https://doi.org/10.15388/vu.thesis.991
https://orcid.org/0000-0003-2449-0158

VILNIUS UNIVERSITY
CENTER FOR PHYSICAL SCIENCES AND TECHNOLOGY

Marijus Ambrozas

Precision of the Drell–Yan process measurement with the CMS Phase–2 upgrade

DOCTORAL DISSERTATION

Natural Sciences,
Physics (N 002)

VILNIUS 2026

The dissertation was prepared between 2020 and 2025 at Vilnius University.

Academic supervisor – Dr. Andrius Juodagalvis (Vilnius University, Natural Sciences, Physics, N 002)

This doctoral dissertation will be defended in a public/closed meeting of the Dissertation Defence Panel:

Chairman – Prof. Dr. Egidijus Anisimovas (Vilnius University, Natural Sciences, Physics, N 002).

Members:

Dr. Darius Jurčiukonis (Vilnius University, Natural Sciences, Physics, N 002),
Dr. Mindaugas Karčiauskas (Center for Physical Sciences and Technology, Natural Sciences, Physics, N 002),
Dr. Mindaugas Šarpis (Vilnius University, Natural Sciences, Physics, N 002),
Dr. Laura Šerkšnytė (CERN, Switzerland, Natural Sciences, Physics, N 002).

The dissertation shall be defended at a public meeting of the Dissertation Defence Panel at 14:00 on the 25th of September 2026 in meeting room A101 of the National Center for Physical Sciences and Technology.
Address: Saulėtekio av. 3, Vilnius, Lithuania

https://doi.org/10.15388/vu.thesis.991
https://orcid.org/0000-0003-2449-0158

VILNIAUS UNIVERSITETAS
FIZINIŲ IR TECHNOLOGIJOS MOKSLŲ CENTRAS

Marijus Ambrozas

CMS detektoriaus II fazės atnaujinimo įtaka Drell–Yan proceso matavimo tikslumui

DAKTARO DISERTACIJA

Gamtos mokslai,
Fizika (N 002)

VILNIUS 2026

Disertacija rengta 2020–2025 metais Vilniaus universitete.

Mokslinis vadovas – dr. Andrius Juodagalvis (Vilniaus universitetas, gamtos mokslai, fizika, N 002).

Gynimo taryba:
Pirmininkas – prof. dr. Egidijus Anisimovas (Vilniaus universitetas, gamtos mokslai, fizika, N 002).
Nariai:
dr. Darius Jurčiukonis (Vilniaus universitetas, gamtos mokslai, fizika, N 002),
dr. Mindaugas Karčiauskas (Fizinių ir technologijos mokslų centras, gamtos mokslai, fizika, N 002),
dr. Mindaugas Šarpis (Vilniaus universitetas, gamtos mokslai, fizika, N 002),
dr. Laura Šerkšnytė (CERN, Šveicarija, gamtos mokslai, fizika, N 002).

Disertacija ginama viešame Gynimo tarybos posėdyje 2026 m.
rugsėjo mėn. 25 d. 14:00 val. Nacionalinio fizinių ir technologijos mokslų centro A101 posėdžių salėje. Adresas: Saulėtekio al. 3, Vilnius, Lietuva.

ABBREVIATIONS

BLUE	Best Linear Unbiased Estimator
BSM	Beyond the Standard Model
CERN	European Organization for Nuclear Research
CKM	Cabibbo–Kobayashi–Maskawa
CMS	Compact Muon Solenoid
CMSSW	CMS Software
CP	Charge-conjugation and parity symmetry
CROC	CMS Readout Chip
CSC	Cathode Strip Chamber
DAC	Digital-to-Analog Converter
DAQ	Data Acquisition
DGLAP	Dokshitzer–Gribov–Lipatov–Altarelli–Parisi
DQM	Data Quality Monitoring
DT	Drift Tube
DTC	Data, Trigger, and Control
DY	Drell–Yan
ECAL	CMS Electromagnetic Calorimeter
EW	Electroweak
FMC	FPGA Mezzanine Card
FPGA	Field-Programmable Gate Array
GEM	Gas Electron Multiplier
GSF	Gaussian Sum Filter
GUT	Grand Unified Theory
HB	HCAL Barrel
HCAL	CMS Hadron Calorimeter
HE	HCAL Endcap
HF	HCAL Forward
HGCAL	High Granularity Calorimeter
HL-LHC	High-Luminosity Large Hadron Collider
HLT	High-Level Trigger
HO	HCAL Outer
ID	Identification
IT	CMS Inner Tracker
L1	Level–1
LHC	Large Hadron Collider
LIPS	Lorentz-Invariant Phase Space

LO	Leading Order
LS	Long Shutdown
lpGBT	Low-Power Gigabit Transceiver
MC	Monte Carlo
MiNNLO	Multi-Scale Improved Next-to-Next-to-Leading Order
MIP	Minimum Ionizing Particle
MTD	MIP Timing Detector
NLO	Next-to-Leading Order
NNLO	Next-to-Next-to-Leading Order
OS	Opposite-Sign
OT	CMS Outer Tracker
PDF	Parton Distribution Function
PF	Particle Flow
Ph2-ACF	Phase-2 Acquisition and Control Framework
POG	Physics Object Group
PS	Proton Synchrotron
PSB	Proton Synchrotron Booster
PU	Pileup
QCD	Quantum Chromodynamics
QED	Quantum Electrodynamics
QFT	Quantum Field Theory
RPC	Resistive Plate Chamber
SCC	Single-Chip Card
SF	Scale Factor
SM	Standard Model
SMEFT	Standard Model Effective Field Theory
SPS	Super Proton Synchrotron
SS	Same-Sign
SUSY	Supersymmetry
TEPX	Tracker Endcap Pixel detector
TMD	Transverse-Momentum-Dependent
ToT	Time-over-Threshold
TxRx	Transceiver
UL	Ultra Legacy
UV	Ultraviolet
VEV	Vacuum Expectation Value
VMUX	Voltage Multiplexer
VTRx+	Versatile Transceiver Plus
WP	Working Point

CONTENTS

ABBREVIATIONS 5
INTRODUCTION 9
Main goal and tasks 12
Novelty and relevance 13
Statements of the thesis 13
THEORETICAL BACKGROUND 15
1 Standard Model of Particle Physics 15
2 SM Calculations 33
3 Drell–Yan Process 48
EXPERIMENTAL FRAMEWORK 56
4 The Large Hadron Collider at CERN 56
5 The Compact Muon Solenoid experiment 61
DRELL–YAN PROCESS MEASUREMENT 74
6 Analysis strategy 74
7 Data and simulated samples 76
8 Corrections 79
9 Event selection 80
10 Background estimation 88
11 Unfolding 118
12 Uncertainty estimation 122
13 Channel combination 125
14 Cross section 127
THE PHASE-2 UPGRADE OF THE CMS DETECTOR 130
15 High-Luminosity LHC 130
16 Phase–2 challenges for the CMS detector 132
17 CMS Phase–2 upgrade 135
PREPARATION WORKS FOR THE INNER TRACKER UPGRADE 143
18 Module test setups 143
19 Ph2-ACF Software 143
20 Prototype testing results 148
PRECISION OF THE DRELL–YAN PROCESS MEASUREMENT WITH THE CMS PHASE–2 UPGRADE 159
21 Simulation setup 159
22 Muon reconstruction efficiency 160
23 Muon and dimuon resolution 161
24 Conclusion 163
CONCLUSIONS 168

REFERENCES 173
SANTRAUKA LIETUVIŲ KALBA 190
CURRICULUM VITAE 207
ACKNOWLEDGEMENTS/PADĖKOS 211
LIST OF WoS PUBLICATIONS 213
LIST OF OTHER PUBLICATIONS 213
LIST OF PRESENTATIONS 214

INTRODUCTION

The Standard Model (SM) of particle physics is a comprehensive theory that describes the known fundamental particles and their interactions, excluding gravity. It classifies elementary particles into quarks, leptons, vector gauge bosons, and the scalar Higgs boson. Within this framework, the electromagnetic, weak, and strong interactions are described at the quantum level, providing the foundation for the current understanding of the fundamental constituents of matter.

The SM has made numerous predictions that have been experimentally confirmed, including the existence of the W and Z bosons [1–4], the top quark [5, 6], and the Higgs boson, discovered at the Large Hadron Collider (LHC) in 2012 [7, 8]. These discoveries have demonstrated the predictive power of the SM and have guided experimental particle physics over the past several decades. Precise measurements of the properties of SM particles continue to provide stringent tests of the theory.

Despite its remarkable success, the SM leaves several fundamental questions unanswered. It does not provide a particle candidate for dark matter, does not explain dark energy, and does not account for the observed matter–antimatter asymmetry in the universe. In its minimal form, the SM also does not explain neutrino oscillations. Furthermore, it does not incorporate gravity and does not provide a unified description of all fundamental interactions. These open questions motivate both new theoretical developments and increasingly precise experimental investigations.

The Large Hadron Collider (LHC) at CERN [9], together with its main experiments CMS [10], ATLAS [11], LHCb [12], and ALICE [13], plays a central role in this experimental program. The discovery of the Higgs boson by the ATLAS and CMS collaborations was a major milestone, completing the set of particles predicted by the SM. Since then, LHC measurements have continued to test the SM with increasing precision, while searches for physics beyond the SM probe a wide range of possible extensions.

The Drell–Yan (DY) process [14], in which charged-lepton pairs are produced through quark-antiquark annihilation via an intermediate virtual photon or Z boson, is one of the most important standard-candle processes at hadron colliders. It has a clean experimental signature and a relatively simple topology, allowing precise measurements over a wide kinematic range. Theoretical predictions for DY production are available at high perturbative precision, including next-to-next-to-leading-order (NNLO) QCD calculations for inclusive and differential observables. Therefore, measurements of DY cross sections

provide stringent tests of the SM.

Precision measurements of DY differential cross sections are also important for constraining parton distribution functions (PDFs), which describe the momentum distributions of quarks and gluons inside the proton. In addition, DY measurements test the perturbative framework of quantum chromodynamics (QCD) and the electroweak (EW) sector of the SM. They also provide important benchmarks for searches for physics beyond the SM, since many new-physics signatures are studied in final states containing charged leptons.

Differential DY cross sections have been measured by the ATLAS and CMS collaborations using proton-proton collision data collected at center-of-mass energies of 7 TeV [15, 16], 8 TeV [17, 18], and 13 TeV [19, 20, A1]. This thesis presents work related to the measurement of the DY differential cross section at a center-of-mass energy of 13 TeV, using data collected by the CMS detector during 2016–2018, corresponding to an integrated luminosity of approximately 137.6 fb^{-1} [B1]. The cross section is measured as a function of the dilepton invariant mass in the range from 40 GeV to 3000 GeV, using both dimuon and dielectron final states. A crucial part of this measurement is the estimation of background contributions, which directly affects the total uncertainty of the measured cross section and is one of the main focus points of this thesis.

The High-Luminosity LHC (HL-LHC) upgrade is designed to increase the collider luminosity by approximately a factor of ten, substantially enlarging the available data samples [21]. This will improve the statistical precision of many measurements, extend the discovery reach of the LHC, and open new possibilities for precision studies of rare SM processes. At the same time, the increased luminosity will lead to significantly higher radiation levels, particle occupancies, and pileup, requiring major upgrades of the LHC experiments.

The CMS Phase–2 upgrade is an extensive enhancement of the CMS detector designed to maintain and improve its performance under the challenging conditions expected at the HL-LHC [22]. The upgrade includes major improvements to the tracking system, calorimeters, muon detectors, trigger, and data acquisition systems. The new tracker will use advanced silicon detector technology with higher granularity, improved radiation hardness, and enhanced readout capabilities.

A critical component of the Phase–2 upgrade is the complete replacement of the CMS Inner Tracker. The upgraded Inner Tracker will use radiation-hard silicon pixel sensors and readout electronics designed to provide precise tracking information in the high-pileup HL-LHC environment. These improvements are essential for maintaining high tracking efficiency, accurate vertex reconstruction, and precise momentum measurements, all of which are important for precision

measurements such as DY production and for searches for new physics.

The upgraded Inner Tracker will use a radiation-hard pixel readout chip developed within the RD53 program [23]. A pre-production prototype for CMS, the RD53B-CMS chip, has been tested using the Ph2-ACF data acquisition and quality-assurance software framework [24]. Contributions to the development of the Ph2-ACF software and to RD53B-CMS prototype testing form another focus of this thesis.

The HL-LHC, together with the CMS Phase–2 upgrade, will provide new opportunities for precision measurements of the DY process. The larger data samples expected at the HL-LHC will reduce statistical uncertainties, while the upgraded detector is expected to improve the resolution and reconstruction performance relevant for dilepton measurements. Studying the expected impact of the upgraded CMS detector on DY measurements provides a connection between the detector-upgrade work and the long-term physics program of CMS.

This thesis is organized as follows. Section "Theoretical Background" introduces the theoretical concepts relevant for the thesis, including the SM, the proton structure described by PDFs, and the DY process. Section "Experimental Framework", describes the CERN accelerator complex, the LHC, and the CMS detector, with emphasis on the detector subsystems, trigger, and data acquisition systems.

Section "Drell-Yan Process Measurement" presents the measurement of the DY differential cross section using CMS Run–2 data. The section describes the event selection, corrections, background estimation, unfolding, systematic uncertainties, and channel combination, with particular emphasis on the background-estimation methods developed and validated by the author. Section "The Phase-2 Upgrade of the CMS Detector" introduces the HL-LHC project and the CMS Phase–2 upgrade. Section "Preparation Works for the Inner Tracker Upgrade" presents the author's contributions to the preparation of the CMS Inner Tracker upgrade, focusing on RD53B-CMS pixel readout chip testing and Ph2-ACF software development. Section "Precision of the Drell-Yan Process Measurement with the CMS Phase-2 Upgrade" discusses the prospects for DY measurements at the HL-LHC with the upgraded CMS detector. Finally, Section "Conclusions" summarizes the main results and conclusions of the thesis.

Main goal and tasks

The main goal of this thesis is twofold. The first part is to contribute to the measurement of the Drell–Yan differential cross section $\mathrm{d}\sigma/\mathrm{d}m$ at a center-of-mass energy of 13 TeV, using proton-proton collision data collected by the CMS detector during 2016–2018. This work was performed in collaboration with the CMS SMP-V Drell–Yan working group, with the main emphasis placed on the estimation of background contributions. Most prompt-lepton backgrounds are estimated using Monte Carlo simulation, while correction factors for the top-quark backgrounds are derived from data using the electron-muon final state. Backgrounds from misidentified or non-prompt leptons are estimated using data-driven methods.

The second part is to contribute to the CMS Phase–2 upgrade by testing the RD53B-CMS pixel readout chip prototype and developing calibration and quality-assurance tools within the Ph2-ACF software framework. Finally, the two parts of the thesis are connected by studying the prospects for DY measurement precision at the HL-LHC with the upgraded CMS detector.

To achieve these goals, the following tasks were completed:

1. Estimated background contributions to the DY signal, with particular emphasis on data-driven methods for misidentified or non-prompt lepton backgrounds.

2. Estimated the statistical and systematic uncertainties associated with the DY background estimates.

3. Contributed to the development of the Ph2-ACF software framework for RD53B-CMS pixel readout chip testing, including the implementation of a new calibration routine and the testing, debugging, and improvement of calibration routines developed by other group members.

4. Participated in RD53B-CMS pixel readout chip prototype testing, including quality-assurance measurements, performance optimization, and analysis of the obtained results.

5. Evaluated the muon reconstruction efficiency, muon transverse momentum and dimuon mass resolution relevant for the Drell-Yan measurement at the HL-LHC with the upgraded CMS detector using Monte Carlo simulations.

Novelty and relevance

The DY process is a well-known SM process that has been studied extensively at hadron colliders. Nevertheless, the measurement presented in this thesis aims to provide one of the most precise DY differential cross-section measurements in a wide dilepton mass range at a center-of-mass energy of 13 TeV. A crucial component of this measurement is the estimation of background contributions, which significantly affects the total uncertainty of the result.

Data-driven background-estimation techniques are well established at the LHC and have been used in many analyses. The novelty of the work presented here lies in the combination and refinement of methods used to estimate background contributions from processes with prompt leptons and from processes with misidentified or non-prompt leptons, in both dielectron and dimuon final states. The methods used in this thesis include improvements over more conventional approaches and allow similar estimation procedures to be used in both final states while maintaining good precision and accuracy.

The CMS Phase–2 upgrade is a major improvement of the CMS detector, designed to maintain and extend its physics performance under the increased data rates, radiation levels, and pileup conditions expected at the HL-LHC. The upgraded detector components will employ novel technologies, including new radiation-hard silicon sensors and readout electronics. This thesis presents tests of several key features of the RD53B-CMS pixel readout chip prototype, which is part of the CMS Inner Tracker upgrade. It also presents contributions to the development of the Ph2-ACF software framework, which supports quality-assurance tasks for CMS Tracker modules.

The DY prospects at the HL-LHC discussed in this thesis are novel because they connect the expected Phase–2 detector performance directly with the precision of future DY measurements. The study is based on recent Monte Carlo simulation workflows and considers the expected high-luminosity running conditions of the upgraded CMS detector.

Statements of the thesis

1. The presented variant of the same-sign method allows the misidentified or non-prompt lepton background to be estimated in both dielectron and dimuon final states with a systematic uncertainty below 50% for dilepton masses below 900 GeV.

2. For dilepton masses above 300 GeV, the uncertainty in the misidentified or non-prompt lepton background estimate is dominated by the statistical

uncertainty of the control samples.

3. The RD53B-CMS pixel readout chip prototype can be tuned to achieve a threshold dispersion approximately two times lower than that obtained with the RD53A demonstrator chip.

4. The signal gain of the RD53B-CMS chip shows linear behavior and does not depend on the hit-detection threshold.

5. The VCal digital-to-analog converter of the RD53B-CMS chip provides an injected-charge step size of approximately 5 electrons.

6. In addition to an expected increase of at least a factor of five in the available DY statistics relative to Run–2, the upgraded CMS detector is expected to improve the muon transverse momentum resolution by at least 30% and the dimuon mass resolution by at least 20% in the studied phase space.

THEORETICAL BACKGROUND

1. Standard Model of Particle Physics

The Standard Model (SM) of particle physics is the most successful theoretical framework currently available for describing the electromagnetic, weak, and strong interactions. It classifies elementary particles into fermions, which form the matter content of the theory, and bosons, which mediate interactions or arise from the Higgs mechanism. The gauge bosons of the SM are the photon, the $W^{\pm}$ and Z bosons, and the gluons, while the Higgs boson is a scalar particle associated with electroweak symmetry breaking.

The modern SM emerged from several major developments in quantum field theory (QFT). Quantum electrodynamics (QED), developed by Feynman, Schwinger, and Tomonaga [25–27], provided the first successful relativistic quantum theory of an interaction. The electroweak theory of Glashow, Salam, and Weinberg [28–30] unified the electromagnetic and weak interactions within a single gauge framework. The quark model, proposed by Gell-Mann and Zweig [31, 32], introduced quarks as the constituents of hadrons. Finally, quantum chromodynamics (QCD) [33], together with the discovery of asymptotic freedom [34, 35], established the gauge theory of the strong interaction.

The following chapter introduces some of the most important aspects of the Standard Model, including the fundamental particles, their interactions, and the theoretical framework that describes them. A more detailed introduction to the SM can be found in Refs. [36–38].

1.1. The main principles

The SM is a quantum field theory that describes elementary particles and their interactions in terms of quantum fields. In this framework, particles are interpreted as excitations of the corresponding fields, while interactions arise from the structure of the Lagrangian and are mediated by gauge bosons.

A central principle of the SM is local gauge symmetry. Requiring the Lagrangian to be invariant under local transformations, called gauge tranformations, imposes strong constraints on the allowed interactions and determines the form of the corresponding gauge fields. The SM is based on the gauge group

$$\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y \,.$$

Here, $\mathrm{SU}(3)_C$ describes the strong interaction, while $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ describes the electroweak interaction before symmetry breaking.

In broad terms, the SM Lagrangian can be separated into three closely related sectors:

1. The electroweak sector, based on the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ gauge symmetry. After electroweak symmetry breaking, this sector gives rise to the weak and electromagnetic interactions.

2. The strong sector, based on the $\mathrm{SU}(3)_C$ gauge symmetry, which describes the interaction of quarks and gluons.

3. The Higgs sector, which is responsible for spontaneous electroweak symmetry breaking and generates the masses of the $W^{\pm}$ and Z bosons. Fermion masses arise through Yukawa interactions with the Higgs field.

Schematically, the SM Lagrangian can therefore be written as

$$\mathcal{L}_{\mathrm{SM}} = \mathcal{L}_{\mathrm{EW}} + \mathcal{L}_{\mathrm{Strong}} + \mathcal{L}_{\mathrm{Higgs}}. \tag{1.1}$$

Each of these sectors is discussed in more detail in the following subsections.

The gauge principle also explains why explicit mass terms for gauge bosons are strongly constrained. In an unbroken gauge theory, such mass terms are forbidden because they would violate gauge invariance. The $\mathrm{SU}(3)_C$ symmetry remains unbroken, and the corresponding gluons are massless. By contrast, the electroweak symmetry is spontaneously broken by the Higgs field, generating massive $W^{\pm}$ and Z bosons while leaving the photon massless. This mechanism is known as the Higgs mechanism and is discussed in Sec. 1.6.

Another important conceptual and practical development in QFT is the description of particle interactions using Feynman diagrams. These diagrams provide a systematic calculational tool for evaluating the corresponding quantum amplitudes. Many predictions of the SM, particularly those involving scattering cross sections and decay rates, can be computed perturbatively using Feynman rules derived from the Lagrangian. The resulting predictions have been tested with high precision in a wide range of experiments, making the SM one of the most successful theories in modern physics.

1.2. Fundamental particles

The SM classifies all known elementary particles into two broad categories: fermions and bosons. These categories are distinguished by spin, which is the intrinsic angular momentum of a particle in units of $\hbar$.

Fermions have half-integer spin and obey the Pauli exclusion principle. All fermions in the SM have spin $\frac{1}{2}$ and form the matter content of the theory. They are divided into two types:

1. Quarks, which participate in the strong, weak, and electromagnetic interactions. Quarks are not observed as isolated particles because of confinement, but they form bound states called hadrons, such as protons and neutrons. There are six quark flavors:

    - Up (u) and down (d) quarks are the lightest quarks and form the valence content of ordinary nuclear matter, such as protons (uud) and neutrons (udd).
    - Charm (c) and strange (s) quarks are heavier and are produced in high-energy processes or in unstable hadrons, eventually decaying through weak interactions.
    - Top (t) and bottom (b) quarks are the heaviest quarks. The top quark is the heaviest known elementary particle and decays before it can form hadrons.

    Up-type quarks (u, c, t) have electric charge $+\frac{2}{3}e$, while down-type quarks (d, s, b) have electric charge $-\frac{1}{3}e$, where e is the positive elementary charge.

2. Leptons, which do not participate in the strong interaction but do interact through the electroweak interaction. There are six leptons:

    - The electron (e), muon (μ), and tau (τ) are charged leptons, each carrying electric charge $-e$. The electron is stable, while the muon and tau are unstable and decay weakly into lighter particles. The tau lepton can decay either leptonically or into hadrons and a tau neutrino.
    - The neutrinos (ν_e, ν_μ, ν_τ) are electrically neutral particles that interact only through the weak interaction in the SM. In the minimal formulation of the SM, neutrinos are massless. However, the observation of neutrino oscillations shows that at least two neutrinos have nonzero masses, which is one of the indications that the minimal SM is incomplete.

The fermions are arranged into three generations, where each generation contains two quarks and two leptons. The particles in the second and third generations have the same gauge quantum numbers as their first-generation counterparts but are generally heavier and unstable.

Bosons have integer spin and include the force carriers of the SM. The gauge bosons are spin-1 particles associated with the gauge symmetries of the theory:

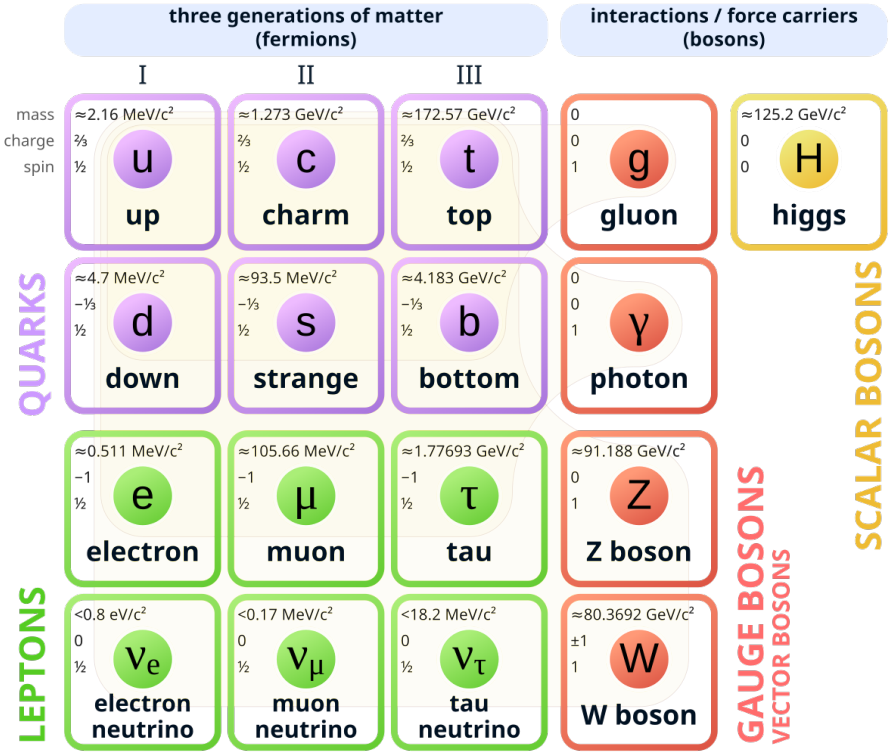


Fig. 1: The fundamental particles in the Standard Model, divided into fermions and bosons. Each of the first three columns represents a generation of fermions, while the second column from the right lists the gauge bosons that mediate the fundamental interactions. The rightmost column contains the Higgs boson [39].

- The photon (γ) mediates the electromagnetic interaction. It is massless and couples to electrically charged particles.
- The gluons (g) mediate the strong interaction between color-charged particles. There are eight gluon fields, corresponding to the eight generators of the $\mathrm{SU}(3)_C$ gauge group. Since gluons themselves carry color charge, they also interact with each other, which is a characteristic feature of non-abelian gauge theories.
- The $W^{\pm}$ and Z bosons mediate the weak interaction. They are massive because of electroweak symmetry breaking. The $W^{\pm}$ bosons mediate charged-current interactions, while the Z boson mediates neutral-current

weak interactions.

In addition to the gauge bosons, the SM contains the Higgs boson (H), a scalar particle with spin 0. It is associated with the Higgs mechanism, through which the $W^{\pm}$ and Z bosons acquire mass. The charged fermion masses are generated through Yukawa interactions with the Higgs field. The full particle content of the SM is summarized in Fig. 1.

1.3. Quantum Electrodynamics

Before discussing the full electroweak theory, it is useful to first consider quantum electrodynamics (QED), which provides the simplest example of how a gauge symmetry gives rise to an interaction. QED is a quantum field theory based on a local $\mathrm{U}(1)_{\mathrm{EM}}$ gauge symmetry and describes the interaction of electrically charged fermions with the photon. In the SM, this electromagnetic gauge symmetry appears as the unbroken subgroup that remains after electroweak symmetry breaking.

We start with the Dirac equation, which describes a free spin-$\frac{1}{2}$ fermion:

$$(i\gamma^{\mu}\partial_{\mu} - m)\psi = 0, \tag{1.2}$$

where ψ is the fermion field, γ^{μ} are the Dirac matrices, $\partial_{\mu} = \frac{\partial}{\partial x^{\mu}}$, and m is the fermion mass. Throughout this section, the Einstein summation convention is used, and natural units with $\hbar = c = 1$ are assumed. Equation (1.2) follows from the free Dirac Lagrangian

$$\mathcal{L}_{\mathrm{Dirac}} = \bar{\psi}(i\gamma^{\mu}\partial_{\mu} - m)\psi, \tag{1.3}$$

where $\bar{\psi} = \psi^{\dagger}\gamma^{0}$ is the Dirac adjoint. This Lagrangian describes a free fermion field and contains no interaction terms.

To introduce electromagnetic interactions, we impose local $\mathrm{U}(1)_{\mathrm{EM}}$ gauge invariance. For a fermion with electric charge $Q_f e$, where Q_f is the charge in units of the positive elementary charge e, the local phase transformation can be written as

$$\psi(x) \to e^{iQ_f\alpha(x)}\psi(x), \tag{1.4}$$

where $\alpha(x)$ is a spacetime-dependent phase. Under this transformation, the ordinary derivative does not transform in the same way as the field:

$$\partial_{\mu}\psi(x) \to e^{iQ_f\alpha(x)}\left[\partial_{\mu}\psi(x) + iQ_f\left(\partial_{\mu}\alpha(x)\right)\psi(x)\right]. \tag{1.5}$$

The additional term proportional to $\partial_{\mu}\alpha(x)$ prevents the free Dirac Lagrangian from being invariant under local phase transformations.

Gauge invariance is restored by replacing the ordinary derivative with the covariant derivative

$$D_\mu \psi = (\partial_\mu - ieQ_f A_\mu)\,\psi, \tag{1.6}$$

where A_μ is the electromagnetic gauge field. For the covariant derivative to transform in the same way as ψ, the gauge field must transform as

$$A_\mu(x) \to A_\mu(x) + \frac{1}{e}\partial_\mu \alpha(x). \tag{1.7}$$

Indeed, substituting Eqs. (1.4) and (1.7) into Eq. (1.6), one obtains

$$\begin{aligned} D_\mu \psi(x) &\to \left(\partial_\mu - ieQ_f \left[A_\mu(x) + \frac{1}{e}\partial_\mu \alpha(x)\right]\right) e^{iQ_f \alpha(x)} \psi(x) \\ &= e^{iQ_f \alpha(x)} \left(\partial_\mu - ieQ_f A_\mu(x)\right) \psi(x) = e^{iQ_f \alpha(x)} D_\mu \psi(x). \end{aligned} \tag{1.8}$$

Thus, imposing local $\mathrm{U}(1)_{\mathrm{EM}}$ gauge invariance naturally introduces the gauge field A_μ, which is identified with the photon field.

Replacing ∂_μ with D_μ in the Dirac Lagrangian gives

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi + eQ_f \bar{\psi}\gamma^\mu A_\mu \psi. \tag{1.9}$$

The first part is the free fermion Lagrangian, while the second term describes the interaction between the charged fermion and the electromagnetic field. For the electron, $Q_f = -1$.

The dynamics of the electromagnetic field are described by the field strength tensor

$$F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{1.10}$$

The corresponding kinetic term is

$$\mathcal{L}_{\text{photon}} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}. \tag{1.11}$$

This term describes the propagation of the free photon field.

Combining the fermion and photon terms, the QED Lagrangian is

$$\mathcal{L}_{\text{QED}} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4} F_{\mu\nu} F^{\mu\nu}. \tag{1.12}$$

This Lagrangian is invariant under local $\mathrm{U}(1)_{\mathrm{EM}}$ gauge transformations and describes the interaction of charged fermions with the photon. A photon mass term of the form $m_A^2 A_\mu A^\mu/2$ is not included, since it would not be invariant under the local gauge transformation in Eq. (1.7). Thus, gauge invariance also explains why the photon is massless in QED.

To describe the weak and electromagnetic interactions in a unified way, the SM uses the gauge symmetry

$$\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y.$$

The $\mathrm{SU}(2)_L$ part acts only on left-handed fermion doublets, while the $\mathrm{U}(1)_Y$ part acts according to the weak hypercharge Y of each field. The physical photon and Z boson appear only after electroweak symmetry breaking as mixtures of the neutral gauge fields of these two groups.

The weak interaction is chiral, meaning that left- and right-handed fermion components transform differently under the electroweak gauge group. The left- and right-handed components of a Dirac spinor are obtained using the projection operators

$$\psi_L = \frac{1}{2}(1 - \gamma^5)\psi, \tag{1.13}$$

and

$$\psi_R = \frac{1}{2}(1 + \gamma^5)\psi, \tag{1.14}$$

where $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$.

In the SM, left-handed fermions are arranged into $\mathrm{SU}(2)_L$ doublets. For the first lepton generation,

$$L_e = \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L,$$

while the right-handed electron is an $\mathrm{SU}(2)_L$ singlet,

$$e_R.$$

Right-handed neutrino fields are not included in the minimal SM. If introduced, they would be singlets under all SM gauge groups and would therefore have no direct gauge interactions.

Similarly, the first-generation quark fields are arranged as

$$Q_L = \begin{pmatrix} u \\ d \end{pmatrix}_L, \qquad u_R, \quad d_R.$$

The same structure is repeated for the second and third generations. This different treatment of left- and right-handed fermions is responsible for parity violation in weak interactions.

The two quantum numbers associated with $\mathrm{SU}(2)_L$ and $\mathrm{U}(1)_Y$ are weak isospin I_a ($a = 1, 2, 3$) and weak hypercharge Y, respectively. The electric charge operator is given by the Gell-Mann–Nishijima relation

$$Q = I_3 + \frac{Y}{2}. \tag{1.15}$$

For example, the left-handed lepton doublet has $Y = -1$. The upper component, the neutrino, has $I_3 = +\frac{1}{2}$ and therefore $Q = 0$, while the lower component, the electron, has $I_3 = -\frac{1}{2}$ and therefore $Q = -1$. The right-handed electron has $I_3 = 0$ and $Y = -2$, again giving $Q = -1$.

Under a local $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ gauge transformation, the left-handed electron doublet transforms as

$$L_e \to \exp\left(i\beta^a(x)\frac{\sigma^a}{2} + i\alpha(x)\frac{Y_L}{2}\right) L_e, \tag{1.16}$$

where $\beta^a(x)$ are the local $\mathrm{SU}(2)_L$ transformation parameters, σ^a are the Pauli matrices, $Y_L = -1$ is the hypercharge of the lepton doublet, and $\alpha(x)$ is the local $\mathrm{U}(1)_Y$ transformation parameter. The Pauli matrices are related to the generators of $\mathrm{SU}(2)_L$ as

$$I_a = \frac{\sigma^a}{2}.$$

The $\mathrm{SU}(2)_L$ transformation can mix the two components of the doublet, whereas the $\mathrm{U}(1)_Y$ transformation changes only the overall phase of the doublet.

The right-handed electron is an $\mathrm{SU}(2)_L$ singlet and transforms only under $\mathrm{U}(1)_Y$:

$$e_R \to \exp\left(i\alpha(x)\frac{Y_{e_R}}{2}\right) e_R, \tag{1.17}$$

with $Y_{e_R} = -2$. The same logic applies to the right-handed quark singlets, with their corresponding hypercharges.

To construct a locally gauge-invariant theory, gauge fields are introduced for both symmetry groups. The $\mathrm{SU}(2)_L$ symmetry introduces three gauge fields W_μ^a ($a = 1, 2, 3$), with coupling constant g. The $\mathrm{U}(1)_Y$ symmetry introduces one gauge field B_μ, with coupling constant g'. The covariant derivative acting on a left-handed doublet is

$$D_\mu L_e = \left(\partial_\mu - ig\frac{\sigma^a}{2}W_\mu^a - ig'\frac{Y_L}{2}B_\mu\right) L_e. \tag{1.18}$$

For a right-handed $\mathrm{SU}(2)_L$ singlet, such as e_R, the covariant derivative contains only the $\mathrm{U}(1)_Y$ interaction:

$$D_\mu e_R = \left(\partial_\mu - ig'\frac{Y_{e_R}}{2}B_\mu\right) e_R. \tag{1.19}$$

The gauge fields themselves must transform in such a way that the covariant derivative transforms like the field on which it acts. For the $\mathrm{SU}(2)_L$ part, defining

$$U(x) = \exp\left(i\beta^a(x)\frac{\sigma^a}{2}\right),$$

the gauge fields transform as

$$W_\mu^a \frac{\sigma^a}{2} \to U\left(W_\mu^a \frac{\sigma^a}{2}\right)U^\dagger + \frac{i}{g}U\partial_\mu U^\dagger. \tag{1.20}$$

With this transformation law, the $\mathrm{SU}(2)_L$ part of the covariant derivative obeys

$$D_\mu L_e \to U D_\mu L_e. \tag{1.21}$$

For the $\mathrm{U}(1)_Y$ part, the gauge field transforms as

$$B_\mu \to B_\mu + \frac{1}{g'}\partial_\mu \alpha(x). \tag{1.22}$$

Consequently, for a field with hypercharge Y,

$$D_\mu \psi \to \exp\left(i\alpha(x)\frac{Y}{2}\right) D_\mu \psi. \tag{1.23}$$

This applies both to left-handed doublets and to right-handed singlets. The transformation is analogous to the $\mathrm{U}(1)_{\mathrm{EM}}$ transformation discussed in the QED subsection, but with weak hypercharge replacing electric charge.

The dynamics of the gauge fields are described by their field strength tensors. For the $\mathrm{SU}(2)_L$ gauge fields,

$$W_{\mu\nu}^a = \partial_\mu W_\nu^a - \partial_\nu W_\mu^a + g\epsilon^{abc} W_\mu^b W_\nu^c, \tag{1.24}$$

where ϵ^{abc} is the antisymmetric Levi-Civita symbol. The last term appears because $\mathrm{SU}(2)_L$ is non-abelian and gives rise to gauge-boson self-interactions.

For the $\mathrm{U}(1)_Y$ gauge field, the field strength tensor is

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu. \tag{1.25}$$

This expression has the same form as the electromagnetic field strength tensor in QED because $\mathrm{U}(1)_Y$ is abelian.

The gauge and fermion-kinetic part of the electroweak Lagrangian for the first lepton generation can then be written as

$$\mathcal{L}_{\mathrm{EW}} = \bar{L}_e i\gamma^\mu D_\mu L_e + \bar{e}_R i\gamma^\mu D_\mu e_R - \frac{1}{4}W_{\mu\nu}^a W^{a\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}. \tag{1.26}$$

The quark terms and the analogous terms for the other lepton generations are omitted here for compactness. At this stage, no explicit mass terms for fermions or gauge bosons have been introduced. Such terms would not be invariant under the full $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ gauge symmetry. The masses of the $W^{\pm}$ and Z bosons, as well as the fermion masses, arise after electroweak symmetry breaking through the Higgs mechanism, which is discussed in Sec. 1.6.

1.5. Strong sector

Quantum chromodynamics (QCD) is the quantum field theory that describes the strong interaction. It is the part of the SM that accounts for the interactions between quarks and gluons and is based on the local gauge symmetry $\mathrm{SU}(3)_C$, where the subscript C denotes color.

Quarks are spin-$\frac{1}{2}$ fermions that appear in six flavors: u, d, s, c, b, and t. Each quark flavor comes in three color states, conventionally labeled as red (r), green (g), and blue (b). For a given quark flavor, the color degrees of freedom can be represented by a color triplet

$$\Psi = \begin{pmatrix} \psi_r \\ \psi_g \\ \psi_b \end{pmatrix}. \tag{1.27}$$

Ignoring the strong interaction for the moment, the free quark Lagrangian has the same form as the Dirac Lagrangian for a massive fermion:

$$\mathcal{L}_{\text{quark}} = \overline{\Psi}\left(i\gamma^{\mu}\partial_{\mu} - m\right)\Psi. \tag{1.28}$$

Here, m denotes the quark mass after electroweak symmetry breaking.

To introduce the strong interaction, we impose local $\mathrm{SU}(3)_C$ gauge invariance. Under a local color rotation, the quark field transforms as

$$\Psi(x) \to U(x)\Psi(x), \qquad U(x) = \exp\left(i\theta^a(x)T^a\right), \tag{1.29}$$

where $\theta^a(x)$ are the local transformation parameters and

$$T^a = \frac{\lambda^a}{2}, \qquad a = 1, \ldots, 8,$$

are the generators of $\mathrm{SU}(3)_C$ in the fundamental representation. The matrices λ^a are the Gell-Mann matrices. The eight generators correspond to the eight gluon fields of QCD.

As in QED and in the electroweak theory, the ordinary derivative does not transform covariantly under local gauge transformations. Gauge invariance is restored by introducing the covariant derivative

$$D_\mu = \partial_\mu - ig_s T^a G^a_\mu, \tag{1.30}$$

where g_s is the strong coupling constant and G^a_μ are the gluon fields.

The gluon fields must transform in such a way that the covariant derivative of the quark field transforms like the quark field itself. In matrix notation, the required transformation is

$$T^a G^a_\mu \to U T^a G^a_\mu U^\dagger + \frac{i}{g_s} U \partial_\mu U^\dagger. \tag{1.31}$$

With this transformation law, one obtains

$$D_\mu \Psi \to U D_\mu \Psi, \tag{1.32}$$

which shows that the quark kinetic term constructed with D_μ is invariant under local $\mathrm{SU}(3)_C$ transformations.

The dynamics of the gluon fields are described by the QCD field strength tensor. In component form, it is

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + g_s f^{abc} G^b_\mu G^c_\nu, \tag{1.33}$$

where f^{abc} are the structure constants of $\mathrm{SU}(3)$, defined by

$$[T^a, T^b] = i f^{abc} T^c. \tag{1.34}$$

The last term in Eq. (1.33) is a direct consequence of the non-abelian nature of $\mathrm{SU}(3)_C$. It leads to gluon self-interactions, unlike in QED where the photon carries no electric charge and does not directly couple to itself.

The QCD Lagrangian for one quark flavor can then be written as

$$\mathcal{L}_{\text{Strong}} = \overline{\Psi} \left(i\gamma^\mu D_\mu - m \right) \Psi - \frac{1}{4} G^a_{\mu\nu} G^{a\mu\nu}. \tag{1.35}$$

For the full SM, this expression is summed over all quark flavors.

The structure of the QCD Lagrangian explains two characteristic features of the strong interaction. At short distances, the strong coupling becomes smaller, allowing perturbative calculations. This behavior is known as asymptotic freedom. At long distances, the strong interaction becomes non-perturbative and leads to confinement, which is why isolated quarks and gluons are not observed as free particles.

In the electroweak theory based on the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ gauge symmetry, explicit mass terms for the gauge bosons and fermions are not allowed before symmetry breaking because they would violate gauge invariance. This is in apparent contradiction with the experimental fact that the $W^{\pm}$ and Z bosons are massive. The Higgs mechanism resolves this problem by introducing spontaneous electroweak symmetry breaking, which generates particle masses while preserving the gauge invariance of the underlying Lagrangian.

To see why explicit gauge-boson masses are problematic, consider a term of the form

$$\Delta\mathcal{L}_{\text{mass}} = \frac{1}{2} m_W^2 W_\mu^a W^{a\mu}. \tag{1.36}$$

Under a local $\mathrm{SU}(2)_L$ transformation, the gauge field transforms as in Eq. (1.20). Because this transformation contains a derivative of the local transformation matrix, the quantity $W_\mu^a W^{a\mu}$ does not remain invariant. Therefore, a mass term written directly for the $\mathrm{SU}(2)_L$ gauge fields would break gauge invariance and unitarity.

A similar issue appears for fermion masses. For example, an electron mass term would have the form

$$\Delta\mathcal{L}_{\text{mass}} = -m_e \left(\bar{e}_L e_R + \bar{e}_R e_L\right). \tag{1.37}$$

However, e_L is part of the left-handed lepton doublet L_e, while e_R is an $\mathrm{SU}(2)_L$ singlet. Since these fields transform differently under $\mathrm{SU}(2)_L$, this term is not invariant under the full electroweak gauge symmetry. A different mechanism is therefore needed to generate masses.

The Higgs mechanism introduces a complex scalar field Φ, called the Higgs field, which transforms as a doublet under $\mathrm{SU}(2)_L$ and has hypercharge $Y_\Phi = 1$:

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}. \tag{1.38}$$

The superscripts denote the electric charges of the components: ϕ^+ has charge $+1$, while ϕ^0 is electrically neutral. Under $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$, the Higgs doublet transforms as

$$\Phi \to \exp\left(i\beta^a(x)\frac{\sigma^a}{2} + i\alpha(x)\frac{Y_\Phi}{2}\right)\Phi. \tag{1.39}$$

The Higgs field is described by the potential

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda \left(\Phi^\dagger \Phi\right)^2, \tag{1.40}$$

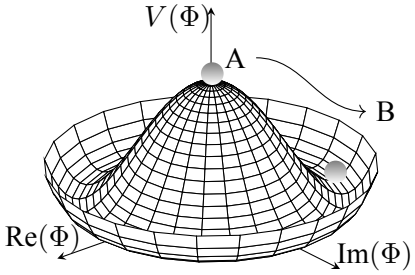


Fig. 2: The Higgs potential $V(\Phi)$ as a function of the Higgs field Φ for $\mu^2 < 0$ and $\lambda > 0$. The circular groove at the bottom has a radius of $v = \sqrt{-\mu^2/\lambda}$ and corresponds to the VEV.

where $\lambda > 0$ ensures that the potential is bounded from below. For $\mu^2 < 0$, the minimum of the potential occurs at a nonzero value of $\Phi^\dagger\Phi$, as shown in Fig. 2. Using the gauge freedom, the vacuum can be chosen in the form

$$\langle\Phi\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v\end{pmatrix}, \tag{1.41}$$

where v is the vacuum expectation value (VEV) of the Higgs field. Substituting this into the potential gives

$$V(v) = \frac{\mu^2 v^2}{2} + \frac{\lambda v^4}{4}. \tag{1.42}$$

The condition for the nontrivial minimum is

$$\frac{\mathrm{d}V}{\mathrm{d}v} = \mu^2 v + \lambda v^3 = 0, \tag{1.43}$$

which gives

$$v^2 = -\frac{\mu^2}{\lambda}, \qquad v = \sqrt{-\frac{\mu^2}{\lambda}}. \tag{1.44}$$

Thus, electroweak symmetry breaking requires $\mu^2 < 0$.

Although the Lagrangian remains invariant under $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$, the chosen vacuum state does not. The nonzero VEV breaks the electroweak symmetry down to the electromagnetic subgroup,

$$\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y \longrightarrow \mathrm{U}(1)_{\mathrm{EM}}.$$

This can be seen by acting on the VEV with the charge operator

$$Q = I_3 + \frac{Y}{2} = \frac{\sigma^3}{2} + \frac{Y}{2}. \tag{1.45}$$

For the Higgs doublet, $Y_\Phi = 1$, and therefore

$$Q\langle\Phi\rangle = \left(\frac{\sigma^3}{2} + \frac{1}{2}\right)\frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v\end{pmatrix} = 0. \tag{1.46}$$

The vacuum is therefore invariant under transformations generated by Q, which defines the unbroken electromagnetic $\mathrm{U}(1)_{\mathrm{EM}}$ symmetry. The other three generators are broken, and their corresponding degrees of freedom become the longitudinal components of the massive $W^\pm$ and Z bosons.

To describe the physical Higgs boson, the field is expanded around the vacuum. In the unitary gauge, this can be written as

$$\Phi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v + h(x)\end{pmatrix}, \tag{1.47}$$

where $h(x)$ is the physical Higgs field. The three remaining scalar degrees of freedom have been absorbed into the longitudinal polarizations of the $W^\pm$ and Z bosons.

The masses of the electroweak gauge bosons arise from the Higgs kinetic term

$$\mathcal{L}_{\text{Higgs, kin}} = \left(D_\mu\Phi\right)^\dagger\left(D^\mu\Phi\right), \tag{1.48}$$

with

$$D_\mu = \partial_\mu - ig\frac{\sigma^a}{2}W_\mu^a - ig'\frac{Y_\Phi}{2}B_\mu. \tag{1.49}$$

Substituting Eq. (1.47) and using $Y_\Phi = 1$, one obtains

$$\mathcal{L}_{\text{Higgs, kin}} = \frac{1}{2}\left(\partial_\mu h\right)\left(\partial^\mu h\right) + \frac{(v+h)^2}{8}\Big[g^2\left(W_\mu^1 W^{1\mu} + W_\mu^2 W^{2\mu}\right) + \left(gW_\mu^3 - g'B_\mu\right)\left(gW^{3\mu} - g'B^\mu\right)\Big]. \tag{1.50}$$

The terms proportional to v^2 are the gauge-boson mass terms:

$$\mathcal{L}_{\text{mass}} = \frac{v^2}{8}\Big[g^2\left(W_\mu^1 W^{1\mu} + W_\mu^2 W^{2\mu}\right) + \left(gW_\mu^3 - g'B_\mu\right)\left(gW^{3\mu} - g'B^\mu\right)\Big]. \tag{1.51}$$

The charged weak boson fields are defined as

$$W_\mu^\pm = \frac{1}{\sqrt{2}}\left(W_\mu^1 \mp iW_\mu^2\right). \tag{1.52}$$

The neutral fields are diagonalized by defining

$$Z_\mu = \frac{gW_\mu^3 - g'B_\mu}{\sqrt{g^2 + g'^2}}, \qquad A_\mu = \frac{g'W_\mu^3 + gB_\mu}{\sqrt{g^2 + g'^2}}. \tag{1.53}$$

In terms of these physical fields, the mass terms become

$$\mathcal{L}_{\text{mass}} = \frac{g^2 v^2}{4} W_\mu^+ W^{-\mu} + \frac{v^2}{8} \left(g^2 + g'^2\right) Z_\mu Z^\mu. \tag{1.54}$$

Comparing with the standard form of massive vector-boson mass terms gives

$$m_W = \frac{gv}{2}, \qquad m_Z = \frac{v}{2}\sqrt{g^2 + g'^2}. \tag{1.55}$$

No mass term is generated for A_μ, so the photon remains massless.

It is useful to define the weak mixing angle, or Weinberg angle, by

$$\tan\theta_{\text{W}} = \frac{g'}{g}. \tag{1.56}$$

Then the physical neutral fields can be written as

$$\begin{cases} Z_\mu = \cos\theta_{\text{W}} W_\mu^3 - \sin\theta_{\text{W}} B_\mu, \\ A_\mu = \sin\theta_{\text{W}} W_\mu^3 + \cos\theta_{\text{W}} B_\mu. \end{cases} \tag{1.57}$$

The weak-boson masses satisfy

$$m_Z = \frac{m_W}{\cos\theta_{\text{W}}}, \tag{1.58}$$

and the elementary electric charge is related to the electroweak gauge couplings by

$$e = g\sin\theta_{\text{W}} = g'\cos\theta_{\text{W}}. \tag{1.59}$$

Fermion masses are generated through Yukawa interactions between fermion fields and the Higgs doublet. For the electron, the Yukawa interaction is

$$\mathcal{L}_{\text{Yukawa}} = -y_e \left(\bar{L}_e \Phi e_R + \bar{e}_R \Phi^\dagger L_e\right), \tag{1.60}$$

where y_e is the electron Yukawa coupling. After substituting Eq. (1.47), this becomes

$$\mathcal{L}_{\text{Yukawa}} = -\frac{y_e}{\sqrt{2}}(v + h)\left(\bar{e}_L e_R + \bar{e}_R e_L\right). \tag{1.61}$$

The term proportional to v is the electron mass term,

$$\mathcal{L}_{\text{mass}} = -m_e \left(\bar{e}_L e_R + \bar{e}_R e_L\right), \tag{1.62}$$

with

$$m_e = \frac{y_e v}{\sqrt{2}}. \tag{1.63}$$

The term proportional to h describes the coupling of the Higgs boson to the electron. Thus, the Higgs coupling to a fermion is proportional to the fermion mass.

The same mechanism generates masses for all charged fermions. For quarks, the Yukawa Lagrangian is

$$\mathcal{L}_{\text{Yukawa}} = -\overline{Q}_L^{\,i} \left(Y_d^{ij} \Phi d_R^j + Y_u^{ij} \widetilde{\Phi} u_R^j \right) + \text{h.c.}, \tag{1.64}$$

where

$$\widetilde{\Phi} = i\sigma^2 \Phi^*, \tag{1.65}$$

Q_L^i are the left-handed quark doublets, u_R^j and d_R^j are the right-handed up-type and down-type quark singlets, and Y_u^{ij} and Y_d^{ij} are Yukawa matrices in generation space. After electroweak symmetry breaking, these terms generate quark mass matrices:

$$\mathcal{L}_{\text{mass}} = -\bar{u}_L^i M_u^{ij} u_R^j - \bar{d}_L^i M_d^{ij} d_R^j + \text{h.c.}, \tag{1.66}$$

where

$$M_u^{ij} = \frac{v}{\sqrt{2}} Y_u^{ij}, \qquad M_d^{ij} = \frac{v}{\sqrt{2}} Y_d^{ij}. \tag{1.67}$$

In general, the Yukawa matrices are not diagonal in generation space. The physical quark mass eigenstates are obtained by unitary rotations of the quark fields:

$$\begin{aligned} u_L'^i &= (U_u)_L^{ij} u_L^j, & u_R'^i &= (U_u)_R^{ij} u_R^j, \\ d_L'^i &= (U_d)_L^{ij} d_L^j, & d_R'^i &= (U_d)_R^{ij} d_R^j. \end{aligned} \tag{1.68}$$

Here, the primed fields denote weak-interaction eigenstates, while the unprimed fields denote mass eigenstates. The unitary matrices are chosen such that

$$(U_u)_L^\dagger M_u (U_u)_R = M_u^{\text{diag}}, \qquad (U_d)_L^\dagger M_d (U_d)_R = M_d^{\text{diag}}. \tag{1.69}$$

The charged-current weak interaction couples the up-type and down-type components of the left-handed quark doublets. Because the rotations that diagonalize the up-type and down-type quark mass matrices are generally different, the charged-current interaction contains the Cabibbo–Kobayashi–Maskawa (CKM) matrix,

$$V_{\text{CKM}} = (U_u)_L^\dagger (U_d)_L. \tag{1.70}$$

This matrix encodes quark flavor mixing in charged-current weak interactions.

In the minimal SM, neutrinos are massless and no neutrino Yukawa term is included. The charged-lepton Yukawa matrix can be diagonalized in the same

way as the quark Yukawa matrices, but no physical lepton mixing appears if neutrinos are massless. The observed neutrino oscillations show that neutrinos do have nonzero masses, requiring an extension of the minimal SM. In such extensions, lepton mixing is described by the Pontecorvo–Maki–Nakagawa–Sakata (PMNS) matrix.

The Higgs boson mass arises from the Higgs potential. Substituting $\Phi = (0, (v+h)/\sqrt{2})^T$ into Eq. (1.40) and using the minimum condition $\mu^2 = -\lambda v^2$ gives

$$V(h) = -\frac{\lambda v^4}{4} + \lambda v^2 h^2 + \lambda v h^3 + \frac{\lambda}{4} h^4. \tag{1.71}$$

The constant term has no effect on the particle dynamics, while the cubic and quartic terms describe Higgs self-interactions. Since the Lagrangian contains $-V(h)$, the quadratic term gives

$$\mathcal{L}_{\text{Higgs mass}} = -\lambda v^2 h^2 \equiv -\frac{1}{2} m_H^2 h^2, \tag{1.72}$$

so that

$$m_H^2 = 2\lambda v^2. \tag{1.73}$$

The Higgs mechanism therefore explains how the $W^\pm$ and Z bosons acquire mass while the photon remains massless, and how charged fermion masses arise through Yukawa interactions. It also predicts the existence of a physical scalar boson, the Higgs boson, which was discovered at the Large Hadron Collider in 2012 [7, 8].

Combining the gauge, fermion, Yukawa, and Higgs terms, the SM Lagrangian can be written compactly as

$$\begin{aligned}
\mathcal{L}_{\text{SM}} = & -\frac{1}{4} G^a_{\mu\nu} G^{a\mu\nu} - \frac{1}{4} W^a_{\mu\nu} W^{a\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \\
& + \sum_i \Big[i \overline{Q}^i_L \gamma^\mu D_\mu Q^i_L + i \bar{u}^i_R \gamma^\mu D_\mu u^i_R + i \bar{d}^i_R \gamma^\mu D_\mu d^i_R \\
& \qquad + i \overline{L}^i_L \gamma^\mu D_\mu L^i_L + i \bar{e}^i_R \gamma^\mu D_\mu e^i_R \Big] \\
& - \Big[\overline{Q}^i_L Y_u^{ij} \widetilde{\Phi} u^j_R + \overline{Q}^i_L Y_d^{ij} \Phi d^j_R + \overline{L}^i_L Y_e^{ij} \Phi e^j_R + \text{h.c.} \Big] \\
& + (D_\mu \Phi)^\dagger (D^\mu \Phi) - \mu^2 \Phi^\dagger \Phi - \lambda (\Phi^\dagger \Phi)^2.
\end{aligned} \tag{1.74}$$

Here, $G^a_{\mu\nu}$, $W^a_{\mu\nu}$, and $B_{\mu\nu}$ are the field strength tensors for the $\mathrm{SU}(3)_C$, $\mathrm{SU}(2)_L$, and $\mathrm{U}(1)_Y$ gauge fields, respectively. The index $i = 1, 2, 3$ labels the fermion generations, and repeated generation, gauge, and spinor indices are summed or contracted. The fields Q^i_L and L^i_L are the left-handed quark and lepton doublets, while u^i_R, d^i_R, and e^i_R are the right-handed fermion singlets. The matrices Y_u^{ij},

Y_d^{ij}, and Y_e^{ij} are the Yukawa coupling matrices. The Higgs doublet is denoted by Φ, and $\widetilde{\Phi} = i\sigma^2\Phi^*$. Natural units with $\hbar = c = 1$ are used throughout.

1.7. Beyond the Standard Model

Although the Standard Model (SM) has achieved remarkable success in describing particle interactions and predicting experimental results, it is widely regarded as an incomplete theory. One of its most fundamental limitations is that it does not incorporate gravity. The SM is a quantum field theory describing the electromagnetic, weak, and strong interactions, while gravity is described separately by general relativity. A consistent quantum theory that includes gravity remains one of the major open problems in fundamental physics.

Another important limitation is the absence of a particle candidate for dark matter. Astronomical and cosmological observations, such as galaxy rotation curves, gravitational lensing, and measurements of the cosmic microwave background, strongly indicate the existence of matter that interacts gravitationally but does not emit or absorb light. No particle in the SM has the required properties to account for this dark matter component. Similarly, the SM does not explain dark energy, which is used to describe the observed accelerated expansion of the universe.

The minimal SM also does not explain neutrino masses. In its original formulation, neutrinos are massless, but neutrino oscillation experiments have shown that at least two neutrinos have nonzero masses. This observation requires physics beyond the minimal SM, for example through the introduction of right-handed neutrinos or other mechanisms that generate small neutrino masses.

Another open question is the observed matter–antimatter asymmetry of the universe. The SM contains charge-conjugation and parity symmetry (CP) violation through the CKM matrix, but the amount of CP violation present in the SM appears insufficient to explain the observed dominance of matter over antimatter. This suggests that additional sources of CP violation or new mechanisms may be required.

