which is used for the POWHEGV2 generator.

The transition from the partonic state produced by the parton shower to the hadronic final state consisting of color-neutral hadrons is simulated in the hadronization step. Unlike the previous steps, the hadronization can not be derived from first principles but has to be described by QCD-inspired models. There are two different main classes of models describing the hadronization [123], known as string models and cluster models.

The hadronization step for the samples used in this analysis is simulated by a string model. In string models, the hadronization is based on linear color confinement (cf. Sec. 2.1.3), which occurs at large distances between color-charged particles. The force field between a quark-antiquark pair is described by a color flux tube also called string, which corresponds to a linearly rising potential $V(r) = \kappa r$, with $\kappa \approx 1 \text{ GeV/fm}$. For growing distance between the quark and antiquark ($q\bar{q}$), the potential energy stored in the string increases up to a point where the energy is large enough to produce a new quark-antiquark pair ($q'\bar{q}'$), which breaks the initial string but introduces two new strings within the two color-neutral systems (qq' and $\bar{q}\bar{q}'$). If the invariant mass of the systems is large enough, this process can repeat multiple times up to a point where the energy is not sufficient to produce additional quark-antiquark pairs. At this point, each color-neutral system corresponds to a hadron. The production of baryons is taken into account by creating two diquarks in a string break instead of a quark-antiquark pair. The hadronization step for the samples used in this analysis is implemented in the PYTHIA8 [133] generator based on the Lund string model [135].

While all steps discussed so far are based on particles produced in the hard process, additional contributions which are not connected to the hard process have to be covered by the MC simulation. Some of these contributions are summarized within the underlying event (UE), which consists of two main sources. The first source, called multiple-parton interactions (MPI), describes additional soft parton-parton interactions which occur within the same proton-proton collision. The second source covers the contribution originating from the beam remnants, which do not interact in the hard process. For the samples used in this analysis, the UE contributions as well as their impact on the color flow, known as color reconnection (CR), are modeled by the PYTHIA8 [133] generator. Similar to the modeling of the hadronization, the contributions from the UE and the impact of CR can not be derived from first principles but are described by phenomenological models, whose set of adjustable parameters are optimized from comparisons to data. The parameter set used in this analysis corresponds to the CP5 tune [136] taking different CR models [137] into account. Other contributions not connected to the hard process correspond to pileup interactions. Such contributions are simulated separately based on soft QCD processes and overlaid with the simulation of the main pp collision.

5.2.3 Detector simulation

The simulation steps discussed in the two previous sections produce events at a stage before the interaction with the particle detector, also known as generator level or MC truth, and are therefore independent of the experiment used for the recording of the collisions. The detector-specific part of the simulation is covered by the detector simulation, which for many high-energy experiments, including the CMS experiment, is supplied by the GEANT4 toolkit [138]. The CMS detector simulation based on this toolkit provides an accurate and comprehensive description of the entire set of complex subdetectors and the magnetic field, discussed in Sec. 3.2.

The simulation is based on sets of physics models, which cover different processes across a wide energy range from a few hundred eV up to TeV energies. These models describe the detector geometry and detector material as well as the interaction between particles and the magnetic field or the detector material, which can lead to the production of secondary particles. In addition, the detector response is generated by simulating hits and energy losses caused by the

particles traversing the detector followed by a simulation of the electrical readout, also known as digitization, and a simulation of the trigger response. Based on these last steps and the interface of the detector simulation with the CMS software framework, simulated events, which contain the same information as the events recorded in real pp collisions, can be generated. This enables the comparison between simulated events and real events, whereas the simulated events also include information regarding the generator level.

The full detector simulation is the most time-consuming step of the simulation chain and takes between a few tens of seconds up to about 100 seconds depending on the type of simulated event and computing hardware [139]. A faster detector simulation is provided by the CMS Fast Simulation [140], which can speed up the detector simulation step by a factor of about 100. However, since for this analysis, an accurate detector simulation is crucial (cf. Sec. 8), only the full simulation based on the GEANT4 toolkit is used.

5.2.4 Simulated samples

In this analysis, different simulated samples describing the SM $t\bar{t}$ process, other SM background contributions as well as a BSM process are employed at different steps of the analysis. For most of these steps, a reasonable agreement between the simulated samples and the recorded collision data is crucial. This calls for additional modifications of the simulated samples, apart from the generation of individual samples per data-taking period to match the corresponding data-taking conditions. In this context, the number of additional pp collisions (pileup) considered in the simulation is reweighted to match the pileup distributions observed in the recorded data, shown in Fig. 3.2, based on a total inelastic pp cross section of 69.2 mb [141]. Here, contributions originating from the same bunch crossing as well as contributions from the bunch crossing before and after a triggered event are taken into account. To derive a proper estimation of events for a given process with cross section σ in a given data set, the simulated events have to be weighted by

$$w = \frac{\sigma \cdot L_{\text{int}}}{N_{\text{gen}}}, \quad (5.2)$$

where N_{gen} denotes the total number of generated events in a sample and L_{int} corresponds to the integrated luminosity of the recorded data set. As already pointed out in Sec. 5.2.1, the cross sections used for this reweighting do not have to match the accuracy of the matrix element used in the simulation process but can be based on higher-order predictions to improve the precision of the normalization of the simulated sample. Additional weights taking into account the trigger efficiencies as well as efficiencies regarding the identification of different objects, like leptons and jets, are applied to the simulated samples and are further discussed in Sec. 5.3 and 6.2.

In the following, the details of the simulation and the normalization of the different samples used in this analysis are summarized. The full list of simulated samples including the cross section values used for their normalization can be found in Appx. 3.

$t\bar{t}(\text{II})$: Simulations of the main process of interest given by $t\bar{t}$ events where both W bosons decay leptonically are used in several parts of the analysis, including, e.g., the correction of detector effects and the interpretation of the final results. The nominal sample used for this process only contains events of the dileptonic decay channel (cf. Sec. 2.3.1) and is produced using the POWHEGV2 generator based on NLO ME accuracy. Events, where at least one of the W bosons decays to a τ lepton are removed from the samples and included in the $t\bar{t}$ other contribution discussed below. An alternative simulation, where the hard process is generated by

the MADGRAPH5_AMC@NLO generator at NLO accuracy, is used as additional independent simulation. In the alternative simulation, all $t\bar{t}$ decay modes are included but only events of the dileptonic channel are selected based on generator-level information, which reduces the available statistical precision of the alternative simulation compared to the nominal dileptonic $t\bar{t}$ sample. Both samples are normalized using cross sections derived at NNLO accuracy including the next-to-next-to-leading-logarithm (NNLL) [142–146] by the TOP++ [69] program assuming a top-quark mass of 172.5 GeV. For the nominal simulation, this cross section is multiplied by the corresponding leptonic branching fractions [37]. In addition, k -factors as a function of the top-quark p_T based on ratios between NNLO QCD + NLO EWK calculations [147] and the nominal signal simulation described above, are applied to all $t\bar{t}$ samples [148]. To derive systematic uncertainties connected to the simulation of this process and $t\bar{t}$ processes with other final states, additional simulated samples are used, as further discussed in Sec. 9.2.

$t\bar{t}$ other: This SM background contribution summarizes all $t\bar{t}$ processes, which do not match the signal process of interest introduced above, namely the all-hadronic and semileptonic decay channel as well as $t\bar{t}$ events where at least one of the W bosons decays to a τ lepton. For all processes, the ME is generated at NLO using the POWHEGV2 generator. Similar to the signal process, the samples are normalized using cross sections derived at NNLO+NNLL accuracy by the TOP++ [69] program and applying the corresponding branching ratios [37].

Single top: Contributions from single top events are split into three different production channels. The largest contributions arise from single top production in association with a W boson (tW), where the samples are generated at NLO accuracy using the POWHEGV2 generator without taking fully hadronic decays of both W bosons into account. The same generator is used for simulating the t -channel production, while the production via the s -channel is generated by the MADGRAPH5_AMC@NLO generator at NLO accuracy. For the tW production and the t -channel production, separate samples for top and anti-top production are used. The cross section applied for the normalization of the tW sample is derived at approximate NNLO [149, 150] taking the corresponding branching fractions of the W boson into account. For the samples covering the t -channel and s -channel, cross sections at NLO accuracy derived with the HATHORV2.1 [151, 152] program are applied. Due to the NLO accuracy used in the simulation of the ME for the tW process, potential overlap between the $t\bar{t}$ process and the tW process has to be taken into account, which is further discussed in Sec. 5.2.5. Unless otherwise stated, the diagram removal scheme is used to describe the tW process.

DY+jets: Contributions from Drell-Yan (DY) production are simulated in two separate invariant mass regions of the two leptons, where the lower region includes the range between 10 and 50 GeV and the higher region covers the mass range above 50 GeV. For both regions, the MADGRAPH5_AMC@NLO generator at NLO accuracy is used. In both samples, the process is simulated in association with up to two additional partons at the ME level and only leptonic decays of the Z/γ^* are considered. The normalization is based on NNLO QCD + NLO EWK cross sections derived with FEWZ3.1 [153, 154] and NLO cross sections derived by the MADGRAPH5_AMC@NLO generator for the high mass and the low mass region, respectively.

Diboson: This contribution combines the simultaneous production of two vector bosons given by the WW, WZ and ZZ processes. For each of the processes, an individual sample is generated based on the PYTHIA8 generator at LO accuracy. The samples are normalized using cross sections derived at NNLO accuracy [155] for the WW process and at NLO accuracy with the MCFM6.6 [156, 157] program for the remaining two processes.

$t\bar{t}W/Z$: Contributions from $t\bar{t}$ production in association with a vector boson (W, Z) are simulated at NLO accuracy using the MADGRAPH5_AMC@NLO generator. For both contributions,

individual samples for the leptonic and hadronic decay of the vector boson are generated. The cross sections used for the normalization of the samples are based on NLO accuracy and are provided by the MADGRAPH5_AMC@NLO generator.

W+jets: Contributions from W boson production in association with jets are simulated with up to four additional partons at LO accuracy using the MADGRAPH5_AMC@NLO generator. Here, only the leptonic decay of the boson is taken into account. The sample is normalized based on a cross section at NNLO QCD + NLO EWK accuracy provided by the FEWZ3.1 program.

BSM: One simulated sample of the stop-pair production, shown in Fig. 2.9, with a stop mass of 525 GeV and a neutralino mass of 350 GeV, available with the full CMS detector simulation, is used in this analysis. The sample is generated at LO accuracy using the MADGRAPH5_AMC@NLO generator. The cross section used for the normalization is derived at approximate NNLO+NNLL accuracy using the NNLL-FAST [158, 159] program.

5.2.5 Interference between tW and $t\bar{t}$

Contributions from single top production in association with a W boson (tW) are simulated based on NLO accuracy at the ME level. Due to the additional real emission generated by the NLO ME, the tW process can produce a final state, which is identical to the final state of the $t\bar{t}$ process given by two top quarks and two W bosons. The corresponding diagrams for both processes leading to the identical final state are shown in Fig. 5.1. In the single-resonant case, the second bottom quark is produced by additional emission, while for the double-resonant case, the invariant mass of the second Wb system is close to the top-quark mass leading to an on-shell top-quark propagator and therefore to a $t\bar{t}$ process at LO.

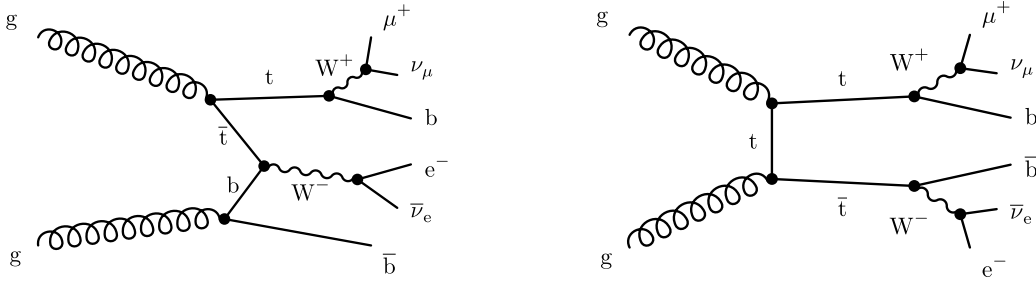


Figure 5.1: Example diagrams for the single-resonant (left) and the double-resonant (right) contribution to the $WWbb$ final state.

While in the NLO $t\bar{t}$ simulation, only the double-resonant contributions are generated, the NLO simulation of the tW process includes both types of contributions, which can lead to a potential double counting of events. In addition, potential interferences between both contributions are not correctly taken into account when using two separate samples for $t\bar{t}$ and tW . The combined amplitude squared of both contributions can be written as

$$|\mathcal{M}|^2 = |\mathcal{M}_{tW} + \mathcal{M}_{t\bar{t}}|^2 = |\mathcal{M}_{tW}|^2 + 2\text{Re}(\mathcal{M}_{tW}\mathcal{M}_{t\bar{t}}^*) + |\mathcal{M}_{t\bar{t}}|^2, \quad (5.3)$$

where $\mathcal{M}_{tW}$ and $\mathcal{M}_{t\bar{t}}$ denote the amplitudes of all single-resonant and all double-resonant contributions, respectively. The interference between both types of contributions is introduced by squaring the combined amplitude.

To avoid the problems described above, two different techniques called diagram removal (DR) and diagram subtraction (DS) [160, 161] are applied in the NLO tW simulation. In the DR approach, the double-resonant contributions given by $\mathcal{M}_{t\bar{t}}$ are completely removed from the tW calculation at NLO leading to an amplitude given by $|\mathcal{M}|_{\text{DR}} = |\mathcal{M}_{tW}|$, which is not gauge-invariant and does not include the interference between both types of contributions. In the DS approach, a local counterterm is added to suppress the double-resonant contributions when the invariant mass of the second Wb system (M_{Wb}) approaches the top-quark mass:

$$|\mathcal{M}|_{\text{DS}}^2 = |\mathcal{M}_{tW} + \mathcal{M}_{t\bar{t}}|^2 - \mathcal{C}_{t\bar{t}}(p_i) \quad \text{with} \quad \mathcal{C}_{t\bar{t}}(p_i) = f(M_{Wb}) |\mathcal{M}_{t\bar{t}}(q_i)|^2. \quad (5.4)$$

The counterterm $\mathcal{C}_{t\bar{t}}$ is a function of the momenta of the external particles p_i and is given by a product of a function f , which equals unity for $M_{Wb} = m_t$, and the double-resonant contributions with reshuffled external momenta q_i . The reshuffling of the external momenta is required to ensure that $\mathcal{M}_{t\bar{t}}(q_i)$ always describes an exactly double-resonant configuration. For f , a ratio of two Breit-Wigner functions is used. This approach is gauge-invariant and takes the interference between both types of contributions into account, but also leads to problems with the appropriate cancellation of collinear singularities [160].

The two different approaches introduced above can lead to significant differences in the simulation of the tW background process, as observed for the normalized distribution as a function of the generated p_T^{miss} (cf. Sec. 6.2.6) shown in Fig. 5.2. While both schemes agree up to a generated p_T^{miss} of around 140 GeV, the simulation based on the DR scheme yields a significantly harder spectrum compared to the DS scheme for the phase space above 140 GeV.

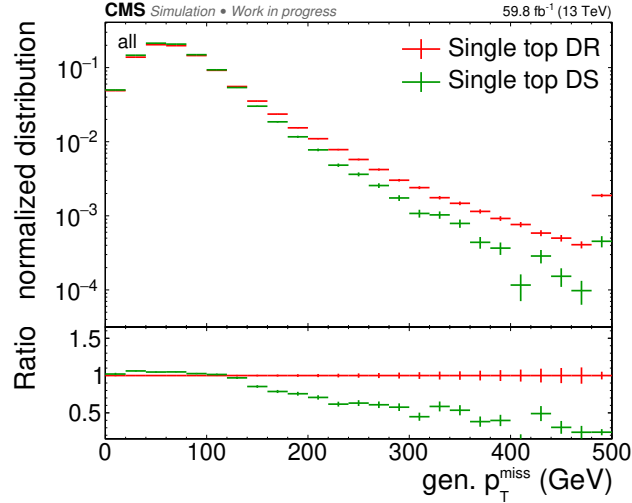


Figure 5.2: Normalized distribution as a function of the generated p_T^{miss} comparing the single top simulation based on the diagram removal (red) and the diagram subtraction scheme (green) for the 2018 data-taking period. Contributions from the t - and s -channel production agree between both results. The lower part shows the ratio between both methods. The error bars indicate the statistical uncertainty of the corresponding simulated sample.

To account for the differences observed in the generated p_T^{miss} spectrum in the final cross section measurements, two different options are considered: The first option corresponds to adding an additional uncertainty to the measurement, which takes the difference between both schemes into account. However, given that both descriptions of the tW process are not fully accurate, this

option would introduce a rather artificial uncertainty. The second option, which is pursued in the present analysis, introduces a second set of measurements, where the $t\bar{t}$ process is described based on the DS scheme rather than the DR scheme. Based on this option, two well-defined sets of measurements, which emphasize the impact of the description of the $t\bar{t}$ process on the measurements, can be provided and studied.

A fully accurate description of both contributions can be achieved by the BB4L generator [162] which combines both contributions at ME level including the correct interference as well as off-shell effects and top-quark decays at NLO. Adequate simulated samples generated by this generator were not available within the scope of this thesis. However, for future measurements of the $t\bar{t}$ cross section, the use of the BB4L generator should be considered.

5.3 Trigger efficiencies

The collision data analyzed in this thesis are collected by a variety of different single-lepton and dilepton triggers, as discussed in Sec. 5.1. The same set of triggers is also available in the MC event samples, where the trigger response is simulated as part of the detector simulation. Thus, identical trigger selections can be applied to the collision data and the simulated event samples. However, potential differences in the trigger efficiencies between data and simulation have to be taken into account and corrected. Therefore, trigger efficiencies are separately measured for data and simulation. Based on the measured trigger efficiencies, scale factors correcting the efficiencies in the trigger simulation are derived taking different sources of systematic uncertainties into account.

5.3.1 Efficiency measurement

The trigger efficiency measurement is performed based on events selected by a set of p_T^{miss} -baseline triggers (cf. Appx. 1), whose efficiency should be independent of the single-lepton and dilepton triggers of interest. Potential correlations between both sets of triggers are taken into account by applying an additional systematic uncertainty as discussed in the next section. The efficiency measurement is performed separately for simulation, using the nominal $t\bar{t}$ signal sample, and for data, where events from the MET PD are analyzed. To avoid any bias by additional selection steps, the trigger efficiency is measured in a phase space that matches the full event selection introduced in Sec. 6.3. In addition, the measurement is split between different data-taking periods and decay channels to account for the different trigger selections used. The trigger efficiency for a set of events is based on

$$\varepsilon = \frac{N_{\text{evt}}(\text{baseline} \cap \text{selection} \cap \text{signal})}{N_{\text{evt}}(\text{baseline} \cap \text{selection})}, \quad (5.5)$$

where the nominator is given by events, which pass the set of baseline and signal triggers and fulfill the event selection, while for the events in the denominator, the requirement regarding the signal triggers is not enforced. The efficiency in the same-flavor channels is measured as a function of the transverse momentum of the leading and subleading lepton, where the former corresponds to the lepton with a larger p_T . For the $e^\pm\mu^\mp$ channel, the measurement is also performed in bins of lepton p_T , but separating the leptons by flavor.

In Fig. 5.3, examples of the measurement performed based on the 2018 data-taking period are shown. For all three channels, a constant efficiency above 95% across most parts of the considered p_T -range is observed in data and simulation. Only at low- p_T bins, smaller inefficiencies caused by the momentum threshold of the lepton triggers, also known as turn-on effect, which are of

the same order in data and simulation, can be found. In general, a good agreement between the efficiencies measured in data and simulation across the considered momentum range is observed. The entire set of results as a function of lepton p_T including the measurements for the remaining data-taking periods, which show similar behavior as discussed above, can be found in Appx. 4.

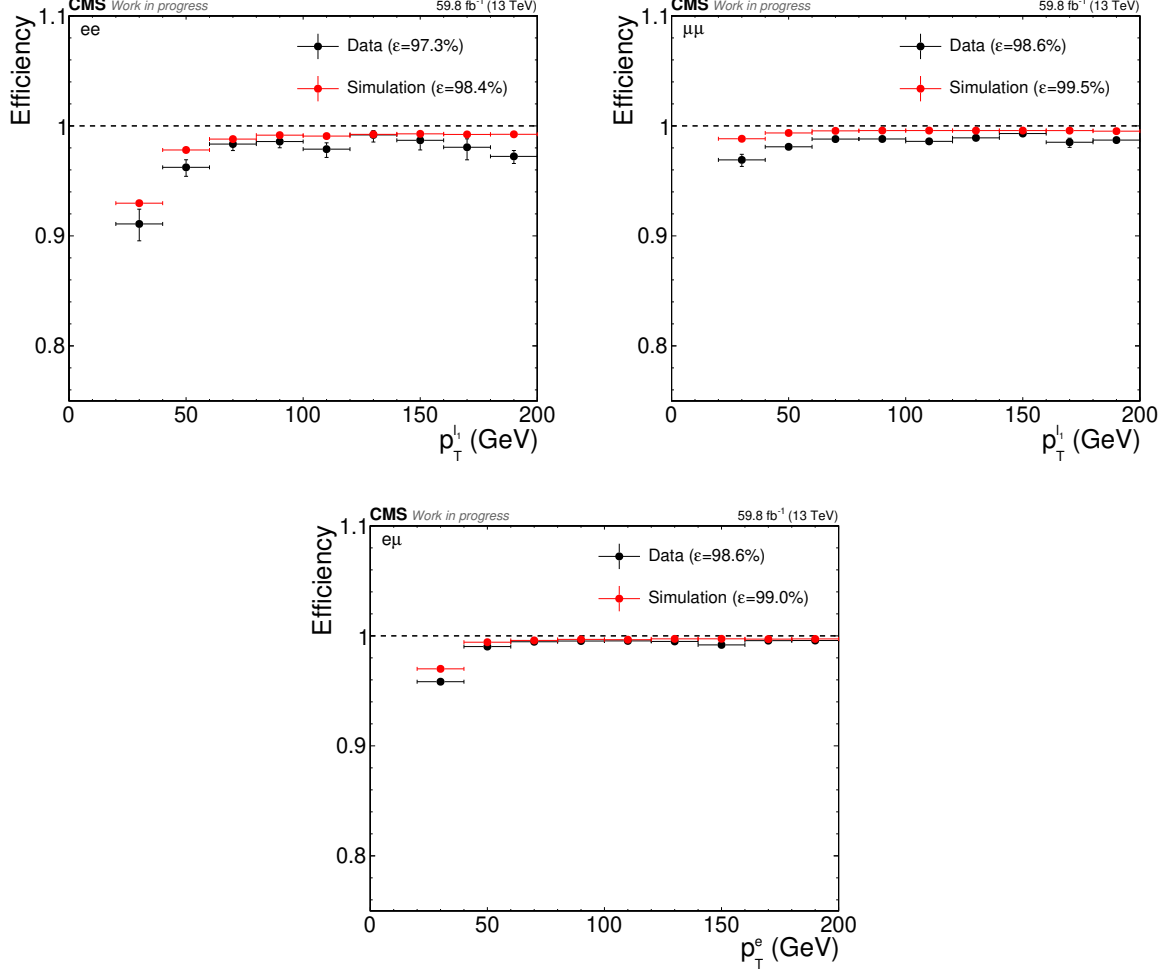


Figure 5.3: Signal trigger efficiencies for 2018 measured in data (black) and simulation (red) for the e^+e^- (top left), $\mu^+\mu^-$ (top right) and $e^\pm\mu^\mp$ (bottom) channel. The efficiencies are shown as a function of the leading lepton p_T for the same-flavor channels and as a function of the electron p_T for the $e^\pm\mu^\mp$ channel. In the legend, the overall efficiencies for data and simulation are given. The vertical lines correspond to the asymmetric statistical uncertainties derived using Clopper-Pearson intervals [163]. The last bin in each distribution corresponds to an overflow bin.

5.3.2 Scale factors

Even though the differences between the trigger efficiencies in data and simulation are small, scale factors (SFs) based on the ratios of the efficiencies are derived to correct the trigger response in the simulated samples. Similar to the measurement of the efficiencies, the scale factors are constructed as a function of lepton p_T and are split between different lepton channels and data-taking periods. The SFs derived for the 2018 data-taking period including their total uncertainties are shown in Fig. 5.4. For all three lepton channels, SFs close to unity across the

entire phase space are observed confirming the good agreement between the efficiencies measured in data and simulation. The SFs for the remaining data-taking periods can be found in Appx. 4.

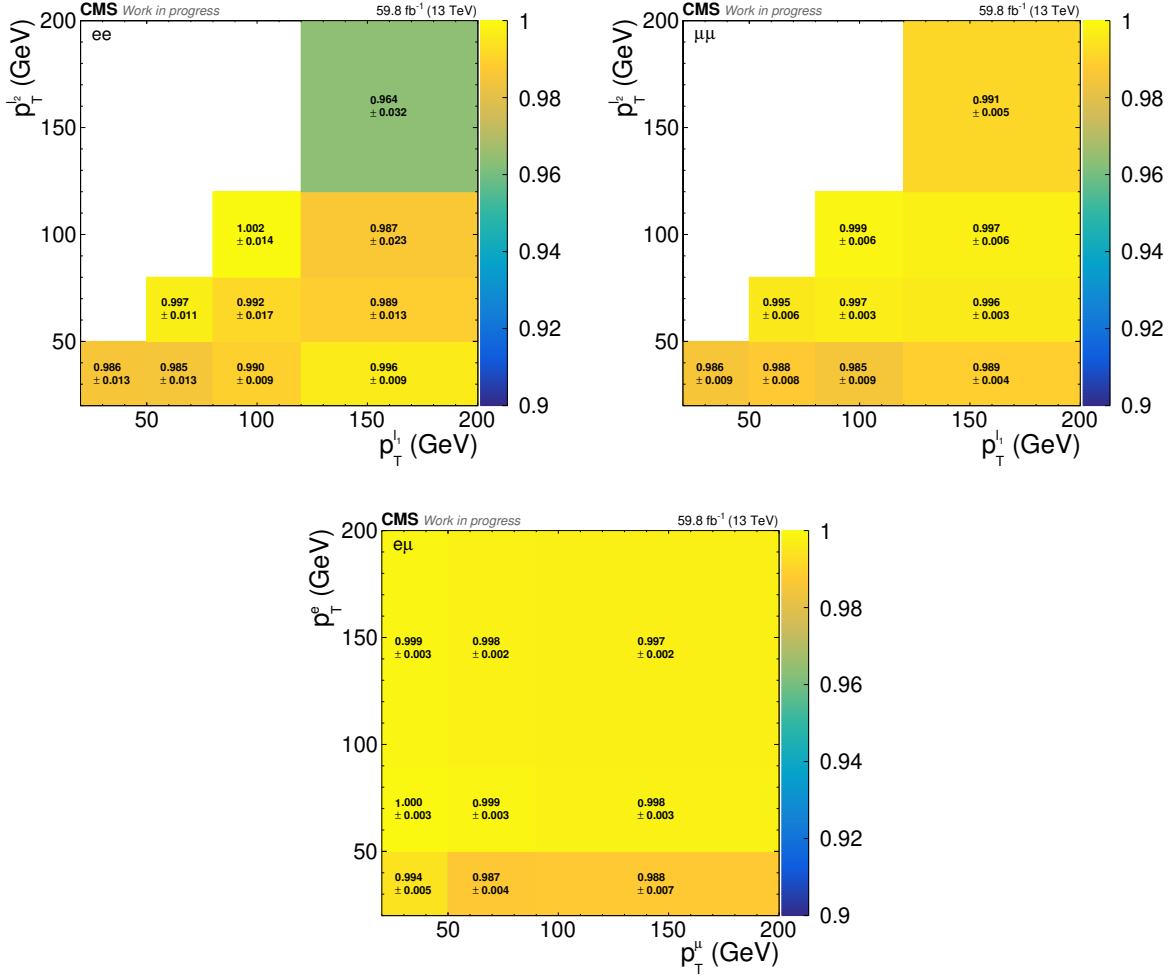


Figure 5.4: Trigger scale factors for 2018 derived for the e^+e^- (top left), $\mu^+\mu^-$ (top right) and $e^\pm\mu^\mp$ (bottom) channel. The SFs are shown as a function of leading and subleading lepton p_T or electron and muon p_T . The values of the SFs including their total uncertainties for each bin are visualized by black digits.

The uncertainties derived for the SFs are based on four different sources, which are further discussed in the following. The total uncertainty, which, similar to the individual contributions, strongly depends on the lepton channel and data-taking period, ranges between 0.2 and 4.8% across different lepton- p_T bins. The impact of the scale-factor uncertainties on the final cross section results can be found in Sec. 9.3.

Run era dependencies cover potential differences between individual run eras within the same data-taking period, which can originate from different trigger combinations used for the eras (cf. Appx. 1) and from changes in the data-taking conditions within the same period. The corresponding uncertainty is derived by measuring the SFs individually for each run era and comparing the luminosity-weighted combination of the individual SFs with the nominal result per bin. This uncertainty varies between 0.1 and 0.6%.

Phase space dependencies are derived to evaluate differences in the trigger efficiencies caused by other observables than the lepton p_T . Here, the SFs are derived individually for six different

phase space regions given by $N_{\text{jets}} < 3$ vs. $N_{\text{jets}} \geq 3$, $p_{\text{T}}^{\text{miss}} < 150 \text{ GeV}$ vs. $p_{\text{T}}^{\text{miss}} \geq 150 \text{ GeV}$ and $N_{\text{PV}} < 30(25)$ vs. $N_{\text{PV}} \geq 30(25)$, where the relaxed cut on the number of primary vertices N_{PV} is used for both 2016 data-taking periods, which contain less pileup. The final contribution from this uncertainty is derived per bin based on the largest difference between the SFs derived for the different phase spaces and the nominal SFs and varies between 0.1 and 4.0%.

Trigger correlations between the lepton signal triggers and the $p_{\text{T}}^{\text{miss}}$ -baseline triggers used for the efficiency measurement could affect the SFs. Their impact is estimated by deriving the correlation between both trigger sets in simulation based on

$$\alpha = \frac{\varepsilon_{\text{lep}} \cdot \varepsilon_{p_{\text{T}}^{\text{miss}}}}{\varepsilon_{\text{lep}+p_{\text{T}}^{\text{miss}}}} , \quad (5.6)$$

where ε_{lep} and $\varepsilon_{p_{\text{T}}^{\text{miss}}}$ denote the efficiencies of the signal and baseline triggers, respectively and $\varepsilon_{\text{lep}+p_{\text{T}}^{\text{miss}}}$ corresponds to the efficiency of requiring both trigger sets. These efficiencies are derived in simulation without the use of another independent set of baseline triggers. The bin-wise uncertainty due to the trigger correlation is given by

$$\sigma_{\text{corr}} = (1 - \alpha) \cdot S_{\text{nom}} , \quad (5.7)$$

where S_{nom} denotes the nominal scale factor. While for a few bins, this contribution is negligible, for the remaining bins it can range up to 0.4%, with the largest impact observed in the e^+e^- channel.

Statistical uncertainties include the impact of the limited statistics available in the data sets and simulated samples used for the efficiency measurement. The uncertainty is derived based on Clopper-Pearson intervals [163] and varies between 0.1 and 3.0%, where the largest contributions originate from the available statistics in the MET PD.

Chapter 6

Event reconstruction and selection

6.1	Particle flow algorithm	65
6.2	Physics object reconstruction	67
6.2.1	Primary vertex	68
6.2.2	Electrons	68
6.2.3	Muons	70
6.2.4	Jets	71
6.2.5	Heavy-flavor jets	72
6.2.6	Missing transverse momentum	73
6.3	Event selection	77
6.3.1	Total yields	78
6.3.2	Control distributions	79

The various particles produced in the pp collisions recorded by the CMS detector introduce different electrical signals in different parts of the detector. These signals are combined to reconstruct the initial physics objects produced in the collisions. This task is performed by the particle-flow (PF) algorithm [164], which is further introduced in Sec. 6.1. The objects reconstructed by the PF algorithm have to fulfill different sets of identification criteria, which depend on the particle type, to increase their selection efficiency and reduce contributions from misidentified objects. This step is discussed in Sec. 6.2. Based on the identified objects and further global observables, an event selection aiming to collect dileptonic $t\bar{t}$ events is defined in Sec. 6.3, followed by first comparisons between the recorded data and event simulations.

6.1 Particle flow algorithm

The complex design of the CMS detector, which includes a large magnetic field, a highly-granular tracker, a highly-segmented ECAL and a hermetic HCAL as well as a precise muon tracking system, allows for a particle-flow reconstruction algorithm to be applied for the first time at a hadron collider [164]. Unlike previous reconstruction algorithms used at hadron colliders, the PF algorithms take advantage of the correlations between signals in different subdetectors leading to an improved event description. In a first step, PF elements consisting of tracks reconstructed in the inner tracker, clusters of calorimeter entries and signatures from the muon system are derived. In a second step, the PF elements from different subdetectors are linked using a set of requirements, creating PF blocks. In the final step, the physics objects, also called

PF candidates, which include muons, electrons/photons as well as charged and neutral hadrons, are reconstructed within each individual PF block. This procedure is sketched in Fig. 6.1. In the following, the three different steps are briefly introduced, while the specificities of the electron and muon track reconstructions are discussed in Sec. 6.2.2 and 6.2.3, respectively.

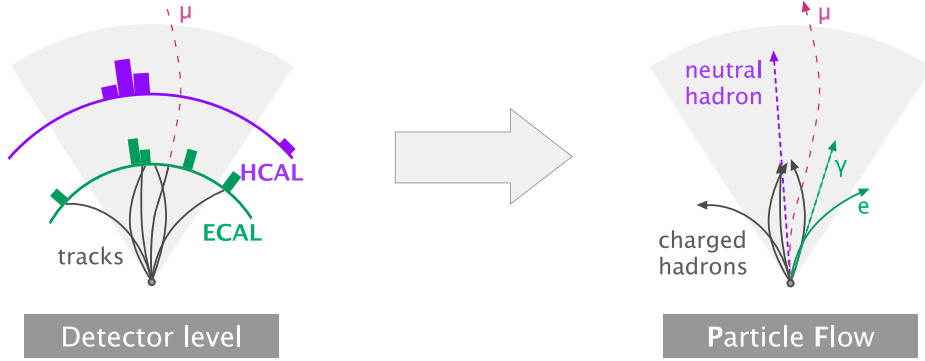


Figure 6.1: Sketch of the procedure applied by the particle flow algorithm [165]. On the left, the individual entries in the tracker (black), ECAL (green), HCAL (violet) and muon system (red) are visualized. The physics objects reconstructed by the PF algorithm based on the detector entries are shown on the right and include charged and neutral hadrons, photons, electrons and muons.

Trajectories

The hits produced by charged particles traversing the pixel and strip detector are used to reconstruct the corresponding trajectories within the tracker system. For the reconstruction, a combinatorial track finder based on Kalman Filtering [166] is applied. The track finder consists of three steps starting with generating track seeds from a few neighboring hits, which are compatible with a charged-particle track hypothesis. Based on the seeds, pattern recognition techniques are used to gather additional matching hits from all tracker layers. In the last step, a final trajectory fit is performed combining all gathered hits to derive the origin, transverse momentum and the direction of the trajectory.

To increase the overall tracking efficiency, the track finder is applied in an iterative way starting with stringent quality criteria enforced in the three steps of the track finding. For the subsequent iterations, the hits associated with tracks by earlier iterations are masked and reduced quality criteria in the track finding are applied. With this iterative procedure, a significantly higher purity in the track reconstruction can be achieved compared to a procedure with one iteration and relaxed quality criteria, which would lead to large contributions from random hit combinations.

Calorimeter clusters

The entries in the calorimeters are clustered to determine the energy and direction of neutral particles (photons and neutral hadrons) as well as to reconstruct the electron energy taking potential photons produced via bremsstrahlung into account and to improve the energy measurement of charged hadrons. For this, a dedicated clustering algorithm is applied separately for ECAL and HCAL clusters. In a first step, cluster seeds are generated based on calorimeter cells with an energy above a given threshold and above the energy recorded by neighboring cells. Here, neighboring cells correspond to four cells, which share one side with the seed candidate, and eight cells, which share at least one corner with the cell candidate, in the ECAL and

HCAL, respectively. Starting from the cluster seeds, topological clusters are formed by adding cells, which have an energy above twice the noise level and share at least one corner with a cell already in the cluster. Finally, individual clusters within a topological cluster are reconstructed by an expectation-maximization algorithm and used as PF elements. Within this algorithm, a model postulating N Gaussian energy deposits in a topological cluster with N cluster seeds is employed. The final energy and position of an individual cluster are then given by the parameters of the corresponding Gaussian function.

Linking and identification

Particles traversing the detector should usually be the origin of multiple PF elements, e.g., charged hadrons producing tracks and HCAL clusters, which motivates a linking between individual PF elements. For this task, a linking algorithm is employed which checks potential links between neighboring elements with respect to the $\eta\phi$ -plane and yields a link distance between linked elements taking the quality of the link into account. Links between trajectories and calorimeter clusters are produced based on the extrapolation of the trajectory from its last measured hit. A link is accepted if the extrapolated track lies within the cluster area and the link quality is defined by the distance between the cluster position and the extrapolated track. If a track can be linked to multiple clusters, only the link with the smallest distance is kept. Links between ECAL and HCAL clusters are generated if the ECAL cluster position lies within the cluster envelope of the HCAL cluster. The quality of the link is given by the distance between the two cluster positions. In case of multiple potential links for the same ECAL cluster, only the link with the smallest distance is kept. The last type of links is given by links between trajectories and entries in the muon system, which are further discussed in Sec. 6.2.3. As a result of the linking algorithm so-called PF blocks are produced, which contain directly- and indirectly-linked PF elements.

Based on the PF blocks, the final PF candidates are identified iteratively per individual block, while after each identification the corresponding PF elements are removed from the PF block. First, muons are identified based on links between trajectories and entries in the muon system. In a second step, isolated photons and electrons are identified based on ECAL clusters and track-ECAL-cluster links, respectively. For the electron identification, contributions from bremsstrahlung are taken into account. In the final step, neutral and charged hadrons as well as non-isolated photons are identified based on the remaining HCAL and ECAL clusters without linked tracks and track-HCAL-cluster links.

6.2 Physics object reconstruction

The PF candidates produced by the PF algorithm build the base of the event selection aiming to collect dileptonic $t\bar{t}$ events. Further selection and identification criteria are applied to the candidates to improve the object reconstruction and reduce contributions from misidentified objects. Furthermore, additional physics objects such as jets or global event observables are constructed based on the PF candidates and used in the event selection. The reconstruction and identification steps for the physics objects used in this thesis¹ are briefly introduced in the following.

¹Photons are not discussed but further details on their reconstruction and identification can be found in [97].

6.2.1 Primary vertex

For an accurate reconstruction of the pp collisions, the produced tracks and particles have to be assigned to the different simultaneous interactions per bunch crossing. To do so, primary interaction vertices are reconstructed based on the tracks of charged particles [167]. The primary vertex reconstruction is performed in three individual steps. First, tracks that are consistent with originating from the primary interaction region and being promptly produced are selected by applying a number of quality criteria such as a minimum required number of hits associated with a track. In the second step, the selected tracks are clustered according to their z -coordinates closest to the beam spot, which corresponds to the center of the luminous region averaged over several events. This task is performed by a deterministic annealing algorithm [168]. The last step is based on applying an adaptive vertex fitter [169] for vertex candidates, which have at least two associated tracks, to reconstruct the exact position of the corresponding vertex. In this analysis, only vertices with at least four associated tracks and a maximal radial and longitudinal distance of 2 cm and 24 cm from the beam spot, respectively, are considered. The vertex of main interest is the vertex whose associated tracks yield the largest sum of squared transverse momenta, also known as leading vertex (LV). The remaining vertices correspond to pileup interactions, which are not further analyzed.

6.2.2 Electrons

Electrons traversing the detector lead to tracks in the inner tracking system and associated clusters in the ECAL motivating the general electron reconstruction employed by the PF algorithm. Due to the significant thickness of the tracker system, the probability of electrons emitting photons via bremsstrahlung is large, which has to be taken into account in the reconstruction of electron tracks and clusters [164]. In case of the ECAL clustering, radiated photons are considered by combining electron clusters with potential photon clusters from bremsstrahlung creating so-called superclusters (SC). To account for the bent electron tracks in the longitudinal magnetic field, which prefer potential radiation with small η - and large ϕ -distance from the electron direction, the superclusters are created with a small longitudinal but a large transverse acceptance. To consider the impact of the radiation with respect to the electron tracks, the corresponding tracks are refitted using a Gaussian-sum filter (GSF) [170], which is better suited to handle tracks with abrupt and large energy losses compared to the nominal iterative tracking introduced above.

To improve the reconstruction and to reduce the probability of misidentified electrons, the electron candidates used in this analysis have to fulfill a set of additional criteria [97], whose parameters are briefly introduced below and listed in Tab. 6.1. Furthermore, only electrons with a transverse momentum of at least 15 GeV reconstructed up to $|\eta| < 2.4$ are considered. Due to inefficiencies in the transition region between the barrel and endcap part of the ECAL, electrons with $1.4442 < |\eta_{\text{SC}}| < 1.5660$, where $|\eta_{\text{SC}}|$ denotes the pseudorapidity of the ECAL supercluster, are not taken into account.

$|\Delta\eta_{\text{seed}}|$, $|\Delta\phi_{\text{in}}|$ measure the distance between the track extrapolated to the ECAL with respect to the cluster seed position in $\Delta\eta$ and with respect to the energy-weighted position of the supercluster in $\Delta\phi$. Both observables quantify the agreement between the track and cluster position and can reduce contributions from wrong PF links.

$|E_{\text{SC}}^{-1} - p^{-1}|$ considers the difference between the energy measured in the supercluster and the track momentum derived in the inner tracker. For nearly massless electrons, where the corresponding track and cluster are linked correctly, this difference should be small.

$\sigma_{i\eta i\eta}$ takes the width of the electromagnetic shower in the ECAL into account. It is derived based on the distribution of the crystal energies with respect to their pseudorapidity for the 5×5 cells around the most energetic crystal in a SC. Requirements on this observable can reduce the contributions from hadrons misidentified as electrons or photons, which tend to induce a broader electromagnetic shower than isolated electrons and photons.

H/E denotes the ratio of hadronic over electromagnetic energy. The nominator corresponds to the energy deposited in the HCAL in a cone of $\Delta R < 0.15$ around the SC direction. The denominator is given by the energy of the electron candidate. Upper limits on this parameter can reduce the contributions from misidentified hadrons. Here, limits depending on the energy E_{SC} of the supercluster and the median transverse energy density ρ per unit area in an event are used to account for different contributions from HCAL noise, pileup and leakage of EM showers to the HCAL.

$|d_0|$, $|d_z|$ denote the transverse and longitudinal distance of the reconstructed track to the primary vertex, respectively. Requirements on these parameters, which are known as impact parameters, can reduce contributions from electrons originating from pileup or secondary vertices.

N_{miss} , **conv. veto** correspond to criteria that reduce contributions from nonprompt electron/positron pair production. The signatures of such contributions usually contain tracks with missing hits in the first layers of the tracking system, whose number is restricted in the electron identification. In addition, a pixel conversion veto [171] aiming to identify electrons from photon conversion is employed.

r_{iso} denotes the relative combined PF isolation, which takes contributions from the transverse momenta of charged hadrons (I_{ch}), photons (I_{γ}), neutral hadrons (I_{n}) and pileup (I_{PU}) inside a cone of $\Delta R < 0.3$ around the electron direction into account. The isolation itself is defined by

$$r_{\text{iso}} = [I_{\text{ch}} + \max(0, I_{\text{n}} + I_{\gamma} - I_{\text{PU}})] / E_{\text{T}}, \quad (6.1)$$

where the maximization in the second summand is motivated by the fact that contributions from neutral components, which can not precisely be associated with a vertex, have to be corrected by pileup contributions. The transverse energy of the electron is given by E_{T} .

Table 6.1: Additional criteria employed for the cut-based electron identification [172] for electrons reconstructed in the barrel (left) and endcap region (right).

Parameter	$ \eta_{\text{SC}} < 1.4442$	$ \eta_{\text{SC}} > 1.5660$
$ \Delta\eta_{\text{seed}} $	< 0.00255	< 0.00501
$ \Delta\phi_{\text{in}} $	< 0.022	< 0.0236
$ E_{\text{SC}}^{-1} - p^{-1} $	$< 0.159 \text{ GeV}^{-1}$	$< 0.0197 \text{ GeV}^{-1}$
$\sigma_{i\eta i\eta}$	< 0.0104	< 0.0353
H/E	$< 0.026 + 1.15/E_{\text{SC}} + 0.0324\rho/E_{\text{SC}}$	$< 0.0188 + 2.06/E_{\text{SC}} + 0.183\rho/E_{\text{SC}}$
$ d_0 $	$< 0.05 \text{ cm}$	$< 0.10 \text{ cm}$
$ d_z $	$< 0.10 \text{ cm}$	$< 0.20 \text{ cm}$
N_{miss}	≤ 1	≤ 1
conv. veto	yes	yes
r_{iso}	$< 0.0287 + 0.506/p_{\text{T}}$	$< 0.0445 + 0.963/p_{\text{T}}$

Differences between the electron identification efficiencies in simulation and data are corrected by applying dedicated scale factors [97, 173] to the simulated events, which are derived as a function of the electron p_T and the pseudorapidity of the ECAL supercluster using a tag-and-probe method. To correct the energy measurement of the reconstructed electrons, scale corrections based on multivariate regressions and comparisons between data and simulation are applied to the electron energies [97]. Since the energy resolution observed in data and simulation does not agree, the electron energies in simulation are additionally smeared to match the resolution observed in data.

6.2.3 Muons

Muons are mainly identified based on links between trajectories in the tracker and entries in the muon system. However, in general, there are three different ways how muons are reconstructed [101, 164]. Standalone muons are built based on entries in the muon system only. Here, hits in the DTs and CSCs are clustered to muon track seeds. To gather hits in all three muon detector types along the muon trajectory, pattern recognition based on a Kalman filtering technique is employed. Global muons are reconstructed by matching the standalone-muon tracks to trajectories reconstructed in the inner tracking system. This matching is performed by comparing the parameters of the two tracks propagated onto a common surface. For the final global-muon track, the hits from both tracks are combined and the trajectory is refitted. Tracker muons are built by inverting the logic used for the global muons. Here, inner tracks above a given momentum threshold are extrapolated to the muon system. A tracker muon is reconstructed if the extrapolated track matches at least one segment² in the muon system. The use of tracker muons leads to an improved efficiency for muons with a transverse momentum below 10 GeV.

Since in this analysis, muons have to fulfill $p_T > 15$ GeV, only global muons are taken into account. In addition, only muons reconstructed within the acceptance of the muon system given by $|\eta| < 2.4$ are considered. Additional identification criteria applied to the reconstructed muons are briefly introduced below and summarized in Tab. 6.2. The observables regarding the impact parameters and the relative isolation match the definition introduced in the discussion of the electron reconstruction (cf. Sec. 6.2.2), except for small differences in the latter concerning the correction of pileup contributions [101].

χ^2/dof denotes the χ^2 value per degree of freedom for the fit of the global-muon track. Limits on this observable reduce the contributions from hadrons escaping the HCAL leading to entries in the muon system (hadronic punch-through) as well as muons originating from the in-flight decay of other particles.

N_{hits}^μ , N_{stat}^μ correspond to the number of hits in the muon system included in the global-muon track and the number of muon stations, in which segments matched to the inner track are found, respectively. The latter implies that the muon is also reconstructed as a tracker muon. These two observables reduce the contributions from hadronic punch-throughs and wrong track-to-segment matches.

$N_{\text{hits}}^{\text{pixel}}$, N_{lay} denote the number of hits in the pixel tracker and the number of tracker layers with a hit, respectively. For both observables, the hits are part of the inner track. While both observables reduce muons from decays in flight, the latter also guarantees a precise measurement of the transverse momentum in the tracker.

²A segment corresponds to a straight-line part of a muon track built within each DT and CSC chamber.

Table 6.2: Additional criteria employed for the muon identification [174]. No distinction between muons reconstructed in the barrel or endcap region is applied.

Parameter	χ^2/dof	N_{hits}^μ	N_{stat}^μ	$N_{\text{hits}}^{\text{pixel}}$	N_{lay}	$ d_0 $	$ d_z $	r_{iso}
Requirement	< 10	> 0	> 1	> 0	> 5	$< 0.2 \text{ cm}$	$< 0.5 \text{ cm}$	< 0.15

Differences between the muon identification efficiencies in simulation and data are corrected by applying dedicated scale factors [101, 175] to the simulated events, which are derived as a function of the muon p_T and the muon pseudorapidity using a tag-and-probe method. The momentum measurements of the reconstructed muons are corrected and the momentum resolution for muons in simulation is smeared to match the resolution observed in data. For this, dedicated calibrations based on $Z/\gamma^* \rightarrow \mu^+ \mu^-$ events are performed [176, 177].

6.2.4 Jets

Hadronic showers produced by the hadronization of quarks and gluons contain multiple individual particles. To properly reconstruct the initial quark or gluon, the individual particles given by PF candidates are clustered and reconstructed as jets. In this analysis, jets clustered with the anti- k_T algorithm [178] provided within the FASTJET [179] package are used. Similar to other sequential recombination jet algorithms, two different distances are defined in the anti- k_T algorithm:

$$d_{ij} = \min(k_{T,i}^{-2}, k_{T,j}^{-2}) \frac{\Delta y_{ij}^2 + \Delta \phi_{ij}^2}{R^2} \quad \text{and} \quad d_{iB} = k_{T,i}^{-2}, \quad (6.2)$$

where $k_{T,i}$ denotes the transverse momentum of particle i , while Δy_{ij}^2 and $\Delta \phi_{ij}^2$ correspond to the squared rapidity distance and azimuthal distance between particle i and j , respectively. The parameter R defines the radius of the jet cone and is set to 0.4 for jets used in this analysis. The clustering itself is performed iteratively employing both types of distances. In each step, the smallest distance is identified. If the smallest distance corresponds to d_{ij} , the entities i and j are clustered together and the distances are derived again. Once d_{iB} is the smallest distance, entity i is found to be a jet and removed from subsequent clustering steps.

Before clustering the jets, contributions from pileup vertices have to be removed to avoid biasing the clustering procedure. Here, two different methods known as charged-hadron subtractions (CHS) [164] and pileup per particle identification (PUPPI) [180] are used. While for the jets employed in this analysis, the CHS method is applied, the PUPPI method is introduced further in the context of the missing transverse momentum in Sec. 6.2.6. In the CHS method, charged hadrons, which are matched to vertices other than the leading vertex are removed from the list of PF candidates before the jet clustering.

In this analysis, only jets with transverse momenta above 30 GeV reconstructed within $|\eta| < 2.4$ are taken into account. In addition, jets reconstructed within $\Delta R < 0.4$ of a selected electron or muon are removed to avoid any overlap between selected jets and leptons. Additional identification criteria [180], which mainly aim to reduce jets originating from noise or failures in the reconstruction, are summarized in Tab. 6.3. For jets with $p_T < 50 \text{ GeV}$, further loose identification criteria [181] discarding contributions from jets originating from pileup vertices, which are derived based on multivariate techniques, have to be fulfilled. In the 2018 data-taking period, two HCAL modules were not available during the second half of the period affecting the jet

reconstruction in the area defined by $-1.57 < \phi < 0.87$ and $-3.0 < \eta < -1.3$. To avoid any bias introduced by this malfunction, events that contain reconstructed jets with $p_T > 15$ GeV in this area are not considered in the affected period.

Table 6.3: Additional criteria employed for the jet identification [182] taking contributions from different particle types to the jet energy and jet multiplicity into account.

Parameter	Requirement
Neutral hadron energy fraction	< 0.9
Neutral EM energy fraction	< 0.9
Number of constituents	> 1
Muon energy fraction	< 0.8
Charged hadron energy fraction	> 0
Charged multiplicity	> 0
Charged EM energy fraction	< 0.8

The energy of the jets, also denoted as jet energy scale (JES), is corrected in multiple steps compensating different effects [183, 184] to achieve an improved agreement between the energy of a reconstructed jet and the corresponding true jet. In a first step, pileup contributions from neutral particles or charged hadrons not associated with any vertex are taken into account event-by-event based on the effective jet area and the average energy density in the event. These corrections are derived using simulation of QCD dijet events with and without pileup overlay. Additional residual effects between data and simulation are taken into account based on a random cone method. In a second step, the jet response as a function of pseudorapidity and transverse momentum is corrected. Here, the corrections are derived by comparing the momenta between reconstructed jets and the corresponding true jets in simulations of QCD dijet events. In a last step, the remaining small differences in the jet energy response between data and simulation are taken into account. Residual corrections derived based on simulations of QCD dijet and multijet events as well as Z/γ^* events are applied to jets in data. To match the jet energy resolution (JER) in simulation to the reduced resolution observed in data, the resolution of jets in simulation is smeared based on a matching between the reconstructed jet and the corresponding jet clustered from generator-level particles [183, 185].

6.2.5 Heavy-flavor jets

The identification of heavy-flavor (HF) jets originating from b quarks is essential for the selection of $t\bar{t}$ events. Three main characteristics of heavy-flavor jets are usually employed in such identifications, also called heavy-flavor tagging [186], which in general hold for jets originating from b and c quarks. The HF hadrons produced in these jets have significantly longer lifetimes compared to light-flavor (LF) hadrons. The increased lifetime, which for b hadrons is of the order of 1.5 ps, can lead to a significant flight distance of the HF hadrons before their decay. This flight distance can be measured based on the reconstruction of secondary vertices, as sketched in Fig. 6.2, using the inner pixel detector. The second characteristic is based on the larger mass of b and c quarks compared to the LF quarks, which leads to a harder fragmentation. Thus, HF jets tend to consist of particles, which on average have a higher transverse momentum relative to the jet axis. Furthermore, HF hadrons are much more likely to contain electrons and muons in their decay chain. Hence, also the presence of charged leptons within a jet can be used in the tagging of HF jets.

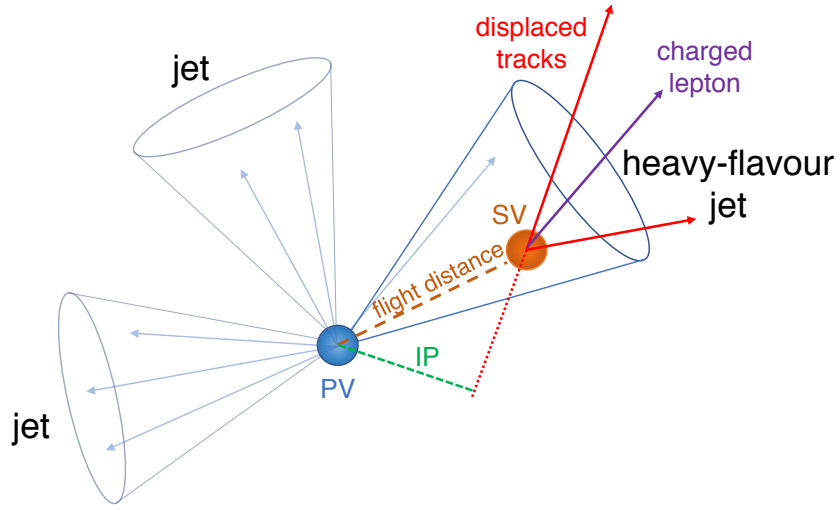


Figure 6.2: Sketch of the identification of a heavy-flavor jet [186] based on the reconstruction of a secondary vertex (SV), which is significantly displaced from the primary vertex (PV), quantified by the large impact parameter (IP). The tracks used for the reconstruction of the SV are shown in red and violet.

In this analysis, the DeepJet tagger [187] is employed to tag jets originating from b quarks. This HF tagger is based on multivariate techniques and uses a total of up to 650 low-level input features, which, among others, also encode the three characteristics discussed above. In more detail, the input features are split into global variables, which contain observables like the jet kinematics and number of constituents in a jet, and into features containing information about the charged and neutral PF candidates of the jet as well as features related to the secondary vertices. The architecture of the DeepJet tagger includes combinations of convolutional, recurrent and fully-connected layers and enables a multi-classification based on six output nodes. Here, the tagger can distinguish between jets originating from b quarks, c quarks, LF quarks as well as from gluons, which also enables the application as a c -jet tagger. The b -quark category is further divided into three sub-categories with jets containing two b hadrons, a leptonically or a hadronically decaying b hadron.