Several theoretical extensions of the SM have been proposed to address these shortcomings. Supersymmetry, grand unified theories (GUTs), extra-dimensional models, and string-theory-inspired scenarios are examples of frameworks that introduce new particles, symmetries, or interactions. Depending on the model, such extensions may provide dark matter candidates, explain neutrino masses, address the hierarchy problem, or move toward a more unified description of the fundamental interactions.

A more model-independent approach is provided by effective field theories.

In particular, the Standard Model Effective Field Theory (SMEFT) extends the SM Lagrangian by adding higher-dimensional operators constructed from SM fields and respecting the SM gauge symmetries. These operators parametrize the possible low-energy effects of heavy, yet-undiscovered particles without assuming a specific underlying theory. SMEFT therefore provides a systematic framework for interpreting precision measurements as constraints on possible beyond-the-Standard-Model (BSM) physics.

At present, no direct experimental evidence has established a specific extension of the SM. Nevertheless, searches for BSM physics continue through several complementary approaches: direct searches for new particles at colliders, precision measurements of SM processes, studies of neutrino properties, direct and indirect searches for dark matter, and cosmological observations. Future experimental results will be essential for determining which, if any, of the proposed extensions provides the correct path toward a more complete description of fundamental physics.

2. SM Calculations

In this section, we discuss how the SM Lagrangian is used to compute physical quantities relevant to high-energy physics. The central goal of such calculations is to obtain measurable observables, such as cross sections and decay rates, from the interaction terms of the theory. These predictions can then be compared with experimental measurements, providing tests of the SM and sensitivity to possible new physics.

The connection between the Lagrangian and measurable quantities is made through scattering amplitudes. In perturbative quantum field theory, these amplitudes are calculated using Feynman rules, which are derived from the terms of the Lagrangian and represented graphically by Feynman diagrams. The diagrams are not only a convenient visualization of particle interactions, but also a compact bookkeeping tool for constructing the mathematical expression for a given process.

We first introduce the general relation between matrix elements, cross sections, and decay rates. Then, we discuss Feynman rules and illustrate their use with a simple Drell–Yan-like process. After that, we briefly introduce renormalization, which is required to obtain finite and physically meaningful predictions beyond leading order. Finally, we discuss parton distribution functions (PDFs), which are essential for connecting parton-level calculations to hadron-collider observables such as those measured in Drell–Yan processes.

In high-energy physics, cross sections and decay rates are the basic quantities used to describe scattering and decay processes. A cross section quantifies the probability for a scattering process to occur under specified initial-state conditions, while a decay rate quantifies the probability per unit time for an unstable particle to decay into a particular final state. Both quantities are directly related to the transition amplitude between an initial and a final quantum state.

To understand how these quantities are computed, it is useful to distinguish between quantum states and quantum fields. The initial and final particles in a process are described by states in a Fock space, characterized by their momenta, spins, and other quantum numbers. The fields appearing in the Lagrangian are operator-valued objects that create, annihilate, and act on these particle states. The interaction terms in the Lagrangian therefore determine the amplitudes for transitions between different initial and final states.

This relation is formalized using the scattering matrix, or S-matrix. The S-matrix connects asymptotic initial states in the distant past to asymptotic final states in the distant future. It is conventionally written as

$$S = \mathbb{1} + iT, \tag{2.1}$$

where $\mathbb{1}$ is the identity operator and T is the transition matrix. In a free theory, $S = \mathbb{1}$, meaning that no transitions between different states occur.

The transition amplitude between an initial state $|i\rangle$ and a final state $|f\rangle$ is given by

$$S_{fi} \equiv \langle f|S|i\rangle = \langle f|i\rangle + i\,\langle f|T|i\rangle\,. \tag{2.2}$$

For a nontrivial scattering or decay process, the initial and final states are different (orthonormal), so the identity contribution $\langle f|i\rangle$ vanishes. The invariant matrix element $\mathcal{M}_{fi}$ is then defined by

$$\langle f|iT|i\rangle = i(2\pi)^4\delta^{(4)}(p_f - p_i)\mathcal{M}_{fi}, \tag{2.3}$$

where p_i and p_f are the total four-momenta of the initial and final states. The four-dimensional delta function enforces energy and momentum conservation. All the dynamical information about the process is contained in $\mathcal{M}_{fi}$.

In perturbation theory, the matrix element is calculated by expanding the S-matrix in powers of the relevant coupling constants. The resulting terms can be represented by Feynman diagrams and evaluated using Feynman rules. At leading order, only the diagrams with the lowest relevant power of the coupling constants are included. Higher-order corrections are obtained by including diagrams with additional vertices, loops, or real radiation.

For a general $2 \to n$ scattering process, the differential cross section can be written as

$$\mathrm{d}\sigma = \frac{1}{(2E_1)(2E_2)\,|\vec{v}_1 - \vec{v}_2|}\,\overline{|\mathcal{M}_{fi}|^2}\,\mathrm{d}\Pi_n, \tag{2.4}$$

where E_1 and E_2 are the energies of the two incoming particles, and $\vec{v}_1$ and $\vec{v}_2$ are their velocities in the center-of-mass frame. The denominator is the incoming flux factor. The bar over $|\mathcal{M}_{fi}|^2$ indicates that the squared matrix element is averaged over the initial spin and color states and summed over the final spin and color states, as appropriate.

The factor $\mathrm{d}\Pi_n$ is the n-body Lorentz-invariant phase-space element,

$$\mathrm{d}\Pi_n = (2\pi)^4 \delta^{(4)}\left(p_i - \sum_{j=1}^{n} p_j\right) \prod_{j=1}^{n} \frac{\mathrm{d}^3 p_j}{(2\pi)^3} \frac{1}{2E_j}, \tag{2.5}$$

where p_j and E_j are the four-momentum and energy of the j-th final-state particle. If identical particles are present in the final state, an additional symmetry factor must be included to avoid overcounting physically indistinguishable configurations.

The total cross section is obtained by integrating over the allowed final-state phase space:

$$\sigma = \frac{1}{(2E_1)(2E_2)\,|\vec{v}_1 - \vec{v}_2|} \int \overline{|\mathcal{M}_{fi}|^2}\,\mathrm{d}\Pi_n. \tag{2.6}$$

In practice, one often integrates only over part of the phase space, leaving the result differential in one or more observables. Such differential cross sections contain more information about the process kinematics than the fully integrated cross section and are therefore particularly important for precision measurements.

Decay rates are calculated in a closely related way. For a particle of mass M decaying into n final-state particles, the differential decay rate in the rest frame of the decaying particle is

$$\mathrm{d}\Gamma = \frac{1}{2M}\,\overline{|\mathcal{M}_{fi}|^2}\,\mathrm{d}\Pi_n. \tag{2.7}$$

The total decay rate is therefore

$$\Gamma = \frac{1}{2M} \int \overline{|\mathcal{M}_{fi}|^2}\,\mathrm{d}\Pi_n. \tag{2.8}$$

Here, the phase-space element has the same form as in Eq. (2.5), but the initial four-momentum corresponds to the decaying particle:

$$p_i = (M, 0, 0, 0) \tag{2.9}$$

in its rest frame.

In natural units, the lifetime τ of an unstable particle is related to its total decay rate by

$$\tau = \frac{1}{\Gamma}. \tag{2.10}$$

If several decay channels are possible, the total decay rate is the sum of the partial decay rates,

$$\Gamma_{\text{tot}} = \sum_i \Gamma_i, \tag{2.11}$$

and the branching fraction for a channel i is

$$\mathcal{B}_i = \frac{\Gamma_i}{\Gamma_{\text{tot}}}. \tag{2.12}$$

Thus, cross-section and decay-rate calculations share the same basic ingredients: the invariant matrix element and the Lorentz-invariant phase space. The main difference is the initial state: scattering processes involve an incoming flux of particles, while decay processes involve a single unstable particle. The calculation of the matrix element $\mathcal{M}_{fi}$ using Feynman rules is discussed in the next subsection.

2.2. Feynman Rules

Feynman rules provide a systematic way to calculate the invariant matrix element $\mathcal{M}_{fi}$ for a given process. They form the practical link between the Lagrangian formulation of the SM and measurable quantities such as cross sections and decay rates. The rules can be derived from the Lagrangian using, for example, canonical quantization or the path-integral formalism. The derivation involves expanding the S-matrix in powers of the interaction Lagrangian and using tools such as Wick's theorem and the LSZ reduction formula. The resulting expressions can be represented graphically by Feynman diagrams.

For the purposes of this thesis, it is not necessary to derive or list all SM Feynman rules. A comprehensive summary can be found in Ref. [40]. Here, we only list the rules needed for the leading-order neutral-current Drell–Yan process, to be introduced more thoroughly in Sec. 3. The external fermion lines denote the usual Dirac spinors $u(p)$, $\bar{u}(p)$, $v(p)$, and $\bar{v}(p)$, depending on whether the external particle is an incoming or outgoing fermion or antifermion.

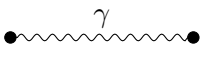
The fermion propagator is



$$= \quad \frac{i(\not{p} + m)}{p^2 - m^2 + i\epsilon}, \tag{2.13}$$

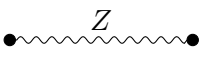
where p is the fermion momentum, m is its mass, and $\not{p} \equiv \gamma^\mu p_\mu$.

In the Feynman gauge, the photon and Z-boson propagators are



$$= \quad \frac{-ig_{\mu\nu}}{p^2 + i\epsilon}, \tag{2.14}$$

and



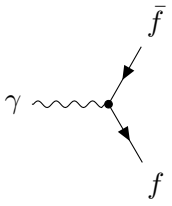
$$= \quad \frac{-ig_{\mu\nu}}{p^2 - m_Z^2 + i\epsilon}, \tag{2.15}$$

where $g_{\mu\nu}$ is the Minkowski metric tensor. Close to the Z-boson pole, the replacement

$$p^2 - m_Z^2 + i\epsilon \to p^2 - m_Z^2 + i m_Z \Gamma_Z$$

is commonly used to account for the finite Z-boson width.

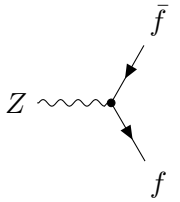
The fermion–fermion–photon vertex is



$$= \quad -ieQ_f\gamma^\mu, \tag{2.16}$$

where Q_f is the electric charge of the fermion in units of the positive elementary charge e.

The fermion–fermion–Z vertex is



$$= \quad -i\frac{g}{2\cos\theta_{\mathrm{W}}}\gamma^\mu\left(v_f - a_f\gamma^5\right), \tag{2.17}$$

where

$$v_f = I_3^f - 2Q_f \sin^2\theta_{\mathrm{W}}, \qquad a_f = I_3^f \tag{2.18}$$

are the vector and axial-vector couplings, and I_3^f is the weak isospin of the fermion. For example,

$$v_u = \frac{1}{2} - \frac{4}{3}\sin^2\theta_{\mathrm{W}}, \quad a_u = \frac{1}{2}, \qquad v_\mu = -\frac{1}{2} + 2\sin^2\theta_{\mathrm{W}}, \quad a_\mu = -\frac{1}{2}.$$

Here, θ_{W} is the Weinberg angle, defined in eq. (1.56).

To demonstrate Feynman rules in action, let's consider the partonic process

$$u(p_1)\bar{u}(p_2) \to \mu^-(p_3)\mu^+(p_4), \tag{2.19}$$

which contributes to the neutral-current Drell–Yan process. At leading order, two diagrams dominate: one mediated by a virtual photon and one mediated by a Z boson, as shown in Fig. 3.

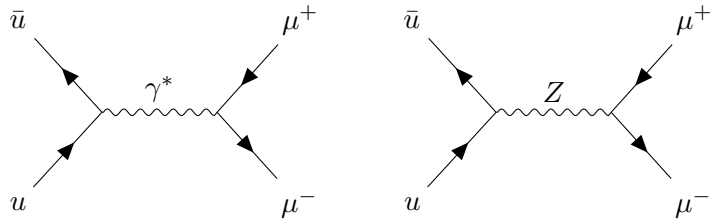


Fig. 3: Tree-level Feynman diagrams contributing to the $u\bar{u} \to \mu^+\mu^-$ process.

Using the rules above, the photon-mediated contribution can be written as

$$i\mathcal{M}_\gamma = [\bar{v}(p_2)(-ieQ_u\gamma^\mu)u(p_1)]\,\frac{-ig_{\mu\nu}}{\hat{s}}\,[\bar{u}(p_3)(-ieQ_\mu\gamma^\nu)v(p_4)]\,, \quad (2.20)$$

where we have neglected the $i\epsilon$ term because the virtual photon is far from its pole, and

$$\hat{s} = (p_1 + p_2)^2$$

is the partonic center-of-mass energy squared. The Z-mediated contribution is

$$\begin{aligned} i\mathcal{M}_Z = &\left[\bar{v}(p_2)\left(-i\frac{g}{2\cos\theta_\mathrm{W}}\gamma^\mu(v_u - a_u\gamma^5)\right)u(p_1)\right]\frac{-ig_{\mu\nu}}{\hat{s} - m_Z^2 + im_Z\Gamma_Z} \\ &\times\left[\bar{u}(p_3)\left(-i\frac{g}{2\cos\theta_\mathrm{W}}\gamma^\nu(v_\mu - a_\mu\gamma^5)\right)v(p_4)\right]. \end{aligned} \quad (2.21)$$

The full leading-order matrix element is therefore

$$\mathcal{M} = \mathcal{M}_\gamma + \mathcal{M}_Z. \quad (2.22)$$

After summing over final-state spins and averaging over initial-state spins and colors, the massless result can be written in a compact form as

$$\overline{|\mathcal{M}|^2} = \frac{e^4}{N_c}\left[(1+\cos^2\theta)F_0^q(\hat{s}) + 2\cos\theta\,F_1^q(\hat{s})\right], \quad (2.23)$$

where $N_c = 3$ is the number of quark colors and θ is the angle between the incoming quark and the outgoing negatively charged lepton in the partonic center-of-mass frame. The two functions F_0^q and F_1^q contain the photon contribution, the Z contribution, and their interference:

$$F_0^q(\hat{s}) = Q_q^2Q_\ell^2 + 2Q_qQ_\ell v_q v_\ell\,\mathrm{Re}\,\chi_Z(\hat{s}) + (v_q^2+a_q^2)(v_\ell^2+a_\ell^2)|\chi_Z(\hat{s})|^2, \quad (2.24)$$

$$F_1^q(\hat{s}) = 2Q_q Q_\ell a_q a_\ell \,\mathrm{Re}\,\chi_Z(\hat{s}) + 4v_q a_q v_\ell a_\ell |\chi_Z(\hat{s})|^2. \tag{2.25}$$

Here,

$$\chi_Z(\hat{s}) = \frac{\hat{s}}{4\sin^2\theta_\mathrm{W}\cos^2\theta_\mathrm{W}\left(\hat{s} - m_Z^2 + im_Z\Gamma_Z\right)}. \tag{2.26}$$

If the finite Z width is neglected, one simply sets $\Gamma_Z = 0$. For the specific process $u\bar{u} \to \mu^+\mu^-$, the charges are

$$Q_u = \frac{2}{3}, \qquad Q_\mu = -1,$$

and the corresponding vector and axial-vector couplings are those given below Eq. (2.18).

Using the standard two-body phase-space expression, the differential partonic cross section in the massless approximation is

$$\frac{\mathrm{d}\hat{\sigma}}{\mathrm{d}\cos\theta} = \frac{1}{32\pi\hat{s}}\overline{|\mathcal{M}|^2}. \tag{2.27}$$

Substituting Eq. (2.23) gives

$$\frac{\mathrm{d}\hat{\sigma}}{\mathrm{d}\cos\theta} = \frac{\pi\alpha^2}{2N_c\hat{s}}\left[(1+\cos^2\theta)F_0^q(\hat{s}) + 2\cos\theta\, F_1^q(\hat{s})\right], \tag{2.28}$$

where $\alpha = e^2/(4\pi)$. The term proportional to F_0^q is symmetric in $\cos\theta$, while the term proportional to F_1^q is antisymmetric. The latter integrates to zero in the total cross section but gives rise to the forward–backward asymmetry.

Integrating over $\cos\theta$ gives the leading-order partonic cross section

$$\hat{\sigma} = \frac{4\pi\alpha^2}{3N_c\hat{s}}F_0^q(\hat{s}). \tag{2.29}$$

For pure photon exchange, this reduces to the familiar result

$$\hat{\sigma}_\gamma = \frac{4\pi\alpha^2}{3N_c\hat{s}}Q_q^2Q_\ell^2 = \frac{4\pi\alpha^2}{9\hat{s}}Q_q^2 \tag{2.30}$$

for charged leptons, where $Q_\ell^2 = 1$ and $N_c = 3$.

This example illustrates the general procedure used in perturbative calculations. First, all diagrams contributing to a given initial and final state are drawn to the desired order in the coupling constants. Then, the corresponding amplitudes are written using the Feynman rules, summed coherently, squared, and integrated over the relevant phase space. For precision predictions, higher-order diagrams must also be included. These introduce loop integrals and real-emission contributions, whose divergences are treated through renormalization and cancellation between real and virtual corrections. The main idea of renormalization is introduced in the next subsection.

2.3. Renormalization

Renormalization is the procedure that makes quantum field theories predictive beyond leading order in perturbation theory. Loop corrections to propagators and vertices often contain divergent integrals, especially in the ultraviolet (UV) region where the loop momentum becomes large. The key idea of renormalization is to absorb these divergences into redefinitions of the fields and parameters of the theory, such that physical observables remain finite.

For simplicity, consider QED. The bare QED Lagrangian can be written as

$$\mathcal{L}_0 = \bar{\psi}_0 \left(i\partial\!\!\!/ - m_0 \right) \psi_0 - \frac{1}{4} F^0_{\mu\nu} F^{0\mu\nu} - e_0 \bar{\psi}_0 \gamma^\mu \psi_0 A^0_\mu, \tag{2.31}$$

where ψ_0, A^0_μ, m_0, and e_0 are bare fields and bare parameters. These quantities are not directly observable. They are related to renormalized fields and parameters by introducing renormalization constants:

$$\psi_0 = \sqrt{Z_2}\psi, \qquad A^0_\mu = \sqrt{Z_3} A_\mu, \qquad m_0 = Z_m m, \qquad e_0 = Z_e e. \tag{2.32}$$

Here, ψ and A_μ are renormalized fields, while m and e are renormalized parameters. The constants Z_2, Z_3, Z_m, and Z_e are fixed order by order in perturbation theory according to a chosen renormalization scheme. Writing

$$Z_i = 1 + \delta_i, \tag{2.33}$$

the quantities δ_i act as counterterms.

Substituting Eq. (2.32) into Eq. (2.31), the bare Lagrangian can be separated into a renormalized part and a counterterm part:

$$\mathcal{L}_0 = \mathcal{L}_\mathrm{R} + \mathcal{L}_\mathrm{C}. \tag{2.34}$$

The renormalized Lagrangian has the same form as the original QED Lagrangian, but is written in terms of renormalized fields and parameters:

$$\mathcal{L}_\mathrm{R} = \bar{\psi} \left(i\partial\!\!\!/ - m \right) \psi - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} - e \bar{\psi} \gamma^\mu \psi A_\mu. \tag{2.35}$$

The remaining terms are collected in the counterterm Lagrangian,

$$\mathcal{L}_\mathrm{C} = i\delta_2 \bar{\psi} \partial\!\!\!/ \psi - (\delta_m + \delta_2) m \bar{\psi}\psi - \left(\delta_e + \delta_2 + \frac{\delta_3}{2} \right) e \bar{\psi} \gamma^\mu \psi A_\mu - \frac{\delta_3}{4} F_{\mu\nu} F^{\mu\nu}. \tag{2.36}$$

These counterterms generate additional Feynman rules. Their role is to cancel the divergent parts of loop diagrams and to impose the chosen definitions of the physical or renormalized parameters.

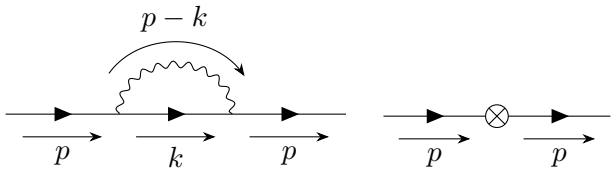


Fig. 4: One-loop electron self-energy diagram (left) and the corresponding counterterm diagram (right).

The procedure is conceptually the same in the full SM, but with more fields and parameters. The SM requires renormalization constants for the gauge fields, fermion fields, Higgs field, gauge couplings, Yukawa couplings, Higgs-potential parameters, and flavor-mixing parameters.

As an illustration, consider the one-loop correction to the electron propagator in QED, shown on the left side of Fig. 4. In the Feynman gauge, the corresponding loop contribution to the electron self-energy can be written schematically as

$$-i\Sigma_2(p) = (-ie)^2 \int \frac{\mathrm{d}^4k}{(2\pi)^4}\gamma^\mu \frac{i(\not{k}+m)}{k^2-m^2+i\epsilon}\gamma^\nu \frac{-ig_{\mu\nu}}{(p-k)^2+i\epsilon}. \tag{2.37}$$

Here, p is the external electron momentum and k is the loop momentum. For large k, this integral is UV divergent and therefore requires renormalization.

The counterterm Lagrangian in Eq. (2.36) provides the counterterm diagram shown on the right side of Fig. 4. For the electron two-point function, the relevant counterterm contribution is

$$i\delta_2\not{p} - i(\delta_m+\delta_2)m. \tag{2.38}$$

The renormalized self-energy is obtained by combining the loop contribution with the counterterm contribution. The one-loop corrected electron propagator becomes:

$$S_F(p) = \frac{i}{\not{p}-m+\Sigma(p)}, \tag{2.39}$$

where $\Sigma(p) = \Sigma_2(p) + \delta_2\not{p} - (\delta_m+\delta_2)m$ is the renormalized self-energy. The constants δ_2 and δ_m are then chosen so that the final propagator satisfies the desired renormalization conditions.

For example, in the on-shell renormalization scheme, the renormalized electron mass is defined as the pole of the full propagator, and the residue of that pole is required to be unity. Schematically, these conditions can be written as

$$\Sigma_2(p)\Big|_{p^2=m^2} - \delta_m m = 0, \quad \frac{\mathrm{d}\Sigma_2(p)}{\mathrm{d}\not{p}}\Big|_{p^2=m^2} + \delta_2 = 0. \tag{2.40}$$

The first condition fixes the mass counterterm, while the second fixes the field-strength counterterm. After these subtractions, the physical predictions obtained from the propagator are finite.

The renormalization procedure also introduces the concept of scale dependence. In schemes such as $\overline{\text{MS}}$, the subtraction of divergences introduces an arbitrary renormalization scale μ_R. Physical observables cannot depend on this arbitrary scale when calculated to all orders:

$$\frac{\mathrm{d}\mathcal{O}}{\mathrm{d}\mu_R} = 0, \tag{2.41}$$

where $\mathcal{O}$ denotes a physical observable such as a cross section or decay rate. However, calculations are always performed at finite perturbative order, and a residual dependence on μ_R remains. This residual dependence is commonly used as an estimate of the uncertainty due to missing higher-order corrections.

Although physical observables are independent of μ_R at all orders, the renormalized parameters themselves depend on the scale. For a generic coupling g, this running is described by the renormalization group equation

$$\mu_R \frac{\mathrm{d}g}{\mathrm{d}\mu_R} = \beta(g), \tag{2.42}$$

where $\beta(g)$ is the beta function. In QED, the effective electromagnetic coupling increases slowly with energy. In QCD, the strong coupling decreases at high energies, leading to asymptotic freedom. This scale dependence of couplings is essential for precision collider predictions and for understanding the behavior of the SM at different energy scales.

2.4. Parton Distribution Functions

In high-energy hadron collisions, such as proton-proton collisions at the LHC (described in Sec. 4), the hard interaction occurs between the constituents of the colliding hadrons. These constituents are called partons and include quarks, antiquarks, and gluons. Parton distribution functions (PDFs) describe how the longitudinal momentum of a hadron is distributed among its partons. They are essential non-perturbative inputs for calculating cross sections involving hadrons in the initial state.

In a simple parton-model interpretation, a PDF $f_i(x, \mu_F^2)$ can be viewed as the probability density of finding a parton of type i carrying a fraction x of the proton's longitudinal momentum when probed at the factorization scale μ_F, which separates long and short-distance physics. More precisely, PDFs are universal, scale-dependent functions whose definition depends on the chosen

factorization scheme. They cannot be calculated entirely from first principles using perturbative QCD, because they encode long-distance, non-perturbative physics. Instead, PDFs are extracted from global fits to experimental data from processes such as deep inelastic scattering, Drell–Yan production, vector-boson production, and jet production.

Although PDFs are non-perturbative objects, they obey important sum rules. The momentum sum rule states that the total longitudinal momentum carried by all partons must equal the proton momentum:

$$\sum_i \int_0^1 \mathrm{d}x\, x f_i(x, \mu_F^2) = 1, \tag{2.43}$$

where the sum runs over all active parton species included in the PDF set, including quarks, antiquarks, and gluons. The valence-quark number sum rules express the net quark content of the proton:

$$\begin{aligned} \int_0^1 \mathrm{d}x \left[f_u(x, \mu_F^2) - f_{\bar{u}}(x, \mu_F^2) \right] &= 2, \\ \int_0^1 \mathrm{d}x \left[f_d(x, \mu_F^2) - f_{\bar{d}}(x, \mu_F^2) \right] &= 1, \\ \int_0^1 \mathrm{d}x \left[f_s(x, \mu_F^2) - f_{\bar{s}}(x, \mu_F^2) \right] &= 0. \end{aligned} \tag{2.44}$$

These relations show that the proton has the valence-quark content uud, namely two up quarks and one down quark, which determines its net flavor quantum numbers. In addition to these valence contributions, the proton contains sea quark and antiquark distributions generated by non-perturbative dynamics and perturbative splittings. Since sea quarks appear together with the corresponding antiquark distributions, they do not change the net valence-quark numbers of the proton. Analogous relations apply to other hadrons, with the right-hand side giving their valence-quark content.

The central theoretical result that allows PDFs to be used in collider predictions is the factorization theorem. It separates the short-distance part of the collision, which can be calculated perturbatively, from the long-distance structure of the incoming hadrons, which is described by PDFs. For a proton-proton process producing a final state X, the cross section can be written schematically as

$$\sigma_{pp\to X}(s, \mu_F^2, \mu_R^2) = \sum_{i,j} \int_0^1 \mathrm{d}x_i \int_0^1 \mathrm{d}x_j\, f_i(x_i, \mu_F^2) f_j(x_j, \mu_F^2)$$

$$\times \hat{\sigma}_{ij\to X}\left(\hat{s}, \mu_F^2, \mu_R^2\right), \qquad (2.45)$$

where

$$\hat{s} = x_i x_j s$$

is the partonic center-of-mass energy squared. The quantity $\hat{\sigma}_{ij\to X}$ is the perturbatively calculable hard-scattering cross section for partons i and j producing the final state X. It depends on the factorization scale μ_F, which separates long- and short-distance contributions, and on the renormalization scale μ_R, which enters through the running couplings.

The factorization scale is arbitrary, similarly to the renormalization scale. When all perturbative orders are included, a physical cross section cannot depend on μ_F:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\mu_F} = 0. \qquad (2.46)$$

At finite perturbative order, however, a residual dependence on μ_F remains. Varying μ_F and μ_R is therefore commonly used to estimate the uncertainty due to missing higher-order corrections.

The dependence of PDFs on the factorization scale is governed by the Dokshitzer–Gribov–Lipatov–Altarelli–Parisi (DGLAP) evolution equations [41]:

$$\mu_F^2 \frac{\partial f_i(x, \mu_F^2)}{\partial \mu_F^2} = \sum_j \int_x^1 \frac{\mathrm{d}y}{y} P_{ij}\left(\frac{x}{y}, \alpha_s(\mu_R^2)\right) f_j(y, \mu_F^2). \qquad (2.47)$$

The functions P_{ij} are the splitting functions, which describe the probability for parton j to radiate or split into parton i carrying a fraction x/y of its momentum. They are calculable perturbatively as expansions in the strong fine-structure constant $\alpha_s \equiv g_s^2/4\pi$.

PDFs are usually parameterized at an initial scale $\mu_{F,0}$ of order 1 GeV and then evolved to the scales relevant for collider processes using the DGLAP equations. Different PDF sets, such as CT18 [42], MSHT [43], and NNPDF [44], differ in their input data sets, fitting methodology, parametrization choices, treatment of uncertainties, and theoretical assumptions. Examples of PDFs are shown in Fig. 5. A widely used tool for accessing modern PDF sets is the `LHAPDF` library [45].

The PDFs discussed above are usually called collinear PDFs because they describe only the longitudinal momentum fraction of a parton and neglect its transverse momentum with respect to the parent hadron. For observables that are sensitive to transverse parton motion, transverse-momentum-dependent (TMD) PDFs are needed. TMD PDFs depend not only on x and the scale, but also on the parton's transverse momentum k_T. They are particularly relevant

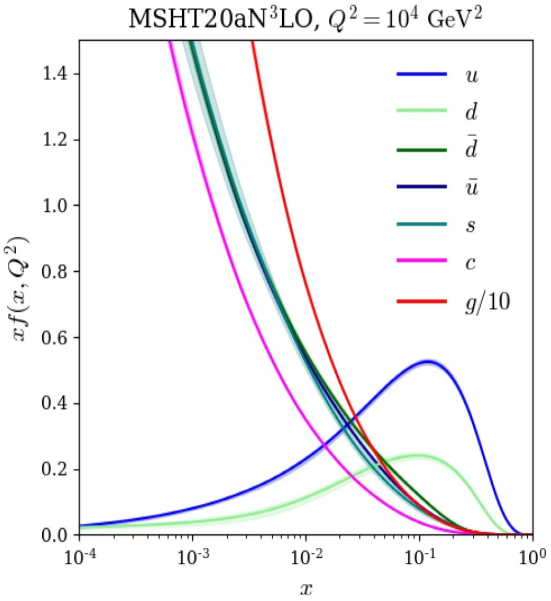


Fig. 5: Examples of parton distribution functions from the MSHT collaboration [43] at a factorization scale of 100 GeV. The PDFs are shown for different parton types, including quarks, antiquarks, and gluons.

for processes such as semi-inclusive deep inelastic scattering and Drell–Yan production at low dilepton transverse momentum. Collections of TMD PDF sets are available through tools such as `TMDlib` [46].

PDFs therefore play a central role in connecting perturbatively calculable partonic cross sections to experimentally measured hadronic observables. Their precise determination and scale evolution are essential for precision SM predictions, for measurements of processes such as Drell–Yan production, and for searches for physics beyond the SM.

2.5. Monte Carlo Simulation

Calculations of hadronic cross sections in high-energy collider physics often involve integrals over high-dimensional phase spaces, representing all possible final-state configurations. These integrals quickly become analytically intractable, especially for processes with many particles in the final state, complicated selection requirements, or higher-order corrections. Monte Carlo (MC) methods provide a powerful numerical approach to evaluating such integrals by replacing analytic integration with statistical sampling.

The basic idea is to approximate a phase-space integral by randomly sam-

pling points according to a chosen probability distribution. Schematically, a cross section can be written as

$$\sigma \propto \int |\mathcal{M}|^2 \, \mathrm{d}\Pi_{\mathrm{LIPS}} \approx \frac{1}{N} \sum_{i=1}^{N} \frac{|\mathcal{M}(p_i)|^2}{f_{\mathrm{sampling}}(p_i)}, \tag{2.48}$$

where p_i denotes the sampled phase-space configuration, $|\mathcal{M}(p_i)|^2$ is the squared matrix element evaluated for that configuration, and $f_{\mathrm{sampling}}(p_i)$ is the sampling probability density. The sampling distribution is usually chosen to approximate the most important features of the integrand, improving the numerical efficiency of the calculation.

As the number of sampled points N increases, the Monte Carlo estimate converges to the exact integral, with a statistical uncertainty that scales as $1/\sqrt{N}$. This convergence is relatively slow, but importantly it depends only weakly on the number of integration dimensions. For this reason, MC integration is particularly useful for the high-dimensional integrals that occur in collider physics.

Each sampled phase-space configuration can be interpreted as a possible collision event. By generating many such events, MC methods produce simulated datasets that can be analyzed in the same way as experimental data. Depending on the generator and the method used, events may be assigned weights proportional to their probability, or they may be unweighted so that each event contributes equally to distributions. Histograms of weighted or unweighted simulated events can then be used to estimate total and differential cross sections.

A complete MC simulation of a proton-proton collision contains several stages beyond the calculation of the hard-scattering matrix element. After the hard process is generated, additional QCD and QED radiation is simulated through initial- and final-state radiation and parton showers. Multi-parton interactions and beam-remnant effects model the additional activity accompanying the primary hard interaction, collectively referred to as the underlying event. Colored partons then undergo hadronization, a non-perturbative process in which quarks and gluons form color-neutral hadrons. Unstable particles subsequently decay, and the resulting stable particles can be passed through a detector simulation to model the response, resolution, and acceptance of a real experiment. Each stage introduces modeling assumptions and corresponding uncertainties, which must be validated against data. The main stages of a simulated proton-proton collision, excluding the detector response, are illustrated in Fig. 6.

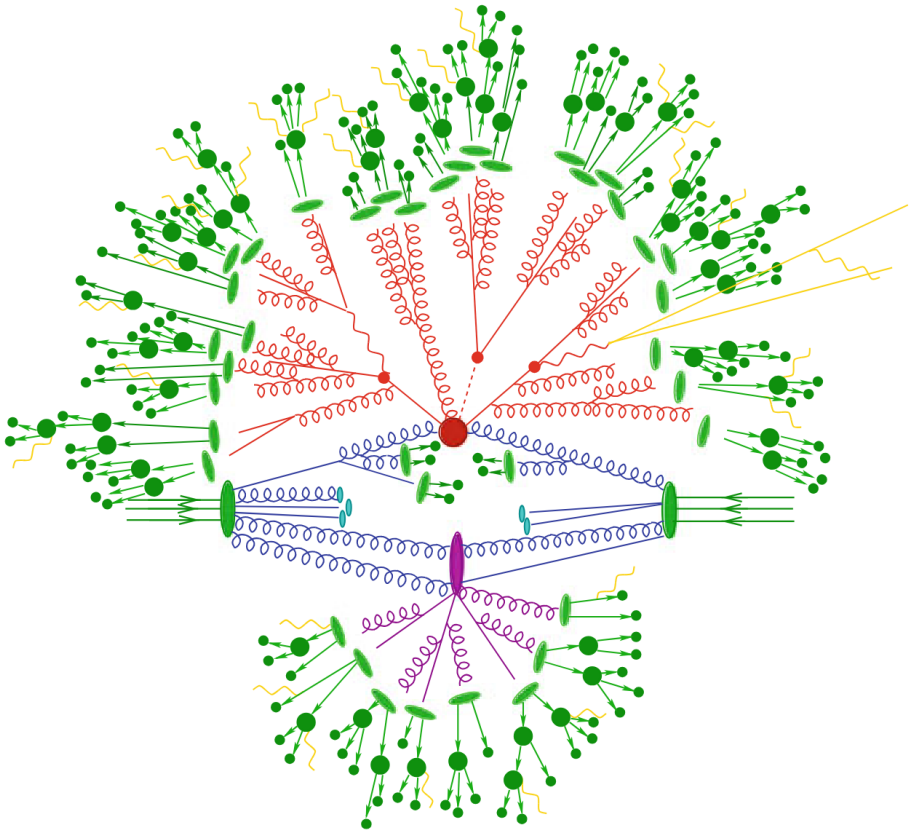


Fig. 6: A graphic representation of a proton-proton collision event as produced by an MC generator. Two large green ellipses, with three green arrows entering each, represent the colliding protons. Partons from the colliding protons are shown as blue lines, two of which meet at a large red circle representing the hard interaction. The red lines represent particles produced in the hard interaction, some of which decay promptly, as indicated by the small red circles. The outgoing colored partons then undergo hadronization, shown by the light-green ellipses, and unstable hadrons decay into longer-lived particles, represented by smaller dark-green circles. The purple ellipse represents the underlying event, while the light-blue ellipses show the proton remnants. Photon radiation, shown as yellow lines, may occur at different stages of the event [47].

Many MC tools are used in collider physics, including Pythia 8 [48], Herwig 7 [49, 50], Sherpa [51], MadGraph5_aMC@NLO [52, 53], Powheg Box [54–56], and FEWZ [57]. General-purpose event generators, such as Pythia, Herwig, and Sherpa, simulate many stages of the event, including hard-process generation, parton showers, hadronization, particle decays, and underlying-event activity. Their built-in hard-process matrix elements are often available at leading order, although modern workflows can combine them with higher-order matrix-element calculations.

Other tools specialize in high-precision calculations of the hard scattering. MadGraph5_aMC@NLO automates the calculation of matrix elements and can provide predictions at leading order or next-to-leading order in QCD, and for selected processes also in the electroweak theory. Powheg Box provides a framework for matching Next-to-Leading Order (NLO) calculations, and in some cases Next-to-Next-to-Leading Order (NNLO) -improved calculations, to parton showers while generating events with positive weights. FEWZ is a fixed-order precision tool for Drell–Yan and vector-boson production, providing cross sections at NNLO in QCD with electroweak corrections. Such specialized calculations are often interfaced with general-purpose generators when a realistic description of the full final state is required.

Monte Carlo simulations therefore play a central role in modern collider physics. They connect theoretical calculations to experimentally observable final states, provide predictions for signal and background processes, support detector-performance studies, and enable the extraction of physical parameters from data. Without MC simulations, it would be impossible to make reliable predictions for the complex multi-particle final states encountered in high-energy collisions at the LHC.

3. Drell–Yan Process

The Drell–Yan (DY) process [14] refers to the production of a charged lepton–antilepton pair in hadron collisions through the annihilation of a quark and an antiquark. In proton-proton (pp) or proton-antiproton ($p\bar{p}$) collisions, the neutral-current DY process proceeds through an intermediate virtual photon or Z boson:

$$q\bar{q} \to \gamma^*/Z \to \ell^+\ell^-,$$

where q and $\bar{q}$ denote a quark and an antiquark, and ℓ denotes a charged lepton, such as an electron, muon, or tau. A leading-order Feynman diagram for this process is shown in Fig. 7. It is analogous to the diagram shown previously in Fig. 3, but written for a general quark flavor and charged-lepton final state.

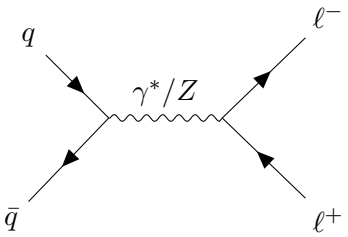


Fig. 7: Leading-order Feynman diagram for the neutral-current Drell–Yan process.

The theoretical foundations of the DY process were introduced by Sidney Drell and Tung-Mow Yan in 1970 [14]. Their work investigated the production of massive lepton pairs in hadron collisions within the quark-parton model [58]. At the time, the internal structure of hadrons and the universality of parton distributions were central open questions. The DY process provided an important test of whether the same partonic picture used to describe deep inelastic scattering could also describe hadron-hadron collisions.

The original Drell–Yan calculation focused on electromagnetic production through an intermediate virtual photon, γ^*. The contribution from the Z boson became relevant later, after the development and experimental confirmation of the electroweak theory, including the discovery of the W and Z bosons in the early 1980s [1–4].

The first experimental observation of the DY process followed shortly after the original proposal, in a fixed-target experiment using proton-uranium collisions at the Alternating Gradient Synchrotron at Brookhaven National Laboratory [59]. These early measurements provided important evidence for the partonic structure of hadrons and helped establish the quark-parton description of high-energy scattering.

Since then, the DY process has become one of the most extensively studied hard-scattering processes in high-energy hadron collisions. It is theoretically well understood and experimentally clean, especially in the electron and muon final states. For this reason, it plays several important roles in particle physics. It provides precision tests of perturbative QCD and electroweak theory, constrains parton distribution functions, serves as a standard candle for detector calibration and performance studies, and forms an important background in many searches for physics beyond the SM.

In the literature, the process mediated by γ^*/Z exchange is often called neutral-current DY production. A related charged-current DY process is also

defined:

$$q\bar{q}' \to W^{\pm} \to \ell^{\pm}\nu_{\ell},$$

where q and $\bar{q}'$ are quarks of different flavors, and ν_{ℓ} denotes the corresponding neutrino or antineutrino. Charged-current DY production is also very important at hadron colliders and is widely used in precision electroweak measurements. However, because this thesis focuses on charged dilepton final states, the term "Drell–Yan process" will hereafter refer specifically to neutral-current DY production.

3.1. Cross Section and Differential Distributions

According to the factorization theorem introduced in Sec. 2.4, the hadronic cross section for the DY process can be expressed as a convolution of parton distribution functions (PDFs) and perturbatively calculable partonic cross sections. For neutral-current DY production in proton-proton collisions, this can be written schematically as

$$\sigma_{pp\to\ell^+\ell^-}(s,\mu_F^2,\mu_R^2) = \sum_{i,j}\int_0^1 \mathrm{d}x_i \int_0^1 \mathrm{d}x_j\, f_i(x_i,\mu_F^2) f_j(x_j,\mu_F^2) \times \hat{\sigma}_{ij\to\ell^+\ell^-}\left(\hat{s},\mu_F^2,\mu_R^2\right). \quad (3.1)$$

At leading order, the neutral-current DY process proceeds through quark-antiquark annihilation,

$$q\bar{q} \to \gamma^*/Z \to \ell^+\ell^-.$$

The corresponding tree-level partonic cross section was derived in Sec. 2.2. At higher orders, additional partonic subprocesses contribute, including real-emission processes such as

$$q\bar{q} \to \ell^+\ell^- g, \qquad qg \to \ell^+\ell^- q$$

(where g denotes a gluon), as well as virtual loop corrections. These higher-order contributions are essential for precision predictions and for describing observables such as the dilepton transverse momentum.

The convolution of the partonic cross section with the PDFs in Eq. (3.1) makes the hadronic DY process sensitive not only to the electroweak couplings and perturbative QCD corrections, but also to the internal partonic structure of the proton. While the leading-order partonic process $q\bar{q} \to \ell^+\ell^-$ can be described by a small number of kinematic variables, the experimentally relevant hadronic DY process is more naturally expressed in terms of variables describing

the dilepton system and the angular orientation of the final-state leptons. The fully differential neutral-current DY cross section is commonly written as

$$\frac{\mathrm{d}^5\sigma_{pp\to\ell^+\ell^-}}{\mathrm{d}M\,\mathrm{d}y\,\mathrm{d}p_\mathrm{T}\,\mathrm{d}\cos\theta^*\,\mathrm{d}\phi^*}, \tag{3.2}$$

where M, y, and p_T describe the invariant mass, rapidity, and transverse momentum of the dilepton system, while θ^* and ϕ^* describe the lepton angular distribution in a chosen dilepton rest frame.

The invariant mass of the dilepton system is defined as

$$M \equiv \sqrt{p_{\ell\ell}^2} = \sqrt{E_{\ell\ell}^2 - |\vec{p}_{\ell\ell}|^2}, \tag{3.3}$$

where

$$p_{\ell\ell} = p_{\ell^+} + p_{\ell^-}.$$

The rapidity of the dilepton system is

$$y = \frac{1}{2}\ln\left(\frac{E_{\ell\ell} + p_{z,\ell\ell}}{E_{\ell\ell} - p_{z,\ell\ell}}\right), \tag{3.4}$$

and its transverse momentum is

$$p_\mathrm{T} = \sqrt{p_{x,\ell\ell}^2 + p_{y,\ell\ell}^2}. \tag{3.5}$$

The angles θ^* and ϕ^* are often defined in the Collins–Soper frame [60], a dilepton rest frame designed to reduce the ambiguity associated with the unknown initial-state quark direction in hadron collisions. This frame is illustrated in Fig. 8.

In practice, experimental measurements often integrate over some of the five variables in Eq. (3.2). For example, measurements may be performed as one-dimensional distributions in M or p_T, two-dimensional distributions in (M, y), or angular measurements in bins of the dilepton kinematics. The choice of observables depends on the available data sample, detector acceptance, resolution, and the physics question being addressed. Invariant-mass distributions are particularly important for testing the overall DY spectrum and searching for high-mass deviations, and together with rapidity and angular distributions provide strong sensitivity to PDFs and electroweak parameters. The dilepton transverse-momentum distribution is especially sensitive to QCD radiation and soft-gluon resummation effects.

Modern theoretical predictions for the DY process include high-order QCD and electroweak corrections. The process is known at NLO in the electroweak theory [62–64], at NNLO in QCD [57, 65–68], and with mixed QCD–EW

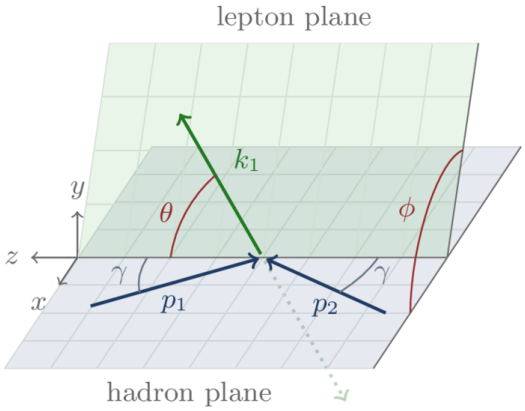


Fig. 8: An illustration of the Collins–Soper frame [61]. It is a dilepton center-of-mass frame, where k_1 and k_2 denote the lepton momenta and p_1 and p_2 denote the momenta of the incoming protons. The z-axis is chosen to bisect the angle between p_1 and $-p_2$, while the x-axis is chosen to bisect the angle between p_1 and p_2. The y-axis is perpendicular to the x–z plane. The polar angle θ^* and azimuthal angle ϕ^* describe the lepton orientation in this frame.

corrections [69–75]. Several calculations of N^{3}LO QCD corrections are also available for inclusive cross sections and some differential observables [76–81]. Tools such as FEWZ, DYTurbo, and Powheg Box are commonly used to obtain precise theoretical predictions for DY observables, often with different compromises between fixed-order accuracy, resummation, and event-level simulation.

An example comparison of mass-dependent one-dimensional DY differential distributions,

$$\Sigma(Q^2) = \Sigma(M^2) = M^2 \frac{\mathrm{d}\sigma}{\mathrm{d}M^2} = \frac{M}{2} \frac{\mathrm{d}\sigma}{\mathrm{d}M}, \tag{3.6}$$

calculated at different perturbative orders is shown in Fig. 9.

3.2. Current State of Drell–Yan Measurements

Experimental measurements of the Drell–Yan process play a crucial role in particle physics, as they directly probe the electroweak sector of the SM and provide important information about the partonic structure of hadrons. Experiments typically measure integrated or differential cross sections in proton-proton or proton-antiproton collisions by selecting events with charged dilepton final states, most commonly e^+e^- and $\mu^+\mu^-$. The $\tau^+\tau^-$ final state is experimentally

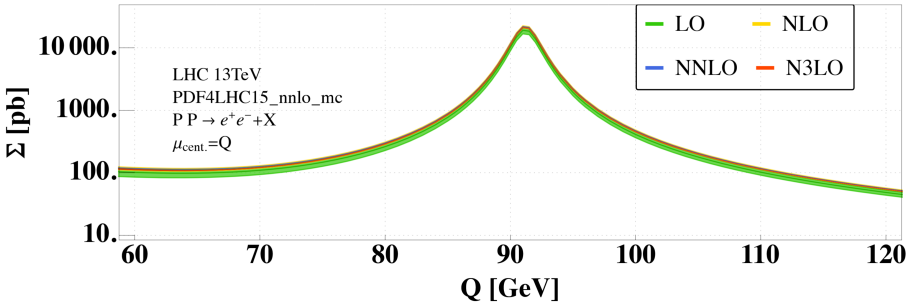


Fig. 9: A comparison of one-dimensional Drell–Yan differential cross sections $\Sigma(Q^2)$, Eq. (3.6) at the proton-proton collision energy $\sqrt{s} = 13$ TeV in the dilepton mass range $60 \leq M \leq 120$ GeV, calculated at different accuracies from LO to N^{3}LO in QCD [77].

more challenging because of the short lifetime of the tau lepton and the presence of neutrinos in many tau decays. For this reason, measurements involving tau leptons are usually performed separately from the electron and muon channels.

Drell–Yan measurements serve several important purposes. First, they provide precision tests of perturbative calculations in the SM. By comparing measured distributions with theoretical predictions, experiments can test QCD and electroweak calculations at different perturbative orders. As discussed in Sec. 3.1, theoretical predictions for DY production are available at NNLO in QCD, NLO in the electroweak theory, and with mixed QCD–EW corrections, while N^{3}LO QCD calculations are becoming available for increasingly more differential observables. To fully exploit this theoretical precision, experimental measurements must achieve comparably well-controlled statistical and systematic uncertainties. The DY process is particularly well suited for such tests because the electron and muon final states provide clean experimental signatures, with precisely reconstructed leptons and relatively low backgrounds.

Another key application of DY measurements is the extraction and validation of proton PDFs. The DY process is sensitive to both valence- and sea-quark distributions, since it involves a quark and an antiquark in the initial state. The relative contributions of up-type and down-type quarks differ because of their different electroweak couplings, making DY measurements sensitive to the flavor composition of the proton. Low- and intermediate-mass DY production provides important constraints on quark and antiquark PDFs at relatively small and moderate values of the parton momentum fraction x. High-mass DY production probes larger values of x, where PDF uncertainties are typically larger. In addition, DY events with nonzero dilepton transverse momentum are sensitive

to QCD radiation and therefore provide information relevant to higher-order QCD effects and, indirectly, to the gluon PDF through subprocesses involving initial-state gluons.

The DY process is also an important background in many searches for physics beyond the SM. In searches for heavy resonances decaying into charged lepton pairs, such as hypothetical Z' bosons or gravitons predicted in extra-dimensional models, DY production is usually the dominant irreducible background [82–89]. Accurate measurements of the DY cross section are therefore essential for validating background modeling and reducing uncertainties in BSM searches.

Although the fully differential DY cross section is five-dimensional, as discussed in Sec. 3.1, most experimental measurements are performed in one, two, or occasionally three dimensions. This reduction is usually necessary because of limited event yields, detector acceptance, and the increasing complexity of unfolding and uncertainty treatment in higher-dimensional distributions. Common observables include the dilepton invariant mass, rapidity, transverse momentum, and angular variables.

Over the past decade, precise measurements of the DY process have been performed by the ATLAS and CMS collaborations at the LHC. Both collaborations have reported differential cross-section measurements at $\sqrt{s} = 7$ TeV [15, 16], 8 TeV [17, 18], and 13 TeV [19, 20, A1, 90]. These measurements include one-dimensional [15–17, 19, 20, 90], two-dimensional [15, 17, A1, 90], and, in some cases, three-dimensional [18] distributions. An example of the one-dimensional differential cross-section measurement $\mathrm{d}\sigma/\mathrm{d}M$ at $\sqrt{s} = 13$ TeV performed by the CMS collaboration is shown in Fig. 10.

The CMS collaboration has also performed measurements of the Drell–Yan forward–backward asymmetry [91–93] and of the weak mixing angle [94], both of which are closely related to the angular structure of the DY process. In addition to proton-proton measurements, the ALICE, ATLAS, and CMS collaborations have measured DY production in proton-lead collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV [95–97] and $\sqrt{s_{\mathrm{NN}}} = 8.16$ TeV [98, 99]. These measurements are particularly useful for studying nuclear modifications of PDFs. Further measurements of DY production have also been performed at lower-energy experiments, such as PHENIX at RHIC [100], SeaQuest at Fermilab [101], and COMPASS at CERN [102].

Together, these measurements have significantly improved our understanding of the DY process. They provide stringent tests of perturbative QCD and electroweak predictions, constrain the partonic structure of the proton and nuclei, and improve the modeling of an important background in searches for physics

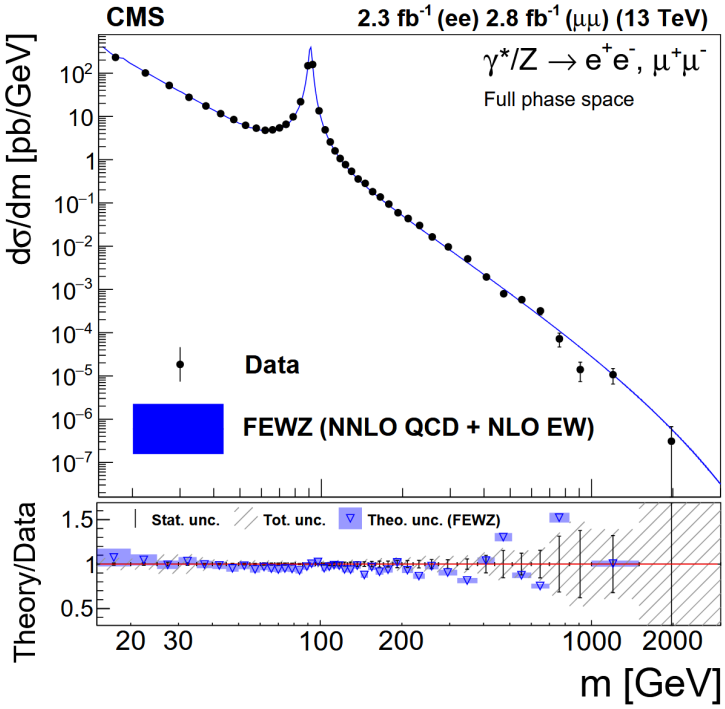


Fig. 10: The differential cross section $\mathrm{d}\sigma/\mathrm{d}M$ of the Drell–Yan process measured by the CMS collaboration at $\sqrt{s} = 13$ TeV [19]. A comparison with the theoretical prediction calculated with FEWZ at NNLO in QCD and NLO in the electroweak theory is also shown.

beyond the SM.

EXPERIMENTAL FRAMEWORK

4. The Large Hadron Collider at CERN

The Large Hadron Collider (LHC) at CERN is the highest-energy particle collider currently in operation [9]. It is installed in a 27 km underground tunnel crossing the French–Swiss border near Geneva, at depths between approximately 50 and 175 m. The LHC is designed to collide proton beams at a center-of-mass energy of up to 14 TeV and to reach instantaneous luminosities above 10^{34} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

The first LHC physics run, Run–1, took place in 2010–2012 at center-of-mass energies of 7 and 8 TeV and culminated in the discovery of the Higgs boson by the ATLAS and CMS collaborations [7, 8]. Run–2, from 2015 to 2018, operated at 13 TeV and provides the proton-proton collision data used for the Drell–Yan measurement presented in this thesis. Run–3 began in 2022 at 13.6 TeV and, at the time of writing, is scheduled to continue until mid-2026. It will be followed by the third long shutdown (LS3), during which the accelerator and experiments will undergo major upgrades for the HL-LHC program. The start of Run–4, corresponding to the beginning of HL-LHC operation, is currently planned for 2030 [103].

4.1. Particle acceleration and CERN's accelerator complex

The LHC reaches its collision energy through a chain of accelerators in the CERN accelerator complex. Protons are first accelerated in LINAC4 to 160 MeV. They are then passed through the Proton Synchrotron Booster (PSB), the Proton Synchrotron (PS), and the Super Proton Synchrotron (SPS), reaching energies of 2 GeV, 26 GeV, and 450 GeV, respectively. Finally, the beams are injected into the LHC, where they are accelerated to their final energy. During Run–2, the beam energy was 6.5 TeV, corresponding to a proton-proton collision energy of 13 TeV. During Run–3, the beam energy is 6.8 TeV, corresponding to 13.6 TeV collisions.

Before Run–3, the first stage of the proton injector chain was LINAC2, which accelerated protons to approximately 50 MeV. LINAC4 replaced LINAC2 as part of the LHC Injector Upgrade program, improving the injector performance for Run–3 and the HL-LHC era. Heavy ions can also be accelerated through the CERN accelerator chain, although the initial stages differ from those used for protons.

4.2. Proton beams at the LHC

The LHC uses superconducting dipole and quadrupole magnets to guide and focus the proton beams. The dipole magnets bend the beams along the circular trajectory of the accelerator, while quadrupole magnets focus the beams and reduce their transverse size at the interaction points. The main dipole magnets operate at a temperature of about 1.9 K, cooled by superfluid helium, which allows them to produce the magnetic fields required to keep multi-TeV proton beams on their orbit.

The proton beams are organized into bunches. Each bunch contains approximately 10^{11} protons and has a length of several centimeters. In nominal 25 ns operation, bunch crossings occur at a rate of up to 40 MHz. The LHC has four main interaction points where the beams are brought into collision, hosting the ATLAS [11], CMS [10], ALICE [13], and LHCb [12] experiments.

4.3. Luminosity

Luminosity characterizes the collision rate delivered by a collider. For a process with cross section σ, the event rate R is related to the instantaneous luminosity $\mathcal{L}$ by

$$R = \mathcal{L}\sigma. \tag{4.1}$$

The instantaneous luminosity is typically expressed in units of $\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

For a collider with Gaussian beams, the instantaneous luminosity can be written approximately as

$$\mathcal{L} = \frac{n_b\, N_1\, N_2\, f_{\mathrm{rev}}\, \gamma}{4\pi \varepsilon_n \beta^*} F, \tag{4.2}$$

where n_b is the number of bunches per beam, N_1 and N_2 are the numbers of protons per bunch in the two beams, f_{rev} is the revolution frequency, γ is the Lorentz factor of the beams, ε_n is the normalized transverse beam emittance, β^* is the beta function at the interaction point, and F is a geometric reduction factor accounting for the crossing angle and other beam-geometry effects.

During Run–2, the LHC achieved peak instantaneous luminosities of approximately 2×10^{34} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$ in proton-proton operation. The integrated luminosity,

$$\mathcal{L}_{\mathrm{int}} = \int \mathcal{L}(t)\, \mathrm{d}t, \tag{4.3}$$

determines the total number of events available for analysis. In LHC measurements, it is usually expressed in inverse femtobarns, fb^{-1}. During Run–2, the

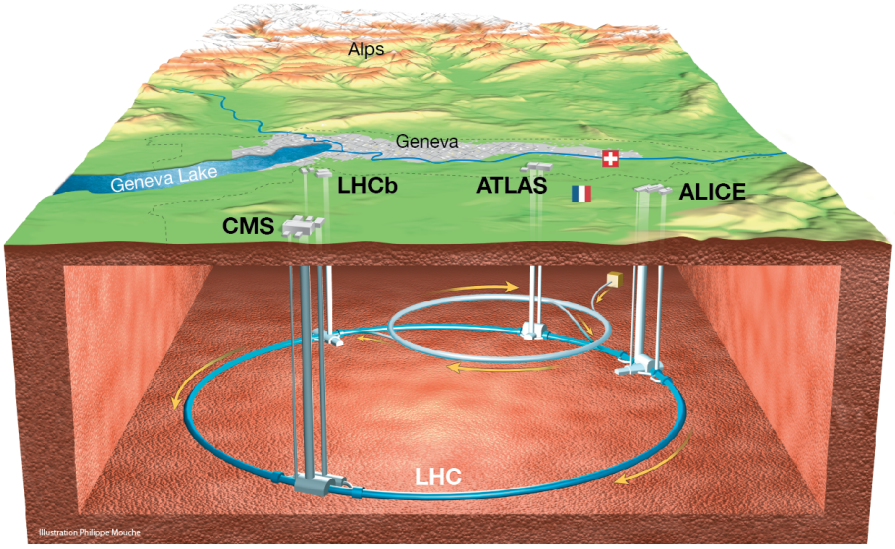


Fig. 11: Schematic drawing of the LHC and its experiments [105].