Within this analysis, the DeepJet tagger is operated at a loose working point, which yields a probability of around 10% for jets from LF quarks and gluons to be misidentified as b jets. For jets from c quarks, this probability increases to around 40 to 45%, while for the b -tagging itself, an efficiency of up to 95% is observed in simulated $t\bar{t}$ events.

Differences in the b -tagging efficiency between data and simulation are taken into account by applying additional weights to the simulated events. The weights are derived based on the probabilities for a given configuration of tagged and non-tagged jets in data and simulation [188]. These probabilities take analysis-specific tagging efficiencies, which are derived on simulation and depend on the true flavor of a jet, as well as centrally provided scale factors (SFs) per data-taking period into account.

6.2.6 Missing transverse momentum

One of the main observables employed in this analysis is the missing transverse momentum, as already discussed in Sec. 2.3.3. Before the collision, the incoming protons have a negligible transverse momentum, which due to momentum conservation should also hold for the vectorial

momentum sum of all particles produced in the same pp collision. Particles not detected by the detector, such as neutrinos or BSM particles, which only interact weakly, introduce an imbalance in this momentum sum, which can be quantified by the missing transverse momentum:

$$p_T^{\text{miss}} = \left| \vec{p}_T^{\text{miss}} \right| = \left| - \sum_i^N \vec{p}_{T_i} \right|, \quad (6.3)$$

where the sum runs over all particles reconstructed in an event. Thus, for a signature with only one undetected particle, p_T^{miss} corresponds to the transverse momentum of the undetected particle, while in general, p_T^{miss} equals the transverse momentum of the system of all undetected particles. Since every reconstructed particle contributes to the sum in Eq. (6.3), the missing transverse momentum is highly sensitive to the momentum resolution of different particle types. Hence, mismeasurements of any kind as well as detector artifacts can have a large impact on the p_T^{miss} reconstruction and have to be taken into account. Such contributions are also known as fake p_T^{miss} , while contributions originating from weakly-interacting particles escaping detection are called genuine p_T^{miss} .

In this thesis, four different reconstruction methods for the missing transverse momentum are considered. One of these methods, also called DNN p_T^{miss} , has been developed within the scope of this thesis and is discussed in detail in Sec. 7. The remaining three reconstruction methods including applied corrections are introduced in the following:

Particle flow p_T^{miss}

This reconstruction of the missing transverse momentum [189], which is denoted as PF p_T^{miss} in the following, is the most straightforward method and is entirely based on the use of candidates reconstructed by the PF algorithm. Here, the sum in Eq. (6.3) runs over all PF candidates reconstructed in a given event. In a following step, the result of this rather simple approach is corrected based on the JES corrections introduced in Sec. 6.2.4. The JES corrections are propagated by

$$\vec{p}_T^{\text{miss}} = \vec{p}_T^{\text{miss,raw}} - \sum_{\text{jets}} (\vec{p}_{T,\text{jet}}^{\text{corr}} - \vec{p}_{T,\text{jet}}), \quad (6.4)$$

where $\vec{p}_T^{\text{miss,raw}}$, $\vec{p}_{T,\text{jet}}^{\text{corr}}$ and $\vec{p}_{T,\text{jet}}$ denote the uncorrected PF p_T^{miss} , the jet p_T corrected by the JES corrections and the uncorrected jet p_T , respectively. For the sum, only jets with $p_T > 15 \text{ GeV}$ are considered and overlaps between jets and charged leptons or photons are taken into account. The impact of these corrections, also known as type-1 corrections, is shown in Fig. 6.3. By propagating the JES corrections, the p_T^{miss} response can be significantly improved.

A second set of corrections is applied to the PF p_T^{miss} in order to reduce modulations in $\phi(p_T^{\text{miss}})$, which are observed with a period of around 2π and whose amplitude scales roughly linearly with the number of observed vertices N_{vtx} . The potential cause of these modulations can be found in anisotropic detector response, detector misalignment or the displacement of the beam spot. Since in general the azimuthal distribution of p_T^{miss} should be flat, dedicated corrections as a function of N_{vtx} are derived [190] and applied on top of the type-1 corrections.

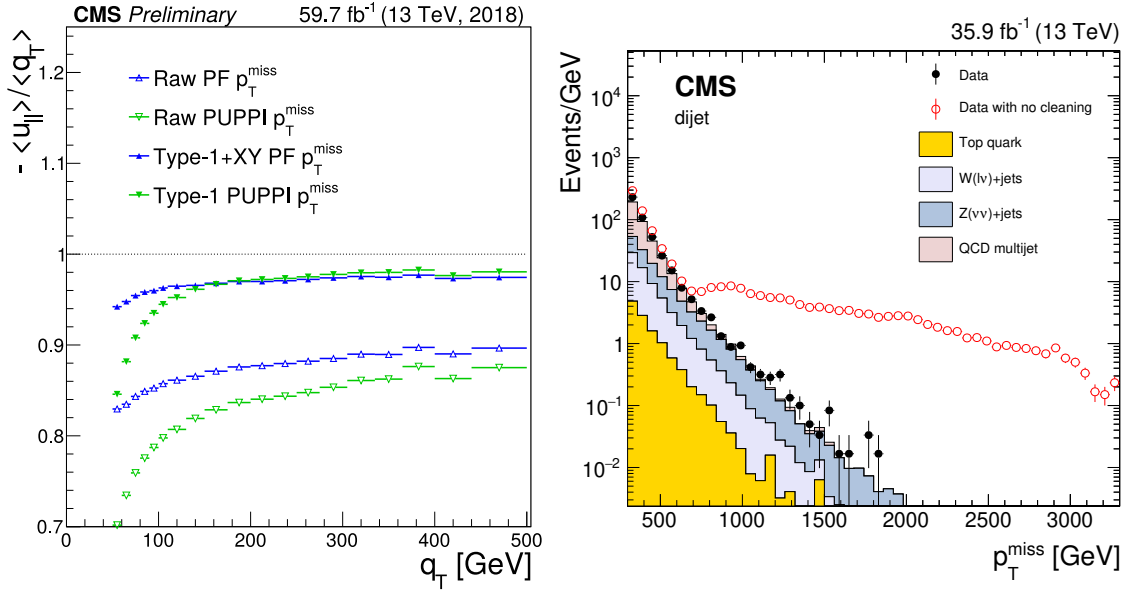


Figure 6.3: Left plot: Response of different p_T^{miss} reconstructions measured in γ +jets events as a function of the photon transverse momentum q_T [191]. The response is derived based on a method taking the hadronic recoil parallel to the photon axis into account [189]. The response for the uncorrected (empty markers) and corrected (filled markers) for the PF (blue) and PUPPI (green) p_T^{miss} reconstruction are compared. Right plot: Comparison between the PF p_T^{miss} spectrum of QCD dijet events in simulation and recorded data with (black) and without (red) anomalous p_T^{miss} filters applied [189].

PUPPI p_T^{miss}

An alternative reconstruction method of the missing transverse momentum is based on the pileup per particle identification (PUPPI) [180], which aims to mitigate the impact of pileup on observables of clustered hadrons, such as p_T^{miss} or jets. In this method, Eq. (6.3) is altered by assigning a weight w_i between 0 and 1 to each PF candidate, which rescales the corresponding momentum in the sum. This weight quantifies the probability of a particle originating from PU ($w_i = 0$) or from the LV ($w_i = 1$). For charged particles, the assignment of the weight is straightforward and based on the association of the corresponding charged particle track with the LV or a PU vertex. This weight assignment does not hold for charged hadrons without an associated vertex. Here, the closest distance to the LV along the z-axis is taken into account. For distances below (above) 0.3 cm, a weight of 1(0) is applied to the charged hadron. The main part of the PUPPI algorithm is dedicated to the weight assignment for neutral particles, where the local particle distribution around the particle of interest is taken into account. The probability of a neutral particle i originating from a PU collision is encoded in

$$\alpha_i = \log \sum_{j \neq i, \Delta R_{ij} < R_0} \left(\frac{p_{T,j}}{\Delta R_{ij}} \right)^2 \begin{cases} \text{for } |\eta_i| < 2.5, & j = \text{charged particles from LV,} \\ \text{for } |\eta_i| > 2.5, & j = \text{all reconstructed particles.} \end{cases} \quad (6.5)$$

Here, the sum runs over all particles j in a cone of $\Delta R < 0.4$ around the particle of interest i . In the acceptance region of the tracker ($|\eta| < 2.5$), only charged particles from the LV are taken into account to describe the local particle distribution around particle i , while for the remaining region all reconstructed particles are taken into account. Particles originating from a parton

shower are expected to be surrounded by other particles from the same shower leading to a large α_i . This does not hold for particles, which are diffusely radiated from PU collisions. Based on the observable α_i , the weight w_i is derived using an expected α distribution derived from PU contributions in an event. Finally, an additional condition for neutral particles as a function of N_{vtx} based on tuned parameters is applied to remove residual dependencies on the number of interactions.

Similar to the PF reconstruction, additional corrections are applied taking into account the JES corrections, where jets reconstructed based on the PUPPI algorithm are considered (cf. Fig. 6.3). In addition, corrections regarding the $\phi(p_T^{\text{miss}})$ modulation, which are significantly smaller compared to PF p_T^{miss} , are used.

Deep p_T^{miss}

While the two results of the methods introduced above (PF and PUPPI p_T^{miss}) and the DNN p_T^{miss} reconstruction, developed in the scope of this thesis, are employed in the final $t\bar{t}$ cross section measurement, the results of the Deep p_T^{miss} reconstruction [192, 193] are only used for comparisons discussed in Sec. 7.6. This reconstruction is based on a deep neural network approach deriving weights (w_i) and biases ($b_{x,i}, b_{y,i}$) for the momentum of each PF candidate i , which are combined to reconstruct the x - and y -component of p_T^{miss} separately:

$$p_x^{\text{miss}} = \sum_i (w_i \cdot p_{x,i} + b_{x,i}) \quad \text{and} \quad p_y^{\text{miss}} = \sum_i (w_i \cdot p_{y,i} + b_{y,i}) , \quad (6.6)$$

where $p_{x,i}$ and $p_{y,i}$ denote the two components of the transverse momentum of the PF candidate. A total of eleven features per PF candidate are used as inputs for the network. These include inputs regarding the four-momenta of the candidates, impact parameters, the particle type, the electric charge as well as the weights derived by the PUPPI algorithm and the association with the LV, if available. The network itself is trained on simulation of $Z/\gamma^* + \text{jets}$ and $t\bar{t}$ events. The Deep p_T^{miss} reconstruction is provided based on two different tunes, which differ in the loss function used for the training of the network. For the resolution tune, only the mean squared errors between the reconstructed and true values of p_x^{miss} and p_y^{miss} are considered. For the response tune, additional terms, which aim to balance the response at different scales of the transverse momentum, are added.

Generated p_T^{miss} and $p_T^{\nu\nu}$

The three different observables introduced above all correspond to different reconstruction methods of the missing transverse momentum and thus serve as a proxy of the true missing transverse momentum present in an event. In this thesis, two different definitions of the true p_T^{miss} in an event are employed. The generated p_T^{miss} describes the missing transverse momentum at generator level (cf. Sec. 5.2.3). Here, the contributions from all particles, which interact only weakly and escape the detector, are taken into account. In case of dileptonic $t\bar{t}$ events, this includes neutrinos originating from the two W-boson decays but also potential additional neutrinos produced, e.g., in the decays of b mesons usually present in the hadronization of b quarks.

The second definition, denoted as $p_T^{\nu\nu}$, corresponds to the transverse component of the vectorial momentum sum of two neutrinos and is therefore only defined for events with two neutrinos in the final state. For dileptonic $t\bar{t}$ events, these two neutrinos correspond to the neutrinos from the W-boson decays. Unlike for the generated p_T^{miss} , any additional non-prompt neutrinos produced in the hadronization process are not considered in the $p_T^{\nu\nu}$ observable. Hence, the two definitions

of the true p_T^{miss} only agree for events without additional non-prompt neutrinos or contributions from other undetected weakly-interacting particles.

Anomalous p_T^{miss} filters

To remove events with anomalous high p_T^{miss} caused by failures in the reconstruction or artifacts in the detector, additional event filters are applied. ECAL-related filters consider contributions from noisy sensors or spurious signals from particle interactions with the photodetectors. Similar effects are taken into account by HCAL-related filters. In addition, contributions from particles produced by proton collisions upstream of the detector, also called beam halo, as well as events including poorly reconstructed muons are removed. The impact of the cleaning based on these filters is shown in Fig. 6.3. The agreement between data and simulation as a function of the missing transverse momentum for dijet events is significantly improved. This confirms the high efficiency of the filters of around 85-90% combined with a small misidentification probability of less than 0.1% [189].

6.3 Event selection

The event selection introduced in the following aims to collect a set of events containing a large number of dileptonic $t\bar{t}$ events while keeping the contributions from other SM processes as small as possible. The selection is mainly guided by matching the final state of the dileptonic $t\bar{t}$ process, as introduced in Sec. 2.3.1, which includes two charged leptons of opposite electrical charge, two neutrinos as well as two b quarks.

The different steps in the event selection are summarized in Tab. 6.4, distinguishing between the same-flavor (SF) channels $e^+e^-/\mu^+\mu^-$ and the different-flavor (DF) channel $e^\pm\mu^\mp$. The requirements regarding the transverse momenta of the two selected leptons are driven by the lepton triggers used to collect the collision data (cf. Sec. 5.1). Since for dileptonic $t\bar{t}$ events, no additional charged leptons are expected, events with a third charged lepton with $p_T > 15$ GeV are vetoed. The lower limits on the invariant mass of the two leptons are motivated by reducing contributions from low-mass DY processes and heavy-flavor resonances. In the SF channels, resonant contributions from the DY process are removed by vetoing events with an invariant mass around the Z boson mass. For all channels, at least two reconstructed jets are required to match

Table 6.4: Summary of the different detector level event selection criteria for the same-flavor channels ($e^+e^-/\mu^+\mu^-$) and the different-flavor channel ($e^\pm\mu^\mp$). The two selected leptons have to be oppositely charged.

Parameter	$e^+e^-/\mu^+\mu^-$	$e^\pm\mu^\mp$
Number of leptons (N_ℓ)	= 2	= 2
Leading lepton p_T	> 25 GeV	> 25 GeV
Subleading lepton p_T	> 20 GeV	> 20 GeV
Third lepton veto	yes	yes
Invariant mass ($m_{\ell\ell}$)	> 20 GeV $\notin [76,106]$ GeV	> 20 GeV
Number of jets (N_{jets})	≥ 2	≥ 2
Number of b jets (N_b)	≥ 1	≥ 1
Missing transverse momentum (p_T^{miss})	> 40 GeV	–

the $t\bar{t}$ final state while also considering $t\bar{t}$ events with additional jets from FSR/ISR. Out of these jets, at least one jet has to be b-tagged by the DeepJet tagger. No second b-tagged jet is required due to the limited efficiency of the tagger and the presence of other SM contributions, which can not be further reduced by higher requirements on the b-jet multiplicity. The final selection step, which is only applied to the SF channels, contains a lower limit on the missing transverse momentum to reduce contributions from DY events, where no genuine p_T^{miss} is expected. This last step is performed based on the DNN p_T^{miss} introduced in the next chapter.

6.3.1 Total yields

The total number of expected events derived from simulation as well as the number of events observed in collision data after the full event selection are shown in Tab. 6.5. The event selection yields a rather pure set of dileptonic $t\bar{t}$ events, which sum up to a contribution of 77.5% across all channels. The DF channel yields a $t\bar{t}$ signal fraction of almost 80%, while for the SF channels, this fraction is slightly reduced due to larger contributions from the DY+jets background process, which reaches up to more than 8% for the SF channels. The largest background contributions in all channels correspond to events from the $t\bar{t}$ other processes (cf. Sec. 5.2.4). Background contributions originating from single top production make up for around 4.5% of the events in each of the three channels, while the remaining SM background processes account for only less than one percent of the total event count. Comparisons of the total number of events between the three channels show the expected results. The largest number of signal events is observed in the DF channel due to combinatorics, which in theory yield a factor of two with respect to each of the SF channels, and due to the relaxed requirements used in the DF channel (cf. Tab. 6.4).

Table 6.5: Total number of events per channel after the full event selection combining all four data-taking periods. The upper part of the table shows the individual contributions to the total event prediction derived from simulation (Sum MC). The ratio of the individual contributions and the total prediction in percent is shown in round brackets. The lower part of the table contains the uncertainty on the total prediction, which includes systematic and statistical contributions, as well as the number of events in data and the ratio between data and simulation (Data/MC).

Channel		e^+e^-	$\mu^+\mu^-$	$e^\pm\mu^\mp$	all
Contribution	$t\bar{t}$ (ll)	115513.6 (74.7)	221061.4 (73.7)	530937.9 (79.9)	867513.0 (77.5)
	$t\bar{t}$ other	18156.9 (11.7)	38517.6 (12.8)	90397.1 (13.6)	147071.7 (13.1)
	Single top	7118.7 (4.6)	13370.8 (4.5)	31522.6 (4.7)	52012.1 (4.6)
	DY+jets	12602.9 (8.1)	24752.4 (8.3)	6867.2 (1.0)	44222.4 (4.0)
	Diboson	700.0 (0.5)	1235.8 (0.4)	2429.0 (0.4)	4364.9 (0.4)
	$t\bar{t}W/Z$	476.7 (0.3)	862.6 (0.3)	1766.2 (0.3)	3105.5 (0.3)
	W+jets	75.4 (0.0)	75.1 (0.0)	881.6 (0.1)	1032.1 (0.1)
Sum MC		154644.2	299875.7	664801.8	1119321.6
$\sigma(\text{Sum MC})$		9702.6	13844.8	27151.5	47269.8
Data		151568	302168	649504	1103240
Data/MC (%)		98.0	100.8	97.7	98.6

The difference between both SF channels is a consequence of the reduced electron reconstruction efficiency compared to the muon reconstruction.

For all three channels as well as for the combined yield, good agreement between the number of events predicted by simulation and observed in data is found. The differences between both yields for all channels range well below one standard deviation with respect to the total uncertainty of the predicted number of events. The individual contributions to the systematic uncertainty are further introduced in Sec. 9. The total number of events separated by the four data-taking periods can be found in Appx. 5. Here, similar conclusions compared to the combined yields can be drawn.

In Fig. 6.4, the comparison between the total number of events in simulation and data is shown for different steps of the event selection. Starting with the selection of two oppositely-charged leptons, the selected set of events is dominated by contributions from the DY+jets process. Only after applying the criteria regarding the invariant dilepton mass and the number of jets, the relative contribution from the dileptonic $t\bar{t}$ process becomes dominant and is significantly increased by the following requirements concerning the number of b-tagged jets and the missing transverse momentum. Across all selection steps, a good agreement between simulation and data can be observed, implying a reasonable normalization of the different SM contributions. Similar histograms split by $t\bar{t}$ decay channel are shown in Appx. 5.

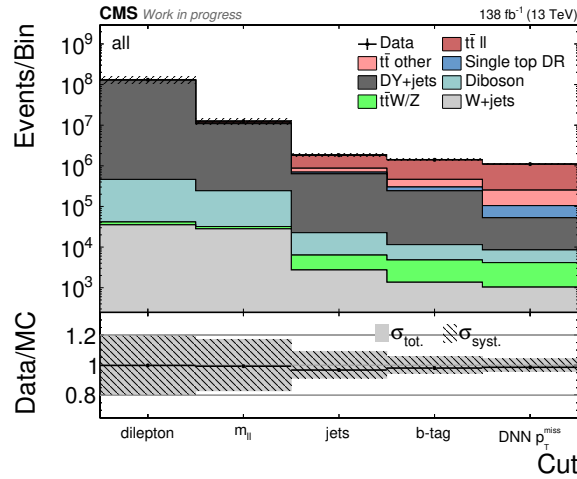


Figure 6.4: Total number of events after different steps of the event selection for all channels and data-taking periods combined. In the upper part of the plot, the individual contributions to the yield predicted by simulation are shown including the total uncertainty (hatched area) and the observed data yield (black). The lower part of the plot shows the ratio between data and simulation as well as the systematic and total uncertainties on the predicted yields.

6.3.2 Control distributions

Basic validations of the agreement between observed data and simulation can be performed based on simple control distributions, which take the different reconstructed physics objects into account, as shown in Fig. 6.5 and 6.6. Any significant differences between data and simulation could hint at missing or faulty corrections applied to the simulated events or contributions from additional SM processes not taken into account in the simulation. The agreement in distributions sensitive to potential BSM contributions, discussed in Sec. 10, is only studied once all analysis ingredients are fixed following a strict blinding strategy.

Figure 6.5 shows the control distributions based on the transverse momenta of the reconstructed leptons, which can be used to validate the corrections regarding the lepton identification and lepton energy. Even though the simulated spectra are slightly harder compared to the observed data, for both distributions reasonable agreement between data and simulation taking into account the corresponding uncertainties is observed. The simulated and observed distributions of the leading and subleading transverse jet momenta are also compared in Fig. 6.5. Here, good agreement between data and simulation is observed indicating reasonable corrections of the jet energy scale and jet energy resolution. Only in the high- p_T range of the leading jet, smaller differences between data and simulation are found, though well within the systematic uncertainties which increase up to more than 20% in this region.

Comparisons regarding the multiplicities of the reconstructed jets and b-tagged jets are shown in Fig. 6.6. Based on the multiplicity of jets, which shows a good agreement between data and simulation up to eight reconstructed jets, the simulation of additional FSR and ISR can be validated. Good agreement between data and simulation is also found in the number of b-tagged

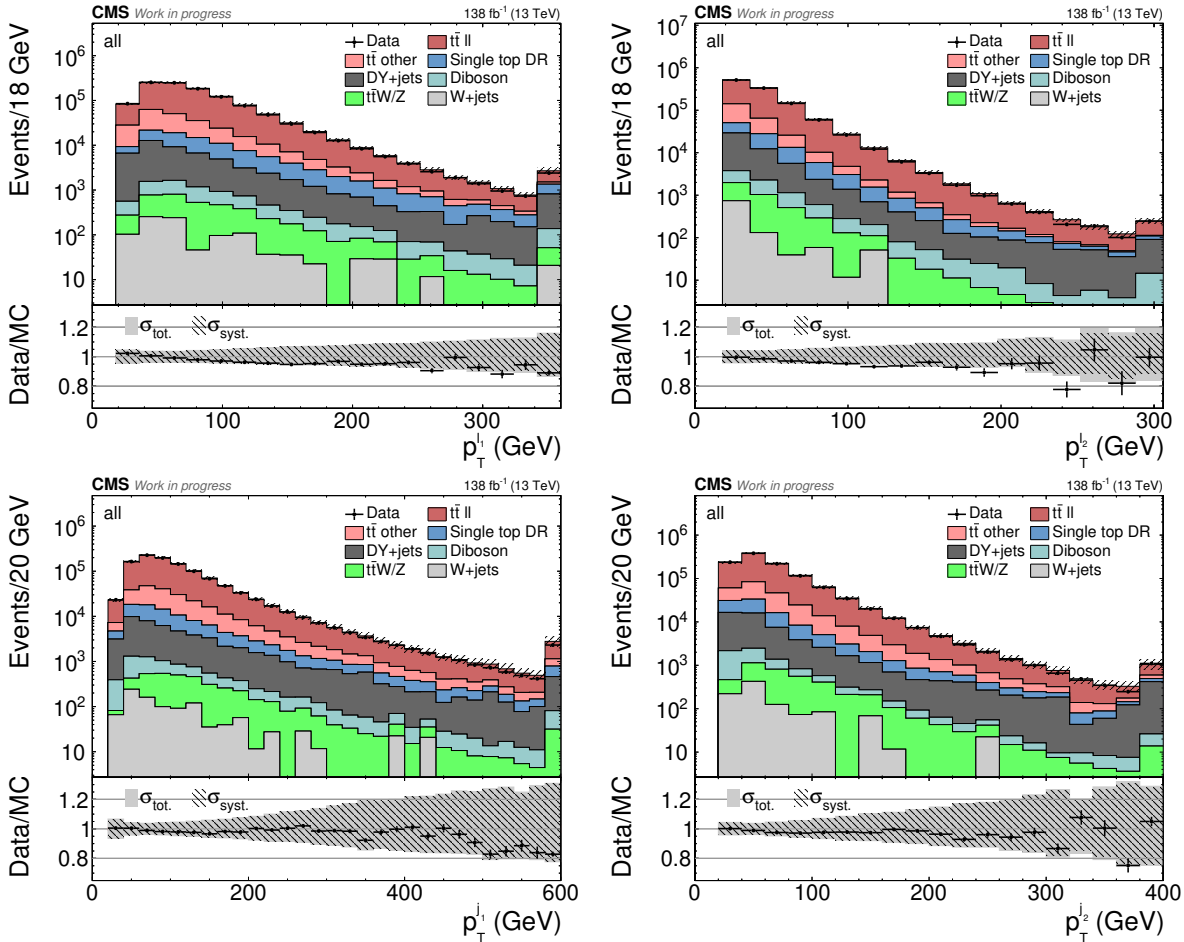


Figure 6.5: Control distributions as a function of the leading (left) and subleading (right) lepton (top) and jet (bottom) transverse momentum for all channels and data-taking periods combined. The upper parts of the plots show the observed data (black) and the contributions predicted by simulation including the total uncertainty (hatched area). The lower parts of the plots show the ratio between data and simulation as well as the systematic and total uncertainty on the predicted distributions. In all distributions, the last bin includes all events above the plotted range.

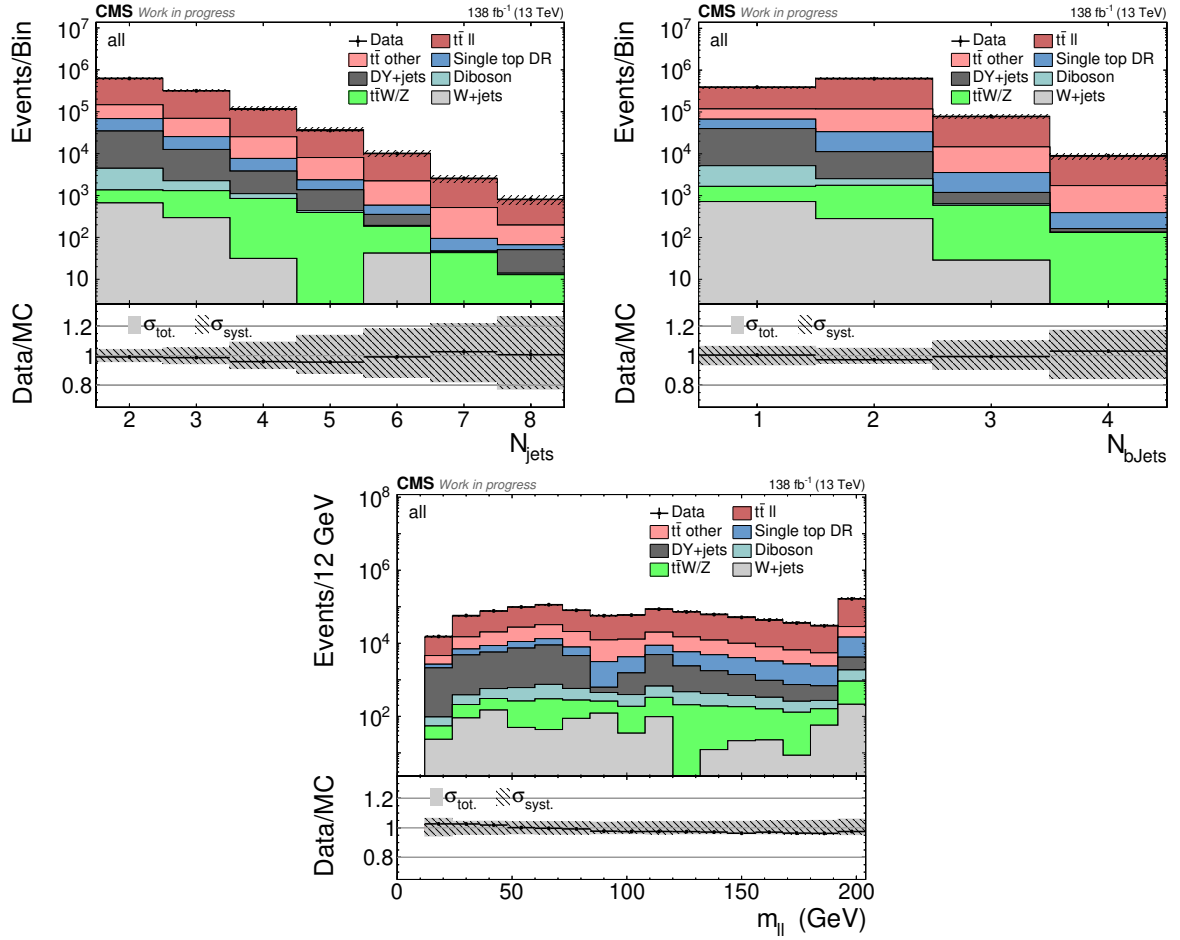


Figure 6.6: Control distributions as a function of the number of selected jets (top left), the number of selected b jets (top right) and the invariant mass of both leptons (bottom) for all channels and data-taking periods combined. The upper parts of the plots show the observed data (black) and the contributions predicted by simulation including the total uncertainty (hatched area). The lower parts of the plots show the ratio between data and simulation as well as the systematic and total uncertainties on the predicted distributions. In all distributions, the last bin includes all events above the plotted range.

jets up to a multiplicity of four jets indicating a reasonable application of the b-tagging scale factors discussed in Sec. 6.2.5. Here, the largest contributions from the signal $t\bar{t}$ process are observed for events with two b-tagged jets, as expected for the dileptonic $t\bar{t}$ signature.

The distribution of the invariant mass of the two reconstructed leptons is also shown in Fig. 6.6. Similar to the other control distributions, a good agreement between data and simulation is found, indicating a reasonable correction of the lepton energy scale. The dip in the distribution around the Z-boson mass is caused by the additional $m_{\ell\ell}$ -requirement applied in the SF channels, which significantly reduces the contribution from resonant DY+jets events.

The selection of control distributions discussed above split by the individual lepton channels can be found in Appx. 6. Here similar conclusion to the combined distributions can be drawn indicating a good agreement between data and simulation also for the individual lepton channels. Thus, in summary, no significant deviations between data and simulation, which would indicate missing corrections or contributions in the simulation, are observed.

Chapter 7

Deep neural network regression for p_T^{miss} correction

7.1	Phase space dependencies of the p_T^{miss} response	84
7.2	Machine learning and neural networks	85
7.2.1	Mathematical foundation	86
7.2.2	Hyperparameters	88
7.3	Initial architecture and training process	90
7.3.1	Training sample reweighting	91
7.4	Input features	92
7.4.1	Impact ordering	93
7.4.2	Goodness-of-fit tests	94
7.5	Hyperparameter optimization	96
7.6	Performance studies	98

The missing transverse momentum, introduced in the previous section, serves as a proxy for the transverse momentum $p_T^{\nu\nu}$ of the two neutrinos produced in dileptonic $t\bar{t}$ events. Thus, this observable plays a key role in the cross section measurements performed in this thesis, which are derived as a function of $p_T^{\nu\nu}$ and $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$. Due to the nature of the p_T^{miss} reconstruction, its resolution and response are affected by the momentum measurement of each PF candidate in an event, leading to a significantly reduced accuracy compared to, e.g., the reconstruction of individual charged leptons. In order to partially recover the resolution and response of the missing transverse momentum, a dedicated regression based on a deep neural network providing corrections for p_T^{miss} measurements in dileptonic $t\bar{t}$ events has been developed in the context of this thesis. Parts of the development and optimization of the neural network presented in the following have been performed within a master thesis [194], supervised in the scope of this Ph.D. thesis. Dedicated studies of correlations between the p_T^{miss} response and different phase space regions are presented in Sec. 7.1. In Sec. 7.2, a brief introduction to the general concept of machine learning and neural networks is given. The following three sections cover the different steps in the development of the deep neural network regression used to correct the p_T^{miss} measurements, including initial studies based on a baseline network, detailed studies of the input variables as well as a dedicated optimization of different network parameters. In the last section of this chapter, the performance of the regression with respect to p_T^{miss} reconstruction methods introduced in Sec. 6.2.6 is discussed.

7.1 Phase space dependencies of the p_T^{miss} response

The response of the missing transverse momentum, defined as the mean difference between the reconstructed and generated p_T^{miss} for a given set of events, is crucial for the cross section measurements performed in the present analysis. A stable initial response close to unity across a wide phase space reduces the impact of the unfolding procedure, which corrects the measured distribution for detector effects, as further discussed in Sec. 8. Furthermore, improved response and resolution allow for a finer binning in the cross section measurements leading to a larger number of measuring points, which can be compared to theory predictions, as further discussed in Sec. 8.2. Both of these aspects become especially important in case of the two-dimensional cross section measurements. Here, the p_T^{miss} response for different regions in $\min[\Delta\phi(p_T^{\text{miss}}, \ell)]$ has to be taken into account, motivating a dedicated study of the p_T^{miss} response as a function of the angular observable.

The results of such a study are summarized in Fig. 7.1, comparing different reconstruction methods and phase space regions. The mean response for all simulated dileptonic $t\bar{t}$ events passing the event selection shows a bias towards larger reconstructed missing transverse momentum (rec. p_T^{miss}) with respect to the generated missing transverse momentum (gen. p_T^{miss}) across the entire angular range. For these events, a weak dependency between the response and $\min[\Delta\phi(p_T^{\text{miss}}, \ell)]$ is observed. This dependency is increased for events with large rec. p_T^{miss} . Here, the response degrades significantly for higher $\min[\Delta\phi(p_T^{\text{miss}}, \ell)]$, indicating larger contributions from fake p_T^{miss} in this phase space region. Comparing the two different reconstruction methods, the observed trend is more pronounced for the PUPPI reconstruction compared to the PF result.

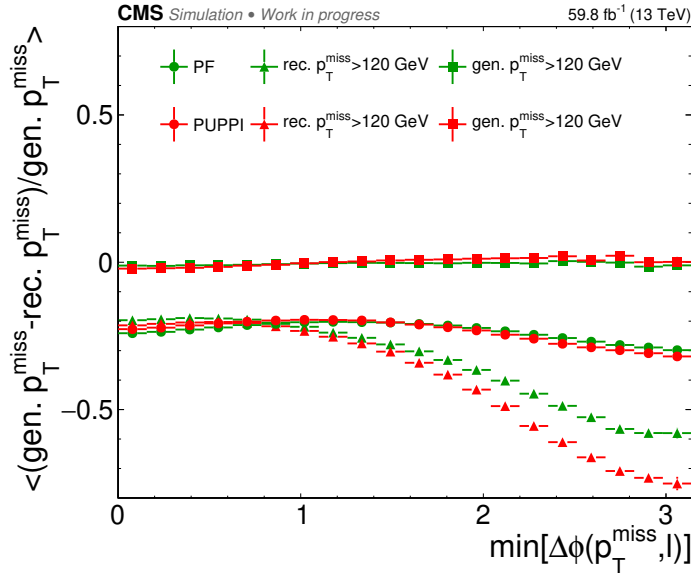


Figure 7.1: Mean relative p_T^{miss} response as a function of the minimal azimuthal distance between rec. p_T^{miss} and a selected lepton in simulated dileptonic $t\bar{t}$ events for the 2018 data-taking period. The results for the PF (green) and the PUPPI (red) reconstruction are compared for all selected events (circles), events with a rec. p_T^{miss} larger than 120 GeV (triangles) and events with a gen. p_T^{miss} larger than 120 GeV (squares).

Selecting only events with a significant amount of gen. p_T^{miss} removes this trend and a flat response centered around zero is observed for both reconstruction methods. This confirms the assumption of larger fake p_T^{miss} contributions at high $\min[\Delta\phi(p_T^{\text{miss}}, \ell)]$, since strict requirements on the gen. p_T^{miss} , in general, reduce the contributions of fake p_T^{miss} as observed in the present study.

The reason for the larger contributions of events with significant fake p_T^{miss} at large separation between the missing transverse momentum and the nearest lepton can be explained based on the sketch shown in Fig. 7.2. For a perfectly measured dileptonic $t\bar{t}$ event without fake- p_T^{miss} contributions, large $\min[\Delta\phi(p_T^{\text{miss}}, \ell)]$ values can only be observed in a setting where the dineutrino system is almost back-to-back with respect to each of the two leptons. Adding the requirement of a large rec. p_T^{miss} increases the preference for events, where the two neutrinos have a small azimuthal distance. Taking into account, that the neutrinos and the charged leptons each do not originate from the same decay, but are produced by two separated W bosons, the sketched setting has a rather small probability to occur in correctly measured dileptonic $t\bar{t}$ events. However, this phase space constraint is significantly reduced for events with large fake p_T^{miss} contributions. Here, the direction of the rec. p_T^{miss} has a weaker correlation with the neutrino directions leading to a larger probability for events with large $\min[\Delta\phi(p_T^{\text{miss}}, \ell)]$ even though the back-to-back relation sketched in Fig. 7.2 is not fulfilled.

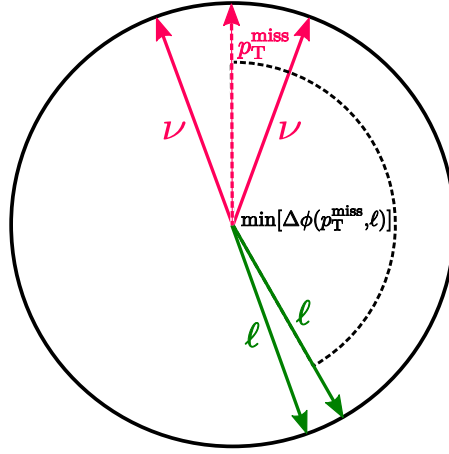


Figure 7.2: A sketch of the transverse plane of a potential dileptonic $t\bar{t}$ signature only showing the two neutrinos (red) and the two charged leptons (green) in a setting, which would lead to a large minimal angular distance $\min[\Delta\phi(p_T^{\text{miss}}, \ell)]$ between the missing transverse momentum and a lepton.

In summary, the phase space at large $\min[\Delta\phi(p_T^{\text{miss}}, \ell)]$ is disproportionally populated by events with significant mismeasurements leading to a degraded p_T^{miss} response in the given phase space. This effect is amplified by selecting events with large rec. p_T^{miss} , which prefers settings with collinear neutrino directions as well as events with larger fake- p_T^{miss} contributions.

The observation of this study combined with the general importance of a stable response and resolution of the missing transverse momentum for the cross section measurements motivates the additional corrections applied to the p_T^{miss} reconstruction derived by the deep neural network regression introduced in the remainder of this chapter.

7.2 Machine learning and neural networks

In general, machine learning [195–197] describes algorithms, which are able to learn from data in a way that their performance to solve a given task is improved by growing experience, i.e., with more data. Such algorithms are usually employed for tasks, which are too complex to be covered by fixed programs implemented by humans. Neural networks (NNs), as used for the regression performed in this thesis, correspond to a specific kind of machine learning algorithms inspired

by biological neural networks. In the following, the mathematical foundation, the training and different settings of such NNs, also known as hyperparameters, are briefly discussed. This discussion mainly focuses on the regression task using neural networks. A detailed review of other machine learning techniques and applications can be found in [195].

7.2.1 Mathematical foundation

Neural networks define a non-linear mapping of a set of input features given by $\mathbf{x} \in \mathbb{R}^n$ to a set of outputs $\mathbf{y} \in \mathbb{R}^k$ approximating some function f^* . This mapping is performed by deriving an approximative function $f(\mathbf{x}, \boldsymbol{\xi}) = \mathbf{y}$, which depends on a set of parameters $\boldsymbol{\xi}$ aiming to produce the best approximation of f^* . The optimization of these parameters, known as training of the network, is achieved by minimizing an objective function, also called loss function, which highly depends on the task performed by the NN. In case of supervised training, the relation $f^*(\mathbf{x}) = \mathbf{y}$ is known for a given data set and the loss function usually encodes some kind of difference between $f^*(\mathbf{x})$ and $f(\mathbf{x}, \boldsymbol{\xi})$. The following discussion mainly focuses on the class of feed-forward neural networks, which correspond to the most fundamental NNs. Here, only one direction of information flow from the input features to the outputs via different intermediate computations without any feedback connections is implemented.

The basic architecture of a feed-forward neural network is sketched in Fig. 7.3. The architecture can be seen as a chain structure of individual functions, where each of the functions corresponds to an individual layer. The first layer of the network (input layer) holds the input features $\mathbf{x}$, while the last layer, also known as output layer, contains the outputs $\mathbf{y}$. The layers in between the input and output layer are called hidden layers. A NN is denoted as a deep neural network (DNN) if it includes multiple hidden layers [197].

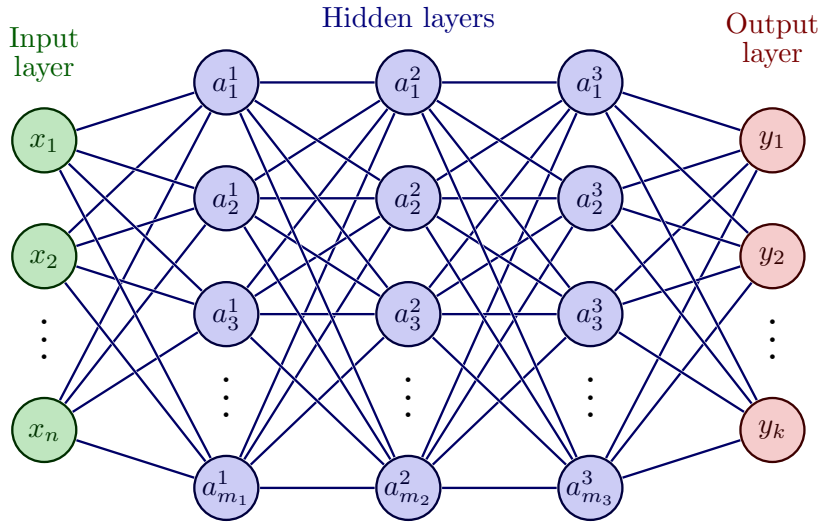


Figure 7.3: Sketch of the basic architecture of a fully-connected feed-forward neural network with three hidden layers (blue) [198]. The green and red circles correspond to the nodes of the input and output layer, respectively. Each connection illustrates a multiplication by a given weight w_{ij}^l and for each node, a dedicated bias b_i^l is applied.

Each layer holds a given number of units, also called nodes or neurons, which correspond to mathematical operations. In case of the input and output layer, the number of nodes has to match the dimension of $\mathbf{x}$ and $\mathbf{y}$, respectively. Each node of a hidden layer receives information

from all nodes of the previous layer and provides information to all nodes of the following layer (fully-connected). The output derived by a node of layer l can be expressed as:

$$a_i^l = \sigma^l \left(\sum_j w_{ij}^l \cdot a_j^{l-1} + b_i^l \right) . \quad (7.1)$$

Here, i and j denote the nodes of layer l and $l - 1$, respectively. Each output of layer $l - 1$ is multiplied with a weight w_{ij}^l , illustrated as a connection between two nodes in Fig. 7.3. The weighted outputs are summed over all nodes of layer $l - 1$. For each node of layer l , an additional bias b_i^l is added to the weighted sum. In order to model non-linear functions with respect to the input features, a non-linear activation function σ^l , which is the same for all nodes of a layer but can differ between different layers, is applied to derive the final output of a node.

The choice of the activation function strongly depends on the task performed by the network. For the layers of the DNN developed in the scope of this thesis a leaky rectified linear unit function (leaky ReLU) [195] given by

$$\sigma_{\text{leaky ReLU}} = \begin{cases} x, & \text{if } x \geq 0 \\ \alpha x, & \text{if } x < 0 \end{cases} , \quad (7.2)$$

is used as activation function, where α denotes the leakage of the ReLU, further discussed in the next section. Only for the output layer, the activation function is changed to a linear function, as required for a regression task.

The optimization of the parameter set $\boldsymbol{\xi}$, which contains the weights w_{ij}^l and biases b_i^l , is performed in the training of the network. Here, a loss function, which encodes a distance¹ between the true values $f^*(\mathbf{x})$ and the predictions $f(\mathbf{x}, \boldsymbol{\xi})$, is minimized. For regression tasks, one of the most popular choices of a loss function is given by the mean squared error (MSE):

$$L_{\text{MSE}} = \frac{1}{N} \sum_{i=1}^N (f(\mathbf{x}_i, \boldsymbol{\xi}) - f^*(\mathbf{x}_i))^2 , \quad (7.3)$$

where the sum runs over all data points of a given data set with size N . However, for the present network, a slightly more sophisticated loss function given by the logarithmic hyperbolic cosine (logcosh) loss [199] is used:

$$L_{\text{logcosh}} = \frac{1}{N} \sum_{i=1}^N \log [\cosh (f(\mathbf{x}_i, \boldsymbol{\xi}) - f^*(\mathbf{x}_i))] . \quad (7.4)$$

For small differences between prediction and true function, the logcosh behaves like the MSE, while for large differences it is proportional to the absolute difference. Thus, compared to the MSE loss, outliers have a smaller impact on the logcosh loss, which, in the context of the present network, reduces the impact of falsely reconstructed events in the training process. The minimization of the loss function is performed based on a gradient descent method, which is widely used in the training of neural networks and ensures a convergence towards a local

¹For outputs with a dimension larger than one, the mean distance across all dimensions is used.

minimum of the loss function. In this method, the parameters are adjusted iteratively following the direction of the gradient of the loss function L with respect to ξ :

$$\xi_{t+1} = \xi_t - \eta_t \cdot \nabla_{\xi} L(\xi_t) , \quad (7.5)$$

where t denotes the step of the iterative procedure and η_t corresponds to a learning rate, which controls the size of the update applied to the parameter set in step t . Since the loss function depends on all data points used in the training of the network, the computation of its gradient becomes computationally expensive, if a large data set is used for the training. To avoid this problem, the present network is trained based on the stochastic gradient descent (SGD) method, where the gradient is approximated based on small subsets (batches) of the total data set. In each step of the SGD method, ξ is adjusted only based on a single batch, which introduces stochasticity to the training procedure reducing the chance for the optimization to get stuck in a single isolated local minimum. Furthermore, the training on individual batches significantly reduces the computing time required for the training. One loop over the entire set of batches is also known as epoch. The calculation of the gradient itself is based on the backpropagation algorithm, which applies the chain rule of partial differentiation exploiting the layered structure of the network.

7.2.2 Hyperparameters

In the training of a network, usually, only the weights between two nodes and the biases per node are optimized. However, even for rather simple fully-connected feed-forward neural networks, there are multiple additional parameters that can highly affect the networks' performance. Such parameters, which are not optimized during the training process, are summarized as hyperparameters. For the present network, a dedicated optimization of the hyperparameters is performed as discussed in Sec. 7.5. The different hyperparameters considered in this optimization are briefly introduced in the following.

Network architecture

The network architecture defined by the number of layers and number of nodes per layer can have a significant impact on the network performance. For rather complex tasks, improved performance is usually observed for larger networks, while for simpler tasks, smaller networks may lead to a better overall performance based on a faster convergence. Two main problems can occur if the architecture does not match the complexity of the task. If the size of the network is too large, the network might perform significantly better on the data set used for the training compared to a new data set, previously unknown to the network. Such a lack of generalization is also known as overtraining. In the opposite case, where the network is too small to match the complexity of the task, the network may not be able to learn all necessary correlations present in a data set. Hence, the architecture of a network should be optimized for the corresponding task.

ReLU leakage

The behavior of the activation function (cf. Eq. (7.2)) used for most of the layers of the present network is partially defined by the leakage parameter α . Compared to the nominal ReLU activation, which corresponds to $\alpha = 0$, the extension made by the leaky ReLU activation aims to

reduce the risks of vanishing gradients. In case of the nominal ReLU, such vanishing gradients can occur for arguments below zero potentially leading to slow convergence of the training procedure. For the leakage, usually, parameters of the order of 10^{-1} to 10^{-2} are employed, to still keep the main characteristics of the ReLU activation.

Learning rate

The learning rate introduced in the context of the gradient descent method defined in Eq. (7.5) can have a significant impact on the convergence of the training process. If too large, the optimization might oscillate around a local minimum without reaching its actual valley. If too small, only a slow convergence can be achieved increasing the computing time of the training process. In the training of the present network, the learning rate is adapted during the training procedure using the Adam optimizer [200]. This optimizer introduces an additional hyperparameter, called step size, which can be seen as an approximate upper bound for the size of the parameter updates. Thus, similar considerations compared to the learning rate also hold for the step size parameter of the Adam optimizer.

Batch size

The SGD method used for the optimization of the network parameters during training is based on splitting the data set used for the training into individual batches. The size of the batches is a priori not fixed and should be optimized to achieve a balance between speed and accuracy of the gradient approximation within the SGD method.

Regularization

In general, all strategies aiming to reduce overtraining can be collectively summarized as regularization. However, in the context of the present network, the regularization corresponds to a norm penalty term, which is added to the loss function:

$$L' = L_{\log\cosh} + \lambda \|\xi'\|_2^2, \quad (7.6)$$

where $\|\xi'\|_2^2$ denotes the L_2 norm [195] of the reduced parameter set only containing the weights of the network and no biases. The strength of the regularization is controlled by the regularization parameter λ , which should be optimized for a given network. The addition of the norm penalty to the loss effectively constrains the magnitude of the weights, penalizing large weights in the training process. For the biases usually no penalty term is added, since controlling only one respective node the biases are easier to optimize with less data, while the optimization of the weights connecting two respective nodes is more involved.

Dropout

The dropout method represents an additional regularization technique, which aims to prevent overtraining by reducing the number of artificial correlations between nodes. In this method, individual nodes are ignored (dropped out) in each step of the SGD with a probability p_D . The corresponding training step is only performed for the reduced network lowering the reliance on individual nodes and improving the generalization of the network. To adequately propagate the

method to the network prediction the trained weights have to be multiplied by p_D resulting in an averaged weight over all reduced networks.

7.3 Initial architecture and training process

An additional correction applied to the reconstructed missing transverse momentum is derived for dileptonic $t\bar{t}$ events using a regression based on a deep neural network, also called DNN regression in the following. Since the observables employed in the cross section measurements not only depend on the magnitude of p_T^{miss} but also on the azimuthal angle $\phi(p_T^{\text{miss}})$, the regression is performed by a fully-connected feed-forward neural network with two output nodes. The two output nodes are trained to predict the difference between the reconstructed PUPPI p_T^{miss} and the generated p_T^{miss} split into the corresponding x - and y -component:

$$f^*(\mathbf{x}) = \begin{pmatrix} [\vec{p}_T^{\text{miss}}(\text{PUPPI}) - \vec{p}_T^{\text{miss}}(\text{gen.})]_x \\ [\vec{p}_T^{\text{miss}}(\text{PUPPI}) - \vec{p}_T^{\text{miss}}(\text{gen.})]_y \end{pmatrix}. \quad (7.7)$$

This split into two spatial components allows for a simultaneous correction of the magnitude and azimuthal angle derived in the PUPPI reconstruction. By choosing the difference between the rec. and the gen. p_T^{miss} , as defined in Sec. 6.2.6, the regression aims to correct the p_T^{miss} reconstruction for effects of detector resolution and response. Any further differences between the gen. p_T^{miss} and the $p_T^{\nu\nu}$ observable used in the cross section measurements, which can, for example, be caused by additional non-prompt neutrinos in the event, are taken into account in the unfolding procedure discussed in Sec. 8.

For some of the initial studies of the DNN regression, a baseline network with a total of 54 input features, listed in Appx. 7, and manually fixed hyperparameters, summarized in Tab. 7.1, is considered. For the final DNN regression, the number of input features is significantly reduced and a dedicated optimization of the hyperparameters is performed as discussed in Sec. 7.4 and 7.5, respectively.

The network is trained using the TENSORFLOW [201] backend combined with the KERAS [202] interface making use of several SCIKIT-LEARN [203] and PANDAS [204] packages. The data points used during the training correspond to simulated dileptonic $t\bar{t}$ events from the alternative sample based on the MADGRAPH5_AMC@NLO generator. The alternative sample is chosen to avoid introducing any bias by using the nominal $t\bar{t}$ signal sample, which is also employed in other parts of the analysis. The simulated events have to fulfill the full event selection to be used in the training process, except for the p_T^{miss} requirement. Ignoring the p_T^{miss} requirement allows to perform the said requirement in the event selection based on p_T^{miss} corrected by the DNN regression. The full set of simulated events available for the training procedure is randomly split into three separate subsets. The training sample, which corresponds to the sample actually used in the SGD method of the training, accounts for 60% of all events. The validation sample, which accounts for 20% of the events, is used to validate the performance of the network during the training as well as during the hyperparameter optimization. The remaining 20% of the events correspond to the test sample, which is only used to measure the network performance after all training and optimization steps. This commonly used split of event samples allows to detect overtraining by comparing the performance between training and validation sample as well as to derive an unbiased estimate of the network performance based on the test sample.

In general, the present network is trained for a total of up to 400 epochs. However, to get an additional handle against overtraining and to reduce the training time for rapidly converging networks, an early stopping mechanism [195] is employed during the training. This mechanism terminates the training procedure as soon as one of the following requirements is fulfilled and restores the network with the smallest value of the loss function evaluated with the validation sample (validation loss):

- The validation loss does not decrease over a range of ten epochs.
- The relative difference between the validation loss and the loss function evaluated with the training sample (training loss) exceeds 2% for ten epochs in a row.
- The absolute difference between validation and training loss increases ten epochs in a row.

To account for potential differences between the four data-taking periods, which can, for example, originate from the upgrade of the pixel detector performed before the 2017 data-taking period (cf. Sec. 3.2.3), different networks are individually trained for each data-taking period. The selection of input features and the optimized hyperparameters do not differ between these networks.

7.3.1 Training sample reweighting

The distribution of the gen. p_T^{miss} for the dileptonic $t\bar{t}$ events used in the training sample is shown in Fig. 7.4. As expected, the events are not evenly distributed with respect to the gen. p_T^{miss} but peak around 50 GeV, followed by a steeply falling spectrum. The phase space at high gen. p_T^{miss} with momenta above 300 GeV has a rather poor representation. Given that the gen. p_T^{miss} is part of the target defined in Eq. (7.7) and thus also used in the loss function, the observed features of the gen. p_T^{miss} distribution are expected to have an impact on the training process.

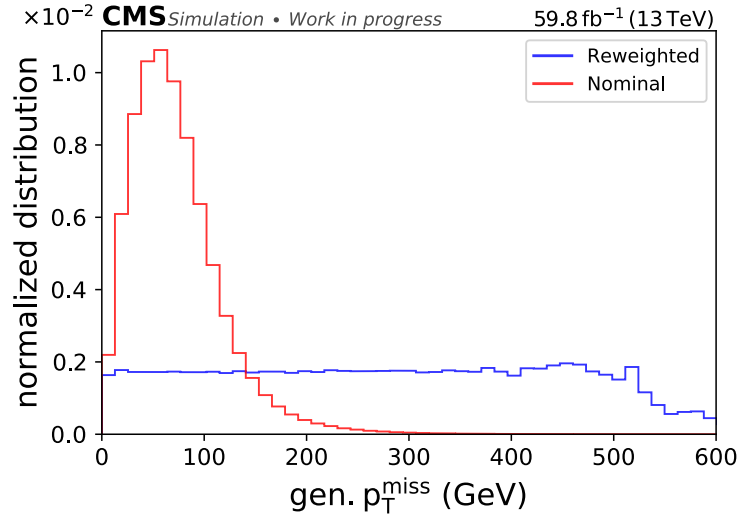


Figure 7.4: Normalized gen. p_T^{miss} distributions for the nominal training sample (red) and the reweighted training sample (blue) as used in the network training for the 2018 data-taking period. The bin widths used in the two histograms do not match the widths used in the reweighting method applied to the training sample.

This assumption is confirmed based on the results shown in Fig. 7.5. While the PUPPI reconstruction yields a flat response around zero for the region with gen. $p_T^{\text{miss}} > 100$ GeV, the response

for p_T^{miss} corrected by the DNN regression leads to a bias starting already below 100 GeV. Here, p_T^{miss} is significantly underestimated for most parts of the studied phase space. This behavior can be explained by the gen. p_T^{miss} distribution of the training sample. The best response is observed at around 50 GeV, which matches the peak position of the distribution. This phase space is represented by a large fraction of events leading to a significant impact on the loss function used during the training procedure. The opposite holds for the region at high gen. p_T^{miss} . Here, only a few events are considered in each step of the SGD leading to a significantly reduced performance at high gen. p_T^{miss} .

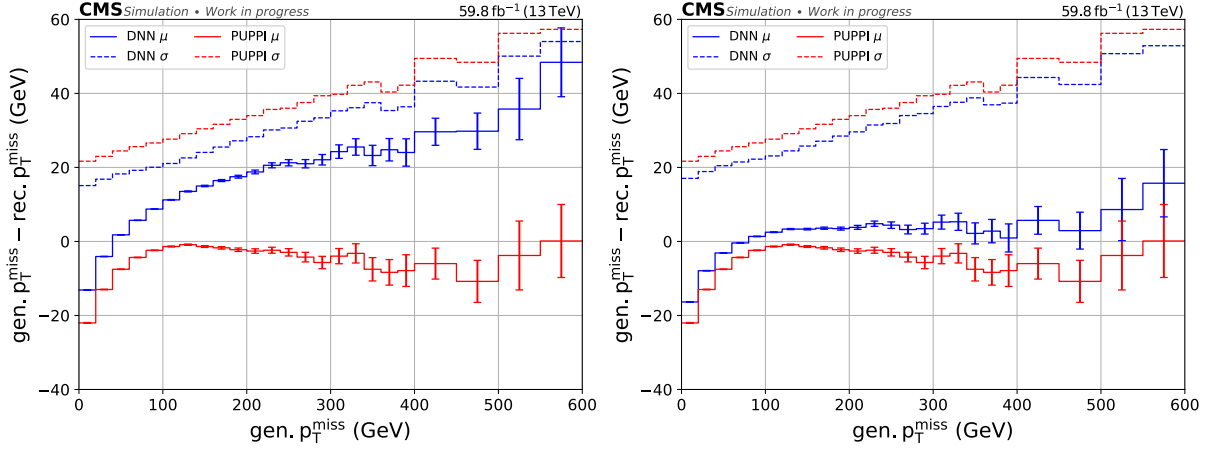


Figure 7.5: Performance of the PUPPI p_T^{miss} (red) and the DNN p_T^{miss} (blue), trained with the nominal (left) and with the reweighted training sample (right), as a function of the gen. p_T^{miss} . The solid (dashed) line corresponds to the mean (standard deviation) of the difference between the gen. p_T^{miss} and the rec. p_T^{miss} in each gen. p_T^{miss} bin. The error bars correspond to the statistical uncertainty on the mean. The result for PUPPI p_T^{miss} agrees between both plots. For both DNN p_T^{miss} results, the baseline network has been used.

To improve the performance of the network towards a uniform response with respect to a large gen. p_T^{miss} range, the training sample is manipulated in the following way. The events with gen. $p_T^{\text{miss}} < 500$ GeV are separated into gen. p_T^{miss} bins with equal widths of 8 GeV. For each of the bins, a weight is derived such that the weighted event count is the same for all bins. Events with gen. $p_T^{\text{miss}} \geq 500$ GeV are allocated the same weight as derived for the last bin considered in the reweighting method. This manipulation leads to a reweighted training sample with a flat gen. p_T^{miss} spectrum up to gen. $p_T^{\text{miss}} = 500$ GeV followed by a smooth decline, as shown in Fig. 7.4.

The impact on the performance of the DNN regression using the reweighted training sample can be seen in Fig. 7.5. Due to the uniform population of events across a wide gen. p_T^{miss} range, a flat response for the region above 100 GeV is now also observed for the DNN p_T^{miss} removing the significant bias found before. Even though the reweighting procedure slightly worsens the resolution, the DNN p_T^{miss} still yields an improved resolution compared to PUPPI p_T^{miss} across the studied phase space.

7.4 Input features

The DNN used for the regression task is based on a variety of different input features. These features contain global event variables, like the missing transverse momentum derived by the PF

and PUPPI reconstruction or the number of jets and pileup vertices in an event. In addition, features of the final state particles, e.g., lepton and jet kinematics and the flavor of the selected charged leptons, are employed. The final group of features corresponds to variables combining multiple reconstructed objects, such as angular distances between jets and leptons or invariant masses of different final state systems. Input features describing some kind of momentum are usually split into the corresponding x - and y -components to match the structure of the network's output nodes. The full set of all 54 input features considered in the development of the network is shown in Appx. 7.

In order to decrease the complexity of the network, the impact of all 54 input features on the DNN output is studied in the following aiming to remove redundant input features as well as features with only small impacts on the output. Since the network is exclusively trained on simulated events, any mismodeling of the input features in the simulation could potentially lead to a bias in the network prediction. To reduce the risk of such biases and to improve confidence in the prediction of the network, the modeling of all input features used in the final network architecture is thoroughly tested by comparing observed data and simulation, as discussed in Sec. 7.4.2.

7.4.1 Impact ordering

The impact of the individual input features on the prediction of the network, also known as feature attribution, is measured in order to rank the features with respect to their impact aiming to reduce the total number of input features. For the feature attribution, a method based on the concept of Shapley values [205] originally developed in the context of game theory is applied. Here, each input feature corresponds to a player in a game, where the outcome of the game relates to the prediction of the network. The Shapley values quantify the contribution of each player to the outcome, thus, estimating the contribution of each individual input feature to the prediction of the network. The Shapley values analyzed in the following are derived using the expected gradient method [206], which is based on gradients of the prediction $f(\mathbf{x}, \xi)$ with respect to the individual input features $\mathbf{x}_i$.

The Shapley values for the 20 input features with the largest impact based on the baseline network with manually fixed hyperparameters (cf. Sec. 7.3) are shown in Fig. 7.6. The impact is derived separately for the two components of the target. For most of the input features split into their x - and y -component, a strong correlation with the corresponding direction of the target is found. The largest overall impact on the network prediction is observed for features related to the PUPPI and PF p_T^{miss} as well as for both transverse components of the vectorial momentum sum $p_{x,y}^{\text{all j}}$ of all reconstructed jets. Other input features showing a significant impact on the prediction correspond to the transverse momenta of the leading jet and leading lepton and the invariant mass of the system including all selected jets ($m_{\text{all j}}$) and all selected jets plus both selected leptons ($m_{\text{ll+all j}}$).

To get an estimate of the number of input features required for a reasonable performance in the DNN regression, the baseline network is trained starting with the full set of features. In the following steps, the number of features is iteratively reduced starting with the lowest-impact features. After each of these steps, the validation loss is analyzed and compared to the result derived in the previous step. For the networks based on a number of input features in the range from 20 to 54, no significant change in the validation loss is observed. However, restricting the set of input features to the ten features with the largest impact results in a significant increase in the validation loss and thus in a reduced performance of the network. This indicates that the first 20 input features include enough information for a reasonable performance of the network

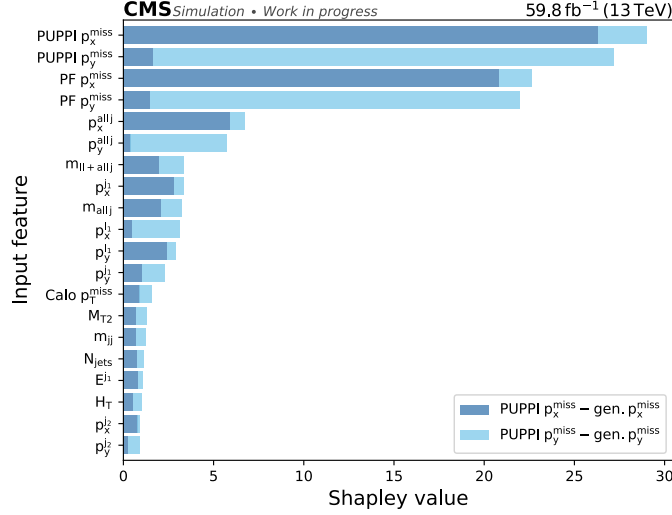


Figure 7.6: Shapley values for the first 20 input features with the largest impact on the network prediction using the baseline network with 54 input features and manually defined hyperparameters trained for the 2018 data-taking period. The impact is derived separately for the x - (dark blue) and the y -component (light blue) of the target difference. The individual results are stacked on top of each other for each individual feature.

while the remaining 34 features would only add redundancy. Thus, for the following steps, only the 20 features shown in Fig. 7.6 are considered.

7.4.2 Goodness-of-fit tests

The present network is exclusively trained on simulated events, which might introduce a potential bias if mismodeling of input features is present in the simulation. To reduce the risk of such a bias impacting the network prediction, the observed and simulated distributions for each of the remaining 20 input features are compared. Similar studies are performed for every possible combination of two input features to ensure a good description of the correlations between different features. In each of these studies, a goodness-of-fit (GoF) test is carried out based on a binned maximum likelihood fit using the CMS combine tool [207, 208]. The corresponding likelihood function is defined as

$$\mathcal{L}(\text{data}|\mu, \theta) = \prod_i [\text{Poisson}(n_i|\mu \cdot s_i(\theta) + b_i(\theta))] \cdot p(\tilde{\theta}|\theta) \quad (7.8)$$

$$\text{with Poisson}(n_i|\mu \cdot s_i + b_i) = \frac{(\mu s_i + b_i)^{n_i}}{n_i!} e^{-\mu s_i - b_i},$$

where n_i , s_i and b_i denote the observed, predicted signal and predicted background yield in bin i , respectively. The predicted number of signal events is scaled by the signal strength parameter μ . Both predicted yields depend on a set of nuisance parameters θ , which integrate systematic and statistical uncertainties into the likelihood function via prior nuisance probability density functions $p(\tilde{\theta}|\theta)$, where $\tilde{\theta}$ corresponds to the best prior estimate of the nuisance parameters. For uncertainties, which affect the rate in each bin i in the same way, e.g., uncertainties on the measured luminosity, the prior probability is described by a log-normal distribution. Uncertainties, which also affect the shape of the distributions, are considered with a template morphing proce-

ture [209], where a vertical interpolation between two templates, corresponding to a $\pm 1\sigma$ shift of the nuisance parameter, is performed. Statistical uncertainties due to the limited number of simulated events are taken into account based on the Barlow-Beeston-lite approach [208, 210].

For the GoF tests of the individual one- and two-dimensional distributions, a ratio of two likelihoods following the Neyman-Pearson lemma [211] is used in the test statistic:

$$\lambda = \frac{\mathcal{L}(\text{data}|\hat{\mu},\hat{\theta})}{\mathcal{L}_{\text{sat}}} . \quad (7.9)$$

Here, $\hat{\mu}$ and $\hat{\theta}$ denote the signal strength parameter and the set of nuisance parameters, which maximize the likelihood in Eq. (7.8), respectively. Since for the given problem, no alternative hypothesis is defined, a so-called saturated model $\mathcal{L}_{\text{sat}}$ [212] is constructed, where in each bin, the model prediction equals the number of observed events. For sufficient large statistics in each bin, the test statistic $q = -2\ln(\lambda)$ follows a χ^2 distribution. Based on this, a p -value can be defined quantifying the probability of obtaining a value of the test statistic as least as large as the value observed in data. Thus, small (large) p -values indicate a poor (good) agreement between the simulated and the measured distribution.

In each of the GoF tests, an expected distribution $f(q)$ of the test statistic is derived based on 500 toy data sets sampled from the expected distribution of the input feature or input feature combination of interest. Utilizing $f(q)$, the p -value for each test can be estimated by

$$p = \int_{q_{\text{obs}}}^{\infty} f(q) dq \approx \frac{N_{\text{toy}}(q_{\text{toy}} > q_{\text{obs}})}{N_{\text{toy}}} , \quad (7.10)$$

where N_{toy} denotes the number of toys and $q_{\text{toy}}(q_{\text{obs}})$ corresponds to the value of the test statistic for the toy (observed) data. The tests are performed individually for each of the four data-taking periods combining the three different decay channels. The full set of systematic uncertainties, discussed in Sec. 9, is taken into account in the binned maximum likelihood fits. An example of a GoF-test result for one possible combination of two input features is shown in Fig. 7.7. Here, the p -value of around 0.8 indicates a good agreement between the observed and expected distribution. The sampled test statistic for the toy data sets can be described by a χ^2 distribution matching the number of bins used in the tested distribution.

The p -values for all one- and two-dimensional distributions individually derived for the four data-taking periods are summarized in Fig. 7.8. In the one-dimensional tests, ignoring potential correlations between input features, good agreement between the observed and expected distributions of all 20 input features in each of the four data-taking periods is observed. In the two-dimensional tests, however, a few of the input feature combinations yield p -values below 5%, indicating poor modeling of the correlations in simulation. In a majority of these combinations, either the missing transverse momentum derived from calorimeter entries (Calo $p_{\text{T}}^{\text{miss}}$), the transverse mass $M_{\text{T}2}$ [213, 214] or the number of jets N_{jets} is involved, where the latter holds mainly for the 2017 data-taking period. To avoid the propagation of these poorly modeled correlations to the network training, the three corresponding input features are removed leading to a total of 17 input features used in the final network architecture. For the final set of input features, only four input feature combinations yield p -values below 5%. The features used in these combinations are kept for the following reasons. None of these input features shows a systematic mismodeling, where multiple combinations indicate poor modeling across multiple data-taking periods. In addition, some of the combinations are based on strongly correlated features, e.g.,

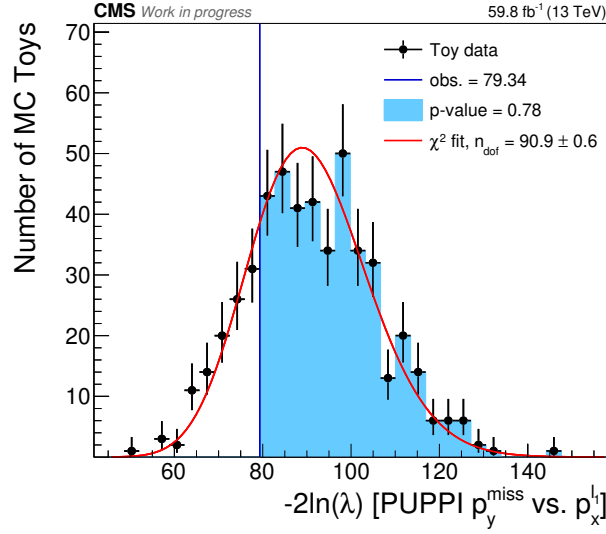


Figure 7.7: Example distribution of the test statistic used for the GoF tests performed for a two-dimensional distribution based on the y -component of PUPPI p_T^{miss} and the x -component of the leading lepton momentum using a total of 90 bins. The result based on the toy data is shown in black. The observed value corresponds to the vertical blue line. A χ^2 distribution is fitted to the histogram derived from the toy data results. The value of the observed test statistics, the corresponding p -value and the number of degrees of freedom from the χ^2 fit are given in the legend.

the combination of PUPPI p_x^{miss} and PF p_x^{miss} , where parts of the two-dimensional phase space are only populated by a small number of events, which can lead to potential problems in the maximum likelihood fit. Finally, a few individual tests yielding small p -values are expected when performing a total of almost 900 individual GoF tests.

7.5 Hyperparameter optimization

The performance of a neural network can significantly depend on the choice of hyperparameters, which are not optimized during the training process but have to be defined before the actual training of the network. A good choice of hyperparameters can usually not be generalized but has to be individually found for different tasks. While for the baseline network the hyperparameters, shown in Tab. 7.1, were chosen manually based on prior knowledge and assumptions, the parameters used in the final network are subject to a dedicated optimization aiming to improve the performance of the DNN regression.

Since a network usually depends on a larger set of hyperparameters, e.g., a total of seven parameters for the present network (cf. Sec. 7.2.2), an individual optimization of every single parameter, requiring a large number of individual training and evaluation steps, is usually not feasible due to limited computing resources. Furthermore, an individual optimization would neglect potential correlations between hyperparameters, which might affect the performance of the network. Thus, for the present network, the hyperparameters are optimized simultaneously using a Bayesian optimization algorithm [215, 216].

As suggested by its name, the Bayesian optimization algorithm is based on Bayes' theorem taking the outcome of previous optimization steps into account. For the objective function $f(\mathbf{u})$ given by the loss of the network as a function of the set of hyperparameters $\mathbf{u}$ a probabilistic model is constructed using a Gaussian process, which is updated in each step of the optimization.

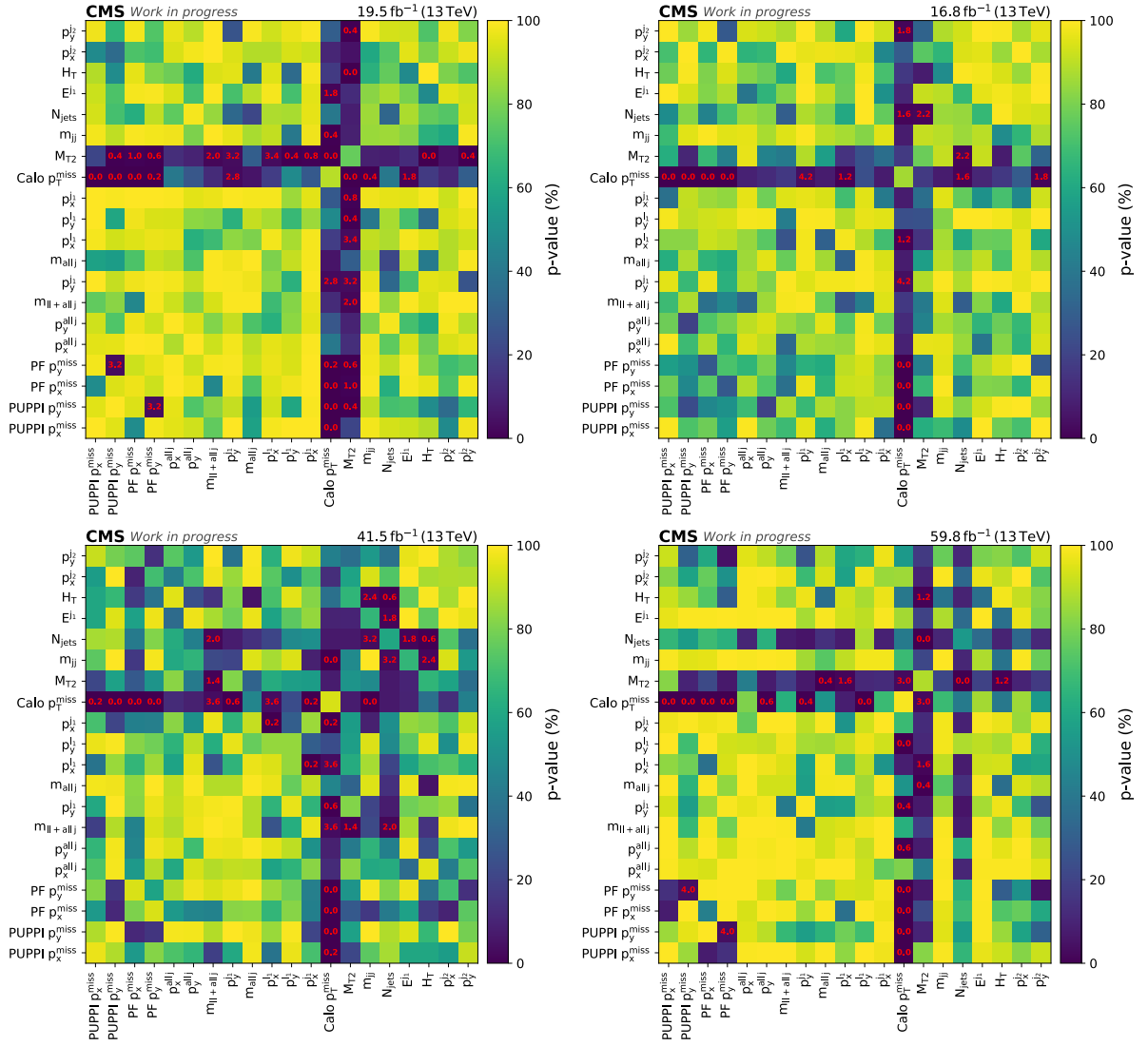


Figure 7.8: Observed p -values for the GoF tests performed for all 20 input features (diagonal elements) and possible combinations between two input features (off-diagonal elements). The results are shown for the 2016preVFP (top left), 2016postVFP (top right), 2017 (lower left) and 2018 (lower right) data-taking period. For tests, which yield a p -value below 5%, the corresponding p -value is shown in red digits.

This model corresponds to the prior probability $P(f_i)$ in Bayes' theorem. In each step of the optimization, the objective function is evaluated for a given hyperparameter set $\mathbf{u}_i$, which yields an observation given by $\mathcal{D}_i = \{\mathbf{u}_i, f(\mathbf{u}_i)\}$ and the corresponding likelihood $P(\mathcal{D}_i|f_i)$ taking the current version of the Gaussian process f_i into account. With each iteration, the Gaussian process and therefore the prior $P(f_i)$ is updated using the information from evaluation $\mathcal{D}_i$ of the objective function. For the selection of the hyperparameter set $\mathbf{u}_{i+1}$ used in the next step of the optimization a posterior based on Bayes' theorem is derived:

$$P(f_i|\mathcal{D}_i) \propto P(\mathcal{D}_i|f_i)P(f_i) . \quad (7.11)$$

Based on the posterior, an acquisition function is constructed and used to determine the next point of evaluation $\mathbf{u}_{i+1}$ in the hyperparameter space. For the present optimization, the Upper Confidence Bound [215] is employed as acquisition function allowing for a trade-off between regions, where the objective function has a large uncertainty (exploration), and regions, where the objective function is expected to be large (exploitation).

The optimization of the present network is performed based on the seven hyperparameters introduced in Sec. 7.2.2 using the negative validation loss of the network as the objective function. For the optimization, a total of 200 iterations, where each iteration includes a training step and an evaluation step, are performed. The first nine iterations are based on random guesses of hyperparameters within reasonable ranges, while for the remaining iterations, the Bayesian optimization algorithm is employed. The optimization is performed based on simulated events from the 2018 data-taking period. Comparing the validation loss of the baseline and the optimized network, an improvement of only around 1% is achieved by the optimization. However, as shown by the optimized parameters in Tab. 7.1, the optimization yields a much smaller network with only two instead of six hidden layers as used for the baseline network. Thus, the main outcome of the optimization does not correspond to a significant reduction of the loss, but rather a large reduction of the network complexity while keeping a stable performance, which significantly decreases the CPU time required for training and evaluating the network.

Table 7.1: Choice of hyperparameters for the baseline and the optimized network. The numbers of nodes are given per layer starting with the first hidden layer of the network.

Hyperparameter	Baseline network	Optimized network
Number of hidden layers	6	2
Number of nodes	1024, 1024, 1024, 512, 256, 128	512, 256
ReLU leakage	0.2	0.12
Step size (Adam optimizer)	$1.0 \cdot 10^{-4}$	$4.4 \cdot 10^{-4}$
Batch size	5000	666
Regularization weight	0.05	0.0026
Dropout probability	0.35	0.35

The optimized hyperparameters, as listed in Tab. 7.1, are derived based on the 2018 data-taking period but also applied to the networks used in the remaining data-taking periods. This practice is motivated by keeping a uniform approach for all periods and additionally justified by comparing the validation loss of the baseline network and the optimized network in each of the four periods, which yields a similar improvement as observed for the 2018 data-taking period.

7.6 Performance studies

The final DNN regression based on 17 input features and the optimized hyperparameters is studied on the $t\bar{t}$ test sample not used in the training or optimization process. The performance is compared to the four different p_T^{miss} reconstruction methods introduced in Sec. 6.2.6. The comparisons with respect to the magnitude and the azimuthal angle for the 2018 data-taking period are shown in Fig. 7.9. In both cases, the DNN regression significantly improves the performance compared to the PUPPI and PF reconstruction. In contrast to PUPPI p_T^{miss} , which by construction acts as the base of the DNN p_T^{miss} (cf. Eq. (7.7)), the resolution is improved by more than 15% for the magnitude and around 11.5% for the azimuthal angle. This simultaneous improvement verifies the approach of using two output nodes in the DNN regression. Further-

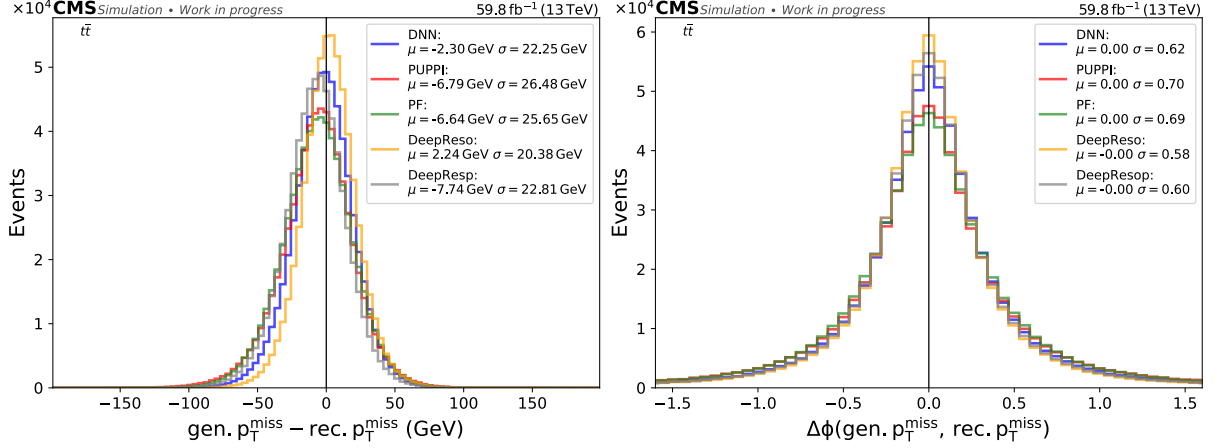


Figure 7.9: Difference between generated and reconstructed p_T^{miss} with respect to the magnitude (left) and the azimuthal angle (right) comparing DNN (blue), PUPPI (red) and PF (green) p_T^{miss} as well as the resolution (orange) and response tune (grey) of the Deep p_T^{miss} reconstruction. The results are derived based on the 2018 data-taking period using the $t\bar{t}$ test sample. The mean and the standard deviation of each distribution are listed in the legend.

more, the bias in the response of the magnitude present in the PUPPI and PF reconstruction is significantly reduced, while no additional bias is introduced with respect to the angle. The performance studies for the remaining data-taking periods can be found in Appx. 8. Comparing the DNN p_T^{miss} to the two Deep p_T^{miss} reconstructions yields an improvement with respect to the response tune, while the resolution tune seems to outperform all other reconstructions.

The main drawback of the Deep p_T^{miss} resolution tune, however, becomes evident when studying its performance as a function of the generated p_T^{miss} , as shown in Fig. 7.10. While the four remaining reconstructions yield a flat response around zero across the entire phase space above 100 GeV,

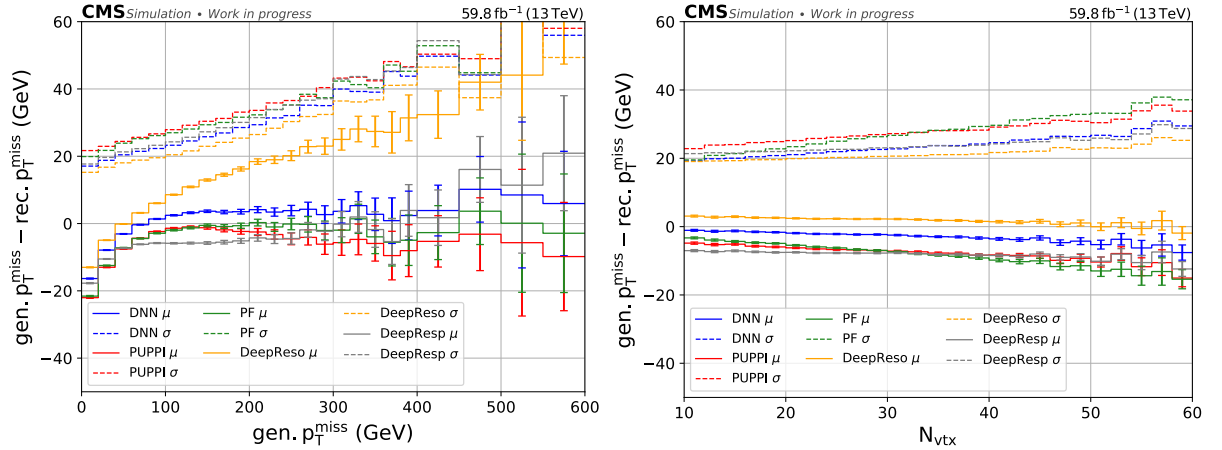


Figure 7.10: Response (solid lines) and resolution (dashed lines) of different p_T^{miss} reconstructions as a function of the gen. p_T^{miss} (left) and the number of vertices (right). The results are based on the 2018 data-taking period using the $t\bar{t}$ test sample and shown for DNN (blue), PUPPI (red) and PF (green) p_T^{miss} as well as the resolution (orange) and response tune (grey) of the Deep p_T^{miss} reconstruction. The error bars correspond to the statistical uncertainty on the mean.

the resolution tune shows a significant bias in the response which reaches up to 40 GeV in the most extreme phase space. Comparing the resolution of the four remaining results shows that the DNN p_T^{miss} outperforms the other reconstructions across the entire gen. p_T^{miss} range indicating a stable performance for different phase space regions.

The resolution and response of p_T^{miss} as a function of additional vertices in an event are shown in Fig. 7.10 as well. Here, the positive effects of the pileup-related corrections used in the PUPPI reconstruction, as discussed in Sec. 6.2.6, become visible. The response and resolution of PUPPI p_T^{miss} indicate a more stable performance as a function of the number of vertices compared to the PF reconstruction. This positive effect is maintained by the DNN correction, which yields a similar slope compared to the PUPPI p_T^{miss} with respect to the response and resolution, while still improving the resolution across the studied range.

The response achieved by the DNN regression as a function of $\min[\Delta\phi(p_T^{\text{miss}}, \ell)]$, which presents one of the main motivations for the additional correction, as discussed in Sec. 7.1, is shown in Fig. 7.11. Compared to the PUPPI and PF reconstruction, the response is significantly shifted towards unity by the DNN regression taking into account the full event selection as well as events with a rec. p_T^{miss} above 120 GeV. The decrease in response at larger angular distances, which is most prominent in the PUPPI reconstruction, is still present after the DNN correction but partially reduced.

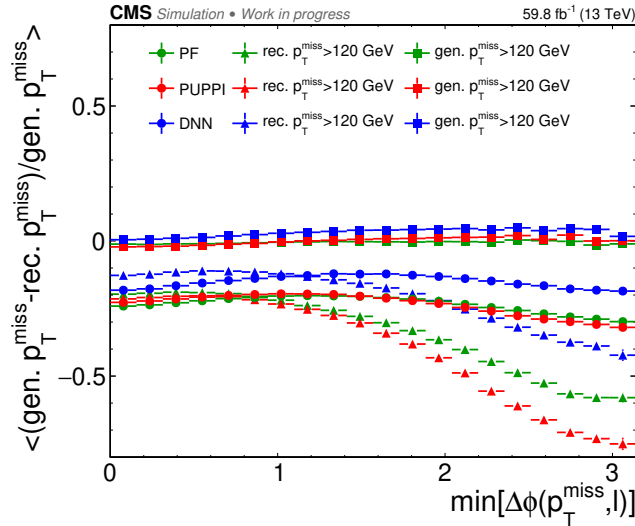


Figure 7.11: Mean relative p_T^{miss} response as a function of the minimal azimuthal distance between rec. p_T^{miss} and a selected lepton in simulated dileptonic $t\bar{t}$ events for the 2018 data-taking period. The results for the PF (green) and the PUPPI (red) reconstruction as well as for the DNN correction (blue) are compared for all selected events (circles), events with a rec. p_T^{miss} larger than 120 GeV (triangles) and events with a gen. p_T^{miss} larger than 120 GeV (squares).

For the final cross section measurements, the DNN regression is applied to data and does therefore not only affect dileptonic $t\bar{t}$ events but also SM background events and potential BSM contributions. To exclude that the regression introduces any additional bias for contributions other than the $t\bar{t}$ signal events, its performance is also studied for other simulated samples as shown in Fig. 7.12. For the performance on single top events similar conclusions to the study on $t\bar{t}$ signal events, where the DNN regression significantly improves the resolution and response compared to the PUPPI and PF reconstruction, can be drawn. This behavior is expected since

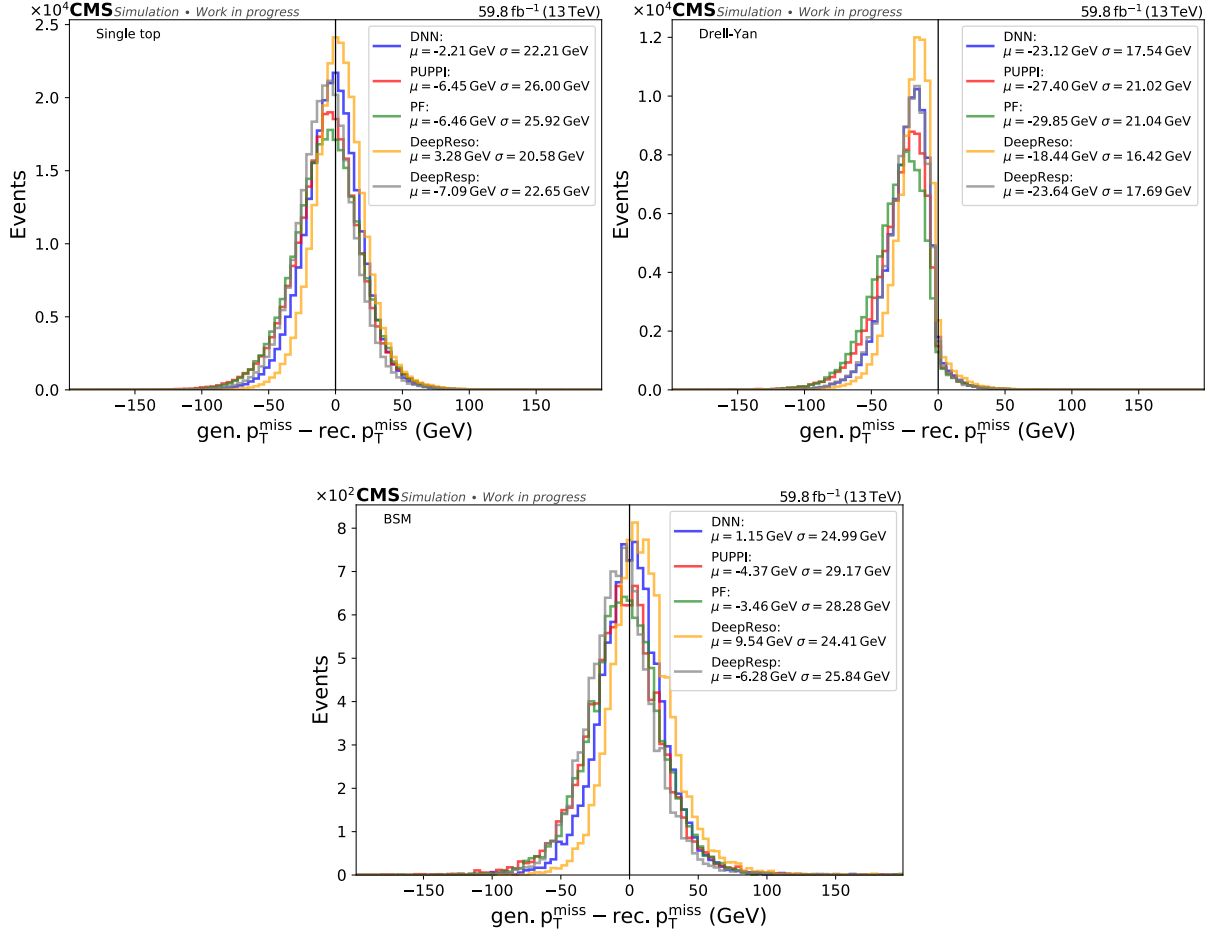


Figure 7.12: Difference between generated and reconstructed missing transverse momentum with respect to the magnitude comparing DNN (blue), PUPPI (red) and PF (green) p_T^{miss} as well as the resolution (orange) and response tune (grey) of the Deep p_T^{miss} reconstruction. The results are derived based on simulation of the single top process (top left), the DY process (top right) and the stop-pair production (bottom). The mean and the standard deviation of each distribution are listed in the legend.

the single top signature does not differ too much from the signature of the $t\bar{t}$ signal events. For the DY events, in contrast, no genuine p_T^{miss} is expected, yielding a different signature compared to the signal events. However, also for the DY events similar improvements with respect to the response and resolution can be observed after the DNN regression. Apart from studying the performance on SM events, the DNN regression is also tested on stop-pair production events, where additional contributions to p_T^{miss} arise from the neutralinos in the final state (cf. Sec. 2.3.3). Here, a similar picture to the previous studies, where the DNN regression yields a significant improvement, can be observed, while the bias introduced by the Deep p_T^{miss} resolution tune becomes even more evident. In summary, the DNN seems to learn to correct the detector resolution itself rather than relying on a particular event topology, which makes it suitable for application in this analysis.

Chapter 8

Cross section extraction

8.1	Fiducial phase space	103
8.2	Binning optimization	104
8.2.1	Binning requirements	104
8.2.2	One-dimensional binning	106
8.2.3	Two-dimensional binning	107
8.3	Unfolding procedure	110
8.3.1	Signal extraction	110
8.3.2	Unfolding with TUnfold	111
8.4	Validation tests	113

In case no potential contributions from BSM processes are found in the measurements performed on detector level, this analysis aims to derive differential cross sections of the dileptonic $t\bar{t}$ process. The differential cross section measurements are performed as a function of the transverse momentum $p_T^{\nu\nu}$ of the dineutrino system and the smallest azimuthal distance $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ between the dineutrino system and a lepton based on a fiducial phase space (particle-level) defined in Sec. 8.1. To maximize the number of measuring points, which can be compared to different theory predictions and MC simulations, the binnings used for the cross section measurements are optimized by dedicated algorithms, as discussed in Sec. 8.2. The cross sections are derived based on an unfolding procedure aiming to correct for detector effects and for intrinsic differences between the detector-level and particle-level observables, which blur the true underlying spectrum and affect the acceptance and efficiency of the reconstruction and selection. This procedure as well as the subtraction of SM contributions other than the dileptonic $t\bar{t}$ contribution are introduced in Sec. 8.3. In the final part of this chapter, the unfolding procedure is validated with respect to its statistical properties and biases introduced by potential differences between the true distribution present in data and in simulation.

8.1 Fiducial phase space

The most intuitive approach for a cross section measurement usually corresponds to a measurement using the full phase space of the process of interest, to enable a proper comparison to fixed-order theory predictions. In the context of this analysis, such an approach would correspond to a cross section measurement of the full (dileptonic) $t\bar{t}$ process. However, such measurements strongly rely on the extrapolation from the phase space accessible by the detector to the full

phase space, which can introduce large model dependencies. To avoid this kind of extrapolation, the cross section measurements, obtained in the context of this thesis, are derived for a fiducial phase space, also known as particle level, mainly following the definition discussed in [217, 218]. This particle-level definition is based on stable particles originating from the decay of the (anti-)top quark and subsequent decays without including the simulation of the detector effects. This enables the measurement of differential cross sections as a function of the kinematics of the final state particles, as aimed by this analysis.

In detail, the following requirements are used to reconstruct particle-level objects. Generated particles, considered in the particle-level definition, have to be within the general detector acceptance of $|\eta| < 5.0$. Charged leptons and neutrinos have to be prompt, excluding those originating from hadron decays. To take additional photon radiation into account, the charged leptons are "dressed" by including photons within $\Delta R < 0.1$ around the lepton direction into the corresponding charged lepton momentum. In addition, charged leptons have to fulfill $p_T > 15$ GeV as well as $|\eta| < 2.5$ to match the acceptance of the tracking system. Particle-level jets are clustered with the anti- k_T algorithm using a distance parameter of $R = 0.4$. Non-prompt neutrinos originating from hadron decays are included in the jet clustering. The clustered jets have to be separated from the selected charged leptons by $\Delta R > 0.4$ and have to fulfill $p_T > 30$ GeV and $|\eta| < 2.4$. Heavy-flavor jets are identified based on including decayed heavy hadrons, also known as ghost hadrons, in the jet clustering, but reducing their momentum to be negligible. If a ghost hadron is clustered into a jet, the jet is assigned the flavor of the ghost hadron.

Based on these particle-level objects, the fiducial phase space is defined aiming for the expected signature of dileptonic $t\bar{t}$ events, i.e., the presence of leptons, neutrinos and jets. Events have to contain two prompt oppositely-charged leptons (electrons or muons) with $p_T > 20$ GeV and two jets with $p_T > 30$ GeV containing ghost b hadrons, where leptons and jets are within $|\eta| < 2.4$. In addition, the invariant mass of the two leptons has to exceed 20 GeV. For same-flavor events, the transverse momentum of the dineutrino system has to be larger than 40 GeV matching the p_T^{miss} requirement employed on detector level (cf. Tab. 6.4). Finally, events, where at least one of the charged leptons originates from an intermediate τ lepton, are vetoed.

8.2 Binning optimization

For each of the three differential cross section measurements, a reasonable binning has to be defined. In order to avoid any bias potentially introduced by manually defined binnings and to optimize the binnings, this task is performed by a mostly automatized procedure. Here, an optimized binning is considered to have a maximal number of bins while fulfilling a set of requirements introduced in the following. With this procedure, the number of points to compare with theory predictions can be maximized while keeping a stable unfolding procedure (cf. Sec. 8.3) and a reasonable statistical precision. The optimization discussed in the following is performed using the nominal $t\bar{t}$ simulation for the 2018 data-taking period. In order to keep a uniform approach for each period, the binning results are also applied for the remaining periods.

8.2.1 Binning requirements

In the procedure used for the binning optimization, a set of requirements is considered. These requirements aim to ensure proper conditions for the unfolding process and a large $t\bar{t}$ signal contribution across the entire measured spectrum. In addition, the detector resolution with respect to the observable of interest as well as the statistical precision of the final measurement are taken into account. In detail, the set of requirements includes the following parts:

Migration effects

The finite detector resolution can introduce event migrations between bins on detector and particle level. Such event migrations are, to some degree, corrected by the unfolding procedure introduced in Sec. 8.3. However, to keep the impact of the unfolding as small as possible and to avoid a potentially unstable unfolding, an optimized binning should keep the event migrations for each individual bin at a reasonable level. In this analysis, the migrations are quantified for each bin i by the purity P_i and the stability S_i defined as

$$P_i = \frac{N(\text{gen}_i \& \text{rec}_i)}{N(\text{rec}_i)} \quad S_i = \frac{N(\text{gen}_i \& \text{rec}_i)}{N(\text{gen}_i)}, \quad (8.1)$$

where N denotes the number of events per particle-level bin i , which fulfill the particle-level and detector-level selection, distinguishing between events, where the generated (gen_i), the reconstructed (rec_i) or both observables ($\text{gen}_i \& \text{rec}_i$) match the range of bin i . Thus, $1 - P_i$ is a measure of events migrating into a bin, while $1 - S_i$ is a measure of events migrating out of a bin. To ensure a reasonable level of migrations in the final binnings, a minimal requirement on the purity and stability of 0.5 (0.3) is enforced for each bin of the optimized binnings used for the 1D (2D) measurements.

Detector resolution

In addition to the requirements regarding stability and purity, the detector resolution is directly taken into account to reduce the probability of large event migrations. Here, the detector resolution in each particle-level bin is derived based on a Gaussian fit to the difference between the generated and reconstructed quantity. The width of the Gaussian is taken as an estimate of the detector resolution in a given bin with respect to the observable of interest. In the final binning, the width of each bin is required to exceed the corresponding detector resolution.

Signal efficiency

To ensure that each particle-level bin of the final binning is represented by a reasonable fraction of signal events selected on detector level, each bin has to fulfill a minimal requirement regarding the signal efficiency defined as:

$$\epsilon_i = \frac{N'(\text{gen}_i \& \text{rec})}{N'(\text{gen}_i)}, \quad (8.2)$$

where N' denotes the number of events per particle-level bin i , which fulfill the particle-level selection, distinguishing between all events in bin i (gen_i) and those, which also fulfill the detector-level selection ($\text{gen}_i \& \text{rec}$). For all three measurements, a minimal signal efficiency of 20% is required for each particle-level bin.

Statistical precision

The requirements introduced above can usually be satisfied by small bins, which only hold a few events. Such binnings would introduce large statistical uncertainties on the final cross section measurements. To avoid such outcomes, a minimal requirement regarding the number of

expected events per bin has to be satisfied by each bin of the final binning. Taking into account that the integrated luminosity of the 2018 data-taking period accounts for more than 40% of the total integrated luminosity (cf. Tab. 5.1), a minimum of 50 (10) events per bin is required for the binnings used in the 1D (2D) measurements.

8.2.2 One-dimensional binning

The general concept of the binning optimization is based on the idea of starting with a fine binning and merging bins until the set of requirements defined above is fulfilled. This procedure is sketched in Fig. 8.1 and, in detail, contains the following steps. The measuring range of interest is divided into small equidistant bins with a width of 2 GeV and 0.04 for the $p_T^{\nu\nu}$ and $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ measurements, respectively. The set of requirements is checked for the lowest bin of the selected range, which usually does not fulfill the requirements. In this case, the current bin is merged with its higher neighboring bin and the requirements are checked again. If the requirements fail again, the procedure is repeated until the current bin matches the set of requirements, after which the next higher bin is taken into account. This logic is repeated until the highest bin of the range of interest is reached. If this bin fulfills the requirements, the procedure stops and an optimized binning is found. Otherwise, the merging procedure introduced above is employed again but the merging steps are now performed with the lower neighboring bin until all bins match the requirements.

For the $p_T^{\nu\nu}$ measurement, the procedure additionally accounts for the $p_T^{\nu\nu}$ requirement included in the definition of the fiducial phase space (cf. Sec. 8.1) by enforcing a bin edge at $p_T^{\nu\nu} = 40$ GeV.

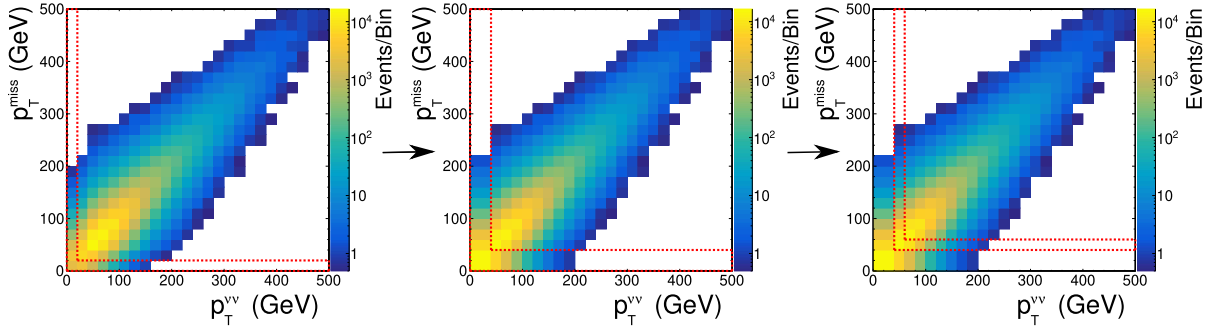


Figure 8.1: Sketch of the different steps used in the 1D binning optimization. The bins on the x - and y -axis correspond to the particle-level and detector-level bins, respectively. The dashed red line indicates the current bin. On the left, the initial equidistant binning is shown with the lowest bin corresponding to the current bin, which, by assumption, does not match the requirements. In the middle, this bin is merged with the higher neighboring bin, which, by assumption, matches the requirements. On the right, the next higher bin is taken to be the current bin.

The results of the binning optimization for the $p_T^{\nu\nu}$ and $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ measurement can be found in Tab. 8.1. For both measurements, the optimization procedure is also performed using the PUPPI p_T^{miss} reconstruction for the detector-level observables to quantify the impact of the DNN p_T^{miss} regression (cf. Sec. 7) on the optimized binnings. By applying the additional DNN p_T^{miss} correction, the $p_T^{\nu\nu}$ measurement can be extended by three additional bins compared to relying on the PUPPI p_T^{miss} reconstruction. Given that fewer migrations are observed in the $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ measurement, the impact of the regression is smaller for the corresponding measurement resulting in one additional bin.

Table 8.1: Results of the binning optimization for the one-dimensional measurements using the corrected DNN p_T^{miss} and the nominal PUPPI p_T^{miss} . The highest bin includes all events exceeding the last bin edge.

Detector level	Total bins	Bin edges (GeV)									
DNN p_T^{miss}	9	0	40	68	100	140	196	260	332	410	
PUPPI p_T^{miss}	6	0	40	72	112	196	340	-	-	-	
DNN $\min[\Delta\phi(p_T^{\text{miss}}, \ell)]$	6	0	0.4	0.88	1.36	1.84	2.36	-	-	-	
PUPPI $\min[\Delta\phi(p_T^{\text{miss}}, \ell)]$	5	0	0.44	0.96	1.52	2.08	-	-	-	-	

The purity, stability and signal efficiency for each individual bin of both optimized 1D binnings are shown in Fig. 8.2. For both measurements, an almost flat signal efficiency of around 35-40% can be achieved. For the $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ measurement, only a slight dependency between both migration measures and the particle-level bins is observed, while for the $p_T^{\nu\nu}$ measurement, the first and the last bin of the range show increased stability and purity. For the latter observable, a few bins in the middle of the spectrum show purities slightly below the required value of 0.5. This small discrepancy is caused by an update of the DNN training (cf. Sec. 7) performed after the binning optimization. However, given that the requirements for the purity were chosen rather conservatively, these differences should have no significant impact on the unfolding procedure.

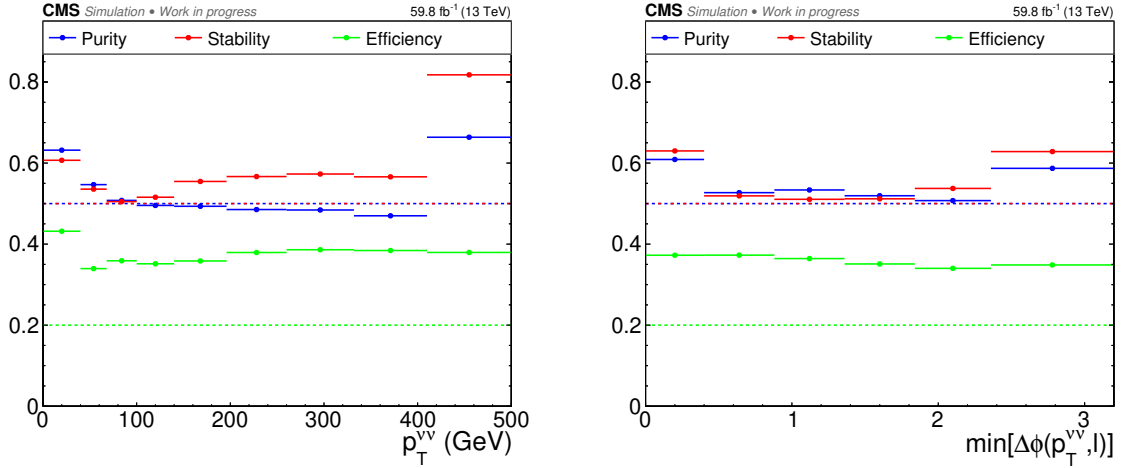


Figure 8.2: Purity (blue), stability (red) and signal efficiency (green) for the optimized binnings used in the $p_T^{\nu\nu}$ (left) and $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ (right) measurement. The dashed green and red-blue lines indicate the minimal requirement for the efficiency and the stability/purity, respectively.

8.2.3 Two-dimensional binning

The algorithm employed for the optimization of the two-dimensional binning is based on the same concept as used for the one-dimensional case. However, in the 2D case, the choice of the higher neighboring bin used in the merging step is not unique anymore. Thus, two additional decision steps, as visualized in Fig. 8.3, have to be added to the optimization procedure for the 2D measurements. The first of these additional steps corresponds to the decision which

neighboring bin should be merged if the current bin does not fulfill the set of requirements. The second additional step takes care of selecting a new current bin once a bin fulfills all requirements.

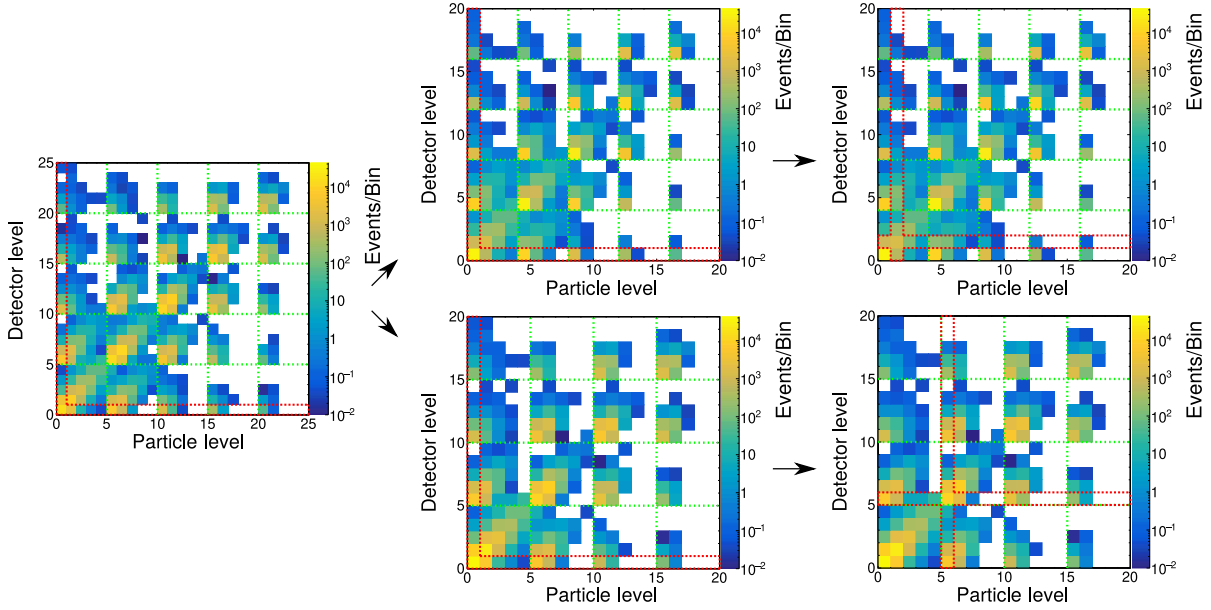


Figure 8.3: Sketch of the steps employed in the 2D binning optimization. The bins on the x - and y -axis correspond to the particle-level and detector-level bins, respectively. The dashed green lines indicate bins of the second dimension while each green block contains several bins of the first dimension. The dashed red line indicates the current bin. On the left, the initial equidistant binning is shown with the lowest bin corresponding to the current bin, which, by assumption, does not match the requirements. The plots in the middle, show the merging with a bin of the first dimension (top) and the second dimension (bottom). The plots on the right show the two options for selecting the next current bin.

The decision logic employed for a merging of two bins depends on which requirement is not fulfilled by the current bin. If the current bin fails the requirement regarding the detector resolution with respect to one of the two observables, it is merged with the higher neighboring bin of the corresponding dimension. If the detector-resolution requirement for both observables is not fulfilled, the bin to merge with is chosen in an arbitrary way¹. If any of the three remaining requirements is not fulfilled by the current bin, a so-called measure of improvement (MOI) for both merging candidates is derived. While different definitions of the MOI have been studied, the definition which yields the final binning used in the 2D measurements is given by:

$$\text{MOI} = \prod_i \left[1 + \underbrace{\left(\frac{(x_i^{\text{new}} - x_i^{\text{old}})(x_i^{\text{min}} - x_i^{\text{old}})}{x_i^{\text{min}} \cdot x_i^{\text{min}}} \right)}_{\Delta x} \right] \cdot \underbrace{\left(\frac{\Delta_{\text{new}} - \Delta_{\text{old}}}{\Delta_{\text{range}}} \right)^{-1}}_{\Delta y}, \quad (8.3)$$

where i loops over the quantities considered by the different requirements, only taking quantities into account which are improved by the given merging step. The values of the quantities before

¹Arbitrary way denotes a decision based on the number of merging steps already performed for the current bin. An even (odd) number leads to a merge with the bin of the first (second) dimension ensuring alternation between both dimensions.

and after the merging step are denoted as x_i^{old} and x_i^{new} , respectively, while the required value is given by x_i^{min} . The first part of the product, given by Δx , estimates the relative improvement of a given quantity relative to its required value. The second part, given by Δy , corresponds to an additional penalty term for merger with large bins, by comparing the difference between the old bin width Δ_{old} and new bin width Δ_{new} to the total range Δ_{range} considered in the measurement. This additional penalty term aims to avoid binnings, which yield only one large bin for one of two dimensions. The merging candidate, which results in a larger MOI, is the one used in the merging step. If both merging candidates fail to improve any of the quantities considered in the requirements, the merging direction is again chosen in an arbitrary way.