ATLAS and CMS experiments each recorded more than 150 fb^{-1} of proton-proton collision data [104]. The Drell–Yan measurement discussed in this thesis uses the CMS data collected in 2016–2018, corresponding to an integrated luminosity of approximately 137.6 fb^{-1}.

4.4. LHC experiments

The LHC hosts four large experiments located at its main interaction points. ATLAS [11] and CMS [10] are general-purpose detectors designed for a broad physics program, including Higgs-boson measurements, precision SM studies, and searches for physics beyond the SM. ALICE [13] is optimized for heavy-ion physics and the study of strongly interacting matter under extreme conditions. LHCb [12] is dedicated primarily to flavor physics, CP violation, and rare decays involving heavy-flavor hadrons.

The CMS detector, which collected the data used in this thesis, is described in detail in Sec. 5. A schematic view of the LHC and its experiments is shown in Fig. 11.

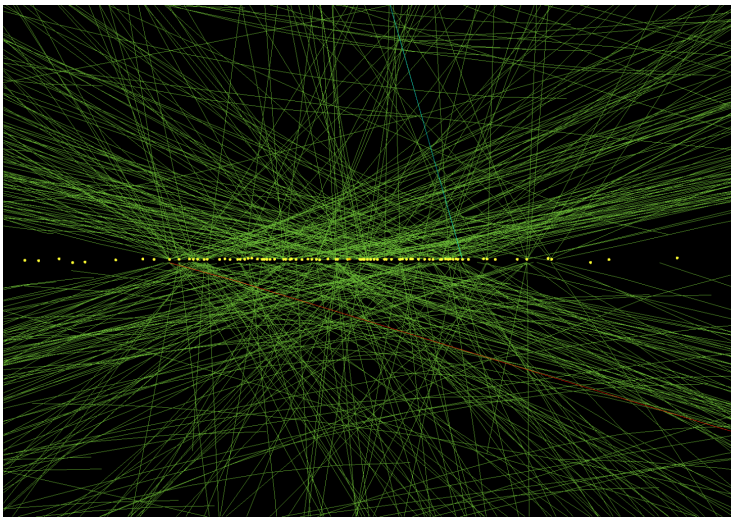


Fig. 12: An example of pileup in the CMS experiment. The image shows a collision event recorded in 2012 with multiple tracks from different interactions in the same bunch crossing. A total of 78 primary vertices were reconstructed [106].

4.5. Pileup

Pileup (PU) refers to additional proton-proton interactions that occur in the same or nearby bunch crossings as the hard interaction of interest. It is a direct consequence of the high instantaneous luminosity of the LHC and presents an important challenge for event reconstruction and physics analyses. PU increases the number of detector hits, tracks, and calorimeter energy deposits, complicating the reconstruction of primary vertices, leptons, jets, and missing transverse momentum. An example of a CMS event with many reconstructed vertices is shown in Fig. 12.

Two types of PU are commonly distinguished. In-time PU consists of additional interactions occurring in the same bunch crossing as the hard interaction. Out-of-time PU arises from signals produced in neighboring bunch crossings that overlap with the detector readout window. Out-of-time PU is particularly relevant for detector systems whose signals extend over several bunch crossings, such as calorimeters.

The mean number of interactions per bunch crossing, denoted by μ, is related to the instantaneous luminosity by

$$\mu = \frac{\mathcal{L}\sigma_{\text{inel}}}{f_{\text{BX}}}, \tag{4.4}$$

where σ_{inel} is the inelastic proton-proton cross section and f_{BX} is the bunch-

crossing frequency. During Run–2, typical PU values were of order several tens of interactions per bunch crossing, depending on the data-taking year and instantaneous luminosity. In the HL-LHC era, the average PU is expected to increase to about 140–200 interactions per bunch crossing, requiring more advanced reconstruction algorithms and detector upgrades.

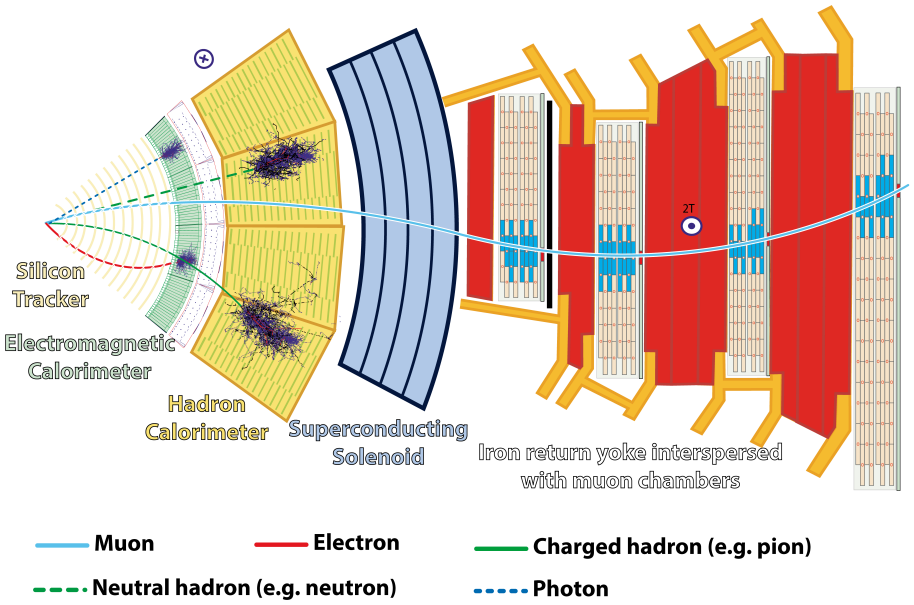


Fig. 13: A cross-sectional slice of the CMS detector [107].

5. The Compact Muon Solenoid experiment

The Compact Muon Solenoid (CMS) is one of the two general-purpose detectors at the LHC [10]. It is located at Point 5 near Cessy, France, approximately 100 m underground. CMS is designed to reconstruct a wide range of final-state particles and to support a broad physics program, including precision SM measurements, Higgs-boson studies, and searches for physics beyond the SM.

A defining feature of CMS is its superconducting solenoid, which provides a magnetic field of 3.8 T. Inside the solenoid volume are the silicon tracker, the electromagnetic calorimeter, and the hadronic calorimeter. The muon system is embedded in the steel return yoke outside the solenoid. The detector is approximately 21 m long, 15 m in diameter, and has a mass of about 14 000 t. A cross-sectional slice of CMS is shown in Fig. 13.

The detector has a cylindrical geometry around the beam axis and is divided into barrel and endcap regions. The barrel region covers particles produced at relatively large angles with respect to the beam direction, while the endcaps extend the detector acceptance to more forward directions. This geometry provides nearly hermetic coverage for reconstructing charged particles, photons, jets, muons, and missing transverse momentum.

5.1. CMS coordinate system

The CMS detector uses a right-handed coordinate system centered at the nominal interaction point. The z-axis is oriented along the counterclockwise beam direction, the x-axis points horizontally toward the center of the LHC ring, and the y-axis points vertically upward.

Particle directions are commonly described using the azimuthal angle ϕ and the polar angle θ. The azimuthal angle is measured in the plane transverse to the beam axis, while the polar angle is measured with respect to the positive z-axis. In collider physics, the polar angle is usually replaced by the pseudorapidity,

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right]. \tag{5.1}$$

With this definition, particles moving close to the positive z-axis have large positive η, while particles moving close to the negative z-axis have large negative η. The central region of the detector corresponds to $\eta \approx 0$.

Pseudorapidity is useful because differences in rapidity are invariant under Lorentz boosts along the beam axis, and for ultra-relativistic particles the pseudorapidity approximates the rapidity,

$$y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right). \tag{5.2}$$

The angular distance between two reconstructed objects is commonly expressed as

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \tag{5.3}$$

The momentum component transverse to the beam axis is denoted by p_{T} and is defined as

$$p_{\mathrm{T}} = \sqrt{p_x^2 + p_y^2}. \tag{5.4}$$

Together with the mass, pseudorapidity, and azimuthal angle, the transverse momentum provides a convenient parameterization of particle four-momenta in collider analyses. The CMS coordinate conventions are illustrated in Fig. 14.

5.2. Tracker

The innermost component of CMS is the silicon tracker, which provides precise measurements of charged-particle trajectories [10]. During Run–2, the tracker covered the pseudorapidity range $|\eta| < 2.5$ and consisted of an inner pixel detector surrounded by a silicon strip detector. The pixel detector, located closest to the interaction point, provides high-granularity measurements essential

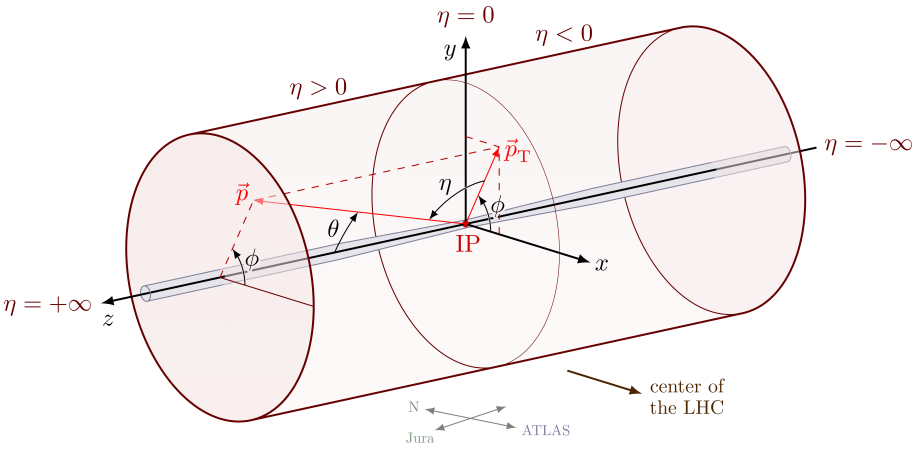


Fig. 14: The coordinate system used by the CMS detector [108].

for primary-vertex reconstruction, impact-parameter measurements, and the identification of secondary vertices. The strip detector extends the tracking volume to larger radii and provides additional measurements for reconstructing charged-particle momenta.

The tracker is required to provide high spatial resolution, fast response, and robust operation in a high-pileup and high-radiation environment, while minimizing the amount of material traversed by particles. Silicon sensors satisfy these requirements by providing compact segmentation, fast charge collection, and sufficient radiation tolerance for operation close to the interaction point.

Silicon detectors operate by collecting charge carriers produced when a charged particle traverses a reverse-biased semiconductor sensor. The reverse bias creates a depleted region with a strongly reduced density of free charge carriers and an electric field across the sensor volume. A charged particle passing through the sensor produces electron-hole pairs through ionization. These charges drift toward the electrodes, where they are collected and read out as an electrical signal. A simplified illustration of silicon-sensor operation is shown in Fig. 15.

The original CMS tracker contained approximately 66 million n^+-in-n silicon pixels with a size of 100×150 μm^2, arranged in three barrel layers and two endcap disks on each side. The surrounding strip detector contained approximately 9.3 million silicon strips arranged in barrel layers and endcap disks [10].

During the Phase–1 upgrade, completed before the main 2017 data-taking period, the pixel detector was replaced to improve tracking performance under the higher instantaneous luminosities of Run–2. The upgraded pixel detector

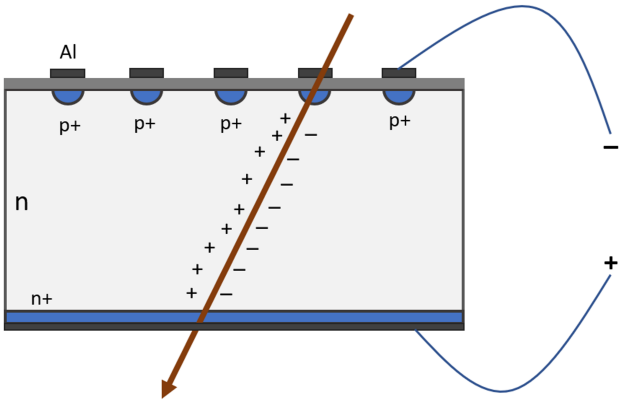


Fig. 15: A sketch of a p-in-n silicon sensor operation. A charged particle traversing the sensor creates electron-hole pairs, which are separated by the electric field and collected by the electrodes. Reversing the doping types and voltage polarity gives the corresponding n-in-p or n^+-in-n sensor configurations.

retained the same pixel size but increased the number of barrel layers from three to four and the number of endcap disks from two to three on each side. The number of pixels increased to approximately 124 million [109]. The upgraded readout chip provided improved rate capability and radiation tolerance, which helped maintain high tracking efficiency and vertexing performance in the presence of increased PU.

The tracker is central to the reconstruction of physics objects used in the Drell–Yan analysis. It provides the momentum measurement for charged particles, supports electron and muon reconstruction, enables vertex identification, and contributes to the suppression of backgrounds from non-prompt particles and PU.

5.3. Electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) surrounds the tracker and measures the energies and positions of electrons and photons [10]. It is a homogeneous calorimeter made of lead tungstate ($PbWO_4$) scintillating crystals. Electrons and photons initiate electromagnetic showers in the crystals and deposit most of their energy in the ECAL. The scintillation light is detected by avalanche photodiodes in the barrel region and vacuum phototriodes in the endcaps.

Lead tungstate was chosen because it is dense, has a short radiation length, and has a fast scintillation response. These properties allow the ECAL to be compact, highly granular, and compatible with the 25 ns bunch spacing of the

LHC. The relatively low light yield of $PbWO_4$ is compensated by sensitive photodetectors and precise calibration.

The ECAL consists of a barrel section and two endcaps. It contains 75 848 $PbWO_4$ crystals and covers the range $|\eta| < 3.0$. The barrel covers $|\eta| < 1.48$, while the endcaps extend the coverage to $|\eta| < 3.0$. The barrel crystals have a granularity of approximately 0.0174×0.0174 in η–ϕ space and a length corresponding to about 25.8 radiation lengths. The endcap crystals have a length of about 24.7 radiation lengths.

In the forward region, $1.65 < |\eta| < 2.6$, the ECAL is preceded by a preshower detector made of lead absorbers and silicon strip sensors. The preshower improves photon identification and helps distinguish single photons from close photon pairs produced in neutral-pion decays. The transition region between the ECAL barrel and endcap at $1.4442 < |\eta| < 1.5660$ has reduced reconstruction performance and is commonly excluded from analyses using electrons or photons.

For electrons and photons with energies around 100 GeV, the ECAL provides an energy resolution at the percent level or better, depending on the detector region and shower containment. The energy resolution is commonly parameterized as

$$\frac{\sigma_E}{E} = \sqrt{\left(\frac{2.8\%}{\sqrt{E\ [\text{GeV}]}}\right)^2 + \left(\frac{12.5\%}{E\ [\text{GeV}]}\right)^2 + (0.3\%)^2}, \tag{5.5}$$

where the three terms describe stochastic, noise, and constant contributions, respectively [110]. The ECAL performance is especially important for the dielectron channel of the Drell–Yan measurement.

5.4. Hadronic calorimeter

The hadronic calorimeter (HCAL) is located outside the ECAL and measures the energies of charged and neutral hadrons [10]. It is essential for reconstructing hadronic jets, estimating missing transverse momentum, and identifying hadronically decaying tau leptons. It also provides the calorimetric information used by the Particle Flow reconstruction.

The HCAL is a sampling calorimeter, using dense absorber material interleaved with scintillator layers. In the barrel and endcap regions, hadrons interact with brass and steel absorbers and produce hadronic showers. The shower particles produce scintillation light in plastic scintillator tiles, which is collected and converted into electrical signals. During the Phase–1 upgrade, the original hybrid photodiodes were replaced by silicon photomultipliers in the

barrel and endcap regions, improving photon detection efficiency, segmentation, calibration, and operation under radiation-damaged conditions [111].

The HCAL is divided into several regions. The HCAL Barrel (HB) covers $|\eta| < 1.3$ and is located inside the solenoid, directly outside the ECAL barrel. The HCAL Endcaps (HE) extend the coverage to $|\eta| < 3.0$. The HCAL Outer (HO), located outside the solenoid in the barrel region, provides additional sampling to recover energy from showers that leak beyond the HB. The HCAL Forward (HF) extends the calorimetric coverage to $3.0 < |\eta| < 5.2$ and uses quartz fibers embedded in a steel absorber matrix, where shower particles produce Cherenkov light.

The HCAL energy resolution is lower than that of the ECAL because hadronic showers have larger intrinsic fluctuations. Nevertheless, the HCAL provides the information needed for jet reconstruction and missing-transverse-momentum measurements. These quantities are important in many CMS analyses and are also used in the background estimation and event characterization of Drell–Yan measurements.

5.5. Solenoid magnet

A central feature of CMS is its superconducting solenoid magnet, which provides a magnetic field of 3.8 T in the inner detector volume. The magnetic field bends the trajectories of charged particles, allowing their momenta to be measured from the curvature of their tracks. This is essential for precise reconstruction of charged leptons and hadrons.

The solenoid is approximately 12.5 m long and 6 m in internal diameter, enclosing the tracker, ECAL, and HCAL. It is constructed from superconducting niobium-titanium coils and operates at cryogenic temperature. The magnetic flux is returned through a large steel yoke, which also serves as the absorber structure for the muon system, described in the next subsection. The combination of the strong magnetic field and the precise silicon tracker gives CMS excellent charged-particle momentum resolution, especially for muons and electrons relevant to the Drell–Yan analysis.

5.6. Muon system

The muon system is the outermost major subsystem of CMS and is designed to identify muons, provide triggering information, and contribute to momentum measurements [10]. Muons are minimum-ionizing particles that typically pass through the tracker and calorimeters with relatively small energy loss. They

are therefore detected in gaseous chambers embedded in the steel return yoke outside the solenoid.

The Run–2 CMS muon system used three detector technologies. Drift Tubes (DTs) cover the barrel region, approximately $|\eta| < 1.2$, and provide precise position measurements where the magnetic field and particle rates are relatively moderate. Cathode Strip Chambers (CSCs) are used in the endcaps, approximately $0.9 < |\eta| < 2.4$, and provide precise spatial and timing measurements where the magnetic field is less uniform and the particle rates are higher. Resistive Plate Chambers (RPCs) cover approximately $|\eta| < 1.9$ and complement the DT and CSC systems by providing fast timing information for the Level–1 trigger.

Gas Electron Multiplier (GEM) detectors are being added to improve the forward muon system for high-luminosity operation [112]. The first GEM station was installed after Run–2 and therefore was not used in the Drell–Yan measurement presented in this thesis. The Phase–2 muon upgrades are discussed in Sec. 14.

The muon system can reconstruct standalone muon tracks using only hits in the muon detectors. For most physics analyses, however, the best performance is obtained by combining muon-system information with tracks reconstructed in the silicon tracker. This combination provides high muon identification efficiency and good momentum resolution over a wide range of transverse momenta. The muon system is therefore essential for the dimuon channel of the Drell–Yan measurement.

5.7. Trigger and data acquisition system

The LHC bunch-crossing rate of up to 40 MHz produces far more collision data than can be permanently stored. At the design luminosity, the raw data rate would be tens of terabytes per second, while only a small fraction of events can be written to permanent storage. CMS therefore uses a two-level trigger system to select events of interest in real time and reduce the data rate to a manageable level [10, 113].

The first level, the Level–1 (L1) trigger, is implemented in custom hardware and uses coarse information from the calorimeters and muon detectors. During Run–2, the L1 trigger reduced the event rate from the bunch-crossing rate to at most about 100 kHz within a fixed latency of approximately $4\ \mu$s [114]. The L1 trigger reconstructs simplified physics objects, including electron/photon candidates, muons, jets, hadronically decaying tau candidates, and missing transverse momentum. These objects are passed to the Global Trigger, which

applies programmable selection requirements and decides whether the event should be read out.

The L1 trigger consists of three main blocks: the Calorimeter Trigger, the Muon Trigger, and the Global Trigger The Calorimeter Trigger processes trigger primitives from the ECAL and HCAL to reconstruct energy deposits and global energy sums. The Muon Trigger combines information from the Drift Tube, Cathode Strip Chamber, and Resistive Plate Chamber systems to reconstruct muon candidates. The Global Trigger receives the reconstructed calorimeter and muon objects and applies the final L1 trigger menu.

The L1 trigger underwent a major Phase–1 upgrade before the 2016 data-taking period. The upgrade replaced the original electronics with FPGA-based processors and high-bandwidth optical links, improving the granularity and flexibility of the trigger algorithms. It also improved calorimeter clustering, introduced upgraded regional muon track finders, and enabled the use of additional correlations between trigger objects. These changes improved trigger efficiency and background rejection under the higher luminosity and pileup conditions of Run–2 [114].

Events accepted by the L1 trigger are processed by the High-Level Trigger (HLT), a software trigger running on a large processor farm. The HLT has access to the full detector readout and uses reconstruction algorithms similar to those used offline, although optimized for real-time processing. This allows the HLT to apply more precise selections based on reconstructed leptons, photons, jets, missing transverse momentum, and event-level quantities. During Run–2, the HLT reduced the event rate from approximately 100 kHz to about 1 kHz for permanent storage.

The HLT selection is organized into trigger paths, each targeting a particular event signature. Examples include single-lepton, dilepton, photon, jet, and missing-transverse-momentum triggers. Each path applies a sequence of object-reconstruction and selection steps, such as transverse-momentum thresholds, identification requirements, isolation criteria, or topological selections. The HLT menu is updated and validated regularly to maintain high efficiency for physics analyses while keeping the output rate and computing time within operational limits.

The Data Acquisition (DAQ) system is responsible for reading out, assembling, and transferring events accepted by the trigger system [10, 113]. Upon receiving an L1 accept signal, detector front-end electronics transmit buffered data through Front-End Drivers. The event builder assembles fragments from different subdetectors into complete event records, which are then processed by the HLT farm. Events accepted by the HLT are transferred to permanent

storage and subsequently reconstructed offline.

The DAQ system is designed to sustain the full L1 output rate while maintaining synchronization between detector subsystems and preventing data loss. To protect against buffer overflows and backpressure, the trigger-throttling system provides rapid feedback to the trigger and front-end electronics, dynamically limiting the trigger rate when necessary.

For the Drell–Yan measurement presented in this thesis, the trigger system is especially important because the selected events are required to contain high-quality electrons or muons. The analysis relies on lepton trigger paths whose efficiencies must be measured and corrected for in simulation, as discussed in Sec. 5.8.6.

5.8. Event reconstruction

Event reconstruction transforms the detector signals recorded by CMS into physics objects that can be used in analyses. These objects include electrons, muons, photons, hadronic jets, and missing transverse momentum. The reconstruction combines information from multiple subdetectors to achieve high efficiency and good resolution while suppressing backgrounds from PU, detector noise, and misidentified particles.

5.8.1. Particle Flow

The Particle Flow (PF) algorithm [115] is the main event reconstruction strategy used by CMS. It combines information from all subdetectors to identify and reconstruct each long-lived particle in the event. Compared with reconstruction based on individual subdetectors alone, PF improves the measurement of jets, missing transverse momentum, charged leptons, and photons by using the detector component that provides the best resolution for each particle type.

The PF reconstruction starts from basic elements: tracks reconstructed in the silicon tracker, tracks and segments reconstructed in the muon system, and energy clusters reconstructed in the ECAL and HCAL. These elements are linked using spatial and kinematic compatibility. The resulting PF candidates are classified as charged hadrons, neutral hadrons, photons, electrons, or muons.

Charged hadrons are primarily measured using tracks in the silicon tracker, which provide better momentum resolution than the calorimeters for many charged particles. Photons are reconstructed mainly from ECAL clusters without associated tracks. Neutral hadrons are reconstructed from calorimeter energy deposits not associated with charged particles or photons. Electrons and muons

are reconstructed using dedicated algorithms that combine tracker information with ECAL clusters or muon-system hits, respectively.

Pileup is an important challenge for PF reconstruction. Charged particles associated with PU vertices can often be identified using tracking information and removed or down-weighted. Neutral particles from PU are more difficult to identify and require additional correction methods. These mitigation techniques are essential for stable jet, isolation, and missing-transverse-momentum reconstruction in high-PU conditions.

5.8.2. Electron and photon reconstruction

Electron and photon reconstruction in CMS relies primarily on the tracker and ECAL. Both particles produce electromagnetic showers in the ECAL, but electrons also leave tracks in the silicon tracker, while photons are electrically neutral and do not leave tracks unless they convert into an electron-positron pair in the tracker material.

Photon reconstruction [116] starts from energy clusters in the ECAL. Because photons and electrons may radiate or convert before reaching the ECAL, their energy deposits can spread mainly in the azimuthal direction due to the magnetic field. CMS therefore groups ECAL clusters into superclusters, which collect the energy of the electromagnetic shower and associated radiated photons. Photon candidates are built from superclusters without a matching electron track and are identified using shower-shape and isolation variables.

Electron reconstruction [117] is more involved because electrons can lose a significant fraction of their energy through bremsstrahlung while traversing the tracker material. Electron candidates are reconstructed by matching ECAL superclusters to tracks in the silicon tracker. Two complementary seeding approaches are used: ECAL-based seeding, which starts from an ECAL supercluster and searches for compatible tracker hits, and tracker-based seeding, which starts from reconstructed tracks and searches for compatible ECAL clusters.

Because bremsstrahlung leads to non-Gaussian energy losses, electron tracks are fitted using the Gaussian Sum Filter (GSF) [118]. The final electron candidate is formed by associating a GSF track with an ECAL supercluster. Electron identification then uses variables describing the track-cluster matching, ECAL shower shape, track quality, and isolation. Precise electron reconstruction is essential for the dielectron channel of the Drell–Yan measurement.

5.8.3. Muon reconstruction

Muon reconstruction [119] combines information from the silicon tracker and the muon system. Several complementary muon reconstruction types are used in CMS, providing high efficiency across a broad momentum range and detector acceptance.

Standalone muons are reconstructed using only hits and segments in the muon system. This approach provides muon reconstruction outside the tracker acceptance, but the momentum resolution is limited compared with methods that use the inner tracker. Tracker muons are reconstructed by extrapolating tracks from the silicon tracker to the muon system and searching for compatible muon segments. This method is particularly efficient for low-p_{T} muons that may not leave enough hits to form a standalone muon track.

Global muons are reconstructed by matching a standalone muon track to a compatible tracker track and refitting the combined set of hits. This provides the best momentum resolution for many analysis applications, especially for high-p_{T} muons. Additional identification requirements, based on the quality of the tracker and muon-system tracks, the number of hits, and compatibility with the primary vertex, are used to suppress backgrounds from hadronic punch-through, cosmic rays, and misreconstructed tracks.

The Drell–Yan measurement in the dimuon channel relies on high-quality reconstructed muons with well-measured momenta, good identification efficiency, and controlled trigger and reconstruction efficiencies.

5.8.4. Jet reconstruction

Jets are collimated sprays of particles produced by the hadronization of quarks and gluons. In CMS, jets are reconstructed by clustering PF candidates, which allows the tracker, ECAL, and HCAL information to be combined consistently. Muons, electrons, and isolated photons are identified before jet clustering so that reconstructed jets predominantly represent hadronic activity.

Jets are typically clustered using the anti-k_{T} algorithm [120]. The distance between two PF candidates i and j is defined as

$$d_{ij} = \min\left(k_{\mathrm{T},i}^{-2},\ k_{\mathrm{T},j}^{-2}\right)\frac{\Delta R_{ij}^2}{R^2}, \tag{5.6}$$

where $k_{\mathrm{T},i}$ and $k_{\mathrm{T},j}$ are the transverse momenta of the candidates, ΔR_{ij} is their angular separation, and R is the jet radius parameter, commonly chosen as $R = 0.4$ in CMS analyses. The distance to the beam is

$$d_{iB} = k_{\mathrm{T},i}^{-2}. \tag{5.7}$$

The algorithm iteratively combines the pair with the smallest distance or declares a candidate to be a final jet if d_{iB} is the smallest distance. The anti-k_{T} algorithm is infrared and collinear safe and produces nearly conical jets that are robust against soft radiation.

After clustering, jet energy corrections are applied to account for detector response, PU, and residual differences between data and simulation. In simulation, jet energy resolution smearing is also applied to match the resolution observed in data [121]. Jets are not the primary objects in the Drell–Yan signal selection, but they are important for background rejection and calculations and corrections of the missing transverse momentum and particle isolation.

5.8.5. Missing transverse momentum reconstruction

In proton-proton collisions, the total transverse momentum of the initial state is approximately zero. Particles that do not interact with the detector, such as neutrinos, can therefore be inferred from an imbalance in the reconstructed transverse momentum. The missing transverse momentum vector is defined as the negative vector sum of the transverse momenta of all PF candidates in the event:

$$\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}} = -\sum_{i=1}^{N_{\mathrm{PF}}} \vec{p}_{\mathrm{T},i}. \tag{5.8}$$

Pileup can produce spurious missing transverse momentum through additional soft particles, detector resolution effects, and vertex misassignment. CMS therefore applies PU mitigation and event-cleaning procedures to maintain stable $p_{\mathrm{T}}^{\mathrm{miss}}$ performance. Missing transverse momentum is particularly important for analyses involving neutrinos or invisible particles. In the Drell–Yan measurement, it is also useful for estimating the non-prompt or misidentified lepton backgrounds.

5.8.6. Particle isolation

Particle isolation quantifies how much additional activity surrounds a reconstructed lepton or photon. Prompt leptons from electroweak boson decays are often isolated, while non-prompt leptons produced inside jets, or particles misidentified as leptons, are usually accompanied by nearby hadronic activity. Isolation is therefore an important discriminant in Drell–Yan event selection and background suppression.

A relative isolation variable can be defined as

$$I_i^{\text{rel}} = \frac{1}{p_{\text{T}}^i} \sum_{j \neq i,\, \Delta R_{ij} < R} p_{\text{T}}^j, \tag{5.9}$$

where the sum runs over reconstructed particles within a cone of radius R around the candidate direction. Different isolation definitions can be constructed depending on which particles or detector signals are included in the sum. Track isolation uses charged-particle tracks, calorimeter isolation uses energy deposits in the ECAL and HCAL, and PF isolation uses reconstructed PF candidates.

Pileup contributes additional particles inside the isolation cone and must be corrected for. A commonly used PF relative isolation for leptons is

$$I_{\text{PF}}^{\text{rel}} = \frac{1}{p_{\text{T}}^{\ell}} \left[\sum_{\text{chg.}} p_{\text{T}} + \max\left(0, \sum_{\text{ntr. had.}} p_{\text{T}} + \sum_{\gamma} p_{\text{T}} - \Delta\beta \sum_{\text{chg., PU}} p_{\text{T}} \right) \right], \tag{5.10}$$

where “chg.” stands for “charged”, and “ntr. had.” stands for “neutral hadrons”. All sums are evaluated inside the isolation cone. The term proportional to $\Delta\beta$ estimates the neutral PU contribution from the charged-PU activity, with $\Delta\beta$ commonly taken to be 0.5.

Isolation variables are used in electron, muon, and photon identification and are especially important for suppressing backgrounds with non-prompt or misidentified leptons. They will therefore be used again in the Drell–Yan event selection and background-estimation discussion.

This section presents the measurement of the Drell–Yan (DY) differential cross section as a function of the dilepton invariant mass, using proton-proton collision data collected by the CMS experiment during 2016–2018 at a center-of-mass energy of $\sqrt{s} = 13$ TeV. The analyzed dataset corresponds to an integrated luminosity of 137.6 fb^{-1} and was processed using the CMS Ultra Legacy reconstruction.

The analysis was carried out within the CMS SMP-V Drell–Yan working group, to which the author contributed. The author's primary contribution was the development, implementation, validation, and uncertainty evaluation of the background-estimation methods described in Sec. 10. Other components of the analysis, including detector and physics corrections, unfolding, channel combination, and final cross-section extraction, were developed within the working group. They are summarized in this section to provide the full context of the measurement.

The most detailed description of this analysis is documented in CMS analysis note AN-22-177 [B1], an internal CMS document containing contributions from all members of the working group. The corresponding journal publication is in preparation at the time of writing.

6. Analysis strategy

The measurement targets the neutral-current DY process,

$$pp \to Z/\gamma^* \to \ell^+\ell^-,$$

where ℓ denotes an electron or a muon. The measured observable is the differential cross section as a function of the dilepton invariant mass,

$$\frac{\mathrm{d}\sigma}{\mathrm{d}M_{\ell\ell}}.$$

The analysis uses events containing two isolated, oppositely charged leptons and is performed independently in the dielectron (ee) and dimuon ($\mu\mu$) channels.

The data are divided into four data-taking periods: 2016 pre-APV, 2016 post-APV, 2017, and 2018. These periods are treated separately to account for changes in detector conditions, trigger menus, reconstruction algorithms, and calibration constants. The results are later combined across eras and between the two lepton channels.

The primary result is measured in a fiducial phase space defined at generator level. This fiducial region is chosen to closely match the detector acceptance

while minimizing model dependence from extrapolation outside the experimentally accessible phase space. Generator-level leptons are defined in their dressed form by adding the four-momenta of photons within $\Delta R(\ell, \gamma) < 0.1$ (see eq. (5.3)) to the corresponding lepton. The common fiducial requirements for both channels are:

- $p_{\mathrm{T}}^{\mathrm{lead}} > 20$ GeV and $p_{\mathrm{T}}^{\mathrm{sub}} > 15$ GeV, where $p_{\mathrm{T}}^{\mathrm{lead}}$ and $p_{\mathrm{T}}^{\mathrm{sub}}$ are the transverse momenta of the leading and subleading dressed leptons;
- $|\eta_\ell| < 2.4$ for both dressed leptons;
- oppositely charged lepton pair.

As will be explained in Sec. 9, the reconstructed event selections differ between the electron and muon channels because of trigger and object-reconstruction requirements. In particular, the reconstructed electron selection uses tighter transverse-momentum thresholds than the fiducial definition. The unfolding procedure corrects for these selection and detector effects and extrapolates the reconstructed spectra to the common fiducial phase space.

The invariant mass range of the measurement is

$$40 < M_{\ell\ell} < 3000 \text{ GeV}. \tag{6.1}$$

The mass binning follows the previous CMS DY $\mathrm{d}\sigma/\mathrm{d}M$ measurement at 13 TeV [19], with refinements where the larger Run–2 dataset provides sufficient statistical precision. The bin edges are:

40, 45, 50, 55, 60, 64, 68, 72, 76, 81, 86, 91, 96, 101, 106, 110, 115, 120, 126, 133, 141, 150, 160, 171, 185, 200, 220, 243, 273, 320, 380, 440, 510, 600, 700, 830, 1000, 1500, 2000, 3000 GeV.

Signal simulation is used to model the detector response, derive correction factors, build the unfolding response matrices, and evaluate systematic uncertainties. Background contributions are estimated using a combination of simulation and data-driven methods. Most prompt-lepton backgrounds are taken from simulation, while top-quark backgrounds are modeled using simulation with a data-driven correction derived from the $e\mu$ control region. Backgrounds with misidentified or non-prompt leptons are estimated using same-sign data control regions, as described in Sec. 10.

After background subtraction, the measured reconstructed mass distributions are unfolded to correct for detector efficiency, acceptance within the fiducial region, and bin-to-bin migration due to finite resolution. The unfolding procedure is described in Sec. 11. The electron and muon channel results, together with

Table 2: Summary of the data samples used in the analysis.

Era	Integrated luminosity [fb^{-1}]	Trigger streams	
		Muon channel	Electron channel
2016, pre-APV	19.5	`SingleMuon`, `DoubleMuon`	`DoubleEG`
2016, post-APV	16.8	`SingleMuon`, `DoubleMuon`	`DoubleEG`
2017	41.5	`SingleMuon`, `DoubleMuon`	`DoubleEG`
2018	59.8	`SingleMuon`, `DoubleMuon`	`EGamma`

their covariance matrices, are then combined using the procedure described in Sec. 13.

In addition to the fiducial cross section, a result extrapolated to the full phase space is also produced using acceptance corrections derived from simulation. Since this extrapolation depends on the simulated acceptance outside the fiducial region, the fiducial cross section is the less model-dependent result. The final fiducial and full phase-space results are presented in Sec. 14 and compared with theoretical predictions.

7. Data and simulated samples

This analysis uses the full Run–2 proton-proton collision dataset recorded by CMS in 2016–2018 at $\sqrt{s} = 13$ TeV and processed with the Ultra Legacy reconstruction. The 2016 dataset is divided into pre-APV and post-APV periods, corresponding to different operating conditions of the silicon strip tracker readout. These periods are processed and analyzed separately.

For the muon channel, events are selected from data streams based on single-muon and double-muon triggers. To avoid double-counting events, events satisfying the single-muon trigger requirements are taken from the `SingleMuon` stream, while the `DoubleMuon` stream contributes only events that do not satisfy the single-muon trigger selection. In 2018, data samples including an improved alignment for high-p_{T} muons are used.

For the electron channel, the `DoubleEG` data stream is used in 2016 and 2017. In 2018, CMS combined the single- and double-electron streams into the `EGamma` dataset, which is used for that year. The integrated luminosities correspond to the standard CMS certified datasets for each data-taking period. A summary of the data samples is given in Table 2.

Simulated Monte Carlo (MC) samples are used to model the DY signal, estimate several background contributions, derive detector and physics corrections,

construct the unfolding response matrices, and evaluate systematic uncertainties. All MC samples are produced centrally by the CMS collaboration in the Run–2 Ultra Legacy configuration and are analyzed in the CMS NanoAOD format.

The signal samples, $Z/\gamma^* \to ee$ and $Z/\gamma^* \to \mu\mu$, are generated with POWHEG [55, 56] using the MiNNLO$_{PS}$ method, which provides NNLO QCD accuracy matched to a parton shower [68]. Parton showering, hadronization, and initial-state photon radiation are modeled with PYTHIA8 [48] using a modified CP5 tune [122], while QED final-state radiation is simulated with PHOTOS [123]. The NNPDF 3.1 PDF set is used [124]. The detector simulation in this and all the following MC datasets was performed with GEANT4 [125].

To improve statistical precision across the wide dilepton mass range, additional DY signal samples are generated in bins of dilepton mass below 50 GeV and above 100 GeV. An alternative DY sample generated with MADGRAPH5_aMC@NLO interfaced to PYTHIA8 is used to assess modeling uncertainties.

The $Z/\gamma^* \to \tau^+\tau^-$ process is treated as a background. For this sample, only events in which at least one τ lepton decays to an electron or muon are kept, and the corresponding branching fraction is applied in the normalization. Other simulated background samples include top-quark production ($t\bar{t}$, tW, and $\bar{t}W$), diboson production (WW, WZ, and ZZ), and photon-induced dilepton production ($\gamma\gamma \to \ell\ell$).

The $t\bar{t}$ sample is generated with POWHEG and normalized to the inclusive NNLO cross section recommended by the LHC Top Working Group, with the relevant leptonic branching fractions applied. Single-top tW and $\bar{t}W$ samples are also generated with POWHEG. The simulated top-quark backgrounds are later corrected using a data-driven reweighting derived from the $e\mu$ control region, as described in Sec. 10.

Diboson samples are generated with POWHEG or MADGRAPH5_aMC@NLO and normalized to the corresponding generator cross sections. Photon-induced dilepton production is simulated using the LPAIR formalism [126, 127] as implemented in CEPGEN [128], interfaced to PYTHIA6 for the inelastic components. Fully elastic, single-dissociative, and fully inelastic configurations are considered.

Backgrounds with misidentified or non-prompt leptons are estimated from data. A W+jets MC sample, generated with MADGRAPH5_aMC@NLO interfaced to PYTHIA8, is used only in the construction and validation of these data-driven methods.

The simulated signal and background samples used in the analysis are summarized in Tables 3 and 4.

Table 3: MC samples used as signal in the analysis and corresponding cross sections (σ) in pb.

Process	Mass range [$M_{\ell\ell}$ in GeV]	σ [pb]
$Z/\gamma^* \to \mu\mu$, $Z/\gamma^* \to ee$	10–50	7012.53
	> 50 (inclusive)	2006.65
	100–200	77.95
	200–400	2.78
	400–500	0.15
	500–700	0.084
	700–800	0.013
	800–1000	0.011
	1000–1500	0.0060
	1500–2000	0.00081
	> 2000	0.00020

Table 4: MC samples used as background in the analysis and corresponding cross sections (σ) in pb.

Process	Description	σ [pb]
$Z/\gamma^* \to \tau\tau$	$M_{\tau\tau} > 50$ GeV, at least one τ decays to e or μ	1164.31
$t\bar{t} \to \ell\ell\nu\nu$	Dileptonic $t\bar{t}$	88.51
tW	Inclusive decays	39.65
$\bar{t}W$	Inclusive decays	39.65
WW	$WW \to 2\ell 2\nu$	11.09
WZ	$WZ \to 2\ell 2q$	6.431
	$WZ \to 3\ell\nu$	4.664
ZZ	$ZZ \to 2\ell 2q$	3.698
	$ZZ \to 2\ell 2\nu$	0.9738
	$ZZ \to 4\ell$	1.325
$\gamma\gamma \to ee$, elastic	$10 < M_{ee} < 30$ GeV	11.39
	$30 < M_{ee} < 50$ GeV	0.67
	$50 < M_{ee} < 200$ GeV	0.246
	$200 < M_{ee} < 1500$ GeV	5.6×10^{-3}
	$M_{ee} > 1500$ GeV	3.8×10^{-6}
$\gamma\gamma \to ee$, single-dissociative	$10 < M_{ee} < 30$ GeV	10.39
	$30 < M_{ee} < 50$ GeV	0.89
	$50 < M_{ee} < 200$ GeV	0.368

	$200 < M_{ee} < 1500$ GeV	1.1×10^{-2}
	$M_{ee} > 1500$ GeV	8.8×10^{-6}
$\gamma\gamma \to ee$, double-dissociative	$10 < M_{ee} < 30$ GeV	10.48
	$30 < M_{ee} < 50$ GeV	1.16
	$50 < M_{ee} < 200$ GeV	0.563
	$200 < M_{ee} < 1500$ GeV	2.2×10^{-2}
	$M_{ee} > 1500$ GeV	2.0×10^{-5}
$\gamma\gamma \to \mu\mu$, elastic	$10 < M_{\mu\mu} < 30$ GeV	11.39
	$30 < M_{\mu\mu} < 50$ GeV	0.67
	$50 < M_{\mu\mu} < 200$ GeV	0.247
	$200 < M_{\mu\mu} < 1500$ GeV	5.6×10^{-3}
	$M_{\mu\mu} > 1500$ GeV	3.8×10^{-6}
$\gamma\gamma \to \mu\mu$, single-dissociative	$10 < M_{\mu\mu} < 30$ GeV	10.52
	$30 < M_{\mu\mu} < 50$ GeV	0.90
	$50 < M_{\mu\mu} < 200$ GeV	0.369
	$200 < M_{\mu\mu} < 1500$ GeV	1.1×10^{-2}
	$M_{\mu\mu} > 1500$ GeV	8.8×10^{-6}
$\gamma\gamma \to \mu\mu$, double-dissociative	$10 < M_{\mu\mu} < 30$ GeV	10.40
	$30 < M_{\mu\mu} < 50$ GeV	1.12
	$50 < M_{\mu\mu} < 200$ GeV	0.562
	$200 < M_{\mu\mu} < 1500$ GeV	2.2×10^{-2}
	$M_{\mu\mu} > 1500$ GeV	2.0×10^{-5}
W+jets	$W \to \ell\nu$, used in misidentified or non-prompt lepton estimates	61526.7

8. Corrections

Several corrections are applied to simulated events to improve the agreement between data and simulation and to account for known detector and reconstruction effects. Most of these corrections follow the standard CMS Run–2 recommendations provided by the corresponding Physics Object Groups (POGs).

Simulated events are first reweighted to reproduce the PU distribution observed in data. The data pileup profile is obtained using the recommended inelastic proton-proton cross section of 69.2 mb, and each simulated event is assigned a weight based on the ratio of the data and simulation distributions as a function of the true number of interactions.

In 2016 and 2017, timing effects in the ECAL and muon trigger systems caused some Level–1 trigger primitives to be assigned to the previous bunch

crossing, leading to an event loss known as pre-firing. Since this effect is not fully present in simulation, dedicated pre-firing weights are applied to simulated events. The effect is most relevant at large lepton pseudorapidity and high transverse momentum.

Residual differences between data and simulation in lepton reconstruction, identification, isolation, and trigger efficiencies are corrected with scale factors. These scale factors are generally measured with the tag-and-probe method in DY events and are applied as functions of lepton p_{T} and η. Muon identification, isolation, and trigger scale factors are taken from the CMS Muon POG. Electron reconstruction, identification, and trigger scale factors are derived using dedicated studies based on DY events and are applied consistently across data-taking eras.

Accurate lepton momentum and energy calibration is essential because miscalibration can shift events between invariant mass bins. For muons, Rochester momentum corrections [129] are applied to improve the agreement between data and simulation in the reconstructed dimuon mass spectrum. For electrons, dedicated energy scale and resolution corrections are applied to improve the modeling of the dielectron invariant mass distribution, especially around the Z peak.

Jet energy corrections and jet energy resolution smearing are applied to reconstructed jets following the recommendations of the CMS Jets and Missing Energy POG. These corrections affect the b-jet veto, missing transverse momentum, and background estimates. Additional scale factors are applied to correct differences between data and simulation in the efficiency and mistag rate of the DeepJet b-tagging algorithm.

The uncertainties associated with all corrections are propagated to the final cross section as described in Sec. 12.

9. Event selection

This section summarizes the event selection criteria applied to reconstructed objects in the muon and electron channels. The selections are designed to retain high efficiency for DY events over the full invariant-mass range while suppressing backgrounds from top-quark production and from misidentified or non-prompt leptons.

Two additional requirements are used in both channels to improve the signal purity. First, a veto on events containing b-tagged jets is applied to suppress top-quark backgrounds. Before applying this veto, the $t\bar{t}$ contribution becomes sizable above the Z peak and reduces the signal purity to about 65% around

$M_{\ell\ell} = 300$ GeV. The b-jet veto suppresses the top-quark background by up to about 55% in the most relevant mass region. The corresponding DY signal loss is at most about 4% in the electron channel and 2% in the muon channel, and only about 0.8% near the Z peak. As a result, the minimum signal purity improves to about 80% around $M_{\ell\ell} = 300$ GeV in both channels.

Second, a requirement on the ratio between the dilepton invariant mass and the subleading-lepton transverse momentum is applied. Events with large values of $M_{\ell\ell}/p_{\mathrm{T}}^{\mathrm{sub}}$ were found to be dominated by top-quark backgrounds and by backgrounds with misidentified or non-prompt leptons, while the DY signal contribution is negligible above the chosen thresholds. The requirement is therefore used to reject events with high invariant mass and relatively low subleading-lepton p_{T}, which are not characteristic of the DY signal.

9.1. Muon channel

The event selection in the muon channel begins with the trigger. A logical OR of single-muon and double-muon triggers is used to achieve high efficiency across the full invariant-mass range. Single-muon triggers provide high efficiency near the Z peak and at higher masses, while double-muon triggers improve the acceptance at low masses, where one or both muons can have relatively low p_{T}. The full list of triggers used in each data-taking era is given in Table 5.

Table 5: HLT triggers used in the muon channel.

Era	Trigger name
2016, pre-APV	`HLT_IsoMu24` `HLT_IsoTkMu24` `HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL` `HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL`
2016, post-APV	`HLT_IsoMu24` `HLT_IsoTkMu24` `HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ` `HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ`
2017	`HLT_IsoMu27` `HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass8`
2018	`HLT_IsoMu24` `HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8`

Muon candidates are reconstructed using the Particle Flow (PF) algorithm, described in Sec. 5.8.1. Events are required to contain at least two muons

Table 6: Scheme for the trigger logic and application of trigger efficiency scale factors in the muon channel. Numbers in parentheses are the values used for the 2017 data.

$p_{\mathrm{T}}^{\mathrm{lead}}$ condition	Trigger condition	Action
$\geqslant 26$ (29) GeV	Pass single-μ trigger	Apply single-μ SF
	Pass only double-μ trigger	Apply double-μ SF
< 26 (29) GeV	Pass only single-μ trigger	Reject event
	Pass double-μ trigger	Apply double-μ SF

satisfying the Tight identification (ID) working point (WP) defined by the CMS Muon POG [119]. A key requirement of this WP is that the muon must be reconstructed as a global muon. The muons must also pass the Tight PF isolation requirement, corresponding to a relative PF isolation smaller than 0.15.

Both muons are required to lie within $|\eta_\mu| < 2.4$. The leading and subleading muons must satisfy $p_{\mathrm{T}}^{\mathrm{lead}} \geqslant 20$ GeV and $p_{\mathrm{T}}^{\mathrm{sub}} \geqslant 15$ GeV, respectively. Only oppositely charged muon pairs are used to reconstruct the DY candidate. To suppress backgrounds with misidentified or non-prompt leptons, the additional requirement

$$M_{\mu\mu}/p_{\mathrm{T}}^{\mathrm{sub}} < 10$$

is applied.

All corrections introduced in Sec. 8 are applied at this stage. The trigger efficiency scale factors are applied using a dedicated scheme that exploits the high efficiency of the single-muon triggers on their efficiency plateau. The logic is summarized in Table 6. Single-muon trigger scale factors are applied only when the leading muon lies on the trigger-efficiency plateau, defined as starting 2 GeV above the trigger threshold. Thus, the plateau requirement is $p_{\mathrm{T}}^{\mathrm{lead}} \geqslant 26$ GeV for the `HLT_IsoMu24` triggers used in 2016 and 2018, and $p_{\mathrm{T}}^{\mathrm{lead}} \geqslant 29$ GeV for the `HLT_IsoMu27` trigger used in 2017. Events in which the leading muon fires only the single-muon trigger but lies below the corresponding plateau are rejected to avoid poorly measured trigger-efficiency regions.

Jets are reconstructed from PF candidates using the anti-k_{T} algorithm with a distance parameter of $R = 0.4$ [120]. They must have $p_{\mathrm{T}} \geqslant 50$ GeV and lie within $|\eta| < 2.5$. Pileup contributions to jet momenta are mitigated using charged-hadron subtraction and the recommended jet energy corrections. Since the jet energy resolution is worse in data than in simulation, simulated jets are smeared using the recommended hybrid method, in which reconstructed jets are matched to generator-level jets within $\Delta R < 0.2$.

After jet reconstruction, events containing at least one b-tagged jet are

rejected. The veto uses the DeepJet algorithm at the medium WP [130, 131]. Differences between data and simulation in b-tagging efficiencies and mistag rates are corrected with scale factors recommended by the CMS b-tagging and Vertexing POG.

A summary of the full muon-channel selection is given in Table 7.

Table 7: Summary of event selection for the muon channel.

Category	Condition
Trigger	Single-μ OR double-μ trigger (Table 5)
b jet	No b jet (DeepJet medium WP)
Single muon	$p_{\mathrm{T}}^{\mathrm{lead}} \geqslant 20$ GeV, $p_{\mathrm{T}}^{\mathrm{sub}} \geqslant 15$ GeV Low-efficiency single-μ trigger region removed (Table 6) $\lvert\eta_\mu\rvert < 2.4$ Cut-based Tight ID WP Tight PF isolation, $I_{\mathrm{PF}}^{\mathrm{rel}} < 0.15$
Dimuon	Oppositely charged muon pair $M_{\mu\mu}/p_{\mathrm{T}}^{\mathrm{sub}} < 10$

Control plots comparing data and the sum of signal and background MC samples after the event selection are shown in Figs. 16 and 17 for the leading- and subleading-muon kinematic distributions. The plots are shown for 2018 data as a representative example; similar agreement is observed in the other data-taking eras. The dimuon invariant mass distributions, which constitute the main observable of the analysis, are shown for all eras in Fig. 18. The data-driven background estimation procedures described in Sec. 10 have not yet been applied in these plots. Nevertheless, good agreement between data and simulation is observed across the full mass range, with differences typically below the 5% level except in the most extreme regions of phase space.

9.2. Electron channel

The event selection in the electron channel follows the same general strategy as the muon-channel selection, with differences arising from the electron trigger, reconstruction, identification, and fiducial acceptance requirements.

The electron channel uses double-electron triggers. In 2016, the analysis relies on the unprescaled `HLT_Ele23_Ele12_CaloIdL_TrkIdL_IsoVL_DZ` trigger. The `DZ` requirement imposes a selection on the difference between the longitudinal impact parameters of the two electrons, helping ensure that the two trigger objects originate from the same interaction vertex. In 2017 and 2018,

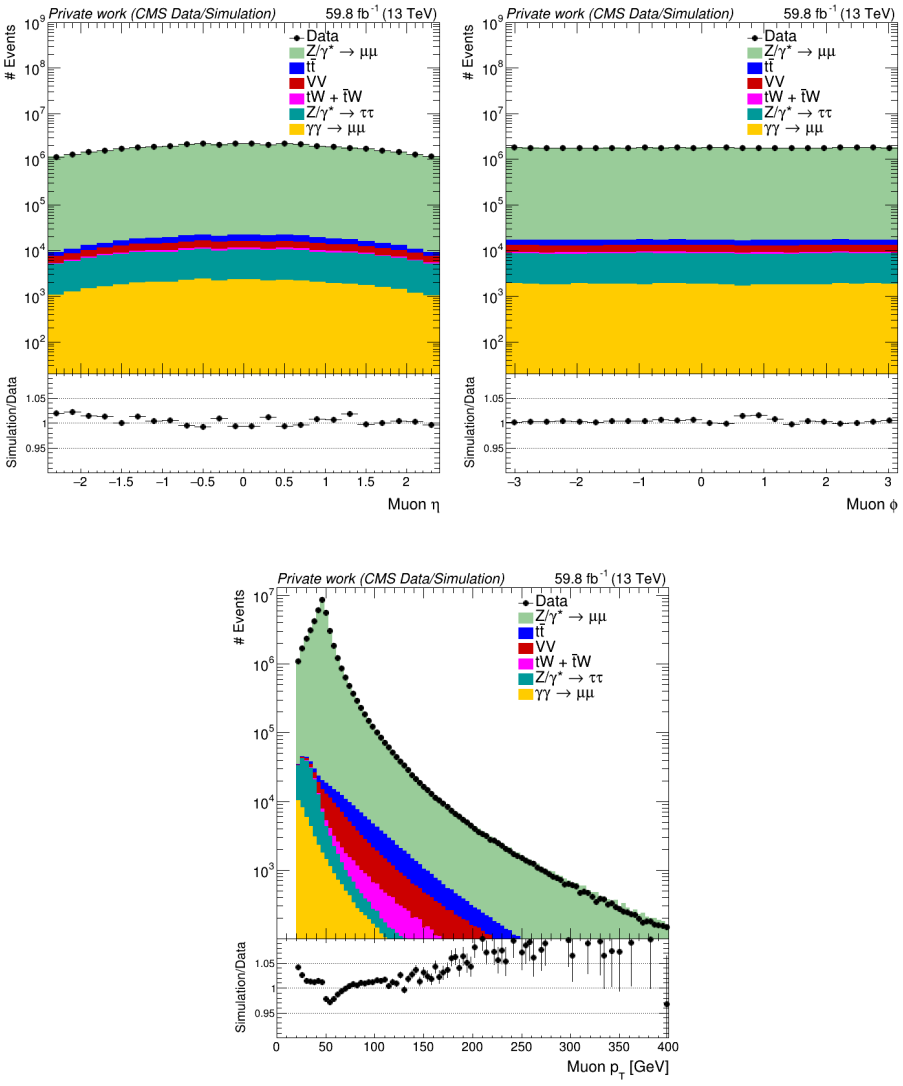


Fig. 16: Leading-muon kinematic distributions for 2018 data and simulation after event selection and corrections.

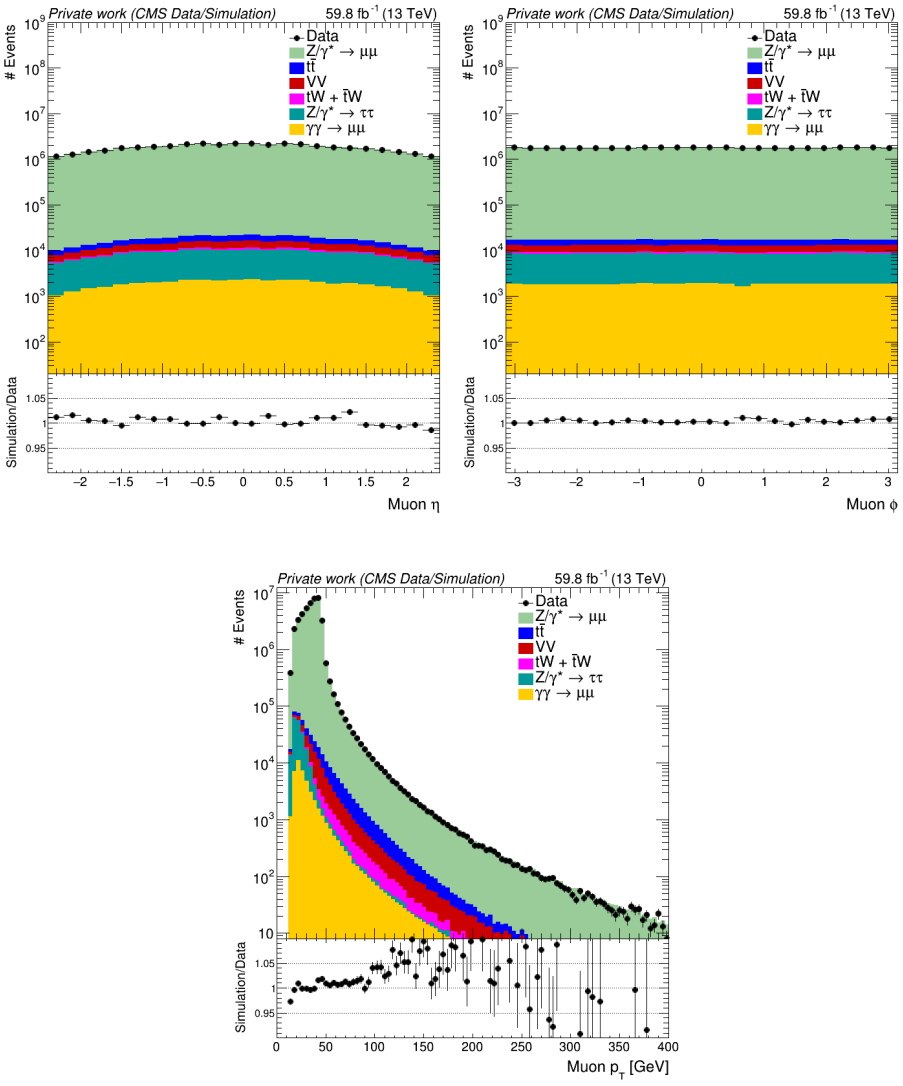


Fig. 17: Subleading-muon kinematic distributions for 2018 data and simulation after event selection and corrections.

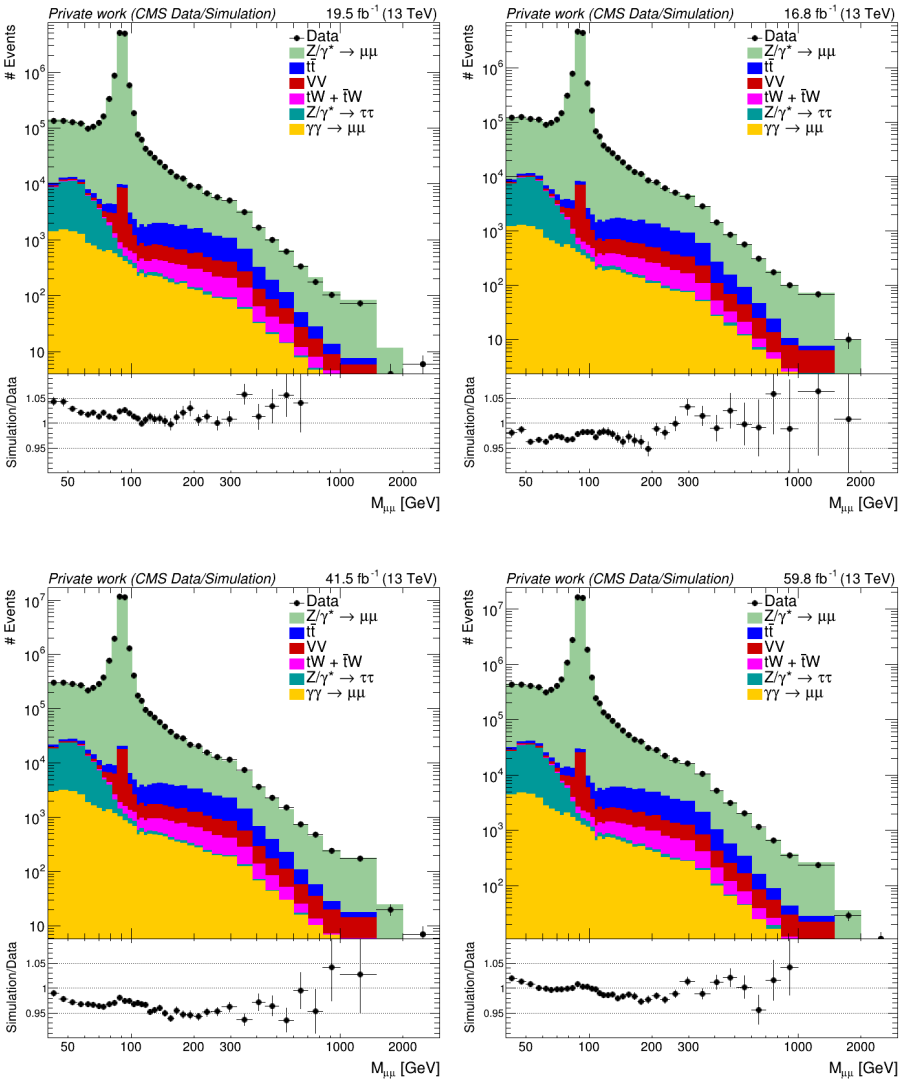


Fig. 18: Dimuon invariant mass distributions for all eras after event selection and corrections.

the same trigger without the DZ requirement remained unprescaled and is used instead. The trigger paths used in each era are listed in Table 8.

Table 8: HLT triggers used in the electron channel.

Era	Trigger name
2016	HLT_Ele23_Ele12_CaloIdL_TrkIdL_IsoVL_DZ
2017, 2018	HLT_Ele23_Ele12_CaloIdL_TrkIdL_IsoVL

Electrons are reconstructed using the PF algorithm. To ensure a well-modeled detector response, only electrons with ECAL supercluster pseudorapidity $|\eta_{\text{SC}}| < 2.5$ are considered. Electrons in the ECAL barrel-endcap transition region,

$$1.4442 < |\eta_{\text{SC}}| < 1.566,$$

are excluded because the detector response is less reliable there. The leading and subleading electrons are required to satisfy $p_{\text{T}}^{\text{lead}} \geqslant 28$ GeV and $p_{\text{T}}^{\text{sub}} \geqslant 20$ GeV, respectively. These thresholds are higher than those used in the fiducial definition because of the double-electron trigger requirements. The unfolding procedure corrects the reconstructed distribution to the common fiducial phase space defined in Sec. 6.

Electron identification is performed using the cut-based Medium ID WP defined by the CMS Electron and Photon POG. This selection includes requirements on shower shape, track-cluster matching, hadronic leakage, and isolation. No separate isolation requirement is applied beyond the isolation included in the Medium ID WP. Only oppositely charged electron pairs are used to reconstruct the DY candidate.

As in the muon channel, backgrounds with misidentified or non-prompt leptons are further suppressed by requiring

$$M_{ee}/p_{\text{T}}^{\text{sub}} < 8.$$

Top-quark backgrounds are suppressed using the same b-jet veto, jet reconstruction, calibration, smearing, and b-tagging procedure as described in Sec. 9.1. A summary of the electron-channel selection is given in Table 9.

All corrections described in Sec. 8 are applied in the electron channel. Control plots comparing data with the stacked simulation for the leading- and subleading-electron kinematic distributions are shown in Figs. 19 and 20, respectively, using 2018 data as a representative example. The dielectron invariant mass distributions for all eras are shown in Fig. 21. As in the muon channel, data-driven background estimates described in Sec. 10 are not included in these

Table 9: Summary of event selection for the electron channel.

Category	Condition
Trigger	Double-electron trigger (Table 8)
b jet	No b jet (DeepJet medium WP)
Single electron	$p_{\mathrm{T}}^{\mathrm{lead}} \geqslant 28$ GeV, $p_{\mathrm{T}}^{\mathrm{sub}} \geqslant 20$ GeV $\lvert\eta_{\mathrm{SC}}\rvert \leqslant 1.4442$ or $1.566 \leqslant \lvert\eta_{\mathrm{SC}}\rvert < 2.5$ Cut-based Medium ID WP
Dielectron	Oppositely charged electron pair $M_{ee}/p_{\mathrm{T}}^{\mathrm{sub}} < 8$

plots. Nevertheless, reasonably good agreement between data and simulation is observed across the full mass range, with differences typically below the 10% level except in the most extreme regions of phase space.

10. Background estimation

The development, implementation, validation, and uncertainty evaluation of the background-estimation methods described in this section were the author's primary contribution to the present Drell–Yan differential cross section measurement.

The background processes considered in this analysis can be grouped into two broad categories. The first category contains processes with prompt leptons in the final state, including leptonic decays of top-antitop quark pairs ($t\bar{t}$), associated production of a single top quark and a W boson (tW and $\bar{t}W$), diboson production (WW, WZ, and ZZ), DY $\rightarrow \tau^+\tau^-$ production, and photon-induced dilepton production ($\gamma\gamma \rightarrow \ell\ell$). These backgrounds are modeled primarily using MC simulation. The top-quark backgrounds are additionally corrected using a data-driven reweighting derived from the $e\mu$ control region.

The second category contains backgrounds with misidentified or non-prompt leptons. These arise mainly from QCD multijet and W+jets events, where at least one reconstructed lepton originates from a hadron decay inside a jet, a heavy-flavor decay, or a jet misidentified as a lepton. In the following, these contributions are referred to as misidentified/non-prompt lepton backgrounds. They are estimated from data using same-sign (SS) dilepton control regions, where the two leptons have the same charge signs instead of the opposite ones.

The same overall strategy is used in the muon and electron channels, with channel-specific modifications. In the muon channel, the main complication is the mixture of QCD multijet and W+jets events in the SS control region,

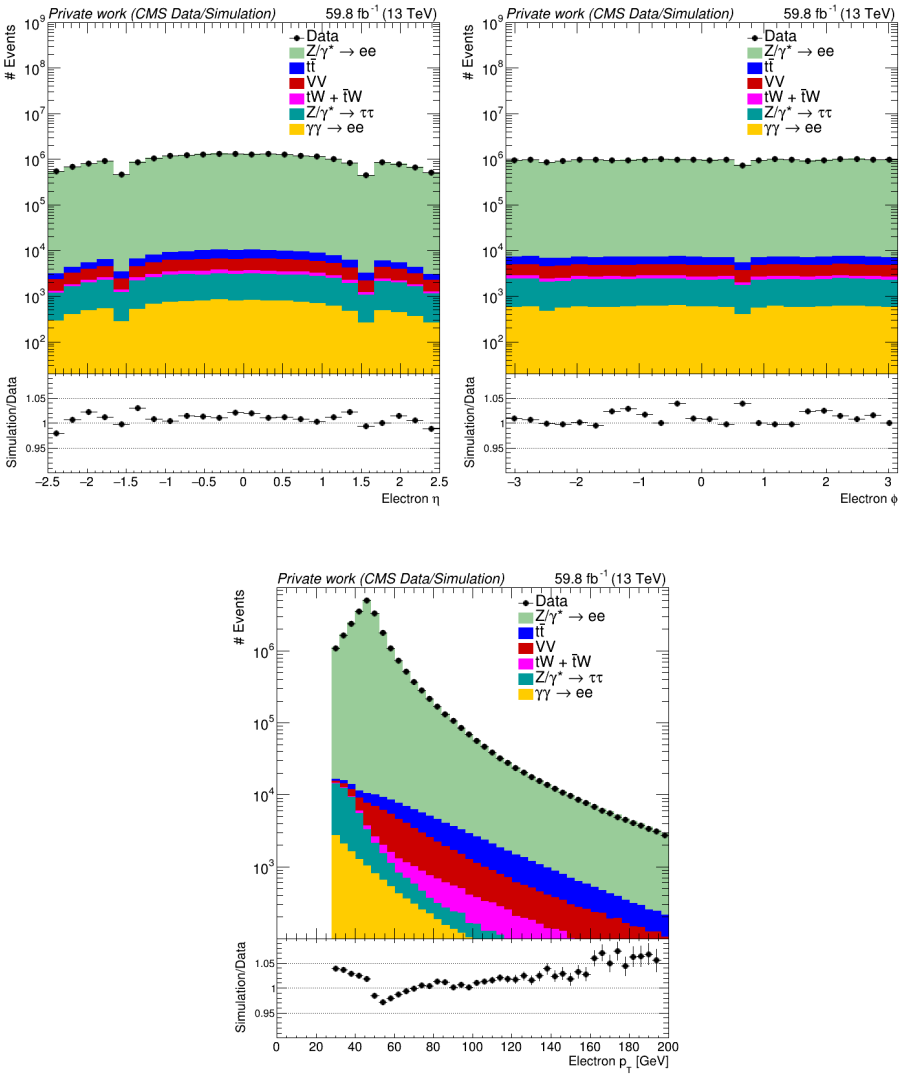


Fig. 19: Leading-electron kinematic distributions for 2018 data and simulation after event selection and corrections.

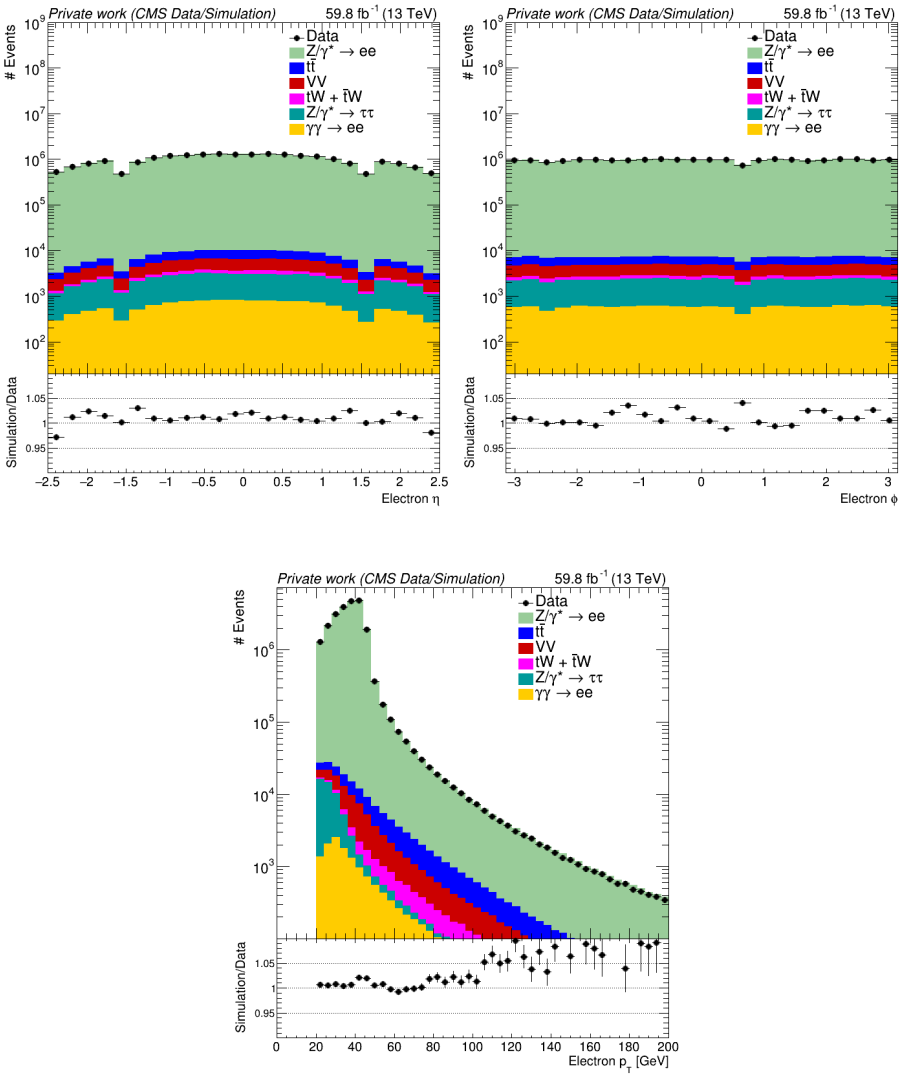


Fig. 20: Subleading-electron kinematic distributions for 2018 data and simulation after event selection and corrections.

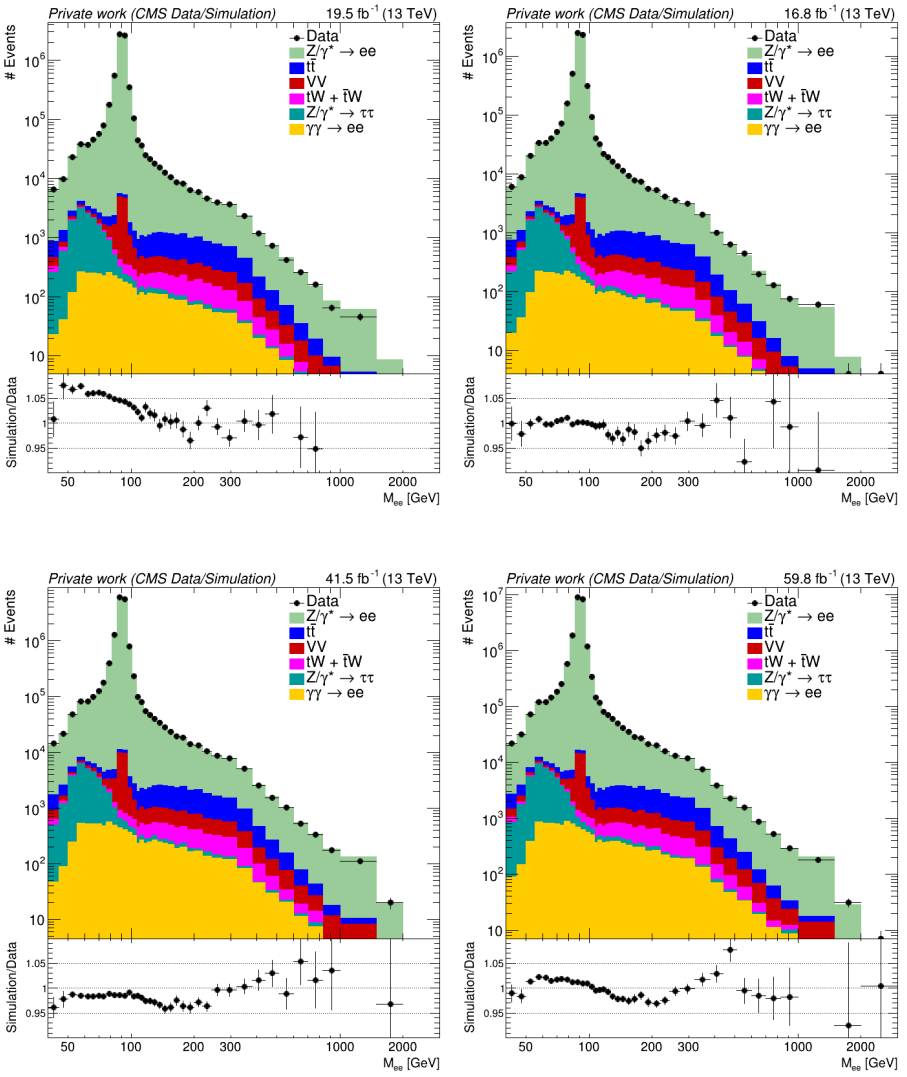


Fig. 21: Dielectron invariant mass distributions for all eras after event selection and corrections.

which requires a kinematics-dependent OS/SS transfer factor. In the electron channel, the higher probability of electron charge misidentification must also be accounted for before the SS region can be used to estimate the misidentified/non-prompt contribution.

A similar same-sign approach was used in the CMS DY $\mathrm{d}^2\sigma/\mathrm{d}M\,\mathrm{d}p_{\mathrm{T}}$ measurement [A1], to which the author contributed through discussions of the background-estimation strategy and review of the paper draft. In the present analysis, the method is extended and refined, and it is applied consistently in both the dimuon and dielectron channels.

10.1. Muon channel

In the dimuon channel, the dominant backgrounds at high dimuon invariant masses are top-quark and diboson processes with two prompt leptons in the final state. At low dimuon masses, the main prompt-lepton background comes from DY $\to \tau^+\tau^-$ production, where both τ leptons decay into muons. In addition, QCD multijet and W+jets events contribute through misidentified/non-prompt muons, especially at low masses and low subleading-muon p_{T}.

10.1.1. Misidentified/non-prompt muon background estimation

The misidentified/non-prompt muon background is estimated from a data control region in which all event selection criteria are identical to those of the signal region, except that the two muons are required to have the same electric charge ($\mu^{\pm}\mu^{\pm}$) instead of opposite charge. Prompt DY production, $t\bar{t}$, and diboson processes predominantly produce opposite-sign (OS) lepton pairs and are therefore strongly suppressed in this same-sign region. In contrast, QCD multijet and W+jets events can populate the SS region through non-prompt muons or particles misidentified as muons. The SS data sample is therefore enriched in the background component that needs to be estimated.

The starting point is the number of SS events observed in data, $N^{\mathrm{SS}}_{\mathrm{data}}$. Simulated contributions from prompt-lepton processes are subtracted to obtain the SS misidentified/non-prompt component:

$$N^{\mathrm{SS}}_{\mathrm{misid./non\text{-}prompt}} = N^{\mathrm{SS}}_{\mathrm{data}} - N^{\mathrm{SS}}_{\mathrm{prompt\,MC}}, \tag{10.1}$$

where

$$N^{\mathrm{SS}}_{\mathrm{prompt\,MC}} = N^{\mathrm{SS}}_{\mathrm{DY}\mu\mu\,\mathrm{MC}} + N^{\mathrm{SS}}_{\mathrm{DY}\tau\tau\,\mathrm{MC}} + N^{\mathrm{SS}}_{t\bar{t}\,\mathrm{MC}} + N^{\mathrm{SS}}_{tW\,\mathrm{MC}} + N^{\mathrm{SS}}_{\mathrm{VV\,MC}} + N^{\mathrm{SS}}_{\gamma\gamma\,\mathrm{MC}}. \tag{10.2}$$

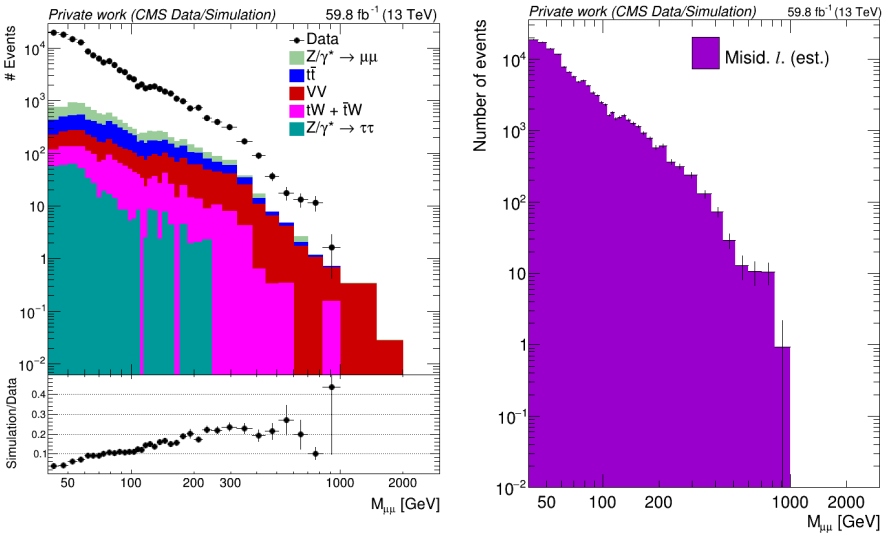


Fig. 22: Left: comparison between data and simulated prompt-lepton backgrounds in the SS dimuon control region for 2018 data. The excess of data over the sum of simulated distributions is attributed to events with misidentified/non-prompt muons. Right: estimated distribution of misidentified/non-prompt muons in the same region, obtained by subtracting the simulated prompt-lepton backgrounds from data.

Here, tW includes both tW and $\bar{t}W$ production, and VV denotes the diboson backgrounds.

The left plot in Fig. 22 shows the comparison between data and the sum of simulated prompt-lepton backgrounds in the SS dimuon control region for 2018 data. The other eras exhibit similar behavior. A sizable excess of data over the prompt-lepton simulation is observed and is interpreted as the SS misidentified/non-prompt muon contribution according to Eq. (10.1). The corresponding distribution is shown in the right plot in Fig. 22. In this and the following figures, the legend label "Misid. ℓ" denotes the misidentified/non-prompt lepton background.

To obtain the misidentified/non-prompt muon contribution in the OS signal region, the SS estimate is multiplied by an OS/SS transfer factor $R_{\text{OS/SS}}$:

$$N^{\text{OS}}_{\text{misid./non-prompt}} = R_{\text{OS/SS}} \left(N^{\text{SS}}_{\text{data}} - N^{\text{SS}}_{\text{prompt MC}} \right). \tag{10.3}$$

In many analyses, $R_{\text{OS/SS}}$ is approximated by a constant value between 1 and 2 [A1, 91, 92, 132, 133]. In the present analysis, however, the SS region contains a mixture of QCD multijet and W+jets events, which have different charge

asymmetries and kinematic distributions. A single constant transfer factor is therefore not sufficient to describe the full dimuon mass spectrum.

The OS/SS transfer factor for W+jets, $R_{\text{OS/SS}}^{W\text{+jets}}$, is measured from simulation and is typically between 1.5 and 2.5. For QCD multijet events, the corresponding factor $R_{\text{OS/SS}}^{\text{QCD}}$ is determined from a control region where both muons fail the isolation requirement, which strongly enriches the sample in QCD events. After subtracting the residual simulated prompt-lepton contribution, $R_{\text{OS/SS}}^{\text{QCD}}$ is found to be consistent with unity within uncertainties in all data-taking eras and across the studied mass range. Since $R_{\text{OS/SS}}^{W\text{+jets}}$ and $R_{\text{OS/SS}}^{\text{QCD}}$ differ significantly, the effective transfer factor in the signal region depends on the relative fractions of W+jets and QCD multijet events, as well as on their kinematics.

To account for this dependence, a variable transfer factor $R_{\text{OS/SS}}(M_{\mu\mu}, p_{\text{T}}^{\text{miss}})$ is used, where $M_{\mu\mu}$ is the dimuon invariant mass and $p_{\text{T}}^{\text{miss}}$ is the missing transverse momentum. The transfer factor is determined in a sideband region chosen to reproduce as closely as possible the mixture of W+jets and QCD multijet events expected in the signal region. Three sideband definitions were tested:

1. both muons fail the PF isolation requirement,
2. exactly one muon fails the PF isolation requirement,
3. the event fails the $M_{\mu\mu}/p_{\text{T}}^{\text{sub}}$ requirement,

with all other event selection requirements following the SS control-region selection. The sideband in which exactly one muon fails the isolation requirement was chosen because it provides the best agreement between the expected W+jets fraction and the total data-driven misidentified/non-prompt estimate.

In this sideband, the OS and SS distributions of $M_{\mu\mu}$ and $p_{\text{T}}^{\text{miss}}$ are compared after subtracting residual prompt-lepton contributions using simulation. Events in the Z-peak region, $61 < M_{\mu\mu} < 121$ GeV, are removed to avoid artifacts from imperfect modeling of DY events with non-isolated leptons. The resulting OS and SS misidentified/non-prompt muon distributions, shown in Fig. 23 for 2018 data, are used to construct the two-dimensional ratio $R_{\text{OS/SS}}(M_{\mu\mu}, p_{\text{T}}^{\text{miss}})$.

The measured two-dimensional transfer factor is parameterized with a smooth empirical function,

$$R_{\text{OS/SS}}(M_{\mu\mu}, p_{\text{T}}^{\text{miss}}) = [a_0 + a_1 \log_{10}(M_{\mu\mu})] \left[a_2 - a_3 p_{\text{T}}^{\text{miss}} \exp\left(-a_4 p_{\text{T}}^{\text{miss}}\right)\right], \tag{10.4}$$

where the parameters a_i are obtained from a fit to the measured sideband ratio. This functional form has no theoretical interpretation; it was chosen because it provided a good description of the observed OS/SS dependence among the

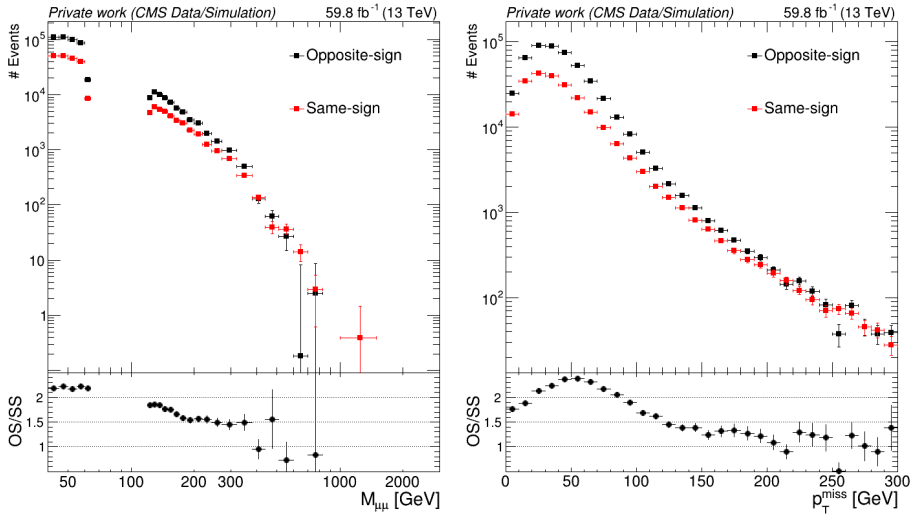


Fig. 23: Comparison of OS and SS dimuon invariant mass (left) and missing transverse momentum $p_{\mathrm{T}}^{\mathrm{miss}}$ (right) distributions for misidentified/non-prompt muons in the data sideband where exactly one muon fails the isolation requirement. The distributions are shown for 2018 data; the other eras show similar behavior.

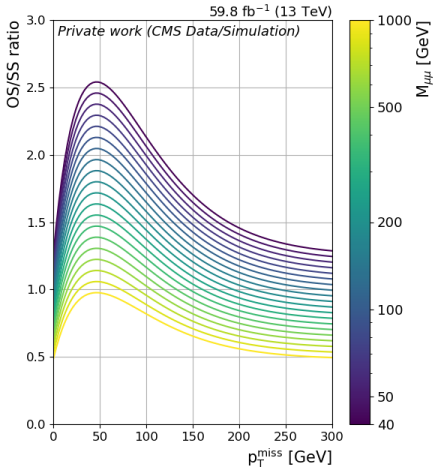


Fig. 24: One-dimensional slices of the fitted OS/SS transfer factor $R_{\mathrm{OS/SS}}(M_{\mu\mu}, p_{\mathrm{T}}^{\mathrm{miss}})$ in the sideband for 2018 data.

tested alternatives. An illustration of the fitted function for 2018 data, shown as one-dimensional slices in $p_{\mathrm{T}}^{\mathrm{miss}}$, is given in Fig. 24. The shapes obtained in the other eras are similar.

The fitted transfer factor $R_{\mathrm{OS/SS}}(M_{\mu\mu}, p_{\mathrm{T}}^{\mathrm{miss}})$ is then applied event-by-event to SS data in the main analysis region to obtain the misidentified/non-prompt muon background prediction in the OS signal region. A closure test is performed in the sideband by applying the same reweighting procedure to SS events and comparing the resulting distribution with the OS data. The dimuon mass and $p_{\mathrm{T}}^{\mathrm{miss}}$ closure tests for 2018 data are shown in Fig. 25. No systematic non-closure is observed in the mass distribution; the remaining differences are compatible with the statistical uncertainty of the closure test. A separate closure uncertainty is therefore not assigned, while the fit-parameter and functional-form uncertainties described below cover the observed level of variation.

In the final estimate, the simulated top-quark contribution subtracted from the SS region is first corrected using the $e\mu$-based reweighting described in Subsection 10.1.2. The corrected prompt-lepton contribution is then subtracted from SS data, and the remaining SS distribution is reweighted with $R_{\mathrm{OS/SS}}(M_{\mu\mu}, p_{\mathrm{T}}^{\mathrm{miss}})$ to obtain the OS misidentified/non-prompt muon prediction. The resulting distributions are shown in Fig. 26 for each data-taking era.

Several sources of systematic uncertainty are considered for this estimate:

1. the statistical uncertainty of the SS data used to extract the misidenti-

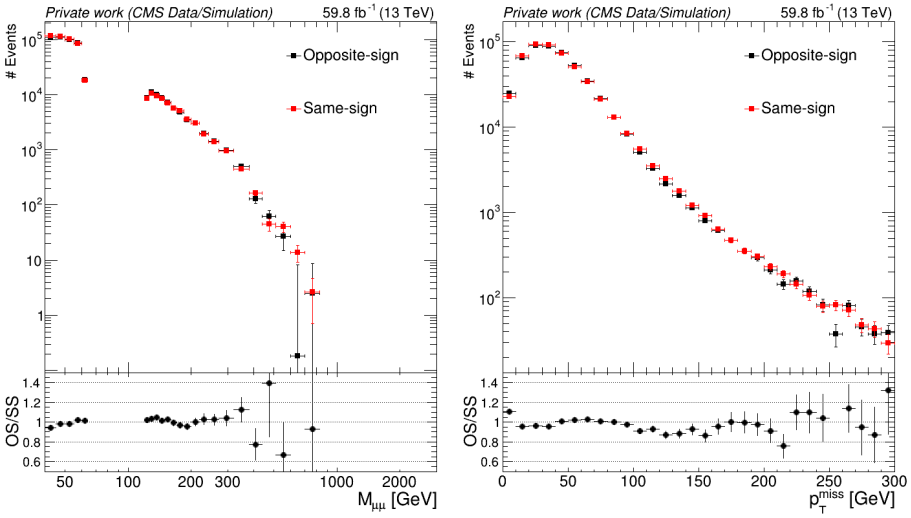


Fig. 25: Closure tests for the misidentified/non-prompt muon background estimate in the data sideband. Shown are the dimuon mass (left) and missing transverse momentum $p_{\mathrm{T}}^{\mathrm{miss}}$ (right) distributions for OS data and reweighted SS data for 2018.

fied/non-prompt component,

2. the uncertainty on the fitted transfer factor $R_{\mathrm{OS/SS}}(M_{\mu\mu}, p_{\mathrm{T}}^{\mathrm{miss}})$, including correlations between fit parameters,

3. the choice of functional form used to parameterize $R_{\mathrm{OS/SS}}(M_{\mu\mu}, p_{\mathrm{T}}^{\mathrm{miss}})$,

4. the uncertainty on the top-quark background correction in the SS region.

For each source, an alternative estimate is produced by modifying the corresponding part of the procedure. The systematic uncertainty associated with that source is taken as the absolute difference between the nominal and alternative estimates,

$$\Delta N_{\mathrm{bkg.}} = \left| N_{\mathrm{bkg.}}^{\mathrm{nominal}} - N_{\mathrm{bkg.}}^{\mathrm{alternative}} \right| . \tag{10.5}$$

When several alternative estimates are available for the same source, the variation producing the largest difference with respect to the nominal result is used.

The impact of the fit-parameter uncertainties is evaluated by coherently varying the parameters according to their covariance matrix. The uncertainty due to the functional form is estimated by comparing the nominal result with

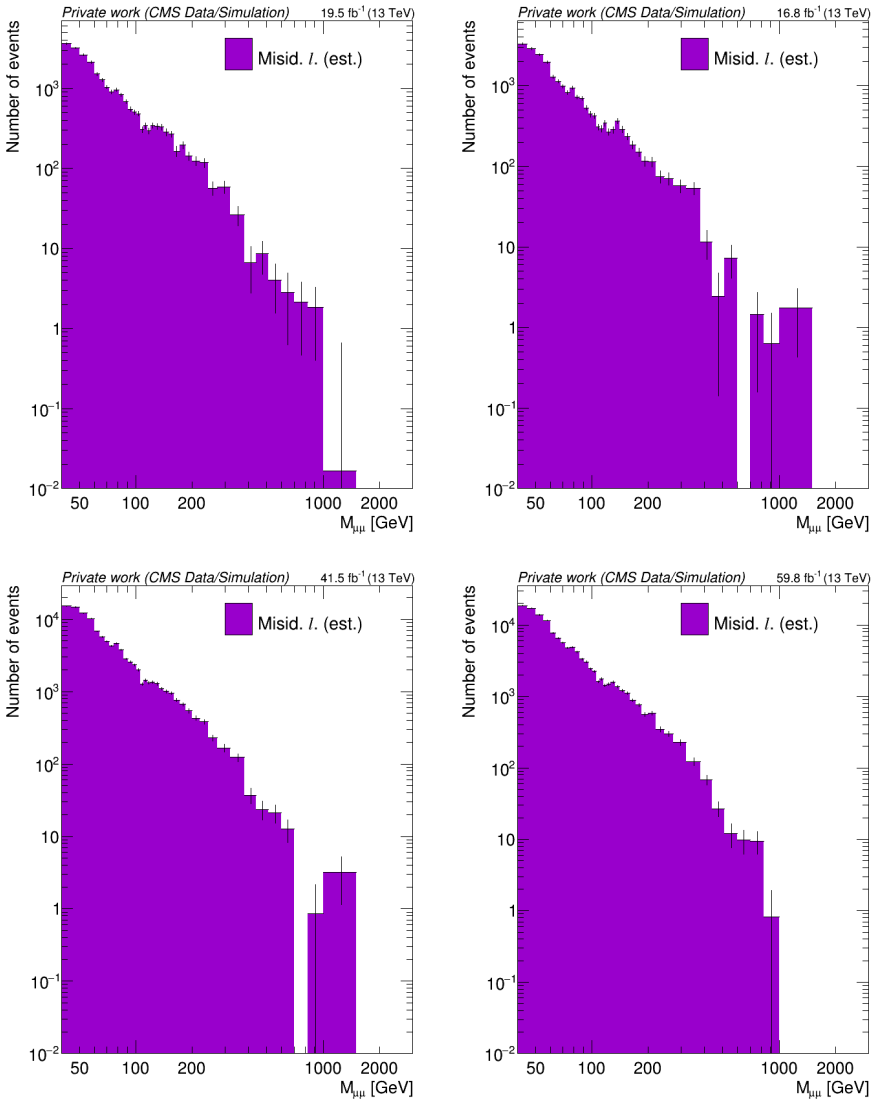


Fig. 26: Estimated distributions of misidentified/non-prompt muon backgrounds in the OS dimuon signal region for each data-taking era, obtained by reweighting the SS data with $R_{\mathrm{OS/SS}}(M_{\mu\mu}, p_{\mathrm{T}}^{\mathrm{miss}})$.

an alternative estimate obtained using a constant transfer factor $R_{\mathrm{OS/SS}} = 1.2$. This value is chosen as a representative alternative motivated by studies in which the sideband was split into regions with different isolation values; those studies indicated that transfer factors closer to 1.2 can be obtained when the sideband has a different isolation composition. The uncertainty from the top-quark correction is evaluated by comparing the nominal result with one obtained without applying the $e\mu$-based reweighting to the top-quark backgrounds. The total systematic uncertainty is obtained by adding all contributions in quadrature.

The resulting relative uncertainties as a function of $M_{\mu\mu}$ are shown in Fig. 27 for each era. They are dominated by the choice of parameterization for $R_{\mathrm{OS/SS}}(M_{\mu\mu}, p_{\mathrm{T}}^{\mathrm{miss}})$ for $M_{\mu\mu} < 250$ GeV, and by the statistical uncertainty at higher masses. The systematic uncertainty remains below 50% for $M_{\mu\mu} < 900$ GeV, while the total (statistical $\oplus$ systematic) uncertainty stays below 50% for $M_{\mu\mu} < 400$ GeV. At very high masses, the uncertainties increase because of limited statistics in the control and signal regions.

When added to the simulated prompt-lepton backgrounds, the estimated misidentified/non-prompt muon contribution improves the agreement between data and prediction in the signal region. Figure 28 shows the comparison between data and the sum of simulated prompt-lepton backgrounds plus the misidentified/non-prompt muon estimate. This background accounts for roughly 2–5% of the total background at low dimuon masses, 1–2% at high masses, and less than 0.05% in the Z-peak region.

10.1.2. Top-quark background correction using the $e\mu$ method

Events with one electron and one muon in the final state provide a useful control region for processes with two prompt leptons, especially top-quark backgrounds. Such events are largely produced by $t\bar{t}$, single-top, diboson, and DY $\to \tau^+\tau^-$ processes, while the DY signal contribution to same-flavor light lepton pairs is absent by construction. This technique is commonly referred to as the $e\mu$ method.

An $e\mu$ control sample is selected with requirements similar to those of the dimuon signal region, but with one electron and one muon instead of two muons. The selection is summarized in Table 10. Since the sample is collected with single-muon triggers, the muon p_{T} threshold is higher than in the dimuon signal region. The stability of the method with respect to this difference is checked later in this subsection.

In the idealized case where simulation perfectly models all relevant backgrounds, the same-flavor yield from processes that also produce $e\mu$ events

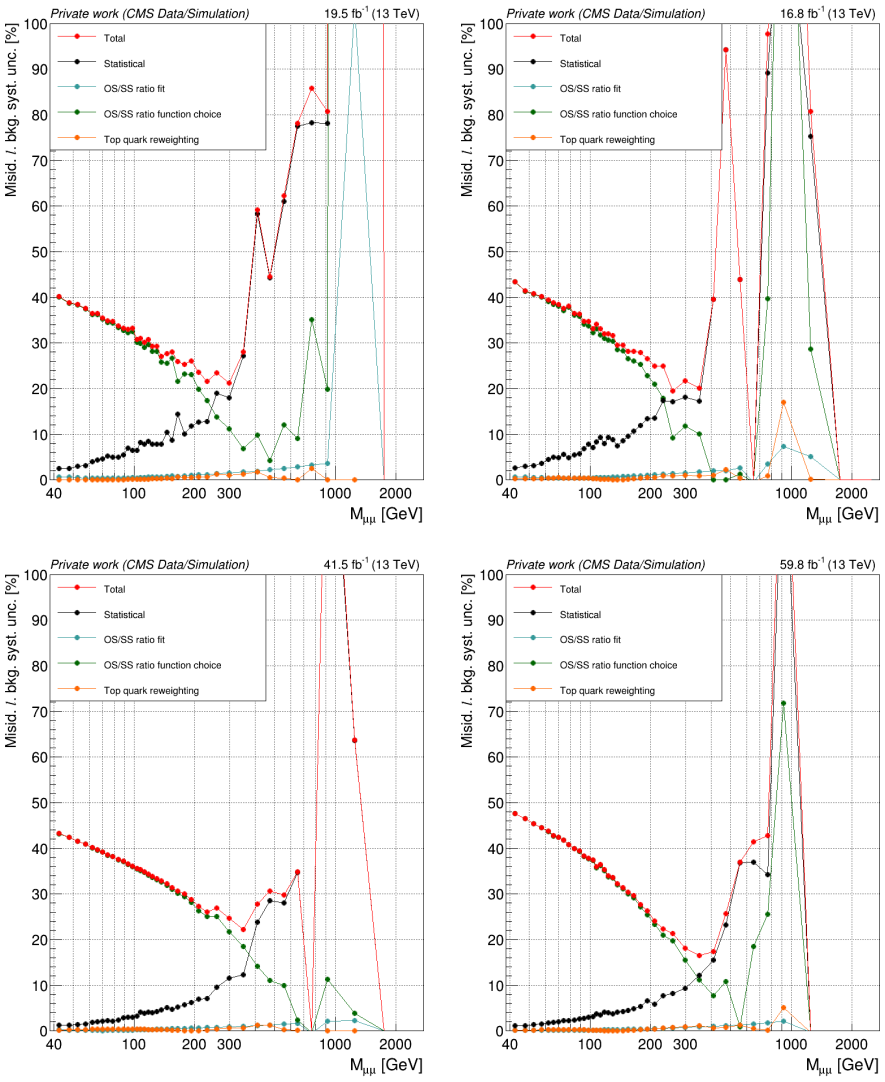


Fig. 27: Relative uncertainties on the misidentified/non-prompt muon background estimate for each data-taking era. Points with 0% uncertainty correspond to empty bins. Very large uncertainties at the highest masses are due to limited statistics in both the control and signal regions.

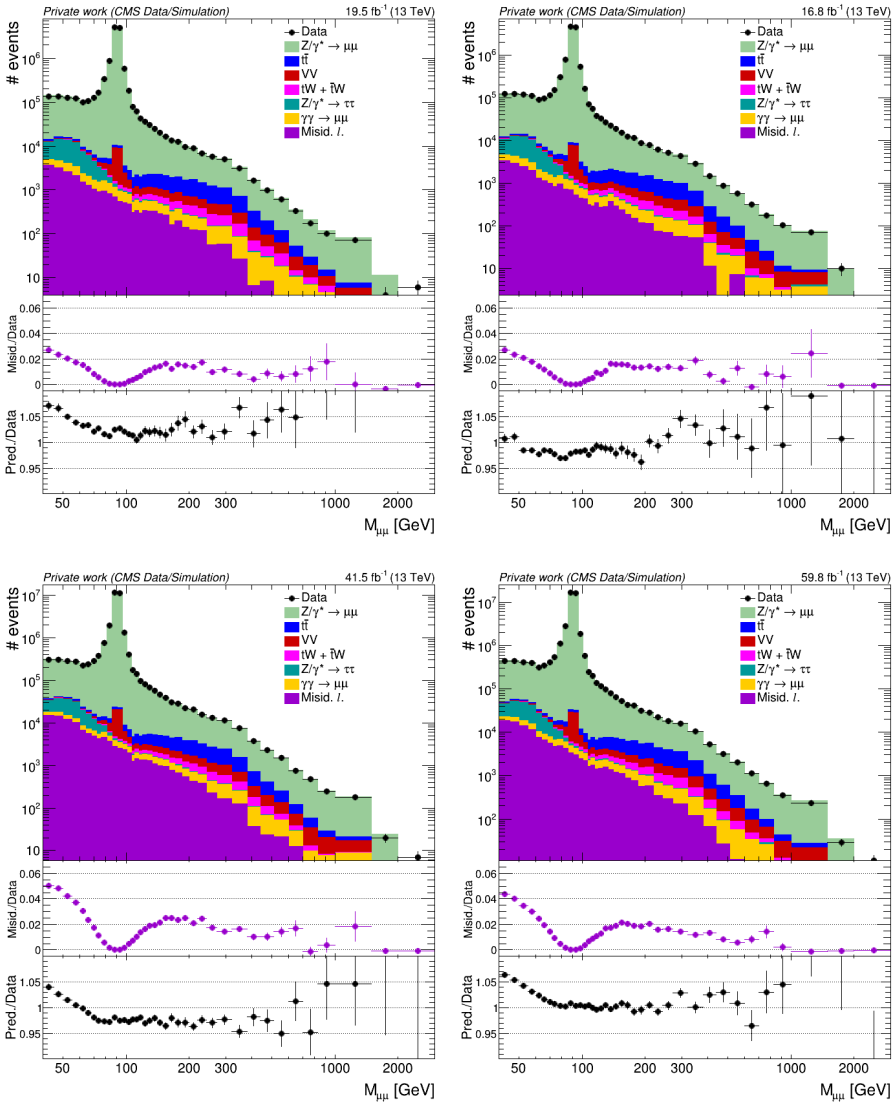


Fig. 28: Comparison between data and the sum of simulated prompt-lepton backgrounds plus the misidentified/non-prompt muon estimate in the OS dimuon signal region for each era. The middle panel shows the fraction of the observed data attributed to misidentified/non-prompt muon backgrounds, while the bottom panel shows the ratio of the total signal plus background yield prediction and the observed data.

Table 10: Summary of event selection for the $e\mu$ control sample.

Category	Condition
Trigger	Single-muon trigger (Table 5)
b jet	No b jet (DeepJet medium WP)
Muon	2016 and 2018: $p_{\mathrm{T}} \geqslant 26$ GeV 2017: $p_{\mathrm{T}} \geqslant 29$ GeV $\lvert\eta_{\mu}\rvert < 2.4$ Cut-based Tight ID WP Tight PF isolation, $I_{\mathrm{PF}}^{\mathrm{rel}} < 0.15$
Electron	$p_{\mathrm{T}} \geqslant 20$ GeV $\lvert\eta_{\mathrm{SC}}\rvert \leqslant 1.4442$ or $1.566 \leqslant \lvert\eta_{\mathrm{SC}}\rvert \leqslant 2.5$ Cut-based Medium ID WP
Dilepton	Opposite-sign $e\mu$ pair $M_{e\mu}/p_{\mathrm{T}}^{\mathrm{sub}} < 10$

could be predicted directly from MC. The $e\mu$ method instead uses the ratio between data and simulation in the $e\mu$ control sample to correct the same-flavor prediction:

$$N_{\mathrm{est}}^{\ell\ell} = N_{\mathrm{MC}}^{\ell\ell} \, \frac{N_{\mathrm{obs}}^{e\mu}}{N_{\mathrm{MC}}^{e\mu}}, \tag{10.6}$$

where $\ell\ell$ denotes either the $\mu^+\mu^-$ or e^+e^- final state. Here, $N_{\mathrm{MC}}^{e\mu}$ and $N_{\mathrm{MC}}^{\ell\ell}$ are the simulated yields of processes contributing to both $e\mu$ and same-flavor final states, while $N_{\mathrm{obs}}^{e\mu}$ is the corresponding yield in data after subtracting backgrounds with misidentified/non-prompt leptons.

The misidentified/non-prompt contribution in the $e\mu$ sample is estimated using the same-sign method, following the same logic as in the dimuon channel. The left plot in Fig. 29 shows the SS $e\mu$ sample used to extract this contribution. After subtracting simulated prompt-lepton backgrounds, the remaining SS data are scaled to the OS region using a constant transfer factor $R_{\mathrm{OS/SS}} = 1.27$. This value is determined in a sideband where the electron fails the cut-based identification requirement and is found to be compatible across all data-taking eras within uncertainties. As shown in the right plot of Fig. 29, the OS/SS ratio in this sideband is approximately flat as a function of $M_{e\mu}$, justifying the use of a constant transfer factor in the $e\mu$ channel.

Figure 30 shows the comparison between data and simulation in the OS $e\mu$ control region, including the data-driven estimate of misidentified/non-prompt leptons. A systematic trend is observed in all eras: the simulation overestimates the data at high $M_{e\mu}$. Similar behavior has been reported in a CMS $t\bar{t}$ differential

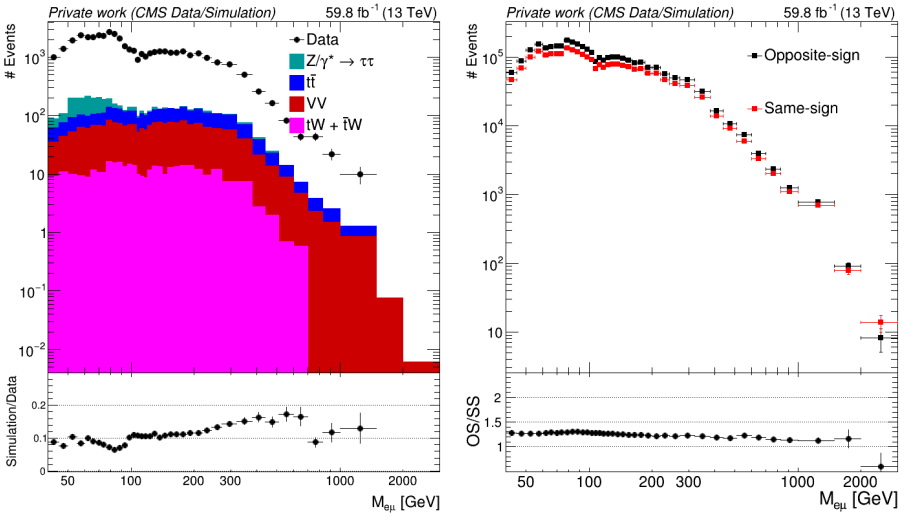


Fig. 29: Left: comparison of data and simulated prompt-lepton backgrounds in the SS $e\mu$ sample. The excess of data is attributed to misidentified/non-prompt leptons. Right: comparison of OS and SS $e\mu$ invariant mass distributions in the sideband where the electron fails the Medium ID requirement. The approximate flatness of the ratio motivates the use of a constant OS/SS transfer factor, $R_{\mathrm{OS/SS}} = 1.27$, in the $e\mu$ channel. Both plots are shown for 2018 data; the other eras exhibit similar behavior.

cross section measurement [134], and it is not observed in processes without top quarks. The $e\mu$ method is therefore used as a data-driven correction to the simulated top-quark backgrounds rather than as a correction to all prompt-lepton backgrounds.

To isolate the top-quark contribution, the method is modified by subtracting from data the simulated non-top backgrounds and the estimated misidentified/non-prompt contribution:

$$N^{\ell\ell}_{t\bar{t}+tW\text{ est}} = N^{\ell\ell}_{t\bar{t}+tW\text{ MC}} \frac{N^{e\mu}_{\text{data}} - N^{e\mu}_{0t\text{ MC}} - N^{e\mu}_{\text{misid./non-prompt est}}}{N^{e\mu}_{t\bar{t}+tW\text{ MC}}}. \tag{10.7}$$

Here, $N^{\ell\ell}_{t\bar{t}+tW\text{ MC}}$ and $N^{e\mu}_{t\bar{t}+tW\text{ MC}}$ are the simulated yields of $t\bar{t}$, tW, and $\bar{t}W$ processes in the same-flavor and $e\mu$ channels, respectively. The term $N^{e\mu}_{0t\text{ MC}}$ denotes simulated non-top prompt-lepton backgrounds,

$$N^{e\mu}_{0t\text{ MC}} = N^{e\mu}_{\text{DY}\tau\tau\text{ MC}} + N^{e\mu}_{\text{VV MC}} + N^{e\mu}_{\gamma\gamma\text{ MC}}. \tag{10.8}$$

The ratio between the modified $e\mu$ method prediction and the nominal simulated top-quark background is shown in Fig. 31. It exhibits a smooth, approximately linear decrease as a function of $\log_{10}(M_{\ell\ell})$. To reduce sensitivity to binning fluctuations, this ratio is parameterized as

$$w(M_{\ell\ell}) = c_0 - c_1 \log_{10}(M_{\ell\ell}), \tag{10.9}$$

where $w(M_{\ell\ell})$ is the event weight applied to simulated top-quark events. In the dimuon channel, this weight is evaluated using $M_{\ell\ell} = M_{\mu\mu}$. The same parameterization is later used in the dielectron channel with $M_{\ell\ell} = M_{ee}$.

The stability of the result with respect to the lepton p_{T} thresholds in the $e\mu$ sample is tested by repeating the procedure with higher muon and/or electron p_{T} requirements. The fitted parameters c_0 and c_1 are found to vary only within their statistical uncertainties, indicating that the reweighting is not sensitive to the low-p_{T} region that is inaccessible because of the trigger thresholds.

Two main sources of systematic uncertainty are considered for the top-quark reweighting:

1. the uncertainty on the fit parameters,

2. the uncertainty on the misidentified/non-prompt lepton contribution in the $e\mu$ sample, propagated through the modified $e\mu$ method.

The fit-parameter uncertainty is evaluated by coherently varying c_0 and c_1 according to their covariance matrix. The uncertainty on the misidentified/non-prompt contribution is estimated by varying the OS/SS transfer factor in the

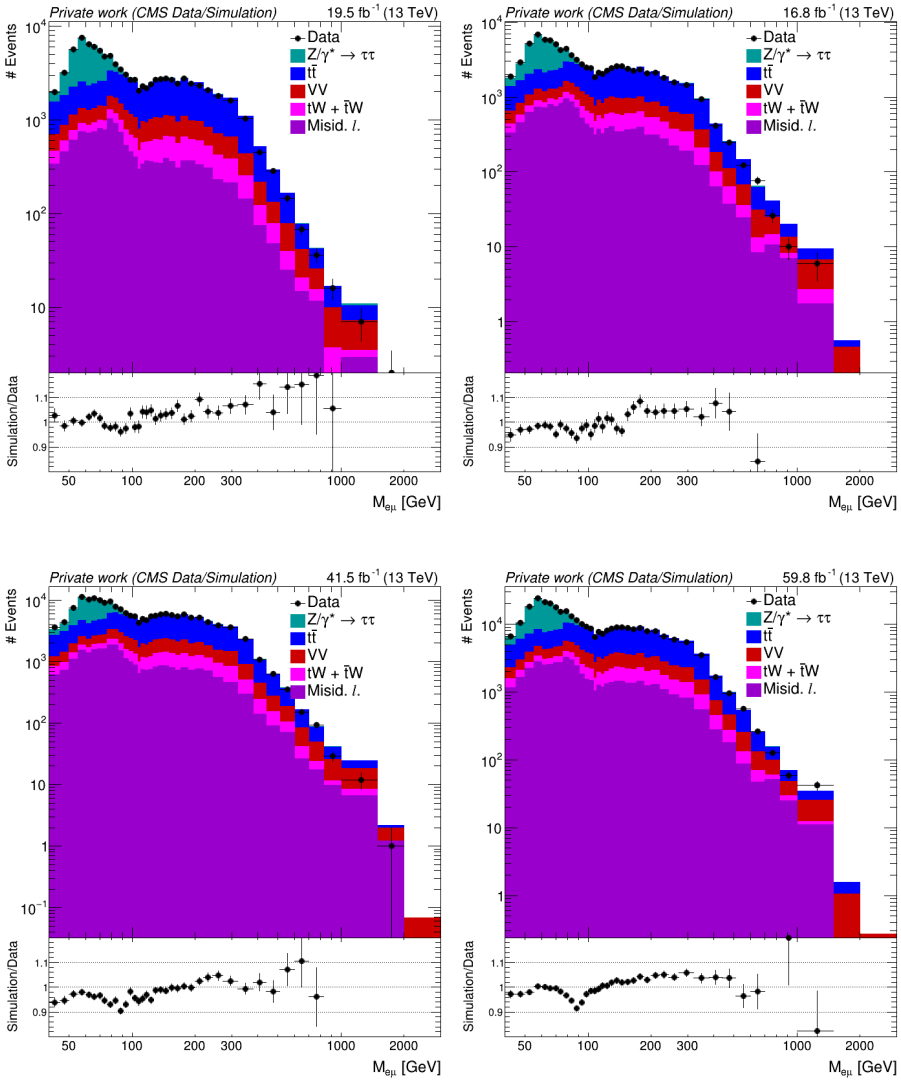


Fig. 30: Comparison between data and the sum of simulated backgrounds plus the misidentified/non-prompt lepton estimate in the OS $e\mu$ control region for each era. A systematic overestimate of the total background at high invariant mass is visible.

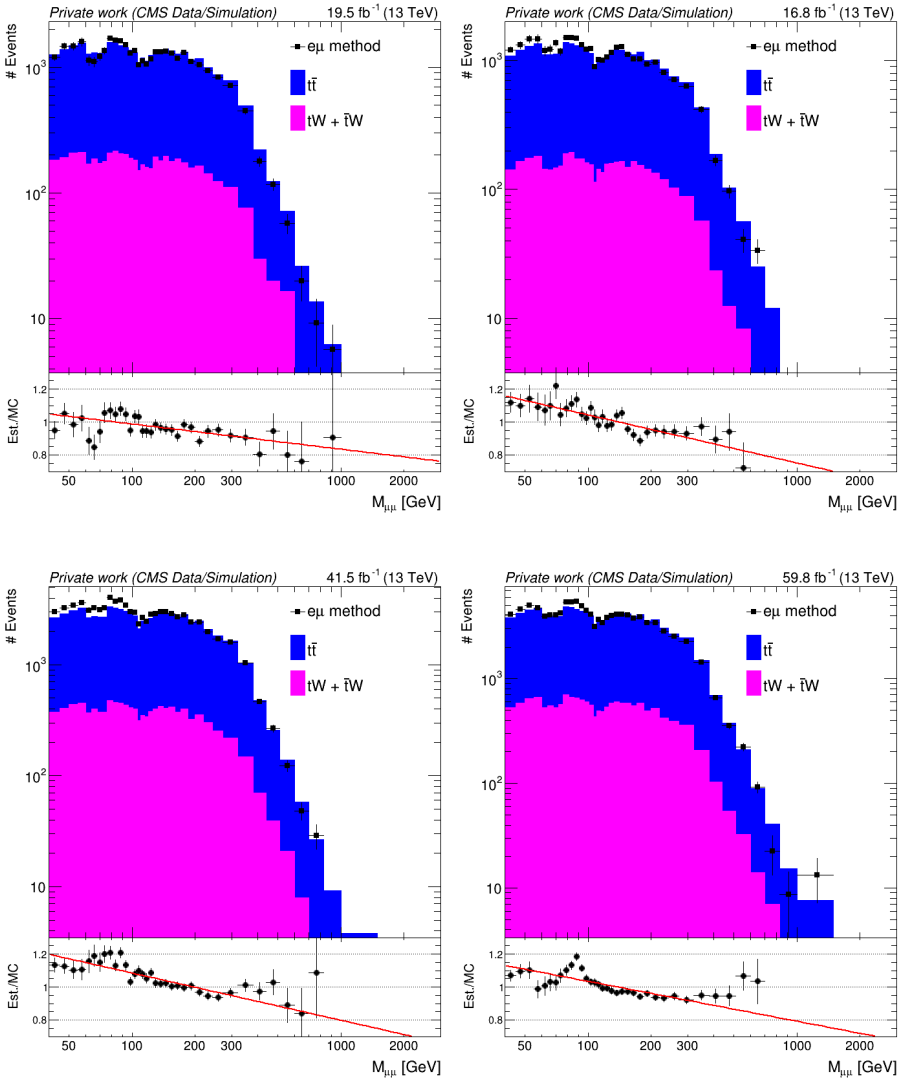


Fig. 31: Ratio of the modified $e\mu$ method prediction to the nominal simulated top-quark background in the dimuon channel for each era. The red line shows the linear fit in $\log_{10}(M_{\mu\mu})$ used to define the event weight in Eq. (10.9).