Once the optimization procedure is finished, the optimized bins have to be tested again with respect to the requirements. This additional step is only necessary for the 2D optimization, since here, optimized bins can be still altered by the following steps of the optimization procedure.

The optimization for the 2D measurements starts based on an initial binning with a bin width of 5 GeV and 0.04 in the $p_T^{\nu\nu}$ and $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ direction, respectively. The optimization is performed multiple times using different MOI definitions and different options regarding the selection of the next current bin (always use the first dimension or always use the second dimension) as well as regarding the general direction of the procedure. In the latter case, the procedure is not only performed starting with the lowest bin of each dimension but also starting with the highest bin and considering the lower neighboring bins as merging candidates. The final binning used for the 2D measurements, shown in Tab. 8.2, is based on the MOI defined in Eq. (8.3), starting with the highest bin of both dimensions and always selecting the bin in the $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ direction as the next current bin. As already observed in the optimization of the 1D binning, the DNN p_T^{miss} corrections allow for additional bins to be used in the measurements. For the 2D measurements, these additional bins are rather crucial since they allow for additional measuring points in the region of main interest at large momentum and large angular separation as discussed in Sec. 2.3.3.

Table 8.2: Results of the binning optimization for the 2D measurements using the corrected DNN p_T^{miss} and the nominal PUPPI p_T^{miss} . The highest bin includes all events exceeding the last bin edge.

Detector level	Total bins	Bin edges						
DNN p_T^{miss}	12	$p_T^{\nu\nu}$ (GeV)	0	70	125	200	-	
		$\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$	0	0.56	1.08	-	-	
PUPPI p_T^{miss}	10	$p_T^{\nu\nu}$ (GeV)	0	45	70	125	180	
		$\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$	0	0.76	-	-	-	

The purity, stability and signal efficiency for each individual bin of the optimized 2D binning are shown in Fig. 8.4. Similar to the 1D measurements, a rather flat signal efficiency of around 35 to 40% is observed for most of the bins. Only the highest bins at large angular separation show a slightly reduced efficiency. The stability and purity increase in the first two $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ bins with increasing momentum. In the highest $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ bin, this trend is reversed for the purity, which can be explained by the reduced p_T^{miss} resolution at large angular separation leading to additional event migrations, as discussed in Sec. 7.1. Again, not all optimized bins match the purity requirement of 0.3. Here, the same arguments, as discussed for the 1D measurements, hold.

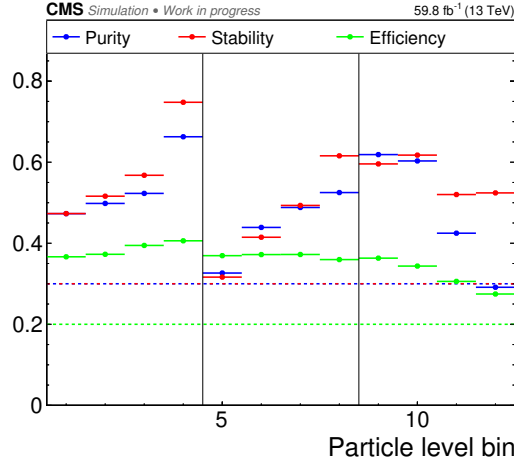


Figure 8.4: Purity (blue), stability (red), and signal efficiency (green) for the optimized binning used in the 2D measurements. The vertical black lines correspond to the bin edges in $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$, while inside each of these bins, several $p_T^{\nu\nu}$ bins are shown in increasing order. The dashed green and red-blue lines indicate the minimal requirement for the efficiency and the stability/purity, respectively.

8.3 Unfolding procedure

Based on the detector-level distributions extracted from the observed events fulfilling the full detector-level selection (cf. Sec. 6.3), differential cross section results are derived on particle level using an unfolding procedure introduced in the following. This procedure can be divided into two main steps. In the first step, the yield of dileptonic $t\bar{t}$ events in each bin of the detector-level distributions is derived based on the observed number of events and estimations of the contributions from other SM processes. In the second step, the impact of the detector acceptance and resolution is corrected, which corresponds to the actual unfolding of the detector-level distributions to the particle level including the transition from the detector-level observables p_T^{miss} and $\min[\Delta\phi(p_T^{\text{miss}}, \ell)]$ to the particle-level observables $p_T^{\nu\nu}$ and $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$. For the unfolding, a variety of different methods are available [219] out of which a least square method implemented in the TUnfold package [220] is used in this analysis.

The steps of the unfolding procedure are performed separately for each data-taking period combining the three lepton channels to achieve sufficient statistical precision. For this purpose, all systematic uncertainties, introduced in Sec. 9, are treated as fully correlated between the different lepton channels within each data-taking period. The measurements of the individual data-taking periods are combined after the unfolding procedure taking correlations between systematic uncertainties into account, to derive the final results based on all four data-taking periods.

8.3.1 Signal extraction

The number of dileptonic $t\bar{t}$ signal events in each bin of the detector-level distributions is derived prior to the unfolding step by removing contributions from all other SM processes. Here, two different methods are applied differing between SM background contributions originating from

the $t\bar{t}$ process, introduced as $t\bar{t}$ other² in Sec. 5.2.4, and all other SM background processes. In detail, the final number of signal events per bin i is given by

$$N_{i,\text{data}}^{\text{sig}} = (N_{i,\text{data}} - N_{i,\text{MC}}^{\text{bkg}}) \cdot f_i^{\text{sig}}, \quad (8.4)$$

where $N_{i,\text{data}}$ denotes the total number of events observed in data and $N_{i,\text{MC}}^{\text{bkg}}$ is the estimated contribution from SM background processes derived from simulation, which are not part of the $t\bar{t}$ other background. The signal fraction f_i^{sig} , defined as

$$f_i^{\text{sig}} = \frac{N_{i,\text{MC}}^{\text{sig}}}{N_{i,\text{MC}}^{\text{sig}} + N_{i,\text{MC}}^{\text{t}\bar{t}\text{ other}}}, \quad (8.5)$$

takes the contribution from $t\bar{t}$ other into account. Here, $N_{i,\text{MC}}^{\text{sig}}$ and $N_{i,\text{MC}}^{\text{t}\bar{t}\text{ other}}$ denote the number of signal events and the contribution from $t\bar{t}$ other events, respectively. Both quantities are estimated from simulation. By applying the signal fraction instead of subtracting the estimated $t\bar{t}$ other background yield, the same scale in normalization as observed for the dileptonic $t\bar{t}$ signal process is applied to the $t\bar{t}$ other background contribution. This strategy takes into account that both of these contributions originate from the same $t\bar{t}$ production processes.

8.3.2 Unfolding with TUnfold

The differential and multi-differential cross sections derived in the scope of this analysis are defined based on the number of unfolded signal events $\hat{M}_j^{\text{unf}}$ per particle-level bin j , obtained after the signal extraction using the unfolding method introduced below. The detailed definition of the absolute differential cross sections is given by

$$\frac{d\sigma_j}{dx} = \frac{1}{\Delta x_j} \frac{\hat{M}_j^{\text{unf}}}{\mathcal{L}}, \quad (8.6)$$

where $\mathcal{L}$ denotes the integrated luminosity of a given data set and Δx_j corresponds to the bin width with respect to the observable of interest x . For the two-dimensional measurements, the division by the bin width is always performed based on the $p_T^{\nu\nu}$ observable. In addition to the absolute cross sections, normalized results are obtained by dividing Eq. (8.6) by the total fiducial cross section σ , derived by integrating over the entire measured spectrum of interest.

The problem to be solved in the unfolding step, which aims to correct for differences between the reconstructed (detector-level) and generated (particle-level) quantities caused by mismeasurements in the reconstruction as well as limited detector acceptance and intrinsic differences between both quantities, e.g., the difference between the generated p_T^{miss} and $p_T^{\nu\nu}$ (cf. Sec. 6.2.6), can be written as

$$M_i^{\text{sig}} = \sum_{j=1}^m A_{ij} M_j^{\text{true}}, \quad 1 \leq i \leq n. \quad (8.7)$$

²In the context of the unfolding procedure, $t\bar{t}$ other also includes dileptonic $t\bar{t}$ events, which are selected by the detector-level selection, but originate outside of the fiducial phase space defined in Sec. 8.1.

Here, M_i^{sig} denotes the expected number of signal events on detector level in bin i and M_j^{true} corresponds to the expected number of signal events on particle level in bin j . The total number of detector- and particle-level bins is given by n and m , respectively. The elements of the response matrix A_{ij} describe the migrations from the j th bin on particle level to the i th bin on detector level considering the detector acceptance and selection efficiency. The ranges covered by the bins on detector- and particle-level are the same. However, for the measurements performed in this analysis, most of the particle-level bins are split into two bins with equal width to obtain the detector-level bins. The only exception to this procedure is given by the highest $p_T^{\nu\nu}$ bins, which are the same for both levels to avoid large statistical fluctuations. This condition, corresponding to $n > m$, usually allows for a more stable operation by overconstraining the unfolding problem.

In this analysis, the problem defined in Eq. (8.7), is solved based on a least square method implemented in the TUnfold package [220]. The result of this solution is an estimate of the true number of signal events in each particle-level bin, which serves as $\hat{M}_j^{\text{unf}}$ in the definition of the differential cross sections (cf. Eq. (8.6)). The method is based on the following Lagrangian:

$$\begin{aligned} \mathcal{L} = & \left(\vec{N}^{\text{sig}} - \mathbf{A} \vec{M}^{\text{unf}} \right)^T \mathbf{V}^{-1} \left(\vec{N}^{\text{sig}} - \mathbf{A} \vec{M}^{\text{unf}} \right) \\ & + \tau^2 \left(\vec{M}^{\text{unf}} - \vec{M}^{\text{true}} \right)^T \left(\mathbf{L}^T \mathbf{L} \right) \left(\vec{M}^{\text{unf}} - \vec{M}^{\text{true}} \right) \\ & + \lambda \left(Y - \vec{e}^T \vec{M}^{\text{unf}} \right) \end{aligned} \quad (8.8)$$

with $Y = \sum_i N_i^{\text{sig}}$ and $e_i = \sum_j A_{ij}$.

The first term of the Lagrangian corresponds to the least square problem, where $\vec{N}^{\text{sig}}$ denotes the observed number of signal events on detector level after the signal extraction, with the corresponding diagonal covariance matrix $\mathbf{V}$ encoding the statistical uncertainties. The response matrix is denoted as $\mathbf{A}$ and is, in the context of this analysis, derived from the nominal $t\bar{t}$ signal simulation. The second term introduces an optional regularization (reg.) controlled by the reg. parameter τ and the reg. condition $\mathbf{L}$ aiming to reduce potential fluctuations in the unfolded results introduced by statistical fluctuations in the observed number of signal events. For the present measurements, the second derivative is used as reg. condition while the reg. parameter is obtained by minimizing the average global correlation as defined in [220]. The third term of the Lagrangian introduces an additional so-called area constraint, controlled by a Lagrangian parameter λ . This constraint aims to match the total number of observed signal events, corrected by acceptance and efficiency effects, to the total number of unfolded events. The final estimate $\hat{M}_j^{\text{unf}}$ as well as the corresponding covariance matrix of the number of unfolded signal events are extracted by minimizing the Lagrangian with respect to the observed number of signal events $\vec{N}^{\text{sig}}$.

The response matrices for the three individual measurements based on the 2018 data-taking period are shown in Fig. 8.5. For both one-dimensional measurements, large diagonal elements of around 0.5 to 0.6 can be observed indicating a good resolution for both observables. The off-diagonal elements on both sides of the diagonal show rather significant differences for both 1D measurements. For most parts of the measured phase space, the elements above the diagonal are larger caused by a small bias towards larger reconstructed values with respect to the generated values. In case of the $p_T^{\nu\nu}$ measurement, this behavior has been already observed in the study of the DNN p_T^{miss} response in Fig. 7.10. However, only the elements close to the diagonal are affected indicating migrations to neighboring bins but no migrations across large

phase space regions. For the 2D measurements, similar effects can be observed with respect to the migrations across p_T bins, which, compared to the migrations across angular bins, dominate most parts of the phase space in terms of migrations. The migrations are maximal in the phase space of large p_T and large angular separation, caused by the decreasing p_T^{miss} resolution in this phase space, as already discussed in Sec. 7.1 and 8.2.

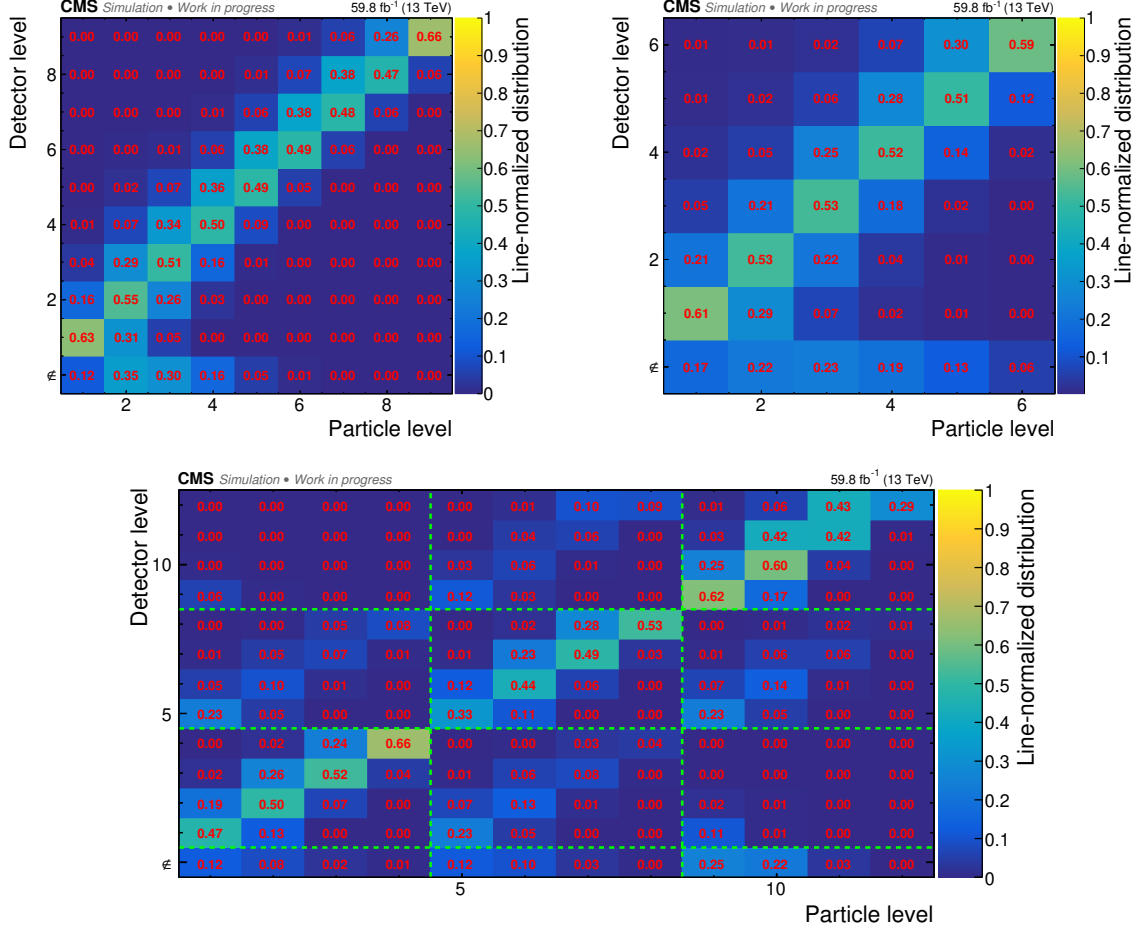


Figure 8.5: Response matrices for the $p_T^{\nu\nu}$ (top left), the $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ (top right) and 2D measurements (bottom) based on the 2018 data-taking period. Each entry is normalized by the integral of the corresponding horizontal line leading to diagonal elements, which show the purity. The bins on the x - and y -axis correspond to the particle-level and detector-level bins, respectively. Bins marked with $\notin$ correspond to events, which originate from the fiducial phase space but do not fulfill the detector-level selection. The dashed green lines in the response matrix for the 2D measurements indicate $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ bins, which each include several $p_T^{\nu\nu}$ bins.

8.4 Validation tests

To evaluate the validity of the unfolding procedure and to decide on the detailed approach employed in the unfolding method, dedicated validation tests are performed. In these tests, three different unfolding approaches are compared. The first approach, denoted as nominal approach, corresponds to the least square method expressed in Eq. (8.8) without including the optional regularization ($\tau = 0$). For the second approach, denoted as regularized approach, the

regularization is enabled. Apart from reducing statistical fluctuations in the final result, this additional regularization can also introduce a potential bias if the regularization is too strong. The last approach corresponds to a naive bin-by-bin (BBB) unfolding, where only the diagonal elements of the response matrices, shown in Fig. 8.5, are considered, neglecting any migration effects. For this approach, a simple correction factor is derived for each particle-level bin and the detector-level bins are chosen to match the particle-level bins.

Truth test

The truth test aims to check the setup used for the unfolding procedure. Here, the nominal response matrices based on the dileptonic $t\bar{t}$ signal sample are applied to pseudo data derived from the same sample. Thus, by construction, the migration and acceptance effects, present in the pseudo data and described by the response matrix, are the same. A valid unfolding setup should therefore yield an unfolded result, which exactly matches the particle-level distribution predicted by the dileptonic $t\bar{t}$ signal sample.

As an example of the outcome of the truth test, the results for the $p_T^{\nu\nu}$ measurement are shown in Fig. 8.6. All unfolding approaches yield unfolded distributions, which exactly match the expected distribution indicating valid unfolding setups. The same observations can be made based on the results for the $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ and the 2D measurement, which are both shown in Appx. 9.

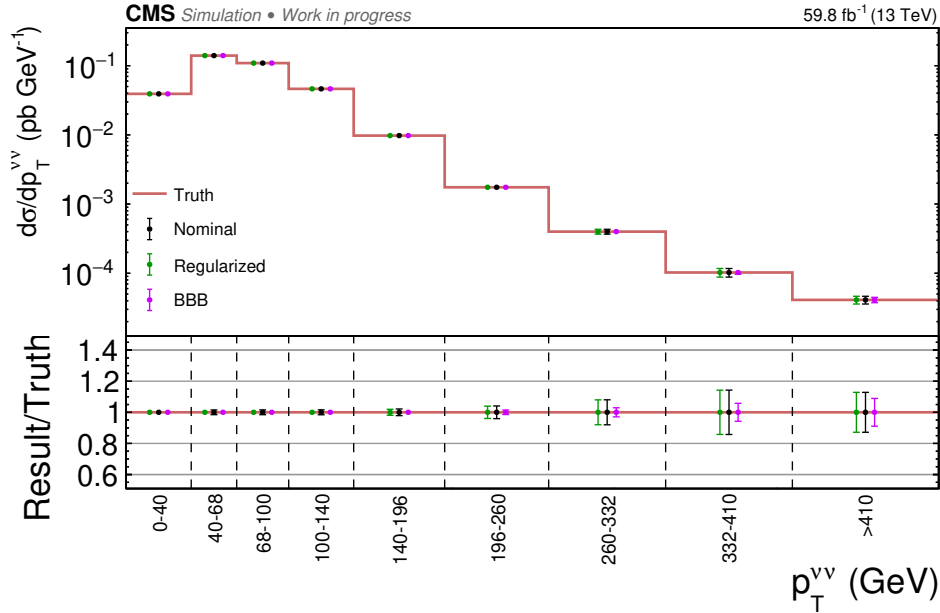


Figure 8.6: Result of the truth test performed for the $p_T^{\nu\nu}$ measurement derived with the nominal (black), the regularized (green) and the bin-by-bin unfolding (purple) based on the 2018 data-taking period. The unfolded distributions are compared to the expected distribution (red) based on the dileptonic $t\bar{t}$ signal sample. The error bars correspond to the statistical uncertainty. The lower part shows the ratios of the unfolded and the expected distributions.

Toy test

To validate the statistical properties of the unfolding procedures, tests based on toy data sets drawn from the detector-level distribution of the $t\bar{t}$ signal sample are performed. The toy data

sets, which are used as pseudo data, are generated by smearing the yield in each detector-level bin based on Poisson statistics. For each toy data set, the unfolding procedure is performed and the unfolded distribution is compared to the non-smearing result, derived from the truth test introduced above, taking the covariance from the non-smearing result into account. Here, the χ^2 results between smeared and non-smearing unfolded distributions as well as the pulls between both types of distributions for each individual bin are studied.

The results of the toy test for the $p_T^{\nu\nu}$ measurement using a total of 5000 toys are shown in Fig. 8.7. The distributions of the χ^2 results for each of the three approaches fit the expected χ^2 function, where the degrees of freedom match the number of bins used in the corresponding measurement. The distributions of the pulls, where the pull is derived individually for each bin of the particle-level distribution, show the expected behavior and follow a Gaussian function with a mean of zero and a width of unity. The toy tests performed for the remaining measurements, which can be found in Appx. 9, show similar results, indicating valid statistical properties of all three unfolding approaches.

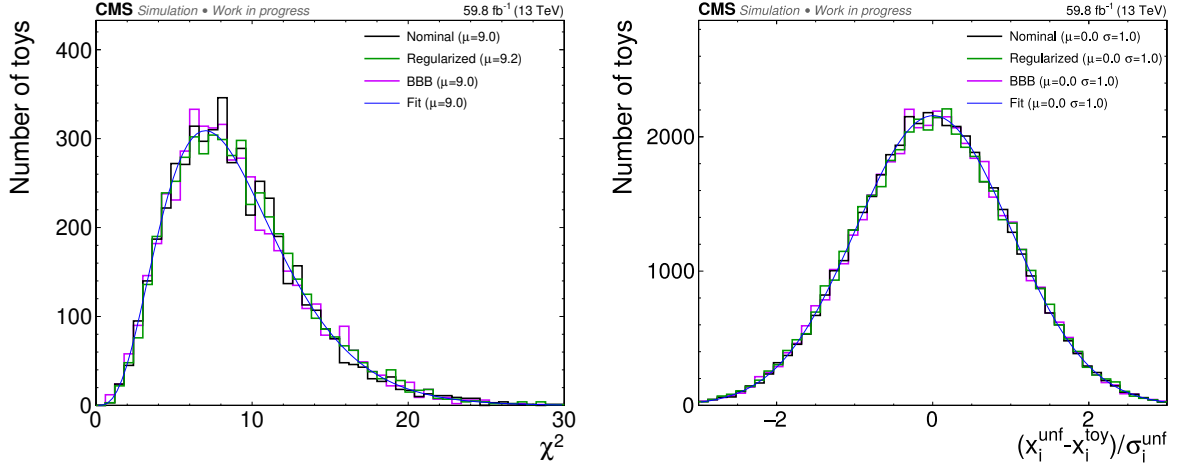


Figure 8.7: Results of the toy test performed for the $p_T^{\nu\nu}$ measurement derived with the nominal (black), the regularized (green) and the bin-by-bin unfolding (purple) based on the 2018 data-taking period. The left plot shows the distribution of the χ^2 values derived for each individual toy data set fitted by a χ^2 function (blue). The right plot shows the pull derived for each bin in each toy data set fitted by a Gaussian function (blue). Both fits are performed based on the nominal result.

Reweighting test

In the truth test as well as in the toy test, the particle-level distributions used to derive the response matrix and the underlying true distribution of the pseudo data match. The reweighting test, in contrast, aims to study the potential bias introduced by a mismatch between both distributions, which corresponds to a more realistic scenario expected for the actual measurements performed on observed data. To mimic this scenario, the underlying particle-level distribution of the pseudo data is reweighted event-wise based on

$$w(x) = \max \left[(1 + (x - m) \cdot s)^2, 0.1 \right], \quad (8.9)$$

where x denotes the particle-level observable and m and s control the nature of the reweighting. The altered pseudo data is unfolded with the nominal response matrix and compared to the expected reweighted particle-level distribution.

The results of the reweighting test performed for the $p_T^{\nu\nu}$ measurement with $s = 0.001$ and $m = 100$ GeV can be found in Fig. 8.8. All three unfolded distributions seem to follow the expected distribution. However, in the result of the simple bin-by-bin unfolding, a bias towards the particle-level distribution used in the response matrix is observed, which is also indicated by the large χ^2/dof . This bias is mainly introduced by ignoring migration effects in this approach. The remaining two approaches, which take the migrations into account, indicate better performance. However, especially at low $p_T^{\nu\nu}$, small fluctuations can be observed. These fluctuations are fully covered by the uncertainties in the nominal approach, while the χ^2/dof indicates a slightly worse agreement for the regularized approach. Similar conclusions can be drawn from the test results for the remaining two measurements shown in Appx. 9.

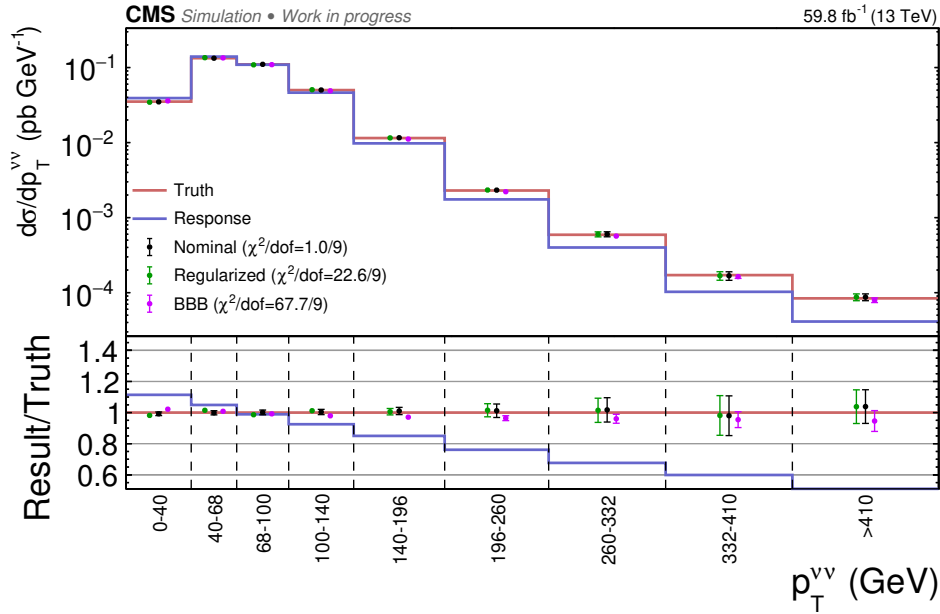


Figure 8.8: Result of the reweighting test performed for the $p_T^{\nu\nu}$ measurement derived with the nominal (black), the regularized (green) and the bin-by-bin unfolding (purple) based on the 2018 data-taking period. The unfolded distributions are compared to the expected distribution (red) based on the reweighted dileptonic $t\bar{t}$ sample with the corresponding χ^2/dof values given in the legend. The unweighted distribution, used for the response matrix, is shown in blue. The error bars correspond to the statistical uncertainty. The lower part shows the ratios of unfolded and expected distributions.

Alternative MC test

In addition to the reweighting test, the impact of a potential bias introduced by the unfolding procedure is studied by utilizing the alternative $t\bar{t}$ signal simulation (cf. Sec. 5.2.4). In this test, the pseudo data correspond to the detector-level distributions derived from the alternative signal simulation. These pseudo data are unfolded based on the response matrix derived from the nominal signal simulation, again mimicking the situation with potential differences between the measured data and the MC predictions.

The results of this test performed for the $p_T^{\nu\nu}$ measurement are shown in Fig. 8.9. Here, similar conclusions as already drawn in the reweighting test can be stated. The result derived with the bin-by-bin approach shows a small bias towards the distribution of the response matrix, especially at large $p_T^{\nu\nu}$, while the remaining two approaches yield a good agreement with the expected distribution despite small fluctuations around the expected distribution at large $p_T^{\nu\nu}$. The results for the remaining measurements, which confirm the observations discussed above, can be found in Appx. 9.

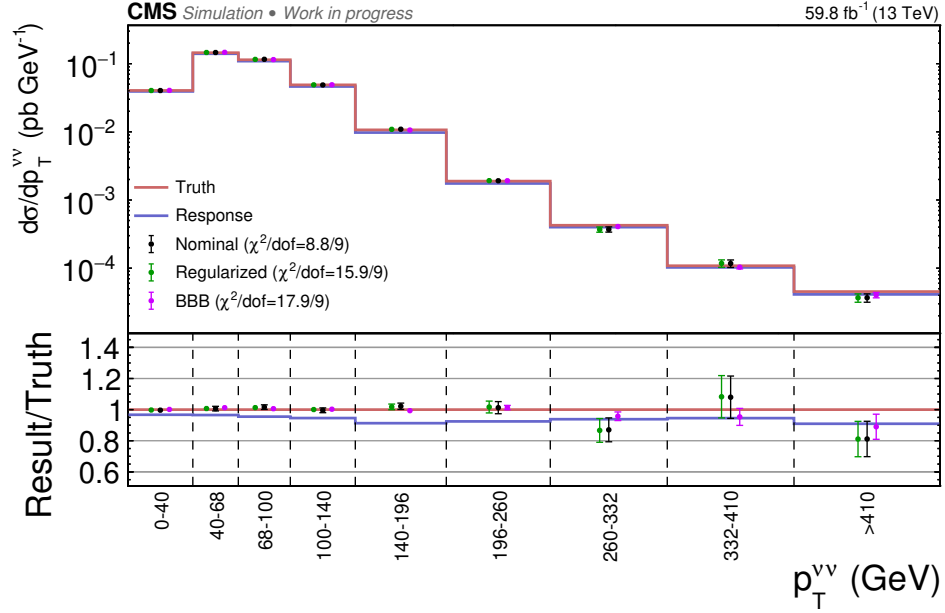


Figure 8.9: Result of the alternative MC test performed for the $p_T^{\nu\nu}$ measurement derived with the nominal (black), the regularized (green) and the bin-by-bin unfolding (purple) based on the 2018 data-taking period. The unfolded distributions are compared to the expected distribution (red) based on the alternative dileptonic $t\bar{t}$ signal sample with the corresponding χ^2/dof values given in the legend. The nominal distribution, used for the response matrix, is shown in blue. The error bars correspond to the statistical uncertainty. The lower part shows the ratios of the unfolded and the expected distributions.

Summarizing the different test results, all three approaches are based on a valid unfolding setup and yield the expected statistical properties. In the two tests focusing on potential biases due to differences between the underlying distributions in the observed data and simulation, the results from the simple bin-by-bin approach show some tendency towards the distribution used in the response matrices. For the two remaining approaches, good agreement with the expected distributions can be observed. For the final cross section measurements, the nominal approach is used in the unfolding procedure, since this approach yields the smallest χ^2/dof in all tests and does not rely on additional regularization.

Chapter 9

Systematic uncertainties

9.1	Experimental sources	119
9.2	Theoretical sources	123
9.3	Impact on differential measurements	125
9.4	Definition of covariance matrices	127

The differential cross sections derived in the context of this analysis are subject to various sources of systematic uncertainties, which have to be carefully evaluated to achieve valid results. The different contributions to the total systematic uncertainty can be separated into experimental sources, discussed in Sec. 9.1, and theoretical sources, which mainly concern the modeling of the signal and background processes, discussed in Sec. 9.2.

Unless otherwise stated, the impact of each individual uncertainty is obtained by repeating the analysis twice based on an upward and downward shift of the corresponding uncertainty, which can affect the event selection, the background estimation as well as the unfolding procedure in terms of changes in the response matrix. The difference between the nominal result and the result obtained with the shifts in place yields the estimate of the corresponding uncertainty. This procedure is performed individually for each data-taking period. As discussed in Sec. 8.3, the final cross section results are derived by combining the results from the individual periods taking the corresponding uncertainty correlations, described in this chapter, into account.

The impact of the individual uncertainties on the absolute and normalized differential cross section measurements is further discussed in Sec. 9.3, followed by a definition of the covariance matrices of the final results in Sec. 9.4.

9.1 Experimental sources

The experimental sources of systematic uncertainties include effects regarding the particle and event reconstruction and the event selection. In addition, uncertainty sources related to the background prediction as well as potential detector anomalies are taken into account. The majority of these uncertainties are derived by varying different sets of scale factors, which aim to improve the agreement between observed data and simulation, within their uncertainties employing standard recipes used within the CMS collaboration. The full set of experimental uncertainties including the methods to derive their impact on the differential cross section measurements is discussed in the following.

b-jet tagging

The identification of jets originating from b quarks, as introduced in Sec. 6.2.5, is crucial for the selection of $t\bar{t}$ events. Differences in the identification efficiency for b jets and misidentification probabilities for non-b jets between observed data and simulation are corrected by applying weights to the simulated events [186]. These weights are obtained as a function of jet p_T and $|\eta|$ based on centrally-provided scale factors and efficiency and misidentification probabilities derived for the analysis-specific phase space.

Uncertainties regarding the b-jet tagging are obtained by varying the scale factors for heavy-flavor (b and c jets) and light-flavor jets individually within their corresponding uncertainties. In addition, the uncertainties are broken down into parts that are fully correlated between the different data-taking periods and uncorrelated parts.

In general, most variations introduced in the following could impact the tagging efficiencies, leading to an overestimation of the corresponding uncertainty if the nominal weights, mentioned above, are used. To avoid this, the efficiencies and misidentification probabilities are derived separately with the shifts regarding the jet energy scale and resolution, the jet pileup identification and all modeling uncertainties except for the PDF variations in place. When estimating the corresponding uncertainties, the nominal weights are replaced by weights based on these alternative efficiencies and misidentification probabilities. This modification is not applied for systematic uncertainties regarding the lepton reconstruction or trigger efficiencies, since for these shifts the jet kinematics are not expected to change significantly.

Background estimation

Contributions from SM processes apart from the dileptonic $t\bar{t}$ signal process are removed from the observed event yields based on estimates from simulation, as discussed in Sec. 8.3.1. The normalizations of the corresponding simulated samples are performed using higher-order cross sections (cf. Sec. 5.2.4), which are themselves subject to uncertainties. To take this into account, the following uncertainties are considered for the different SM background processes. Contributions from the $t\bar{t}$ other processes are attributed with a normalization uncertainty of 5%, while for processes including only one top quark, a 30% normalization uncertainty is considered. Contributions from the DY+jets process are assumed to have a normalization uncertainty of 20%, whereas all remaining contributions including W+jets, $t\bar{t}W/Z$ and diboson processes are assigned an uncertainty of 30%.

Background normalization uncertainties concerning one individual contribution are assumed to be fully correlated between the different data-taking periods.

Jet energy scale

The energy scale of reconstructed jets is corrected by applying a dedicated procedure [183, 184] introduced in Sec. 6.2.4. Uncertainties regarding this procedure are provided as a function of jet p_T and η and include seven individual sources, which are either fully correlated or uncorrelated between the different data-taking periods. The sources consider, e.g., uncertainties on the impact of pileup contributions or on the simulated jet momentum response. The impact of the jet energy scale (JES) uncertainty on the final cross section results is derived by repeating the analysis with the jet momenta scaled up and down by the corresponding JES uncertainty. In addition, the change in the jet momenta is propagated to the missing transverse momentum. Thus, unlike uncertainties, which are only derived based on varying event weights, this uncertainty can also affect the event selection itself.

Jet energy resolution

In addition to corrections of the JES, the jet energy resolution (JER) is smeared in simulation to match the resolution observed in data (cf. Sec. 6.2.4) [183, 185]. To derive the uncertainty related to this smearing, the JER is shifted within its uncertainty individually for jets with $|\eta| < 1.93$ and with $1.93 \leq |\eta| < 2.5$. This uncertainty source is assumed to be uncorrelated between the different data-taking periods.

Jet pileup identification

Another jet-related uncertainty is introduced by applying identification criteria, which aim to discard contributions from jets originating from pileup vertices [181], as introduced in Sec. 6.2.4. To correct for potential efficiency differences between data and simulation, dedicated scale factors are derived and varied within their uncertainties to obtain their impact on the cross section results. The uncertainty regarding this identification criterion is assumed to be fully correlated between the different data-taking periods.

L1-trigger prefiring

Timing shifts in the ECAL, which were not correctly propagated to the L1 trigger during the 2016 and 2017 data-taking periods, introduced the so-called trigger prefiring [104]. Due to the prefiring, trigger signals were wrongly associated with the previous bunch crossing, leading to a veto of the original bunch crossing, since two consecutive bunch crossings are not allowed to both trigger the L1 trigger. Similar effects were also present in the muon system in all four data-taking periods [102]. To correct for these effects, additional scale factors are applied to the simulation. The uncertainty of this correction is derived by shifting the scale factors within their uncertainties. The resulting uncertainty is taken to be fully correlated between the different data-taking periods.

Lepton reconstruction

To improve the modeling of the lepton reconstruction and selection in simulation, multiple scale factors are applied to the reconstructed muons [101, 175] and electrons [97, 173] as a function of p_T and η (cf. Sec. 6.2.2 and 6.2.3). To derive the uncertainty originating from these scale factor applications, the corresponding scale factors, distinguishing between reconstruction and identification for the electrons and identification and isolation for the muons, are individually varied within their uncertainties. Since the scale factors are derived based on DY events but are predominantly applied to $t\bar{t}$ events, an additional uncertainty of 1% and 0.5% per electron and muon, respectively, is applied to cover the impact of the extrapolation.

Uncertainties regarding the muon scale factors are further divided into a systematic part, which is assumed to be fully correlated between the different data-taking periods, and a statistical part, which is taken to be uncorrelated. Uncertainties regarding the electron scale factors are considered as fully correlated between the different data-taking periods.

Further corrections are applied to the lepton momentum scale and resolution, which introduce additional uncertainties [97, 176, 177]. These uncertainties are derived individually per lepton flavor by varying the corrections within their uncertainties. The shifts in the lepton momenta are, similar to the JES, propagated to the missing transverse momentum. Uncertainties regarding the lepton scale and resolution are assumed to be fully correlated between the different data-taking periods.

Luminosity measurement

The integrated luminosity is measured individually for each year of data-taking. Since similar methods are applied in each of these measurements, the uncertainties on the luminosities are partially correlated. The corresponding correlation scheme is shown in Tab. 9.1. The uncertainties on the luminosities for the data-taking periods 2016preVFP and 2016postVFP are assumed to be fully correlated.

Table 9.1: Breakdown of the correlated and uncorrelated luminosity uncertainties per year of data-taking. The individual uncertainties are shown in the last row [116–118].

Year	2016 (%)	2017 (%)	2018 (%)
Uncorrelated 16	1.0	—	—
Uncorrelated 17	—	2.0	—
Uncorrelated 18	—	—	1.5
Correlated 16, 17, 18	0.6	0.9	2.0
Correlated 17, 18	—	0.6	0.2
Individual result	1.2	2.3	2.5

Pileup

The uncertainty regarding pileup contributions originates from the reweighting of the pileup distributions in simulation to the expected pileup distribution in data, which is based on the total inelastic pp cross section of 69.2 mb [141]. The uncertainty is derived by repeating the reweighting with the total inelastic pp cross section shifted by $\pm 4.6\%$ and is assumed to be fully correlated between the different data-taking periods.

Trigger efficiencies

The uncertainties of the scale factors accounting for the difference in trigger efficiencies between data and simulation have been already discussed in detail in Sec. 5.3.2. The final uncertainties on the trigger efficiencies, which are taken to be uncorrelated between the different data-taking periods, are derived by varying the scale factors within their uncertainties.

Unclustered energy

In addition to propagating the shifts in the JES and the lepton momentum scale and resolution to the missing transverse momentum, an uncertainty related to the unclustered energy [189] is taken into account. For this uncertainty, the deposited energies of charged and neutral hadrons as well as photons and hadrons, which are reconstructed in the forward part of the HCAL (cf. Sec. 3.2.5), are varied within their corresponding energy resolutions and propagated to the missing transverse momentum. The impact on the final result is obtained by repeating the analysis based on these shifted p_T^{miss} vectors. This uncertainty is assumed to be uncorrelated between the different data-taking periods.

9.2 Theoretical sources

The systematic uncertainties originating from theoretical sources mainly concern the modeling of the signal and background processes as used in the individual steps of the generation of simulated events, introduced in Sec. 5.2. These uncertainties are derived by changing the corresponding modeling parameters using additional simulated samples with altered parameters or applying additional weights to the nominal samples. In both cases, the total cross sections used to normalize the samples are kept constant. The uncertainties introduced below are either applied to all simulated samples, to the $t\bar{t}$ processes including all decay channels or only to the dileptonic $t\bar{t}$ signal sample. In contrast to the experimental sources, all theoretical uncertainties are assumed to be fully correlated between the different data-taking periods.

b-hadron branching fraction

The semileptonic branching fraction of b hadrons can have a significant impact on the energy response of b jets due to the non-prompt neutrinos produced in the decay chain, which escape the detector undetected. To account for the uncertainty on the branching fraction and differences between the branching fractions stated by the Particle Data Group [221] and the value used in the PYTHIA8 generator [133], additional weights are applied to the simulation of the $t\bar{t}$ processes. These weights cover the envelope of the branching fraction uncertainty and the differences between the two branching fractions mentioned above.

Color reconnection

For the nominal $t\bar{t}$ simulations, a color reconnection (CR) model (cf. Sec. 5.2.2), based on multi-parton interactions (MPI) with the option for early resonance decays (ERD) disabled [137], is applied. To account for differences between different CR models, additional $t\bar{t}$ simulations with three alternative CR models are used. These models include a QCD-inspired model, a gluon move model and an MPI-based model with the ERD enabled. The analysis is repeated separately for each of the additional simulated samples and the uncertainty attributed to the choice of the CR model is derived based on the envelope of the different final cross section results compared to the nominal result.

ME scale

Uncertainties due to the renormalization scale μ_R and the factorization scale μ_F in the matrix element generators are derived by applying dedicated event weights to all simulated samples. Here, a total of six different shifts are taken into account:

- μ_R fixed and μ_F varied by 2 (0.5) for up (down) variation
- μ_F fixed and μ_R varied by 2 (0.5) for up (down) variation
- simultaneously varying μ_R and μ_F by 2 (0.5) for up (down) variation

The combined ME-scale uncertainty on the final cross section results is obtained based on the envelope of the six different shifts with respect to the nominal result.

ME-PS matching

In the POWHEGV2 generator used to simulate the $t\bar{t}$ processes, the cross section for real emissions is scaled by a damping function, which controls the merging between the matrix element (ME) and the parton shower (PS) and regulates high- p_T radiations. The damping function depends on the parameter h_{damp} [136], which is varied within its uncertainty to obtain the uncertainty related to the ME-PS matching. This variation is realized by using two alternative $t\bar{t}$ simulations, which include the shift in h_{damp} , and repeating the analysis separately for both alternative simulations.

Parton distribution functions

The uncertainty originating from the parton distribution functions (PDFs) is derived by applying event weights for the dileptonic $t\bar{t}$ sample based on 100 eigenvectors of the NNPDF3.1 NNLO hessian error PDF set [134]. The combined uncertainty is obtained from the envelope of the shifted final results with respect to the nominal result. In addition, the value of the strong coupling constant α_s , applied in the PDF set, is varied within its uncertainty using two additional sets of event weights for the dileptonic $t\bar{t}$ sample.

Parton shower

To account for the uncertainty of α_s in the parton shower, the renormalization scale for QCD emissions in the final-state and initial-state (FSR and ISR) is varied. This variation is performed based on event weights, which separately shift the scale of the FSR and ISR by a factor of two and one half. The analysis is repeated four times to consider the individual shifts connected to both types of radiation. This uncertainty is derived for all simulated samples including non- $t\bar{t}$ processes.

Top-quark mass

To account for the uncertainty on the top-quark mass m_t , the analysis is repeated with two alternative sets of $t\bar{t}$ simulations assuming a top-quark mass of 169.5 and 175.5 GeV, while for the nominal simulated samples $m_t = 172.5$ GeV is used. The final uncertainty related to the top-quark mass is estimated by linearly scaling the variations induced by the two alternative sets of simulations assuming an uncertainty on the top-quark mass of 1 GeV.

Top-quark p_T reweighting

To improve the agreement between data and simulation with respect to the top-quark p_T spectrum, additional k -factors are applied to all $t\bar{t}$ simulations, as discussed in Sec. 5.2.4. The uncertainty originating from this additional reweighting is derived by repeating the analysis without the reweighting in place and comparing the result to the nominal measurement leading to a one-sided uncertainty.

Underlying event

The uncertainty related to the simulation of the underlying event (cf. Sec. 5.2.2), is derived based on two alternative $t\bar{t}$ simulations, where the parameters of the CP5 tune, describing the underlying event in the simulation, are varied within their corresponding uncertainties, as described in [136].

9.3 Impact on differential measurements

The impact of the systematic uncertainties on the final differential cross section measurements is derived by repeating the analysis multiple times, combining the individual results for the different data-taking periods taking proper correlations into account and comparing the shifted with the nominal results, as discussed in the previous sections. The statistical uncertainty in the final results is derived from the statistical precision of the observed number of events entering the unfolding procedure, introduced in Sec. 8.3. The total uncertainty in each bin of the final results is derived by computing the quadratic sum of all uncertainty contributions.

In Fig. 9.1, the breakdown of the uncertainty contributions for the $p_T^{\nu\nu}$ measurement is shown. The total uncertainty increases with increasing momenta from below 5% to almost 40% caused by larger contributions from the statistical and systematic uncertainties. In the phase space between 100 and 400 GeV, the largest contributions originate from the jet energy scale uncertainty caused by its propagation to the missing transverse momentum, which serves as a proxy for the $p_T^{\nu\nu}$ observable. The second and third as well as the last bin of the measurement are dominated by the uncertainty on the background prediction. For the second and third bin, this effect mainly originates from the large DY+jets contributions. These contributions are not present in the first bin of the measurement due to the $p_T^{\nu\nu}$ requirement applied to same-flavor events within the particle-level definition (cf. Sec. 8.1). The background uncertainties as well as the matrix element scale uncertainties in the last bin of the measurement, in contrast, are mainly caused by the tW contributions, which dominate the background in this part of the phase space. This observation can be explained based on the event selection applied in this analysis, introduced in Sec. 6.3. Due to the requirement of at least two reconstructed jets, tW events with an additional jet from ISR or FSR are predominantly selected. In case of the presence of an ISR jet, the

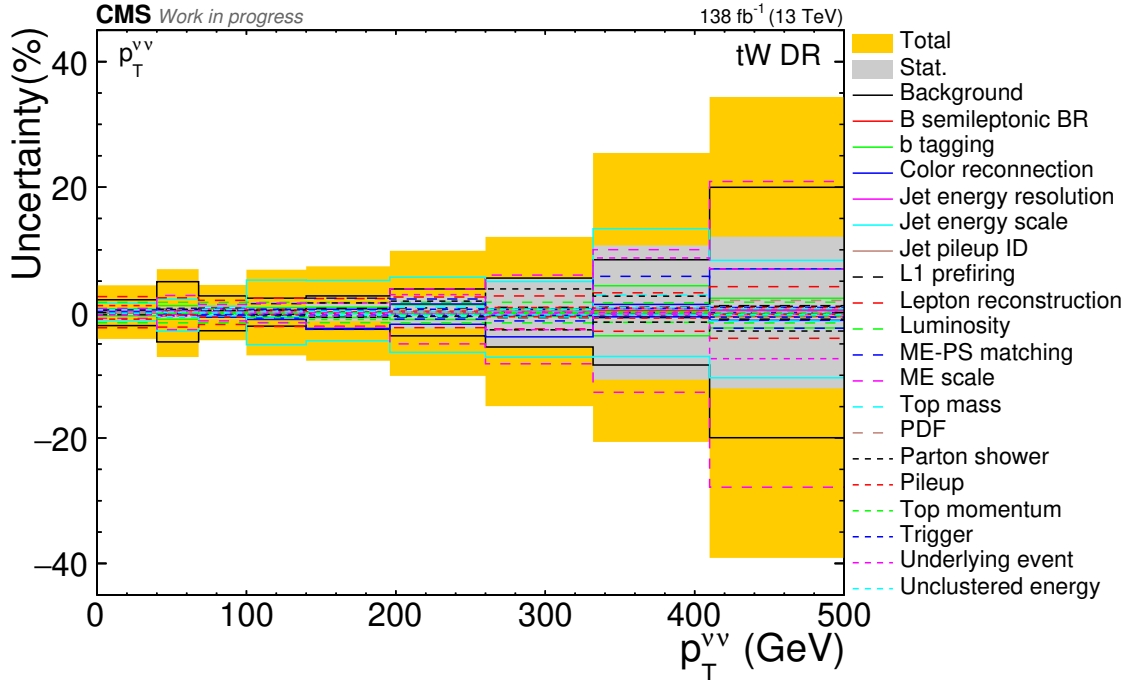


Figure 9.1: Breakdown of the uncertainties for the $p_T^{\nu\nu}$ measurement. The statistical uncertainty is shown as gray-shaded area. The orange-shaded area indicates the total uncertainty, given by the quadratic sum of all systematic contributions and the statistical uncertainty.

$t\bar{W}$ system can be boosted, allowing for higher neutrino momenta compared to the $t\bar{t}$ topology, which leads to larger $t\bar{W}$ contributions at high $p_T^{\nu\nu}$.

The total uncertainty derived for the $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ measurement ranges between 3.5 and 7.0%. The corresponding breakdown of the different uncertainty contributions, which significantly differs from the one observed for the $p_T^{\nu\nu}$ measurement, is shown in Fig. 9.2. For the majority of the measured phase space, the dominant uncertainty contributions originate from the lepton reconstruction, caused by the lepton kinematics affecting the measured observable. The largest part of this contribution originates from the uncertainty on the electron reconstruction. For the phase space with small angular separation, the contributions from the jet energy scale and the background prediction have a larger impact. The increased impact from the background prediction is caused by a larger contribution from DY +jets events in this phase space. The contributions from the jet energy scale, in contrast, originate from events with an increased jet activity or additional jets which tend to yield signatures with smaller angular separation due to the additional boost of the $t\bar{t}$ system.

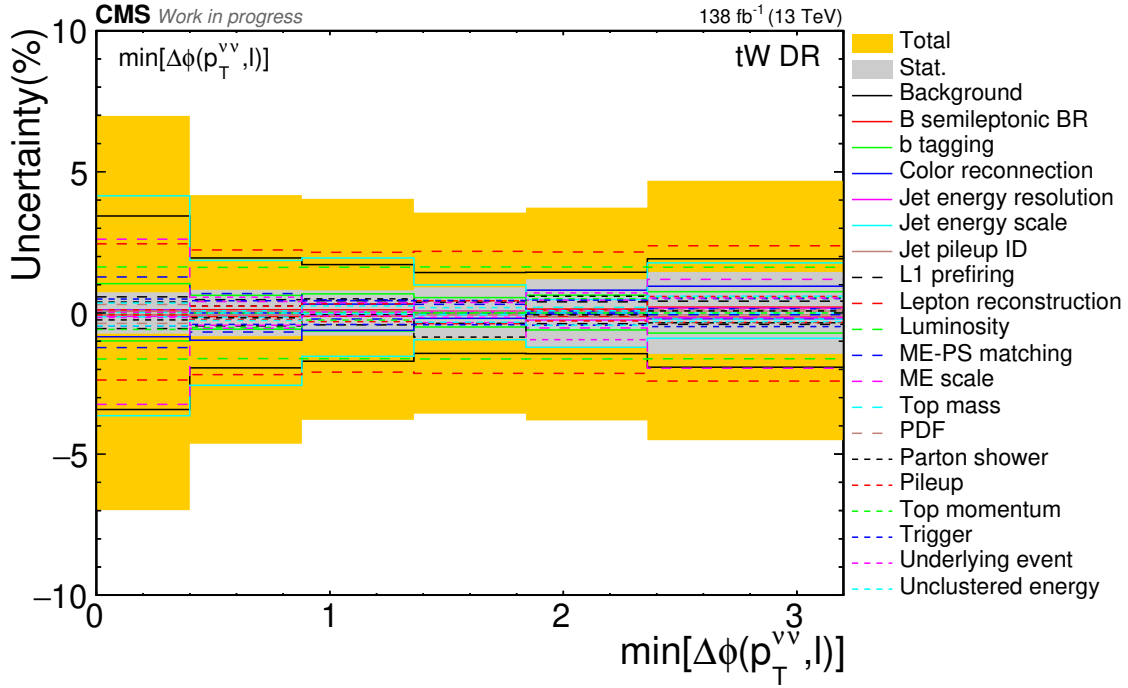


Figure 9.2: Breakdown of the uncertainties for the $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ measurement. The statistical uncertainty is shown as gray-shaded area. The orange-shaded area indicates the total uncertainty, given by the quadratic sum of all systematic contributions and the statistical uncertainty.

The uncertainty breakdown for the 2D measurement, where the total uncertainties vary between around 3.5 and 27.5%, depending on the phase space, can be found in Fig. 9.3. The breakdown in the first angular bin is rather similar to the one observed in the 1D $p_T^{\nu\nu}$ measurement discussed above. In the first (last) bin of the spectrum, the uncertainties from the background prediction have a significant impact caused by large DY +jets ($t\bar{W}$) contributions in the corresponding phase space. At high $p_T^{\nu\nu}$, the $t\bar{W}$ background leads to larger contributions from the matrix element scale, while for the medium- p_T range the uncertainty is dominated by the jet energy scale. These observations do not hold for the other two angular bins. Here, the contributions

from the background predictions are significantly reduced and the dominant uncertainty in the majority of the measured bins originates from the jet energy scale.

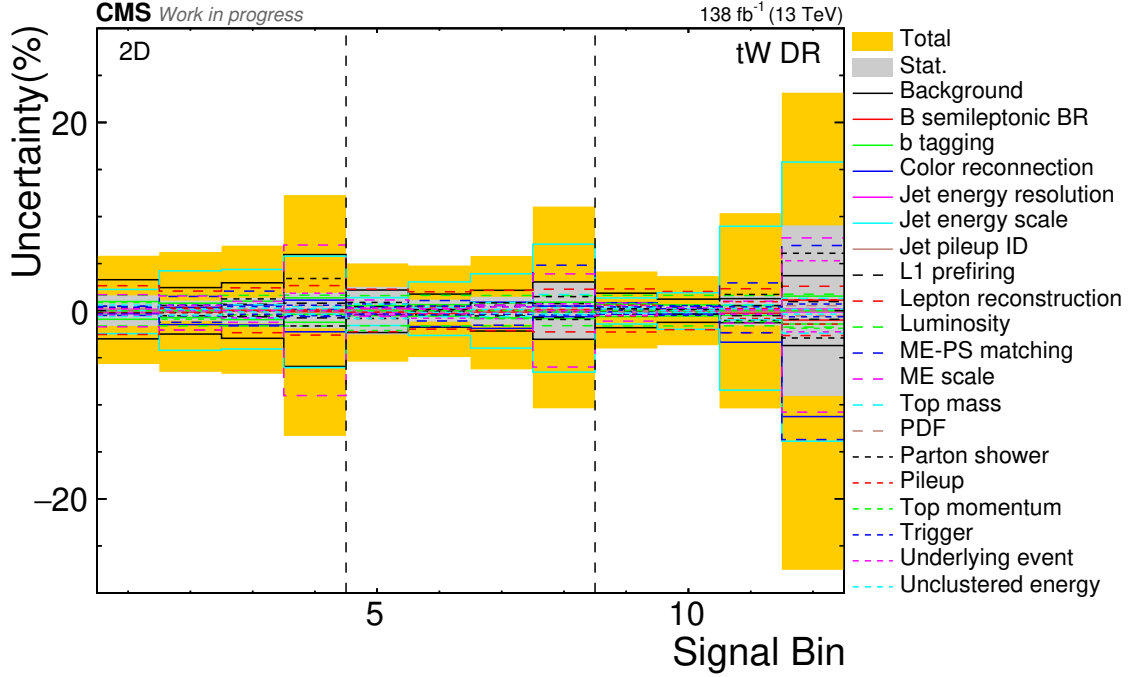


Figure 9.3: Breakdown of the uncertainties for the 2D measurement. The statistical uncertainty is shown as gray-shaded area. The orange-shaded area indicates the total uncertainty, given by the quadratic sum of all systematic contributions and the statistical uncertainty. The vertical black lines correspond to the bin edges in $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$, while inside each of these bins, several $p_T^{\nu\nu}$ bins are shown in increasing order.