$e\mu$ channel by $\pm 10\%$, recomputing the misidentified/non-prompt estimate, and repeating the full top-quark correction procedure. The total systematic uncertainty is obtained by adding both contributions in quadrature.

The resulting relative uncertainties on the reweighted top-quark background are shown in Fig. 32. The systematic uncertainty does not exceed 6% over the full mass range. For dimuon masses below 600 GeV, the total uncertainty remains below 5%; at higher masses, it is dominated by the statistical uncertainty of the $e\mu$ control sample.

Finally, the reweighted top-quark backgrounds and the misidentified/non-prompt muon estimate are combined with the remaining simulated backgrounds in the dimuon channel. Figure 33 shows the resulting comparison between data and the total prediction. Comparing this result with Figs. 18 and 28, one can see that the $e\mu$-based top-quark reweighting significantly flattens the data-to-prediction ratio at high dimuon mass and improves the overall agreement.

10.2. Electron channel

In the dielectron channel, the dominant backgrounds at high dielectron invariant masses are top-quark and diboson processes with two prompt leptons in the final state. Below the Z peak, the main prompt-lepton background comes from DY $\rightarrow \tau^+\tau^-$ production, where both τ leptons decay into electrons. As in the muon channel, QCD multijet and W+jets events also contribute through misidentified/non-prompt electrons.

The background-estimation strategy in the dielectron channel follows the same general logic as in the dimuon channel, with data-driven methods used to estimate the misidentified/non-prompt lepton background and to correct the simulated top-quark backgrounds. The main additional complication is the higher probability of electron charge misidentification, especially at high p_{T} and in the endcap region. This effect causes prompt OS dielectron events, dominated by DY production, to enter the SS control region and must therefore be corrected before the SS region can be used to estimate the misidentified/non-prompt background.

10.2.1. Misidentified/non-prompt electron background estimation

The misidentified/non-prompt electron background is estimated using a same-sign method analogous to the one described for the muon channel in Subsection 10.1.1. A control region is defined by applying the same selection as in the dielectron signal region, except that the two electrons are required to have the same electric charge rather than opposite charge. Prompt-lepton processes

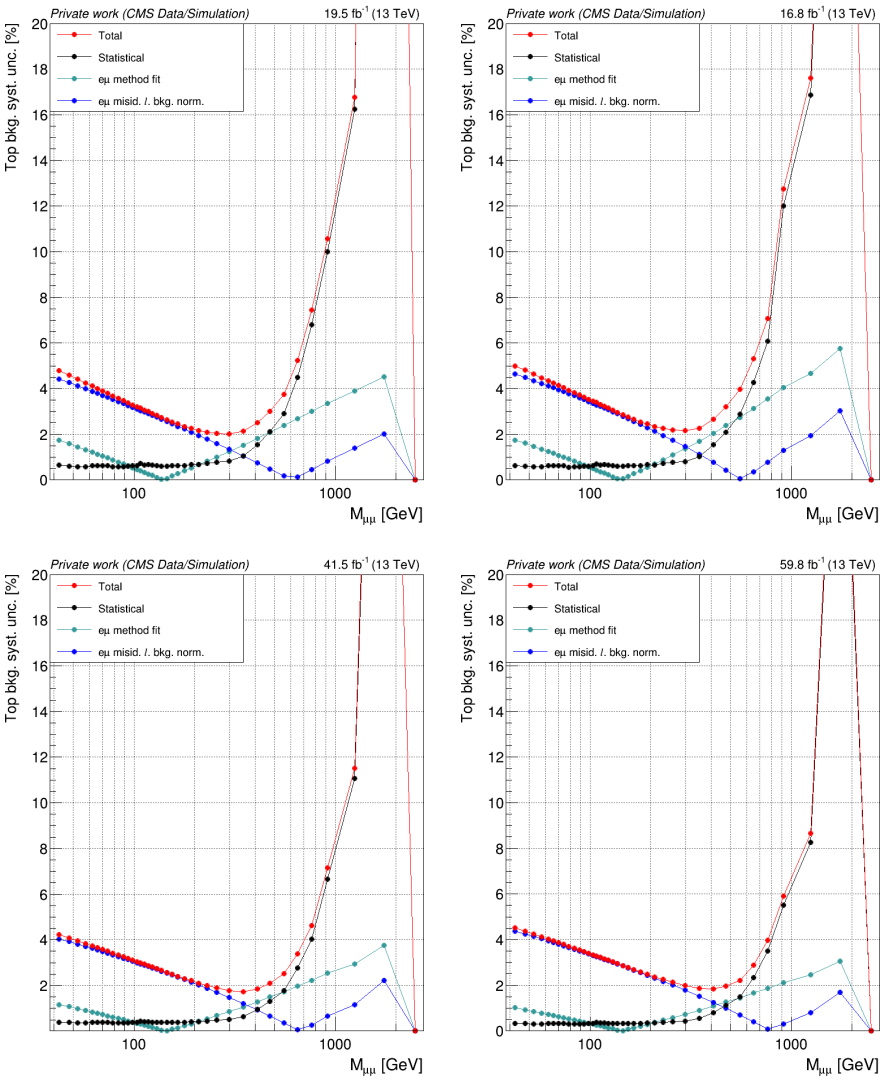


Fig. 32: Relative uncertainties on the reweighted top-quark background prediction in the dimuon channel for each era. Points with 0% uncertainty at very high mass correspond to empty bins. A single point at very high mass in each plot lies outside the vertical axis range due to limited statistics.

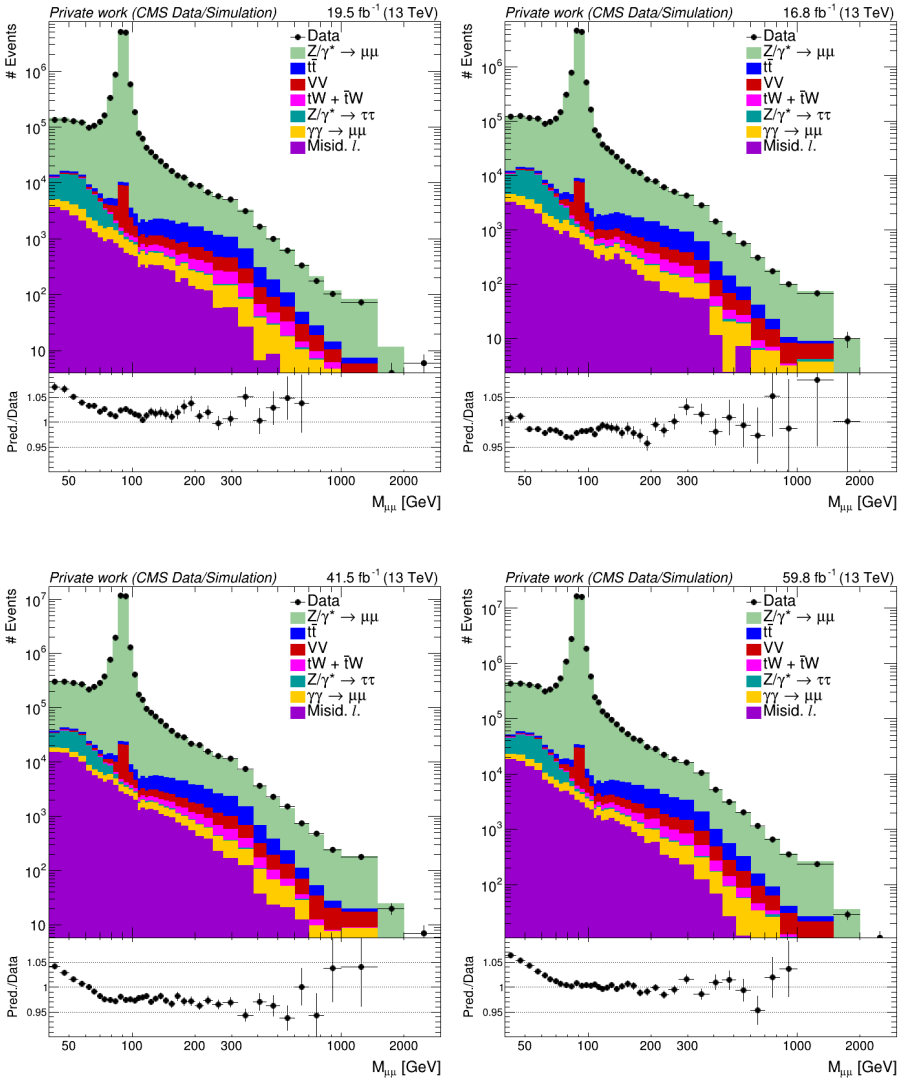


Fig. 33: Comparison between data and the total background prediction in the OS dimuon signal region, after applying the $e\mu$-based reweighting of top-quark backgrounds and including the misidentified/non-prompt muon estimate, for each data-taking era.

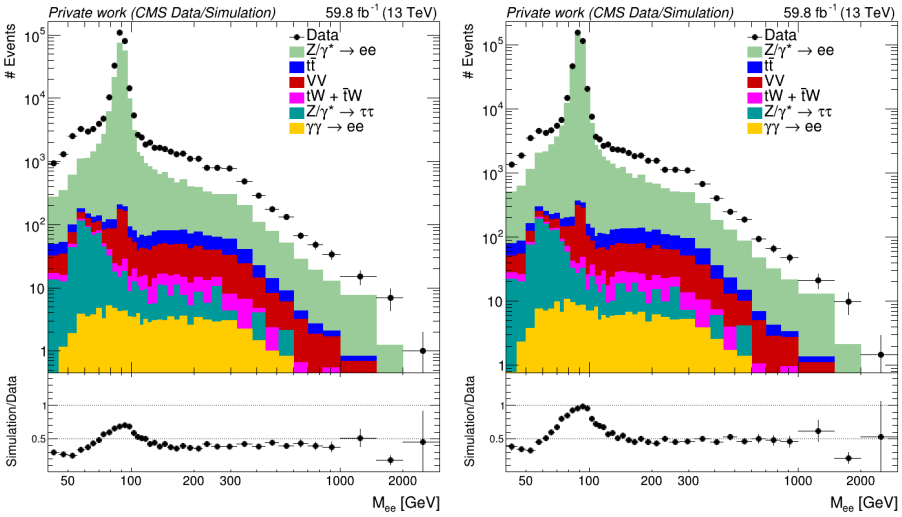


Fig. 34: Comparison between data and simulated prompt-lepton backgrounds in the SS dielectron control region before (left) and after (right) applying the charge misidentification correction, for 2018 data.

predominantly produce OS lepton pairs and are therefore suppressed in this region. The SS dielectron region is enriched in QCD multijet and W+jets events, where jets or non-prompt particles can be reconstructed as electron candidates.

In contrast to the muon channel, prompt DY production contributes non-negligibly to the SS dielectron control region because of electron charge misidentification. This effect is not perfectly modeled in simulation, particularly in the 2017 and 2018 samples. To improve the description of prompt SS dielectron events, charge misidentification weights derived by the CMS SMP-V DY forward-backward asymmetry analysis group [93] are applied to simulated DY events. These weights are derived from electron charge-misidentification probabilities and are applied to simulated electrons whose reconstructed charge differs from the generator-level charge.

Figure 34 shows the comparison between data and simulated prompt-lepton backgrounds in the SS dielectron control region for 2018 data, before and after applying the charge misidentification correction. The correction substantially improves the agreement between data and simulation. The behavior in the other data-taking eras is similar. After applying this correction, the remaining excess of data over the prompt-lepton simulation is interpreted as the SS misidentified/non-prompt electron contribution.

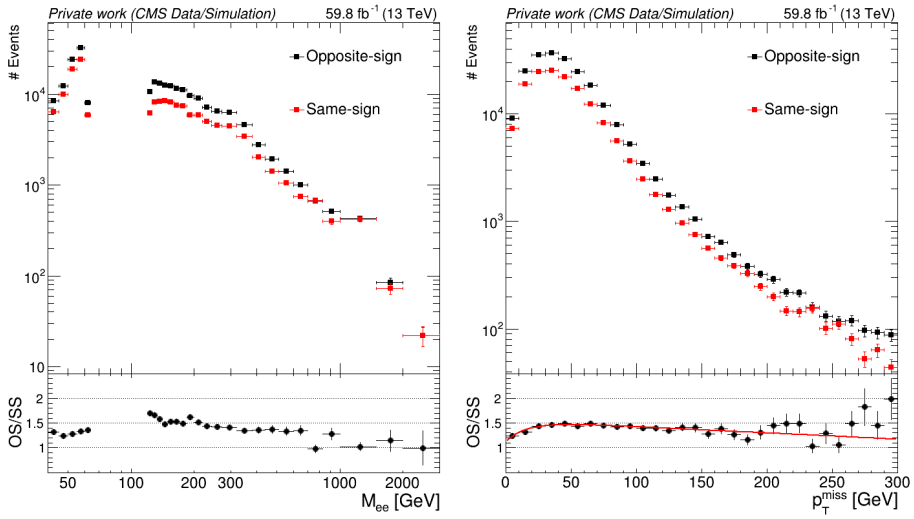


Fig. 35: Comparison of OS and SS dielectron invariant mass (left) and missing transverse momentum $p_{\mathrm{T}}^{\mathrm{miss}}$ (right) distributions for misidentified/non-prompt electrons in the sideband where one electron fails the Medium ID requirement, for 2018 data. The red line shows the functional fit used to define the event weight in Eq. (10.10).

As in the muon channel, the misidentified/non-prompt electron contribution in the OS signal region is obtained by rescaling the SS estimate with an OS/SS transfer factor $R_{\mathrm{OS/SS}}$ according to Eq. (10.3). The transfer factor depends on the relative contributions and kinematic properties of W+jets and QCD multijet events. To identify the sideband that best reproduces the W+jets-to-QCD mixture expected in the signal region, the same set of studies as in the muon channel is performed. In the electron channel, the most representative sideband is found to be the region where one of the two electrons fails the Medium cut-based ID requirement.

In this sideband, OS and SS distributions are compared after subtracting prompt-lepton contributions using simulation. As in the muon channel, events in the Z-peak region, $61 < M_{ee} < 121$ GeV, are removed to avoid artifacts from imperfect modeling of DY events in the sideband region. Figure 35 shows the corresponding OS and SS comparisons for the dielectron invariant mass and $p_{\mathrm{T}}^{\mathrm{miss}}$ distributions in 2018 data.

The $p_{\mathrm{T}}^{\mathrm{miss}}$ variable is found to provide good discrimination between W+jets and QCD multijet events in the electron channel. Unlike in the muon channel,

extending the transfer factor to depend on additional variables does not improve the closure of the method. Therefore, the OS/SS transfer factor is parameterized as a one-dimensional function of $p_{\mathrm{T}}^{\mathrm{miss}}$,

$$R_{\mathrm{OS/SS}}(p_{\mathrm{T}}^{\mathrm{miss}}) = a_0 - a_1 p_{\mathrm{T}}^{\mathrm{miss}} - \exp\left(a_2 - a_3 p_{\mathrm{T}}^{\mathrm{miss}}\right), \tag{10.10}$$

where the parameters a_i are obtained from a fit to the measured OS/SS ratio in the sideband. This is an empirical parameterization chosen to describe the observed dependence in the range relevant for the analysis. The fitted function is then applied event-by-event to SS data using the reconstructed $p_{\mathrm{T}}^{\mathrm{miss}}$ of each event.

A closure test is performed by applying the reweighting procedure in the sideband and comparing the reweighted SS distributions with the OS data. Examples for the dielectron mass and $p_{\mathrm{T}}^{\mathrm{miss}}$ distributions are shown in Fig. 36 for 2018 data. The other eras show similar behavior. The closure is adequate for the one-dimensional $\mathrm{d}\sigma/\mathrm{d}M$ measurement performed in this thesis. For some angular variables, such as electron pseudorapidity and dielectron rapidity, larger non-closure can be observed in certain regions; this is relevant mainly for potential multidimensional measurements and is not a limiting effect for the present one-dimensional mass measurement.

The fitted transfer factor $R_{\mathrm{OS/SS}}(p_{\mathrm{T}}^{\mathrm{miss}})$ is applied event-by-event to SS data in the main analysis region to obtain the OS misidentified/non-prompt electron prediction. As in the muon channel, the simulated top-quark contribution subtracted from the SS region is first corrected using the $e\mu$-based reweighting described in Subsection 10.1.2. The resulting OS misidentified/non-prompt electron mass distribution for 2018 data is shown in the left plot of Fig. 37.

The estimate obtained directly from the SS region shows residual structures near the Z peak. These structures are not expected for the misidentified/non-prompt electron background and arise from the large prompt DY contribution in the SS region, together with the finite precision of the charge misidentification correction. Previous CMS Drell–Yan measurements [15, 17, 19] show that the misidentified/non-prompt electron background is expected to be a smoothly falling function of the dielectron invariant mass, without a resonant feature at the Z peak.

To restore this expected behavior, the prediction in the range $76 < M_{ee} < 106$ GeV is replaced by a linear interpolation between the neighboring mass regions. The interpolation is performed in terms of event densities, $n(M) = \mathrm{d}N/\mathrm{d}M$, measured in events per GeV:

$$n_{\text{interp.}}(M) = n(M_1) + \frac{n(M_2) - n(M_1)}{M_2 - M_1}(M - M_1), \tag{10.11}$$

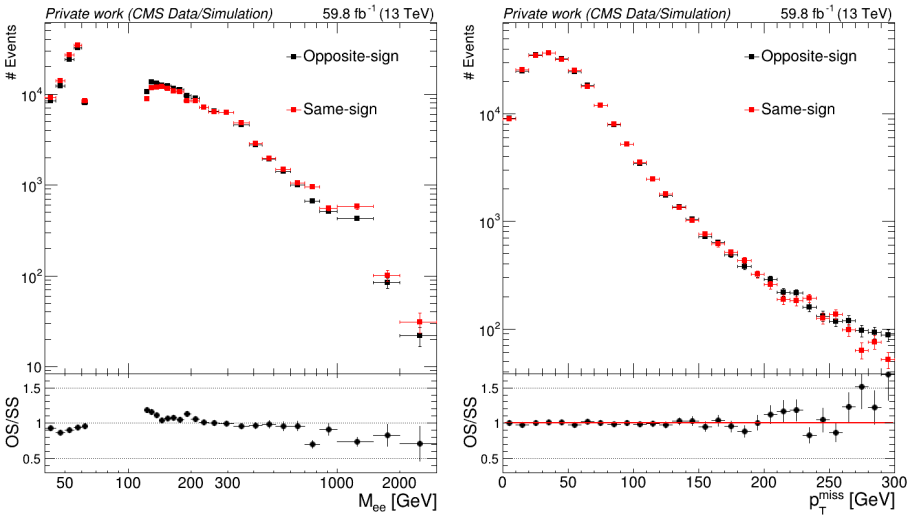


Fig. 36: Closure tests for the misidentified/non-prompt electron background in the sideband where one electron fails the Medium ID requirement. Shown are comparisons of OS data and reweighted SS data for the dielectron mass (left) and missing transverse momentum $p_{\mathrm{T}}^{\mathrm{miss}}$ (right), for 2018 data.

where M_1 and M_2 represent the mass points below and above the Z peak. To reduce sensitivity to statistical fluctuations, $n(M_1)$ and $n(M_2)$ are computed as average event densities in the ranges $50 < M_{ee} < 76$ GeV and $106 < M_{ee} < 170$ GeV, respectively. The statistical uncertainty of the interpolated value is obtained from the uncertainties on these averages:

$$\Delta n_{\text{interp.}}(M) = \sqrt{\left(\frac{M_2 - M}{M_2 - M_1}\Delta n(M_1)\right)^2 + \left(\frac{M - M_1}{M_2 - M_1}\Delta n(M_2)\right)^2}. \tag{10.12}$$

The final interpolated estimate for 2018 data is shown in the right plot of Fig. 37. Similar behavior is observed in the other data-taking eras.

Several sources of systematic uncertainty are considered for the misidentified/non-prompt electron background estimate:

1. the statistical uncertainty of the charge misidentification weights,
2. the uncertainty on the fitted transfer factor $R_{\mathrm{OS/SS}}(p_{\mathrm{T}}^{\mathrm{miss}})$, including correlations between fit parameters,
3. the choice of functional form used to parameterize $R_{\mathrm{OS/SS}}(p_{\mathrm{T}}^{\mathrm{miss}})$,

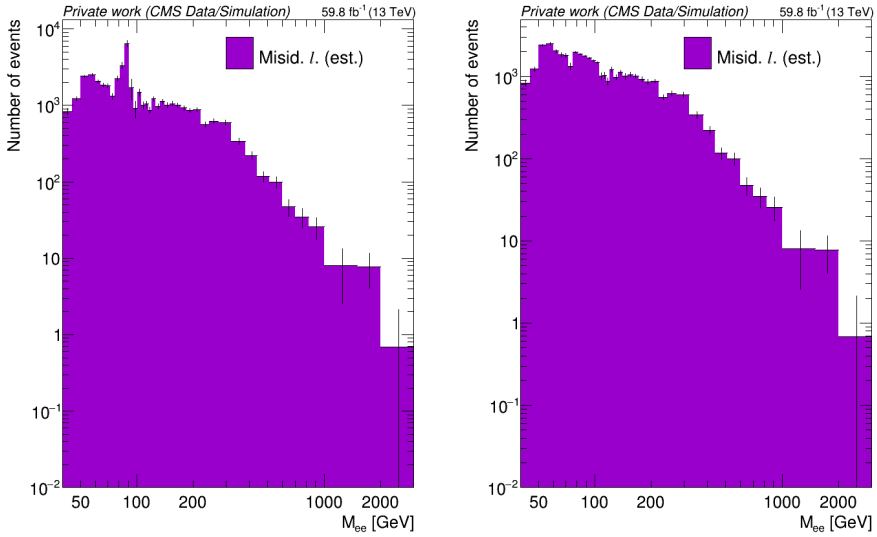


Fig. 37: Estimated misidentified/non-prompt electron background in the OS dielectron signal region for 2018 data. Left: distribution before replacing the Z-peak region; right: distribution after applying the linear interpolation.

4. the uncertainty associated with the interpolation in the Z-peak region,

5. the uncertainty on the top-quark background correction in the SS region.

For each source, an alternative estimate of the misidentified/non-prompt electron background is obtained by modifying the corresponding part of the procedure, and the associated uncertainty is taken as the absolute difference with respect to the nominal estimate, following Eq. (10.5).

The uncertainty due to limited statistics in the charge misidentification correction is evaluated by coherently varying the correction weights within their statistical uncertainties and taking the variation that leads to the largest deviation. The impact of the $R_{\mathrm{OS/SS}}(p_{\mathrm{T}}^{\mathrm{miss}})$ fit-parameter uncertainties is evaluated by coherently varying the fit parameters according to their covariance matrix. The uncertainty from the functional form of $R_{\mathrm{OS/SS}}(p_{\mathrm{T}}^{\mathrm{miss}})$ is assessed by comparing the nominal result to one obtained with a constant transfer factor. The constant is chosen to be $R_{\mathrm{OS/SS}} = 1.4$ for 2016 and $R_{\mathrm{OS/SS}} = 1.2$ for 2017 and 2018, based on representative values observed in the corresponding sidebands. The uncertainty associated with the Z-peak interpolation is evaluated by replacing the linear interpolation with a cubic spline and taking the difference between the two results. Finally, the uncertainty from the top-quark correction is evaluated by comparing the nominal estimate with one obtained without applying the

$e\mu$-based reweighting to the top-quark contribution in the SS region. The total systematic uncertainty is obtained by adding all contributions in quadrature.

The resulting relative uncertainties as a function of M_{ee} are shown in Fig. 38 for each data-taking era. The statistical uncertainty and the uncertainty from the functional form of $R_{\mathrm{OS/SS}}(p_{\mathrm{T}}^{\mathrm{miss}})$ dominate over most of the mass range. In the Z-peak region, the uncertainty is dominated by the interpolation procedure. Overall, the systematic uncertainty remains below 40% for $M_{ee} < 900$ GeV, while the total uncertainty stays below 40% for $M_{ee} < 400$ GeV. At very high masses, the uncertainties increase because of limited statistics in both the control and signal regions.

When added to the simulated prompt-lepton backgrounds, the estimated misidentified/non-prompt electron contribution improves the agreement between data and prediction in the dielectron signal region. Figure 39 shows the comparison between data and the sum of simulated prompt-lepton backgrounds plus the misidentified/non-prompt electron estimate. This background contributes approximately 2–4% of the total background at both low and high masses, and about 0.1% in the Z-peak region.

10.2.2. Top-quark background correction using the $e\mu$ method

The top-quark background correction in the dielectron channel uses the same modified $e\mu$ method as described for the dimuon channel in Subsection 10.1.2. The $e\mu$ control sample, the subtraction of non-top and misidentified/non-prompt contributions, and the construction of the event weight $w(M_{\ell\ell})$ follow the same procedure. In the dielectron channel, the weight is evaluated using $M_{\ell\ell} = M_{ee}$ and applied to simulated $t\bar{t}$, tW, and $\bar{t}W$ events.

The only practical difference is that in the dielectron channel the lepton p_{T} thresholds in the $e\mu$ control sample can be matched to those of the signal region. It is therefore unnecessary to repeat the stability tests on the p_{T} thresholds that were needed for the muon channel. The systematic uncertainties associated with the $e\mu$ reweighting are evaluated in the same way as in Subsection 10.1.2, with comparable relative size.

Figure 40 shows the comparison between data and the total background prediction in the dielectron signal region after applying the $e\mu$-based top-quark reweighting and including the misidentified/non-prompt electron estimate. Compared with Fig. 39, which shows the prediction before the top-quark reweighting, the ratio between data and prediction becomes flatter at high mass in the 2016 data. In the 2017 and 2018 eras, the reweighting improves the agreement but does not fully remove the upward trend at high masses, suggesting that additional

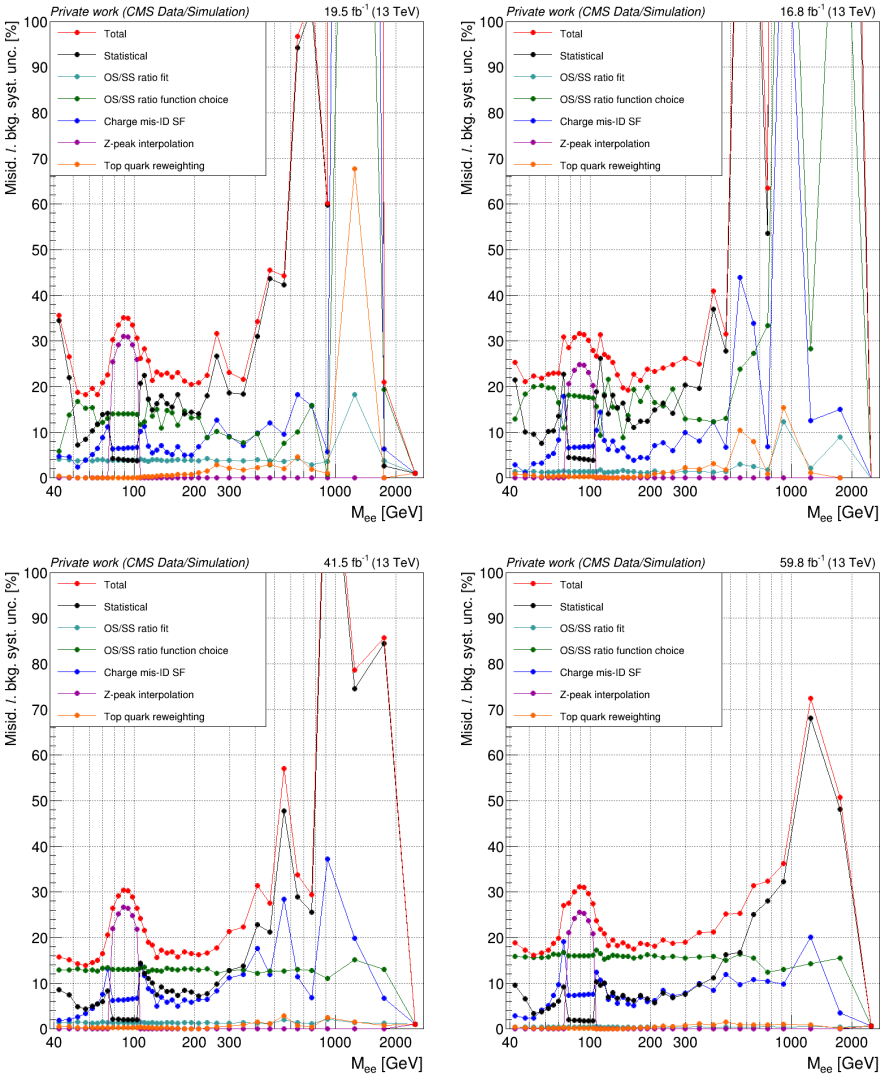


Fig. 38: Relative uncertainties on the misidentified/non-prompt electron background estimate for each data-taking era. Points with 0% uncertainty correspond to empty bins. Large uncertainties at the highest masses arise from limited statistics in the control and signal regions.

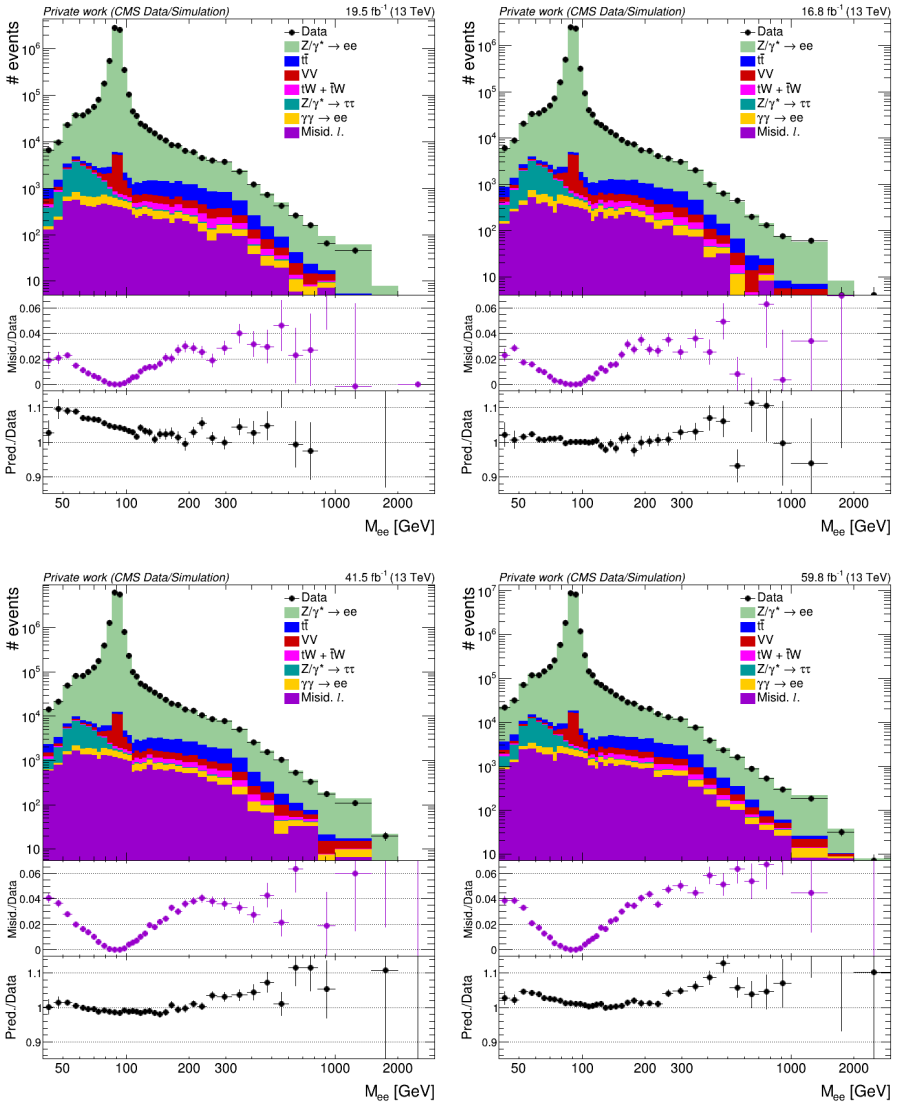


Fig. 39: Comparison between data and the sum of simulated prompt-lepton backgrounds plus the misidentified/non-prompt electron estimate in the OS dielectron signal region for each data-taking era. The middle panel shows the fraction of the observed data attributed to misidentified/non-prompt electron backgrounds, while the bottom panel shows the ratio of the total signal plus background prediction and the observed data.

electron-channel modeling effects not captured by the top-quark reweighting remain in those eras.

11. Unfolding

Measurements performed with a real detector are affected by finite resolution, reconstruction inefficiencies, and imperfect detector acceptance. As a result, the reconstructed dilepton mass distribution does not directly represent the underlying particle-level distribution in the fiducial phase space. To correct for these effects, an unfolding procedure is applied.

The detector response is described by a migration matrix $\mathbf{A}$, which relates the generator-level ("true" for short) distribution $\vec{x}$ to the reconstructed distribution $\vec{y}$:

$$\vec{y} = \mathbf{A}\,\vec{x}. \tag{11.1}$$

The matrix $\mathbf{A}$ is derived from simulated signal events and encodes the probability for an event generated in a given fiducial mass bin to be reconstructed in a given reconstructed mass bin. It includes the effects of detector resolution, reconstruction efficiency, acceptance within the fiducial definition, event weights, and the relation between reconstructed leptons and dressed generator-level leptons. Background contributions are subtracted from the reconstructed data distribution before unfolding.

The inversion of Eq. (11.1) can become unstable if the response matrix is ill-conditioned. In such cases, statistical fluctuations in the reconstructed distribution may be amplified in the unfolded result, and regularization is often introduced to stabilize the solution. In this analysis, the migration matrices are well-conditioned, with condition numbers below 4 in both channels. Therefore, no regularization is used.

The unfolding is performed with the `TUnfold` package [135]. The unfolded distribution is obtained by minimizing

$$L = (\vec{y} - \mathbf{A}\,\vec{x})^T \mathbf{V}_{yy}^{-1} (\vec{y} - \mathbf{A}\,\vec{x}), \tag{11.2}$$

where $\vec{y}$ is the background-subtracted reconstructed distribution, $\vec{x}$ is the unfolded distribution, and $\mathbf{V}_{yy}$ is the covariance matrix of the reconstructed distribution. Optional regularization and normalization constraints are supported by `TUnfold`, but are not used in this measurement.

Unfolding is performed separately in the dielectron and dimuon channels. Within each channel, the four data-taking eras are unfolded simultaneously to a single common particle-level spectrum. This is achieved by extending the

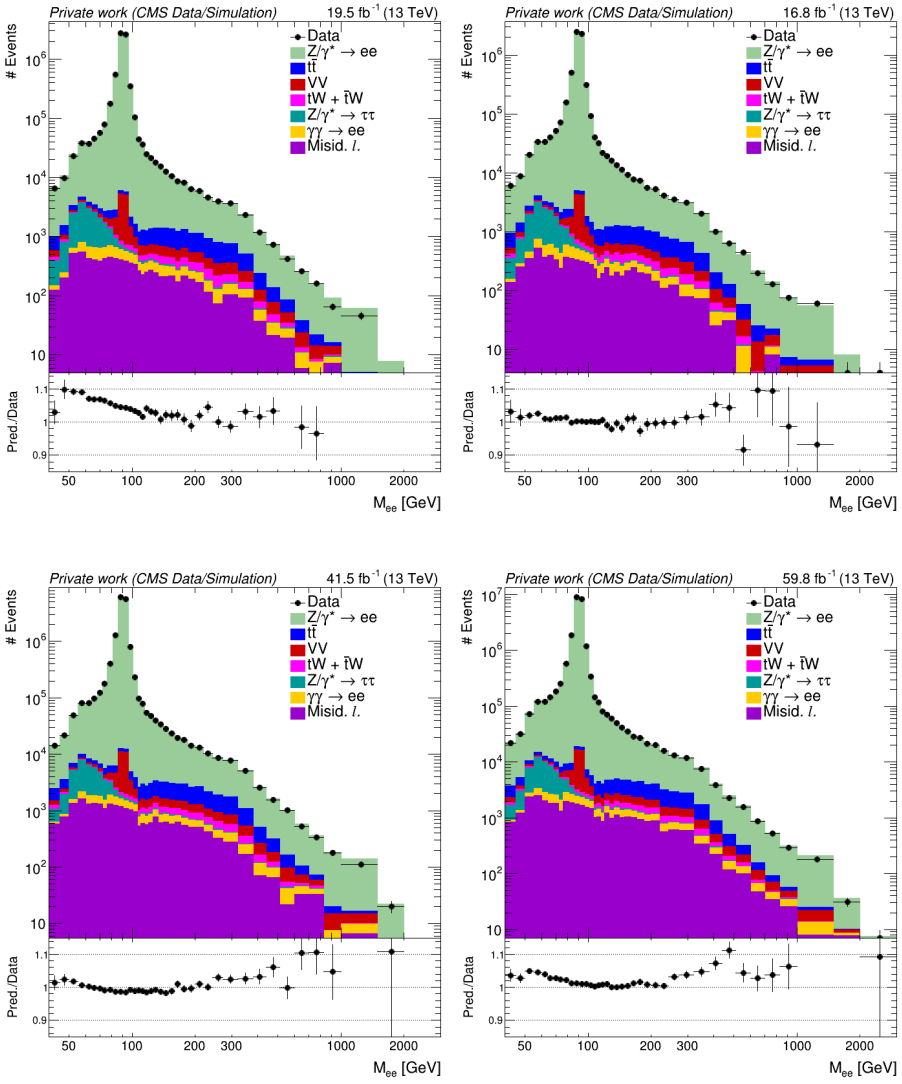


Fig. 40: Comparison between data and the total background prediction in the OS dielectron signal region after applying the $e\mu$-based top-quark reweighting and including the misidentified/non-prompt electron estimate, for each data-taking era: 2016 pre-APV (top left), 2016 post-APV (top right), 2017 (bottom left), and 2018 (bottom right).

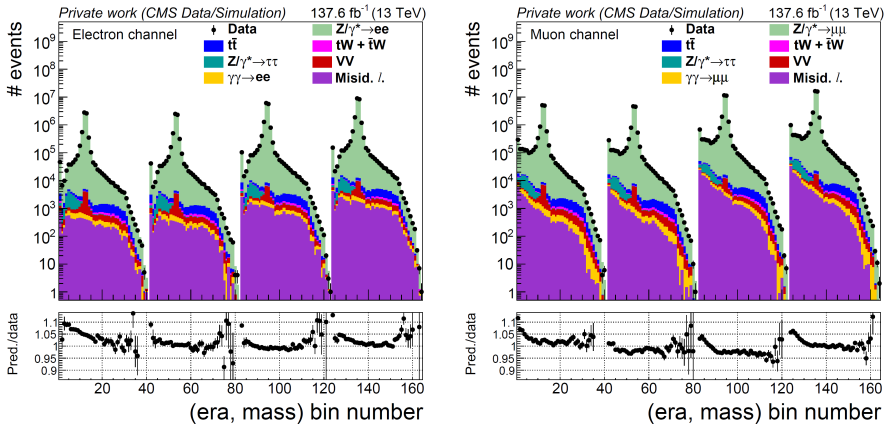


Fig. 41: Reconstructed invariant mass distributions for electrons (left) and muons (right), split by data-taking era. Each Z boson peak corresponds to a different era.

reconstructed axis of the response matrix with an additional era dimension: the same reconstructed mass binning is repeated for the 2016 pre-APV, 2016 post-APV, 2017, and 2018 periods, while the true axis contains only one common set of fiducial mass bins. This strategy satisfies the `TUnfold` requirement that the reconstructed distribution contain more bins than the true distribution, without artificially splitting the mass bins. It also allows the four eras to be combined naturally while preserving their separate detector responses and uncertainties.

The fiducial mass range of the measurement is $40 < M_{\ell\ell} < 3000$ GeV. Events may migrate into or out of this range because of detector resolution. The response matrix therefore includes underflow and overflow regions to track events reconstructed below 40 GeV or above 3000 GeV. Events generated outside the fiducial region but reconstructed inside the analysis selection are handled by `TUnfold` as non-fiducial contributions.

The reconstructed distributions for both channels are shown in Fig. 41, and the corresponding response matrices are shown in Fig. 42. The diagonal block structure of the response matrices reflects the era-split reconstructed axis.The figures in this and the remaining subsections of this section were produced within the CMS SMP-V Drell–Yan working group.

Two validation checks are performed. First, a direct closure test is carried out by constructing the response matrix and reconstructed distribution from the same simulated sample. Applying the unfolding procedure to this reconstructed distribution reproduces the corresponding generator-level distribution exactly,

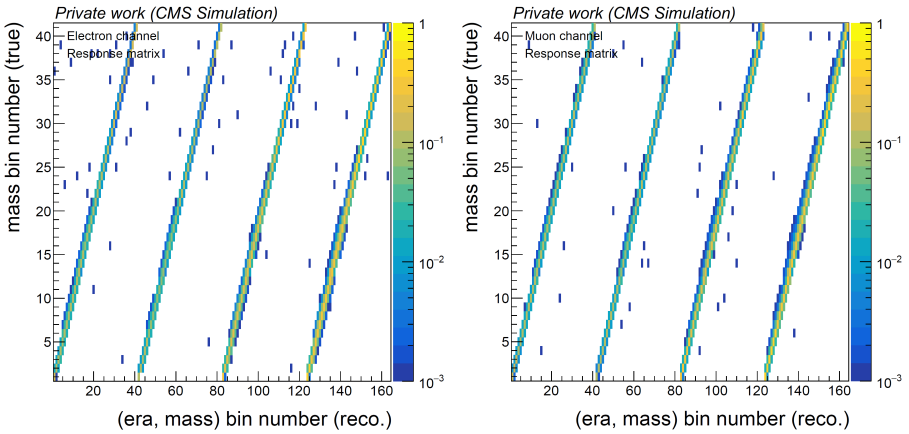


Fig. 42: Response matrices for electrons (left) and muons (right). Each diagonal block corresponds to a different data-taking era on the reconstructed axis, while the true axis corresponds to the common fiducial mass spectrum.

as expected. Second, a validation test is performed by splitting the simulated events into two statistically independent samples: one sample is used to build the response matrix, while the other is unfolded. In this case, the unfolded result is found to agree with the generator-level distribution within statistical uncertainties, confirming that the procedure is stable when applied to statistically independent data.

Finally, the unfolding is performed on data. Figure 43 shows the unfolded spectra compared with the generator-level prediction from the nominal signal simulation. Good agreement is observed across the full mass range.

The unfolding depends on the response matrix derived from signal simulation and is therefore sensitive to the modeling of the DY process. This dependence is treated as an unfolding model uncertainty. The nominal response matrix is built using the POWHEG MiNNLO signal sample. To estimate the model uncertainty, the unfolding is repeated using an alternative response matrix derived from a MADGRAPH5_AMC@NLO DY sample. The uncertainty is taken as the absolute bin-by-bin difference between the unfolded data spectra obtained with the nominal and alternative response matrices, normalized to the nominal result. The resulting model uncertainty remains below 2% over most of the mass range.

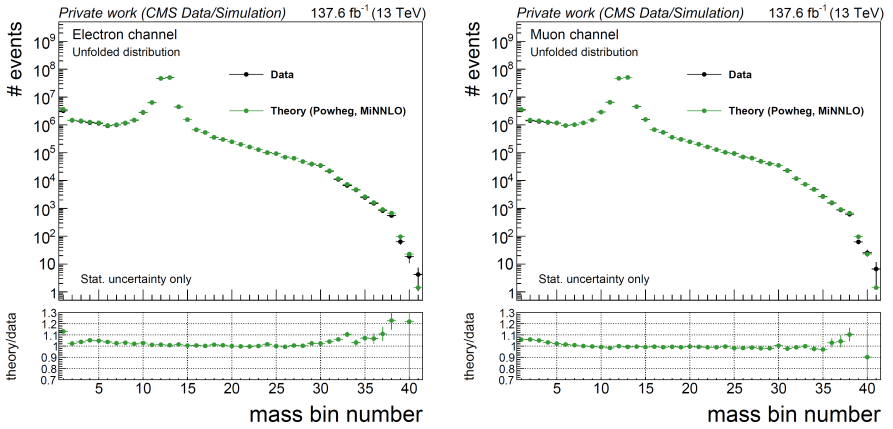


Fig. 43: Comparison of unfolded data with the generator-level prediction from the nominal signal simulation for electrons (left) and muons (right), with all data-taking eras combined.

12. Uncertainty estimation

This section summarizes the statistical and systematic uncertainties considered in the measurement. Each uncertainty source is propagated to the unfolded differential cross section, and a covariance matrix is constructed according to the assumed correlations between mass bins, data-taking eras, and lepton channels. The total covariance matrix is obtained by summing the individual covariance matrices.

The uncertainty sources considered in this analysis are:

1. data and simulation statistical uncertainties, including finite sample sizes in the migration matrix and background estimates,
2. PU reweighting,
3. L1 pre-firing correction,
4. lepton efficiency scale factors for reconstruction, identification, isolation, and trigger efficiencies,
5. lepton momentum or energy calibration,
6. data-driven background estimates,
7. theoretical inputs evaluated through PDF, α_s, and scale variations of the simulated samples,

8. integrated luminosity,

9. unfolding model uncertainty.

Statistical uncertainties from data, signal simulation, and background estimates are propagated through the unfolding using the covariance-matrix tools provided by `TUnfold`. The statistical uncertainty is small over most of the mass range but becomes important at the highest dilepton masses, where the event yields are limited.

The PU uncertainty is evaluated by varying the assumed inelastic proton-proton cross section by its uncertainty and propagating the resulting change in PU weights through the signal response, background predictions, and unfolding. This source is treated as correlated between data-taking eras and between lepton channels.

The uncertainty associated with the L1 pre-firing correction is evaluated using the event-by-event up and down variations provided in the input samples and propagating them through the full analysis chain. It is treated as correlated between eras and uncorrelated between the electron and muon channels, since the effect is driven by different detector subsystems in the two channels.

Lepton efficiency scale factor uncertainties are separated into statistical and systematic components. The statistical component originates from the finite size of the calibration samples used to measure the scale factors and is propagated by generating ensembles of randomized scale-factor maps, recalculating the cross section for each variation, and constructing the corresponding covariance. These statistical components are treated as uncorrelated between eras and between channels. Systematic components are evaluated using the dedicated alternative scale-factor variations provided by the calibration procedures and are treated as correlated between eras within a given channel.

Uncertainties from the muon momentum and electron energy calibrations are evaluated using the alternative calibration variations provided with the corresponding correction schemes. Replica-based statistical variations are propagated through ensembles and converted into covariance matrices, while systematic variations are propagated one by one and combined in quadrature. These sources are treated as correlated between eras within each channel.

Uncertainties associated with data-driven background estimates are propagated by varying the corresponding background estimates within their uncertainties and repeating the cross-section extraction. Statistical components are propagated through `TUnfold` as background-related covariance matrices, while systematic components are obtained from alternative estimates and treated according to the correlations of the underlying inputs. The details of the

background-related uncertainties are given in Sec. 10.

Theoretical-input uncertainties are evaluated by reweighting the simulated samples and repeating the full cross-section extraction. Three sources are considered: PDF uncertainties, α_s variations, and renormalization- and factorization-scale variations. For each variation, the relevant MC distributions, including simulated background distributions and migration matrices, are recomputed using the corresponding event weights. The full analysis chain is then repeated, and the resulting differential cross section is compared with the nominal result.

The PDF uncertainty is evaluated using the Hessian sets of the NNPDF 3.1 NNLO PDF set. The uncertainty in each mass bin is obtained from the quadrature sum of the deviations from the nominal result obtained with the individual Hessian sets. The α_s uncertainty is evaluated using PDF variations corresponding to shifted α_s values around the nominal value. The full cross-section extraction is repeated for both variations, and the larger bin-by-bin deviation from the nominal result is taken as the uncertainty.

The scale uncertainty is evaluated in an analogous way using the standard set of renormalization- and factorization-scale variations. The largest bin-by-bin deviation from the nominal result among the considered scale variations is taken as the uncertainty. Theoretical-input uncertainties are treated as correlated between eras and between the electron and muon channels. Correlations between mass bins are assigned according to whether the corresponding variations move the bins in the same or opposite direction with respect to the nominal result.

The integrated luminosity uncertainty affects both the normalization of simulated backgrounds and the normalization of the measured cross section. It is propagated using the correlated and uncorrelated luminosity components provided for the Run–2 data-taking periods. The corresponding covariance matrix is constructed by summing the contributions from the individual luminosity components.

The unfolding model uncertainty is evaluated as described in Sec. 11, by repeating the unfolding with an alternative response matrix derived from MadGraph5_aMC@NLO and comparing the result with the nominal unfolding based on Powheg MiNNLO. This uncertainty is treated as correlated between eras and between channels.

For each source, the resulting covariance matrix is added to the total covariance matrix. The total correlation matrix ρ_{ij} is obtained from the total covariance matrix $\mathbf{V}_{ij}$ as

$$\rho_{ij} = \frac{\mathbf{V}_{ij}}{\sqrt{\mathbf{V}_{ii}\mathbf{V}_{jj}}}. \tag{12.1}$$

The final relative uncertainties and the total correlation matrices for the

electron and muon channels are shown in Fig. 44. In the dimuon channel, the total uncertainty is approximately 2% at the Z peak and remains at or below about 6% over the mass range $40 < M_{\ell\ell} < 1500$ GeV. At low masses, below about 70 GeV, the uncertainty is dominated by muon efficiency scale factors. Around the Z peak and above it, the luminosity uncertainty becomes the dominant contribution, while in the range $200 < M_{\ell\ell} < 500$ GeV it becomes comparable to the theoretical-input uncertainty. At masses above about 600 GeV, the total uncertainty is increasingly dominated by the statistical uncertainty of the data.

In the dielectron channel, the total uncertainty is approximately 3% at the Z peak and remains at or below about 10% over the mass range $60 < M_{\ell\ell} < 1500$ GeV. At low masses, below about 60 GeV, the uncertainty is dominated by theoretical-input uncertainties. Around the Z peak, the luminosity uncertainty is the leading contribution. In the intermediate mass range, approximately $250 < M_{ee} < 700$ GeV, electron efficiency scale factors dominate, while at the lowest and highest masses the total uncertainty is driven mainly by statistical uncertainties.

13. Channel combination

The dielectron and dimuon differential cross sections are combined after verifying their compatibility. The two channels are first compared bin by bin using their unfolded fiducial cross sections and covariance matrices. A quantitative compatibility test is performed using

$$\chi^2 = \vec{y}^T M^{-1} \vec{y}, \tag{13.1}$$

where

$$y_i = \left(\frac{\mathrm{d}\sigma}{\mathrm{d}M}\right)_i^{ee} - \left(\frac{\mathrm{d}\sigma}{\mathrm{d}M}\right)_i^{\mu\mu} \tag{13.2}$$

is the difference between the dielectron and dimuon results in mass bin i, and M is the covariance matrix of this difference. Uncertainty sources that are fully correlated between the two channels, such as luminosity, theoretical inputs, and unfolding model uncertainty, cancel in the difference and are therefore not included in the compatibility covariance. The resulting value is $\chi^2/\text{ndf} = 1.08$, with 39 degrees of freedom, indicating good compatibility between the two channels.

The combination is performed using a hybrid method based on an inverse-variance weighted mean for the central values and a BLUE-style [136] propagation of the covariance matrix. In the first step, the central value of the combined

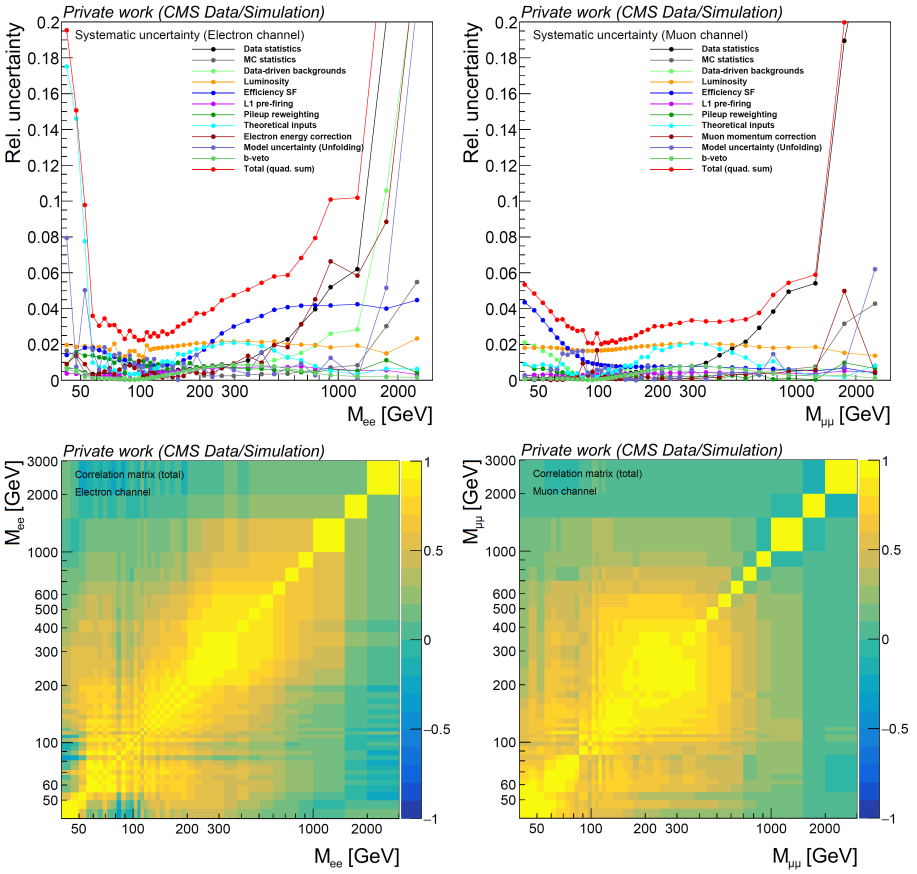


Fig. 44: Summary of the total relative uncertainties (top) and the total correlation matrices (bottom) for the electron (left) and muon (right) channels as a function of dilepton invariant mass.

result is computed independently in each mass bin as

$$\left(\frac{\mathrm{d}\sigma}{\mathrm{d}M}\right)_i^{\mathrm{comb}} = \lambda_i^{ee}\left(\frac{\mathrm{d}\sigma}{\mathrm{d}M}\right)_i^{ee} + \lambda_i^{\mu\mu}\left(\frac{\mathrm{d}\sigma}{\mathrm{d}M}\right)_i^{\mu\mu}, \tag{13.3}$$

where the weights are

$$\lambda_i^{ee} = \frac{(\mathrm{unc}_i^{ee})^{-2}}{(\mathrm{unc}_i^{ee})^{-2} + (\mathrm{unc}_i^{\mu\mu})^{-2}}, \qquad \lambda_i^{\mu\mu} = 1 - \lambda_i^{ee}. \tag{13.4}$$

Here, unc_i^{ee} and $\mathrm{unc}_i^{\mu\mu}$ are the total per-bin uncertainties of the dielectron and dimuon measurements, evaluated for the purpose of determining the weights. This choice gives a larger weight to the more precise channel in each mass bin.

In the second step, the covariance matrix of the combined result is constructed by propagating the full covariance information of the two input channels:

$$\mathrm{cov}^{\mathrm{comb}} = \lambda M_{2N} \lambda^T. \tag{13.5}$$

Here, N is the number of mass bins, λ is an $N \times 2N$ matrix containing the weights assigned to the dielectron and dimuon results, and M_{2N} is the full $2N \times 2N$ block covariance matrix of the two channels. The diagonal blocks of M_{2N} contain the dielectron and dimuon covariance matrices, while the off-diagonal blocks contain cross-channel covariance terms. These off-diagonal terms are included for uncertainty sources that are correlated between the channels, such as luminosity, theoretical inputs, and unfolding model uncertainty. Other uncertainty sources are treated according to the correlation assumptions described in Sec. 12.

The combined cross section and its uncertainty are compared with the individual channel results in Fig. 45. The combined spectrum lies between the two individual measurements and is typically closer to the dimuon result, reflecting its smaller uncertainty in many mass bins. The combination reduces the total uncertainty with respect to the individual channels, especially in regions where statistical or channel-specific systematic uncertainties dominate.

14. Cross section

The final result of the analysis is the Drell–Yan differential cross section measured as a function of the dilepton invariant mass. The primary result is the fiducial cross section, defined in the common generator-level fiducial phase space described in Sec. 6. This result is the least model-dependent measurement, since it avoids extrapolation to regions outside the experimental acceptance.

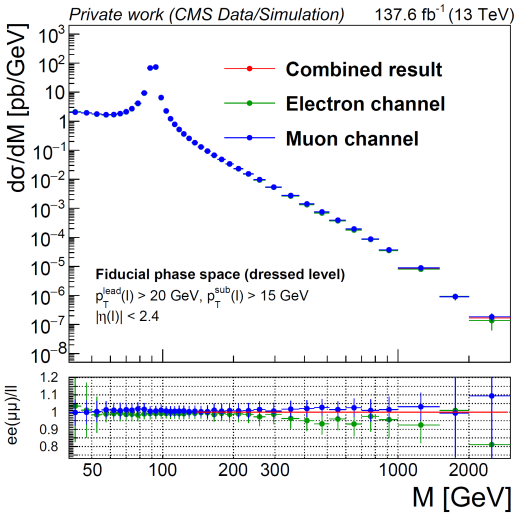


Fig. 45: Comparison between the combined and individual channel results for the measured Drell–Yan differential cross section.

For each mass bin i, the fiducial cross section can be written schematically as

$$\left(\frac{\mathrm{d}\sigma}{\mathrm{d}M}\right)_i^{\text{fid}} = \frac{N_i^{\text{unf}}}{\mathcal{L}_{\text{int}}\,\Delta M_i}, \tag{14.1}$$

where N_i^{unf} is the unfolded signal yield in the fiducial region, $\mathcal{L}_{\text{int}}$ is the integrated luminosity, and ΔM_i is the width of the mass bin.