The breakdowns of the uncertainties for the normalized cross section measurements and for the results derived with the DS scheme (cf. Sec. 5.2.5) can be found in Appx. 10 and 11, respectively. For the normalized $p_T^{\nu\nu}$ and the normalized 2D measurements, the impacts of the individual uncertainties are slightly reduced compared to the absolute measurements, while the dominant contributions in the different phase space regions are mainly unchanged. For the normalized $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ measurement, however, the total uncertainty is significantly reduced compared to the absolute measurement. This change is mainly caused by the uncertainty on the lepton reconstruction, which mostly affects the overall normalization rather than the shape of the measurement. As a consequence, most parts of the range considered in the normalized $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ measurement are dominated by the uncertainty on the jet energy scale.

9.4 Definition of covariance matrices

In the discussion of the final results, presented in Sec. 10, the measured differential cross sections are compared to different theoretical predictions. To account for the correlations present between individual bins of the final results, these comparisons take the full covariance matrices of the measurements, whose definition is introduced in the following, into account.

The covariance matrices $\mathbf{C}$ of the final measurements, based on the combined results of all data-taking periods, include two separate parts:

$$\mathbf{C} = \mathbf{C}^{\text{stat}} + \mathbf{C}^{\text{syst}}, \quad (9.1)$$

where $\mathbf{C}^{\text{stat}}$ denotes the covariance matrix encoding the statistical uncertainties after the unfolding, which, unlike the covariance of the observed input data (cf. Sec. 8.3), can have off-diagonal elements originating from correlations introduced in the unfolding procedure. The second part $\mathbf{C}^{\text{syst}}$ of the total covariance matrix includes the contributions from the systematic uncertainties.

While the statistical covariance $\mathbf{C}^{\text{stat}}$ is derived based on the sum of the individual statistical covariances of each data-taking period obtained within the TUnfold framework [220], the systematic covariance matrix $\mathbf{C}^{\text{syst}}$ is estimated directly from the systematic uncertainties on the combined results (cf. Sec. 9.3), following the method used in [222]. In detail, the calculation of the systematic covariance matrix elements is given by:

$$\mathbf{C}_{i,j}^{\text{syst}} = \sum_{k=1}^M \frac{1}{N_k} \sum_{l=1}^{N_k} C_{k,l,i} C_{k,l,j} \quad \text{with } 1 \leq i, j \leq m. \quad (9.2)$$

The outer sum runs over all M systematic uncertainty sources introduced in Sec. 9.1 and 9.2. The inner sum corresponds to a mean across all N_k variations considered for the present systematic uncertainty source k . For the majority of the sources, $N_k = 2$ holds, given that an upward and downward variation is taken into account. However, for sources including more than two variations, e.g., the matrix element scale or color reconnection, N_k has to be adapted. The argument of the inner sum contains the shifts $C_{k,l,i}$ introduced in the i th particle-level bin by variation l of the uncertainty source k . With this construction, the diagonal elements of the covariance hold the squared total uncertainties, while the off-diagonal elements encode the correlations leading to a valid covariance matrix.

The comparisons between the measured cross section results and different theoretical predictions, taking the full covariance matrices into account, are performed based on a χ^2 test:

$$\chi^2 = \mathbf{D}^T \mathbf{C}^{-1} \mathbf{D}, \quad (9.3)$$

where $\mathbf{D}$ corresponds to the vector holding the difference between the measurement and the prediction for each of the m particle-level bins and $\mathbf{C}^{-1}$ denotes the inverse of the total covariance matrix, as defined in Eq. 9.1. For the normalized cross section results, one degree of freedom is absorbed due to the normalization of the results reducing the degrees of freedom used in the χ^2 test to $m-1$. In addition, the covariance matrix becomes singular due to the normalization, which prevents the construction of its inverse $\mathbf{C}^{-1}$. To avoid this problem, one of the bins considered in Eq. (9.3) is discarded, which leads to a non-singular covariance matrix, whose inverse is well-defined. Due to the additional correlation between the bins introduced by the normalization, the result of this reduced χ^2 test is independent of the choice of the discarded bin, which has been tested for the normalized results presented in this thesis.

Chapter 10

Results

10.1 One-dimensional cross sections	130
10.1.1 Measurements in $p_T^{\nu\nu}$	130
10.1.2 Measurements in $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$	133
10.2 Two-dimensional cross sections	137

The different steps towards a reliable differential cross section measurement of the dileptonic $t\bar{t}$ process, including a reasonable event selection, a correction of the missing transverse momentum, the definition of the cross section extraction as well as a detailed estimation of the systematic uncertainties were introduced in the previous chapters. The results of these measurements, performed on the full Run 2 data set, corresponding to an integrated luminosity of 138 fb^{-1} , recorded by the CMS experiment combining all three leptonic decay channels, are presented in the following. The first part of this chapter focuses on the results derived for the one-dimensional measurements as a function of $p_T^{\nu\nu}$ and $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$. In the second part, the absolute and normalized two-dimensional measurements, employing both particle-level observables, are discussed. For each differential cross section measurement, two results using two different schemes (DR and DS) to describe the tW background contribution in simulation (cf. Sec. 5.2.5) are provided. The results presented in the following are denoted as "CMS Work in progress", which marks results, that are not yet published in a journal paper but are endorsed by the CMS collaboration for use in a Ph.D. thesis.

Following the strategy discussed in Sec. 2.3.3, once all analysis ingredients are fixed, the detector-level distributions used to obtain the cross sections are studied regarding the agreement between observed data and simulation to account for potential BSM contributions (cf. Appx. 12) before performing the unfolding procedure. Only if a good agreement between data and simulation is observed, the unfolding procedure is performed to derive the desired differential cross sections. The differential cross sections are compared to four different theoretical predictions. Two of these predictions are based on the MC simulations used throughout this analysis (cf. Sec. 5.2.4). The first of these predictions corresponds to the nominal dileptonic $t\bar{t}$ signal simulation generated with the POWHEGV2 generator and is denoted as POWHEG in the following. The second prediction based on MC simulations corresponds to the alternative signal simulation generated with the MADGRAPH5_AMC@NLO generator, denoted as MADGRAPH in the following. Both of these predictions are based on NLO matrix element accuracy, use the NNPDF3.1 PDF set at NNLO accuracy and are interfaced with the PYTHIA8 generator to describe the parton shower. The normalization of both samples is based on the inclusive cross section derived at

NNLO+NNLL accuracy. The uncertainty on both MC-based predictions includes a 5% normalization uncertainty and matrix element scale uncertainties, as introduced in Sec. 9.2.

The second set of predictions used to interpret the cross section measurements corresponds to fixed-order calculations [30] based on the methods described in [29]. Here, one prediction based on full NNLO accuracy in QCD using the Narrow Width Approximation (NWA) is provided. The second fixed-order calculation corresponds to NLO accuracy in QCD and is also based on the NWA approach. The uncertainty of the fixed-order calculations covers the matrix element scales using the same scheme as discussed in Sec. 9.2 as well as uncertainties regarding the finite size of MC samples used in the calculation of the fixed-order predictions. For the NNLO prediction, additional uncertainty contributions originating from the PDFs are taken into account.

10.1 One-dimensional cross sections

10.1.1 Measurements in $p_T^{\nu\nu}$

The differential cross section measurements on particle level as a function of the transverse momentum $p_T^{\nu\nu}$ of the dineutrino system are performed based on the detector-level distributions of the corrected missing transverse momentum DNN p_T^{miss} , introduced in Sec. 7. The comparisons between observed data and simulation for the given detector-level distributions, based on all lepton channels and data-taking periods, are shown in Fig. 10.1. For the distribution using the DR scheme, the region up to a missing transverse momentum of around 300 GeV yields a good agreement between data and simulation. For the region above this value, the simulation predicts a harder spectrum compared to the observed distribution. However, this tension is covered by the uncertainty attributed to the predicted distribution, which, in this phase space, is dominated by the normalization and ME-scale uncertainties of the tW background prediction (cf. Sec. 9.3).

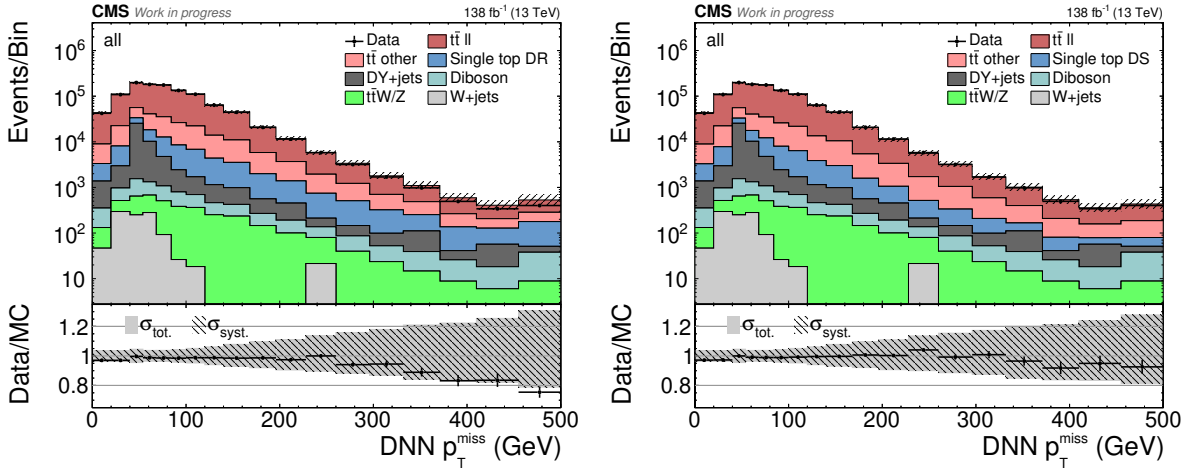


Figure 10.1: Detector-level distributions as a function of DNN p_T^{miss} for all channels and data-taking periods combined using the DR (left) and the DS scheme (right). The upper parts of the plots show the observed data (black) and the contributions predicted by simulation including the total uncertainty (hatched area). The lower parts show the ratio between data and simulation as well as the systematic and total uncertainties on the predictions. The last bin includes all events above the plotted range.

For the distribution using the DS scheme, the agreement between data and simulation in the high- p_T^{miss} region is significantly improved. The reason for this improvement is the following. As shown in Fig. 10.1, the largest background contribution in the high- p_T^{miss} region originates from

the single top process, which is dominated by the tW production mechanism. The comparison of the generated- p_T^{miss} spectra between the single top simulation using the DR and the DS scheme, studied in Sec. 5.2.5, yields large differences between both spectra in the given phase space, where the DR scheme predicts a significantly harder spectrum. Thus, describing the tW background based on the DS scheme yields an overall softer spectrum in simulation, which results in a better agreement with the observed data. However, based on both detector-level distributions no indications of potential BSM contributions, which would be expected in the high- p_T^{miss} region, as discussed in Sec. 2.3.3, can be found.

Given there is no evidence of potential BSM contributions, differential cross sections as a function of $p_T^{\nu\nu}$ are derived, as shown in Fig. 10.2. Comparing the different theory predictions to the measured spectrum using the DR scheme indicates that all four predictions describe the data

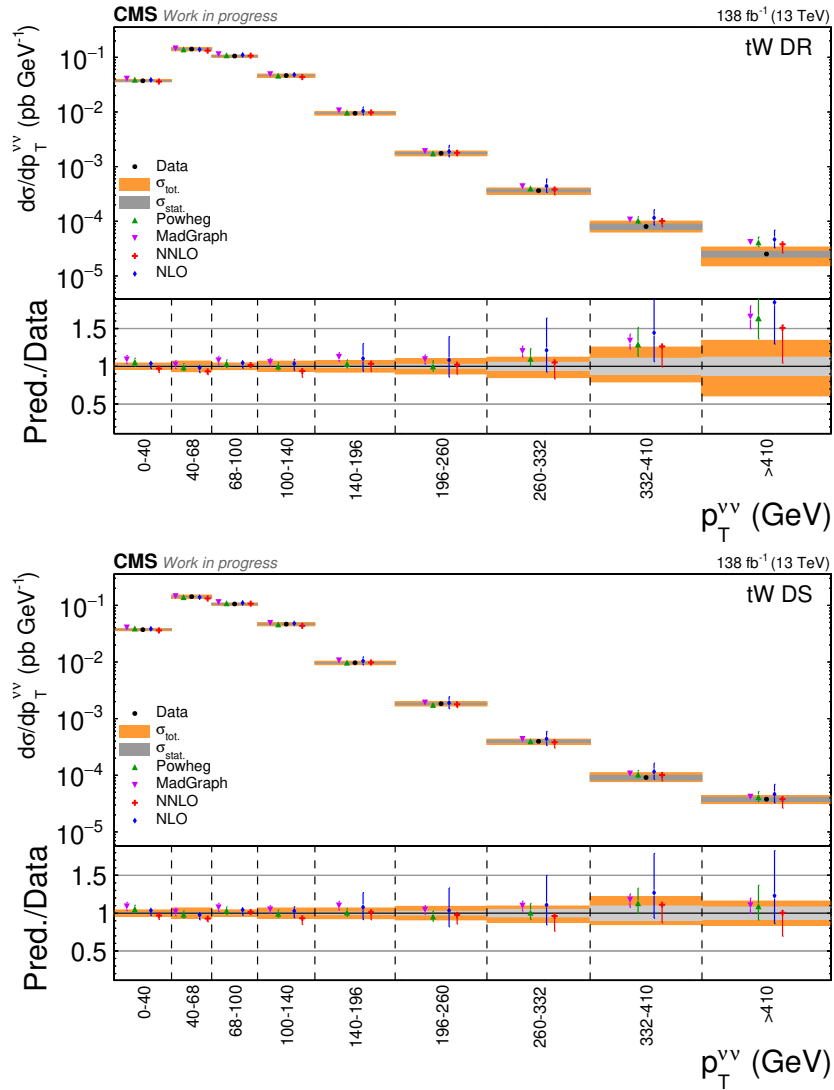


Figure 10.2: Differential cross section results as a function of $p_T^{\nu\nu}$ based on the DR (top) and the DS scheme (bottom). The total (statistical) uncertainty on the measurement is shown as an orange (grey) band. The results are compared to NLO (blue) and NNLO (red) fixed-order predictions as well as predictions from the POWHEG (green) and MADGRAPH (purple) generators. The lower parts show the ratio between the different predictions and the measurement.

reasonably well up to the region of 250 to 300 GeV. Only the MADGRAPH prediction yields a small difference in overall normalization. Above this region, all four predictions show a harder spectrum compared to the measured cross section, where the smallest tension is observed for the NNLO fixed-order prediction. As expected from the detector-level distributions, the result using the DS scheme shows a significantly improved agreement between the predictions and the measured cross section in the high- $p_T^{\nu\nu}$ region. In addition, the systematic uncertainty in the last two bins is reduced compared to the DR results. This change is caused by the smaller tW background contribution predicted by the DS scheme since the corresponding uncertainty on the background prediction is the dominating contribution at high- $p_T^{\nu\nu}$ (cf. Sec. 9.3).

While these observations neglect the correlations between the measured bins, the results of the χ^2 tests take the proper correlations into account and are shown in Tab. 10.1. Considering not only the uncertainty on the measurements but also on the predictions, the tests indicate an excellent agreement between the measurements and all four predictions. The large impact of the correlations between the measured bins can be observed for the MADGRAPH prediction. While the comparisons of the measured and predicted spectra based on Fig. 10.2 indicate an agreement that is slightly worse compared to the other predictions, the χ^2 test neglecting the uncertainties in the predictions yields the best result for the MADGRAPH prediction. The differences between the cross section measurements using the DR and the DS scheme, discussed above, can be also observed in the χ^2 tests, which yield smaller results for the comparisons between all four predictions and the measurement based on the DS scheme.

Table 10.1: Results of the χ^2 tests (cf. Sec. 9.4) performed for the absolute and normalized measurements as a function of $p_T^{\nu\nu}$ between the measured cross section and the four predictions. The nominal results take only the covariance of the measurement into account, while the values in round brackets include the uncertainties on the predictions. The results for the measurements based on the DR and the DS scheme are shown. For each measurement, the degrees of freedom (dof) are given.

Measurement	dof	POWHEG	MADGRAPH	NNLO	NLO
$p_T^{\nu\nu}$ (tW DR)	9	14.8 (4.0)	12.7 (6.7)	20.2 (5.5)	17.9 (2.8)
norm. $p_T^{\nu\nu}$ (tW DR)	8	14.2 (4.9)	10.8 (7.7)	20.1 (7.6)	16.6 (2.7)
$p_T^{\nu\nu}$ (tW DS)	9	9.2 (1.7)	9.0 (3.7)	13.9 (2.6)	12.5 (1.5)
norm. $p_T^{\nu\nu}$ (tW DS)	8	8.9 (3.8)	6.6 (5.7)	14.7 (5.6)	12.1 (1.8)

The results of the normalized differential cross section as a function of $p_T^{\nu\nu}$ can be found in Fig. 10.3. Here, similar conclusions regarding slight tensions at high $p_T^{\nu\nu}$ for all four predictions in the measurement using the DR scheme, as already seen in the absolute measurement, can be drawn. Again, the tension is significantly reduced for the measurement based on the DS scheme. Apart from this, all four predictions show a good agreement with both measured spectra. The largest difference compared to the nominal measurement is given by the comparison to the MADGRAPH prediction, where the difference in the overall normalization is removed. The results of the χ^2 tests for the normalized measurements are listed in Tab. 10.1. Ignoring the uncertainties on the predictions, the best overall description of the normalized spectra is observed for both MC-based predictions, while both fixed-order calculations still yield a reasonable description of both measured spectra.

The comparison between both $p_T^{\nu\nu}$ measurements shows that the high- $p_T^{\nu\nu}$ phase space is highly sensitive to the description of the tW process. In this specific observable, the description based on the DS scheme seems to be preferred by the measured data.

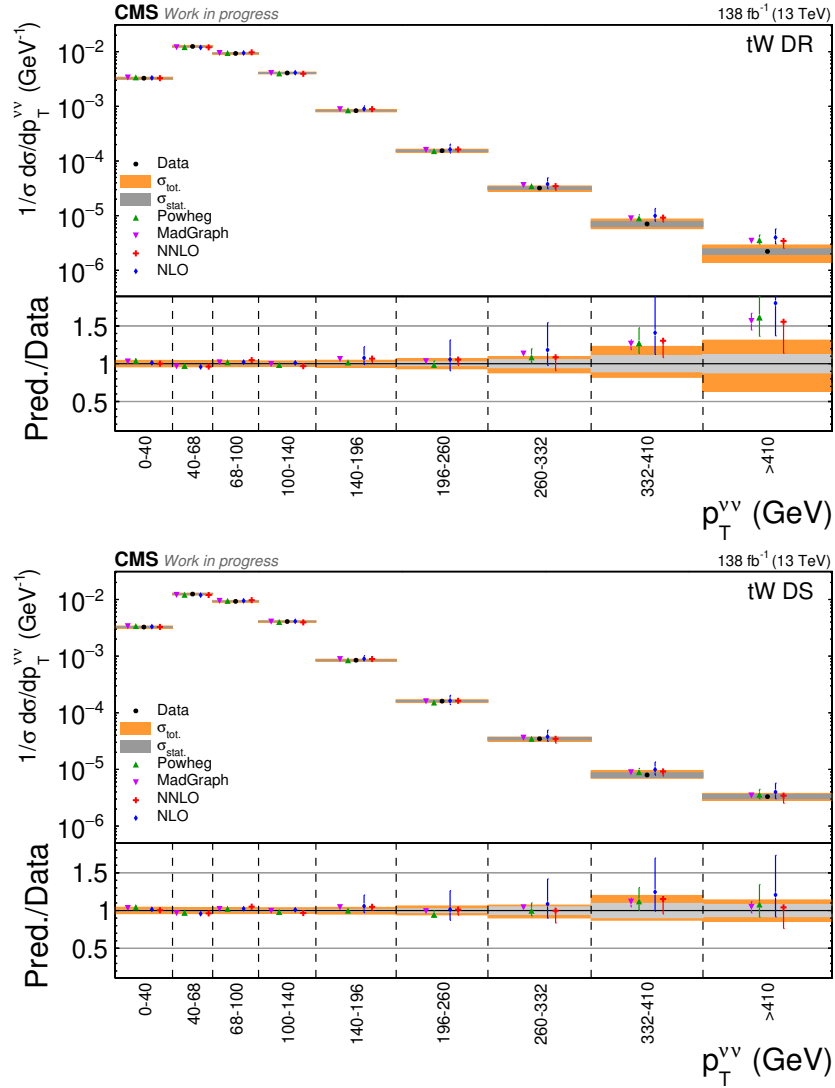


Figure 10.3: Normalized differential cross section results as a function of $p_T^{\nu\nu}$ based on the DR (top) and the DS scheme (bottom). The total (statistical) uncertainty on the measurement is shown as an orange (grey) band. The results are compared to NLO (blue) and NNLO (red) fixed-order predictions as well as predictions from the POWHEG (green) and MADGRAPH (purple) generators. The lower parts show the ratio between the different predictions and the measurement.

10.1.2 Measurements in $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$

The differential cross sections as a function of the minimal azimuthal distance $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ between the dineutrino system and a lepton are based on the detector-level distributions of the minimal azimuthal distance $\min[\Delta\phi(p_T^{\text{miss}}, \ell)]$ between the missing transverse momentum and a lepton, which are shown in Fig. 10.4. Comparisons between the observed data and the simulation indicate a reasonable agreement for both detector-level distributions, taking the total uncertainty of the predictions into account. Here, no significant differences between the distributions using the two different schemes for the tW simulation are found. However, for both distributions, a small shape difference between data and simulation can be observed across the entire measured region, where the data yields smaller (larger) contributions at small (large) angular separations. Similar to the detector-level distributions used in the previous measurements,

no indications for potential BSM contributions, which would be mainly expected at large angular separation (cf. Sec. 2.3.3), can be found.

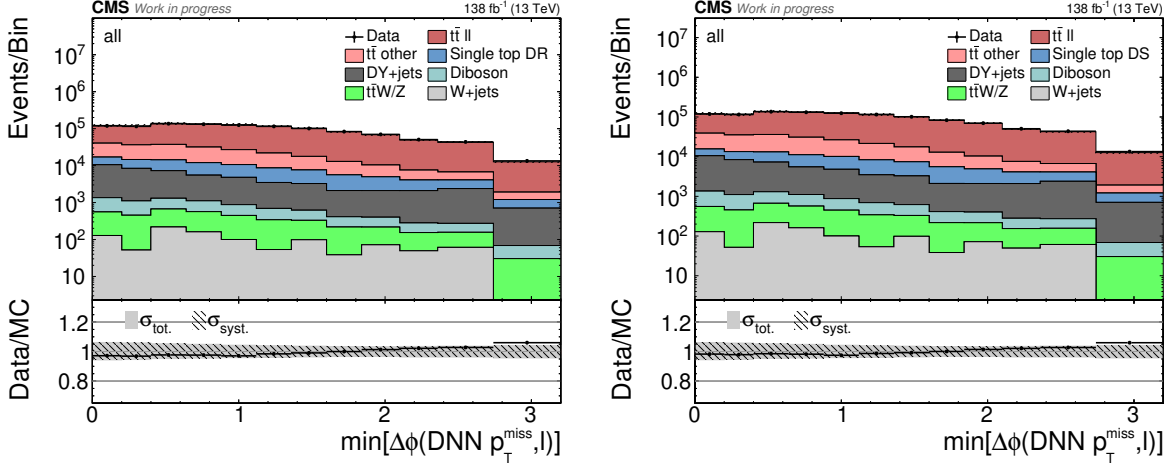


Figure 10.4: Detector-level distributions as a function of $\min[\Delta\phi(p_T^{\text{miss}}, \ell)]$ using DNN p_T^{miss} for all channels and data-taking periods combined based on the DR (left) and the DS scheme (right). The upper parts of the plots show the observed data (black) and the contributions predicted by simulation including the total uncertainty (hatched area). The lower parts show the ratio between data and simulation as well as the systematic and total uncertainties on the predicted distribution. The last bin includes all events above the plotted range.

The results for the differential cross section as a function of $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ can be found in Fig. 10.5. In general, all four predictions show a reasonable agreement with both measured spectra. However, similar to the situation observed on detector level, all predictions indicate a small tension with the measured data regarding the shape of the spectrum. While for the MC-based predictions, this tension is well within the uncertainties across the entire range, both fixed-order predictions show a small discrepancy in the last bin of the measured spectra. The best overall agreement between data and predictions is observed for the POWHEG prediction, which yields a better description of the phase space of small angular separation compared to the MADGRAPH prediction. This observation still holds when taking the full covariance matrix of the measurements into account, as shown by the χ^2 -test results in Tab. 10.2. While all four predictions show good agreement when considering the corresponding uncertainties on the

Table 10.2: Results of the χ^2 tests (cf. Sec. 9.4) performed for the absolute and normalized measurements as a function of $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ between the measured cross section and the four predictions. The nominal results take only the covariance of the measurement into account, while the values in round brackets include the uncertainties on the predictions. The results for the measurements based on the DR and the DS scheme are shown. For each measurement, the degrees of freedom (dof) are given.

Measurement	dof	POWHEG	MADGRAPH	NNLO	NLO
$\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ (tW DR)	6	7.9 (1.9)	12.5 (3.2)	11.7 (4.4)	12.9 (2.9)
norm. $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ (tW DR)	5	7.7 (7.5)	10.4 (9.7)	11.0 (5.7)	12.2 (6.0)
$\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ (tW DS)	6	8.1 (1.6)	12.5 (2.7)	12.3 (4.1)	13.3 (2.6)
norm. $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ (tW DS)	5	7.9 (7.8)	10.2 (9.5)	12.1 (6.5)	12.7 (5.9)

predictions, the POWHEG prediction yields the smallest χ^2 values for both results. As already observed for the detector-level distributions, the scheme used for the description of the tW process in simulation has no significant impact on the $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ measurement, which is also indicated by the similar results of the χ^2 tests.

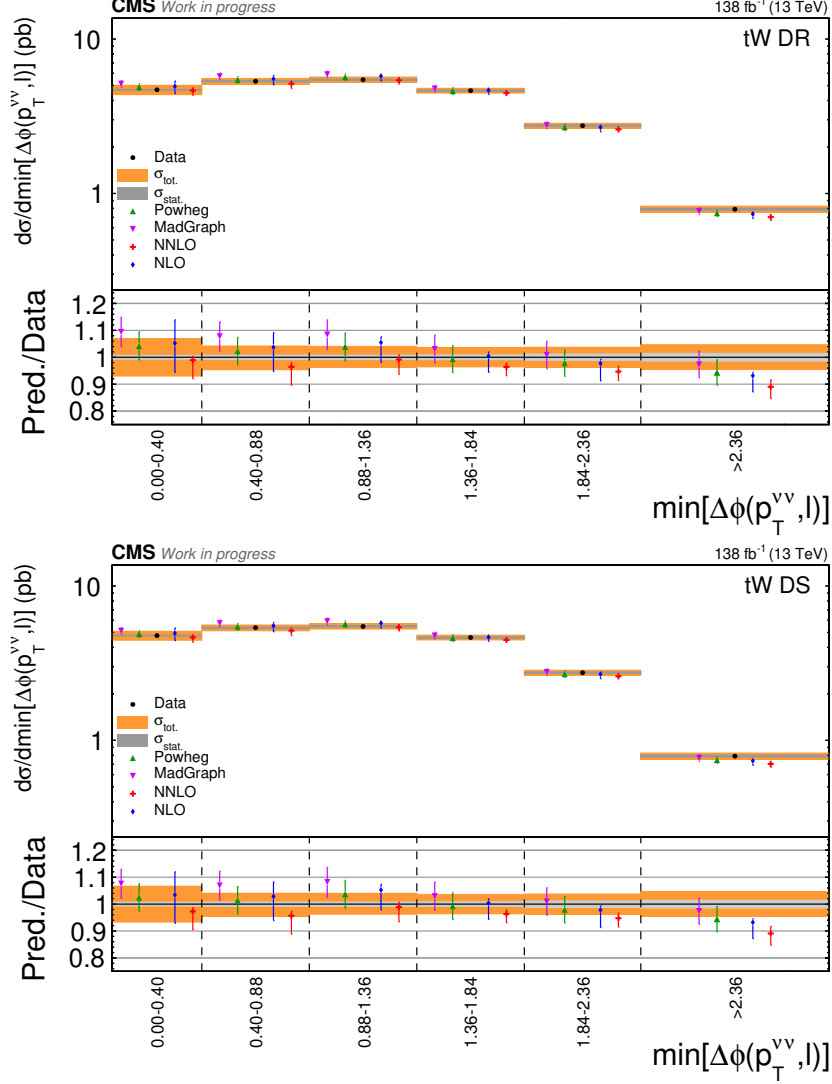


Figure 10.5: Differential cross section results as a function of $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ based on the DR (top) and the DS scheme (bottom). The total (statistical) uncertainty on the measurement is shown as an orange (grey) band. The results are compared to NLO (blue) and NNLO (red) fixed-order predictions as well as predictions from the POWHEG (green) and MADGRAPH (purple) generators. The lower parts show the ratio between the different predictions and the measurement.

The corresponding normalized differential cross sections are shown in Fig. 10.6. As already discussed in Sec. 9.3, the systematic uncertainties of the measurements are significantly reduced compared to the absolute cross section results. The same holds for the uncertainties of the four theory predictions. Unlike the situation for the absolute measurements, in both normalized measurements, all four predictions show a significant tension to the measured data in the last bin of the spectra. Differences between the measured shapes and predicted shapes are still present, but for the normalized measurements, the best description of the measured shapes

seems to be provided by the POWHEG and the NNLO fixed-order prediction. The results of the χ^2 tests (cf. Tab. 10.2), indicate that the shape differences are at least partially covered by the correlations present in the measurements.

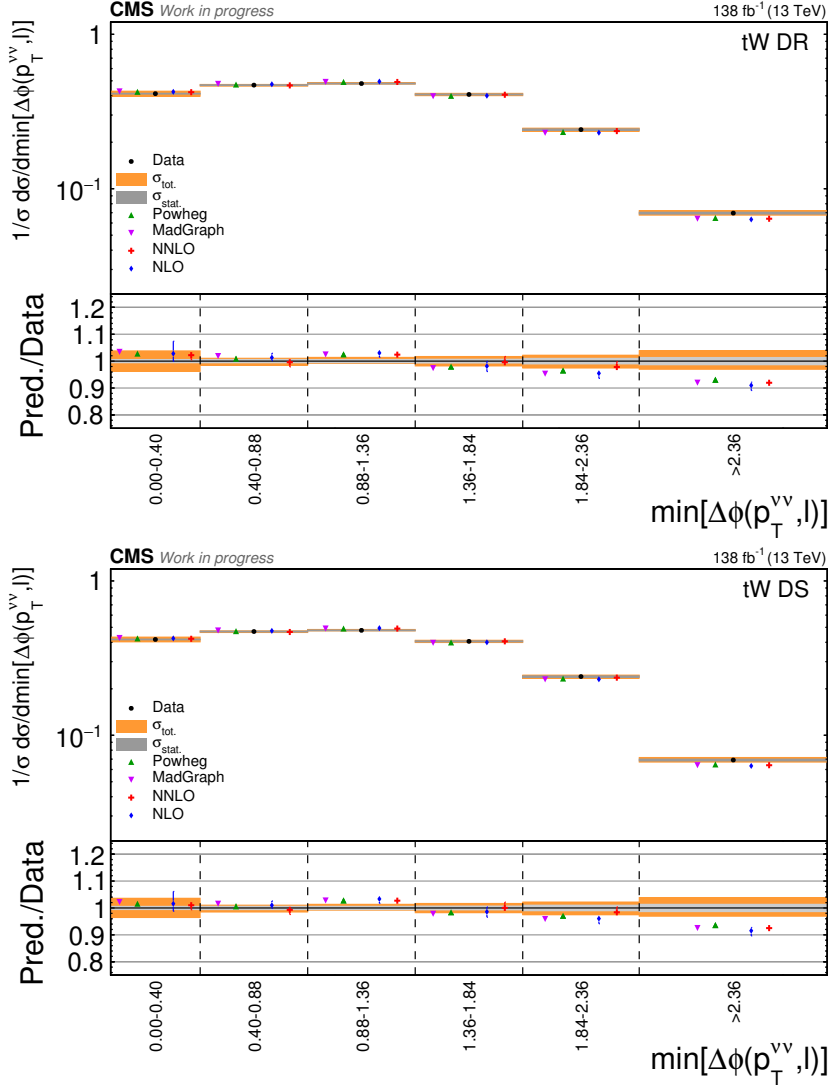


Figure 10.6: Normalized differential cross section results as a function of $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ based on the DR (top) and the DS scheme (bottom). The total (statistical) uncertainty on the measurement is shown as an orange (grey) band. The results are compared to NLO (blue) and NNLO (red) fixed-order predictions as well as predictions from the POWHEG (green) and MADGRAPH (purple) generators. The lower parts show the ratio between the different predictions and the measurement.

The shape differences observed for the normalized and absolute measurements as a function of $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ match the observations made in earlier differential cross section measurements of the dileptonic $t\bar{t}$ process based on the azimuthal angle $\Delta\phi(\ell, \ell)$ between both leptons performed by ATLAS [223] and CMS [224]. Here, the measured results exceed various predictions, which include the two MC-based predictions used in this analysis, at small angular separations between both leptons. Given that the phase space of large $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ is mostly populated by events with a small distance between both leptons, as discussed in Sec. 7.1, the tensions observed in the $\Delta\phi(\ell, \ell)$ measurements match the discrepancies observed in the present results.

10.2 Two-dimensional cross sections

The detector-level distributions used in this analysis, which are the most sensitive to potential BSM contributions, correspond to the two-dimensional combination of the DNN p_T^{miss} and the $\min[\Delta\phi(p_T^{\text{miss}}, \ell)]$ observable. The comparisons between observed data and prediction for these particular distributions can be found in Fig. 10.7. For the majority of the measured range, a good agreement between data and simulation can be observed in both detector-level distributions taking the uncertainty of the prediction into account. For the distribution using the DR scheme,

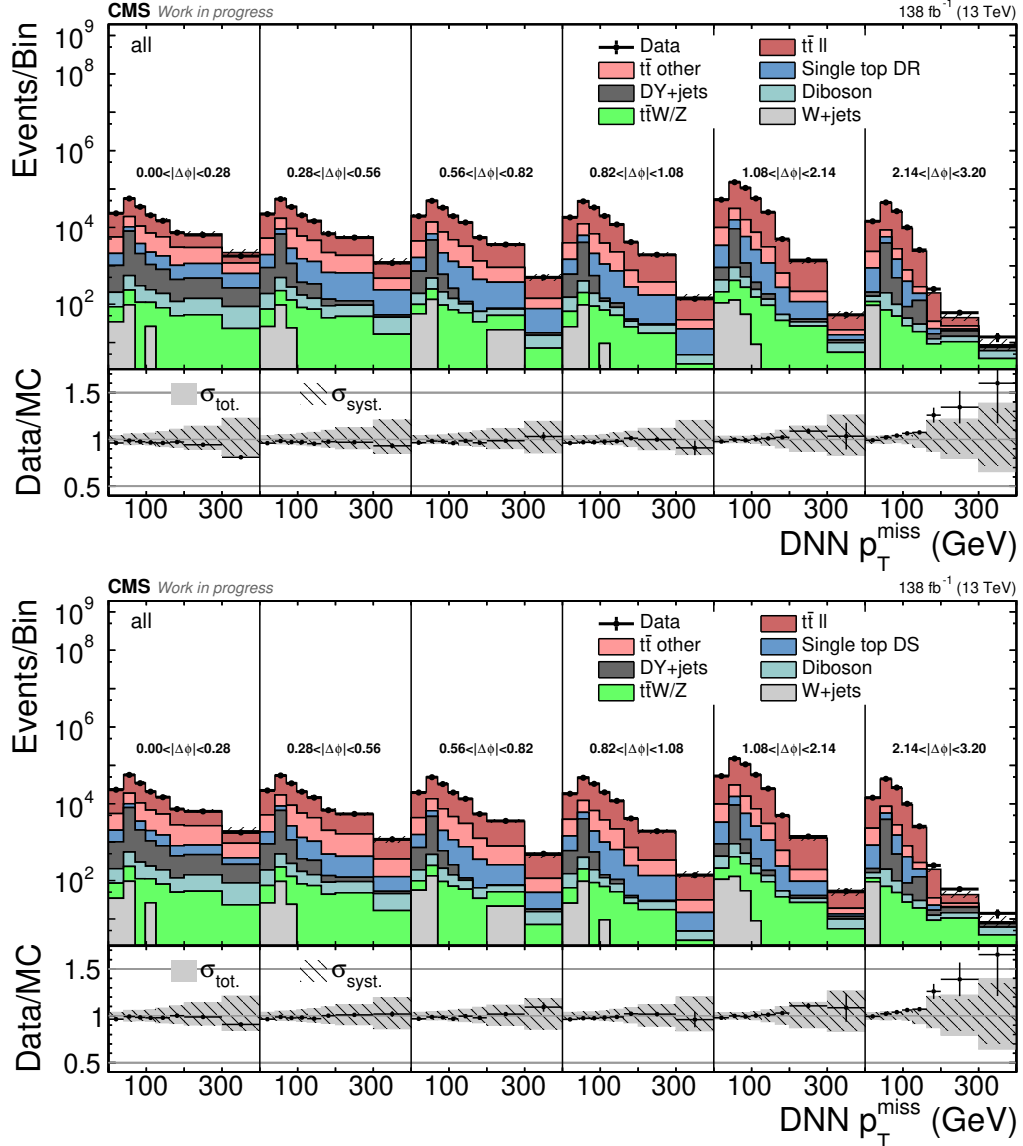


Figure 10.7: Detector-level distributions as a function of $\min[\Delta\phi(p_T^{\text{miss}}, \ell)]$ and $\text{DNN } p_T^{\text{miss}}$ for all channels and data-taking periods combined based on the DR (top) and the DS scheme (bottom). The vertical lines indicate bins in $\min[\Delta\phi(p_T^{\text{miss}}, \ell)]$. The upper parts of the plots show the observed data (black) and the contributions predicted by simulation including the total uncertainty (hatched area). The lower parts show the ratio between data and simulation as well as the systematic and total uncertainties on the predicted distribution. The last p_T^{miss} bin in each $\min[\Delta\phi(p_T^{\text{miss}}, \ell)]$ bin includes all events above the plotted range.

the predicted p_T^{miss} spectrum in the first two angular bins is slightly harder than the spectrum observed in data, which matches the observation made in the one-dimensional DNN p_T^{miss} distribution, shown in Fig. 10.1. This difference is not present in the distribution based on the DS scheme, which yields good agreement between data and simulation in the first two angular bins. For both distributions, the shape of the p_T^{miss} spectrum from the third to the fifth angular bin is well described by the simulation. The largest difference between data and simulation in both distributions can be found in the highest angular bin, where a small excess in data is observed in the high- p_T^{miss} region. However, this excess is still within the uncertainties due to the large statistical uncertainties in the given phase space. Thus, even though this particular phase space is expected to have the largest sensitivity to the BSM signatures, discussed in Sec. 2.3.3, no significant indications for potential BSM contributions are found in the given detector-level distributions.

Based on the conclusion drawn from the detector-level distributions, two-dimensional differential cross sections are derived as a function of $p_T^{\nu\nu}$ and $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$, which can be found in Fig. 10.8. For both measurements, all predictions yield a reasonable description of the measured spectra across most parts of the measured range. In the result using the DR scheme, a small difference regarding the $p_T^{\nu\nu}$ shape in the first angular bin, similar to the one found in the one-dimensional measurement, is observed for all four predictions. This difference seems to be maximal for the NLO prediction, while the POWHEG prediction yields the best description of the shape. As expected from the one-dimensional result, the description of the $p_T^{\nu\nu}$ shape is improved using the DS scheme, which is the only significant difference between the results using the two different schemes. In the second angular bin, both measured spectra are almost perfectly described by the NNLO calculation, whereas the remaining predictions fail to match the data in the first $p_T^{\nu\nu}$ bin. In addition, the MADGRAPH and NLO predictions tend to overpredict the measurements across the entire angular bin. Both spectra in the last angular bin are described reasonably well by all four predictions taking the corresponding uncertainties into account. However, the comparisons with the NNLO prediction indicate a small difference regarding normalization. Furthermore, in the last bin, the measured cross sections exceed all four predictions, even though still covered by the given uncertainties.

The results of the corresponding χ^2 tests, shown in Tab. 10.3, confirm the observations discussed above. The POWHEG prediction yields the best description of the measured spectra, followed by the NNLO calculation. Taking the uncertainties of the different predictions into account, all four predictions provide a good agreement with the measured spectra. The measurement based

Table 10.3: Results of the χ^2 tests (cf. Sec. 9.4) performed for the absolute and normalized 2D measurements between the measured cross section and the four predictions. The nominal results take only the covariance of the measurement into account, while the values in round brackets include the uncertainties on the predictions. The results for the measurements based on the DR and the DS scheme are shown. For each measurement, the degrees of freedom (dof) are given.

Measurement	dof	POWHEG	MADGRAPH	NNLO	NLO
2D (tW DR)	12	26.6 (5.6)	42.1 (10.9)	32.9 (6.1)	39.9 (4.0)
norm. 2D (tW DR)	11	23.4 (19.4)	37.2 (30.6)	27.4 (14.1)	35.6 (9.6)
2D (tW DS)	12	20.4 (4.4)	38.1 (8.7)	25.4 (4.4)	35.1 (3.4)
norm. 2D (tW DS)	11	19.3 (16.2)	34.4 (27.1)	23.4 (11.1)	33.3 (8.0)

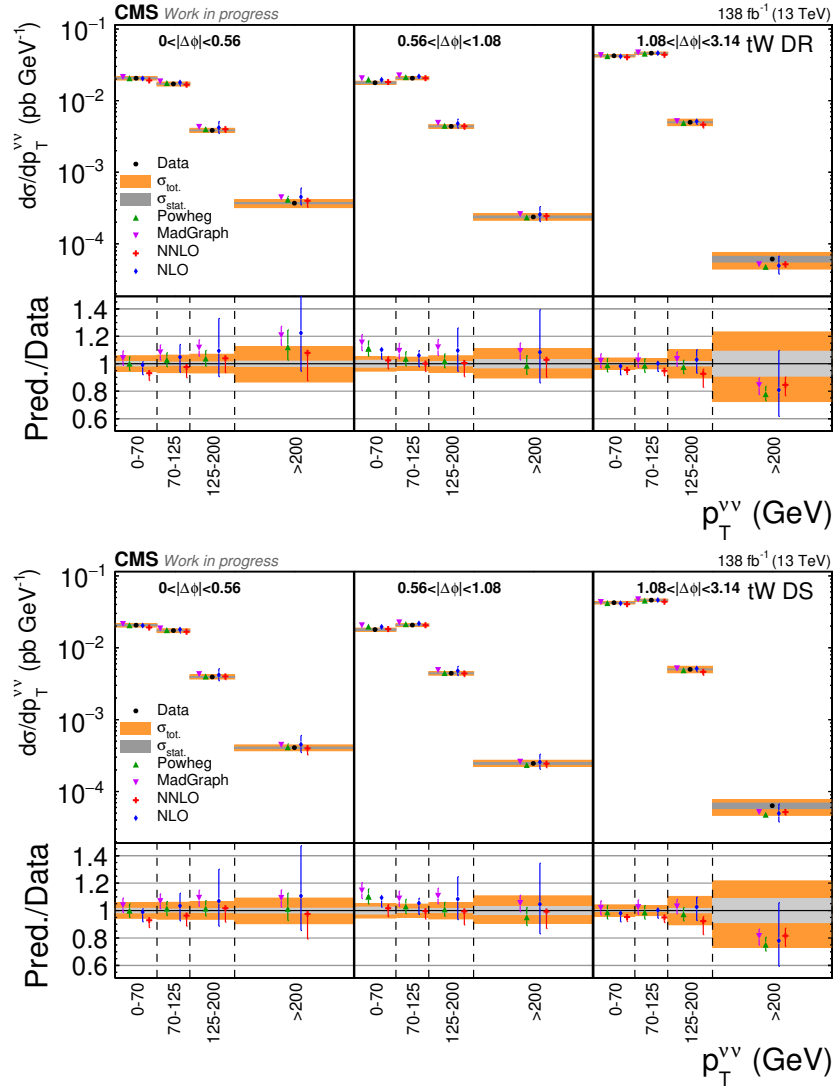


Figure 10.8: Differential cross section results as a function of $p_T^{\nu\nu}$ and $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ based on the DR (top) and the DS scheme (bottom). The vertical lines indicate bins in $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$. The total (statistical) uncertainty on the measurement is shown as an orange (grey) band. The results are compared to NLO (blue) and NNLO (red) fixed-order predictions as well as predictions from the POWHEG (green) and MADGRAPH (purple) generators. The lower parts show the ratio between the different predictions and the measurement.

on the DS scheme yields reduced χ^2 values mainly caused by the improved agreement in the first angular bin.

The results of the normalized differential cross sections for the two-dimensional measurements can be found in Fig. 10.9. Similar to the absolute measurements, a reasonable agreement between the predictions and both measurements is observed for most parts of the phase space. The conclusions drawn for the first angular bin of the absolute measurements mostly hold for the normalized measurements, where the result based on the DS scheme yields a better agreement with the four predictions. The most significant tensions between predictions and data in both measurements are observed in the first $p_T^{\nu\nu}$ bin of the second angular region, where fixed-order (MC-based) predictions yield a result, which is around 5 (10)% larger than the mea-

surements. In the last angular bin, all predictions slightly underestimate the measured results for most parts of the $p_T^{\nu\nu}$ spectra, where the best description of the highest $p_T^{\nu\nu}$ region is provided by the NNLO prediction.

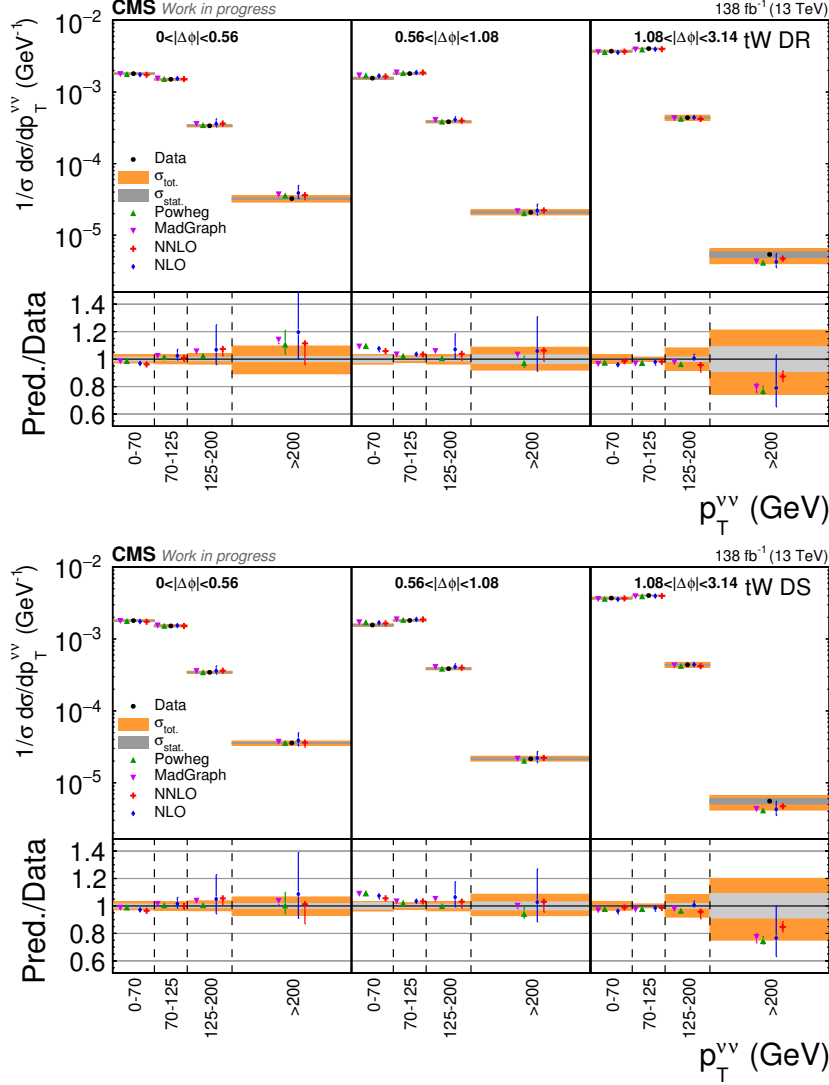


Figure 10.9: Normalized differential cross section results as a function of $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ and $p_T^{\nu\nu}$ based on the DR (top) and the DS scheme (bottom). The vertical lines indicate bins in $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$. The total (statistical) uncertainty on the measurement is shown as an orange (grey) band. The results are compared to NLO (blue) and NNLO (red) fixed-order predictions as well as predictions from the POWHEG (green) and MADGRAPH (purple) generators. The lower parts show the ratio between the different predictions and the measurement.

Taking the correlations of the normalized results as well as the uncertainties on the predictions into account, the results of the χ^2 tests, shown in Tab. 10.3, indicate a good overall agreement between the normalized measurements and three out of the four predictions. Only the MADGRAPH prediction leads to χ^2/dof results significantly larger than unity ($\chi^2/\text{dof} \sim 2.5$), which are the largest values observed among all other measurements and predictions. However, given the rather challenging observables and specific phase space studied by the two-dimensional measurements, all four predictions provide a high level of agreement with the measured spectra.

Chapter 11

Conclusions and outlook

In this thesis, an analysis aiming to combine the primary aspects of a BSM search and a SM precision measurement, focusing on pp-collision signatures with a top and an anti-top quark in the final state, was presented. The analysis relies on detector-level observables, which separate potential BSM contributions from the dileptonic $t\bar{t}$ process, given by the missing transverse momentum p_T^{miss} and the smallest azimuthal distance $\min[\Delta\phi(p_T^{\text{miss}}, \ell)]$ between the missing transverse momentum and a lepton. Unlike nominal BSM searches, which usually focus on specific phase space regions, in the present analysis, the entire spectrum of the given observables was measured enabling differential cross section measurements of the dileptonic $t\bar{t}$ process in case no significant BSM contributions were found. These measurements were performed as a function of particle-level observables, which match the detector-level observables for the SM $t\bar{t}$ process, namely the transverse momentum $p_T^{\nu\nu}$ of the dineutrino system and the minimal azimuthal distance $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ between the dineutrino system and a lepton. Additional cross section measurements were performed based on the two-dimensional combination of both particle-level observables. This analysis corresponds to the first measurement of the dileptonic $t\bar{t}$ process based on these two observables at the LHC.

The analysis is based on pp-collision data recorded by the CMS experiment between 2016 and 2018 at a center-of-mass energy of 13 TeV, which correspond to a total integrated luminosity of 138 fb^{-1} . Following the expected signatures for the dileptonic $t\bar{t}$ process, events were selected based on final states with two oppositely-charged leptons, two jets, out of which at least one jet was required to be identified with a b jet, and significant missing transverse momentum, originating from the two neutrinos. The results were derived based on the combination of three leptonic decay channels, given by the e^+e^- , $\mu^+\mu^-$ and $e^\pm\mu^\mp$ channels.

Since both detector-level observables employed throughout the analysis rely on a good reconstruction of the missing transverse momentum, dedicated studies of its response and resolution were performed showing a reduced performance of the p_T^{miss} reconstruction at large p_T^{miss} and large $\min[\Delta\phi(p_T^{\text{miss}}, \ell)]$. To account for this effect and to improve the overall performance of the p_T^{miss} reconstruction for dileptonic $t\bar{t}$ events, an additional p_T^{miss} correction based on a deep neural network regression was developed within the scope of this thesis. Using a feed-forward neural network with two output nodes enabled the simultaneous correction of the magnitude and the direction of the missing transverse momentum. Dedicated studies of the impact and the description in simulation for various potential input features were conducted. Furthermore, an optimization of the hyperparameters was performed, yielding a final network, which was able to improve the overall resolution of the magnitude and the azimuthal angle by 15% and 11.5%, respectively. Due to an additional reweighing of the training sample, this gain in performance was achieved for a wide p_T^{miss} range including the high- p_T region above $\sim 200 \text{ GeV}$, which was of main

interest for the present analysis. One of the main impacts from the additional p_T^{miss} correction was observed in the optimization of the binnings used for the final cross section measurements. For this optimization, dedicated algorithms were used to maximize the number of measuring points in the final results while keeping migration effects induced by the detector resolution at a reasonable level and ensuring a small statistical uncertainty. Within the optimization, binnings with up to three additional bins were realized based on the DNN p_T^{miss} correction compared to relying on the nominal PUPPI reconstruction.

The differential cross section results for the dileptonic $t\bar{t}$ process were derived based on the corresponding detector-level distributions after contributions from other SM processes were removed using an unfolding procedure based on an unregularized least square method. This choice of unfolding method was supported by different validation tests, in which the statistical properties and the potential bias introduced by differences in the response between observed data and simulation were studied. Within these tests, the best performance was observed for the selected method compared to a regularized least square approach and a simple bin-by-bin unfolding.

Initial studies of the tW background process in NLO simulation, where the overlap between the simulated samples for the tW and $t\bar{t}$ process is considered using the diagram removal or the diagram subtraction scheme, showed significant differences in the generated p_T^{miss} spectrum between both schemes. To account for these differences and to study their potential impact on the differential cross section results, two sets of results using the different schemes for the tW description in simulation were produced.

Before extracting the differential cross sections, all three detector-level distributions were studied concerning potential BSM contributions after all analysis ingredients had been fixed. Here, a reasonable agreement between the observed data and the SM prediction based on MC simulation was observed for all distributions. Small tensions between the data and prediction were found in the p_T^{miss} distribution using the DR scheme at high p_T^{miss} and in the 2D distributions, at high p_T^{miss} and large angular separation. In addition, the $\min[\Delta\phi(p_T^{\text{miss}}, \ell)]$ distributions revealed small shape differences between data and prediction. However, all of these differences are well covered by the corresponding uncertainties. Thus, given that no significant contributions from BSM physics were observed, all selected detector-level distributions were used to derive the corresponding differential cross sections on particle level for the dileptonic $t\bar{t}$ process.

The obtained differential cross section results were compared to two theory predictions based on MC simulations using NLO matrix element accuracy produced by the POWHEG V2 and the MADGRAPH5_AMC@NLO generator. In addition, two fixed-order calculations, which correspond to NLO and NNLO accuracy in QCD, were compared to the cross section results.

For the $p_T^{\nu\nu}$ measurements, all four theory predictions provide a good description of the measured spectra up to the region of around 300 GeV taking the uncertainties of the measurements into account, which range between 5 and 40% for the absolute results. Above this region, a softer spectrum compared to the predictions was observed for the result based on the DR scheme, which shows the smallest tensions with the NNLO fixed-order prediction at high $p_T^{\nu\nu}$. The result using the DS scheme, in contrast, yields a significantly improved agreement between measurement and predictions in the high- $p_T^{\nu\nu}$ region. Thus, in this particular measurement, the data seem to prefer the description based on the diagram subtraction scheme. The best overall description of both measured $p_T^{\nu\nu}$ spectra is provided by the MC-based predictions.

The measured $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ distributions are reasonably described by all four predictions, although small tensions with respect to the shape of the spectrum were observed between data and all theory predictions, which predict a shift towards smaller angular separations. The best overall description of the measured spectra is provided by the POWHEG prediction. The precision of the angular measurements varies between 3.5 and 7.0% for the absolute result and is significantly improved in the normalized result, which yields an uncertainty of only 1.0 to 3.8%. No signifi-

cant differences between the results derived with the two different schemes for the $t\bar{t}W$ process in simulation were observed. The difference in shape, present in both results, matches observations made in earlier differential cross section measurements of the dileptonic $t\bar{t}$ process as a function of the azimuthal angle $\Delta\phi(\ell, \ell)$ between both leptons, which is strongly correlated to the $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ observable.