A result extrapolated to the full phase space is also provided. The extrapolation is performed using an acceptance correction derived from the nominal signal simulation. In each mass bin, the acceptance is defined as

$$A_i = \frac{N_i^{\text{fid}}}{N_i^{\text{full}}}, \tag{14.2}$$

where N_i^{fid} is the number of generated signal events passing the fiducial requirements and N_i^{full} is the corresponding number of generated events in the full phase space. The full phase-space cross section is then obtained as

$$\left(\frac{\mathrm{d}\sigma}{\mathrm{d}M}\right)_i^{\text{full}} = \frac{1}{A_i}\left(\frac{\mathrm{d}\sigma}{\mathrm{d}M}\right)_i^{\text{fid}}. \tag{14.3}$$

Since this extrapolation depends on the simulated acceptance outside the fiducial region, the full phase-space result carries additional model dependence compared to the fiducial measurement.

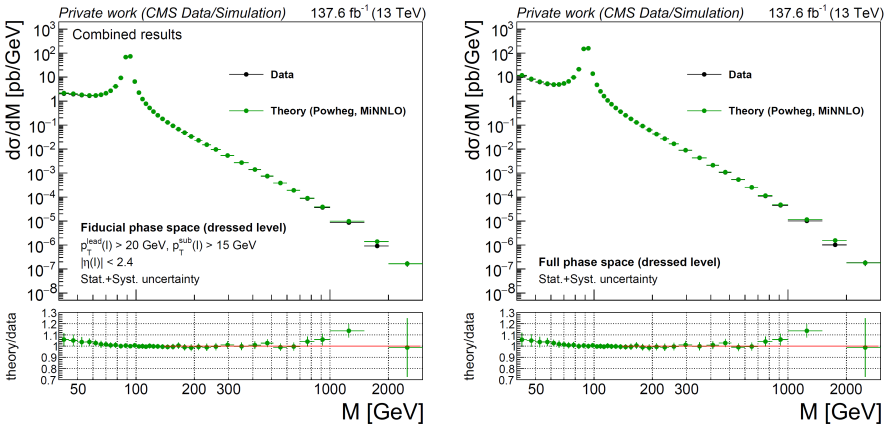


Fig. 46: Comparison of the combined Drell–Yan differential cross section with the SM prediction from POWHEG MiNNLO in the fiducial phase space (left) and extrapolated to the full phase space (right).

The combined differential cross sections in the fiducial and full phase space are shown in Fig. 46, together with the corresponding SM predictions from the POWHEG MiNNLO DY simulation. The measured cross sections are generally compatible with the SM predictions within the total uncertainties over the full mass range. A mild tension is observed in the $1000 < M_{\ell\ell} < 2000$ GeV region, at the level of approximately 2–3 standard deviations, but it is compatible with a statistical fluctuation given the limited event yield in this high-mass region.

15. High-Luminosity LHC

At the time of writing, LHC Run–3 is scheduled to continue until mid-2026, after which the LHC will enter the third Long Shutdown (LS3). During LS3, the accelerator complex and the major LHC experiments will undergo extensive upgrades for the High-Luminosity LHC (HL-LHC) program. The start of HL-LHC operation, corresponding to Run–4, is currently planned for June 2030 [103].

The HL-LHC is designed to increase the instantaneous luminosity delivered to ATLAS and CMS to about 5×10^{34} cm^{-2}s^{-1}, with an ultimate scenario reaching approximately 7.5×10^{34} cm^{-2}s^{-1} [21]. This will allow each experiment to collect several hundred inverse femtobarns per year and up to about 4000 fb^{-1} over the full HL-LHC program.

The increase in luminosity will be achieved through major upgrades of the accelerator complex. The inner triplet quadrupole magnets in the high-luminosity interaction regions will be replaced with stronger, radiation-tolerant Nb_3Sn magnets with larger aperture, allowing a smaller value of the beta function β^* (the focal length of the proton beam) at the interaction points. New corrector magnets, separation dipoles, absorbers, and superconducting radiofrequency crab cavities will also be installed. The crab cavities are needed to compensate for the larger beam crossing angle and recover the overlap of the colliding bunches [21].

To control the experimental conditions, the HL-LHC will use luminosity leveling, keeping the instantaneous luminosity approximately constant during a fill. At a luminosity of 5×10^{34} cm^{-2}s^{-1}, the average number of proton-proton interactions per bunch crossing is expected to be about 140. In the ultimate luminosity scenario, this value may reach about 200. These pileup conditions, together with the much larger integrated luminosity, are the main drivers of the Phase–2 upgrades of the ATLAS and CMS detectors.

The radiation environment will also become significantly more challenging. In the innermost regions of the CMS tracker, the expected 1 MeV neutron-equivalent fluence reaches approximately 2.3×10^{16} n_{eq}/cm^2, with a total ionizing dose of about 12 MGy for the full HL-LHC program [137]. Such radiation levels require new detector technologies, radiation-hard electronics, and upgraded cooling, power, and readout systems.

The ATLAS experiment will undergo a similar Phase–2 upgrade program to the CMS experiment, including a new all-silicon inner tracker, timing detector,

calorimeter and muon upgrades, and a redesigned trigger and data acquisition system [138]. The remainder of this section focuses on the CMS upgrades relevant for this thesis.

15.1. Physics goals

The HL-LHC will substantially extend the physics reach of the LHC by increasing the available proton-proton collision dataset by approximately an order of magnitude. This larger dataset will improve the statistical precision of rare-process measurements, extend the sensitivity of direct searches for new particles, and enable more stringent tests of the SM.

A central goal of the HL-LHC physics program is the precision study of the Higgs boson. The increased dataset will improve measurements of Higgs boson couplings to fermions and vector bosons, including rare or experimentally challenging channels such as the coupling to muons. It will also improve sensitivity to Higgs boson pair production, which provides direct access to the Higgs self-coupling and the structure of the Higgs potential.

The HL-LHC will also enable more precise studies of electroweak processes, including vector-boson scattering and rare multiboson production. These processes probe the mechanism of electroweak symmetry breaking and are sensitive to possible deviations from SM predictions. The larger dataset will also improve indirect searches for physics beyond the SM through precision measurements and interpretations in frameworks such as the SMEFT.

Direct searches for new particles will benefit from both the larger integrated luminosity and the upgraded detector capabilities. The HL-LHC will improve the sensitivity to a wide range of scenarios, including supersymmetry, additional gauge bosons, long-lived particles, and other signatures of physics beyond the SM. Even in the absence of direct discoveries, the increased precision of SM measurements will provide stronger constraints on possible new physics.

Drell–Yan measurements will also benefit from the HL-LHC program. The much larger dataset will reduce statistical uncertainties, especially at high dilepton invariant mass, while the upgraded detector will improve reconstruction performance in the challenging high-PU environment. More precise DY measurements can constrain parton distribution functions, test electroweak and QCD corrections, improve measurements of angular observables such as the forward-backward asymmetry, and provide important background constraints for searches involving dilepton final states. The impact of the CMS Phase–2 upgrade on the precision of DY measurements is discussed in Sec. “Precision of the Drell-Yan Process Measurement with the CMS Phase-2 Upgrade”.

16. Phase–2 challenges for the CMS detector

The HL-LHC environment will be significantly more challenging than that of Run–2 and Run–3. The bunch-crossing rate will remain up to 40 MHz, but with an average PU of about 140–200 interactions per crossing, the total rate of proton-proton interactions can reach several GHz. This will increase detector occupancies, trigger rates, event sizes, radiation damage, and the complexity of online and offline reconstruction.

Radiation damage is one of the main challenges for the Phase–2 detector. Particles produced in proton-proton collisions interact with detector materials and electronics, causing both ionizing and non-ionizing energy loss. Charged particles produce ionization and may initiate hadronic cascades in detector material. Photons from neutral-pion decays and other electromagnetic processes produce electromagnetic showers. Neutrons can travel through the detector volume and contribute to a mixed radiation field. Over the full HL-LHC program, these effects will degrade sensors, front-end electronics, optical links, and other detector components unless they are replaced with radiation-hard technologies.

Pileup is another major challenge. In-time PU increases the number of reconstructed tracks and vertices, complicates the assignment of particles to the correct interaction vertex, and degrades the reconstruction of isolated leptons, photons, jets, and missing transverse momentum. Out-of-time PU is particularly relevant for detector systems with long signal shaping times, such as calorimeters.

Figure 47 illustrates the challenge posed by high PU by comparing an event with a typical-PU level to a special high-PU event with about 100 reconstructed vertices. At the HL-LHC, the PU level will be even higher, requiring improved detector granularity, timing information, and reconstruction algorithms.

The higher particle rates and occupancies also impose strong requirements on the CMS trigger and data acquisition systems. The Level–1 trigger must retain low thresholds for electrons, muons, photons, jets, and missing transverse momentum while keeping the output rate manageable. To achieve this, the Phase–2 trigger will use tracking information at Level–1 for the first time in CMS. This requires a new tracker readout architecture capable of providing track information at the full bunch-crossing rate.

The increased event complexity will also affect the High-Level Trigger, offline reconstruction, simulation, and data processing. Reconstruction algorithms must remain efficient and precise in events with up to 200 PU interactions, while computing resources must scale to larger event sizes and higher data rates. These challenges motivate the broad CMS Phase–2 upgrade program, described

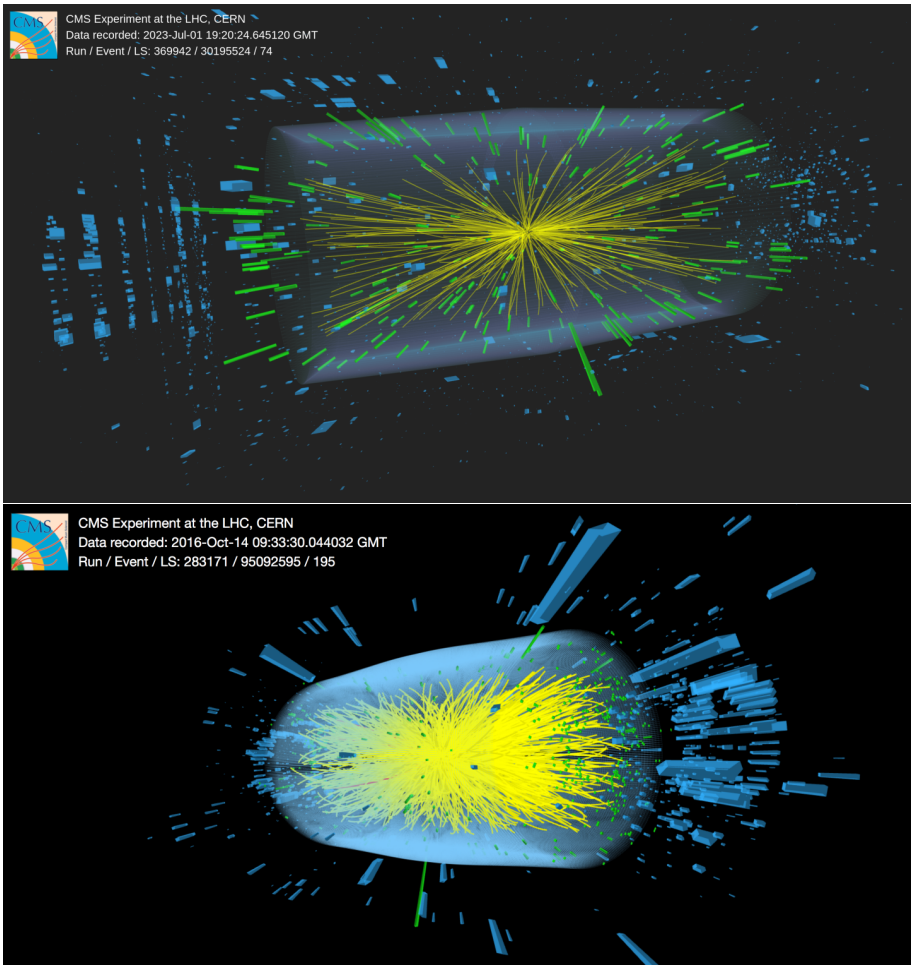


Fig. 47: Comparison of a regular-pileup event with 39 reconstructed vertices (top) [139] and a high-pileup event with approximately 100 reconstructed vertices (bottom) [140].

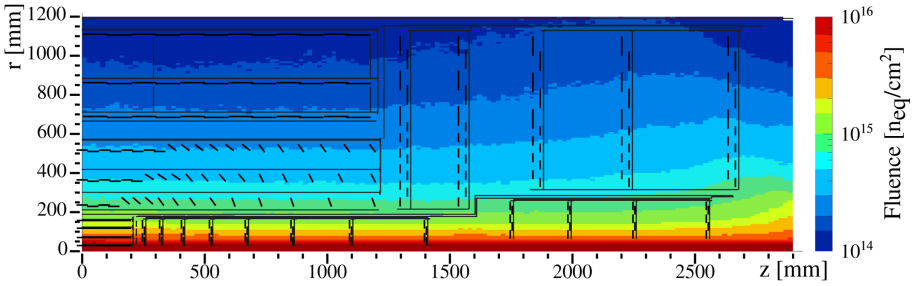


Fig. 48: Integrated particle fluence in 1 MeV neutron equivalent per cm^2 in the CMS Tracker for an integrated luminosity of 3000 fb^{-1} at the HL-LHC [22]. The distance from the proton beam in the transverse direction is denoted by r, while the longitudinal distance from the collision point in the center of the CMS detector is shown by z [137].

in the following sections, with particular emphasis on the tracking detector and its readout system.

16.1. Requirements for the CMS Tracker

The upgraded tracker must sustain complete efficiency up to an integrated luminosity of 3000 fb^{-1} and beyond, if the ultimate scenario were to be achieved [137]. A margin of roughly 50% is aimed for to account for uncertainties in simulations and the potential for additional integrated luminosity. This performance goal must be met without maintenance for the CMS Outer Tracker. For the Inner Tracker, the current accessibility design should continue, allowing extraction of the subdetector during regular shutdowns for module and component replacements as they degrade past the acceptable margin due to radiation damage. CMS simulations suggest radiation exposure in different detector regions will be about ten times higher than the levels aimed for in designing the current tracker [141, 142], reaching a 1 MeV neutron equivalent fluence of 2.3×10^{16} n_{eq}/cm^2 in the innermost sections. The particle fluence is primarily influenced by the transverse distance from the interaction point, with only minor variations along the longitudinal axis, as shown in Fig. 48.

For effective tracking performance under high-PU conditions, channel occupancy must be kept around or below the percent level for the Outer Tracker and the per mille level for the Inner Tracker, necessitating a higher number of channels than currently [137]. Benchmarking of the detector's performance uses target values of 140 and 200 collisions per bunch crossing.

The current tracker's ability to find tracks in highly energetic jets is limited

by hit overlap in the pixel detector. Enhancing two-track separation is essential to fully utilize the extensive collision data during high luminosity operations. Track finding under high-PU conditions also becomes increasingly complex and time-consuming. The upgraded tracker should be designed to enable rapid and efficient track finding, particularly for the HLT [137].

Selecting key physics events at the L1 trigger level is highly challenging at high luminosity due to both increased event rates and the reduced efficiency of selection algorithms under high PU. The CMS trigger will need to operate with substantially higher latency and output rates, requiring the tracker to be compatible with these demands. Additionally, to maintain and potentially enhance performance across a wide range of physics channels, CMS will incorporate tracking information into the L1 event selection, integrating part of the reconstruction currently performed at the HLT [137].

Extending the acceptance range of the tracker and calorimeters in the forward region could greatly enhance the overall physics capabilities of CMS [22]. The upgraded tracking system should ensure effective tracking up to approximately $|\eta| = 4$. A lighter tracker will significantly improve the physics potential at high luminosity. The material used in the present tracker impacts its performance and also affects the calorimeters and overall event reconstruction within CMS.

17. CMS Phase–2 upgrade

The CMS Phase–2 upgrade is designed to maintain and improve the detector performance in the HL-LHC environment [22]. The upgrade program includes the complete replacement of the tracking system, the replacement of the endcap calorimeters, extensions and improvements to the muon system, the installation of a precision timing detector, and major upgrades of the trigger, data acquisition, and computing systems.

The tracking detector will be completely replaced because the present system will not withstand the radiation levels expected during the HL-LHC program. The Phase–2 Tracker will have higher granularity, reduced material budget, extended pseudorapidity coverage, and radiation-hard sensors and electronics [137]. A key new feature is the use of tracking information in the L1 trigger. The Outer Tracker modules will provide local track segments, or stubs, for particles with p_{T} above approximately 2 GeV, enabling track-based trigger decisions at the full bunch-crossing rate. The Inner Tracker will use smaller pixels and thinner sensors to improve impact-parameter resolution, vertex reconstruction, and two-track separation in dense environments. Since the Tracker upgrade is central to the work presented in this thesis, it is described in more detail in

Sec. 17.1.

The endcap electromagnetic and hadronic calorimeters will be replaced by the High Granularity Calorimeter (HGCAL) [143]. The HGCAL is designed to withstand the high radiation levels in the forward region and to provide fine transverse and longitudinal segmentation. This allows three-dimensional reconstruction of electromagnetic and hadronic showers, improving particle identification, PU mitigation, and energy reconstruction in the endcap region.

The muon system will also be upgraded to maintain efficient triggering and reconstruction in the forward region. Additional Gas Electron Multiplier (GEM) and improved Resistive Plate Chamber (RPC) detectors will provide increased redundancy, better timing, and improved spatial resolution in regions with high background rates [112, 144]. These upgrades are particularly important for maintaining low trigger thresholds and good muon identification performance under HL-LHC pileup conditions.

The MIP Timing Detector (MTD) will provide precision timing measurements for charged particles with a resolution of about 30–40 ps [145]. This timing information adds a fourth dimension to event reconstruction and helps separate particles originating from different PU vertices. The MTD is expected to improve vertex reconstruction, PU mitigation, lepton isolation, and the reconstruction of low-p_{T} particles in dense environments.

The beam background monitoring and luminosity measurement systems will be upgraded to operate reliably under the high radiation levels expected at the HL-LHC [146]. These systems provide essential information for detector protection, luminosity calibration, and monitoring of machine-induced backgrounds.

The trigger and data acquisition systems will undergo major upgrades to handle the increased event complexity and data rates [147, 148]. The L1 trigger latency will increase to 12.5 μs, allowing time for hardware track reconstruction and for matching tracks to calorimeter and muon information. The L1 trigger output rate will increase to 500 kHz for PU conditions around 140, with the system designed to support up to 750 kHz for PU values near 200. The upgraded DAQ system will provide the bandwidth and computing resources needed to read out and process these higher rates.

The software and computing model must also evolve to cope with the larger event sizes, higher data rates, and more complex reconstruction expected at the HL-LHC [149]. Improvements in reconstruction algorithms, data formats, storage strategies, and distributed computing workflows are required to keep online and offline processing feasible. Together, these detector, trigger, DAQ, and computing upgrades are essential for preserving the physics performance of CMS throughout the HL-LHC program.

The CMS tracking system will be completely replaced for Phase–2 operation in order to maintain efficient charged-particle reconstruction in the HL-LHC environment [137]. The upgraded Tracker is divided into two main systems: the Outer Tracker (OT) and the Inner Tracker (IT). The OT will consist of six barrel layers and five endcap disks on each side, covering a silicon surface of approximately 190 m^2 with about 13 200 modules and 213 million readout channels. It is designed to withstand fluences up to about 10^{15} n_{eq}/cm^2 in its innermost layers.

A key feature of the OT is its ability to provide information to the L1 trigger. The OT modules are built from two closely spaced silicon sensors. By correlating hits in the two sensors, the front-end electronics can form local track segments, called stubs. Since the displacement between the two hits depends on the curvature of the charged-particle trajectory in the CMS magnetic field, the stub selection can reject low-p_T particles. Only stubs compatible with tracks above approximately 2 GeV are transmitted to the L1 trigger, significantly reducing the data volume while preserving information relevant for trigger decisions.

Two types of OT modules are used. The 2S modules contain two strip sensors, while the PS modules contain one strip sensor and one macro-pixel sensor. This design provides the granularity required to keep occupancies below the percent level while enabling stub reconstruction for the trigger. Compared with the present tracker, the OT will also have a reduced material budget, improving track momentum resolution and reducing the rate of photon conversions and hadronic interactions in the tracking volume.

The IT will consist entirely of silicon pixel modules and will cover the region closest to the interaction point. It is designed to provide precise vertexing, impact-parameter measurements, and track seeding in the most radiation-intense and highest-occupancy part of CMS. The IT will contain approximately two billion pixel channels over about 5 m^2 of silicon and will extend the tracking acceptance to nearly $|\eta| < 4$. The innermost regions must withstand total ionizing doses up to about 1.2 Grad and hit rates up to approximately 3.2 GHz/cm^2.

The IT pixel pitch will be reduced to 25×100 μm^2. The smaller pixel size is required to keep occupancies low, improve two-track separation in dense environments, and maintain high tracking efficiency in boosted jets. The material budget will also be reduced compared with the present tracker, improving tracking performance and reducing the impact on calorimeter measurements.

Figures 49 and 50 compare the layout of the current Phase–1 Tracker with

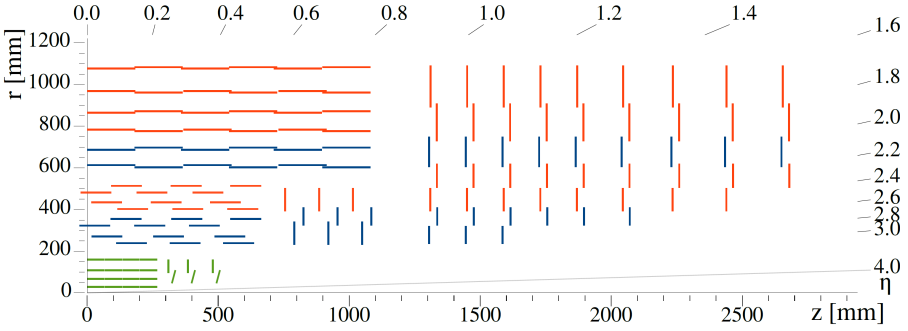


Fig. 49: A one-quarter slice of the current Phase–1 CMS Tracker in the r–z plane [137]. The silicon pixel detector modules are shown in green, while single-sided and double-sided silicon strip detector modules are shown in red and blue, respectively.

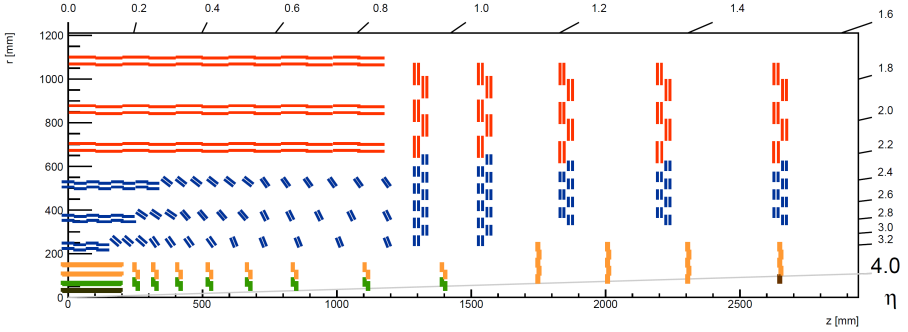


Fig. 50: A one-quarter slice of the CMS Phase–2 Tracker in the r–z plane [137]. The red and blue lines show modules of the Outer Tracker, while the yellow, orange, green, and brown lines denote modules of the Inner Tracker.

the planned Phase–2 Tracker. The Phase–2 layout extends the tracking coverage further into the forward region and uses a different module layout optimized for high PU, trigger compatibility, and reduced material.

Simulations indicate that the upgraded Tracker will maintain high tracking efficiency and good momentum resolution over a broad pseudorapidity range, even under HL-LHC pileup conditions [137]. The improved tracking performance is essential for nearly all CMS physics analyses, including the reconstruction of electrons, muons, jets, missing transverse momentum, and displaced vertices.

The IT data acquisition system must read out approximately two billion pixel channels and maintain stable operation under extreme hit rates and radiation levels. The innermost pixel layer, located at a radius of about 3 cm, is expected to experience hit rates up to approximately 3.2 GHz/cm^2. The full IT system will contain roughly 14 000 pixel readout chips, which must be synchronized, configured, monitored, and read out with high reliability.

The IT DAQ architecture is designed to minimize off-detector bandwidth, material budget, and power consumption while maintaining sufficient readout capacity for the HL-LHC environment [137]. Pixel modules transmit data through electrical links, called e-links, to a Low-Power Gigabit Transceiver (lpGBT) [150]. The lpGBT aggregates the data and sends them optically to the back-end electronics through the Versatile Transceiver Plus (VTRx+) electro-optical transceiver [151]. Data are transmitted from the detector to the back end at 10.24 Gb/s, while clock, trigger, command, and configuration data are transmitted in the opposite direction at 2.56 Gb/s.

The complete IT readout chain is illustrated in Fig. 51. Depending on the detector region, a pixel module contains either two or four readout chips. These chips are connected to one lpGBT, either directly or through local data aggregation within the module. This architecture reduces the number of required electrical links and off-detector optical fibers, helping to limit the material budget while preserving the required readout bandwidth.

At the back end, optical data are received by ATCA-compliant Data, Trigger, and Control (DTC) boards. Each DTC handles a fixed subset of detector channels and is responsible for communication with the front-end electronics, trigger distribution, event building, buffering, and data transfer to the central CMS DAQ. The IT readout system uses 28 DTC boards in total. Each board serves up to 72 optical inputs and reads out approximately 700 pixel readout chips, corresponding to about 70 million channels. The DTC boards use high-performance field-programmable gate arrays (FPGAs) and a System-on-Chip co-processor to provide local processing and high-speed communication with the central DAQ system.

The IT DAQ system is directly relevant to the work presented in Sec. “Preparation Works for the Inner Tracker Upgrade”, where contributions to readout-chip testing and Ph2-ACF software development are described.

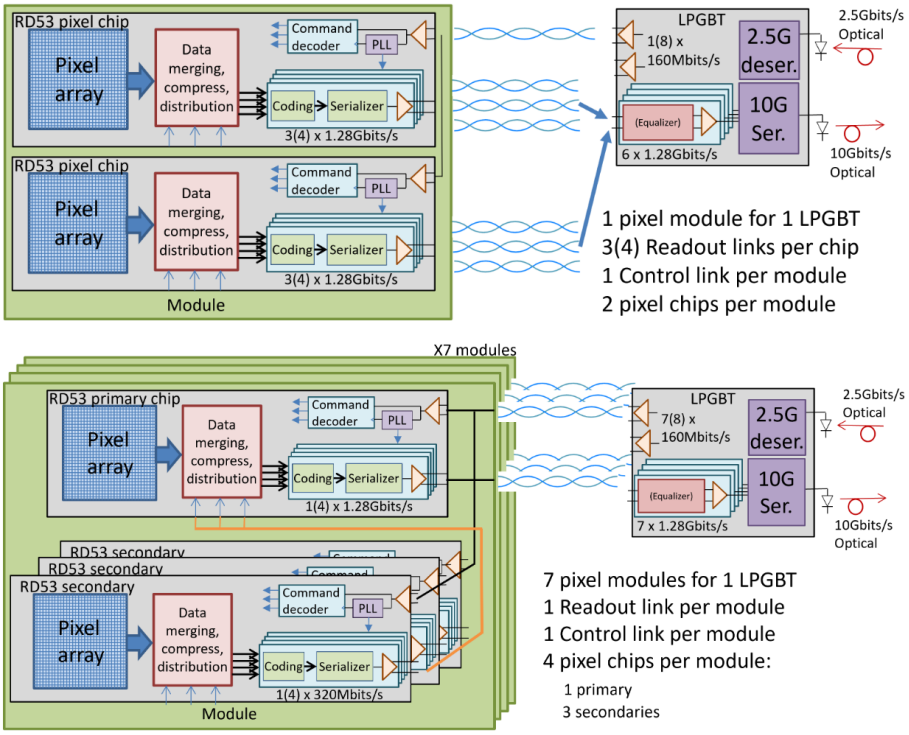


Fig. 51: The IT module readout chain using e-links. Top: Inner Barrel layer module with two pixel chips connected to one lpGBT. Bottom: outer IT layer module with four pixel chips per module and local data aggregation. Adapted from Ref. [137].

The extreme hit rates, radiation levels, and bandwidth requirements of the Phase–2 Inner Tracker require a dedicated pixel readout chip. To meet these requirements, the RD53 Collaboration [23], a joint effort between CMS and ATLAS, developed a family of radiation-hard pixel readout chips in 65 nm CMOS technology [152].

The first large-scale prototype, RD53A, was developed as a technology demonstrator and used to evaluate several analog front-end architectures and digital readout concepts [153, 154]. Based on the results of this program, CMS selected the linear front-end architecture [155]. This design was further developed in the CMS-specific RD53B-CMS prototype, also known as CROCv1 [156]. The final CMS production chip, RD53C-CMS, also known as CROCv2, is the readout chip intended for installation in the Phase–2 CMS Inner Tracker [A2]. A photograph in Fig. 52 shows a Single-Chip Card (SCC) with a CROCv1 chip.

The main features of the CROCv2 readout chip include:

- fabrication in 65 nm CMOS technology, optimized for radiation tolerance;
- a pixel matrix of 336×432 pixels, corresponding to 145 152 channels per chip;
- a pixel size of 25×100 μm^2;
- zero-suppressed readout;
- charge measurement using 4-bit Time-over-Threshold (ToT) digitization;
- dual-slope gain to extend the dynamic range;
- a global hit-detection threshold with per-pixel 5-bit threshold trimming;
- data compression based on binary-tree encoding;
- a 160 Mb/s input link for clock, trigger, command, and configuration signals;
- four 1.28 Gb/s output data links using the Aurora protocol [157];
- internal charge-injection circuitry for calibration and monitoring.

The readout architecture is optimized to reduce data volume before transmission off the module. Pixels are organized and encoded in a way that reduces

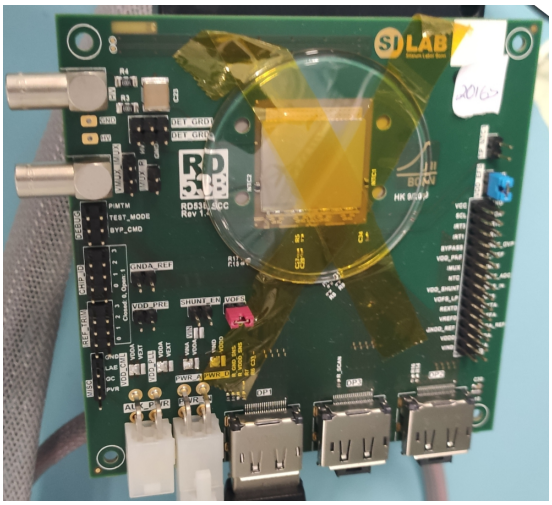


Fig. 52: A Single-Chip-Card with a CROCv1 chip without a sensor under a protective glass.

address overhead, and hit data are transmitted only when a pixel signal exceeds threshold. Each chip provides up to four high-speed Aurora output lanes, allowing sufficient bandwidth for the innermost detector regions. Within a module, data from multiple readout chips can also be merged before transmission to the lpGBT. This local data aggregation reduces the number of required electrical links and helps minimize the detector material budget.

The RD53B-CMS prototype and the final CROCv2 chip are central to the Inner Tracker qualification program. Their calibration, tuning, data readout, and quality-assurance procedures require dedicated software and test systems. The author's contributions to these activities, including work with the Ph2-ACF framework and RD53B-CMS prototype testing, are presented in Sec. "Preparation Works for the Inner Tracker Upgrade".

18. Module test setups

Testing and characterization of Inner Tracker pixel modules require a flexible laboratory setup that can be installed in different locations and connected to a standard Linux computer. The basic setup, illustrated in Fig. 53, consists of a computer running the Ph2-ACF software, a back-end DTC board, and one or more pixel modules or Single-Chip Cards (SCCs). Communication between the computer and the back-end board is performed using the IPbus protocol [158].

The setup used for prototype testing is based on a custom μTCA DTC board (μDTC), equipped with a Xilinx Kintex-7 FPGA, two FPGA Mezzanine Card (FMC) connectors, and DDR3 memory. A commercial digital-I/O FMC board can be used to provide external trigger and clock signals, while a custom FMC board provides the electrical connection to the module through a DisplayPort cable. The same general architecture can also be adapted for optical readout by replacing the electrical interface with an optical one.

The firmware running on the FPGA provides the low-level functionality needed for chip operation. It encodes and decodes commands, distributes clock and trigger signals, controls calibration injections, stores the received data in memory buffers, and hides hardware-specific communication details from the higher-level software. This abstraction allows the same software framework to operate different front-end chips, including RD53A and CROC, using a common control structure.

19. Ph2-ACF Software

The Phase-2 Acquisition and Control Framework (Ph2-ACF) is the main software framework used for controlling and testing CMS Phase-2 Tracker front-end electronics [24, 159]. It is implemented in C++ and provides a common interface for configuring back-end boards, communicating with front-end chips, running calibration procedures, collecting data, and producing monitoring output. Although Ph2-ACF was developed primarily for module testing and characterization, its structure follows the same general principles required by a large distributed detector readout system.

A schematic view of the Ph2-ACF architecture is shown in Fig. 54. The lower software layers provide access to the firmware and hardware registers. The middleware layer translates hardware operations into C++ interface classes, while the higher-level tools implement calibration scans, data-taking procedures,

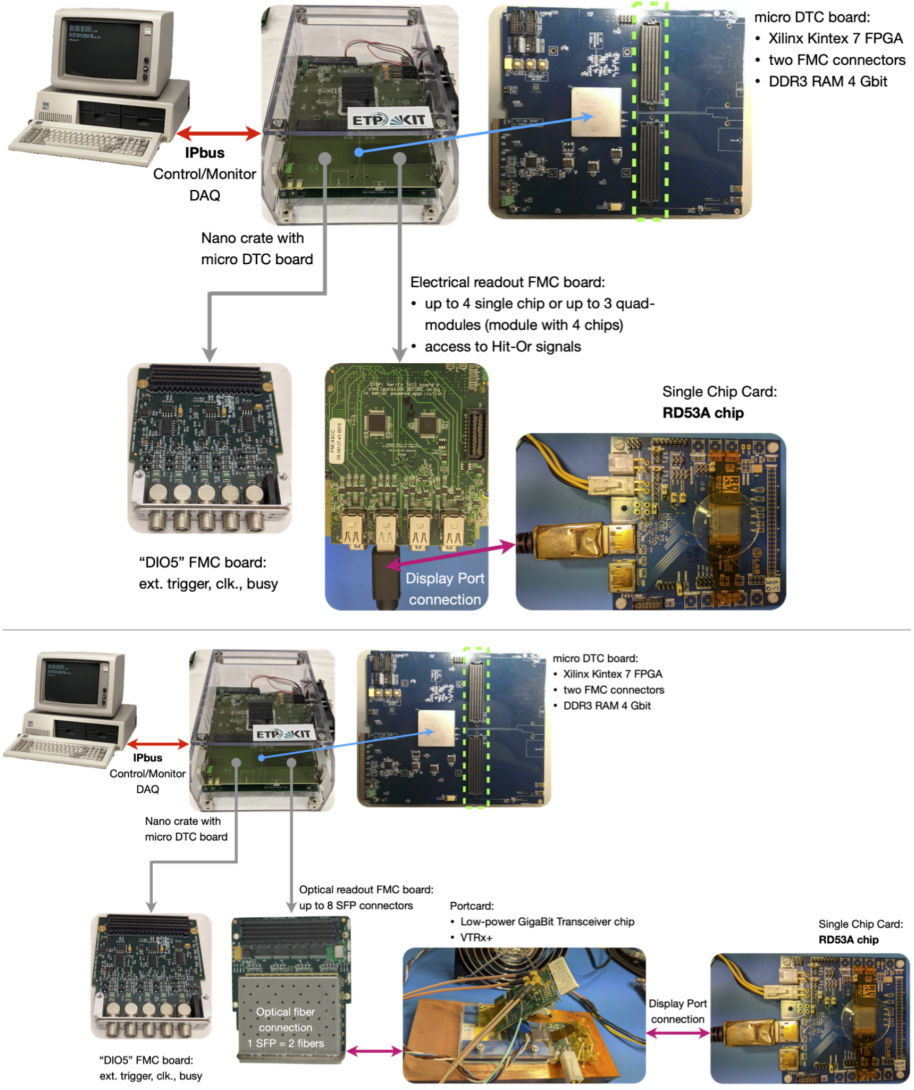


Fig. 53: Main components of the module testing and characterization setup. Adapted from Ref. [159].

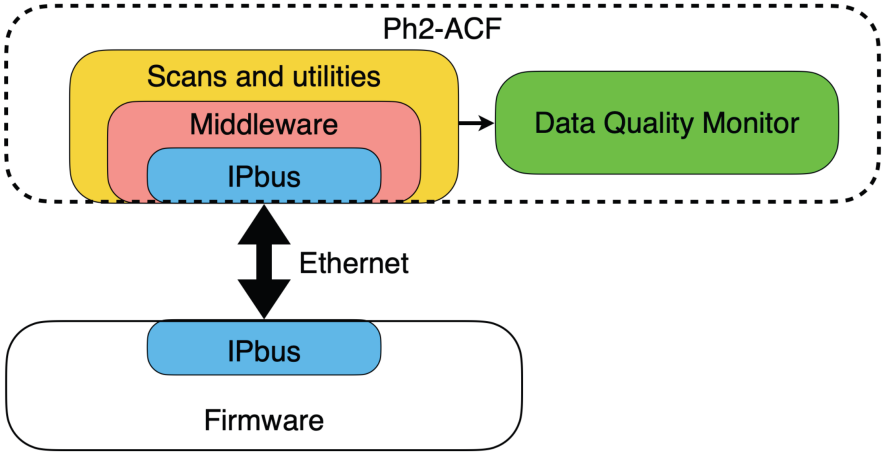


Fig. 54: Ph2-ACF software architecture [24].

and monitoring. The Data Quality Monitoring (DQM) component converts the measured data into histograms and summary plots that can be used to evaluate the quality of chips, modules, and readout links.

The framework represents the hardware using a hierarchical detector description. Objects corresponding to boards, optical or electrical links, modules, and readout chips are created from XML configuration files and stored in container structures that mirror the detector layout. The same hierarchy is used to store calibration results, such as occupancies, thresholds, gain parameters, masked pixels, or monitoring quantities. This design makes it possible to run the same calibration code on a single chip, a module, or a larger multi-module setup.

A central feature of Ph2-ACF is the separation between hardware-specific communication and calibration logic. Low-level interface classes handle register access, command transmission, and data reception, while calibration classes operate on abstract detector objects. As a result, many scans can be reused for different chip generations or readout configurations with only minimal changes. Ph2-ACF also supports multi-threaded execution, which allows calibration, data decoding, and monitoring tasks to run concurrently while protecting hardware access through synchronization mechanisms.

The author contributed to Ph2-ACF through debugging, testing, development of calibration and diagnostic procedures, and validation of the framework on prototype Inner Tracker hardware. In particular, the Data Transmission Test described below was developed by the author. The author also wrote most of the Ph2-ACF User Guide [160], from which parts of the calibration descriptions

in this thesis were adapted.

19.1. Calibrations

The calibration procedures implemented in Ph2-ACF are used to configure the readout chip, verify that the pixel matrix and data links operate correctly, and determine the settings needed for stable data taking. Most calibrations follow the same basic pattern. A selected chip register or calibration parameter is varied, a known test charge is injected into the pixel front-end by the internal charge-injection circuit, triggers are sent by the firmware, and the chip response is read out and summarized as occupancies, signal amplitude (ToT) values, or link-quality metrics.

The injected charge is controlled through the calibration DACs of the readout chip. In many scans, the relevant quantity is

$$\Delta\texttt{VCal} = \texttt{VCAL\_HIGH} - \texttt{VCAL\_MED}, \tag{19.1}$$

which sets the amplitude of the internal calibration pulse. The conversion between $\Delta\texttt{VCal}$ and the injected charge in electrons is chip-dependent and is determined later in this section using both electrical and X-ray calibration measurements.

Several calibrations are particularly important for the results presented in Sec. 20. The first is the PixelAlive scan, which verifies that each pixel responds to a charge injection. A charge well above the expected threshold is injected into each pixel, and the fraction of successful readouts is measured. Pixels with low response efficiency can be identified and masked, while noisy pixels can be detected if they produce more hits than expected. An example occupancy map produced by the PixelAlive scan is shown in Fig. 55. Because of its simple structure, PixelAlive also serves as the basis for several other scans that only differ in which chip setting is varied.

The S-curve scan measures the response efficiency of each pixel as a function of injected charge. For each value of $\Delta\texttt{VCal}$, the same pixel is injected several times, and its hit efficiency is measured. The resulting efficiency curve has the characteristic shape of an error function, as shown in Fig. 56. The charge at which the efficiency reaches 50% is interpreted as the pixel threshold, while the width of the transition region provides an estimate of the equivalent noise. S-curve measurements are therefore used to quantify the noise distribution and threshold dispersion of the chip.

Threshold tuning is performed in two steps. First, the global threshold DAC, denoted here as GDAC, is adjusted to place the average pixel threshold near the

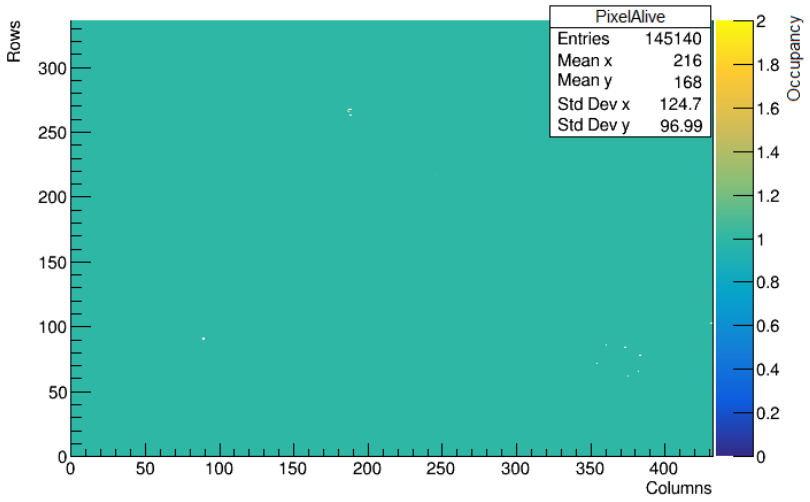


Fig. 55: Example of a two-dimensional pixel occupancy map produced by the PixelAlive scan.

desired value. Then, the per-pixel trimming DACs, denoted as TDAC, are tuned to reduce the pixel-to-pixel threshold spread. The strength of the trimming correction is controlled by the LDAC register. Conceptually, the threshold of a given pixel V_{thr}^{i} can be written as the sum of a global component V_{mean}, a pixel-dependent variation caused by production differences V_{rnd}^{i}, and a trimming correction V_{trim}^{i}:

$$V_{\text{thr}}^{i} = V_{\text{mean}}(\texttt{GDAC}) + V_{\text{rnd}}^{i} - V_{\text{trim}}^{i}(\texttt{TDAC}^{i}, \texttt{LDAC}). \tag{19.2}$$

The purpose of threshold equalization is to choose the TDAC values such that the pixel-dependent variations are compensated as well as possible.

The Gain scan measures the relation between injected charge and ToT. For each pixel, calibration pulses with different ΔVCal values are injected, and the corresponding ToT values are recorded. In the linear front-end, the response can be approximated by

$$\text{ToT}(\Delta\texttt{VCal}) = a\,\Delta\texttt{VCal} + b, \tag{19.3}$$

where a is the gain slope and b is the intercept. An example of the resulting gain curves is shown in Fig. 57. The gain depends strongly on the Krummenacher current I_{Krum}, which controls the discharge of the feedback capacitor in the preamplifier. Tuning this current changes the dynamic range and ToT resolution of the chip.

For data taking with external particles or radioactive sources, Ph2-ACF provides the Physics scan. In this mode, the chip is configured for normal

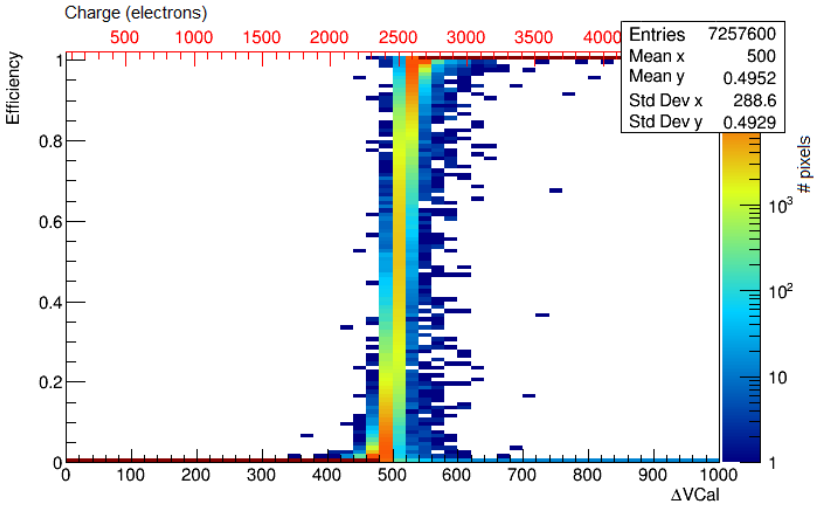


Fig. 56: Example of superimposed S-curves for all pixels on a chip, produced by the S-curve scan.

readout, and data are collected from external triggers, internally generated triggers, or self-trigger-like configurations, depending on the setup. The Physics scan records the position and ToT of the detected hits and can store the raw binary data for later event-by-event analysis. This mode is used in Sec. 20 for the X-ray calibration measurement with a ^{241}Am source.

The framework also includes several additional calibration and diagnostic procedures that are important for module operation, although they are not discussed in detail here. These include latency and injection-delay scans, threshold minimization, gain optimization, voltage tuning, bit-error-rate tests, data-readback optimization, optical-link diagnostics, and VTRx+ light-yield scans. The Data Transmission Test, developed by the author, is one such diagnostic scan. It quantifies the quality of a data link by varying the output-driver amplitude setting and locating the transition between failed and error-free transmission. These procedures provide the technical basis for the stable operation of prototype modules and for the validation studies presented in the next section.

20. Prototype testing results

This section presents selected results obtained with the Ph2-ACF framework during the characterization of prototype CMS pixel readout chips. The tests were performed on SCC equipped with CROCv1 and RD53A readout chips. Most measurements were carried out in the laboratory setup in Building 186

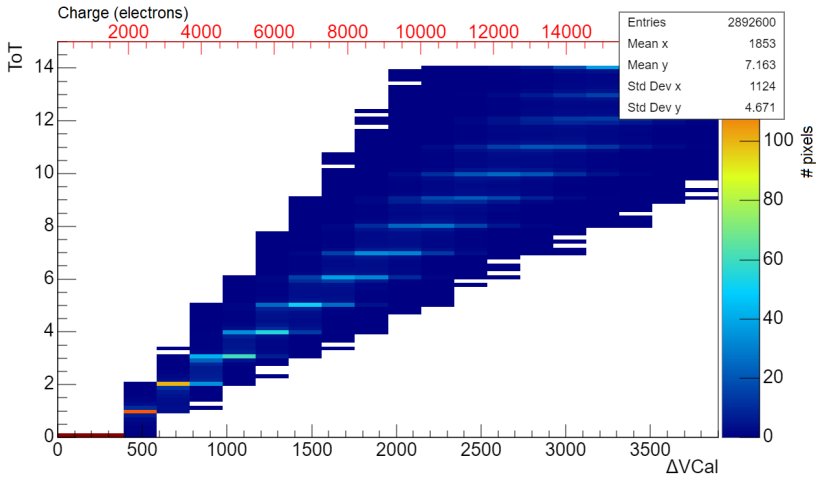


Fig. 57: Example of superimposed gain curves for all pixels on a chip, produced by the Gain scan without ToT compression.

at CERN, while a subset was performed in the Tracker Endcap Pixel detector (TEPX) module testing laboratory at the Experimental Nuclear and Particle Physics Center at Vilnius University.

The test setup used at Vilnius University is shown in Fig. 58 and is analogous to the one used at CERN. It consisted of a Linux computer running Ph2-ACF, an FC7 back-end board connected to the computer via Ethernet, a custom FMC card installed in the L12 slot of the FC7, an SCC connected to the FMC through a DisplayPort cable, a low-voltage power supply for powering the SCC, and a high-voltage power supply for sensor biasing.

The measurements presented here were performed as part of the CROCv1 characterization and calibration program. The author contributed to this work through Ph2-ACF development, debugging, calibration validation, and prototype testing. The results shown in this section were measured and produced by the author, unless stated otherwise. These results have been published in Refs. [A2, B2, B3]. Additional published results to which the author contributed through calibration development, procedure validation, and Ph2-ACF testing can be found in Refs. [A3, A4].

The selected results focus on three aspects of the CROCv1 performance. First, the threshold tuning performance is studied and compared with the RD53A demonstrator chip. Second, the ToT response is measured as a function of threshold and Krummenacher current to verify the behavior of the linear front-end. Third, the calibration of the internal charge-injection circuit is determined

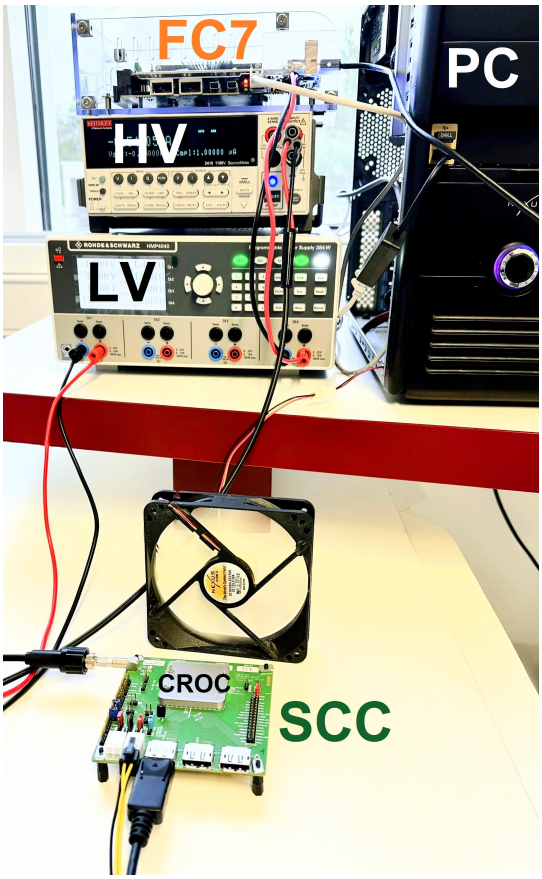


Fig. 58: Test setup used for prototype testing at Vilnius University. “FC7” denotes the back-end board, “PC” the Linux computer, “HV” the high-voltage power supply, “LV” the low-voltage power supply, “SCC” the Single-Chip Card, and “CROC” the CROCv1 readout chip mounted on the SCC.

using both electrical measurements and an independent ^{241}Am X-ray calibration. Unless stated otherwise, the results correspond to a non-irradiated CROCv1 chip bump-bonded to a planar sensor.

20.1. Threshold tuning

The hit-detection threshold of each pixel is controlled by a combination of global and per-pixel settings, as described conceptually in Eq. (19.2). The global threshold setting, denoted as `GDAC`, shifts the average threshold of the full pixel matrix, while the per-pixel trimming values, `TDAC`, compensate pixel-to-pixel variations. The trimming strength is controlled by the `LDAC` register.

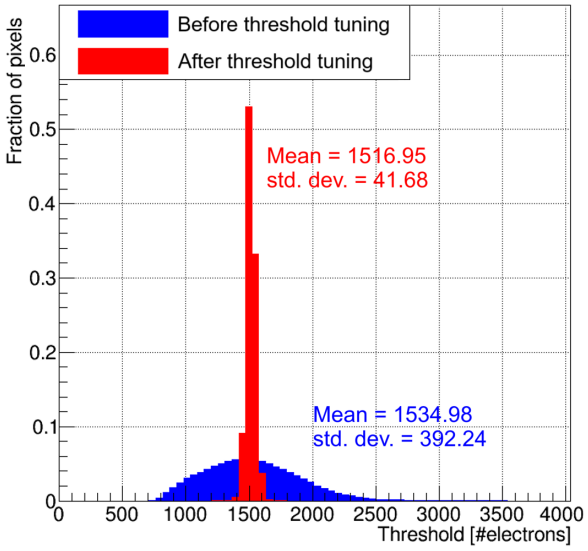


Fig. 59: Threshold distribution of a tested CROCv1 chip before (blue) and after (red) TDAC tuning at an average threshold of approximately 1500 electrons [B3].

The CROCv1 chip provides 5-bit TDAC trimming, one bit more than RD53A. This allows a finer compensation of pixel-to-pixel threshold variations. Figure 59 shows the threshold distribution of a tested CROCv1 chip before and after TDAC tuning at an average threshold of approximately 1500 electrons. The threshold distributions were measured with the S-curve scan, while the tuning itself was performed with the Threshold Equalization procedure implemented in Ph2-ACF.

After tuning, the threshold dispersion is reduced to about 40 electrons [B3]. This is approximately a factor of two improvement compared with the best threshold dispersion achieved with the RD53A demonstrator chip. The comparison is shown in Fig. 60, where the best achievable threshold dispersions for CROCv1 and RD53A are shown as functions of the LDAC register value [A2, B2, B3]. The scan was performed with a custom script that repeated the Threshold Equalization procedure at different LDAC settings and collected the resulting threshold dispersions from the Ph2-ACF output.

The stability of the threshold dispersion with respect to changes in the global threshold was also tested. Figure 61 shows the threshold distribution before and after changing the global threshold while keeping the TDAC values fixed. A 40% change in the global threshold changes the threshold dispersion by less than 10% without retuning [B3]. If the TDAC values are retuned after changing the global threshold, the dispersion can be restored to its optimal value of about

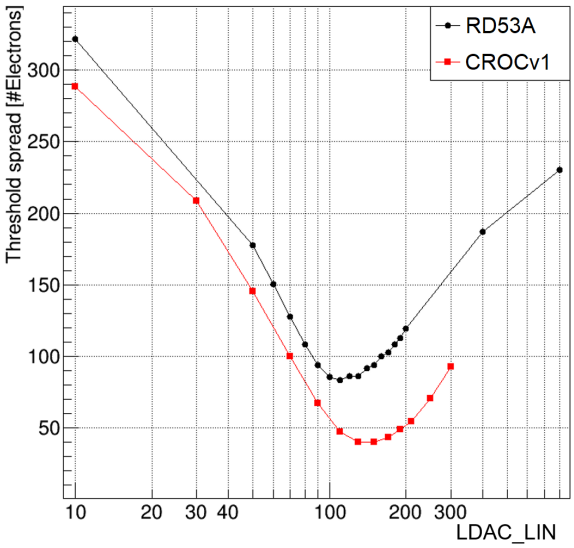


Fig. 60: Comparison of the threshold dispersion achieved with RD53A (black) and CROCv1 (red) chips at different LDAC values [A2, B2, B3].

40 electrons.

These measurements demonstrate that the CROCv1 threshold trimming system provides stable and precise threshold equalization. The achieved threshold dispersion satisfies the requirements for efficient operation at low threshold while keeping the pixel-to-pixel threshold spread under control.

20.2. Gain response

The signal amplitude in the RD53/CROC readout chip is measured using the Time-over-Threshold, defined as the number of 25 ns clock cycles during which the signal remains above threshold. The ToT response depends on the analog front-end settings, especially the Krummenacher current I_{Krum}, which controls the discharge of the preamplifier feedback capacitor. By changing I_{Krum}, one can tune the gain and choose a compromise between charge resolution and dynamic range.

The dependence of the ToT response on the hit-detection threshold was tested by repeating the Gain scan for several global threshold settings at a fixed Krummenacher current of 100 DAC units. The resulting average gain curves are shown in Fig. 62, both without and with 6-to-4-bit ToT compression. As expected for the linear analog front-end, the ToT gain slope is essentially independent of the absolute threshold value. Changing the threshold shifts the

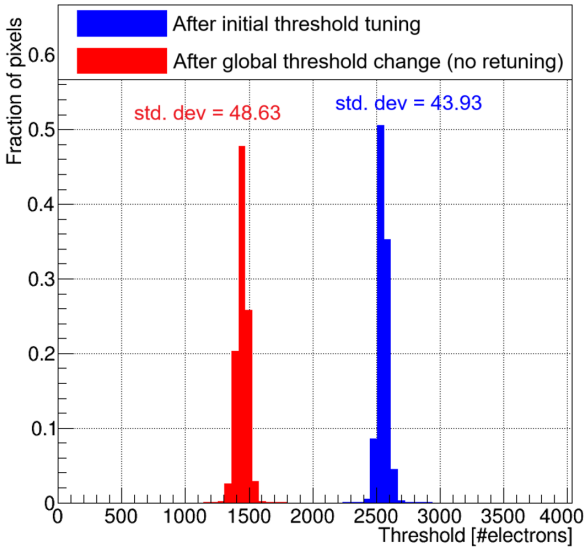


Fig. 61: Threshold distribution before (blue) and after (red) changing the global threshold while keeping the TDAC values fixed [B3].

point at which the signal becomes detectable, but does not significantly modify the slope of the response once the signal is above threshold.

The dependence on the Krummenacher current was studied by repeating the Gain scan for several I_{Krum} values at a fixed threshold of approximately 2000 electrons. The results are shown in Fig. 63. As expected, increasing I_{Krum} reduces the ToT gain slope and increases the dynamic range. The response becomes approximately linear for Krummenacher current values below about 60 DAC units, or for input charges sufficiently above threshold, consistent with the intended behavior of the linear front-end [155, B3].

For very large Krummenacher current values, the dynamic range becomes so large that the ToT compression region cannot be reached using the maximum charge available from the internal calibration circuit. This behavior confirms the expected role of I_{Krum} as the main global handle controlling the ToT gain. The observed independence of the gain slope from the hit-detection threshold is also important for calibration stability, because it allows threshold tuning and gain tuning to be treated largely independently.