Similar conclusions to the one-dimensional results were also drawn for the two-dimensional measurements. Overall, the four theory predictions yield a reasonable description of the measured spectra. In the result using the DR scheme, the first angular bin shows similar tensions at high $p_T^{\nu\nu}$, as observed in the one-dimensional $p_T^{\nu\nu}$ measurements, which are reduced in the result based on the DS scheme. The $p_T^{\nu\nu}$ spectra in the second angular bin are precisely described by the NNLO fixed-order prediction. The cross sections measured in the highest $p_T^{\nu\nu}$ bin of the last angular region slightly exceed all four predictions in both results. Nevertheless, these small tensions are usually well covered by the present uncertainties taking the full covariance of the result into account. The only exception is the MADGRAPH prediction, where the agreement in the normalized 2D measurements is slightly worse compared to the remaining three predictions. The precision achieved in the two-dimensional measurements ranges between 3.5 and 27.5% for the absolute cross section and is only slightly improved in the normalized results.

In conclusion, given the rather exotic observables and phase space regions considered in this analysis, remarkable agreement between the different theory predictions and the measured differential cross sections was observed. In the measurements including the $p_T^{\nu\nu}$ observable, the description of the $t\bar{t}W$ background based on the DS scheme seemed to be preferred by the measured data. For most of the measurements, the best overall description is provided by the POWHEG prediction. However, the differences between the four predictions are mostly small, such that none of the predictions is significantly disfavored by the measured differential cross sections.

The LHC is currently conducting the so-called Run 3, colliding protons at a center-of-mass energy of 13.6 TeV. For this run period, which is planned to be completed by the end of 2025, a recorded data set with an integrated luminosity roughly twice the integrated luminosity of the data set analyzed in the present analysis is expected. Given the larger expected data set and the slightly increased collision energy, which will lead to a larger production rate of $t\bar{t}$ events, smaller improvements regarding the statistical precision could be achieved based on a new iteration of the present analysis using the data set recorded during Run 3. Larger improvements can be expected for the era of the High-Luminosity LHC (HL-LHC), which is scheduled to start operation towards the end of the current decade after a longer upgrade phase of the detectors and the collider. While the center-of-mass energy is planned to only slightly increase to 14 TeV, the instantaneous luminosity is expected to increase by a factor of more than three reaching peak luminosities of around $7.5 \cdot 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. Thus, by the end of the expected HL-LHC lifetime in the 2040s, a combined data set of the order of 3000 to 4000 fb^{-1} is expected to be recorded by each of the two multi-purpose detectors. Repeating the present analysis based on this data set would allow for significant improvements. Due to the vast increase in statistical precision, more bins and therefore more measuring points in the currently investigated phase space, which can be compared to theory predictions, will become available. In addition, given that the high- $p_T^{\nu\nu}$ bins of the current analysis show non-negligible statistical uncertainties, the $p_T^{\nu\nu}$ range studied in the analysis could be enlarged, leading to an increased sensitivity to potential BSM contributions. Furthermore, additional BSM-sensitive observables could be introduced to derive measurements in three or even four dimensions, which would again significantly push the sensitivity to BSM contributions but also to specific aspects of the SM $t\bar{t}$ process. The increase in instantaneous luminosity, however, will also lead to significantly larger pileup, which is expected to reach up to 200 simultaneous collisions per bunch crossing. Given that the present analysis strongly relies

on a good resolution of the missing transverse momentum, which will decrease due to the larger pileup, new techniques for its reconstruction and proper corrections, like the machine learning approach presented in this thesis, have to be developed for a potential new iteration of the analysis.

Apart from the opportunities provided by the HL-LHC, one of the potential key improvements regarding the present analysis is connected to the tW process. Contributions from this process, on the one hand, dominate the SM background as well as the systematic uncertainty at high $p_T^{\nu\nu}$. On the other hand, in simulation, there is no exact solution to handle the overlap between the $t\bar{t}$ process and the tW process at NLO if both processes are considered to be independent. In addition, especially the $p_T^{\nu\nu}$ measurement was shown to be highly sensitive to the description of the tW process, which led to the study of two different sets of cross section results. All of these aspects could be solved by adapting the analysis towards a simultaneous measurement of the $t\bar{t}$ and tW processes. With new MC generators, which describe both processes simultaneously including proper interference and off-shell effects, becoming available, this adaption could be an additional motivation for a new iteration of the analysis based on the data set recorded during Run 3 of the LHC.

Appendix 1: Signal and baseline triggers

The triggers used to collect the pp-collision data in each of the four data-taking periods including their corresponding run eras, which are usually defined by a period of stable data-taking conditions, are shown in Tab. A1.1-A1.3. For each decay channel, different sets of triggers, which are combined with a logical "or", are employed. In addition to the lepton triggers, various baseline p_T^{miss} triggers are employed to select an unbiased set of events used in the measurements of the trigger efficiencies.

Table A1.1: List of triggers used to collect pp-collision data in the 2016preVFP and 2016postVFP data-taking periods, including their corresponding channels, eras and trigger paths.

Channel	Eras	Path
e^+e^-	B-H	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v
	B-H	HLT_Ele27_WPTight_Gsf_v
$\mu^+\mu^-$	B-H	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_v
	B-H	HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ_v
	B-G	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_v
	B-G	HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_v
	B-H	HLT_IsoMu24_v
	B-H	HLT_IsoTkMu24_v
$e^\pm\mu^\mp$	B-H	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v
	B-H	HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v
	B-G	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_v
	B-G	HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_v
	B-H	HLT_IsoMu24_v
	B-H	HLT_IsoTkMu24_v
	B-H	HLT_Ele27_WPTight_Gsf_v
Baseline	B-H	HLT_PFHT300_PFMET110_v
	B-H	HLT_PFMET120_PFMHT120_IDTight_v
	B-H	HLT_PFMET170_HBHECleaned_v
	B-H	HLT_PFMET300_v
	B-H	HLT_PFMETNoMu120_PFMHTNoMu120_IDTight_v
	B-H	HLT_MET200_v

Table A1.2: List of triggers used to collect pp-collision data in the 2017 data-taking period, including their corresponding channels, eras and trigger paths.

Channel	Eras	Path
e^+e^-	B-F	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v
	B-F	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_v
	B-F	HLT_Ele32_WPTight_Gsf_L1DoubleEG_v
$\mu^+\mu^-$	B	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_v
	C-F	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8_v
	B-F	HLT_IsoMu27_v
$e^\pm\mu^\mp$	B-F	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_v
	B-F	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v
	B-F	HLT_Mu12_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v
	B-F	HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v
	B-F	HLT_IsoMu27_v
	B-F	HLT_Ele32_WPTight_Gsf_L1DoubleEG_v
Baseline	B-F	HLT_PFMET200_HBHECleaned_v
	B-F	HLT_PFMET200_HBHE_BeamHaloCleaned_v
	B-F	HLT_PFMETTypeOne200_HBHE_BeamHaloCleaned_v
	B-F	HLT_PFMET120_PFMHT120_IDTight_v
	B-F	HLT_PFMETNoMu120_PFMHTNoMu120_IDTight_v
	B-F	HLT_PFMET120_PFMHT120_IDTight_PFHT60_v
	B-F	HLT_PFMETNoMu120_PFMHTNoMu120_IDTight_PFHT60_v
	B-F	HLT_PFHT500_PFMET100_PFMHT100_IDTight_v
	B-F	HLT_PFHT700_PFMET85_PFMHT85_IDTight_v
	B-F	HLT_PFHT800_PFMET75_PFMHT75_IDTight_v

Table A1.3: List of triggers used to collect pp-collision data in the 2018 data-taking period, including their corresponding channels, eras and trigger paths.

Channel	Eras	Path
e^+e^-	A-D	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v
	A-D	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_v
	A-D	HLT_Ele32_WPTight_Gsf_v
$\mu^+\mu^-$	A-D	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8_v
	A-D	HLT_IsoMu24_v
$e^\pm\mu^\mp$	A-D	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_v
	A-D	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v
	A-D	HLT_Mu12_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v
	A-D	HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v
	A-D	HLT_IsoMu24_v
	A-D	HLT_Ele32_WPTight_Gsf_v
Baseline	B-F	HLT_PFMET200_HBHECleaned_v
	A-D	HLT_PFMET200_HBHE_BeamHaloCleaned_v
	A-D	HLT_PFMETTypeOne200_HBHE_BeamHaloCleaned_v
	A-D	HLT_PFMET120_PFMHT120_IDTight_v
	A-D	HLT_PFMETNoMu120_PFMHTNoMu120_IDTight_v
	A-D	HLT_PFMET120_PFMHT120_IDTight_PFHT60_v
	A-D	HLT_PFMETNoMu120_PFMHTNoMu120_IDTight_PFHT60_v
	A-D	HLT_PFHT500_PFMET100_PFMHT100_IDTight_v
	A-D	HLT_PFHT700_PFMET85_PFMHT85_IDTight_v
	A-D	HLT_PFHT800_PFMET75_PFMHT75_IDTight_v

Appendix 2: Primary data sets

The list of primary data sets (PDs) for all four data-taking periods used in the present analysis including the corresponding DBS paths can be found in Tab. A2.1-A2.4. The DOUBLEMUON, DOUBLEEG and MUONEG PDs contain the events selected by dilepton triggers. To increase the overall trigger efficiency, the SINGLEMUON and SINGLEELECTRON PDs are used, which contain events selected by single-lepton triggers. For the measurement of the trigger efficiencies, events selected by p_T^{miss} -triggers are used, which are included in the MET PD. In 2018, the DOUBLEEG and the SINGLEELECTRON PDs were combined into the EGAMMA PD.

Table A2.1: Primary data sets for the 2016preVFP data-taking period with the corresponding DBS paths. The asterisk symbol has to be replaced by MINIAOD. For eras with similar paths, the paths are summarized and the eras are denoted in round brackets.

Primary data set	DBS path
DOUBLEMUON	/DoubleMuon/Run2016B-ver2_HIPM_UL2016_MiniAODv2-v1/* /DoubleMuon/Run2016(C-F)-HIPM_UL2016_MiniAODv2-v1/*
DOUBLEEG	/DoubleEG/Run2016B-ver2_HIPM_UL2016_MiniAODv2-v3/* /DoubleEG/Run2016(C-F)-HIPM_UL2016_MiniAODv2-v1/*
MUONEG	/MuonEG/Run2016B-ver2_HIPM_UL2016_MiniAODv2-v2/* /MuonEG/Run2016(C-F)-HIPM_UL2016_MiniAODv2-v2/*
SINGLEMUON	/SingleMuon/Run2016B-ver2_HIPM_UL2016_MiniAODv2-v2/* /SingleMuon/Run2016(C-F)-HIPM_UL2016_MiniAODv2-v2/*
SINGLEELECTRON	/SingleElectron/Run2016B-ver2_HIPM_UL2016_MiniAODv2-v2/* /SingleElectron/Run2016(C-D)-HIPM_UL2016_MiniAODv2-v2/* /SingleElectron/Run2016E-HIPM_UL2016_MiniAODv2-v5/* /SingleElectron/Run2016F-HIPM_UL2016_MiniAODv2-v2/*
MET	/MET/Run2016B-ver2_HIPM_UL2016_MiniAODv2-v2/* /MET/Run2016(C-F)-HIPM_UL2016_MiniAODv2-v2/*

Table A2.2: Primary data sets for the 2016postVFP data-taking period with the corresponding DBS paths. The asterisk symbol has to be replaced by MINIAOD. For eras with similar paths, the paths are summarized and the eras are denoted in round brackets.

Primary data set	DBS path
DOUBLEMUON	/DoubleMuon/Run2016(F-G)-UL2016_MiniAODv2-v1/* /DoubleMuon/Run2016H-UL2016_MiniAODv2-v2/*
DOUBLEEG	/DoubleEG/Run2016(F-H)-UL2016_MiniAODv2-v1/*
MUONEG	/MuonEG/Run2016(F-H)-UL2016_MiniAODv2-v2/*
SINGLEMUON	/SingleMuon/Run2016(F-H)-UL2016_MiniAODv2-v2/*
SINGLEELECTRON	/SingleElectron/Run2016(F-H)-UL2016_MiniAODv2-v2/*
MET	/MET/Run2016(F-H)-UL2016_MiniAODv2-v2/*

Table A2.3: Primary data sets for the 2017 data-taking period with the corresponding DBS paths. The asterisk symbol has to be replaced by MINIAOD. For eras with similar paths, the paths are summarized and the eras are denoted in round brackets.

Primary data set	DBS path
DOUBLEMUON	/DoubleMuon/Run2017(B-D)-UL2017_MiniAODv2-v1/*
	/DoubleMuon/Run2017E-UL2017_MiniAODv2-v2/*
	/DoubleMuon/Run2017F-UL2017_MiniAODv2-v1/*
DOUBLEEG	/DoubleEG/Run2017B-UL2017_MiniAODv2-v1/*
	/DoubleEG/Run2017C-UL2017_MiniAODv2-v2/*
	/DoubleEG/Run2017(D-E)-UL2017_MiniAODv2-v1/*
	/DoubleEG/Run2017F-UL2017_MiniAODv2-v2/*
MUONEG	/MuonEG/Run2017(B-F)-UL2017_MiniAODv2-v1/*
SINGLEMUON	/SingleMuon/Run2017(B-F)-UL2017_MiniAODv2-v1/*
SINGLEELECTRON	/SingleElectron/Run2017(B-F)-UL2017_MiniAODv2-v1/*
MET	/MET/Run2017(B-F)-UL2017_MiniAODv2-v1/*

Table A2.4: Primary data sets for the 2018 data-taking period with the corresponding DBS paths. The asterisk symbol has to be replaced by MINIAOD. For eras with similar paths, the paths are summarized and the eras are denoted in round brackets.

Primary data set	DBS path
DOUBLEMUON	/DoubleMuon/Run2018(A-D)-UL2018_MiniAODv2-v1/*
EGAMMA	/EGamma/Run2018(A-C)-UL2018_MiniAODv2-v1/*
	/EGamma/Run2018D-UL2018_MiniAODv2-v2/*
MUONEG	/MuonEG/Run2018(A-D)-UL2018_MiniAODv2-v1/*
SINGLEMUON	/SingleMuon/Run2018A-UL2018_MiniAODv2-v3/*
	/SingleMuon/Run2018(B-C)-UL2018_MiniAODv2-v2/*
	/SingleMuon/Run2018D-UL2018_MiniAODv2-v3/*
MET	/MET/Run2018(A-B)-UL2018_MiniAODv2-v2/*
	/MET/Run2018(C-D)-UL2018_MiniAODv2-v1/*

Appendix 3: Simulated samples

The full list of simulated samples for SM and BSM processes including the cross sections used for the normalization can be found in Tab. A3.1. To derive the full path of the individual samples for the different simulation campaigns used in the four data-taking periods, the strings shown in Tab. A3.2 have to be included. The list of simulated samples of the $t\bar{t}$ process used to derive various systematic uncertainties, e.g., uncertainties regarding the underlying event or the color reconnection, is shown in Tab. A3.3.

Table A3.1: Simulated samples used throughout the analysis grouped by the corresponding process. For each sample, the DBS path, where " * " has to be replaced by the corresponding string listed in Tab. A3.2, and the cross section used to normalize the sample are given.

Process	Sample	σ (pb)
$t\bar{t}$	/TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8/*	89.05
	/TTToHadronic_TuneCP5_13TeV-powheg-pythia8/*	377.96
	/TTToSemiLeptonic_TuneCP5_13TeV-powheg-pythia8/*	366.91
	/TTJets_TuneCP5_13TeV-amcatnloFXFX-pythia8/*	831.76
DY+jets	/DYJetsToLL_M-50_TuneCP5_13TeV-amcatnloFXFX-pythia8/*	6077.22
	/DYJetsToLL_M-10to50_TuneCP5_13TeV-amcatnloFXFX-pythia8/*	18610
Single top	/ST_tW_antitop_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8/*	19.56
	/ST_tW_top_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8/*	19.56
	/ST_s-channel_4f_leptonDecays_TuneCP5_13TeV-amcatnlo-pythia8/*	10.32
	/ST_t-channel_top_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8/*	136.02
	/ST_t-channel_antitop_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8/*	80.95
Diboson	/WW_TuneCP5_13TeV-pythia8/*	118.7
	/WZ_TuneCP5_13TeV-pythia8/*	47.13
	/ZZ_TuneCP5_13TeV-pythia8/*	16.523
$t\bar{t}W/Z$	/TTZToQQ_TuneCP5_13TeV-amcatnlo-pythia8/*	0.5297
	/TTZToLLNuNu_M-10_TuneCP5_13TeV-amcatnlo-pythia8/*	0.2529
	/TTWJetsToLNu_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8/*	0.2043
	/TTWJetsToQQ_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8/*	0.4062
W+jets	/WJetsToLNu_TuneCP5_13TeV-madgraphMLM-pythia8/*	61526.7
BSM ($t\bar{t}$)	/SMS-T2tt_mStop-525_mLSP-350_TuneCP5_13TeV-madgraphMLM-pythia8/*	0.458

Table A3.2: DBS paths related to the simulation campaigns used for the different data-taking periods. In each path, the "+" has to be replaced by the corresponding version of the sample and "+" corresponds to MINIAOD, denoting the data format used throughout the analysis.

Period	Simulation campaign
2016preVFP	RunIISummer20UL16MiniAODAPV2-106X_mcRun2_asymptotic_preVFP_v11-v+/+
2016postVFP	RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v+/+
2017	RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v+/+
2018	RunIISummer20UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v+/+

Table A3.3: Simulated samples used to derive systematic uncertainties grouped by the uncertainty source. For each sample the DBS path, where " * " has to be replaced by the corresponding string listed in Tab. A3.2, is given. The cross sections used to normalize the samples match the values shown in Tab. A3.1.

Uncertainty	Sample
ME-PS matching	/TTTo2L2Nu_hdampDOWN_TuneCP5_13TeV-powheg-pythia8/*
	/TTTo2L2Nu_hdampUP_TuneCP5_13TeV-powheg-pythia8/*
	/TTToHadronic_hdampDOWN_TuneCP5_13TeV-powheg-pythia8/*
	/TTToHadronic_hdampUP_TuneCP5_13TeV-powheg-pythia8/*
	/TTToSemiLeptonic_hdampDOWN_TuneCP5_13TeV-powheg-pythia8/*
	/TTToSemiLeptonic_hdampUP_TuneCP5_13TeV-powheg-pythia8/*
Underlying event	/TTTo2L2Nu_TuneCP5down_13TeV-powheg-pythia8/*
	/TTTo2L2Nu_TuneCP5up_13TeV-powheg-pythia8/*
	/TTToHadronic_TuneCP5up_13TeV-powheg-pythia8/*
	/TTToHadronic_TuneCP5down_13TeV-powheg-pythia8/*
	/TTToSemiLeptonic_TuneCP5down_13TeV-powheg-pythia8/*
	/TTToSemiLeptonic_TuneCP5up_13TeV-powheg-pythia8/*
Color reconnection	/TTTo2L2Nu_TuneCP5_erd0N_13TeV-powheg-pythia8/*
	/TTTo2L2Nu_TuneCP5CR1_13TeV-powheg-pythia8/*
	/TTTo2L2Nu_TuneCP5CR2_13TeV-powheg-pythia8/*
	/TTToHadronic_TuneCP5_erd0N_13TeV-powheg-pythia8/*
	/TTToHadronic_TuneCP5CR1_13TeV-powheg-pythia8/*
	/TTToHadronic_TuneCP5CR2_13TeV-powheg-pythia8/*
	/TTToSemiLeptonic_TuneCP5_erd0N_13TeV-powheg-pythia8/*
	/TTToSemiLeptonic_TuneCP5CR1_13TeV-powheg-pythia8/*
Top-quark mass	/TTTo2L2Nu_mtop169p5_TuneCP5_13TeV-powheg-pythia8/*
	/TTTo2L2Nu_mtop175p5_TuneCP5_13TeV-powheg-pythia8/*
	/TTToHadronic_mtop169p5_TuneCP5_13TeV-powheg-pythia8/*
	/TTToHadronic_mtop175p5_TuneCP5_13TeV-powheg-pythia8/*
	/TTToSemiLeptonic_mtop169p5_TuneCP5_13TeV-powheg-pythia8/*
	/TTToSemiLeptonic_mtop175p5_TuneCP5_13TeV-powheg-pythia8/*

Appendix 4: Trigger efficiencies

The trigger efficiencies in data and simulation as a function of lepton p_T for the different data-taking periods can be found in Fig. A4.1-A4.4. The corresponding scale factors are shown in Fig. A4.5 and A4.6.

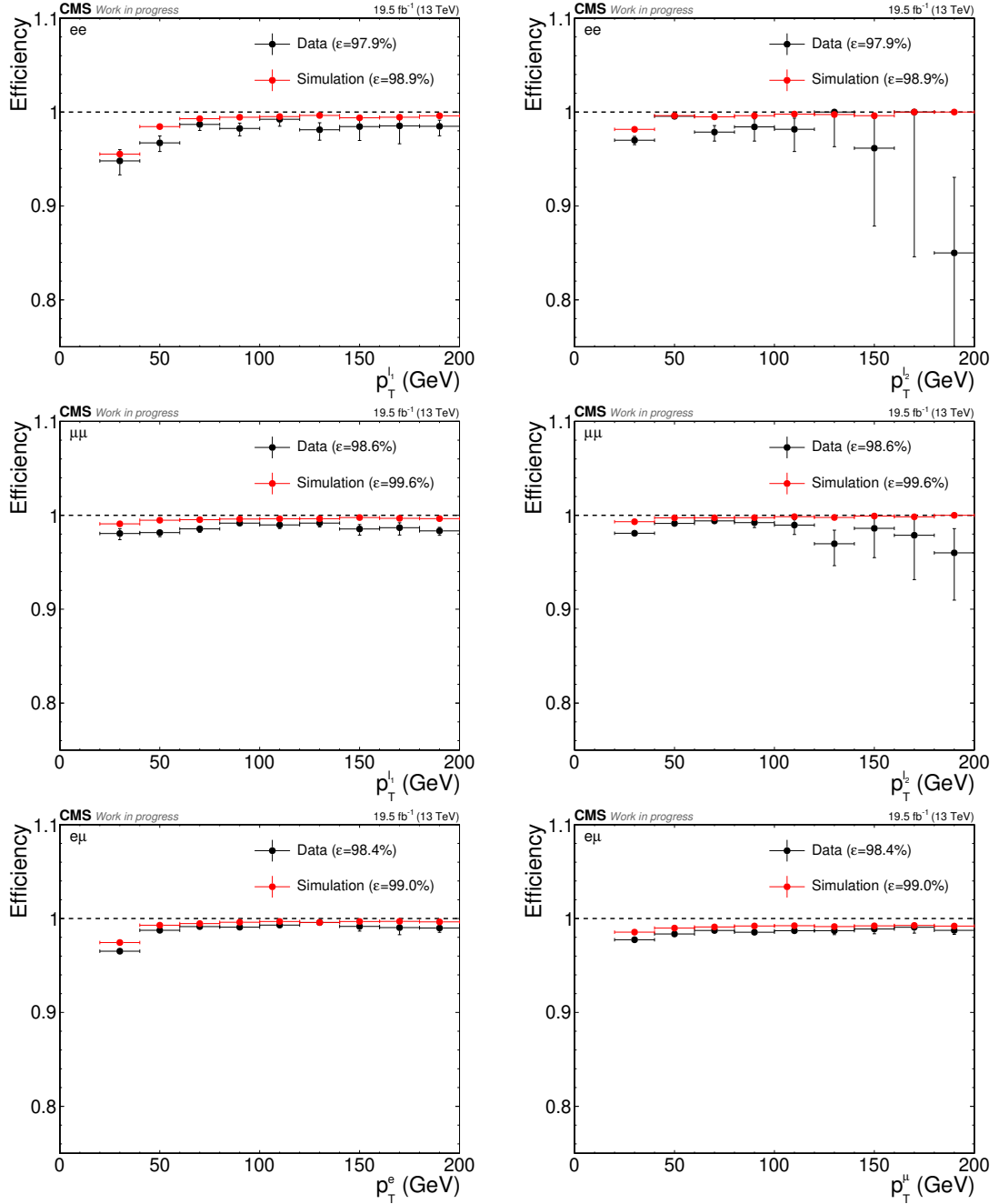


Figure A4.1: Signal trigger efficiencies for 2016preVFP measured in data (black) and simulation (red) for the e^+e^- (top), $\mu^+\mu^-$ (middle) and $e^\pm\mu^\mp$ (bottom) channel. The efficiencies are shown as a function of the (sub-)leading lepton p_T for the same-flavor channels and as a function of the electron and muon p_T for the $e^\pm\mu^\mp$ channel. The overall efficiencies are given in the legend. The vertical lines correspond to the asymmetric statistical uncertainties derived using Clopper-Pearson intervals [163].

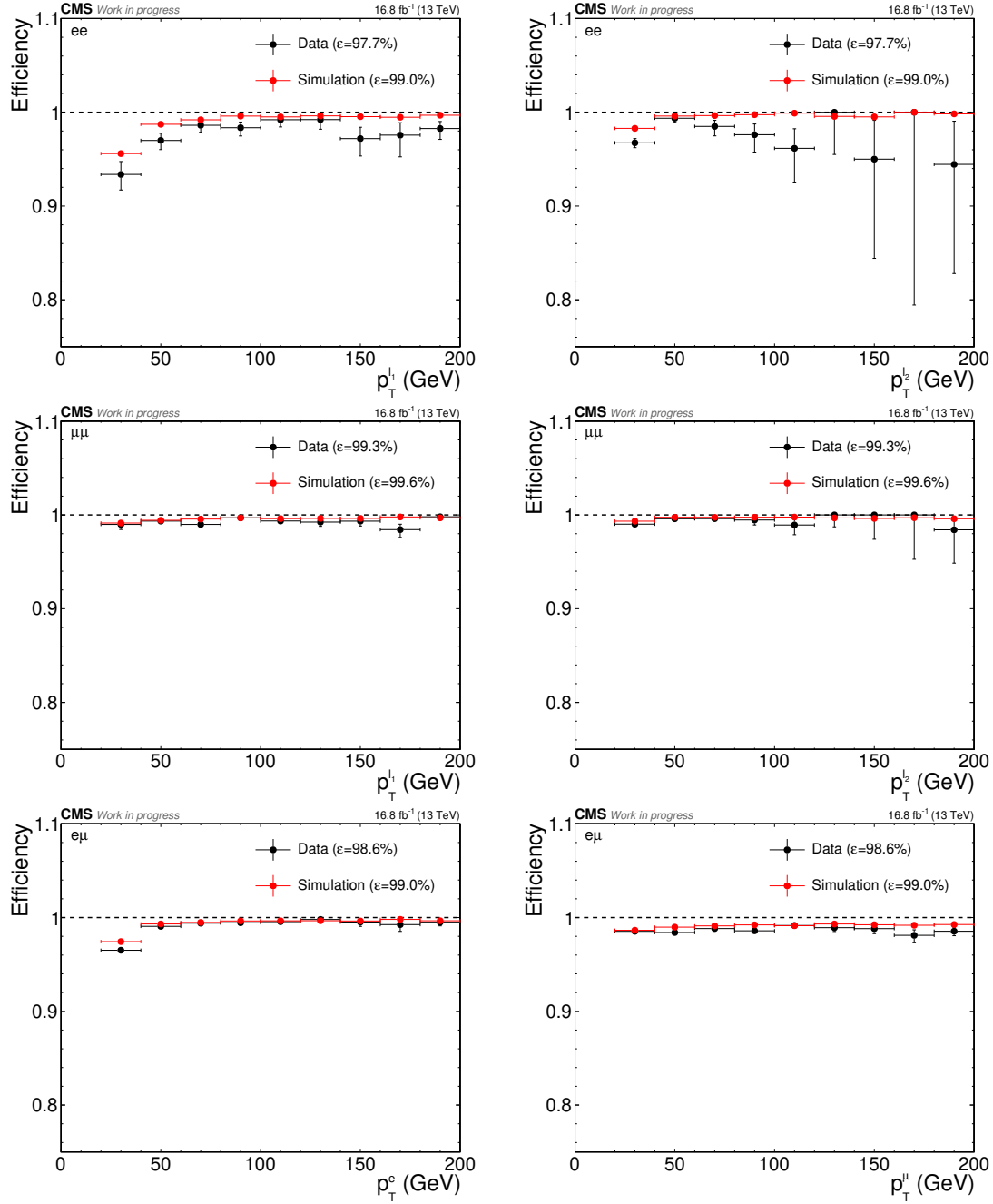


Figure A4.2: Signal trigger efficiencies for 2016postVFP measured in data (black) and simulation (red) for the e^+e^- (top), $\mu^+\mu^-$ (middle) and $e^\pm\mu^\mp$ (bottom) channel. The efficiencies are shown as a function of the (sub-)leading lepton p_T for the same-flavor channels and as a function of the electron and muon p_T for the $e^\pm\mu^\mp$ channel. The overall efficiencies are given in the legend. The vertical lines correspond to the asymmetric statistical uncertainties derived using Clopper-Pearson intervals [163]. The last bin in each distribution corresponds to an overflow bin.

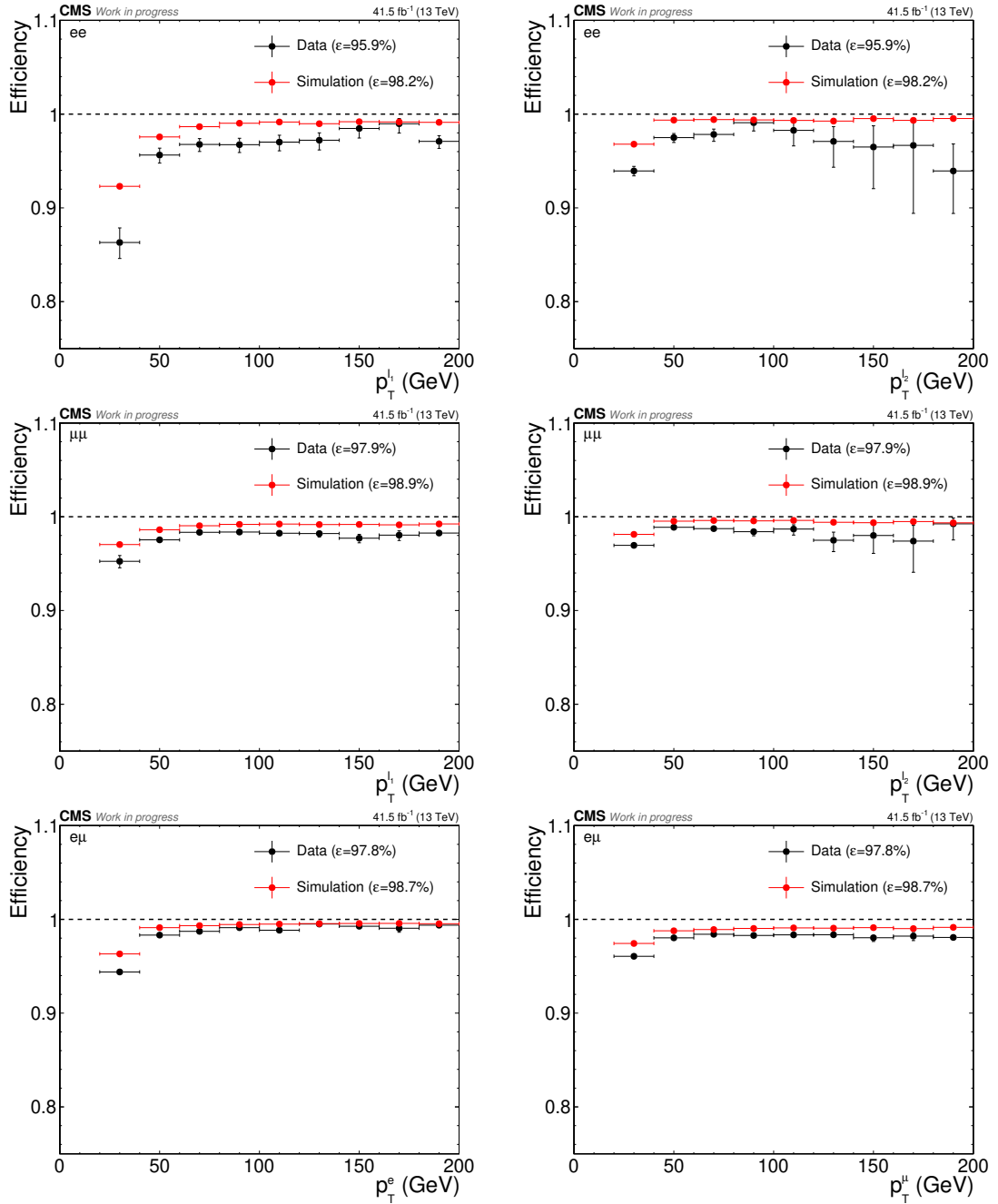


Figure A4.3: Signal trigger efficiencies for 2017 measured in data (black) and simulation (red) for the e^+e^- (top), $\mu^+\mu^-$ (middle) and $e^\pm\mu^\mp$ (bottom) channel. The efficiencies are shown as a function of the (sub-)leading lepton p_T for the same-flavor channels and as a function of the electron and muon p_T for the $e^\pm\mu^\mp$ channel. The overall efficiencies are given in the legend. The vertical lines correspond to the asymmetric statistical uncertainties derived using Clopper-Pearson intervals [163]. The last bin in each distribution corresponds to an overflow bin.

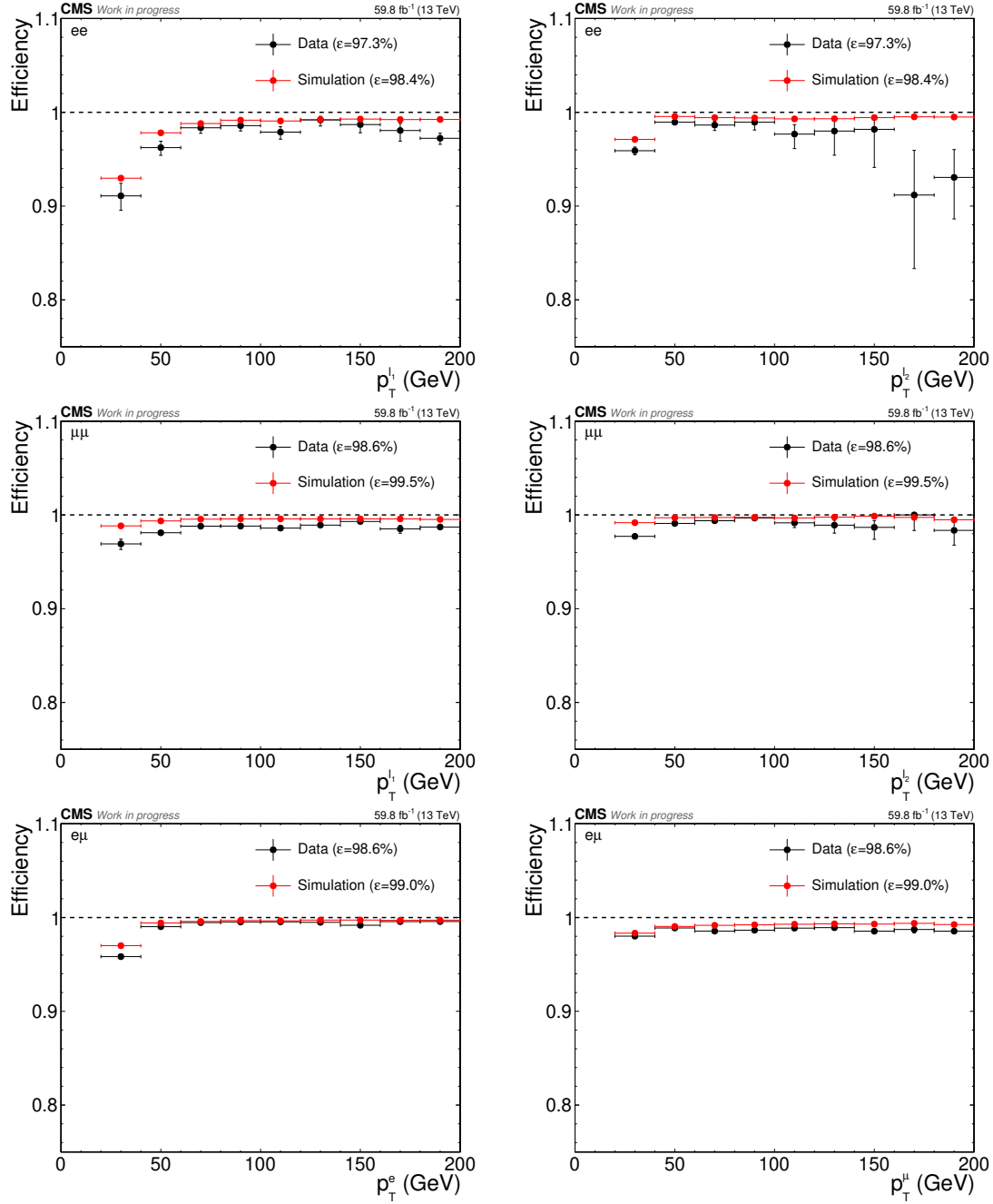


Figure A4.4: Signal trigger efficiencies for 2018 measured in data (black) and simulation (red) for the e^+e^- (top), $\mu^+\mu^-$ (middle) and $e^\pm\mu^\mp$ (bottom) channel. The efficiencies are shown as a function of the (sub-)leading lepton p_T for the same-flavor channels and as a function of the electron and muon p_T for the $e^\pm\mu^\mp$ channel. The overall efficiencies are given in the legend. The vertical lines correspond to the asymmetric statistical uncertainties derived using Clopper-Pearson intervals [163]. The last bin in each distribution corresponds to an overflow bin.

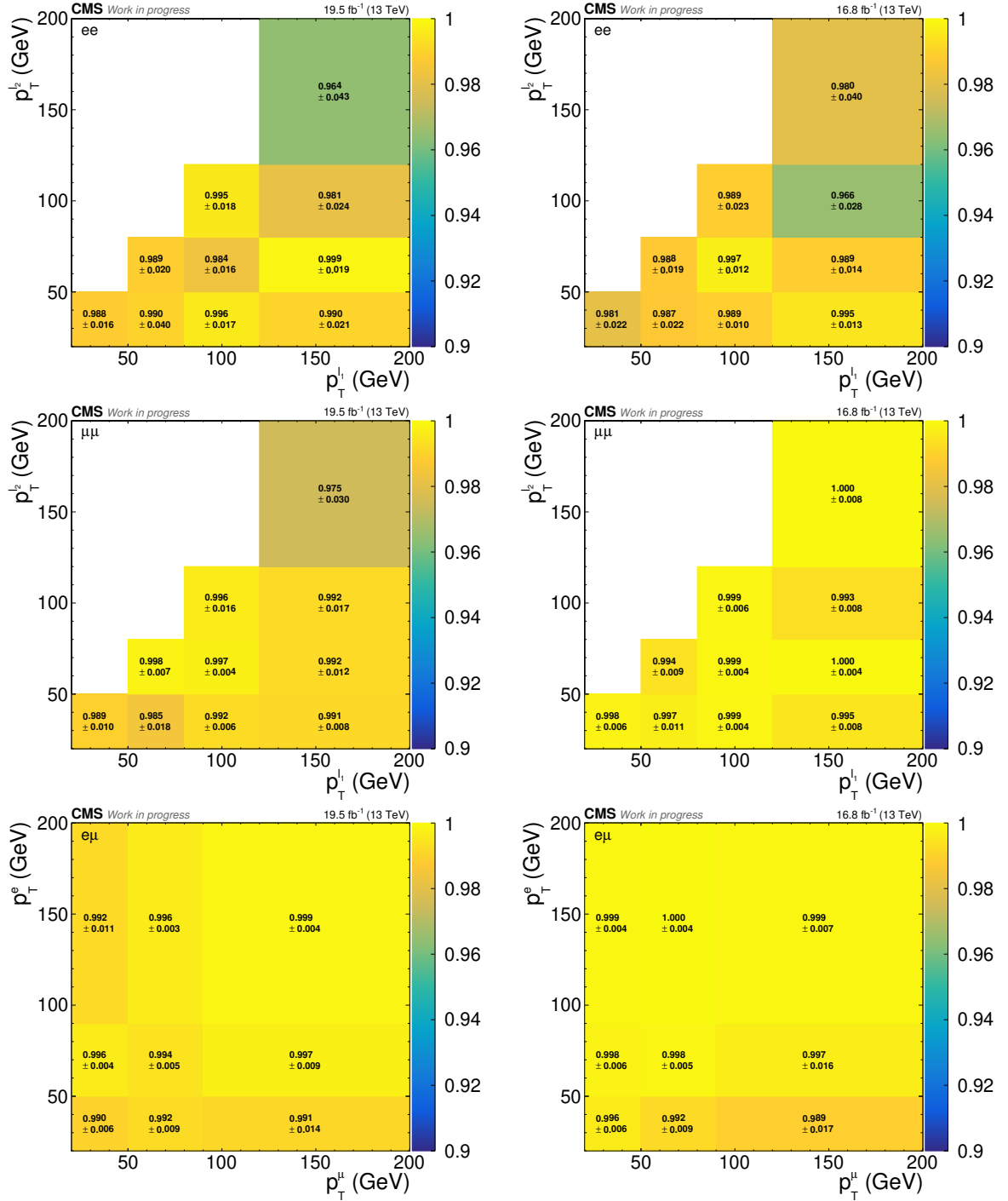


Figure A4.5: Trigger scale factors for 2016preVFP (left) and 2016postVFP (right) derived for the e^+e^- (top), $\mu^+\mu^-$ (middle) and $e^\pm\mu^\mp$ (bottom) channel. The SFs are shown as a function of leading and subleading lepton p_T or electron and muon p_T . The values of the SFs including their total uncertainties for each bin are visualized by black digits.

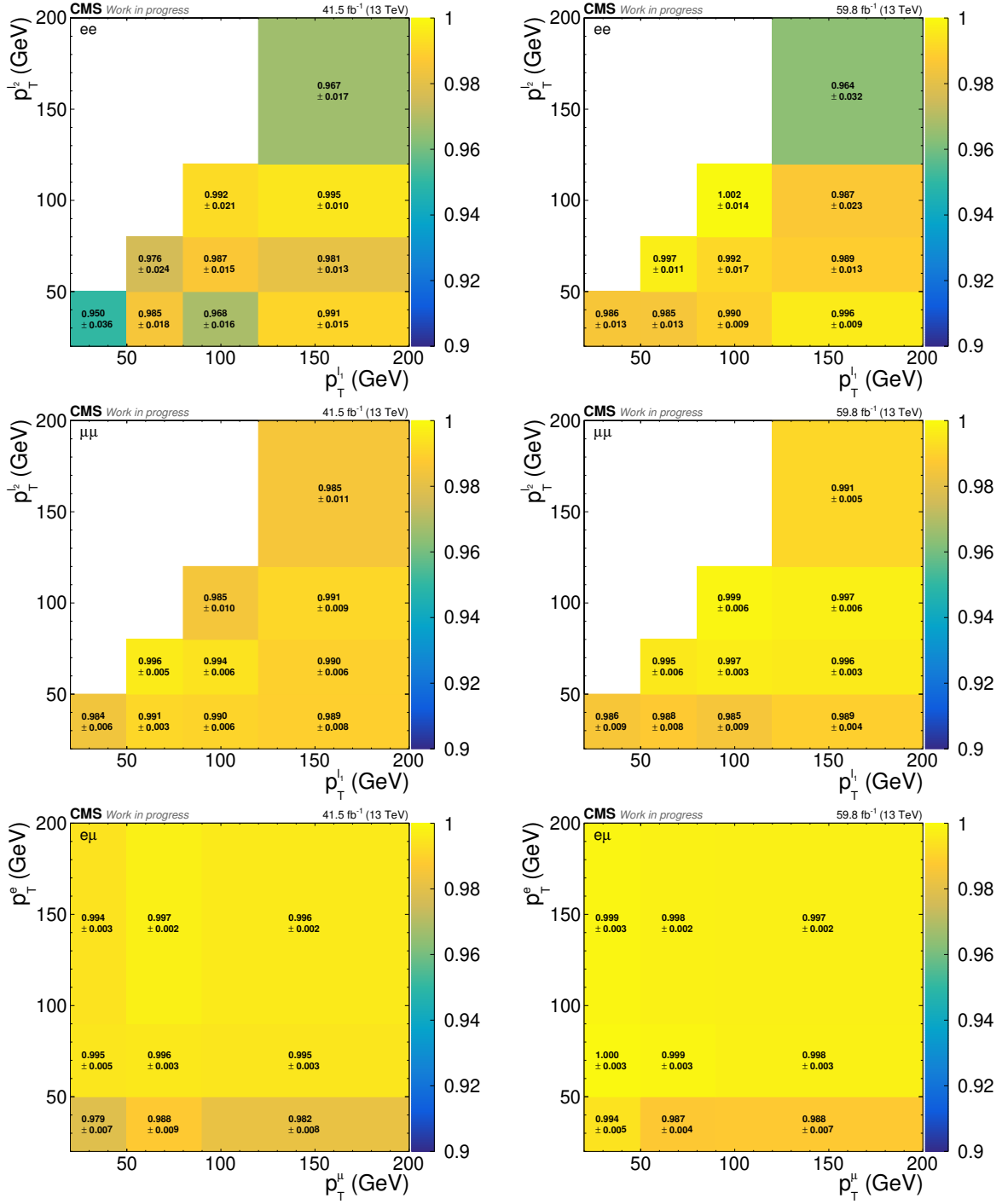


Figure A4.6: Trigger scale factors for 2017 (left) and 2018 (right) derived for the e^+e^- (top), $\mu^+\mu^-$ (middle) and $e^+\mu^-$ (bottom) channel. The SFs are shown as a function of leading and subleading lepton p_T or electron and muon p_T . The values of the SFs including their total uncertainties for each bin are visualized by black digits.

Appendix 5: Total event yields

The total yields for the individual data-taking periods are shown in Tab. A5.1-A5.4. The cutflow diagrams separating the individual decay channels can be found in Fig. A5.1.

Table A5.1: Total number of events per channel after the full event selection for the 2016preVFP data-taking period. The upper part of the table shows the individual contributions to the total event prediction derived from simulation (Sum MC). The ratio of the individual contributions and the total prediction in percent is shown in round brackets. The lower part of the table contains the uncertainty on the total prediction, which includes the systematic and statistical contributions, as well as the number of events in data and the ratio between data and simulation (Data/MC).

Channel		e^+e^-	$\mu^+\mu^-$	$e^\pm\mu^\mp$	all
Contribution	$t\bar{t}$ (ll)	16745.3(75.5)	32656.3 (73.9)	77093.6 (80.0)	126495.2 (77.7)
	$t\bar{t}$ other	2676.6(12.1)	5581.8 (12.6)	13028.6 (13.5)	21287.0 (13.1)
	Single top	1024.3(4.6)	1972.0 (4.5)	4553.1 (4.7)	7549.3 (4.6)
	DY+jets	1555.0(7.0)	3681.0 (8.3)	951.5 (1.0)	6187.5 (3.8)
	Diboson	100.9(0.5)	173.4 (0.4)	337.3 (0.4)	611.6 (0.4)
	$t\bar{t}W/Z$	68.4(0.3)	128.7 (0.3)	258.7 (0.3)	455.8 (0.3)
	W+jets	17.7(0.1)	13.3 (0.0)	143.2 (0.1)	174.1 (0.1)
Sum MC		22188.1	44206.5	96366.0	162760.6
σ (Sum MC)		1484.7	2138.2	3861.1	6914.2
Data		21123	42294	90844	154261
Data/MC [%]		95.2	95.7	94.3	94.8

Table A5.2: Total number of events per channel after the full event selection for the 2016postVFP data-taking period. The upper part of the table shows the individual contributions to the total event prediction derived from simulation (Sum MC). The ratio of the individual contributions and the total prediction in percent is shown in round brackets. The lower part of the table contains the uncertainty on the total prediction, which includes the systematic and statistical contributions, as well as the number of events in data and the ratio between data and simulation (Data/MC).

Channel		e^+e^-	$\mu^+\mu^-$	$e^\pm\mu^\mp$	all
Contribution	$t\bar{t}$ (ll)	14700.3(74.7)	28556.8 (74.3)	67535.4 (79.8)	110792.5 (77.6)
	$t\bar{t}$ other	2325.4(11.8)	4965.9 (12.9)	11494.6 (13.6)	18785.9 (13.2)
	Single top	912.3(4.6)	1716.0 (4.5)	4008.6 (4.7)	6636.8 (4.6)
	DY+jets	1570.9(8.0)	2924.9 (7.6)	843.8 (1.0)	5339.7 (3.7)
	Diboson	91.6(0.5)	156.2 (0.4)	313.9 (0.4)	561.6 (0.4)
	$t\bar{t}W/Z$	60.9(0.3)	112.4 (0.3)	229.4 (0.3)	402.7 (0.3)
	W+jets	28.6(0.1)	18.0 (0.0)	191.9 (0.2)	238.5 (0.2)
Sum MC		19689.8	38450.2	84617.6	142757.6
σ (Sum MC)		1195.4	1674.7	3344.0	5743.2
Data		19558	39231	83229	142018
Data/MC [%]		99.3	102.0	98.4	99.5

Table A5.3: Total number of events per channel after the full event selection for the 2017 data-taking period. The upper part of the table shows the individual contributions to the total event prediction derived from simulation (Sum MC). The ratio of the individual contributions and the total prediction in percent is shown in round brackets. The lower part of the table contains the uncertainty on the total prediction, which includes the systematic and statistical contributions, as well as the number of events in data and the ratio between data and simulation (Data/MC).

Channel	e^+e^-	$\mu^+\mu^-$	$e^\pm\mu^\mp$	all
$t\bar{t}$ (ll)	37078.9(74.5)	70806.2 (73.5)	170863.2 (79.9)	278748.2 (77.4)
$t\bar{t}$ other	5805.5(11.7)	12474.9 (13.0)	29234.1 (13.7)	47514.6 (13.2)
Single top	2256.2(4.5)	4249.7 (4.4)	10172.2 (4.8)	16678.1 (4.6)
DY+jets	4240.7(8.5)	8089.5 (8.4)	2064.1 (1.0)	14394.3 (4.0)
Diboson	225.5(0.5)	389.0 (0.4)	778.4 (0.4)	1392.8 (0.4)
$t\bar{t}W/Z$	151.6(0.3)	277.4 (0.3)	563.8 (0.3)	992.8 (0.3)
W+jets	0.0(0.0)	21.5 (0.0)	275.5 (0.1)	297.0 (0.1)
Sum MC	49758.3	96308.2	213951.4	360017.8
σ (Sum MC)	3490.0	4887.8	9618.6	16837.8
Data	48892	97892	210692	357476
Data/MC [%]	98.3	101.6	98.5	99.3

Table A5.4: Total number of events per channel after the full event selection for the 2018 data-taking period. The upper part of the table shows the individual contributions to the total event prediction derived from simulation (Sum MC). The ratio of the individual contributions and the total prediction in percent is shown in round brackets. The lower part of the table contains the uncertainty on the total prediction, which includes the systematic and statistical contributions, as well as the number of events in data and the ratio between data and simulation (Data/MC).

Channel	e^+e^-	$\mu^+\mu^-$	$e^\pm\mu^\mp$	all
$t\bar{t}$ (ll)	46989.2(74.6)	89042.2 (73.6)	215445.7 (79.8)	351477.1 (77.5)
$t\bar{t}$ other	7349.5(11.7)	15494.9 (12.8)	36639.8 (13.6)	59484.2 (13.1)
Single top	2925.9(4.6)	5433.1 (4.5)	12788.8 (4.7)	21147.8 (4.7)
DY+jets	5236.3(8.3)	10057.0 (8.3)	3007.7 (1.1)	18301.0 (4.0)
Diboson	282.1(0.4)	517.2 (0.4)	999.4 (0.4)	1798.8 (0.4)
$t\bar{t}W/Z$	195.9(0.3)	344.1 (0.3)	714.2 (0.3)	1254.2 (0.3)
W+jets	29.1(0.0)	22.3 (0.0)	271.0 (0.1)	322.5 (0.1)
Sum MC	63007.9	120910.9	269866.8	453785.6
σ (Sum MC)	4357.3	6588.8	12675.6	22346.5
Data	61995	122751	264739	449485
Data/MC [%]	98.4	101.5	98.1	99.1

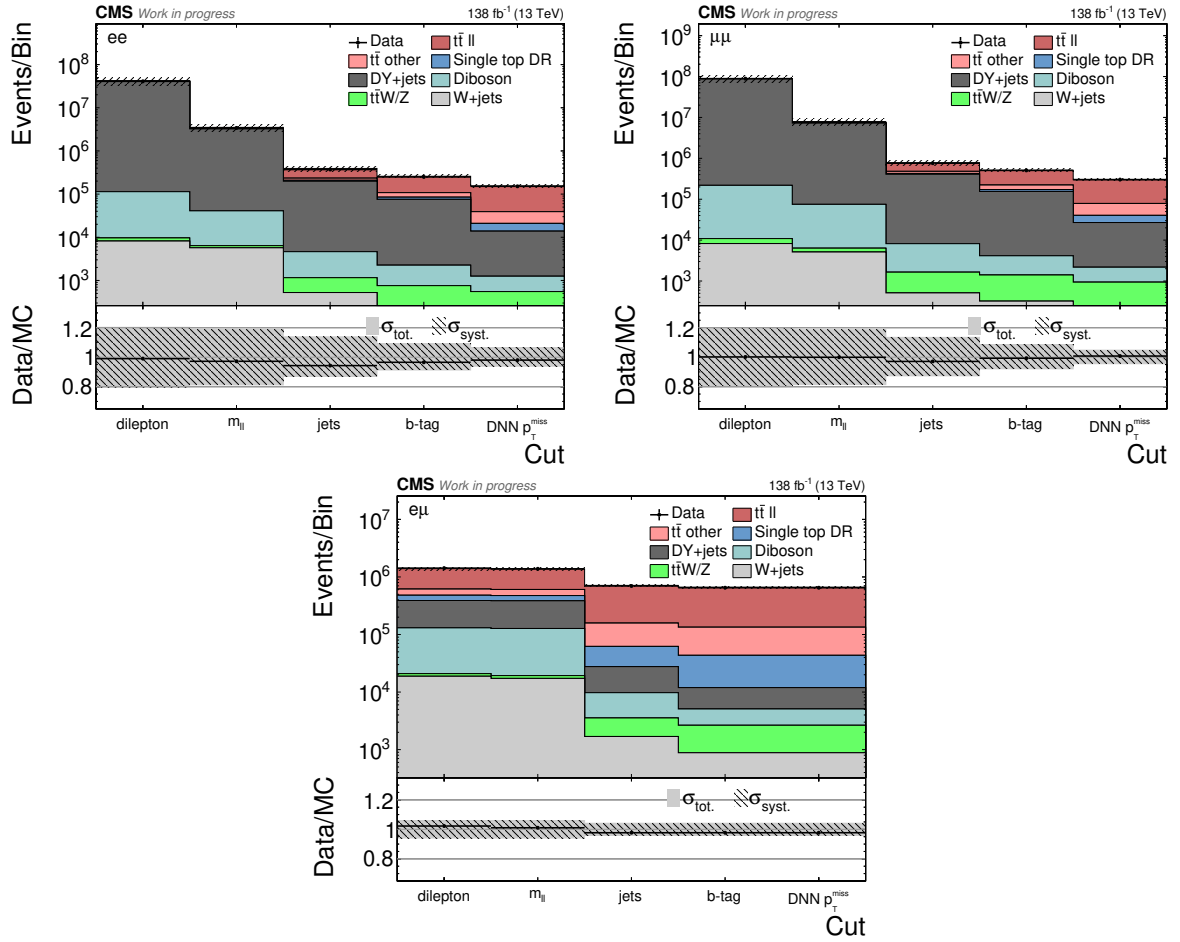


Figure A5.1: Cutflow diagrams showing the total number of events after different steps of the event selection in the e^+e^- (top left), $\mu^+\mu^-$ (top right) and $e^\pm\mu^\mp$ (bottom) channel for all data-taking periods combined. In the upper parts of the plots, the individual contributions to the yield predicted by simulation are shown including the total uncertainty (black-hatched area). The lower parts of the plots show the ratio between data and simulation as well as the systematic and total uncertainty on the predicted yields.

Appendix 6: Control distributions

The control distributions showing the leading and subleading lepton/jet p_T , the number of jets and b-tagged jets and the invariant mass for each decay channel are shown in Fig. A6.1-A6.4.