20.3. Calibration of the injected charge

The S-curve and Gain scans use the chip's internal charge-injection circuit, where the injected charge is controlled by ΔVCal. To express measured thresh-

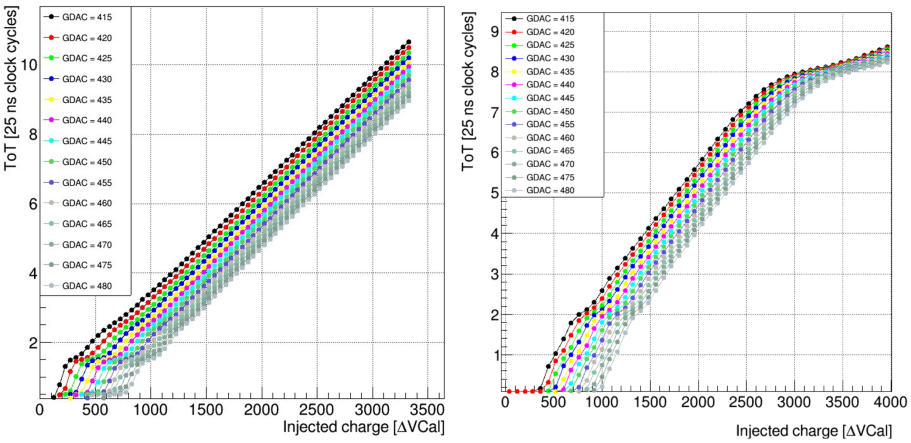


Fig. 62: Average ToT response as a function of injected charge for different hit-detection threshold settings. Left: without 6-to-4-bit compression; right: with 6-to-4-bit compression enabled. The Krummenacher current is fixed to 100 DAC units.

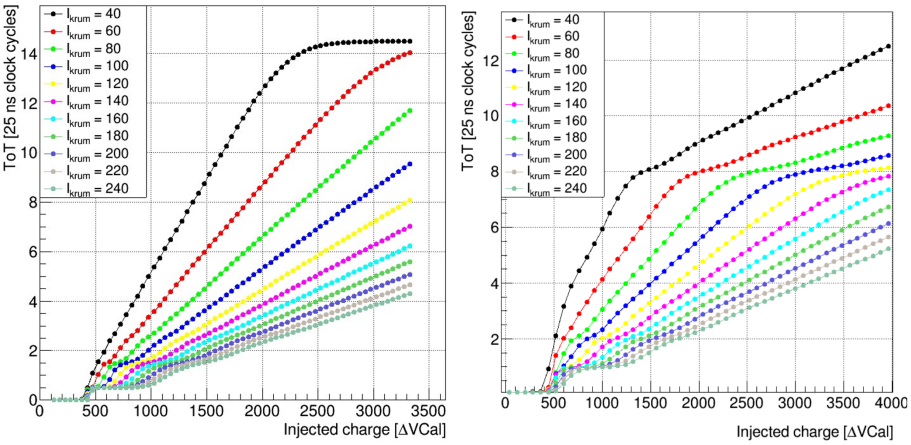


Fig. 63: Average ToT response as a function of injected charge for different Krummenacher current settings. Left: without 6-to-4-bit compression; right: with 6-to-4-bit compression enabled. The hit-detection threshold is fixed to approximately 2000 electrons [B3].

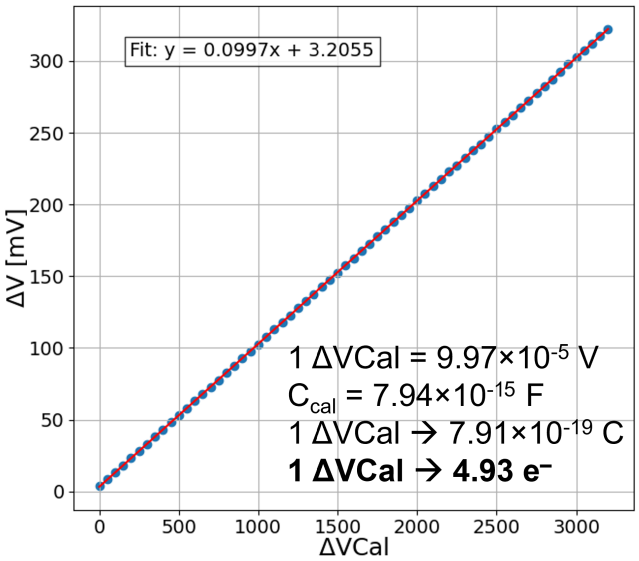


Fig. 64: Calibration voltage measured through the voltage multiplexer as a function of ΔVCal [B3].

olds and signal amplitudes in electrons, the conversion between ΔVCal and injected charge must be determined. Two independent methods were used: an electrical measurement through the chip voltage multiplexer (VMUX), and an X-ray calibration using a monochromatic ^{241}Am source.

In the VMUX method, the voltage corresponding to a given ΔVCal value is routed to the VMUX output and measured directly with a multimeter connected to the SCC. A custom script was used to configure different ΔVCal values through Ph2-ACF and to read the corresponding voltage using a modified version of the CMS POWDER library [161]. The measured voltage dependence is shown in Fig. 64. A linear fit gives a slope of approximately 0.1 mV per ΔVCal unit.

To convert this voltage into charge, the calibration injection capacitance was measured using the chip's built-in capacitance measurement circuit, also accessible through VMUX. For the tested chip, the calibration capacitance was found to be $C_{\text{cal}} \approx 7.94$ fF. Combining the measured voltage step and capacitance gives

$$Q_{\text{cal}} = \frac{C_{\text{cal}} \Delta V_{\text{Cal}}}{e} = \frac{7.94 \times 10^{-15}\ \text{F} \cdot 9.97 \times 10^{-5}\ \text{V}}{1.602 \times 10^{-19}\ \text{C}} = 4.93 \text{ electrons per } \Delta\text{VCal}. \quad (20.1)$$

An independent calibration was performed using a ^{241}Am X-ray source

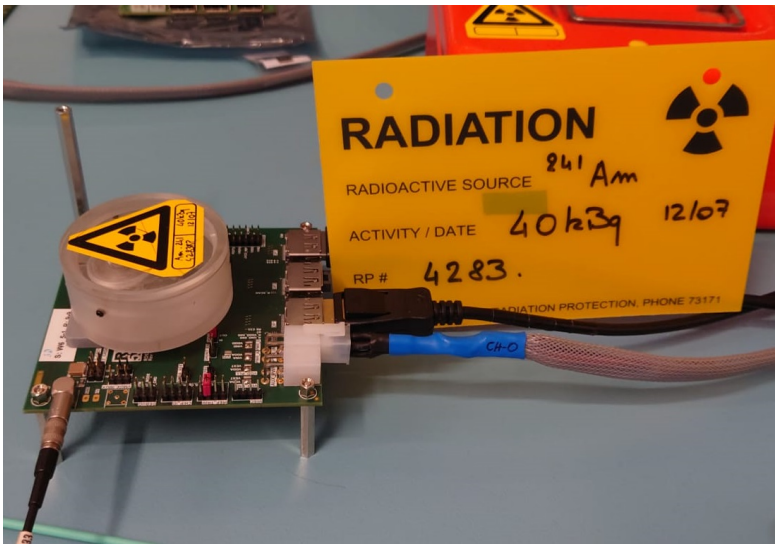


Fig. 65: A ^{241}Am source placed above the CROCv1 chip mounted on an SCC with a planar sensor during the X-ray calibration measurement.

placed above the SCC, as shown in Fig. 65. The source emits 59.54 keV photons, which deposit their energy locally in silicon and produce approximately 16 268 electron-hole pairs on average, assuming 3.66 eV per pair [162]. This provides a known charge signal independent of the internal charge-injection circuit.

The measurement was performed using the Physics scan in Ph2-ACF. A dedicated configuration used the chip's HitOR output to trigger the readout when a hit was detected in any pixel. When a trigger was received, the chip read out the hits associated with the corresponding 25 ns clock cycle. The positions and ToT values of the hits were then stored for event-by-event analysis.

The hit map of the recorded X-ray events is shown in Fig. 66. The ToT of each hit was converted to $\Delta\texttt{VCal}$ using the per-pixel gain calibration obtained from Eq. (19.3):

$$\Delta\texttt{VCal} = \frac{\mathrm{ToT} - b}{a}, \tag{20.2}$$

where a and b are the slope and intercept of the gain curve for the corresponding pixel. Hits in adjacent pixels were grouped into clusters, and the cluster charge was defined as the sum of the $\Delta\texttt{VCal}$ values of the constituent pixels. Only single-pixel clusters were used for the calibration to avoid complications from charge sharing between neighboring pixels.

The resulting $\Delta\texttt{VCal}$ distribution for single-pixel clusters is shown in Fig. 67. The peak was fitted with a Gaussian distribution to extract the mean signal amplitude, equal to 3263 $\Delta\texttt{VCal}$ units, and the standard deviation, equal to

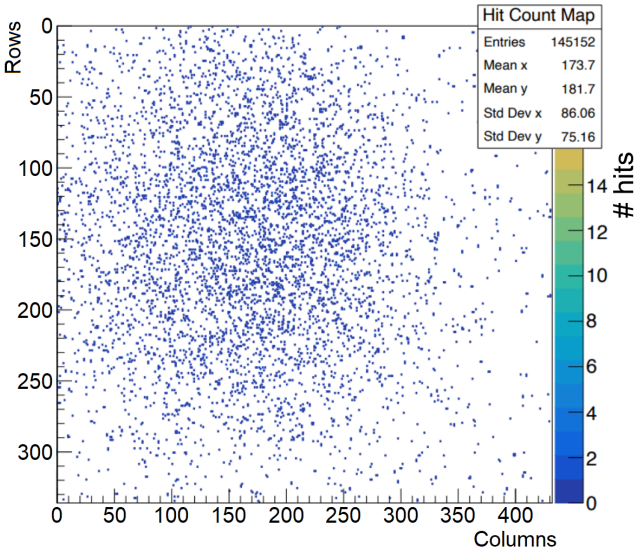


Fig. 66: Hit map of the recorded ^{241}Am X-ray events on the pixel matrix.

169.5 ΔVCal units. The width of the peak is dictated mainly by the limited ToT resolution, which is the largest contribution to the uncertainty of this measurement. Therefore, the standard deviation is used as a conservative estimate of the uncertainty on the measured ΔVCal unit value. Comparing the mean and the standard deviation of the fit with the expected charge of 16 268 electrons gives

$$Q_{\text{cal}} = 4.99 \pm 0.26 \text{ electrons per } \Delta\text{VCal}. \tag{20.3}$$

This result agrees with the VMUX-based calibration within uncertainties.

The two independent methods therefore give consistent conversion factors, and a value of 5 electrons per ΔVCal is adopted. This calibration is used throughout the CMS Phase–2 Tracker group for CROCv1 threshold and gain studies.

20.4. Summary of CROCv1 prototype performance

The measurements presented in this section demonstrate that the CROCv1 prototype satisfies several key expectations for the CMS Phase–2 Inner Tracker readout chip. The threshold-trimming system achieves a threshold dispersion of about 40 electrons, approximately a factor of two smaller than that obtained with the RD53A demonstrator chip. The threshold dispersion remains stable under moderate changes of the global threshold and can be restored to its optimal value by retuning the per-pixel TDAC values.

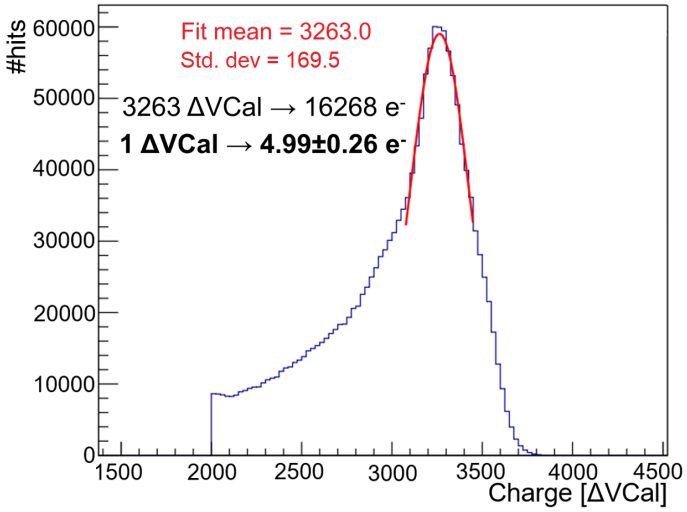


Fig. 67: Signal distribution of the recorded ^{241}Am X-ray hits. Only single-pixel clusters are considered. The signal is measured in ΔVCal units [B3].

The ToT gain response behaves as expected for the linear front-end. The gain slope is largely independent of the hit-detection threshold and is controlled primarily by the Krummenacher current. This confirms that the threshold and gain settings can be tuned with limited cross-dependence, simplifying the calibration strategy.

Finally, the internal charge-injection circuit was calibrated with two independent methods. The VMUX measurement gives 4.93 electrons per ΔVCal, while the ^{241}Am X-ray calibration gives 4.99 ± 0.26 electrons per ΔVCal. Their agreement validates the charge calibration and supports the use of 5 electrons per ΔVCal as the standard conversion factor.

At the time of writing, production of the final CMS pixel readout chip, CROCv2, is underway, and CMS Tracker module production using these chips has begun. The calibration procedures and validation studies described in this section provide part of the technical basis for the quality-assurance tests needed during module production and preparation for installation in the CMS detector.

PRECISION OF THE DRELL–YAN PROCESS MEASUREMENT WITH THE CMS PHASE–2 UPGRADE

To assess the impact of the CMS Phase–2 detector and the HL-LHC pileup environment on DY measurements, a dedicated MC performance study was carried out using a custom simulation chain within the CMSSW framework. Since the DY process, the Run–2 DY measurement, and the CMS Phase–2 upgrade have been discussed in Secs. 3, "Drell-Yan Process Measurement", and 14, respectively, only the ingredients directly relevant to this study are summarized here.

The HL-LHC will also provide a major increase in luminosity. The instantaneous luminosity is expected to increase by at least a factor of five relative to the nominal LHC design, meaning that at least five times more collision statistics can be collected for the same data-taking time.

Approximately 400 fb^{-1} of data are expected to be collected per year during the Phase–2 era, which is nearly three times the 137.6 fb^{-1} dataset used in the Run–2 DY measurement discussed in this thesis. The increased data sample will reduce statistical uncertainties independently of any detector-performance improvements. The present study therefore focuses specifically on detector-related effects, namely the muon reconstruction efficiency and the resolution of muon and dimuon observables.

21. Simulation setup

The custom simulation chain starts from the hard process $pp \to Z/\gamma^* \to \mu^+\mu^-$, generated in the same way as the signal samples used in the Run–2 DY differential cross section measurement described in Sec. "Drell-Yan Process Measurement". Events are generated with Powheg [55, 56] using the MiNNLO$_{PS}$ method, which provides NNLO QCD accuracy matched to a parton shower [68]. The hard process is interfaced to Pythia8 [48] and Photos [123] for parton showering, hadronization, and QED final-state radiation. The NNPDF 3.1 PDF set is used [124]. The detector simulation was performed with Geant4 [125]. A sample of 5 million events was produced and reconstructed using CMS Software (CMSSW) version 16.0.0.pre3.

For the Phase–2 detector configuration, the standard CMS Run–4 simulation setup was used. This includes the Phase–2 detector geometry, realistic detector conditions, the HL-LHC beam spot, and the CMS reconstruction configuration foreseen for Phase–2 studies. Detector aging corresponding to an integrated luminosity of 1000 fb^{-1} was included using the standard Phase–2

aging customization. To emulate the HL-LHC collision environment, events were overlaid with an average PU of 200 additional proton-proton interactions per bunch crossing, using minimum-bias events from the official Phase–2 release-validation samples.

The full standard CMS simulation and reconstruction chain was executed, including event generation, detector simulation, digitization, trigger emulation, raw-data formatting, offline reconstruction, and NanoAOD production. The study therefore probes the full detector and reconstruction response rather than only generator-level or fast-simulation effects.

The resulting custom Phase–2 sample was compared with a Run–2 reference based on the CMS Ultra Legacy 2018 campaign, which is also used in the Run–2 DY measurement discussed in Sec. "Drell-Yan Process Measurement". To ensure a fair comparison of detector and reconstruction effects, the Run–2 reference sample was reduced to 5 million events, matching the size of the custom Phase–2 sample. The comparison was performed using custom-developed ROOT macros.

Only the dimuon channel is considered in this study. The Phase–2 muon reconstruction chain is sufficiently mature to allow a meaningful comparison, while the corresponding electron reconstruction remains under active development. The results should therefore be interpreted as an early estimate of DY performance with the CMS Phase–2 detector. Further improvements may be expected as the detector description, reconstruction algorithms, and calibrations mature.

The defined quantities of interest in the present investigation were:

- Muon reconstruction efficiency as a function of muon $p_{\mathrm{T}} \in [15, 200)$ GeV;
- Muon p_{T} resolution in the range $p_{\mathrm{T}}^{\mu} \in [15, 200)$ GeV;
- Dimuon mass resolution in the range $M_{\mu\mu} \in [50, 1000)$ GeV;
- Dimuon p_{T} resolution in the range $p_{\mathrm{T}}^{\mu\mu} \in [0.1, 1000)$ GeV.

The choices for the ranges were dictated by the generated phase space of the MC sample, HLT options, and available statistics.

22. Muon reconstruction efficiency

The first quantity studied is the global-muon reconstruction efficiency. The efficiency is evaluated on a per-muon basis by matching dressed generator-level

muons to reconstructed global muons. Muons are required to satisfy

$$p_{\mathrm{T}}^{\mu} \geqslant 15\ \mathrm{GeV}, \qquad |\eta_{\mu}| < 2.4.$$

A dressed generator-level muon is considered successfully reconstructed if a global reconstructed muon is found within

$$\Delta R < 0.3, \quad \text{and} \quad \frac{|p_{\mathrm{T}}^{\mathrm{reco}} - p_{\mathrm{T}}^{\mathrm{gen}}|}{p_{\mathrm{T}}} < 0.5,$$

where superscripts "gen" and "reco" stand for the generator-level and the reconstructed quantity, respectively, and ΔR is the angular distance between the generator-level and the reconstructed muon, calculated using (5.3), described in Subsection 5.1.

The efficiency is studied as a function of the muon $p_{\mathrm{T}}^{\mathrm{gen}}$ separately in the barrel and endcap regions. The results are shown in Fig. 68. The global-muon reconstruction efficiency remains high in the Phase–2 environment. Relative to the Run–2 reference, the Phase–2 efficiency is lower by only about 1%, with the difference remaining small over the full kinematic range studied. This indicates that the upgraded detector and reconstruction chain preserve robust muon reconstruction even in the presence of HL-LHC pileup.

23. Muon and dimuon resolution

The second part of the study investigates the resolution of single-muon and dimuon observables. Events are selected using single-muon and dimuon trigger requirements analogous to those used in the Run–2 DY analysis. Reconstructed muons are required to be global muons satisfying

$$p_{\mathrm{T}}^{\mathrm{lead}} \geqslant 20\ \mathrm{GeV}, \qquad p_{\mathrm{T}}^{\mathrm{sublead}} \geqslant 15\ \mathrm{GeV}, \qquad |\eta_{\mu}| < 2.4, \tag{23.1}$$

as well as a relative tracker-isolation requirement below 15%. If more than two muons pass the selection, the OS muon pair with the highest invariant mass is chosen. The selected dimuon system is further required to satisfy $M_{\mu\mu} \geqslant 50$ GeV.

The muon transverse-momentum response is defined as

$$\frac{p_{\mathrm{T}}^{\mathrm{reco}} - p_{\mathrm{T}}^{\mathrm{gen}}}{p_{\mathrm{T}}^{\mathrm{gen}}}.$$

The dimuon invariant-mass and transverse-momentum responses are defined analogously:

$$\frac{M_{\mu\mu}^{\mathrm{reco}} - M_{\mu\mu}^{\mathrm{gen}}}{M_{\mu\mu}^{\mathrm{gen}}}, \qquad \frac{(p_{\mathrm{T}}^{\mu\mu})^{\mathrm{reco}} - (p_{\mathrm{T}}^{\mu\mu})^{\mathrm{gen}}}{(p_{\mathrm{T}}^{\mu\mu})^{\mathrm{gen}}}.$$

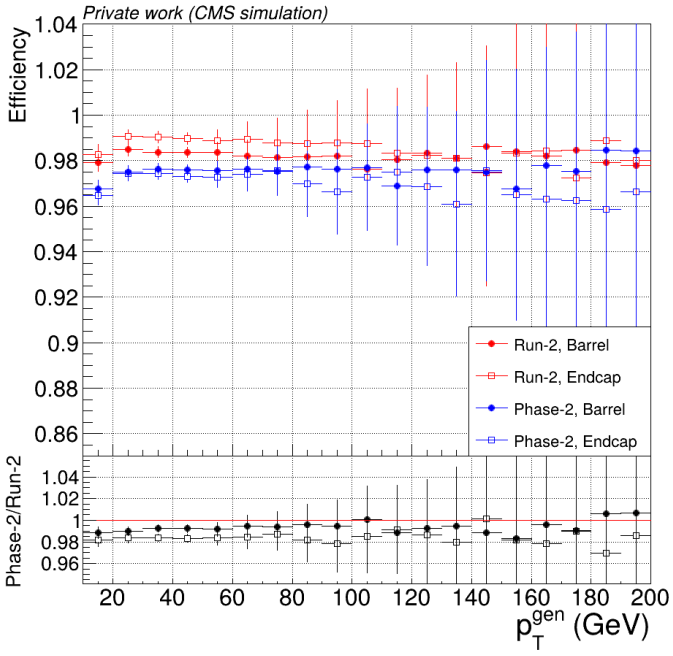


Fig. 68: Comparison of the global-muon reconstruction efficiency in Run–2 and Phase–2 as a function of the generator-level muon transverse momentum. The results are shown separately for the barrel ($|\eta| < 1.2$) and endcap regions. The lower panel shows the ratio of the Phase–2 and Run–2 efficiencies.

For each bin in the corresponding generator-level observable, the response distribution is fitted with a Gaussian function. The fitted mean is used as a measure of the reconstruction scale, while the fitted width σ is used as the resolution estimate.

The results are shown in Figs. 69, 70, and 71 for the muon p_{T}, dimuon invariant mass, and dimuon p_{T}, respectively. In these figures, the marker points show the fitted mean values, and the shaded bands indicate the fitted Gaussian widths. In both the Run–2 and Phase–2 samples, the fitted mean response remains close to zero, indicating that no large reconstruction bias is introduced. The main difference is observed in the resolution. The Phase–2 sample shows an improved muon p_{T} resolution over most of the studied kinematic range, with a typical improvement of approximately 30–40%. The dimuon invariant-mass resolution improves by about 20–40%, depending on the mass region, while the dimuon p_{T} resolution improves by about 15–40%. For the most part, these improvements are made possible by the significantly improved tracker granularity and the MTD, which allows better separation between tracks associated with the primary vertex and those originating from PU interactions.

24. Conclusion

The results presented in this section show that the CMS Phase–2 detector is expected to preserve robust dimuon reconstruction performance despite the much harsher HL-LHC environment. The global-muon reconstruction efficiency decreases by only about 1% relative to the Run–2 reference, even with an average PU of 200. At the same time, the upgraded detector and reconstruction chain provide significantly improved kinematic resolution for muon and dimuon observables.

For precision DY measurements, the improved muon momentum resolution is directly beneficial. It reduces detector smearing in the dilepton invariant mass and transverse momentum distributions, improves unfolding stability, and reduces bin migration. The improved dimuon invariant-mass resolution is particularly important for measurements of $\mathrm{d}\sigma/\mathrm{d}M_{\mu\mu}$, while the improved dimuon p_{T} resolution is important for measurements of $\mathrm{d}\sigma/\mathrm{d}p_{\mathrm{T}}^{\mu\mu}$ and multidimensional distributions such as $\mathrm{d}^2\sigma/\mathrm{d}M\,\mathrm{d}p_{\mathrm{T}}$.

The detector-performance improvements discussed here are independent of the much larger integrated luminosity expected at the HL-LHC. The increased data sample will further reduce statistical uncertainties, especially in the high-mass region where Run–2 measurements are statistically limited. The

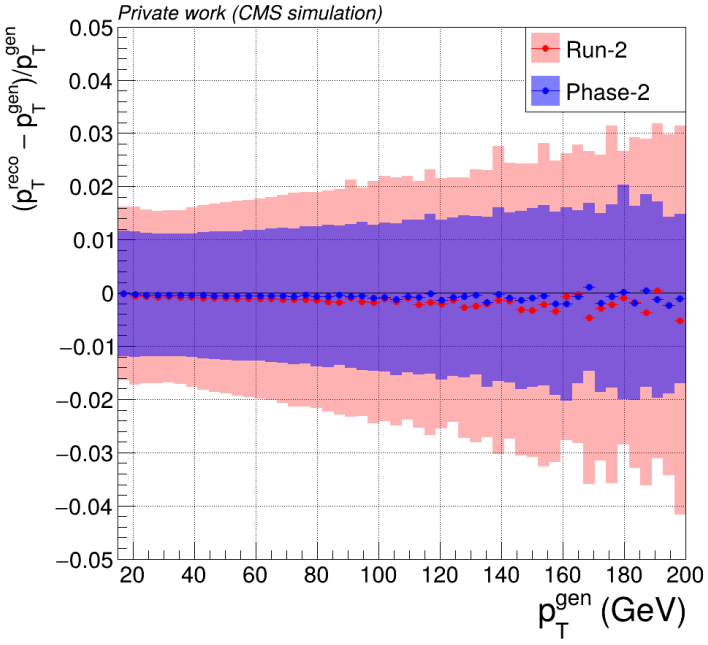


Fig. 69: Comparison of the reconstructed-to-generator response for the muon transverse momentum in Run–2 and Phase–2 as a function of the generator-level dressed muon p_{T}. In each generator-level bin, the response distribution is fitted with a Gaussian function. The marker points show the fitted mean, while the shaded bands indicate the fitted width σ, taken as the resolution.

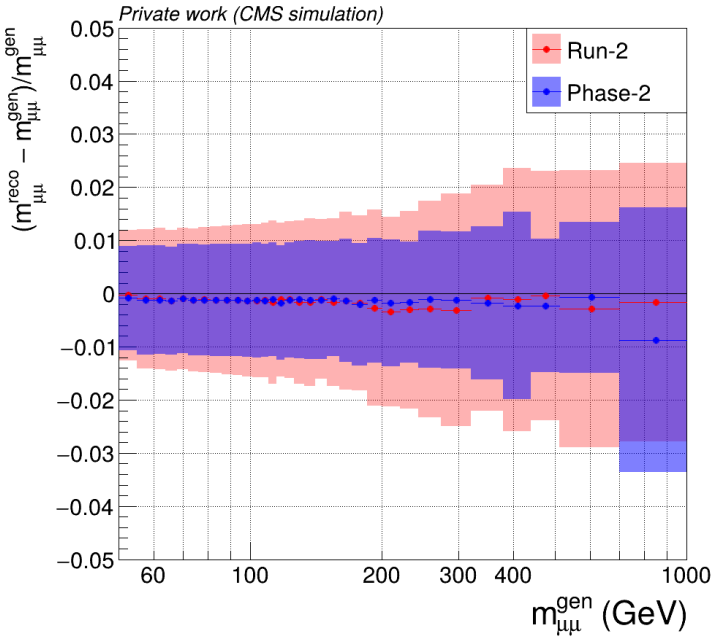


Fig. 70: Comparison of the reconstructed-to-generator response for the dimuon invariant mass in Run–2 and Phase–2 as a function of the generator-level $M_{\mu\mu}$. In each generator-level bin, the response distribution is fitted with a Gaussian function. The marker points show the fitted mean, while the shaded bands indicate the fitted width σ, taken as the resolution.

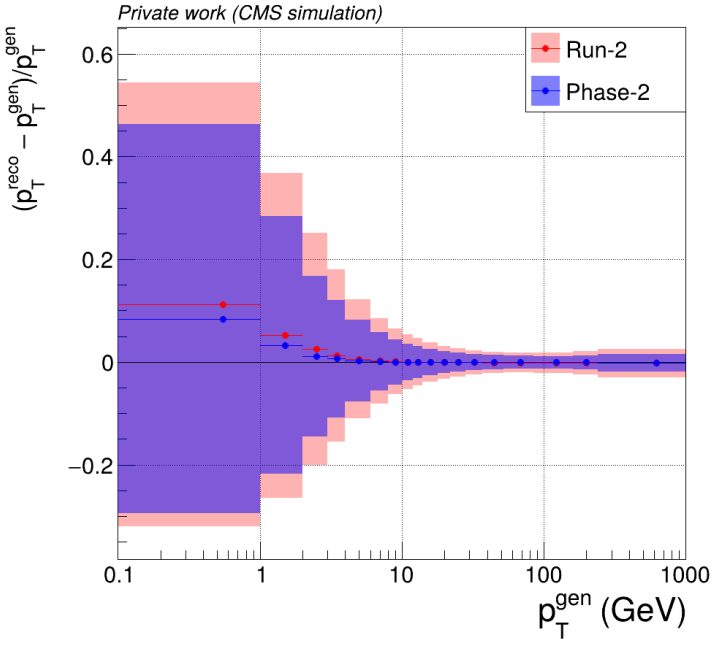


Fig. 71: Comparison of the reconstructed-to-generator response for the dimuon transverse momentum in Run–2 and Phase–2 as a function of the generator-level $p_T^{\mu\mu}$. In each generator-level bin, the response distribution is fitted with a Gaussian function. The marker points show the fitted mean, while the shaded bands indicate the fitted width σ, taken as the resolution.

combination of stable muon reconstruction efficiency, improved momentum and mass resolution, and substantially larger event yields indicates that the dimuon channel will remain a powerful and experimentally robust channel for precision DY measurements in the HL-LHC era.

CONCLUSIONS

This thesis presented work on precision measurements of the Drell–Yan process with the CMS detector and on the preparation of the CMS Inner Tracker for the Phase–2 upgrade. The two parts are connected by the common goal of improving the precision and robustness of measurements at the LHC and the HL-LHC. The first part focused on the measurement of the DY differential cross section as a function of the dilepton invariant mass using Run–2 CMS data. The second part focused on the testing and calibration of prototype pixel readout chips for the CMS Phase–2 Inner Tracker. Finally, a dedicated simulation study was used to assess how the upgraded CMS detector may affect the precision of future DY measurements in the HL-LHC environment.

The theoretical and experimental foundations relevant for the thesis were first reviewed. The SM was introduced as the theoretical framework describing the electromagnetic, weak, and strong interactions, together with the Higgs mechanism responsible for electroweak symmetry breaking. The DY process was discussed as one of the most important standard-candle processes at hadron colliders. Its clean dilepton final state, high theoretical precision, and sensitivity to PDFs make it a powerful probe of perturbative QCD, electroweak physics, and possible effects of physics beyond the SM. The practical ingredients needed for DY predictions and measurements were also discussed, including cross sections, Feynman diagrams, higher-order corrections, PDFs, and MC simulation.

The experimental framework of the thesis was then introduced through a description of the LHC and the CMS detector. The CMS detector combines a high-precision silicon tracker, electromagnetic and hadronic calorimeters, a strong solenoidal magnetic field, and a dedicated muon system. The Particle Flow reconstruction strategy allows information from all subdetectors to be combined into reconstructed physics objects such as electrons, muons, jets, and missing transverse momentum. These reconstruction capabilities are essential for precision DY measurements in both the dielectron and dimuon channels. The trigger and data acquisition systems provide the necessary event selection and data reduction, making it possible to record events of interest from the high collision rate delivered by the LHC.

The main analysis contribution of this thesis was to the CMS Run–2 measurement of the DY differential cross section as a function of $M_{\ell\ell}$ at $\sqrt{s} = 13$ TeV, using the full 2016–2018 dataset corresponding to an integrated luminosity of 137.6 fb^{-1}. The analysis was performed in the dielectron and dimuon channels and measured the cross section in the fiducial mass range $40 < M_{\ell\ell} < 3000$ GeV. The author's primary contribution was the development, imple-

mentation, validation, and uncertainty evaluation of the background-estimation methods. The full analysis is documented in the CMS analysis note AN-22-177 [B1].

A central part of the DY measurement was the estimation of backgrounds with misidentified/non-prompt leptons. These backgrounds arise mainly from QCD multijet and W+jets processes and were estimated using same-sign control regions. In the dimuon channel, a two-dimensional OS/SS transfer factor $R_{\mathrm{OS/SS}}(M_{\mu\mu}, p_{\mathrm{T}}^{\mathrm{miss}})$ was developed to account for the changing mixture of W+jets and QCD multijet events. This transfer factor was derived in a data sideband where exactly one muon fails the isolation requirement and was parameterized with an empirical function of $M_{\mu\mu}$ and $p_{\mathrm{T}}^{\mathrm{miss}}$. Closure tests showed no systematic non-closure in the mass distribution. The resulting systematic uncertainty remained below 50% for $M_{\mu\mu} < 900$ GeV, while the total uncertainty, obtained from the statistical and systematic components, was kept below 50% for $M_{\mu\mu} < 400$ GeV.

The dielectron channel required an additional treatment because electron charge misidentification causes prompt DY events to enter the same-sign control region. Charge misidentification weights were applied to simulated DY events before subtracting prompt backgrounds from the same-sign data. The remaining excess was interpreted as the misidentified/non-prompt electron background. A one-dimensional $R_{\mathrm{OS/SS}}(p_{\mathrm{T}}^{\mathrm{miss}})$ transfer factor was used, and the residual artifact around the Z peak was replaced by a linear interpolation between neighboring mass regions. This procedure produced a smooth and physically motivated estimate of the misidentified/non-prompt electron background, with the systematic uncertainty below 40% for $M_{ee} < 900$ GeV and the total uncertainty below 40% for $M_{ee} < 400$ GeV. Together, these results support **thesis statement no. 1: The presented variant of the same-sign method allows the misidentified or non-prompt lepton background to be estimated in both dielectron and dimuon final states with a systematic uncertainty below 50% for dilepton masses below 900 GeV.** At higher dilepton masses, the uncertainty of the misidentified/non-prompt background estimate becomes increasingly limited by the number of events in the corresponding control regions. This supports **thesis statement no. 2: For dilepton masses above 300 GeV, the uncertainty in the misidentified or non-prompt lepton background estimate is dominated by the statistical uncertainty of the control samples.**

The top-quark background was corrected using a data-driven method based on the $e\mu$ control region. The simulation was observed to overestimate the data at high $M_{e\mu}$, a behavior consistent with earlier CMS studies of top-quark production. A modified $e\mu$ method was therefore used to derive a mass-dependent

correction to simulated $t\bar{t}$, tW, and $\bar{t}W$ events. The correction was parameterized as a linear function of $\log_{10}(M_{\ell\ell})$ and applied to the same-flavor channels. The associated systematic uncertainty was found to remain below 6% over the full mass range, providing a controlled correction to one of the dominant high-mass backgrounds.

After event selection, background subtraction, and corrections, the reconstructed dilepton mass spectra were unfolded using `TUnfold`. The unfolding response was derived from simulated signal events and used to combine the four Run–2 data-taking eras into one common fiducial spectrum per channel. The dielectron and dimuon spectra were then combined using a hybrid weighted-mean and BLUE-style covariance-propagation method. The resulting combined fiducial cross section is the main, least model-dependent output of the measurement, while the full phase-space result was obtained using acceptance corrections derived from simulation. The measured spectra were generally compatible with the SM prediction from POWHEG MiNNLO within uncertainties.

The second major part of the thesis addressed the CMS Phase–2 upgrade, motivated by the HL-LHC environment. The HL-LHC will provide a much larger integrated luminosity, up to about 4000 fb^{-1} over its full program, but it will also introduce much harsher detector conditions. The average pileup is expected to reach about 140–200 interactions per bunch crossing, and the radiation levels in the innermost tracker regions will be substantially higher than in the present detector. To maintain CMS physics performance, the Phase–2 upgrade includes a complete replacement of the Tracker, the installation of the HGCAL, upgrades to the muon system, the addition of the MTD, and major upgrades of the trigger, DAQ, and computing systems.

The thesis focused in particular on the CMS Phase–2 Inner Tracker and its pixel readout chip. The upgraded Inner Tracker will use radiation-hard silicon pixel sensors and readout electronics capable of operating under very high hit rates and radiation doses. The RD53 Collaboration developed the pixel readout chip family for this purpose. The CMS-specific RD53B-CMS prototype, also known as CROCv1, was used to validate the readout architecture, calibration procedures, and analog front-end performance before the final production chip, CROCv2.

A significant part of the author's work was related to the Ph2-ACF software framework and the calibration of prototype readout chips. Ph2-ACF provides the control, configuration, calibration, data-taking, and monitoring infrastructure used for CMS Phase–2 Tracker module testing. The author contributed to debugging, testing, calibration development, diagnostic procedures, and user documentation, including the Ph2-ACF User Guide [160]. The Data Transmis-

sion Test calibration was developed by the author to quantify the quality of data links by locating the transition between failed and error-free transmission.

Prototype testing results were presented for CROCv1 and, where relevant, compared with RD53A. Threshold tuning studies showed that CROCv1 can achieve a threshold dispersion of about 40 electrons after per-pixel `TDAC` tuning. This is approximately a factor of two smaller than the best dispersion obtained with RD53A. The threshold dispersion was also found to remain stable under moderate changes of the global threshold and could be restored to its optimal value by retuning the per-pixel `TDAC` values. These results support **thesis statement no. 3: The RD53B-CMS pixel readout chip prototype can be tuned to achieve a threshold dispersion approximately two times lower than that obtained with the RD53A demonstrator chip.**

The gain response of CROCv1 was studied by measuring the ToT response as a function of injected charge, hit-detection threshold, and Krummenacher current. The gain slope was found to be largely independent of the hit-detection threshold, while the Krummenacher current controlled the dynamic range and ToT response, as expected from the design of the linear analog front-end [B3]. These results support **thesis statement no. 4: The signal gain of the RD53B-CMS chip shows linear behavior and does not depend on the hit-detection threshold.**

The calibration of the internal charge-injection circuit was determined with two independent methods. In the VMUX method, the voltage corresponding to Δ`VCal` was measured directly and combined with the measured calibration capacitance, giving 4.93 electrons per Δ`VCal`. In the second method, a ^{241}Am X-ray source provided a known charge deposition of approximately 16 268 electrons in silicon. The measured X-ray signal gave 4.99 ± 0.26 electrons per Δ`VCal`. The agreement between the two methods validates the calibration of the charge-injection circuit. These results support **thesis statement no. 5: The VCal digital-to-analog converter of the RD53B-CMS chip provides an injected-charge step size of approximately 5 electrons.** These results have been reported in Refs. [A2, B2, B3], while related CMS Tracker upgrade studies to which the author contributed are presented in Refs. [A3, A4].

Finally, the thesis studied the impact of the Phase–2 CMS detector on future DY measurements using a dedicated custom MC simulation chain. A Phase–2 $pp \to Z/\gamma^* \to \mu^+\mu^-$ sample was generated and reconstructed with realistic Run–4 detector settings, detector aging corresponding to 1000 fb^{-1}, and an average pileup of 200 interactions per bunch crossing. The sample was compared with a Run–2 Ultra Legacy reference using the same hard process and the same number of events. The study focused on the dimuon channel, where

the Phase–2 reconstruction is sufficiently mature for a meaningful comparison.

The Phase–2 global-muon reconstruction efficiency was found to remain very high, with only about a 1% decrease relative to the Run–2 reference. At the same time, the reconstructed muon and dimuon resolutions improved substantially. The muon p_{T} resolution improved by approximately 30–40%, the dimuon invariant-mass resolution by about 20–40%, and the dimuon p_{T} resolution by about 15–40%, depending on the kinematic region. These results support **thesis statement no. 6: In addition to an expected increase of at least a factor of five in the available DY statistics relative to Run–2, the upgraded CMS detector is expected to improve the muon transverse momentum resolution by at least 30% and the dimuon mass resolution by at least 20% in the studied phase space.**

The detector-performance improvements studied in this thesis are independent of the increase in available luminosity at the HL-LHC. The increased dataset will reduce statistical uncertainties, especially in the high-mass region where Run–2 DY measurements are statistically limited. The combination of much larger event yields, stable muon reconstruction efficiency, and improved muon and dimuon kinematic resolution indicates that the dimuon DY channel will remain a powerful and experimentally robust probe for precision SM measurements and searches for deviations from SM predictions in the HL-LHC era.

In summary, the work presented in this thesis contributes to both the current and future precision DY program of CMS. The Run–2 analysis work provided data-driven background-estimation methods and uncertainty evaluations for a wide-mass-range DY differential cross-section measurement. The Phase–2 Tracker work validated important calibration and performance aspects of the RD53B-CMS pixel readout chip and contributed to the software tools needed for module testing. The Phase–2 DY simulation study showed that the upgraded CMS detector is expected to preserve excellent muon reconstruction performance and improve the resolution of key DY observables despite the challenging HL-LHC environment. Together, these results demonstrate the continuity between present CMS precision measurements, detector-upgrade preparation, and the future HL-LHC physics program.

REFERENCES

[1] UA1 Collaboration, "Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at s = 540 GeV", *Physics Letters B*, vol. 122, no. 1, pp. 103–116, 1983, ISSN: 0370-2693. DOI: 10.1016/0370-2693(83)91177-2.

[2] UA2 Collaboration, "Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN pp Collider", *Physics Letters B*, vol. 122, no. 5, pp. 476–485, 1983, ISSN: 0370-2693. DOI: 10.1016/0370-2693(83)91605-2.

[3] UA1 Collaboration, "Experimental Observation of Lepton Pairs of Invariant Mass Around 95 GeV/c^2 at the CERN SPS Collider", *Phys. Lett. B*, vol. 126, pp. 398–410, 1983.
DOI: 10.1016/0370-2693(83)90188-0.

[4] UA2 Collaboration, "Evidence for $Z^0 \to e^+e^-$ at the CERN pp Collider", *Physics Letters B*, vol. 129, no. 1, pp. 130–140, 1983, ISSN: 0370-2693. DOI: 10.1016/0370-2693(83)90744-X.

[5] CDF Collaboration, "Observation of Top Quark Production in $p\overline{p}$ Collisions with the Collider Detector at Fermilab", *Physical Review Letters*, vol. 74, no. 14, pp. 2626–2631, 1995, ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.74.2626.

[6] D0 Collaboration, "Observation of the Top Quark", *Physical Review Letters*, vol. 74, no. 14, pp. 2632–2637, 1995, ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.74.2632.

[7] ATLAS Collaboration, "Observation of a New Particle in the Search for the Standard Model Higgs Boson with the ATLAS Detector at the LHC", *Physics Letters B*, vol. 716, no. 1, pp. 1–29, 2012, ISSN: 0370-2693. DOI: 10.1016/j.physletb.2012.08.020.

[8] CMS Collaboration, "Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC", *Physics Letters B*, vol. 716, no. 1, pp. 30–61, 2012, ISSN: 0370-2693. DOI: 10.1016/j.physletb.2012.08.021.

[9] L. Evans and P. Bryant, "LHC Machine", *Journal of Instrumentation*, vol. 3, no. 08, S08001, 2008, ISSN: 1748-0221.
DOI: 10.1088/1748-0221/3/08/S08001.

[10] CMS Collaboration, "The CMS Experiment at the CERN LHC", *Journal of Instrumentation*, vol. 3, no. 08, S08004, 2008, ISSN: 1748-0221. DOI: 10.1088/1748-0221/3/08/S08004.

[11] ATLAS Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider", *Journal of Instrumentation*, vol. 3, no. 08, S08003, 2008. DOI: 10.1088/1748-0221/3/08/S08003.

[12] LHCb Collaboration, "The LHCb Detector at the LHC", *Journal of Instrumentation*, vol. 3, no. 08, S08005, 2008, ISSN: 1748-0221. DOI: 10.1088/1748-0221/3/08/S08005.

[13] ALICE Collaboration, "The ALICE Experiment at the CERN LHC", *Journal of Instrumentation*, vol. 3, no. 08, S08002, 2008, ISSN: 1748-0221. DOI: 10.1088/1748-0221/3/08/S08002.

[14] S. D. Drell and T.-M. Yan, "Massive Lepton-Pair Production in Hadron-Hadron Collisions at High Energies", *Physical Review Letters*, vol. 25, no. 5, pp. 316–320, 1970. DOI: 10.1103/PhysRevLett.25.316, Corrected in "Massive Lepton-Pair Production in Hadron-Hadron Collisions at High Energies", *Physical Review Letters*, vol. 25, no. 13, pp. 902–902, 1970. DOI: 10.1103/PhysRevLett.25.902.2.

[15] CMS Collaboration, "Measurement of the Differential and Double-Differential Drell-Yan Cross Sections in Proton-Proton Collisions at $\sqrt{s} = 7$ TeV", *Journal of High Energy Physics*, vol. 2013, no. 12, p. 30, 2013, ISSN: 1029-8479. DOI: 10.1007/JHEP12(2013)030.

[16] ATLAS Collaboration, "Measurement of the Low-Mass Drell-Yan Differential Cross Section at $\sqrt{s} = 7$ TeV Using the ATLAS Detector", *Journal of High Energy Physics*, vol. 2014, no. 6, p. 112, 2014, ISSN: 1029-8479. DOI: 10.1007/JHEP06(2014)112.

[17] CMS Collaboration, "Measurements of Differential and Double-Differential Drell-Yan Cross Sections in Proton-Proton Collisions at 8 TeV", *European Physical Journal C*, vol. 75, no. 4, p. 147, 2015, ISSN: 1434-6044, 1434-6052. DOI: 10.1140/epjc/s10052-015-3364-2.

[18] ATLAS Collaboration, "Measurement of the Drell-Yan Triple-Differential Cross Section in pp Collisions at $\sqrt{s} = 8$ TeV", *Journal of High Energy Physics*, vol. 2017, no. 12, p. 59, 2017, ISSN: 1029-8479. DOI: 10.1007/JHEP12(2017)059.

[19] CMS collaboration, "Measurement of the Differential Drell-Yan Cross Section in Proton-Proton Collisions at $\sqrt{s} = 13$ TeV", *Journal of High Energy Physics*, vol. 2019, no. 12, p. 59, 2019, ISSN: 1029-8479. DOI: 10.1007/JHEP12(2019)059.

[20] ATLAS Collaboration, "Measurement of the Transverse Momentum Distribution of Drell-Yan Lepton Pairs in Proton-Proton Collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector", *European Physical Journal C*, vol. 80, no. 7, p. 616, 2020, ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-020-8001-z.

[21] G. Apollinari *et al.*, "High-Luminosity Large Hadron Collider (HL-LHC): Preliminary Design Report", CERN Yellow Reports: Monographs, 2015. DOI: 10.5170/CERN-2015-005.

[22] D. Contardo *et al.*, "Technical Proposal for the Phase-II Upgrade of the CMS Detector", CERN-LHCC-2015-010, LHCC-P-008, CMS-TDR-15-02. DOI: 10.17181/CERN.VU8I.D59J.

[23] J. Chistiansen and M. Garcia-Sciveres, "RD Collaboration Proposal: Development of Pixel Readout Integrated Circuits for Extreme Rate and Radiation", CERN-LHCC-2013-008, LHCC-P-006, 2013. Available: https://cds.cern.ch/record/1553467.

[24] M. E. Dinardo, "The CMS Inner Tracker DAQ System for the High Luminosity Upgrade of LHC: From Single-Chip Testing, to Large-Scale Assembly Qualification", *EPJ Web of Conferences*, vol. 295, p. 02 028, 2024, ISSN: 2100-014X. DOI: 10.1051/epjconf/202429502028.

[25] R. P. Feynman, "Mathematical Formulation of the Quantum Theory of Electromagnetic Interaction", *Physical Review*, vol. 80, no. 3, pp. 440–457, 1950, ISSN: 0031-899X. DOI: 10.1103/PhysRev.80.440.

[26] J. Schwinger, "Quantum Electrodynamics. I. A Covariant Formulation", *Physical Review*, vol. 74, no. 10, pp. 1439–1461, 1948. DOI: 10.1103/PhysRev.74.1439.

[27] S. Tomonaga, "On a Relativistically Invariant Formulation of the Quantum Theory of Wave Fields*", *Progress of Theoretical Physics*, vol. 1, no. 2, pp. 27–42, 1946, ISSN: 0033-068X. DOI: 10.1143/PTP.1.27.

[28] S. L. Glashow, "Partial-Symmetries of Weak Interactions", *Nuclear Physics*, vol. 22, no. 4, pp. 579–588, 1961, ISSN: 0029-5582. DOI: 10.1016/0029-5582(61)90469-2.

[29] A. Salam, "Weak and Electromagnetic Interactions", in *Selected Papers of Abdus Salam*, ser. World Scientific Series in 20th Century Physics, vol. Volume 5, WORLD SCIENTIFIC, 1994, pp. 244–254, ISBN: 978-981-02-1662-7. DOI: 10.1142/9789812795915_0034.

[30] S. Weinberg, "A Model of Leptons", *Physical Review Letters*, vol. 19, no. 21, pp. 1264–1266, 1967. DOI: 10.1103/PhysRevLett.19.1264.

[31] M. Gell-Mann, "A Schematic Model of Baryons and Mesons", *Physics Letters*, vol. 8, no. 3, pp. 214–215, 1964, ISSN: 0031-9163. DOI: 10.1016/S0031-9163(64)92001-3.

[32] G. Zweig, "An SU_3 Model for Strong Interaction Symmetry and Its Breaking; Version 2", 1964. DOI: 10.17181/CERN-TH-412.

[33] H. Fritzsch, M. Gell-Mann, and H. Leutwyler, "Advantages of the Color Octet Gluon Picture", *Physics Letters B*, vol. 47, no. 4, pp. 365–368, 1973, ISSN: 0370-2693. DOI: 10.1016/0370-2693(73)90625-4.

[34] D. J. Gross and F. Wilczek, "Ultraviolet Behavior of Non-Abelian Gauge Theories", *Physical Review Letters*, vol. 30, no. 26, pp. 1343–1346, 1973. DOI: 10.1103/PhysRevLett.30.1343.

[35] H. D. Politzer, "Reliable Perturbative Results for Strong Interactions?", *Physical Review Letters*, vol. 30, no. 26, pp. 1346–1349, 1973. DOI: 10.1103/PhysRevLett.30.1346.

[36] D. J. Griffiths, *Introduction to Elementary Particles*. Weinheim: Wiley, 2007, ISBN: 978-0-471-60386-3.

[37] M. D. Schwartz, *Quantum Field Theory and Standard Model*. Cambridge: Cambridge University Press, 2013, ISBN: 978-1-108-98503-1. DOI: 10.1017/9781139540940.

[38] M. E. Peskin, *An Introduction To Quantum Field Theory*. Boca Raton: CRC Press, 2018, ISBN: 978-0-429-50355-9. DOI: 10.1201/9780429503559.

[39] Wikimedia Commons, *Standard Model of Elementary Particles*, 2024. [Online]. Available: https://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg (visited on 05/07/2026).

[40] J. C. Romao and J. P. Silva, "A Resource for Signs and Feynman Diagrams of the Standard Model", *International Journal of Modern Physics A*, vol. 27, no. 26, p. 1 230 025, 2012, ISSN: 0217-751X, 1793-656X. DOI: 10.1142/S0217751X12300256.

[41] G. Altarelli and G. Parisi, "Asymptotic Freedom in Parton Language", *Nuclear Physics B*, vol. 126, no. 2, pp. 298–318, 1977, ISSN: 0550-3213. DOI: 10.1016/0550-3213(77)90384-4.

[42] T.-J. Hou *et al.*, "New CTEQ Global Analysis of Quantum Chromodynamics with High-Precision Data from the LHC", *Physical Reviev D*, vol. 103, no. 1, p. 014 013, 2021.
DOI: 10.1103/PhysRevD.103.014013.

[43] J. McGowan *et al.*, "Approximate N^{3}LO Parton Distribution Functions with Theoretical Uncertainties: MSHT20aN^{3}LO PDFs", *European Physical Journal C*, vol. 83, no. 3, p. 185, 2023, ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-023-11236-0.

[44] NNPDF Collaboration, "The Path to N^{3}LO Parton Distributions", *European Physical Journal C*, vol. 84, no. 7, p. 659, 2024, ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-024-12891-7.

[45] A. Buckley *et al.*, "LHAPDF6: Parton Density Access in the LHC Precision Era", *European Physical Journal C*, vol. 75, no. 3, p. 132, 2015, ISSN: 1434-6052.
DOI: 10.1140/epjc/s10052-015-3318-8.

[46] N. A. Abdulov *et al.*, "TMDlib2 and TMDplotter: A Platform for 3D Hadron Structure Studies", *European Physical Journal C*, vol. 81, no. 8, p. 752, 2021, ISSN: 1434-6052.
DOI: 10.1140/epjc/s10052-021-09508-8.

[47] F. Siegert, "Monte-Carlo Event Generation for the LHC", Doctoral thesis, Durham University, 2010.

[48] C. Bierlich *et al.*, "A Comprehensive Guide to the Physics and Usage of PYTHIA 8.3", *SciPost Physics Codebases*, p. 008, 2022, ISSN: 2949-804X. DOI: 10.21468/SciPostPhysCodeb.8.

[49] M. Bähr *et al.*, "Herwig++ Physics and Manual", *European Physical Journal C*, vol. 58, no. 4, pp. 639–707, 2008, ISSN: 1434-6052.
DOI: 10.1140/epjc/s10052-008-0798-9.

[50] J. Bellm *et al.*, "Herwig 7.0/Herwig++ 3.0 Release Note", *European Physical Journal C*, vol. 76, no. 4, p. 196, 2016, ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-016-4018-8.

[51] T. Gleisberg *et al.*, "Event Generation with SHERPA 1.1", *Journal of High Energy Physics*, vol. 2009, no. 02, p. 007, 2009, ISSN: 1126-6708. DOI: 10.1088/1126-6708/2009/02/007.

[52] 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", *Journal of High Energy Physics*, vol. 2014, no. 7, p. 79, 2014, ISSN: 1029-8479. DOI: 10.1007/JHEP07(2014)079.

[53] R. Frederix *et al.*, "The Automation of Next-to-Leading Order Electroweak Calculations", *Journal of High Energy Physics*, vol. 2018, no. 7, p. 185, 2018, ISSN: 1029-8479. DOI: 10.1007/JHEP07(2018)185.

[54] P. Nason, "A New Method for Combining NLO QCD with Shower Monte Carlo Algorithms", *Journal of High Energy Physics*, vol. 2004, no. 11, p. 040, 2004, ISSN: 1126-6708. DOI: 10.1088/1126-6708/2004/11/040.

[55] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD Computations with Parton Shower Simulations: The POWHEG Method", *Journal of High Energy Physics*, vol. 2007, no. 11, p. 070, 2007, ISSN: 1126-6708. DOI: 10.1088/1126-6708/2007/11/070.

[56] S. Alioli *et al.*, "A General Framework for Implementing NLO Calculations in Shower Monte Carlo Programs: The POWHEG BOX", *Journal of High Energy Physics*, vol. 06, p. 043, 2010. DOI: 10.1007/JHEP06(2010)043.

[57] Y. Li and F. Petriello, "Combining QCD and Electroweak Corrections to Dilepton Production in the Framework of the FEWZ Simulation Code", *Physical Review D*, vol. 86, no. 9, p. 094 034, 2012. DOI: 10.1103/PhysRevD.86.094034.

[58] R. P. Feynman, "Very High-Energy Collisions of Hadrons", *Physical Review Letters*, vol. 23, no. 24, pp. 1415–1417, 1969. DOI: 10.1103/PhysRevLett.23.1415.

[59] J. H. Christenson *et al.*, "Observation of Massive Muon Pairs in Hadron Collisions", *Physical Review Letters*, vol. 25, no. 21, pp. 1523–1526, 1970. DOI: 10.1103/PhysRevLett.25.1523.

[60] J. C. Collins and D. E. Soper, "Angular Distribution of Dileptons in High-Energy Hadron Collisions", *Physical Review D*, vol. 16, no. 7, pp. 2219–2225, 1977. DOI: 10.1103/PhysRevD.16.2219.

[61] A. Ridder *et al.*, "Precision Phenomenology with Fiducial Cross Sections in the Triple-Differential Drell-Yan Process", *Journal of High Energy Physics*, vol. 2023, no. 5, p. 2, 2023, ISSN: 1029-8479. DOI: 10.1007/JHEP05(2023)002.

[62] C. M. C. Calame *et al.*, "Precision Electroweak Calculation of the Production of a High Transverse-Momentum Lepton Pair at Hadron Colliders", *Journal of High Energy Physics*, vol. 2007, no. 10, p. 109, 2007, ISSN: 1126-6708. DOI: 10.1088/1126-6708/2007/10/109.

[63] S. Dittmaier and M. Huber, "Radiative Corrections to the Neutral-Current Drell-Yan Process in the Standard Model and Its Minimal Supersymmetric Extension", *Journal of High Energy Physics*, vol. 2010, no. 1, p. 60, 2010, ISSN: 1029-8479. DOI: 10.1007/JHEP01(2010)060.

[64] L. Barzè *et al.*, "Neutral-Current Drell-Yan with Combined QCD and Electroweak Corrections in the POWHEG BOX", *European Physical Journal C*, vol. 73, no. 6, p. 2474, 2013, ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-013-2474-y.

[65] K. Melnikov and F. Petriello, "Electroweak Gauge Boson Production at Hadron Colliders through $\mathcal{O}(\alpha_s^2)$", *Physical Review D*, vol. 74, no. 11, p. 114 017, 2006. DOI: 10.1103/PhysRevD.74.114017.

[66] S. Catani *et al.*, "Vector Boson Production at Hadron Colliders: A Fully Exclusive QCD Calculation at Next-to-Next-to-Leading Order", *Physical Review Letters*, vol. 103, no. 8, p. 082 001, 2009. DOI: 10.1103/PhysRevLett.103.082001.

[67] S. Camarda *et al.*, "DYTurbo: Fast Predictions for Drell–Yan Processes", *European Physical Journal C*, vol. 80, no. 3, p. 251, 2020, ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-020-7757-5.

[68] P. F. Monni *et al.*, "MiNNLOPS: A New Method to Match NNLO QCD to Parton Showers", *Journal of High Energy Physics*, vol. 2020, no. 5, p. 143, 2020, ISSN: 1029-8479. DOI: 10.1007/JHEP05(2020)143.

[69] S. Dittmaier, A. Huss, and C. Schwinn, "Dominant Mixed QCD-electroweak $\mathcal{O}(\alpha_s\alpha)$ Corrections to Drell–Yan Processes in the Resonance Region", *Nuclear Physics B*, vol. 904, pp. 216–252, 2016, ISSN: 0550-3213. DOI: 10.1016/j.nuclphysb.2016.01.006.

[70] R. Bonciani *et al.*, "NNLO QCD × EW Corrections to Z Production in the $q\overline{q}$ Channel", *Physical Review D*, vol. 101, no. 3, p. 031 301, 2020. DOI: 10.1103/PhysRevD.101.031301.

[71] S. Dittmaier, T. Schmidt, and J. Schwarz, "Mixed NNLO QCD×electroweak Corrections of $\mathcal{O}(N_f\alpha_s\alpha)$ to Single-W/Z Production at the LHC", *Journal of High Energy Physics*, vol. 2020, no. 12, p. 201, 2020, ISSN: 1029-8479. DOI: 10.1007/JHEP12(2020)201.

[72] R. Bonciani *et al.*, "Mixed Strong-Electroweak Corrections to the Drell-Yan Process", *Physical Review Letters*, vol. 128, no. 1, p. 012 002, 2022. DOI: 10.1103/PhysRevLett.128.012002.