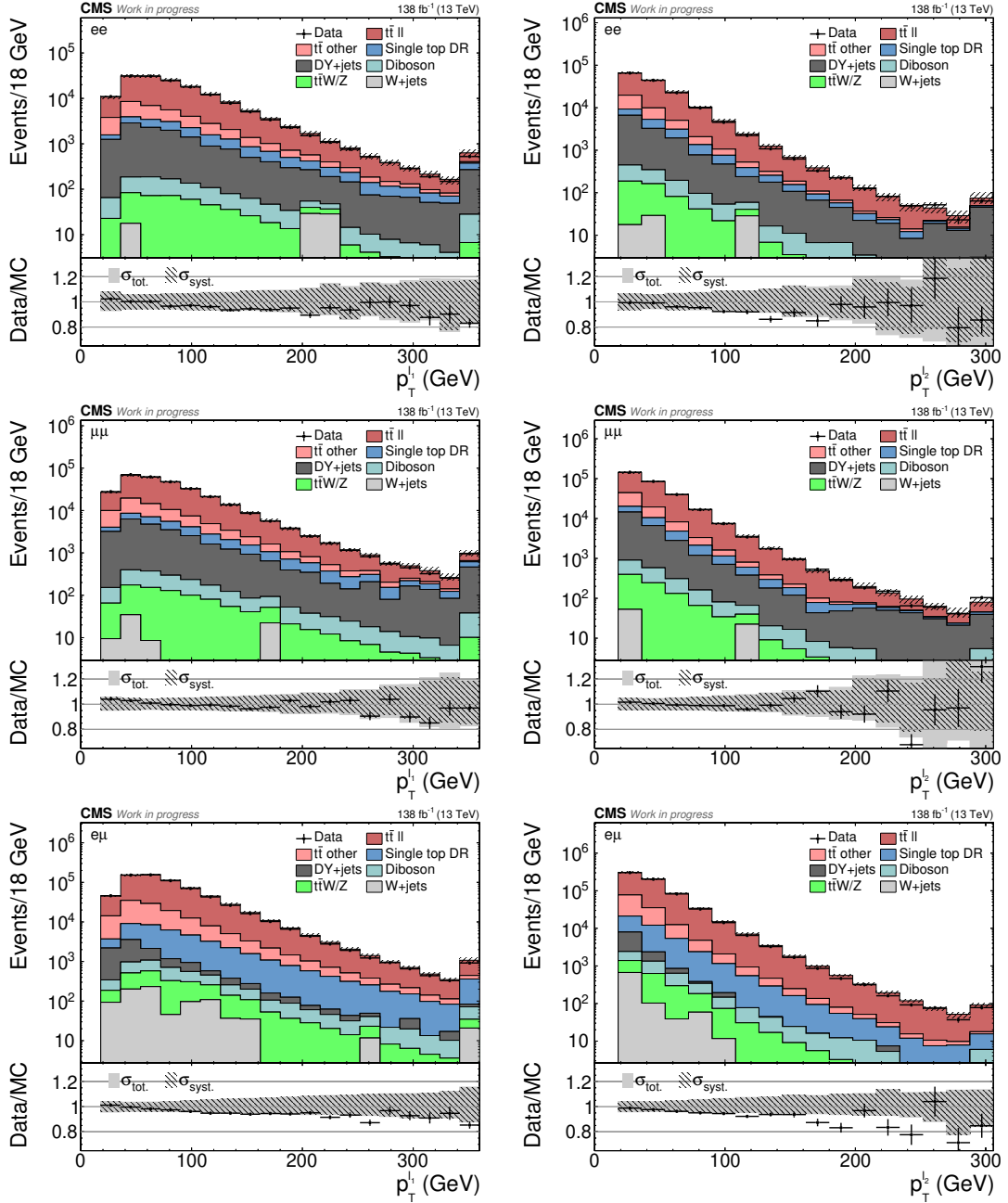


Figure A6.1: Control distributions as a function of the leading (left) and subleading (right) lepton transverse momentum in the e^+e^- (top), $\mu^+\mu^-$ (middle) and $e^\pm\mu^\mp$ (bottom) channel for all data-taking periods combined. The upper parts of the plots show the observed data (black) and the contributions predicted by simulation including the total uncertainty (hatched area). The lower parts of the plots show the ratio between data and simulation as well as the systematic and total uncertainties on the predicted distributions. In all distributions, the last bin includes all events above the plotted range.

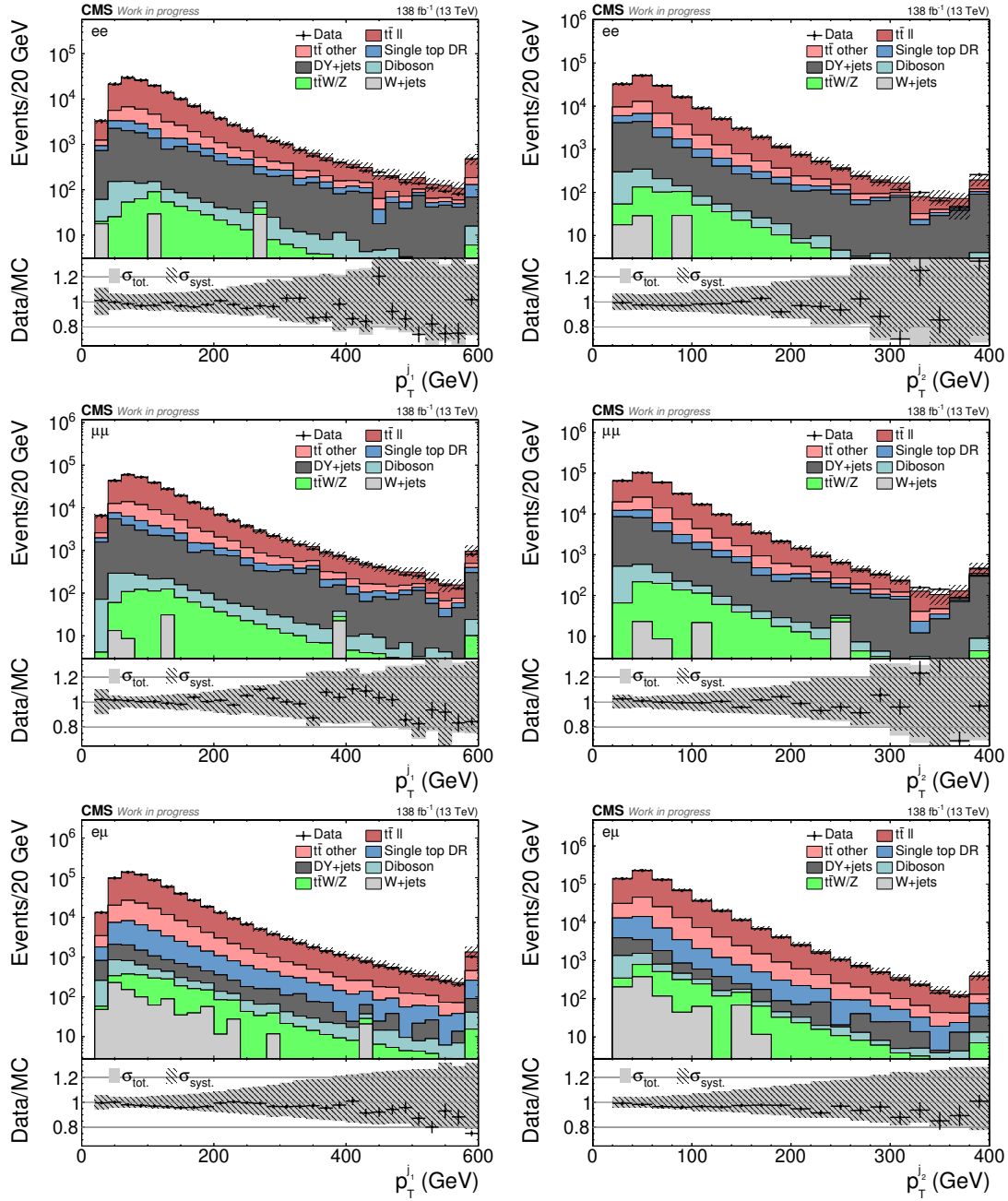


Figure A6.2: Control distributions as a function of the leading (left) and subleading (right) jet transverse momentum in the e^+e^- (top), $\mu^+\mu^-$ (middle) and $e^+\mu^-$ (bottom) channel for all data-taking periods combined. The upper parts of the plots show the observed data (black) and the contributions predicted by simulation including the total uncertainty (hatched area). The lower parts of the plots show the ratio between data and simulation as well as the systematic and total uncertainties on the predicted distributions. In all distributions, the last bin includes all events above the plotted range.

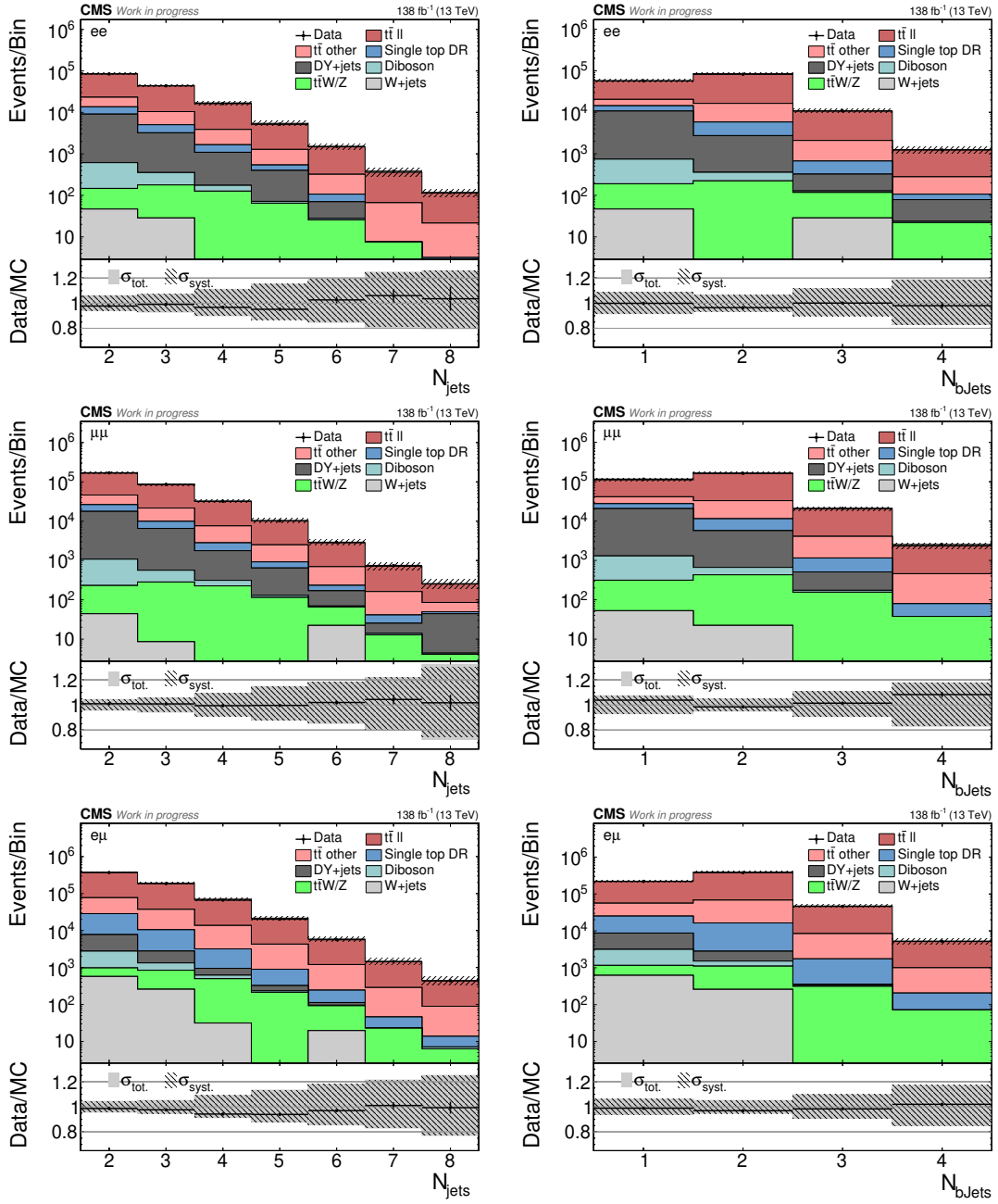


Figure A6.3: Control distributions as a function of the number of reconstructed jets (left) and b-tagged jets (right) in the e^+e^- (top), $\mu^+\mu^-$ (middle) and $e^\pm\mu^\mp$ (bottom) channel for all data-taking periods combined. The upper parts of the plots show the observed data (black) and the contributions predicted by simulation including the total uncertainty (hatched area). The lower parts of the plots show the ratio between data and simulation as well as the systematic and total uncertainties on the predicted distributions. In all distributions, the last bin includes all events above the plotted range.

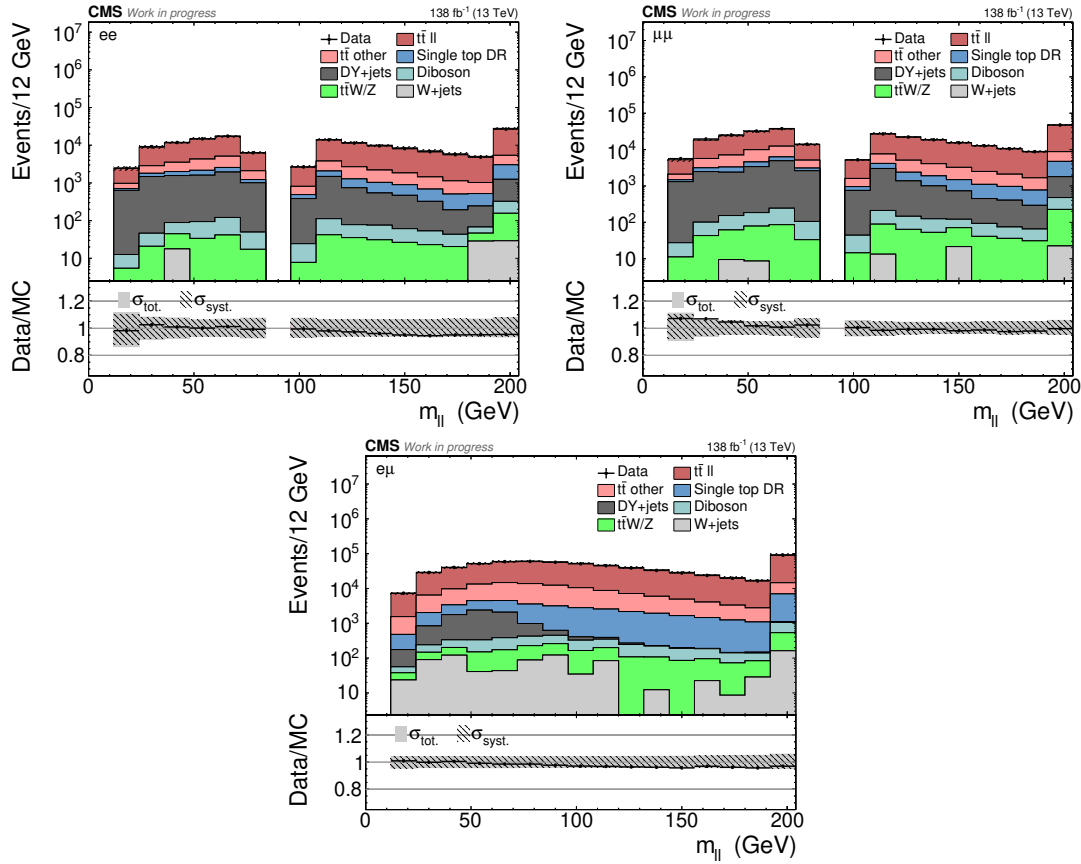


Figure A6.4: Control distributions as a function of the invariant mass of the reconstructed leptons in the e^+e^- (top left), $\mu^+\mu^-$ (top right) and $e^\pm\mu^\mp$ (bottom) channel for all data-taking periods combined. The upper parts of the plots show the observed data (black) and the contributions predicted by simulation including the total uncertainty (hatched area). The lower parts of the plots show the ratio between data and simulation as well as the systematic and total uncertainties on the predicted distributions. In all distributions, the last bin includes all events above the plotted range.

Appendix 7: DNN input features

The full list of all 54 input features considered in the development of the DNN regression including descriptions of the individual inputs is shown in Tab. A7.1.

Table A7.1: Full set of input features considered in the development of the DNN regression.

Symbol	Description
PUPPI $p_{x,y}^{\text{miss}}$	x,y -component of PUPPI p_T^{miss} (cf. Sec.6.2.6)
PF $p_{x,y}^{\text{miss}}$	x,y -component of PF p_T^{miss} (cf. Sec.6.2.6)
Calo p_T^{miss}	p_T^{miss} only derived from calorimeter entries [189]
$\sigma(\text{PUPPI } p_T^{\text{miss}})$	uncertainty on PUPPI p_T^{miss}
$\sigma(\text{PF } p_T^{\text{miss}})$	uncertainty on PF p_T^{miss}
$p_{x,y}^{\text{all jet}}$	momentum in x,y -direction of the system including all jets
$p_T^{\text{all jet}}$	transverse momentum of the system including all jets
$l_{1(2)}^{1(2)}$	x,y -component of the (sub)leading lepton momentum
$j_{1(2)}^{1(2)}$	x,y -component of the (sub)leading jet momentum
$\eta_{1(2)}$	pseudorapidity of the (sub)leading lepton
$\eta_{j_{1(2)}}$	pseudorapidity of the (sub)leading jet
$E_{l_{1(2)}}$	energy of the (sub)leading lepton
$E_{j_{1(2)}}$	energy of the (sub)leading jet
flavor($l_{1(2)}$)	flavor of the (sub)leading lepton ($e \hat{=} 1, \mu \hat{=} 2$)
loose veto	event contains third lepton with looser identification criteria
N_{jets}	number of jets
N_{vtx}	number of reconstructed vertices
$\Delta\phi(\text{PUPPI } p_T^{\text{miss}}, j_{1(2)})$	azimuthal distance between PUPPI p_T^{miss} and the (sub)leading jet
$\Delta\phi(\text{PF } p_T^{\text{miss}}, j_{1(2)})$	azimuthal distance between PF p_T^{miss} and the (sub)leading jet
$\min[\Delta\phi(\text{PUPPI } p_T^{\text{miss}}, j)]$	minimal azimuthal distance between PUPPI p_T^{miss} and any jet
$\min[\Delta\phi(\text{PF } p_T^{\text{miss}}, j)]$	minimal azimuthal distance between PF p_T^{miss} and any jet
$\max[\Delta\phi(\text{PF } p_T^{\text{miss}}, j)]$	maximal azimuthal distance between PF p_T^{miss} and any jet
$\max[\Delta\phi(\text{PUPPI } p_T^{\text{miss}}, j)]$	maximal azimuthal distance between PUPPI p_T^{miss} and any jet
$\Delta\phi(\text{PUPPI } p_T^{\text{miss}}, j_b)$	azimuthal distance between PUPPI p_T^{miss} and the leading b jet
$\Delta\phi(\text{PF } p_T^{\text{miss}}, j_b)$	azimuthal distance between PF p_T^{miss} and the leading b jet
$\Delta\phi(l_1, l_2)$	azimuthal distance between the leading and subleading lepton
$\Delta\phi(j_1, j_2)$	azimuthal distance between the leading and subleading jet
$\Delta\phi(l_1, j_1)$	azimuthal distance between the leading lepton and leading jet
$\Delta\phi(l_1, j_b)$	azimuthal distance between the leading lepton and leading b jet
$m_{\text{all j}}$	invariant mass of the system including all jets
$m_{\text{ll+all j}}$	invariant mass of the system including all jets and leptons
m_{ll}	invariant mass of the system including the two leptons
m_{jj}	invariant mass of the system including the two leading jets
M_T	transverse mass of the leading lepton and PF p_T^{miss}
M_{T2}	stransverse mass [213, 214]
H_T	scalar sum of all jet transverse momenta
$p_T^{\text{ll+all jet}}$	transverse momentum of the system including all jets and leptons
$p_T^{\text{ll}}/p_T^{\text{all jet}}$	p_T ratio between the systems of both leptons and all jets

Appendix 8: DNN performance

The performance of the DNN regression regarding the magnitude and the azimuthal angle for both 2016 and the 2017 data-taking period are shown in Fig. A8.1. The corresponding plots for the 2018 data-taking period can be found in Fig. 7.9.

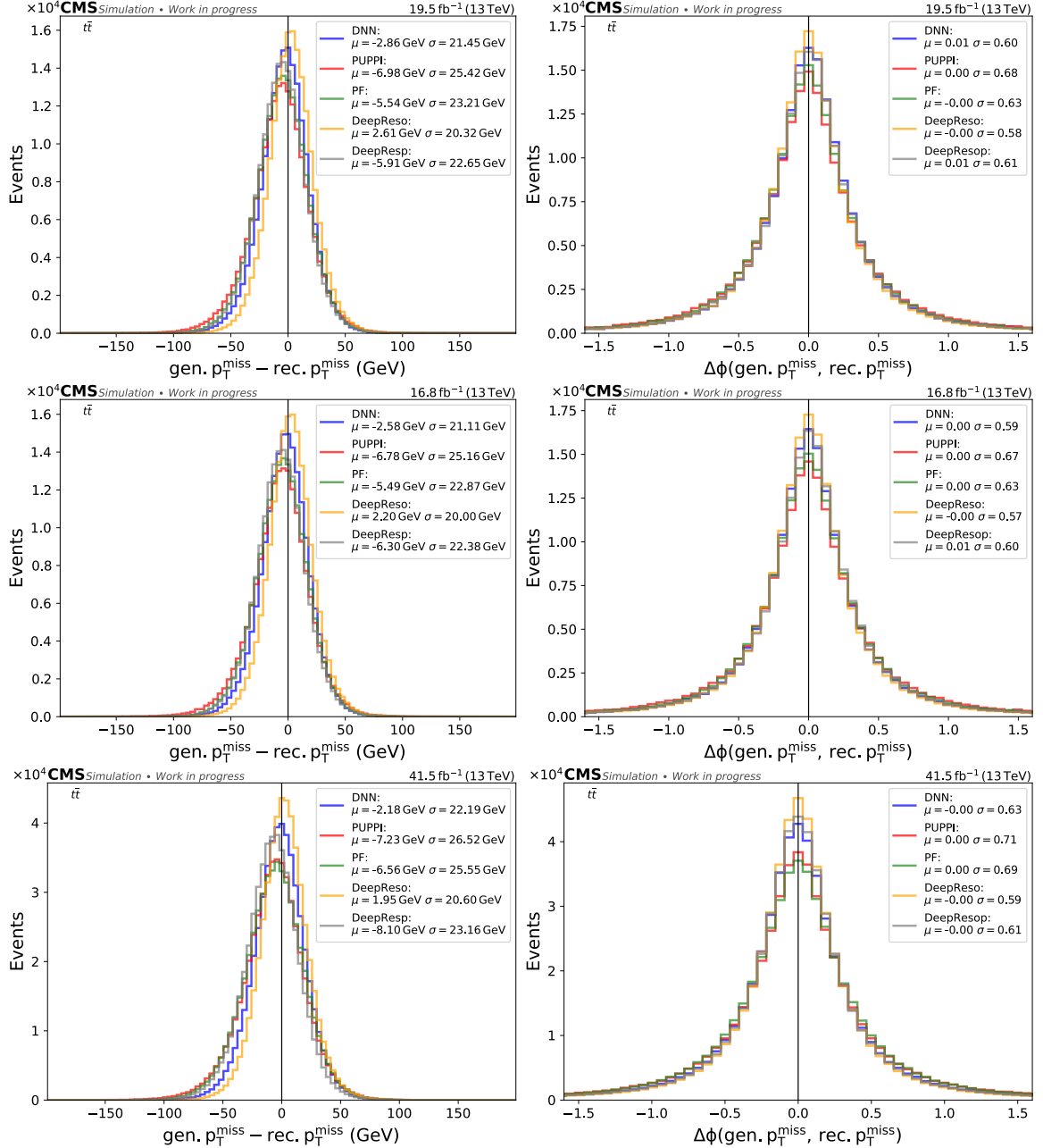


Figure A8.1: Difference between generated and reconstructed p_T^{miss} with respect to the magnitude (left) and the azimuthal angle (right) comparing DNN (blue), PUPPI (red) and PF (green) p_T^{miss} as well as the resolution (orange) and response tune (grey) of the Deep p_T^{miss} reconstruction. The results are derived based on the 2016preVFP (top), 2016postVFP (middle) and 2017 (bottom) data-taking period using the $t\bar{t}$ test sample. The mean and the standard deviation of each distribution are listed in the legend.

Appendix 9: Unfolding validation tests

The validation tests performed for the $p_T^{\nu\nu}$ measurement have been already discussed in Sec. 8.4. The corresponding tests for the $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ and the two-dimensional measurements, which yield similar results to the $p_T^{\nu\nu}$ measurement and therefore support the selection of the unregularized least square approach, can be found in Fig. A9.1-A9.4.

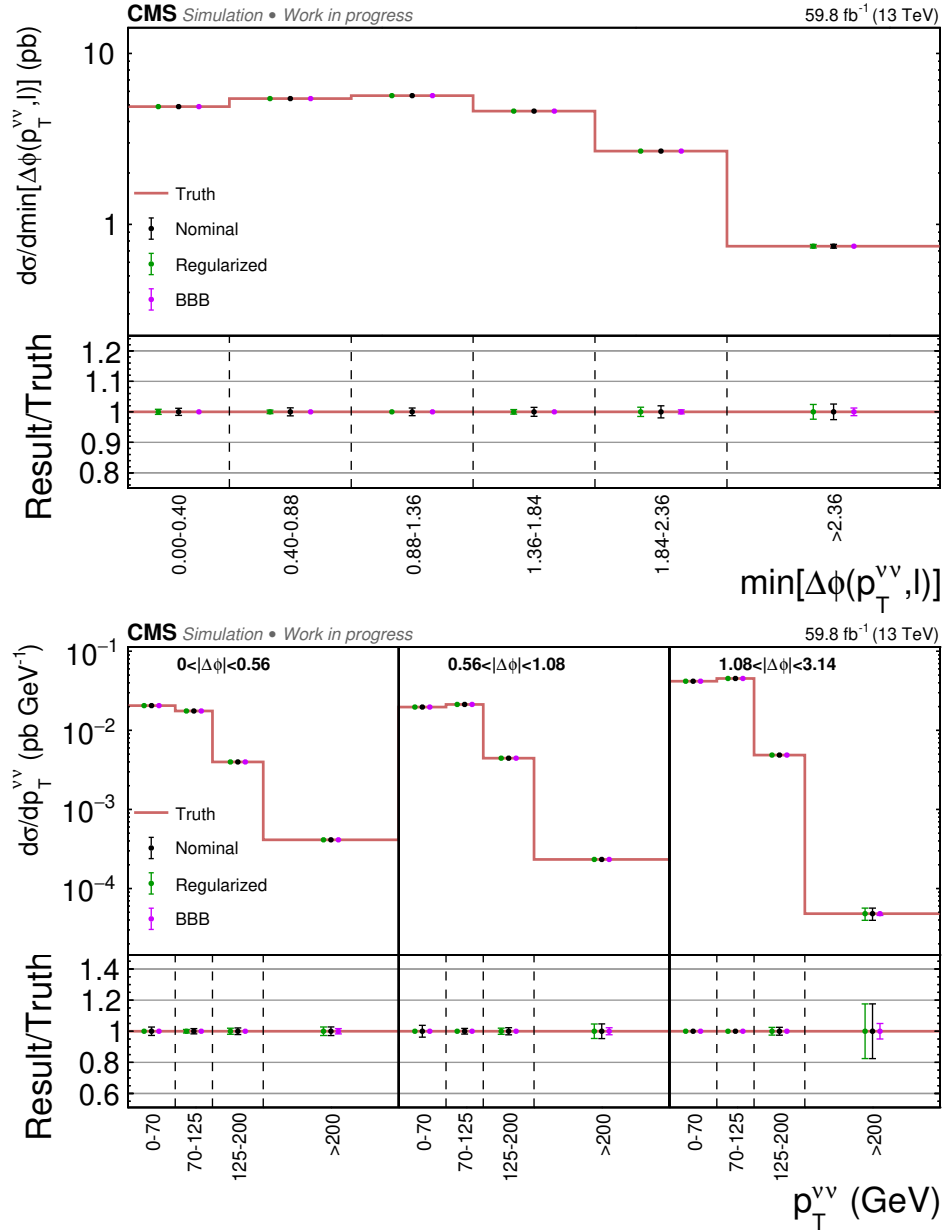


Figure A9.1: Result of the truth test performed for the $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ (top) and 2D measurement (bottom) derived with the nominal (black), the regularized (green) and the bin-by-bin unfolding (purple) based on the 2018 data-taking period. The unfolded distributions are compared to the expected distribution (red) based on the dileptonic $t\bar{t}$ signal sample. The error bars correspond to the statistical uncertainty. The lower parts show the ratios of the unfolded and the expected distributions.

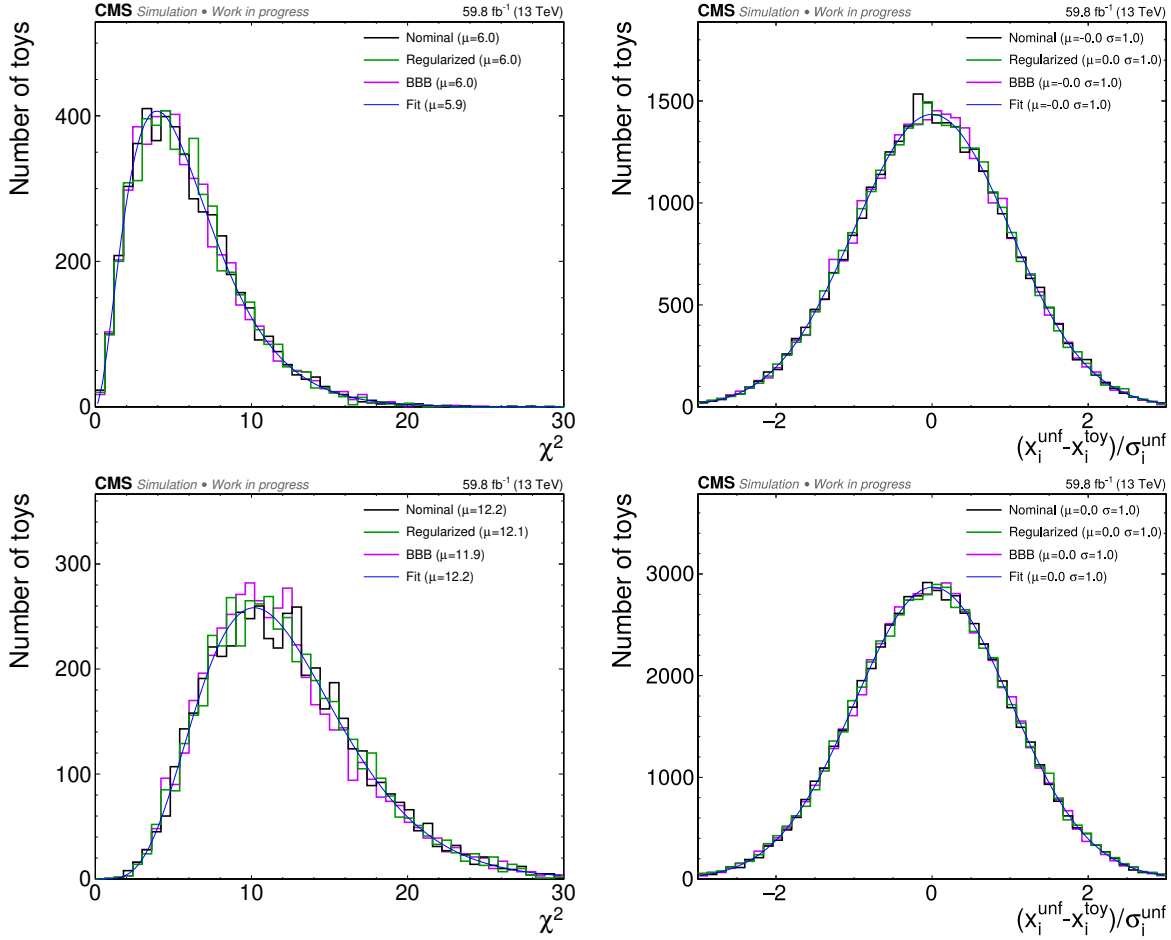


Figure A9.2: Results of the toy test performed for the $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ (top) and 2D measurement (bottom) derived with the nominal (black), the regularized (green) and the bin-by-bin unfolding (purple) based on the 2018 data-taking period. The plots on the left show the distribution of the χ^2 -values derived for each individual toy data set fitted by a χ^2 function (blue). The plots on the right show the pull derived for each bin in each toy data set fitted by a Gaussian function (blue).

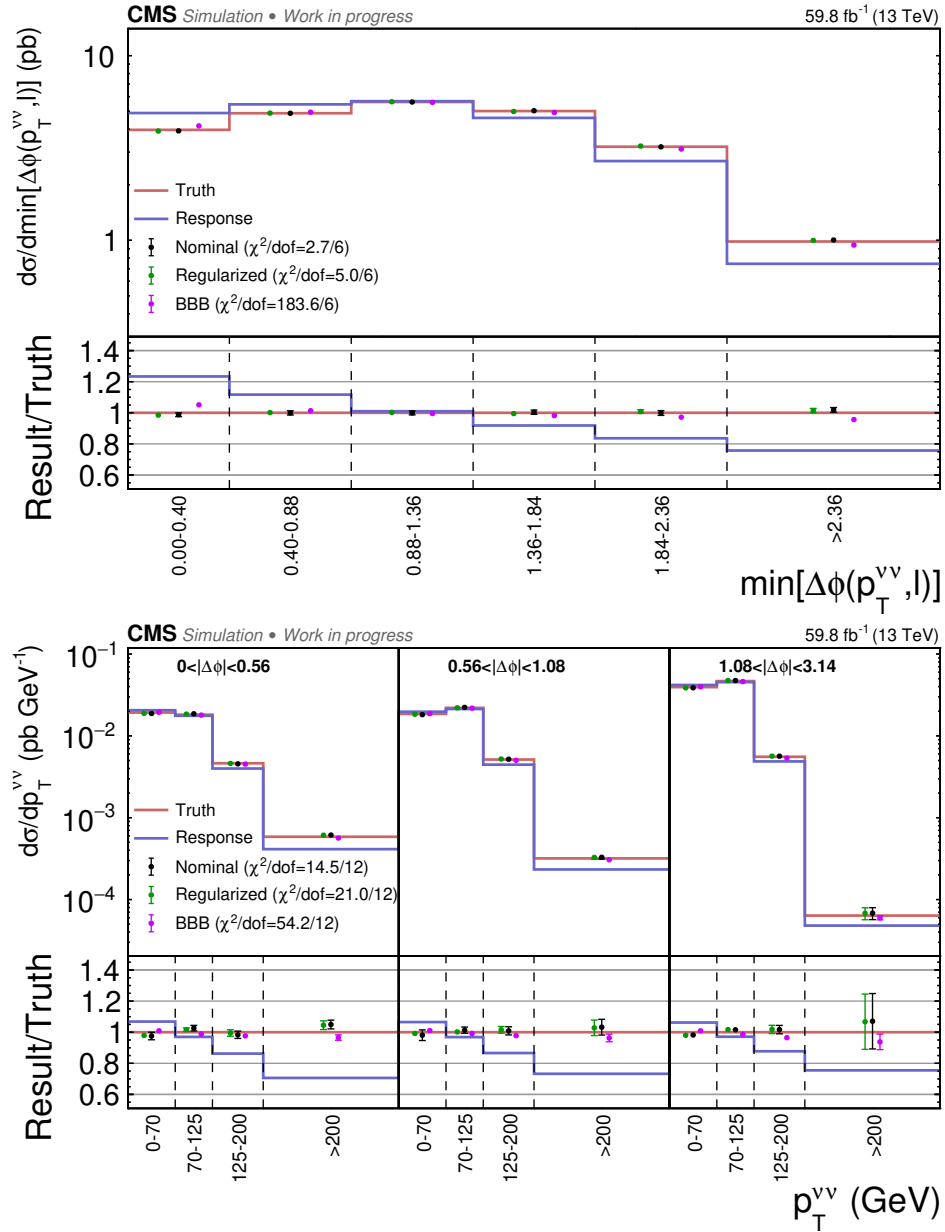


Figure A9.3: Result of the reweighting test performed for the $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ (top) and 2D measurement (bottom) derived with the nominal (black), the regularized (green) and the bin-by-bin unfolding (purple) based on the 2018 data-taking period. The unfolded distributions are compared to the expected distribution (red) based on the reweighted dileptonic $t\bar{t}$ signal sample with the corresponding χ^2/dof values given in the legend. The unweighted distribution, used for the response matrix, is shown in blue. The error bars correspond to the statistical uncertainty. The lower parts show the ratios of the unfolded and the expected distributions.

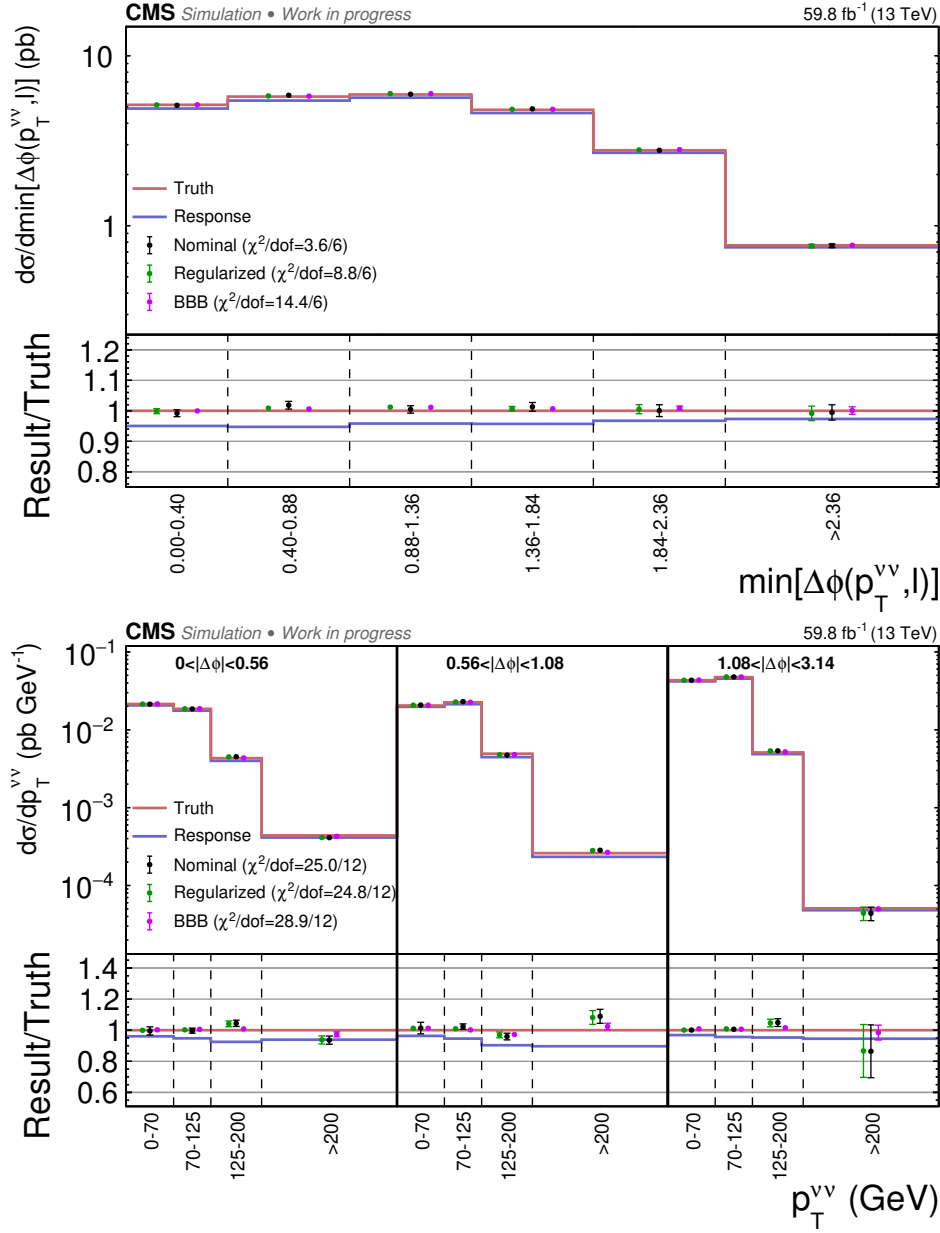


Figure A9.4: Result of the alternative MC test performed for the $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ (top) and 2D measurement (bottom) derived with the nominal (black), the regularized (green) and the bin-by-bin unfolding (purple) based on the 2018 data-taking period. The unfolded distributions are compared to the expected distribution (red) based on the alternative dileptonic $t\bar{t}$ signal sample with the corresponding χ^2/dof values given in the legend. The nominal distribution, used for the response matrix, is shown in blue. The error bars correspond to the statistical uncertainty. The lower parts show the ratios of the unfolded and the expected distributions.

Appendix 10: Uncertainties for normalized results

The breakdowns of the normalized differential cross section measurements as a function of $p_T^{\nu\nu}$, $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ and the 2D combination of both observables can be found in Fig. A10.1-A10.3. For the $p_T^{\nu\nu}$ and the two-dimensional measurement, conclusions similar to the corresponding absolute measurements, discussed in Sec. 9.3, can be drawn. Only for the $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ measurement, a significant difference between the breakdown for the normalized and the absolute cross section can be observed, which is caused by the significant reduction of the lepton reconstruction uncertainty in the normalized result.

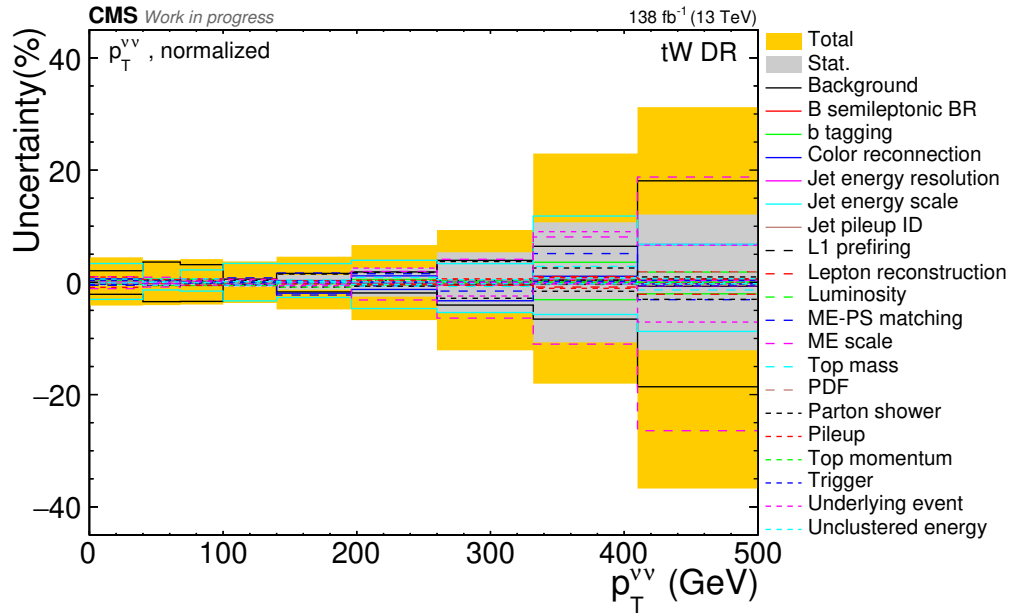


Figure A10.1: Breakdown of the uncertainties for the normalized $p_T^{\nu\nu}$ measurement using the DR scheme. The statistical uncertainty is shown as gray-shaded area. The orange-shaded area indicates the total uncertainty, given by the quadratic sum of all systematic contributions and the statistical uncertainty.

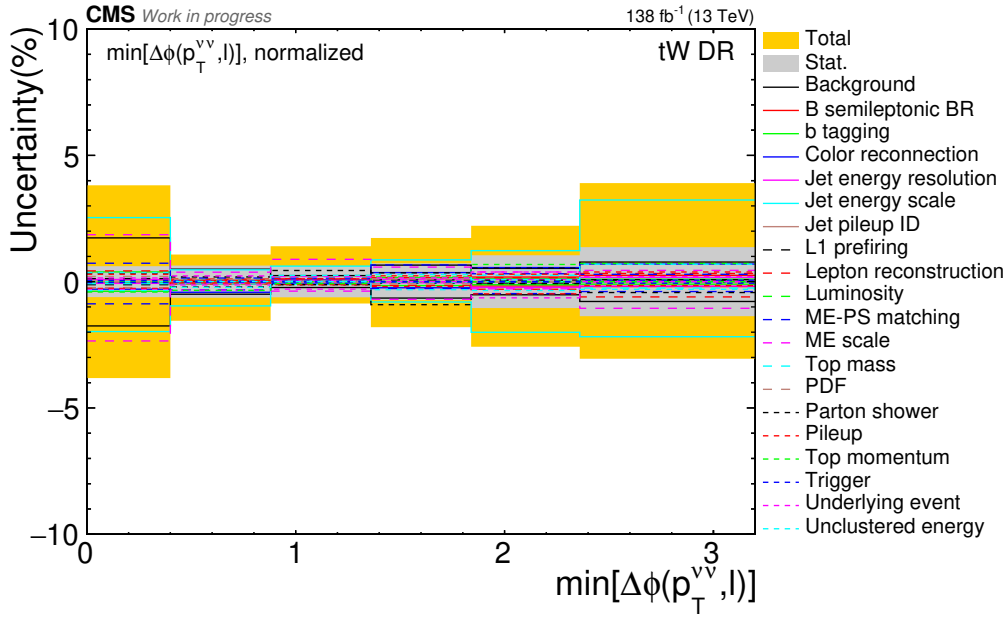


Figure A10.2: Breakdown of the uncertainties for the normalized $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ measurement using the DR scheme. The statistical uncertainty is shown as gray-shaded area. The orange-shaded area indicates the total uncertainty, given by the quadratic sum of all systematic contributions and the statistical uncertainty.

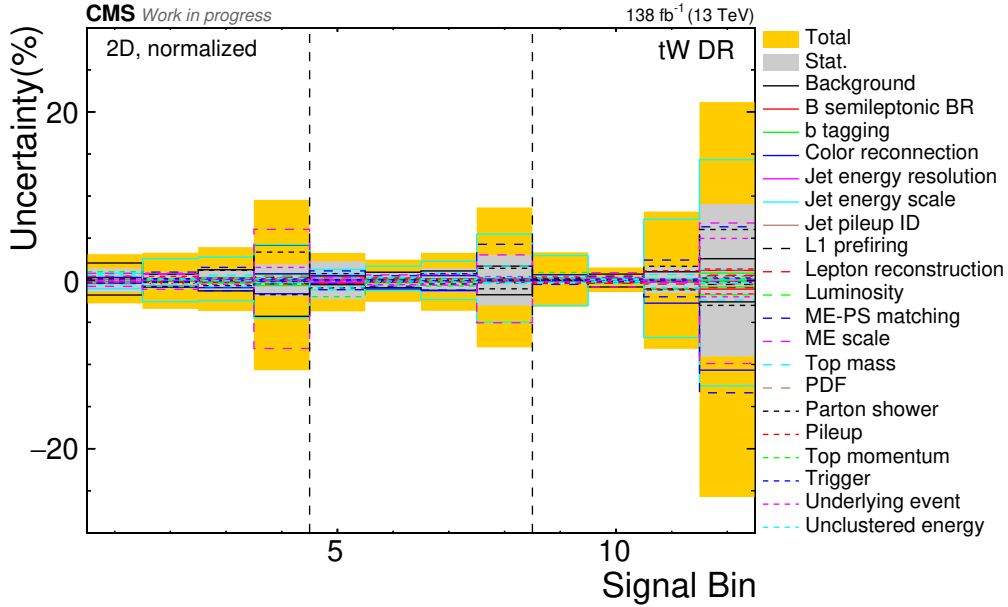


Figure A10.3: Breakdown of the uncertainties for the normalized 2D measurement using the DR scheme. The statistical uncertainty is shown as gray-shaded area. The orange-shaded area indicates the total uncertainty, given by the quadratic sum of all systematic contributions and the statistical uncertainty. The vertical black lines correspond to the bin edges in $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$, while inside each of these bins, several $p_T^{\nu\nu}$ bins are shown in increasing order.

Appendix 11: Uncertainties for results using the DS scheme

The breakdowns of the uncertainties for the absolute and normalized differential cross sections using the DS scheme to describe the $t\bar{W}$ background in simulation can be found in Fig. A11.1-A11.6. The largest difference to the uncertainties derived for the results using the DR scheme, shown in Sec. 9.3 and Appx. 10, are observed for both $p_T^{\nu\nu}$ measurements. Here, the uncertainty on the $t\bar{W}$ background prediction, mainly given by the ME-scale and background normalization uncertainty, is significantly reduced at high- $p_T^{\nu\nu}$ due to the smaller $t\bar{W}$ contribution predicted in the DS scheme. Effects of the same origin can also be observed at high $p_T^{\nu\nu}$ in the first angular bin of the 2D measurements. In the measurements as a function of $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$, however, no significant differences between the results using the two different schemes are found.

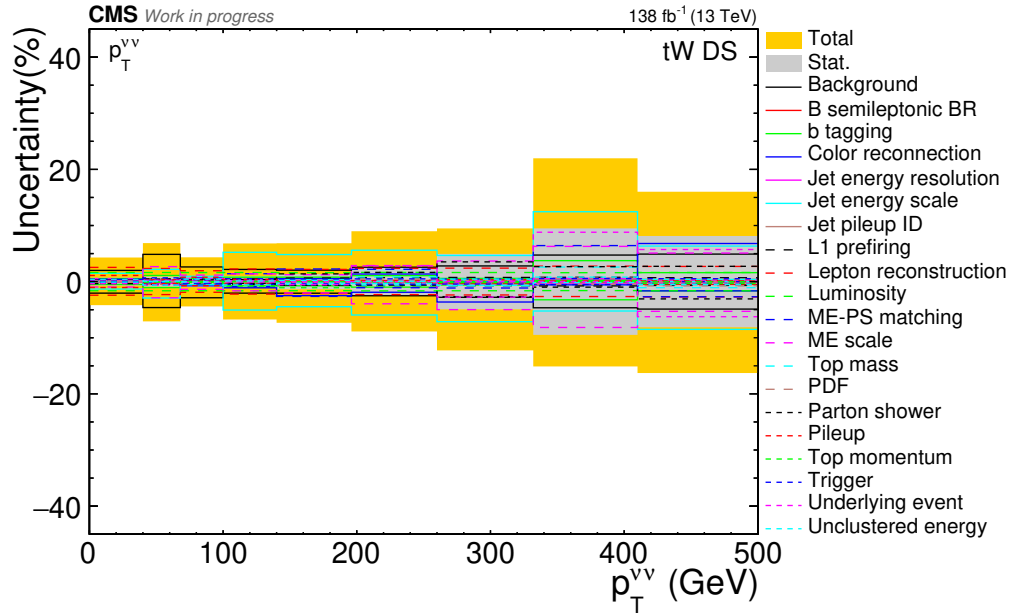


Figure A11.1: Breakdown of the uncertainties for the absolute $p_T^{\nu\nu}$ measurement using the DS scheme. The statistical uncertainty is shown as gray-shaded area. The orange-shaded area indicates the total uncertainty, given by the quadratic sum of all systematic contributions and the statistical uncertainty.

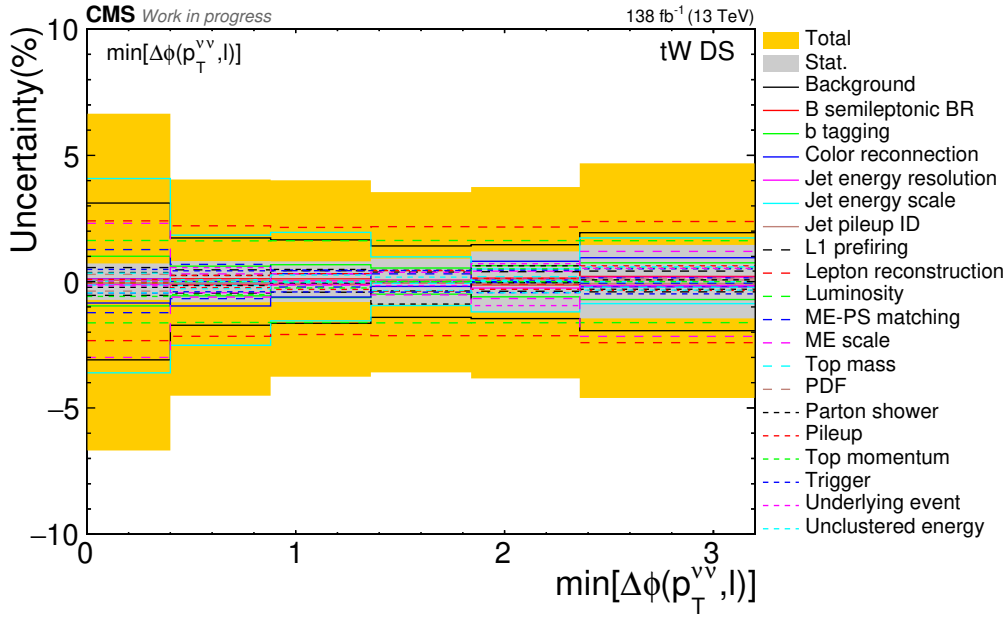


Figure A11.2: Breakdown of the uncertainties for the absolute $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$ measurement using the DS scheme. The statistical uncertainty is shown as gray-shaded area. The orange-shaded area indicates the total uncertainty, given by the quadratic sum of all systematic contributions and the statistical uncertainty.

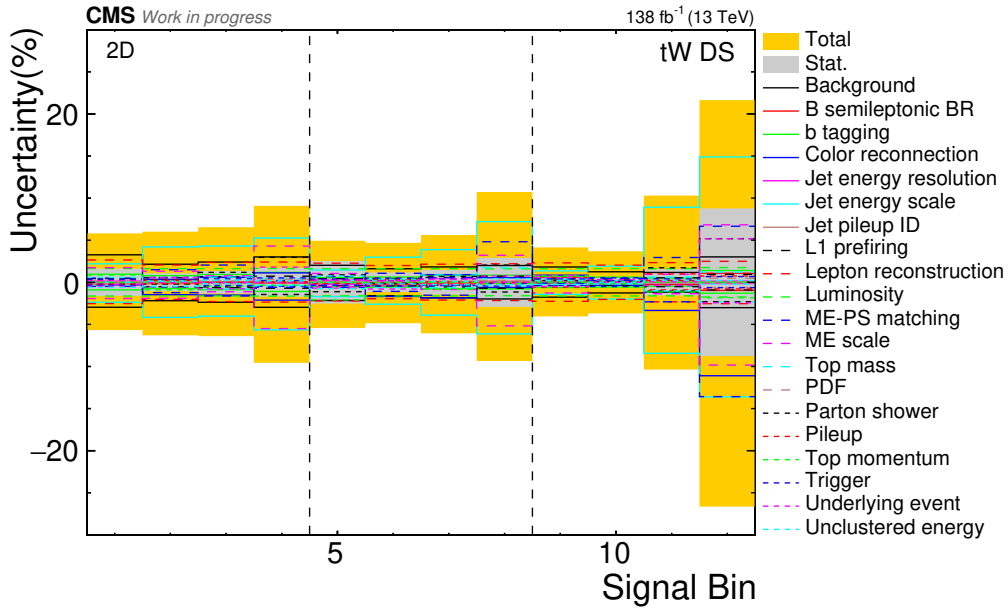


Figure A11.3: Breakdown of the uncertainties for the absolute 2D measurement using the DS scheme. The statistical uncertainty is shown as gray-shaded area. The orange-shaded area indicates the total uncertainty, given by the quadratic sum of all systematic contributions and the statistical uncertainty. The vertical black lines correspond to the bin edges in $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$, while inside each of these bins, several $p_T^{\nu\nu}$ bins are shown in increasing order.

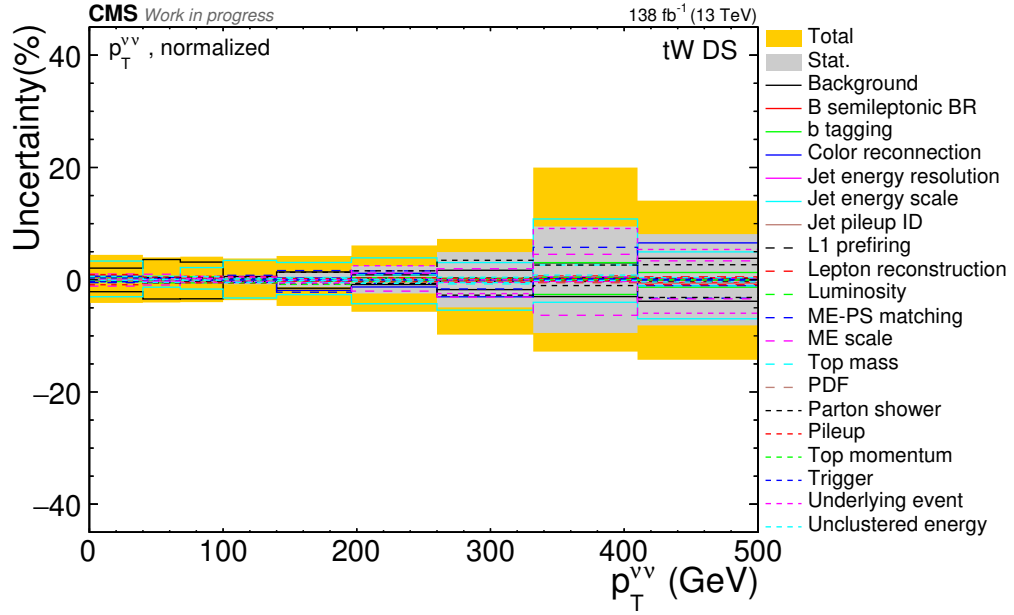


Figure A11.4: Breakdown of the uncertainties for the normalized $p_T^{\nu\nu}$ measurement using the DS scheme. The statistical uncertainty is shown as gray-shaded area. The orange-shaded area indicates the total uncertainty, given by the quadratic sum of all systematic contributions and the statistical uncertainty.

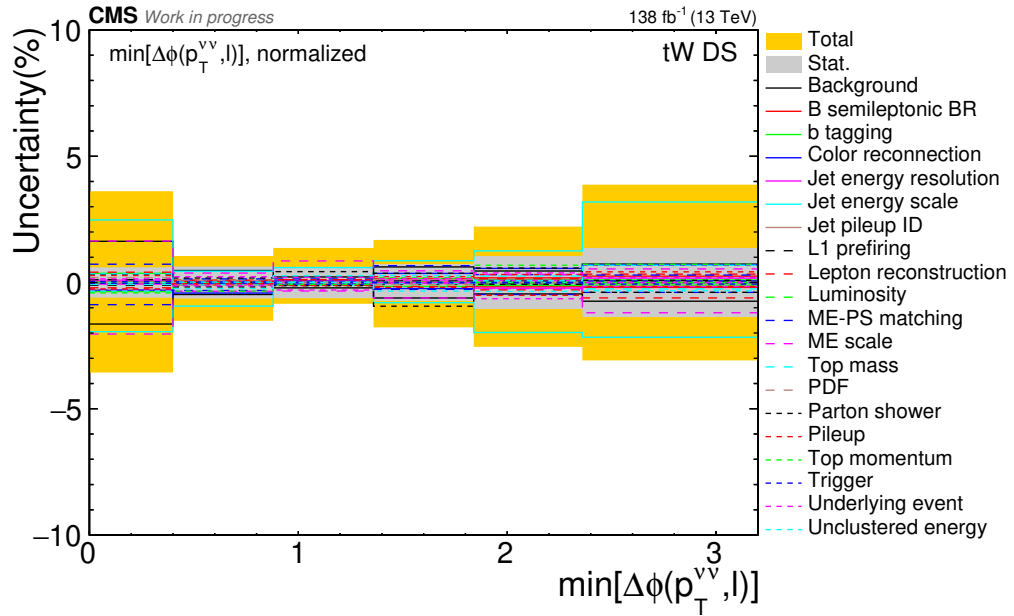


Figure A11.5: Breakdown of the uncertainties for the normalized $\min[\Delta\phi(p_T^{\nu\nu}, l)]$ measurement using the DS scheme. The statistical uncertainty is shown as gray-shaded area. The orange-shaded area indicates the total uncertainty, given by the quadratic sum of all systematic contributions and the statistical uncertainty.

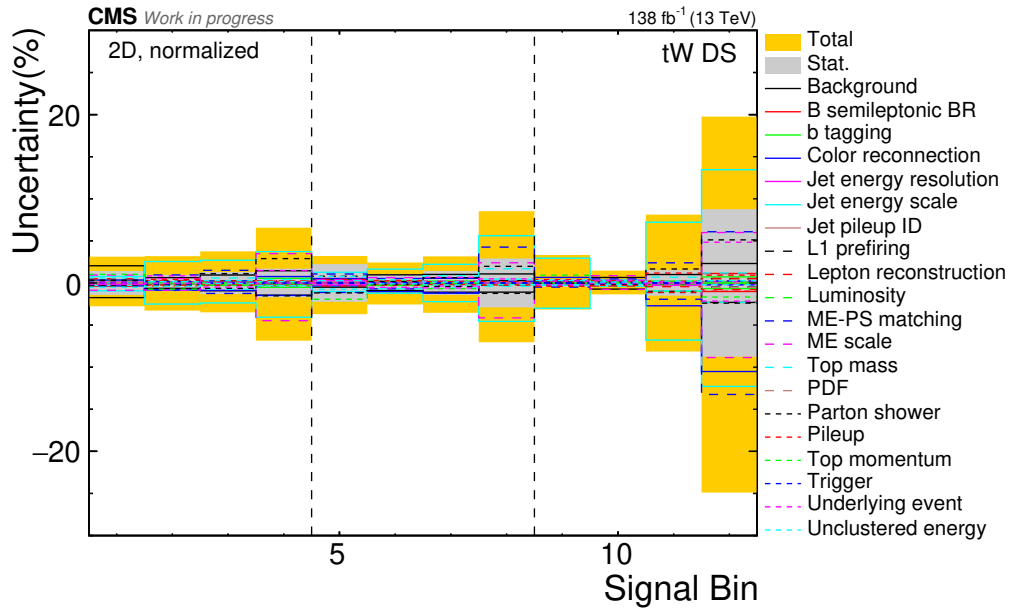


Figure A11.6: Breakdown of the uncertainties for the normalized 2D measurement using the DS scheme. The statistical uncertainty is shown as gray-shaded area. The orange-shaded area indicates the total uncertainty, given by the quadratic sum of all systematic contributions and the statistical uncertainty. The vertical black lines correspond to the bin edges in $\min[\Delta\phi(p_T^{\nu\nu}, \ell)]$, while inside each of these bins, several $p_T^{\nu\nu}$ bins are shown in increasing order.

Appendix 12: BSM content in detector-level distributions

The predicted yields for potential BSM contributions based on a stop-pair production process with stop and neutralino masses of 525 and 350 GeV, respectively, for the three main detector-level distributions are shown in Fig. A12.1. As discussed in Sec. 2.3.3, the largest sensitivity can be observed in the 2D distributions in the phase space of high p_T^{miss} and large angular separation. The BSM yield in the highest $\min[\Delta\phi(p_T^{\text{miss}}, \ell)]$ bin is reduced for the given BSM scenario, due to the rather small mass splitting between the stop and the neutralino, which only leads to moderate additional contributions to the missing transverse momentum from the two neutralinos. Adequate simulated samples for scenarios with a larger mass splitting, which would show a significantly increased sensitivity, were not available in the scope of this thesis.

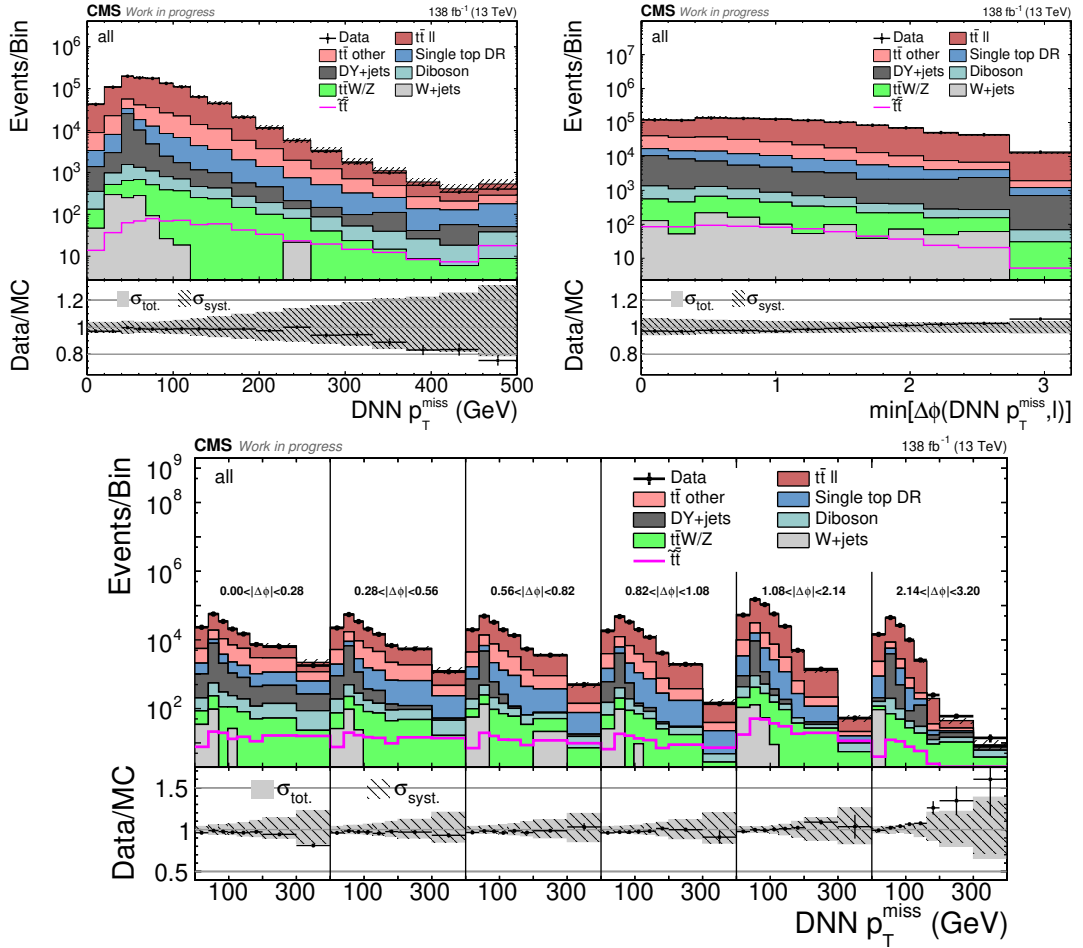


Figure A12.1: Detector-level distributions of $\text{DNN } p_T^{\text{miss}}$ (top left), $\min[\Delta\phi(p_T^{\text{miss}}, \ell)]$ (top right) and the 2D combination (bottom) for all channels and data-taking periods. The upper parts of the plots show the observed data (black), the contributions predicted by simulation including the total uncertainty (hatched area) as well as the predicted yield for a stop-pair production with stop and neutralino masses of 525 and 350 GeV, respectively. The lower parts show the ratio between data and simulation as well as the systematic and total uncertainties on the predicted distribution. The highest p_T^{miss} bins include all events above the plotted range.

References

- [1] J. J. Thomson, “XL. Cathode Rays”, *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **44** (1897), doi:10.1080/14786449708621070.
- [2] E. Rutherford, “The scattering of alpha and beta particles by matter and the structure of the atom”, *Phil. Mag. Ser.6* **21** (1911), doi:10.1080/14786440508637080.
- [3] G. Pancheri and L. Bonolis, “The path to high-energy electron-positron colliders: from Wideroe’s betatron to Touschek’s AdA and to LEP”, 2017. arXiv:1710.09003.
- [4] S. L. Glashow, “Partial Symmetries of Weak Interactions”, *Nucl. Phys.* **22** (1961), doi:10.1016/0029-5582(61)90469-2.
- [5] A. Salam and J. C. Ward, “Electromagnetic and weak interactions”, *Phys. Lett.* **13** (1964), doi:10.1016/0031-9163(64)90711-5.
- [6] S. Weinberg, “A Model of Leptons”, *Phys. Rev. Lett.* **19** (1967), doi:10.1103/PhysRevLett.19.1264.
- [7] A. Salam, “Weak and Electromagnetic Interactions”, *Conf. Proc. C* **680519** (1968), doi:10.1142/9789812795915_0034.
- [8] D. J. Gross and F. Wilczek, “Ultraviolet Behavior of Nonabelian Gauge Theories”, *Phys. Rev. Lett.* **30** (1973), doi:10.1103/PhysRevLett.30.1343.
- [9] H. D. Politzer, “Reliable Perturbative Results for Strong Interactions?”, *Phys. Rev. Lett.* **30** (1973), doi:10.1103/PhysRevLett.30.1346.
- [10] M. Kobayashi and T. Maskawa, “CP Violation in the Renormalizable Theory of Weak Interaction”, *Prog. Theor. Phys.* **49** (1973), doi:10.1143/PTP.49.652.
- [11] CDF Collaboration, “Observation of top quark production in $\bar{p}p$ collisions”, *Phys. Rev. Lett.* **74** (1995), doi:10.1103/PhysRevLett.74.2626, arXiv:hep-ex/9503002.
- [12] D0 Collaboration, “Observation of the top quark”, *Phys. Rev. Lett.* **74** (1995), doi:10.1103/PhysRevLett.74.2632, arXiv:hep-ex/9503003.
- [13] F. Englert and R. Brout, “Broken Symmetry and the Mass of Gauge Vector Mesons”, *Phys. Rev. Lett.* **13** (1964), doi:10.1103/PhysRevLett.13.321.
- [14] P. W. Higgs, “Broken Symmetries and the Masses of Gauge Bosons”, *Phys. Rev. Lett.* **13** (1964), doi:10.1103/PhysRevLett.13.508.
- [15] ATLAS Collaboration, “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”, *Phys. Lett. B* **716** (2012), doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.
- [16] CMS Collaboration, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”, *Phys. Lett. B* **716** (2012), doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.
- [17] K. Garrett and G. Duda, “Dark Matter: A Primer”, *Adv. Astron.* **2011** (2011), doi:10.1155/2011/968283, arXiv:1006.2483.
- [18] S. T. Petcov, “The Nature of Massive Neutrinos”, *Adv. High Energy Phys.* **2013** (2013), doi:10.1155/2013/852987, arXiv:1303.5819.

- [19] SNO Collaboration, “Measurement of the rate of $\nu_e + d \rightarrow p + p + e^-$ interactions produced by ^8B solar neutrinos at the Sudbury Neutrino Observatory”, *Phys. Rev. Lett.* **87** (2001), doi:10.1103/PhysRevLett.87.071301, arXiv:nuc1-ex/0106015.
- [20] Super-Kamiokande Collaboration, “Evidence for oscillation of atmospheric neutrinos”, *Phys. Rev. Lett.* **81** (1998), doi:10.1103/PhysRevLett.81.1562, arXiv:hep-ex/9807003.
- [21] Daya Bay Collaboration, “Observation of electron-antineutrino disappearance at Daya Bay”, *Phys. Rev. Lett.* **108** (2012), doi:10.1103/PhysRevLett.108.171803, arXiv:1203.1669.
- [22] K2K Collaboration, “Measurement of Neutrino Oscillation by the K2K Experiment”, *Phys. Rev. D* **74** (2006), doi:10.1103/PhysRevD.74.072003, arXiv:hep-ex/0606032.
- [23] L. Evans and P. Bryant, “LHC Machine”, *JINST* **3** (2008), doi:10.1088/1748-0221/3/08/S08001.
- [24] ATLAS Collaboration, “The ATLAS Experiment at the CERN Large Hadron Collider”, *JINST* **3** (2008), doi:10.1088/1748-0221/3/08/S08003.
- [25] CMS Collaboration, “The CMS Experiment at the CERN LHC”, *JINST* **3** (2008), doi:10.1088/1748-0221/3/08/S08004.
- [26] 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”, arXiv:2302.01967.
- [27] CMS Collaboration, “Combined search for supersymmetry with photons in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *Phys. Lett. B* **801** (2020), doi:10.1016/j.physletb.2019.135183, arXiv:1907.00857.
- [28] D. Meuser, “Combination of Searches for gauge-mediated Supersymmetry in Events with Photons at CMS”, Master’s thesis, RWTH Aachen University, 2019. <https://cds.cern.ch/record/2671823>.
- [29] M. Czakon, A. Mitov, and R. Poncelet, “NNLO QCD corrections to leptonic observables in top-quark pair production and decay”, *JHEP* **05** (2021), doi:10.1007/JHEP05(2021)212, arXiv:2008.11133.
- [30] Based on private correspondence with Terry Generet and René Poncelet.
- [31] W. Venus, “A LEP summary”, *PoS HEP2001* (2001), doi:10.22323/1.007.0284.
- [32] P. D. Grannis and M. J. Shochet, “The Tevatron Collider Physics Legacy”, *Ann. Rev. Nucl. Part. Sci.* **63** (2013), doi:10.1146/annurev-nucl-102212-170621.
- [33] S. F. Novaes, “Standard model: An Introduction”, in *10th Jorge Andre Swieca Summer School: Particle and Fields*. 1999. arXiv:hep-ph/0001283.
- [34] W. Hollik, “Quantum field theory and the Standard Model”, in *2009 European School of High-Energy Physics*. 2010. arXiv:1012.3883.
- [35] F. Halzen and A. D. Martin, “Quarks and leptones”. John Wiley & Sons, Nashville, TN, January, 1984.

- [36] C. Burgard, “Standard model of physics”.
<https://example.net/tikz/examples/model-physics/>. Accessed 12.06.2023, Modified.
- [37] Particle Data Group Collaboration, “Review of Particle Physics”, *PTEP* **2022** (2022), doi:10.1093/ptep/ptac097.
- [38] T. Nakano and K. Nishijima, “Charge Independence for V-particles”, *Prog. Theor. Phys.* **10** (1953), doi:10.1143/PTP.10.581.
- [39] M. Gell-Mann, “The interpretation of the new particles as displaced charge multiplets”, *Nuovo Cim.* **4** (1956), doi:10.1007/BF02748000.
- [40] CMS Collaboration, “Measurement and QCD analysis of double-differential inclusive jet cross sections in pp collisions at $\sqrt{s} = 8$ TeV and cross section ratios to 2.76 and 7 TeV”, *JHEP* **03** (2017), doi:10.1007/JHEP03(2017)156, arXiv:1609.05331.
- [41] UA1 Collaboration, “Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540$ GeV”, *Phys. Lett. B* **122** (1983), doi:10.1016/0370-2693(83)91177-2.
- [42] UA2 Collaboration, “Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider”, *Phys. Lett. B* **122** (1983), doi:10.1016/0370-2693(83)91605-2.
- [43] UA1 Collaboration, “Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV/c**2 at the CERN SPS Collider”, *Phys. Lett. B* **126** (1983), doi:10.1016/0370-2693(83)90188-0.
- [44] UA2 Collaboration, “Evidence for $Z^0 \rightarrow e^+e^-$ at the CERN $\bar{p}p$ Collider”, *Phys. Lett. B* **129** (1983), doi:10.1016/0370-2693(83)90744-X.
- [45] N. Cabibbo, “Unitary Symmetry and Leptonic Decays”, *Phys. Rev. Lett.* **10** (1963), doi:10.1103/PhysRevLett.10.531.
- [46] CMS Collaboration, “A portrait of the Higgs boson by the CMS experiment ten years after the discovery”, *Nature* **607** (2022), doi:10.1038/s41586-022-04892-x, arXiv:2207.00043.
- [47] S. L. Glashow, J. Iliopoulos, and L. Maiani, “Weak Interactions with Lepton-Hadron Symmetry”, *Phys. Rev. D* **2** (1970), doi:10.1103/PhysRevD.2.1285.
- [48] SLAC-SP-017 Collaboration, “Discovery of a Narrow Resonance in e^+e^- Annihilation”, *Phys. Rev. Lett.* **33** (1974), doi:10.1103/PhysRevLett.33.1406.
- [49] E598 Collaboration, “Experimental Observation of a Heavy Particle J ”, *Phys. Rev. Lett.* **33** (1974), doi:10.1103/PhysRevLett.33.1404.
- [50] V. C. Rubin and W. K. Ford, Jr., “Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions”, *Astrophys. J.* **159** (1970), doi:10.1086/150317.
- [51] E. Corbelli and P. Salucci, “The Extended Rotation Curve and the Dark Matter Halo of M33”, *Mon. Not. Roy. Astron. Soc.* **311** (2000), doi:10.1046/j.1365-8711.2000.03075.x, arXiv:astro-ph/9909252. Modified.

- [52] D. Walsh, R. F. Carswell, and R. J. Weymann, “0957 + 561 A, B: twin quasistellar objects or gravitational lens?”, *Nature* **279** (May, 1979), doi:10.1038/279381a0.
- [53] A. G. Bergmann, V. Petrosian, and R. Lynds, “Gravitational lens models of arcs in clusters”, *The Astrophysical Journal* **350** (February, 1990), doi:10.1086/168359.
- [54] D. Clowe et al., “A direct empirical proof of the existence of dark matter”, *Astrophys. J. Lett.* **648** (2006), doi:10.1086/508162, arXiv:astro-ph/0608407.
- [55] C. L. Bennett et al., “Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: final maps and results”, *The Astrophysical Journal Supplement Series* **208** (2013), doi:10.1088/0067-0049/208/2/20.
- [56] W. J. Percival et al., “Baryon acoustic oscillations in the Sloan Digital Sky Survey Data Release 7 galaxy sample”, *Monthly Notices of the Royal Astronomical Society* **401** (2010), doi:10.1111/j.1365-2966.2009.15812.x.
- [57] VIRGO Consortium Collaboration, “Evolution of structure in cold dark matter universes”, *Astrophys. J.* **499** (1998), doi:10.1086/305615, arXiv:astro-ph/9709010.
- [58] S. P. Martin, “A Supersymmetry primer”, *Adv. Ser. Direct. High Energy Phys.* **18** (1998), doi:10.1142/9789812839657_0001, arXiv:hep-ph/9709356.
- [59] KATRIN Collaboration, “Direct neutrino-mass measurement with sub-electronvolt sensitivity”, *Nature Phys.* **18** (2022), doi:10.1038/s41567-021-01463-1, arXiv:2105.08533.
- [60] M. Drewes, “The Phenomenology of Right Handed Neutrinos”, *Int. J. Mod. Phys. E* **22** (2013), doi:10.1142/S0218301313300191, arXiv:1303.6912.
- [61] J. D. Lykken, “Beyond the Standard Model”, in *CERN Yellow Report CERN-2010-002*. 2010. arXiv:1005.1676.
- [62] H. Murayama, “Supersymmetry phenomenology”, in *ICTP Summer School in Particle Physics*. 2000. arXiv:hep-ph/0002232.
- [63] Super-Kamiokande Collaboration, “Search for proton decay via $p \rightarrow e^+ \pi^0$ and $p \rightarrow \mu^+ \pi^0$ with an enlarged fiducial volume in Super-Kamiokande I-IV”, *Phys. Rev. D* **102** (2020), doi:10.1103/PhysRevD.102.112011, arXiv:2010.16098.
- [64] M. Papucci, J. T. Ruderman, and A. Weiler, “Natural SUSY Endures”, *JHEP* **09** (2012), doi:10.1007/JHEP09(2012)035, arXiv:1110.6926.
- [65] J. L. Feng, “Dark Matter Candidates from Particle Physics and Methods of Detection”, *Ann. Rev. Astron. Astrophys.* **48** (2010), doi:10.1146/annurev-astro-082708-101659, arXiv:1003.0904.
- [66] G. Bertone, D. Hooper, and J. Silk, “Particle dark matter: Evidence, candidates and constraints”, *Phys. Rept.* **405** (2005), doi:10.1016/j.physrep.2004.08.031, arXiv:hep-ph/0404175.
- [67] F.-P. Schilling, “Top Quark Physics at the LHC: A Review of the First Two Years”, *Int. J. Mod. Phys. A* **27** (2012), doi:10.1142/S0217751X12300165, arXiv:1206.4484.
- [68] LHC Top Working Group, “LHCTopWG Summary Plots”.
<https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCTopWGSummaryPlots>.
Accessed 31.03.2023.

- [69] M. Czakon and A. Mitov, “Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders”, *Comput. Phys. Commun.* **185** (2014), doi:10.1016/j.cpc.2014.06.021, arXiv:1112.5675.
- [70] K. Lannon, F. Margaroli, and C. Neu, “Measurements of the Production, Decay and Properties of the Top Quark: A Review”, *Eur. Phys. J. C* **72** (2012), doi:10.1140/epjc/s10052-012-2120-0, arXiv:1201.5873.
- [71] D0 Collaboration, “Useful diagrams of top signals and backgrounds”. <http://www-d0.fnal.gov/Run2Physics/top/toppublicwebpages/topfeynmandiagrams.html>. Accessed 31.03.2023, Modified.
- [72] J. R. Espinosa, “Implications of the top (and Higgs) mass for vacuum stability”, *PoS TOP2015* (2016), doi:10.22323/1.257.0043, arXiv:1512.01222.
- [73] J. Alwall, P. Schuster, and N. Toro, “Simplified Models for a First Characterization of New Physics at the LHC”, *Phys. Rev. D* **79** (2009), doi:10.1103/PhysRevD.79.075020, arXiv:0810.3921.
- [74] LHC New Physics Working Group, “Simplified Models for LHC New Physics Searches”, *J. Phys. G* **39** (2012), doi:10.1088/0954-3899/39/10/105005, arXiv:1105.2838.
- [75] M. R. Buckley, D. Feld, and D. Goncalves, “Scalar Simplified Models for Dark Matter”, *Phys. Rev. D* **91** (2015), doi:10.1103/PhysRevD.91.015017, arXiv:1410.6497.
- [76] CMS Collaboration, “Combined searches for the production of supersymmetric top quark partners in proton–proton collisions at $\sqrt{s} = 13$ TeV”, *Eur. Phys. J. C* **81** (2021), doi:10.1140/epjc/s10052-021-09721-5, arXiv:2107.10892.
- [77] ATLAS Collaboration, “Search for new phenomena in events with two opposite-charge leptons, jets and missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”, *JHEP* **04** (2021), doi:10.1007/JHEP04(2021)165, arXiv:2102.01444.
- [78] W. Buchmuller and D. Wyler, “Effective Lagrangian Analysis of New Interactions and Flavor Conservation”, *Nucl. Phys. B* **268** (1986), doi:10.1016/0550-3213(86)90262-2.
- [79] S. Bruggisser, R. Schäfer, D. van Dyk, and S. Westhoff, “The Flavor of UV Physics”, *JHEP* **05** (2021), doi:10.1007/JHEP05(2021)257, arXiv:2101.07273.
- [80] S. Myers, “The LEP collider, from design to approval and commissioning”, doi:10.5170/CERN-1991-008.
- [81] ALICE Collaboration, “The ALICE experiment at the CERN LHC”, *JINST* **3** (2008), doi:10.1088/1748-0221/3/08/S08002.
- [82] LHCb Collaboration, “The LHCb Detector at the LHC”, *JINST* **3** (2008), doi:10.1088/1748-0221/3/08/S08005.
- [83] E. Mobs, “The CERN accelerator complex - August 2018”. 2018, <http://cds.cern.ch/record/2636343>. Modified.
- [84] CERN, “Linear accelerator 4”. <https://home.cern/science/accelerators/linear-accelerator-4>. Accessed 12.04.2023.

- [85] W. Herr and B. Muratori, “Concept of luminosity”, doi:10.5170/CERN-2006-002.361.
- [86] CERN, “LHC Commissioning with beam”. <https://lhc-commissioning.web.cern.ch/>. Accessed 13.04.2023.
- [87] CMS Collaboration, “Public CMS Luminosity Information”. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>. Accessed 12.04.2023, Modified.
- [88] I. Zurbano Fernandez et al., “High-Luminosity Large Hadron Collider (HL-LHC): Technical design report”, doi:10.23731/CYRM-2020-0010.
- [89] CMS Collaboration, “CMS Physics: Technical Design Report Volume 1: Detector Performance and Software”. 2006, <https://cds.cern.ch/record/922757>.
- [90] CMS Collaboration, T. Sakuma, “Cutaway diagrams of CMS detector”. 2019, <https://cds.cern.ch/record/2665537>. Modified.
- [91] I. Neutelings, “CMS coordinate system”. https://tikz.net/axis3d_cms/. Accessed 13.04.2023.
- [92] CMS Tracker Group Collaboration, “The CMS Phase-1 Pixel Detector Upgrade”, *JINST* **16** (2021), doi:10.1088/1748-0221/16/02/P02027, arXiv:2012.14304.
- [93] CMS Collaboration, CMS Collaboration, “Sketches of the CMS Tracker Detector”. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/DPGResultsTRK>. Accessed 14.04.2023, Modified.
- [94] CMS Collaboration, “Strategies and performance of the CMS silicon tracker alignment during LHC Run 2”, *Nucl. Instrum. Meth. A* **1037** (2022), doi:10.1016/j.nima.2022.166795, arXiv:2111.08757.
- [95] CMS Collaboration, “2017 tracking performance plots”, CMS Detector Performance Note CMS-DP-2017-015, CERN (2017). <https://cds.cern.ch/record/2290524>.
- [96] CMS Collaboration, “Performance and Operation of the CMS Electromagnetic Calorimeter”, *JINST* **5** (2010), doi:10.1088/1748-0221/5/03/T03010, arXiv:0910.3423.
- [97] CMS Collaboration, “Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC”, *JINST* **16** (2021), doi:10.1088/1748-0221/16/05/P05014, arXiv:2012.06888.
- [98] CMS Collaboration, “Calibration of the CMS hadron calorimeters using proton-proton collision data at $\sqrt{s} = 13$ TeV”, *JINST* **15** (2020), doi:10.1088/1748-0221/15/05/P05002, arXiv:1910.00079.
- [99] V. D. Elvira, “Measurement of the Pion Energy Response and Resolution in the CMS HCAL Test Beam 2002 Experiment”, CMS-NOTE-2004-020, CERN, Geneva (2004). <http://cds.cern.ch/record/800406>.
- [100] CMS Collaboration, “Design, Performance, and Calibration of CMS Hadron Endcap Calorimeters”, CMS-NOTE-2008-010, CERN, Geneva (2008). <https://cds.cern.ch/record/1103003>.

- [101] CMS Collaboration, “Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV”, *JINST* **13** (2018), doi:10.1088/1748-0221/13/06/P06015, arXiv:1804.04528.
- [102] CMS Collaboration, “Performance of the CMS muon trigger system in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *JINST* **16** (2021), doi:10.1088/1748-0221/16/07/P07001, arXiv:2102.04790.
- [103] CMS Collaboration, “The CMS trigger system”, *JINST* **12** (2017), doi:10.1088/1748-0221/12/01/P01020, arXiv:1609.02366.
- [104] CMS Collaboration, “Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *JINST* **15** (2020), doi:10.1088/1748-0221/15/10/P10017, arXiv:2006.10165.
- [105] CMS Collaboration, “Performance of the High-Level Trigger System at CMS in LHC Run-2”, *IEEE Trans. Nucl. Sci.* **68** (2021), doi:10.1109/TNS.2021.3087618.
- [106] CMS Collaboration, “Images for CMS Tracker alignment in Run 2”, CMS-PHO-EVENTS-2022-026, CERN (2022). <https://cds.cern.ch/record/2812955>.
- [107] V. Blobel and C. Kleinwort, “A New method for the high precision alignment of track detectors”, in *Conference on Advanced Statistical Techniques in Particle Physics*. 2002. arXiv:hep-ex/0208021.
- [108] V. Blobel, “Software alignment for tracking detectors”, *Nucl. Instrum. Meth. A* **566** (2006), doi:10.1016/j.nima.2006.05.157.
- [109] Millepede, “Millepede II”. <https://www.desy.de/~kleinwrt/MP2/doc/html/index.html>. Accessed 08.05.2023.
- [110] V. Karimaki, A. Heikkinen, T. Lampen, and T. Linden, “Sensor alignment by tracks”, *eConf* **C0303241** (2003), arXiv:physics/0306034.
- [111] V. Karimäki, T. Lampen, and F.-P. Schilling, “The HIP Algorithm for Track Based Alignment and its Application to the CMS Pixel Detector”, CMS-NOTE-2006-018, CERN, Geneva (2006). <https://cds.cern.ch/record/926537>.
- [112] CMS Collaboration, “CMS Tracker Performance results for full Run 2 Legacy reprocessing”, CMS-DP-2020-012, CERN (2020). <http://cds.cern.ch/record/2713208>.
- [113] L. Wilke, V. Chiochia, and T. Speer, “Lorentz angle calibration for the barrel pixel detector”, CMS-NOTE-2008-012, CERN, Geneva (2008). <http://cds.cern.ch/record/1103004>.
- [114] G. Cerminara and B. van Besien, “Automated workflows for critical time-dependent calibrations at the CMS experiment”, *J. Phys. Conf. Ser.* **664** (2015), doi:10.1088/1742-6596/664/7/072009.
- [115] CMS Collaboration, “Tracker Alignment Performance in 2022”, CMS-DP-2022-044, CERN (2022). <https://cds.cern.ch/record/2839739>.
- [116] CMS Collaboration, “Precision luminosity measurement in proton-proton collisions at $\sqrt{s} = 13$ TeV in 2015 and 2016 at CMS”, *Eur. Phys. J. C* **81** (2021), doi:10.1140/epjc/s10052-021-09538-2, arXiv:2104.01927.

- [117] CMS Collaboration, “CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13$ TeV”, CMS-PAS-LUM-17-004, CERN, Geneva (2018).
<https://cds.cern.ch/record/2621960>.
- [118] CMS Collaboration, “CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV”, CMS-PAS-LUM-18-002, CERN, Geneva (2019).
<https://cds.cern.ch/record/2676164>.
- [119] CMS Collaboration, “Simulation of the Silicon Strip Tracker pre-amplifier in early 2016 data”, CMS-DP-2020-04, CERN (2020). <https://cds.cern.ch/record/2740688>.
- [120] M. Jansova, “Search for the supersymmetric partner of the top quark and measurements of cluster properties in the silicon strip tracker of the CMS experiment at Run 2.”. PhD thesis, Strasbourg, IPHC, 2018. <https://cds.cern.ch/record/2647308>.
- [121] CMS Collaboration, “Public CMS Data Quality Information”.
<https://twiki.cern.ch/twiki/bin/view/CMSPublic/DataQuality>.
Accessed 15.05.2023.
- [122] CMS Collaboration, “The CMS dataset bookkeeping service”, *J. Phys. Conf. Ser.* **119** (2008), doi:10.1088/1742-6596/119/7/072001.
- [123] A. Buckley et al., “General-purpose event generators for LHC physics”, *Phys. Rept.* **504** (2011), doi:10.1016/j.physrep.2011.03.005, arXiv:1101.2599.
- [124] P. Nason, “A New method for combining NLO QCD with shower Monte Carlo algorithms”, *JHEP* **11** (2004), doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.
- [125] S. Frixione, P. Nason, and C. Oleari, “Matching NLO QCD computations with Parton Shower simulations: the POWHEG method”, *JHEP* **11** (2007), doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.
- [126] S. Alioli, P. Nason, C. Oleari, and E. Re, “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX”, *JHEP* **06** (2010), doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.
- [127] S. Frixione, P. Nason, and G. Ridolfi, “A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction”, *JHEP* **09** (2007), doi:10.1088/1126-6708/2007/09/126, arXiv:0707.3088.
- [128] E. Re, “Single-top Wt-channel production matched with parton showers using the POWHEG method”, *Eur. Phys. J. C* **71** (2011), doi:10.1140/epjc/s10052-011-1547-z, arXiv:1009.2450.
- [129] S. Alioli, P. Nason, C. Oleari, and E. Re, “NLO single-top production matched with shower in POWHEG: s- and t-channel contributions”, *JHEP* **09** (2009), doi:10.1088/1126-6708/2009/09/111, arXiv:0907.4076. [Erratum: JHEP 02, 011 (2010)].
- [130] J. Alwall et al., “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”, *JHEP* **07** (2014), doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.
- [131] R. Frederix and S. Frixione, “Merging meets matching in MC@NLO”, *JHEP* **12** (2012), doi:10.1007/JHEP12(2012)061, arXiv:1209.6215.

- [132] J. Alwall et al., “Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions”, *Eur. Phys. J. C* **53** (2008), doi:10.1140/epjc/s10052-007-0490-5, arXiv:0706.2569.
- [133] T. Sjostrand, S. Mrenna, and P. Z. Skands, “A Brief Introduction to PYTHIA 8.1”, *Comput. Phys. Commun.* **178** (2008), doi:10.1016/j.cpc.2008.01.036, arXiv:0710.3820.
- [134] NNPDF Collaboration, “Parton distributions from high-precision collider data”, *Eur. Phys. J. C* **77** (2017), doi:10.1140/epjc/s10052-017-5199-5, arXiv:1706.00428.
- [135] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, “Parton Fragmentation and String Dynamics”, *Phys. Rept.* **97** (1983), doi:10.1016/0370-1573(83)90080-7.
- [136] CMS Collaboration, “Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements”, *Eur. Phys. J. C* **80** (2020), doi:10.1140/epjc/s10052-019-7499-4, arXiv:1903.12179.
- [137] CMS Collaboration, “CMS pythia 8 colour reconnection tunes based on underlying-event data”, *Eur. Phys. J. C* **83** (2023), doi:10.1140/epjc/s10052-023-11630-8, arXiv:2205.02905.
- [138] GEANT4 Collaboration, “GEANT4—a simulation toolkit”, *Nucl. Instrum. Meth. A* **506** (2003), doi:10.1016/S0168-9002(03)01368-8.
- [139] M. Hildreth, V. N. Ivanchenko, D. J. Lange, and M. J. Kortelainen, “CMS Full Simulation for Run-2”, *J. Phys. Conf. Ser.* **664** (2015), doi:10.1088/1742-6596/664/7/072022.
- [140] A. Giammanco, “The Fast Simulation of the CMS Experiment”, *J. Phys. Conf. Ser.* **513** (2014), doi:10.1088/1742-6596/513/2/022012.
- [141] CMS Collaboration, “Utilities for Accessing Pileup Information for Data”. https://twiki.cern.ch/twiki/bin/view/CMS/PileupJSONFileforData#Recommended_cross_section. Accessed 14.09.2023.
- [142] M. Cacciari et al., “Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation”, *Phys. Lett. B* **710** (2012), doi:10.1016/j.physletb.2012.03.013, arXiv:1111.5869.
- [143] P. Bärnreuther, M. Czakon, and A. Mitov, “Percent Level Precision Physics at the Tevatron: First Genuine NNLO QCD Corrections to $q\bar{q} \rightarrow t\bar{t} + X$ ”, *Phys. Rev. Lett.* **109** (2012), doi:10.1103/PhysRevLett.109.132001, arXiv:1204.5201.
- [144] M. Czakon and A. Mitov, “NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels”, *JHEP* **12** (2012), doi:10.1007/JHEP12(2012)054, arXiv:1207.0236.
- [145] M. Czakon and A. Mitov, “NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction”, *JHEP* **01** (2013), doi:10.1007/JHEP01(2013)080, arXiv:1210.6832.
- [146] M. Czakon, P. Fiedler, and A. Mitov, “Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$ ”, *Phys. Rev. Lett.* **110** (2013), doi:10.1103/PhysRevLett.110.252004, arXiv:1303.6254.

- [147] M. Czakon et al., “Top-pair production at the LHC through NNLO QCD and NLO EW”, *JHEP* **10** (2017), doi:10.1007/JHEP10(2017)186, arXiv:1705.04105.
- [148] CMS Collaboration, “The modeling of the top quark pT”.
<https://twiki.cern.ch/twiki/bin/view/CMS/TopPtRewighting>.
Accessed 24.05.2023.
- [149] N. Kidonakis, “Two-loop soft anomalous dimensions for single top quark associated production with a W^- or H^- ”, *Phys. Rev. D* **82** (2010),
doi:10.1103/PhysRevD.82.054018, arXiv:1005.4451.
- [150] N. Kidonakis, “Top Quark Production”, in *Helmholtz International Summer School on Physics of Heavy Quarks and Hadrons*. 2014. arXiv:1311.0283.
doi:10.3204/DESY-PROC-2013-03/Kidonakis.
- [151] M. Aliev et al., “HATHOR: HAdronic Top and Heavy quarks crOss section calculatoR”, *Comput. Phys. Commun.* **182** (2011), doi:10.1016/j.cpc.2010.12.040,
arXiv:1007.1327.
- [152] P. Kant et al., “HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions”, *Comput. Phys. Commun.* **191** (2015), doi:10.1016/j.cpc.2015.02.001, arXiv:1406.4403.
- [153] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, “FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order”, *Comput. Phys. Commun.* **182** (2011),
doi:10.1016/j.cpc.2011.06.008, arXiv:1011.3540.
- [154] Y. Li and F. Petriello, “Combining QCD and electroweak corrections to dilepton production in FEWZ”, *Phys. Rev. D* **86** (2012), doi:10.1103/PhysRevD.86.094034,
arXiv:1208.5967.
- [155] T. Gehrmann et al., “ W^+W^- Production at Hadron Colliders in Next to Next to Leading Order QCD”, *Phys. Rev. Lett.* **113** (2014),
doi:10.1103/PhysRevLett.113.212001, arXiv:1408.5243.
- [156] J. M. Campbell and R. K. Ellis, “An Update on vector boson pair production at hadron colliders”, *Phys. Rev. D* **60** (1999), doi:10.1103/PhysRevD.60.113006,
arXiv:hep-ph/9905386.
- [157] J. M. Campbell, R. K. Ellis, and C. Williams, “Vector boson pair production at the LHC”, *JHEP* **07** (2011), doi:10.1007/JHEP07(2011)018, arXiv:1105.0020.
- [158] W. Beenakker et al., “NNLL resummation for stop pair-production at the LHC”, *JHEP* **05** (2016), doi:10.1007/JHEP05(2016)153, arXiv:1601.02954.
- [159] W. Beenakker et al., “NNLL-fast: predictions for coloured supersymmetric particle production at the LHC with threshold and Coulomb resummation”, *JHEP* **12** (2016),
doi:10.1007/JHEP12(2016)133, arXiv:1607.07741.
- [160] S. Frixione et al., “Single-top hadroproduction in association with a W boson”, *JHEP* **07** (2008), doi:10.1088/1126-6708/2008/07/029, arXiv:0805.3067.
- [161] E. Re, “Single-top Wt-channel production matched with parton showers using the POWHEG method”, *Eur. Phys. J. C* **71** (2011),
doi:10.1140/epjc/s10052-011-1547-z, arXiv:1009.2450.

- [162] T. Ježo et al., “An NLO+PS generator for $t\bar{t}$ and Wt production and decay including non-resonant and interference effects”, *Eur. Phys. J. C* **76** (2016), doi:10.1140/epjc/s10052-016-4538-2, arXiv:1607.04538.
- [163] R. D. Cousins, K. E. Hymes, and J. Tucker, “Frequentist evaluation of intervals estimated for a binomial parameter and for the ratio of Poisson means”, *Nucl. Instrum. Meth. A* **612** (2010), doi:10.1016/j.nima.2009.10.156, arXiv:0905.3831.
- [164] CMS Collaboration, “Particle-flow reconstruction and global event description with the CMS detector”, *JINST* **12** (2017), doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.
- [165] N. Bartosik, “Associated Top-Quark-Pair and b-Jet Production in the Dilepton Channel at $\sqrt{s} = 8$ TeV as Test of QCD and Background to $t\bar{t}$ +Higgs Production”. PhD thesis, Universität Hamburg, 2015. Modified, <https://cds.cern.ch/record/2047049>.
- [166] W. Adam, B. Mangano, T. Speer, and T. Todorov, “Track Reconstruction in the CMS tracker”, CMS-NOTE-2006-041, CERN, Geneva (2006). <https://cds.cern.ch/record/934067>.
- [167] CMS Collaboration, “Description and performance of track and primary-vertex reconstruction with the CMS tracker”, *JINST* **9** (2014), doi:10.1088/1748-0221/9/10/P10009, arXiv:1405.6569.
- [168] K. Rose, “Deterministic annealing for clustering, compression, classification, regression, and related optimization problems”, *IEEE Proc.* **86** (1998), doi:10.1109/5.726788.
- [169] R. Frühwirth, W. Waltenberger, and P. Vanlaer, “Adaptive Vertex Fitting”, CMS-NOTE-2007-008, CERN, Geneva (2007). <https://cds.cern.ch/record/1027031>.
- [170] W. Adam, R. Frühwirth, A. Strandlie, and T. Todorov, “Reconstruction of electrons with the Gaussian sum filter in the CMS tracker at LHC”, *eConf* **C0303241** (2003), doi:10.1088/0954-3899/31/9/N01, arXiv:physics/0306087.
- [171] CMS Collaboration, “Performance of Photon Reconstruction and Identification with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV”, *JINST* **10** (2015), doi:10.1088/1748-0221/10/08/P08010, arXiv:1502.02702.
- [172] CMS Collaboration, “Cut Based Electron ID for Run 2”. <https://twiki.cern.ch/twiki/bin/view/CMS/CutBasedElectronIdentificationRun2>. Accessed 31.05.2023.
- [173] CMS Collaboration, “Electron and photon scale factors for Run 2”. <https://twiki.cern.ch/twiki/bin/view/CMS/EgammaUL2016To2018>. Accessed 06.06.2023.
- [174] CMS Collaboration, “Baseline muon selections for Run-II”. <https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideMuonIdRun2>. Accessed 31.05.2023.
- [175] CMS Collaboration, “Muon efficiencies and scale factors for Run-II”. https://twiki.cern.ch/twiki/bin/view/CMS/MuonPOG#User_Recommendations. Accessed 06.06.2023.

- [176] A. Bodek et al., “Extracting Muon Momentum Scale Corrections for Hadron Collider Experiments”, *Eur. Phys. J. C* **72** (2012), doi:10.1140/epjc/s10052-012-2194-8, arXiv:1208.3710.
- [177] CMS Collaboration, “Rochester corrections”.
<https://twiki.cern.ch/twiki/bin/view/CMS/RochcorMuon>. Accessed 31.05.2023.
- [178] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- k_t jet clustering algorithm”, *JHEP* **04** (2008), doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.
- [179] M. Cacciari, G. P. Salam, and G. Soyez, “FastJet User Manual”, *Eur. Phys. J. C* **72** (2012), doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.
- [180] CMS Collaboration, “Pileup mitigation at CMS in 13 TeV data”, *JINST* **15** (2020), doi:10.1088/1748-0221/15/09/P09018, arXiv:2003.00503.
- [181] CMS Collaboration, “Jet identification in high pile-up environment for Ultra Legacy Data”. <https://twiki.cern.ch/twiki/bin/viewauth/CMS/PileupJetIDUL>. Accessed 01.06.2023.
- [182] CMS Collaboration, “Jet Identification for the 13 TeV UL data”.
<https://twiki.cern.ch/twiki/bin/view/CMS/JetID13TeVUL>. Accessed 01.06.2023.
- [183] CMS Collaboration, “Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV”, *JINST* **12** (2017), doi:10.1088/1748-0221/12/02/P02014, arXiv:1607.03663.
- [184] CMS Collaboration, “Recommended Jet Energy Corrections and Uncertainties”.
<https://twiki.cern.ch/twiki/bin/view/CMS/JECDataMC>. Accessed 01.06.2023.
- [185] CMS Collaboration, “Jet Energy Resolution”.
<https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetResolution>. Accessed 01.06.2023.
- [186] CMS Collaboration, “Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV”, *JINST* **13** (2018), doi:10.1088/1748-0221/13/05/P05011, arXiv:1712.07158.
- [187] E. Bols et al., “Jet Flavour Classification Using DeepJet”, *JINST* **15** (2020), doi:10.1088/1748-0221/15/12/P12012, arXiv:2008.10519.
- [188] CMS Collaboration, “Methods to apply b-tagging efficiency scale factors”.
<https://twiki.cern.ch/twiki/bin/view/CMS/BTagSFMethods>. Accessed 01.06.2023.
- [189] CMS Collaboration, “Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13$ TeV using the CMS detector”, *JINST* **14** (2019), doi:10.1088/1748-0221/14/07/P07004, arXiv:1903.06078.
- [190] CMS Collaboration, “MET Corrections and Uncertainties for Run-II”.
<https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETRun2Corrections>. Accessed 02.06.2023.
- [191] CMS Collaboration, “Performance of missing transverse momentum reconstruction in events containing a photon and jets collected by CMS during proton-proton collisions at $\sqrt{s} = 13$ TeV in 2018”, CMS-DP-2020-031, CERN (2020).
<https://cds.cern.ch/record/2723010>.

- [192] Y. Feng, “A New Deep-Neural-Network–Based Missing Transverse Momentum Estimator, and its Application to W Recoil”. PhD thesis, University of Maryland, 2020. <https://cds.cern.ch/record/2744871>.
- [193] CMS Collaboration, “DeepMET”. <https://twiki.cern.ch/twiki/bin/view/CMS/DeepPFMET>. Accessed 02.06.2023.
- [194] P. Nattland, “Determination of missing transverse momentum using multivariate methods for a differential top pair production cross section measurement at CMS”, Master’s thesis, RWTH Aachen University, 2023. <https://cds.cern.ch/record/2863107>.
- [195] I. Goodfellow, Y. Bengio, and A. Courville, “Deep Learning”. MIT Press, 2016. <http://www.deeplearningbook.org>.
- [196] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning”, *Nature* **521** (2015), doi:10.1038/nature14539.
- [197] P. Mehta et al., “A high-bias, low-variance introduction to Machine Learning for physicists”, *Physics Reports* **810** (2019), doi:<https://doi.org/10.1016/j.physrep.2019.03.001>.
- [198] I. Neutelings, “Neural networks”. https://tikz.net/neural_networks/. Accessed 12.06.2023, Modified.
- [199] R. A. Saleh and A. K. M. Ehsanes Saleh, “Statistical Properties of the log-cosh Loss Function Used in Machine Learning”, *arXiv e-prints* (2022), arXiv:2208.04564.
- [200] D. P. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization”, *arXiv e-prints* (2014), arXiv:1412.6980.
- [201] M. Abadi et al., “TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems”. 2015, <https://www.tensorflow.org/>.
- [202] F. Chollet et al., “Keras”. 2015, <https://keras.io>.
- [203] F. Pedregosa et al., “Scikit-learn: Machine Learning in Python”. 2011, <http://jmlr.org/papers/v12/pedregosa11a.html>.
- [204] Wes McKinney, “Data Structures for Statistical Computing in Python”, in *Proceedings of the 9th Python in Science Conference*. 2010. doi:10.25080/Majora-92bf1922-00a.
- [205] L. S. Shapley, “A Value for n-Person Games”, ch. 17, pp. 307–318. Princeton University Press, Princeton, 1953. doi:doi:10.1515/9781400881970-018.
- [206] G. Erion et al., “Improving performance of deep learning models with axiomatic attribution priors and expected gradients”, *Nature Machine Intelligence* **3** (2021), doi:10.1038/s42256-021-00343-w.
- [207] The LHC Higgs Combination Group, “Procedure for the LHC Higgs boson search combination in Summer 2011”, CMS-NOTE-2011-005, CERN, Geneva (2011). <https://cds.cern.ch/record/1379837>.
- [208] CMS Collaboration, “Documentation CMS combine tool”. <http://cms-analysis.github.io/HiggsAnalysis-CombinedLimit/>. Accessed 19.06.2023.

- [209] J. S. Conway, “Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra”, in *PHYSTAT 2011*. 2011. [arXiv:1103.0354](#).
[doi:10.5170/CERN-2011-006.115](#).
- [210] R. J. Barlow and C. Beeston, “Fitting using finite Monte Carlo samples”, *Comput. Phys. Commun.* **77** (1993), [doi:10.1016/0010-4655\(93\)90005-W](#).
- [211] J. Neyman and E. S. Pearson, “IX. On the problem of the most efficient tests of statistical hypotheses”, *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **231** (1933),
[doi:10.1098/rsta.1933.0009](#).
- [212] R. D. Cousins, “Generalization of Chisquare Goodness-of-Fit Test for Binned Data Using Saturated Models , with Application to Histograms”. 2013,
https://www.physics.ucla.edu/~cousins/stats/cousins_saturated.pdf.
- [213] C. G. Lester and D. J. Summers, “Measuring masses of semiinvisibly decaying particles pair produced at hadron colliders”, *Phys. Lett. B* **463** (1999),
[doi:10.1016/S0370-2693\(99\)00945-4](#), [arXiv:hep-ph/9906349](#).
- [214] A. Barr, C. Lester, and P. Stephens, “m(T2): The Truth behind the glamour”, *J. Phys. G* **29** (2003), [doi:10.1088/0954-3899/29/10/304](#), [arXiv:hep-ph/0304226](#).
- [215] J. Snoek, H. Larochelle, and R. P. Adams, “Practical Bayesian Optimization of Machine Learning Algorithms”, *arXiv e-prints* (2012), [arXiv:1206.2944](#).
- [216] E. Brochu, V. M. Cora, and N. de Freitas, “A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning”, *arXiv e-prints* (2010), [arXiv:1012.2599](#).
- [217] CMS Collaboration, “Object definitions for top quark analyses at the particle level”, CMS-NOTE-2017-004, CERN, Geneva (2017). <https://cds.cern.ch/record/2267573>.
- [218] CMS Collaboration, “Measurements of $t\bar{t}$ differential cross sections in proton-proton collisions at $\sqrt{s} = 13$ TeV using events containing two leptons”, *JHEP* **02** (2019),
[doi:10.1007/JHEP02\(2019\)149](#), [arXiv:1811.06625](#).
- [219] V. Blobel, “Unfolding”, ch. 6, pp. 187–225. 2013. [doi:10.1002/9783527653416.ch6](#).
- [220] S. Schmitt, “TUnfold: an algorithm for correcting migration effects in high energy physics”, *JINST* **7** (2012), [doi:10.1088/1748-0221/7/10/T10003](#), [arXiv:1205.6201](#).
- [221] Particle Data Group, “Review of Particle Physics”, *Progress of Theoretical and Experimental Physics* **2020** (08, 2020), [doi:10.1093/ptep/ptaa104](#).
- [222] CMS Collaboration, “Measurement of differential cross sections for the production of top quark pairs and of additional jets in pp collisions at $\sqrt{s} = 13$ TeV”, CMS-PAS-TOP-20-006, CERN, Geneva (2022). <http://cds.cern.ch/record/2803771>.
- [223] ATLAS Collaboration, “Inclusive and differential cross-sections for dilepton $t\bar{t}$ production measured in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector”, *JHEP* **07** (2023), [doi:10.1007/JHEP07\(2023\)141](#), [arXiv:2303.15340](#).
- [224] CMS Collaboration, “Measurements of $t\bar{t}$ differential cross sections in proton-proton collisions at $\sqrt{s} = 13$ TeV using events containing two leptons”, *JHEP* **02** (2019),
[doi:10.1007/JHEP02\(2019\)149](#), [arXiv:1811.06625](#).

Eidesstattliche Erklärung

Danilo Meuser

erklärt hiermit, dass diese Dissertation und die darin dargelegten Inhalte die eigenen sind und selbstständig, als Ergebnis der eigenen originären Forschung, generiert wurden.

Hiermit erkläre ich an Eides statt

1. Diese Arbeit wurde vollständig oder größtenteils in der Phase als Doktorand dieser Fakultät und Universität angefertigt;
2. Sofern irgendein Bestandteil dieser Dissertation zuvor für einen akademischen Abschluss oder eine andere Qualifikation an dieser oder einer anderen Institution verwendet wurde, wurde dies klar angezeigt;
3. Wenn immer andere eigene- oder Veröffentlichungen Dritter herangezogen wurden, wurden diese klar benannt;
4. Wenn aus anderen eigenen- oder Veröffentlichungen Dritter zitiert wurde, wurde stets die Quelle hierfür angegeben. Diese Dissertation ist vollständig meine eigene Arbeit, mit der Ausnahme solcher Zitate;
5. Alle wesentlichen Quellen von Unterstützung wurden benannt;
6. Wenn immer ein Teil dieser Dissertation auf der Zusammenarbeit mit anderen basiert, wurde von mir klar gekennzeichnet, was von anderen und was von mir selbst erarbeitet wurde;
7. Kein Teil dieser Arbeit wurde vor deren Einreichung veröffentlicht.

Datum

Unterschrift