[73] T. Armadillo *et al.*, "Two-Loop Mixed QCD-EW Corrections to Neutral Current Drell-Yan", *Journal of High Energy Physics*, vol. 2022, no. 5, p. 72, 2022, ISSN: 1029-8479. DOI: 10.1007/JHEP05(2022)072.

[74] F. Buccioni *et al.*, "Mixed QCD-electroweak Corrections to Dilepton Production at the LHC in the High Invariant Mass Region", *Journal of High Energy Physics*, vol. 2022, no. 6, p. 22, 2022, ISSN: 1029-8479. DOI: 10.1007/JHEP06(2022)022.

[75] T. Armadillo *et al.*, "Mixed QCD-EW Corrections to the Neutral-Current Drell-Yan Process", *Journal of High Energy Physics*, vol. 2025, no. 141, 2025. DOI: 10.1007/JHEP07(2025)141.

[76] C. Duhr, F. Dulat, and B. Mistlberger, "Drell-Yan Cross Section to Third Order in the Strong Coupling Constant", *Physical Review Letters*, vol. 125, no. 17, p. 172 001, 2020. DOI: 10.1103/PhysRevLett.125.172001.

[77] C. Duhr and B. Mistlberger, "Lepton-Pair Production at Hadron Colliders at N3LO in QCD", *Journal of High Energy Physics*, vol. 2022, no. 3, p. 116, 2022, ISSN: 1029-8479. DOI: 10.1007/JHEP03(2022)116.

[78] J. Baglio *et al.*, "Inclusive Production Cross Sections at N3LO", *Journal of High Energy Physics*, vol. 2022, no. 12, p. 66, 2022, ISSN: 1029-8479. DOI: 10.1007/JHEP12(2022)066.

[79] X. Chen *et al.*, "Dilepton Rapidity Distribution in Drell-Yan Production to Third Order in QCD", *Physical Review Letters*, vol. 128, no. 5, p. 052 001, 2022. DOI: 10.1103/PhysRevLett.128.052001.

[80] X. Chen *et al.*, "Third-Order Fiducial Predictions for Drell-Yan Production at the LHC", *Physical Review Letters*, vol. 128, no. 25, p. 252 001, 2022. DOI: 10.1103/PhysRevLett.128.252001.

[81] T. Neumann and J. Campbell, "Fiducial Drell-Yan Production at the LHC Improved by Transverse-Momentum Resummation at $N^4LL_p + N^3LO$", *Physical Review D*, vol. 107, no. 1, p. L011506, 2023. DOI: 10.1103/PhysRevD.107.L011506.

[82] ATLAS Collaboration, "Search for High-Mass New Phenomena in the Dilepton Final State Using Proton-Proton Collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector", *Physics Letters B*, vol. 761, pp. 372–392, 2016, ISSN: 0370-2693. DOI: 10.1016/j.physletb.2016.08.055.

[83] ATLAS Collaboration, "Search for New High-Mass Phenomena in the Dilepton Final State Using 36 fb^{-1} of Proton-Proton Collision Data at $\sqrt{s} = 13$ TeV with the ATLAS Detector", *Journal of High Energy Physics*, vol. 2017, no. 10, p. 182, 2017, ISSN: 1029-8479. DOI: 10.1007/JHEP10(2017)182.

[84] CMS Collaboration, "Search for Narrow Resonances in Dilepton Mass Spectra in Proton-Proton Collisions at $\sqrt{s} = 13$ TeV and Combination with 8 TeV Data", *Physics Letters B*, vol. 768, pp. 57–80, 2017, ISSN: 0370-2693. DOI: 10.1016/j.physletb.2017.02.010.

[85] CMS Collaboration, "Search for High-Mass Resonances in Dilepton Final States in Proton-Proton Collisions at $\sqrt{s} = 13$ TeV", *Journal of High Energy Physics*, vol. 2018, no. 6, p. 120, 2018, ISSN: 1029-8479. DOI: 10.1007/JHEP06(2018)120.

[86] ATLAS Collaboration, "Search for High-Mass Dilepton Resonances Using 139 fb^{-1} of pp Collision Data Collected at $\sqrt{s} = 13$ TeV with the ATLAS Detector", *Physics Letters B*, vol. 796, pp. 68–87, 2019, ISSN: 0370-2693. DOI: 10.1016/j.physletb.2019.07.016.

[87] CMS Collaboration, "Search for Contact Interactions and Large Extra Dimensions in the Dilepton Mass Spectra from Proton-Proton Collisions at $\sqrt{s} = 13$ TeV", *Journal of High Energy Physics*, vol. 2019, no. 4, p. 114, 2019, ISSN: 1029-8479. DOI: 10.1007/JHEP04(2019)114.

[88] ATLAS Collaboration, "Search for New Non-Resonant Phenomena in High-Mass Dilepton Final States with the ATLAS Detector", *Journal of High Energy Physics*, vol. 2020, no. 11, p. 5, 2020, ISSN: 1029-8479. DOI: 10.1007/JHEP11(2020)005.

[89] CMS Collaboration, "Search for Resonant and Nonresonant New Phenomena in High-Mass Dilepton Final States at $\sqrt{s} = 13$ TeV", *Journal of High Energy Physics*, vol. 2021, no. 7, p. 208, 2021, ISSN: 1029-8479. DOI: 10.1007/JHEP07(2021)208.

[90] CMS Collaboration, "Measurements of Differential Z Boson Production Cross Sections in Proton-Proton Collisions at $\sqrt{s} = 13$ TeV", *Journal of High Energy Physics*, vol. 2019, no. 12, p. 61, 2019, ISSN: 1029-8479. DOI: 10.1007/JHEP12(2019)061.

[91] CMS Collaboration, "Forward-Backward Asymmetry of Drell-Yan Lepton Pairs in pp Collisions at $\sqrt{s} = 7$ TeV", *Physics Letters B*, vol. 718, no. 3, pp. 752–772, 2013, ISSN: 0370-2693. DOI: 10.1016/j.physletb.2012.10.082.

[92] CMS Collaboration, "Forward–Backward Asymmetry of Drell-Yan Lepton Pairs in pp Collisions at $\sqrt{s} = 8$ TeV", *European Physical Journal C*, vol. 76, no. 6, p. 325, 2016, ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-016-4156-z.

[93] CMS Collaboration, "Measurement of the Drell-Yan Forward-Backward Asymmetry at High Dilepton Masses in Proton-Proton Collisions at $\sqrt{s} = 13$ TeV", *Journal of High Energy Physics*, vol. 2022, no. 8, p. 63, 2022, ISSN: 1029-8479. DOI: 10.1007/JHEP08(2022)063.

[94] CMS Collaboration, "Measurement of the Weak Mixing Angle with the Drell-Yan Process in Proton-Proton Collisions at the LHC", *Physical Review D*, vol. 84, no. 11, p. 112 002, 2011. DOI: 10.1103/PhysRevD.84.112002.

[95] ATLAS Collaboration, "Z Boson Production in $p+$Pb Collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV Measured with the ATLAS Detector", *Physical Review C*, vol. 92, no. 4, p. 044 915, 2015. DOI: 10.1103/PhysRevC.92.044915.

[96] CMS Collaboration, "Study of Z Boson Production in pPb Collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV", *Physics Letters B*, vol. 759, pp. 36–57, 2016, ISSN: 0370-2693. DOI: 10.1016/j.physletb.2016.05.044.

[97] ALICE Collaboration, "W and Z Boson Production in P-Pb Collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV", *Journal of High Energy Physics*, vol. 2017, no. 2, p. 77, 2017, ISSN: 1029-8479. DOI: 10.1007/JHEP02(2017)077.

[98] ALICE Collaboration, "Z-Boson Production in p-Pb Collisions at $\sqrt{s_{\mathrm{NN}}}$ = 8.16 TeV and Pb-Pb Collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV", *Journal of High Energy Physics*, vol. 2020, no. 9, p. 76, 2020, ISSN: 1029-8479. DOI: 10.1007/JHEP09(2020)076.

[99] CMS Collaboration, "Study of Drell-Yan Dimuon Production in Proton-Lead Collisions at $\sqrt{s_{\mathrm{NN}}} = 8.16$ TeV", *Journal of High Energy Physics*, vol. 2021, no. 5, p. 182, 2021, ISSN: 1029-8479. DOI: 10.1007/JHEP05(2021)182.

[100] PHENIX Collaboration, "Measurements of $\mu\mu$ Pairs from Open Heavy Flavor and Drell-Yan in $p + p$ Collisions at $\sqrt{s} = 200$ GeV", *Physical Review D*, vol. 99, no. 7, p. 072 003, 2019. DOI: 10.1103/PhysRevD.99.072003.

[101] FNAL E906/SeaQuest Collaboration, "Measurement of Flavor Asymmetry of the Light-Quark Sea in the Proton with Drell-Yan Dimuon Production in $p + p$ and $p + d$ Collisions at 120 GeV", *Physical Review C*, vol. 108, no. 3, p. 035 202, 2023. DOI: 10.1103/PhysRevC.108.035202.

[102] COMPASS Collaboration, "Final COMPASS Results on the Transverse-Spin-Dependent Azimuthal Asymmetries in the Pion-Induced Drell-Yan Process", *Physical Review Letters*, vol. 133, no. 7, p. 071 902, 2024. DOI: 10.1103/PhysRevLett.133.071902.

[103] "Longer Term LHC Schedule". (2024), [Online]. Available: https://lhc-commissioning.web.cern.ch/schedule/LHC-long-term.htm (visited on 05/07/2026).

[104] J. Wenninger, "LHC Status and Performance", in *Proceedings of Prospects for Charged Higgs Discovery at Colliders – PoS(CHARGED2018)*, vol. 339, SISSA Medialab, 2019, p. 001. DOI: 10.22323/1.339.0001.

[105] P. Mouche, "Overall View of the LHC", OPEN-PHO-ACCEL-2014-001, 2014. Available: https://cds.cern.ch/record/1708847.

[106] CMS Collaboration, "78 Reconstructed Vertices in Event from High-Pileup Run 198609", CMS-PHO-EVENTS-2012-006, 2012. Available: https://cds.cern.ch/record/1479324.

[107] D. Barney, "CMS Detector Slice", CMS-PHO-GEN-2016-001, 2016. Available: https://cds.cern.ch/record/2120661.

[108] "CMS Coordinate System – TikZ.Net". (2024), [Online]. Available: https://tikz.net/axis3d_cms/ (visited on 05/07/2026).

[109] A. Dominguez *et al.*, "CMS Technical Design Report for the Pixel Detector Upgrade", CERN-LHCC-2012-016, CMS-TDR-11, 2012. Available: https://cds.cern.ch/record/1481838.

[110] P. Adzic *et al.*, "Energy Resolution Performance of the CMS Electromagnetic Calorimeter", CERN-CMS-NOTE-2006-140, CMS-NOTE-2006-140, 2006. Available: https://cds.cern.ch/record/1000388.

[111] J. Mans *et al.*, "CMS Technical Design Report for the Phase 1 Upgrade of the Hadron Calorimeter", CERN-LHCC-2012-015, CMS-TDR-10, 2012. Available: https://cds.cern.ch/record/1481837.

[112] CMS Collaboration, "The Phase-2 Upgrade of the CMS Muon Detectors", CERN-LHCC-2017-012, CMS-TDR-016, 2017. Available: https://cds.cern.ch/record/2283189.

[113] CMS Collaboration, "CMS The TriDAS Project: Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger", CERN-LHCC-2002-026, CMS-TDR-6, 2002. Available: https://cds.cern.ch/record/578006.

[114] A. Tapper and D. Acosta, "CMS Technical Design Report for the Level-1 Trigger Upgrade", CERN-LHCC-2013-011, CMS-TDR-012, 2013. Available: https://cds.cern.ch/record/1556311.

[115] CMS Collaboration, "Particle-Flow Reconstruction and Global Event Description with the CMS Detector", *Journal of Instrumentation*, vol. 12, no. 10, P10003, 2017, ISSN: 1748-0221. DOI: 10.1088/1748-0221/12/10/P10003.

[116] CMS Collaboration, "Performance of Photon Reconstruction and Identification with the CMS Detector in Proton-Proton Collisions at $\sqrt{s}$ = 8 TeV", *Journal of Instrumentation*, vol. 10, no. 08, P08010, 2015. DOI: 10.1088/1748-0221/10/08/P08010.

[117] CMS Collaboration, "Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s}$ = 8 TeV", *Journal of Instrumentation*, vol. 10, no. 06, P06005, 2015. DOI: 10.1088/1748-0221/10/06/P06005.

[118] W. Adam *et al.*, "Reconstruction of Electrons with the Gaussian-Sum Filter in the CMS Tracker at the LHC", *Journal of Physics G: Nuclear and Particle Physics*, vol. 31, no. 9, N9, 2005. DOI: 10.1088/0954-3899/31/9/N01.

[119] CMS Collaboration, "Performance of the CMS Muon Detector and Muon Reconstruction with Proton-Proton Collisions at $\sqrt{s}$ = 13 TeV", *Journal of Instrumentation*, vol. 13, no. 06, P06015, 2018, ISSN: 1748-0221. DOI: 10.1088/1748-0221/13/06/P06015.

[120] M. Cacciari, G. P. Salam, and G. Soyez, "The Anti-k_t Jet Clustering Algorithm", *Journal of High Energy Physics*, vol. 2008, no. 04, p. 063, 2008. DOI: 10.1088/1126-6708/2008/04/063.

[121] CMS Collaboration, "Jet Energy Scale and Resolution in the CMS Experiment in pp Collisions at 8 TeV", *Journal of Instrumentation*, vol. 12, no. 02, P02014, 2017. DOI: 10.1088/1748-0221/12/02/P02014.

[122] A. M. Sirunyan *et al.*, "Extraction and Validation of a New Set of CMS Pythia8 Tunes from Underlying-Event Measurements", *European Physical Journal C*, vol. 80, no. 1, p. 4, 2020, ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-019-7499-4.

[123] N. Davidson, T. Przedzinski, and Z. Was, "PHOTOS Interface in C++: Technical and Physics Documentation", *Computer Physics Communications*, vol. 199, pp. 86–101, 2016, ISSN: 0010-4655. DOI: 10.1016/j.cpc.2015.09.013.

[124] NNPDF Collaboration, "Parton Distributions from High-Precision Collider Data", *European Physical Journal C*, vol. 77, no. 10, p. 663, 2017, ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-017-5199-5.

[125] S. Agostinelli *et al.*, "Geant4 – a Simulation Toolkit", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 506, no. 3, pp. 250–303, 2003, ISSN: 0168-9002. DOI: 10.1016/S0168-9002(03)01368-8.

[126] J. A. M. Vermaseren, "Two-Photon Processes at Very High Energies", *Nuclear Physics B*, vol. 229, no. 2, pp. 347–371, 1983, ISSN: 0550-3213. DOI: 10.1016/0550-3213(83)90336-X.

[127] S. P. Baranov *et al.*, "LPAIR: A Generator for Lepton Pair Production", in *Workshop on Physics at HERA*, 1991.

[128] L. Forthomme, "CepGen – A generic central exclusive processes event generator for hadron-hadron collisions", *Computer Physics Communications*, vol. 271, p. 108 225, 2022, ISSN: 0010-4655. DOI: 10.1016/j.cpc.2021.108225.

[129] A. Bodek *et al.*, "Extracting Muon Momentum Scale Corrections for Hadron Collider Experiments", *European Physical Journal C*, vol. 72, p. 2194, 2012. DOI: 10.1140/epjc/s10052-012-2194-8.

[130] E. Bols *et al.*, "Jet Flavour Classification Using DeepJet", *Journal of Instrumentation*, vol. 15, no. 12, P12012, 2020, ISSN: 1748-0221. DOI: 10.1088/1748-0221/15/12/P12012.

[131] "Performance Summary of AK4 Jet b Tagging with Data from Proton-Proton Collisions at 13 TeV with the CMS Detector", CERN-CMS-DP-2023-005, CMS-DP-2023-005, 2023. Available: https://cds.cern.ch/record/2854609.

[132] ATLAS Collaboration, "Search for a Heavy Neutral Particle Decaying to $e\mu$, $e\tau$, or $\mu\tau$ in pp Collisions at $\sqrt{s}$ = 8 TeV with the ATLAS Detector", *Physical Review Letters*, vol. 115, no. 3, p. 031 801, 2015. DOI: 10.1103/PhysRevLett.115.031801.

[133] ATLAS Collaboration, "Search for Lepton-Flavour Violation in High-Mass Dilepton Final States Using 139 fb^{-1} of pp Collisions at $\sqrt{s}$ = 13 TeV with the ATLAS Detector", *Journal of High Energy Physics*, vol. 2023, no. 10, p. 82, 2023, ISSN: 1029-8479. DOI: 10.1007/JHEP10(2023)082.

[134] CMS Collaboration, "Differential Cross Section Measurements for the Production of Top Quark Pairs and of Additional Jets Using Dilepton Events from pp Collisions at $\sqrt{s}$ = 13 TeV", *Journal of High Energy Physics*, vol. 2025, no. 2, p. 64, 2025, ISSN: 1029-8479. DOI: 10.1007/JHEP02(2025)064.

[135] S. Schmitt, "TUnfold, an Algorithm for Correcting Migration Effects in High Energy Physics", *Journal of Instrumentation*, vol. 7, no. 10, T10003, 2012, ISSN: 1748-0221. DOI: 10.1088/1748-0221/7/10/T10003.

[136] L. Lyons, D. Gibaut, and P. Clifford, "How to Combine Correlated Estimates of a Single Physical Quantity", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 270, no. 1, pp. 110–117, 1988, ISSN: 0168-9002. DOI: 10.1016/0168-9002(88)90018-6.

[137] CMS Collaboration, "The Phase-2 Upgrade of the CMS Tracker", CERN-LHCC-2017-009, CMS-TDR-014, 2017. DOI: 10.17181/CERN.QZ28.FLHW.

[138] ATLAS Collaboration, "ATLAS Phase-II Upgrade Scoping Document", CERN-LHCC-2015-020, LHCC-G-166, 2015. DOI: 10.17181/CERN.7CRX.AJHP.

[139] CMS Collaboration, "Displays of Events from Run 3 Seen in CMS with Multiple Reconstructed Vertices", CMS-PHO-EVENTS-2023-023-1, 2023. Available: https://cds.cern.ch/record/2867797.

[140] CMS Collaboration and T. Mc Cauley, "Collisions Recorded by the CMS Detector on 14 Oct 2016 During the High Pile-Up Fill", CMS-PHO-EVENTS-2016-008, 2016. Available: https://cds.cern.ch/record/2231915.

[141] CMS Collaboration, "The CMS Tracker System Project: Technical Design Report", CERN-LHCC-98-006, CMS-TDR-5, 1997. Available: https://cds.cern.ch/record/368412.

[142] CMS Collaboration, "The CMS Tracker: Addendum to the Technical Design Report", CERN-LHCC-2000-016, CMS-TDR-5-add-1, 2000. Available: https://cds.cern.ch/record/490194.

[143] CMS Collaboration, "The Phase-2 Upgrade of the CMS Endcap Calorimeter", CERN-LHCC-2017-023, CMS-TDR-019, 2017. DOI: 10.17181/CERN.IV8M.1JY2.

[144] A. Colaleo *et al.*, "CMS Technical Design Report for the Muon Endcap GEM Upgrade", CERN-LHCC-2015-012, CMS-TDR-013, 2015. Available: https://cds.cern.ch/record/2021453.

[145] CMS Collaboration, "A MIP Timing Detector for the CMS Phase-2 Upgrade", CERN-LHCC-2019-003, CMS-TDR-020, 2019. Available: https://cds.cern.ch/record/2667167.

[146] CMS Collaboration, "The Phase-2 Upgrade of the CMS Beam Radiation Instrumentation and Luminosity Detectors", CERN-LHCC-2021-008, CMS-TDR-023, 2021. Available: https://cds.cern.ch/record/2759074.

[147] CMS Collaboration, "The Phase-2 Upgrade of the CMS Level-1 Trigger", CERN-LHCC-2020-004, CMS-TDR-021, 2020. Available: https://cds.cern.ch/record/2714892.

[148] CMS Collaboration, “The Phase-2 Upgrade of the CMS Data Acquisition and High Level Trigger”, CERN-LHCC-2021-007, CMS-TDR-022, 2021. Available: https://cds.cern.ch/record/2759072.

[149] CMS Offline Software and Computing, “CMS Phase-2 Computing Model: Update Document”, Tech. Rep., 2022. Available: https://cds.cern.ch/record/2815292.

[150] lpGBT Design Team, *lpGBT documentation: release*. 2024. [Online]. Available: https://cds.cern.ch/record/2809058 (visited on 04/22/2026).

[151] C. Soós *et al.*, “Versatile Link PLUS Transceiver Development”, *Journal of Instrumentation*, vol. 12, no. 03, p. C03068, 2017, ISSN: 1748-0221. DOI: 10.1088/1748-0221/12/03/C03068.

[152] E. Monteil *et al.*, “A Prototype of a New Generation Readout ASIC in 65nm CMOS for Pixel Detectors at HL-LHC”, *Journal of Instrumentation*, vol. 11, no. 12, p. C12044, 2016, ISSN: 1748-0221. DOI: 10.1088/1748-0221/11/12/C12044.

[153] M. Garcia-Sciveres, “The RD53A Integrated Circuit”, CERN-RD53-PUB-17-001, 2017. Available: https://cds.cern.ch/record/2287593.

[154] M. Garcia-Sciveres, “RD53A Integrated Circuit Specifications”, CERN-RD53-PUB-15-001, 2015. Available: https://cds.cern.ch/record/2113263.

[155] L. Gaioni *et al.*, “Optimization of the 65-Nm CMOS Linear Front-End Circuit for the CMS Pixel Readout at the HL-LHC”, *IEEE Transactions on Nuclear Science*, vol. 68, no. 11, pp. 2682–2692, 2021, ISSN: 1558-1578. DOI: 10.1109/TNS.2021.3117666.

[156] M. Garcia-Sciveres, F. Loddo, and J. Christiansen, “RD53B Manual”, CERN-RD53-PUB-19-002, 2019. Available: https://cds.cern.ch/record/2665301.

[157] Xilinx Inc., *Aurora Protocol Documentation*. [Online]. Available: https://www.xilinx.com/products/intellectual-property/aurora64b66b.html (visited on 05/07/2026).

[158] C. G. Larrea *et al.*, “IPbus: A Flexible Ethernet-based Control System for xTCA Hardware”, *Journal of Instrumentation*, vol. 10, no. 02, p. C02019, 2015, ISSN: 1748-0221. DOI: 10.1088/1748-0221/10/02/C02019.

[159] M. E. Dinardo, “The Data Acquisition System to Test and Characterize the Pixel Detector Modules of the CMS Inner Tracker for the High Luminosity Upgrade of LHC”, in *2021 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, 2021, pp. 1–7. DOI: 10.1109/NSS/MIC44867.2021.9875778.

[160] “Inner Tracker Scans and Calibrations – Ph2_ACF User Guide”. *Non-public document*. (2026), [Online]. Available: https://ph2acf.docs.cern.ch/innerTracker/calibrations/ (visited on 05/07/2026).

[161] “POWDER (GitLab)”. (2022), [Online]. Available: https://gitlab.cern.ch/cms_tk_ph2/power_supply (visited on 05/07/2026).

[162] F. Scholze, H. Rabus, and G. Ulm, “Mean Energy Required to Produce an Electron-Hole Pair in Silicon for Photons of Energies between 50 and 1500 eV”, *Journal of Applied Physics*, vol. 84, no. 5, pp. 2926–2939, 1998, ISSN: 0021-8979. DOI: 10.1063/1.368398.

SANTRAUKA LIETUVIŲ KALBA

Įvadas

Dalelių fizikos Standartinis modelis (SM) yra teorinis pagrindas, šiuo metu tiksliausiai aprašantis žinomas elementariąsias daleles ir jų elektromagnetinę, silpnąją bei stipriąją sąveikas. Per pastaruosius dešimtmečius SM prognozės buvo daug kartų patvirtintos eksperimentiškai, įskaitant W ir Z bozonų [1–4], viršūninio kvarko [5, 6] ir Higso bozono atradimus [7, 8]. Nepaisant šios sėkmės, SM nėra baigtinė fundamentali teorija: jis nepaaiškina tamsiosios medžiagos ir tamsiosios energijos, neutrinų masių, Visatoje stebimos medžiagos ir antimedžiagos asimetrijos bei neaprašo gravitacijos. Todėl didelio tikslumo SM procesų matavimai išlieka vienu svarbiausių būdų ieškoti galimų nukrypimų nuo teorinių prognozių ir netiesioginių naujos fizikos požymių.

Vienas iš svarbiausių etaloninių procesų hadronų greitintuvuose yra Drell–Yan (DY) procesas [14], kurio metu kvarko ir antikvarko anihiliacija sukuria virtualų fotoną arba Z bozoną, iškart suskylantį į priešingo krūvio leptonų porą. Šis procesas pasižymi lengvai aptinkamu eksperimentiniu signalu, aukštu teorinių skaičiavimų tikslumu ir jautrumu protono sandarą aprašančioms partonų pasiskirstymo funkcijoms (PDF). Dėl to DY proceso diferencialinio reakcijos skerspjūvio matavimai leidžia tikrinti perturbatyvios kvantinės chromodinamikos (QCD) ir elektrosilpnosios teorijos prognozes, tikslinti PDF rinkinius bei gerinti fono modeliavimą naujos fizikos paieškose su leptonų poros galutinėmis būsenomis.

Didelio šviesio Didžiojo hadronų greitintuvo (HL-LHC) atnaujinimas yra skirtas maždaug dešimt kartų padidinti LHC šviesį, taip ženkliai padidinant surenkamų duomenų imtis, gerinant matavimų statistinį tikslumą ir plečiant retų SM procesų tyrimų bei naujos fizikos paieškų galimybes [21]. Tuo pačiu didesnis šviesis reikš didesnę radiacinę pažaidą, didesnį detektorių sensorių užimtumą ir daugiau vienalaikių protonų susidūrimų, todėl LHC eksperimentams būtini reikšmingi atnaujinimai. CMS detektoriaus [10] II fazės atnaujinimas yra skirtas pagerinti detektoriaus veikimą HL-LHC sąlygomis modernizuojant trekerį, kalorimetrus, miuonų detektorius, trigerį ir duomenų surinkimo sistemas [22]. Viena svarbiausių šio atnaujinimo dalių yra visiškas CMS Vidinio trekerio pakeitimas radiacijai atspariais silicio pikseliniais sensoriais ir nuskaitymo elektronika, įskaitant RD53 programoje sukurtą RD53B-CMS prototipą [23, 156], kurio testavimui naudojama Ph2-ACF duomenų surinkimo ir kokybės užtikrinimo programinė įranga [24]. Didesnė HL-LHC statistika ir pagerėjęs CMS II fazės detektoriaus veikimas turėtų pagerinti būsimų DY matavimų tiks-

lumą, todėl šioje disertacijoje trekerio atnaujinimo darbai susiejami su ilgalaike CMS didelio tikslumo matavimų programa.

Šioje disertacijoje nagrinėjami du tarpusavyje susiję CMS eksperimento tyrimų programos aspektai. Pirmoji dalis skirta DY proceso diferencialinio reakcijos skerspjūvio matavimui naudojant visą CMS Antrojo duomenų rinkimo etapo (2016-2018 m.) protonų susidūrimų duomenų imtį. Antroji dalis skirta pasirengimo darbams CMS detektoriaus II fazės atnaujinimui, ypatingą dėmesį skiriant CMS Vidinio trekerio pikselinių lustų testavimui, kalibravimui ir Ph2-ACF programinės įrangos taikymui. Galiausiai įvertinama, kaip CMS II fazės atnaujinimas gali paveikti būsimų DY matavimų tikslumą HL-LHC sąlygomis, naudojant Monte Carlo (MC) modeliavimo įrankius.

Tikslas ir uždaviniai

Pagrindinis disertacijos tikslas yra dvejopas. Pirmoji jo dalis yra prisidėti prie DY diferencialinio skerspjūvio $\mathrm{d}\sigma/\mathrm{d}M$ matavimo $\sqrt{s} = 13$ TeV protonų susidūrimuose, naudojant CMS detektoriaus 2016–2018 m. duomenis, ypatingą dėmesį skiriant fono įvertinimui. Antroji dalis yra prisidėti prie CMS II fazės Vidinio trekerio atnaujinimo testuojant RD53B-CMS pikselinio lusto prototipą CROCv1 ir kuriant bei tikrinant Ph2-ACF programinės įrangos kalibravimo bei kokybės užtikrinimo įrankius. Šias dvi dalis jungia trečiasis uždavinys – įvertinti DY proceso matavimo perspektyvas HL-LHC laikotarpiu su modernizuotu CMS detektoriumi.

Tikslui pasiekti buvo atlikti šie darbai: sukurta ir pritaikyta matavimu grįsta fono su klaidingai atpažintais leptonais įvertinimo metodika; įgyvendinta $e\mu$ kontrolinės srities pagrindu paremta viršūninių kvarkų fono pataisa; įvertinti fono sisteminiai ir statistiniai neapibrėžtumai; prisidėta prie Ph2-ACF programinės įrangos kūrimo, tobulinimo, testavimo ir dokumentavimo; atlikti CROCv1 pikselinio lusto prototipo kalibravimo ir veikimo tyrimai; įvertintas DY proceso matavimo tikslumas modeliuojant II fazės CMS detektorių.

Naujumas ir svarba

DY procesas yra gerai žinomas ir plačiai tyrinėtas, tačiau didelio tikslumo diferencialinio skerspjūvio matavimai plačioje invariantinės masės srityje išlieka svarbūs tiek SM patikrinimui, tiek PDF tikslinimui, tiek naujos fizikos paieškoms. Šioje disertacijoje pateiktas matavimas siekia pateikti vieną tiksliausių DY diferencialinio skerspjūvio matavimų plačiame leptonų poros invariantinės masės diapazone, kai masės centro energija yra 13 TeV. CMS Antrojo duomenų rinkimo etapo metu tai buvo rekordinė energija, kurios joks eksperimentas iki

tol nebuvo pasiekęs. Šioje disertacijoje pristatomi fono vertinimo metodai yra svarbūs todėl, kad jie leidžia kontroliuoti vieną iš sudėtingiausių DY matavimo dalių – foną, susijusį su klaidingai atpažintais leptonais – abiejuose (elektronų ir miuonų) kanaluose. Lyginant su įprastesniais metodais, šiame darbe pritaikomi metodai yra patobulinti ir todėl gali būti taikomi universaliau, tuo pačiu išlaikydami gerą tikslumą.

CMS II fazės atnaujinimo dalyje disertacijos naujumas susijęs su RD53B-CMS prototipo testavimu ir Ph2-ACF kalibracijos procedūrų validavimo tikrinimu. Šie darbai yra tiesiogiai susiję su būsimos CMS Vidinio trekerio modulių gamybos ir kokybės užtikrinimo programa. CROCv1 pikselinio lusto signalo aptikimo slenksčio kalibravimo, signalo stiprio ir krūvio injekavimo grandinės kalibravimo tyrimai patvirtina, kad lustas atitinka svarbius II fazės Vidinio trekerio veikimo reikalavimus. II fazės DY modeliavimo tyrimas papildo šį darbą parodydamas, kaip modernizuotas detektorius gali pagerinti būsimų DY matavimų skyrą net sudėtingoje HL-LHC aplinkoje.

Disertacijos teiginiai

Disertacijoje ginami šie teiginiai:

1. Pateiktas vienodo krūvio ženklo (angl. *same-sign*) metodo variantas leidžia klaidingai atpažintų leptonų foną įvertinti tiek elektronų, tiek miuonų galutinėse būsenose taip, kad sisteminis neapibrėžtumas neviršytų 50%, kai leptonų poros invariantinė masė $M_{\ell\ell} < 900$ GeV.

2. Masių srityje $M_{\ell\ell} > 300$ GeV dominuojantis klaidingai atpažintų leptonų fono įverčio neapibrėžtumas yra statistinis.

3. RD53B-CMS pikselinio lusto prototipas gali būti sukalibruotas taip, kad signalo aptikimo slenksčių pasiskirstymo plotis būtų maždaug du kartus mažesnis nei RD53A demonstraciniame luste.

4. RD53B-CMS lusto signalo stipris tiesiškai priklauso nuo pikselyje surinkto krūvio ir nepriklauso nuo signalo aptikimo slenksčio.

5. RD53B-CMS lusto savikalibracijos grandinė leidžia keisti injekuotą krūvį apytiksliai 5 elektronų dydžio žingsniu.

6. Be numatomo bent penkis kartus didesnio DY įvykių skaičiaus, lyginant su CMS Antruoju duomenų rinkimo etapu (2016–2018 m.), modernizuotas CMS detektorius nagrinėtoje fazinėje erdvėje turėtų pagerinti miuono skersinio impulso skyrą bent 30%, o miuonų poros invariantinės masės skyrą bent 20%.

SM yra kvantinė lauko teorija, pagrįsta lokaliomis kalibruotinėmis simetrijomis

$$\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y. \tag{1}$$

Stiprioji sąveika aprašoma QCD teorija, paremta $\mathrm{SU}(3)_C$ simetrija, o elektromagnetinė ir silpnoji sąveikos sujungiamos elektrosilpnojoje teorijoje. Higso mechanizmas paaiškina, kaip W ir Z bozonai bei fermionai įgyja masę, išsaugant kalibruotinį teorijos nuoseklumą. Nors SM yra itin tiksliai patikrinta teorija, jo ribotumai – tamsiosios medžiagos, neutrinų masių, medžiagos ir antimedžiagos asimetrijos bei gravitacijos nepaaiškinimas – motyvuoja tolesnius tikslius matavimus.

Trumpai SM dinamiką galima aprašyti Lagranžiano tankiu

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{fermion}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}}. \tag{2}$$

Čia $\mathcal{L}_{\text{gauge}}$ aprašo kalibruotinių laukų dinamiką ir jų tarpusavio sąveikas, $\mathcal{L}_{\text{fermion}}$ – fermionų (kvarkų ir leptonų) kinetinius narius bei jų sąveiką su kalibruotiniais laukais, $\mathcal{L}_{\text{Higgs}}$ – Higso lauką ir spontaninį elektrosilpnosios simetrijos sulaužymą, o $\mathcal{L}_{\text{Yukawa}}$ – Higso lauko sąveiką su fermionais, lemiančią jų masių atsiradimą. Iš šio Lagranžiano išvedamos Feinmano taisyklės, kurios leidžia apskaičiuoti dalelių sąveikų amplitudes ir iš jų gauti fizikinius stebimus dydžius, tokius kaip reakcijos skerspjūviai ir skilimo tikimybės.

Kvantinėje lauko teorijoje reakcijos skerspjūviai apskaičiuojami keliais etapais. Pirmiausia iš Lagranžiano išvedamos Feinmano taisyklės, kurios nusako, kokie sąveikos viršūnių ir propagatorių nariai atitinka nagrinėjamą procesą. Tuomet nubraižomos atitinkamos Feinmano diagramos ir pagal jas užrašoma proceso matrica, arba amplitudė, $\mathcal{M}$. Pritaikant metodiką protonų susidūrimo procesui, iš amplitudės, susumavus ir suvidurkinus pagal pradines bei galutines būsenas, gaunamas diferencialinis partoninis skerspjūvis:

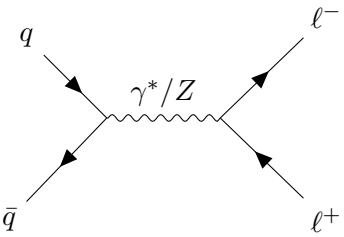
$$\mathrm{d}\hat{\sigma} \propto |\mathcal{M}|^2 \mathrm{d}\Phi_n, \tag{3}$$

kur $\mathrm{d}\Phi_n$ žymi n dalelių fazinės erdvės elementą.

DY procesas yra vienas švariausių hadroninių procesų, naudojamų SM tikslumo patikrinimui. Neutralios srovės DY procesas aprašomas schema

$$q\bar{q} \to \gamma^*/Z \to \ell^+\ell^-. \tag{4}$$

Šiame procese kvarkas ir antikvarkas anihiliuoja į virtualų fotoną arba Z bozoną, kuris vėliau suskyla į priešingo krūvio leptonų porą. Pagrindinė šio proceso Feinmano diagrama pateikta 1 pav.



1 pav.: Drell-Yan proceso medžio lygmens Feinmano diagrama.

Jo partoninis reakcijos skerspjūvis apskaičiuojamas iš virtualaus fotono ir Z bozono mainų amplitudžių bei jų interferencijos. Reakcijos skerspjūvis protonų susidūrimuose $\sigma_{pp\to\ell^+\ell^-}$ apskaičiuojamas kaip partoninio skerspjūvio konvoliucija su protono PDF:

$$\sigma_{pp\to\ell^+\ell^-} = \sum_{i,j} \int \mathrm{d}x_i \mathrm{d}x_j \, f_i(x_i, \mu_F^2) f_j(x_j, \mu_F^2) \hat{\sigma}_{ij\to\ell^+\ell^-}. \tag{5}$$

Čia i, j – galimi partonų tipai (u, d, c, s, b), x_i, x_j – šių partonų nešamos protono impulso santykinės dalys, f_i, f_j – jų pasiskirstymo funkcijos (PDF), $\hat{\sigma}_{ij\to\ell^+\ell^-}$ – partoninis reakcijos skerspjūvis, o μ_F – faktorizacijos skalė, nusakanti ribą tarp perturbatyvaus (nusakomo partoniniu reakcijos skerspjūviu) ir neperturbatyvaus (nusakomu partonų pasiskirstymo funkcijomis) QCD režimų. Ši išraiška pabrėžia, kad DY procesas yra tiesiogiai jautrus protono kvarkų ir antikvarkų pasiskirstymams. Matavimai skirtingose leptonų poros invariantinės masės ir spartos srityse suteikia informaciją apie skirtingas partonų impulso dalies x sritis.

DY diferencialinis reakcijos skerspjūvis gali būti apskaičiuojamas labai aukštu teoriniu tikslumu. QCD pataisos yra apskaičiuotos iki antros eilės (NNLO) daugeliui eksperimentiškai stebimų dydžių, o kai kuriems dydžiams jau pasiekiamas ir trečios eilės (N3LO) tikslumas. Elektrosilpnosios pataisos taip pat svarbios, ypač didelių invariantinių masių srityje. Eksperimentiniu požiūriu DY proceso švarios elektronų ir miuonų galutinės būsenos leidžia atlikti labai tikslius reakcijos skerspjūvio, kampinių pasiskirstymų ir priekinės-atbulinės asimetrijos matavimus. Todėl DY procesas yra vienas pagrindinių įrankių tikrinant SM ir ieškant netiesioginių naujos fizikos požymių.

Drell–Yan proceso tyrimas

DY diferencialinio reakcijos skerspjūvio matavimas šioje disertacijoje atliktas naudojant visą CMS Antrojo duomenų rinkimo etapo (2016–2018 m.) protonų

susidūrimų duomenų imtį, atitinkančią 137.6 fb^{-1} integruotąjį šviesį. Matavimas atliktas esant $\sqrt{s} = 13$ TeV protonų susidūrimo energijai, atskirai elektronų ir miuonų kanaluose. Pagrindinis matuojamas dydis yra diferencialinis reakcijos skerspjūvis dσ/d$M_{\ell\ell}$ invariantinių masių intervale $40 < M_{\ell\ell} < 3000$ GeV.

Tiriamoji fazinės erdvės sritis apibrėžiama naudojant sumodeliuotus leptonus galutinėje būsenoje po radiacinių energijos nuostolių. Jų keturmačiai impulsai buvo susumuoti su jiems artimiausių fotonų, esančių kūgyje $\Delta R(\ell, \gamma) < 0.1$, keturmačiais impulsais. Taip apibrėžti leptonai, angliškai vadinami *dressed leptons*, buvo naudojami tiriamajai fazinės erdvės sričiai apibrėžti. Abiem kanalams taikomi vienodi fazinės srities apibrėžimai: $p_{\mathrm{T}}^{\mathrm{lead}} > 20$ GeV, $p_{\mathrm{T}}^{\mathrm{sub}} > 15$ GeV, $|\eta_\ell| < 2.4$ ir priešingo ženklo leptonų pora. Čia p_{T} žymi skersinį impulsą, „lead“ ir „sub“ – atitinkamai didesnį ir mažesnį p_{T} turinčius leptonus, o η – pseudospartą. Iš detektoriaus užfiksuotų signalų atkurtų fizikinių objektų atranka skiriasi tarp elektronų ir miuonų kanalų dėl skirtingų trigerių ir objektų atkūrimo reikalavimų, tačiau dekonvoliucijos procedūra leidžia rezultatus suvesti į bendrą fazinės erdvės sritį.

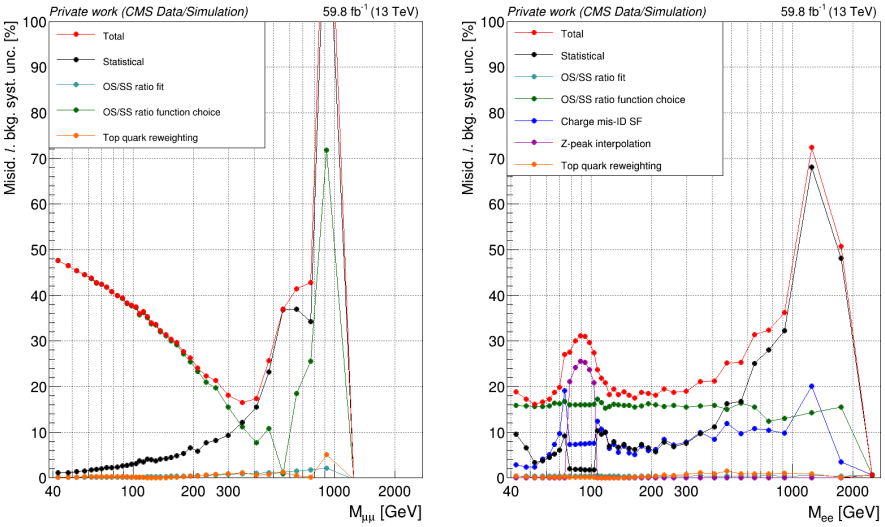
Svarbiausias autoriaus indėlis šioje analizėje buvo fono įvertinimas. Fono procesai, kuriuose rekonstruoti leptonai kyla iš tikrų elektrosilpnųjų bozonų skilimų, tokie, kaip silpnosios sąveikos bozonų poros (WW, WZ, ZZ) susidarymas, DY $\to \tau^+\tau^-$ ir fotonų indukuotas priešingo krūvio (OS) leptonų poros susidarymas, buvo įvertinti naudojant MC modeliavimą. Viršūninių kvarkų fonas taip pat buvo įvertintas naudojant MC modeliavimą, bet pakoreguotas matavimu grįstu $e\mu$ metodu. Foniniai procesai su klaidingai atpažintais leptonais buvo vertinami naudojant vienodo krūvio ženklo kontrolines sritis.

Miuonų kanale vienodo krūvio ženklo (SS) kontrolinėje srityje dominuoja QCD daugelio hadroninių džetų bei W bozono ir džetų (W+jets) įvykiai. Kadangi šių dviejų procesų OS/SS įvykių skaičiaus santykiai ir kinematiniai pasiskirstymai skiriasi, buvo sukurta dvimatė pernašos funkcija $R_{\mathrm{OS/SS}}(M_{\mu\mu}, p_{\mathrm{T}}^{\mathrm{miss}})$. Ji nustatyta atskiroje kontrolinėje srityje, kurioje vienas miuonas nepraeina izoliuotumo reikalavimo. Ši sritis pasirinkta todėl, kad joje esanti W+jets ir QCD tarpusavio proporcija yra panašiausia į signalo sritį. $R_{\mathrm{OS/SS}}(M_{\mu\mu}, p_{\mathrm{T}}^{\mathrm{miss}})$ funkcija buvo panaudota kaip svorinis daugiklis SS įvykiams, taip gaunant fono įvertį priešingo krūvio signalo srityje.

Elektronų kanale svarbus papildomas efektas yra klaidingas elektrono krūvio eksperimentinis nustatymas. Dėl jo tikri DY signalo įvykiai gali patekti į SS sritį, todėl prieš vertinant klaidingai atpažintų elektronų foną buvo pritaikyti klaidingo krūvio nustatymo korekciniai svoriai. Po šios pataisos likęs duomenų perteklius SS srityje interpretuotas kaip klaidingai atpažintų elektronų fonas. Elektronų kanale pernašos daugiklis buvo parametrizuotas kaip $R_{\mathrm{OS/SS}}(p_{\mathrm{T}}^{\mathrm{miss}})$,

o Z maksimumo srityje likęs artefaktas pakeistas tiesine interpoliacija tarp gretimų masių sričių.

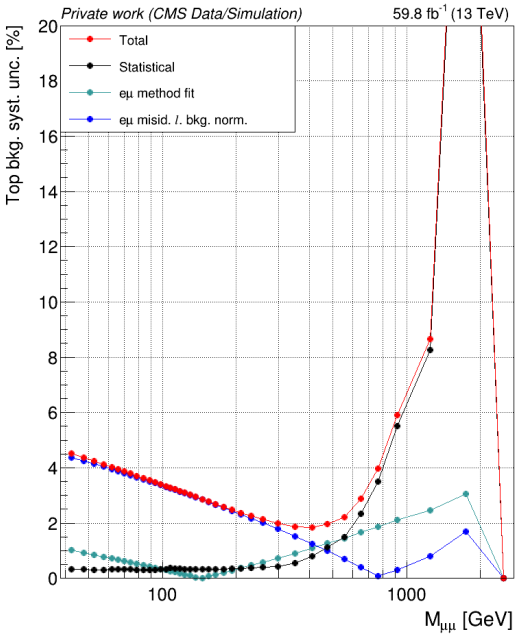
Svarbiausias šių matavimu grįstų metodų validumo patvirtinimo žingsnis yra fono įverčio neapibrėžtumų nustatymas. Miuonų kanale buvo įvertintas SS kontrolinės srities statistinis neapibrėžtumas, $R_{\mathrm{OS/SS}}(M_{\mu\mu}, p_{\mathrm{T}}^{\mathrm{miss}})$ funkcijos parametrų neapibrėžtumai bei pasirinktos $R_{\mathrm{OS/SS}}(M_{\mu\mu}, p_{\mathrm{T}}^{\mathrm{miss}})$ funkcinės išraiškos ir viršūninių kvarkų fono pataisos įtakos. Elektronų kanale papildomai įvertintas klaidingo elektrono krūvio nustatymo korekcinių svorių neapibrėžtumas ir interpoliacijos Z bozono masės srityje neapibrėžtumas. Šios dedamosios į bendrą neapibrėžtumą buvo susumuotos pagal Pitagoro teoremą, o gautos santykinės neapibrėžtys 2018 m. duomenims parodytos 2 pav. Miuonų kanale sisteminis neapibrėžtumas išlieka mažesnis nei 50%, o elektronų kanale – mažesnis nei 40% masių srityje $M_{\ell\ell} < 900$ GeV. Prie didesnių masių neapibrėžtumas sparčiai didėja dėl ribotos kontrolinių sričių statistikos, todėl būtent statistinė dedamoji ima dominuoti virš $M_{\ell\ell} \approx 300$ GeV.



2 pav.: Santykiniai klaidingai atpažintų leptonų fono neapibrėžtumai 2018 m. duomenims miuonų (kairėje) ir elektronų (dešinėje) kanaluose. Kitais duomenų rinkimo laikotarpiais gauti panašūs rezultatai.

Kitas svarbus fonas didelių masių srityje yra viršūninių kvarkų gamyba. Šio fono įvertinimui buvo pritaikytas modifikuotas $e\mu$ metodas. $e\mu$ kontrolinėje srityje pastebėta, kad MC modeliavimas numato per daug įvykių didelės $M_{e\mu}$ srityje. Todėl viršūninių kvarkų fonui buvo pritaikytas nuo masės priklausantis svoris, parametrizuotas kaip tiesinė funkcija nuo $\log_{10}(M_{\ell\ell})$. Ši pataisa tai-

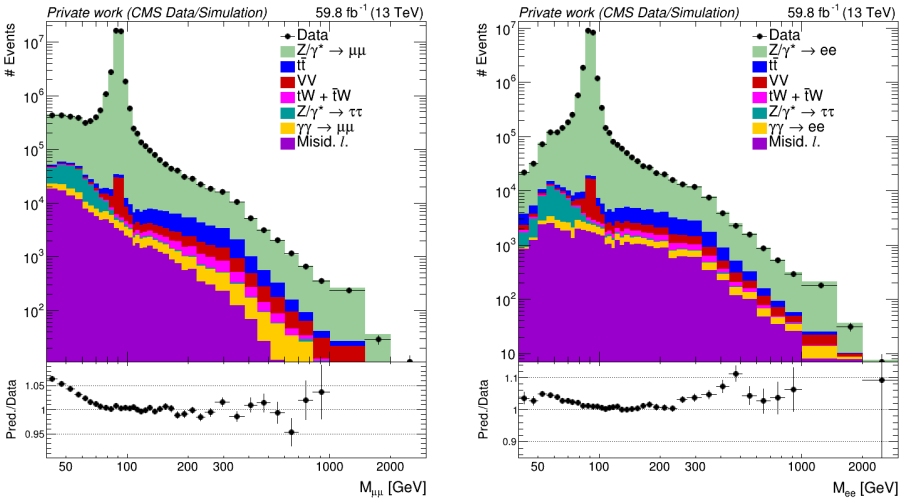
kyta $t\bar{t}$, tW ir $\bar{t}W$ procesams abiejuose kanaluose. Šios pataisos sisteminis neapibrėžtumas buvo įvertintas keičiant pritaikytos funkcijos parametrus jų paklaidų ribose ir varijuojant klaidingai atpažintų leptonų fono neapibrėžtį $e\mu$ kontrolinėje srityje 10% ribose. Gauta santykinė viršūninių kvarkų fono pataisos neapibrėžtis 2018 m. miuonų kanale parodyta 3 pav. Visame masių diapazone ji neviršija 6%.



3 pav.: Santykinė $e\mu$ metodu nustatytos viršūninių kvarkų fono pataisos sisteminė neapibrėžtis dimiuonų kanale 2018 m. duomenims. Kitais duomenų rinkimo laikotarpiais gauti panašūs rezultatai.

Pritaikius klaidingai atpažintų leptonų fono įvertį ir $e\mu$ metodu nustatytą viršūninių kvarkų fono pataisą, pagerėja duomenų ir bendros signalo bei fonų įverčio sutapimas. 4 pav. parodytas 2018 m. duomenų palyginimas su galutine prognoze miuonų ir elektronų kanaluose.

Įvertinus foną, pamatuoti masės pasiskirstymai buvo dekonvoliucijos būdu pataisyti dėl detektoriaus ribotos skyros ir įvykių migracijos tarp masės intervalų naudojant `TUnfold`[135] paketą. Detektoriaus atsako matrica buvo sudaryta pasinaudojant DY MC įvykiais, o keturi duomenų rinkimo laikotarpiai apjungti į vieną bendrą tiriamosios fazinės erdvės spektrą kiekvienam kanalui. Elektronų ir miuonų kanalų rezultatai buvo suderinti tarpusavyje ir sujungti naudojant hibridinį svorinio vidurkio bei BLUE [136] tipo kovariacijų propagavimo



4 pav.: 2018 m. duomenų ir galutinės signalo („$Z/\gamma^* \to \mu\mu$“ arba „$Z/\gamma^* \to ee$“) bei fono prognozės palyginimas miuonų (kairėje) ir elektronų (dešinėje) kanaluose. Kitais duomenų rinkimo laikotarpiais gauti panašūs rezultatai.

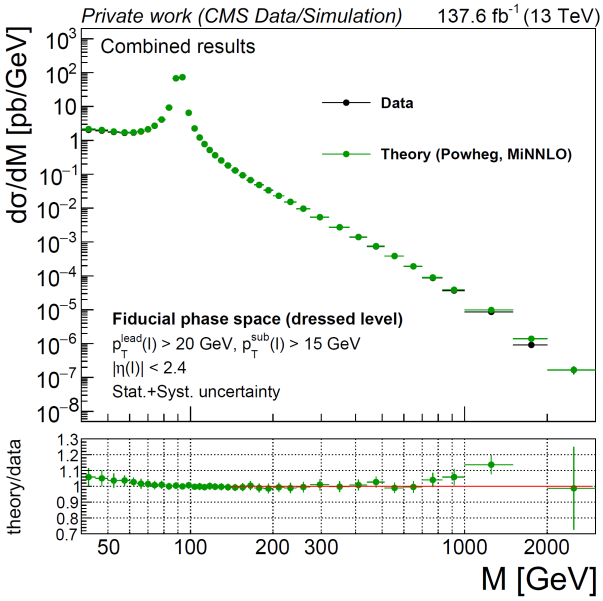
metodą.

Galutinis analizės rezultatas yra DY diferencialinis reakcijos skerspjūvis tiriamojoje fazinės erdvės srityje, pavaizduotas 5 pav. Jis bendrai dera su POWHEG MiNNLO prognoze neapibrėžčių ribose, nors didelių masių srityje statistinės fluktuacijos tampa svarbios. Šis rezultatas dokumentuotas CMS analizės užrašuose AN-22-177 [B1].

CMS detektoriaus antrosios fazės atnaujinimas

HL-LHC programa numato iki maždaug 5×10^{34} cm^{-2}s^{-1} momentinį šviesį, su galimybe pasiekti dar didesnes vertes [21]. Tai leis ATLAS ir CMS eksperimentams per visą programą surinkti protonų susidūrimų duomenų kiekius, atitinkančius maždaug 4000 fb^{-1} integruotąjį šviesį. Toks duomenų kiekis ženkliai pagerins retų procesų matavimus, Higso bozono savybių tyrimus, elektrosilpnųjų procesų analizę ir naujos fizikos paieškas.

Be minėtų privalumų, didesnis šviesis sukurs ir daug sudėtingesnes eksperimentines sąlygas. Vidutinis vienalaikių protonų susidūrimų skaičius sieks apie 140–200 kas 25 ns, o bendras sąveikų dažnis sieks kelis GHz. Taip pat smarkiai padidės radiacinė pažaida, ypač arčiausiai protonų spindulio esančiuose trekerio sluoksniuose. Todėl esamas CMS detektorius turi būti iš esmės



5 pav.: DY diferencialinis reakcijos skerspjūvis $\mathrm{d}\sigma/\mathrm{d}M_{\ell\ell}$ tiriamojoje fazinės erdvės srityje ir jo palyginimas su POWHEG MiNNLO prognoze.

modernizuotas [22].

CMS II fazės atnaujinimas apima viso trekerio pakeitimą, Aukštos erdvinės skyros kalorimetro (HGCAL) įdiegimą CMS antgalių srityse, miuonų detektorių sistemos išplėtimą, Aukštos laikinės skyros detektoriaus (MTD) įrengimą, pirmojo lygio (L1) trigerio ir duomenų nuskaitymo (DAQ) sistemų atnaujinimą bei programinės ir skaičiavimo infrastruktūros modernizavimą [22, 112, 143, 145, 147, 148]. Šioje disertacijoje daugiausia dėmesio skiriama trekeriui, nes būtent jis lemia elektringųjų dalelių trajektorijų, sąveikos taškų, leptonų izoliuotumo ir dalelių impulsų atkūrimo kokybę.

Naujasis II fazės CMS trekeris bus sudarytas iš Išorinio trekerio (OT) ir Vidinio trekerio (IT) [137]. OT moduliai galės pateikti informaciją L1 trigeriui, sudarydami lokalius trajektorijų segmentus, kurie atrenka daleles su p_{T} virš maždaug 2 GeV. IT bus sudarytas tik iš pikselinių modulių ir turės apie du milijardus kanalų. Jis turės išlaikyti aukštą efektyvumą labai didelio vienalaikių susidūrimų skaičiaus ir radiacinės apkrovos sąlygomis, kartu išplečiant trekerio kampinę aprėptį iki beveik $|\eta| < 4$.

IT duomenų nuskaitymo sistema turės valdyti apie 14 000 pikselinių lustų. Duomenys iš modulių bus perduodami per elektrines jungtis į lpGBT lustus [150], kurie duomenis apjungs ir optiškai perduos į duomenų surinkimo

elektroniką per VTRx+ siųstuvus-imtuvus [151]. Iš detektoriaus duomenis priims DTC plokštės, atsakingos už komunikaciją su pikseliniais lustais, trigerio signalų paskirstymą, duomenų pasaugojimą ir perdavimą centrinei CMS DAQ sistemai.

II fazės IT reikalingas naujos kartos radiacijai atsparus pikselinis lustas, galintis veikti esant labai dideliems dalelių srautams ir duomenų perdavimo spartoms. Šiam tikslui RD53 programa, jungianti CMS ir ATLAS eksperimentų pastangas, sukūrė 65 nm CMOS technologija paremtą pikselinių lustų šeimą [23, 152]. Pirmasis reikšmingas prototipas RD53A buvo naudojamas analoginės lusto dalies architektūrų ir skaitmeninės nuskaitymo logikos įvertinimui [153, 154]. CMS pasirinko tiesinę analoginės dalies architektūrą, kuri vėliau buvo įgyvendinta specialiai CMS reikmėms skirtame RD53B-CMS prototipe, dar vadinamame CROCv1 [155, 156]. Galutinė CMS lusto versija – RD53C-CMS, arba CROCv2 – bus skirta naudoti II fazės CMS Vidinio trekerio pikseliniuose moduliuose.

Pasiruošimo CMS vidinio trekerio atnaujinimui veikla

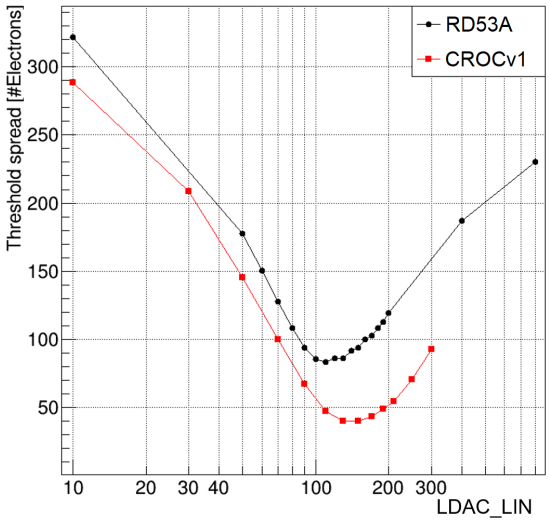
Pasirengimo darbams naudota Ph2-ACF programinė įranga, skirta II fazės IT modulių konfigūravimui, kalibravimui, duomenų kaupimui ir veikimo kokybės stebėsenai [24]. Ph2-ACF veikia kaip tarpinis sluoksnis tarp naudotojo lygmens kalibravimo įrankių ir žemo lygio komunikacijos su pikseliniais lustais bei duomenų nuskaitymo elektronikos plokštėmis. Sistema naudoja hierarchinį detektoriaus aprašymą, leidžiantį tas pačias kalibravimo procedūras taikyti tiek vienam lustui, tiek visam moduliui ar didesniam testavimo stendui.

Autorius prisidėjo prie Ph2-ACF kūrimo, testavimo, tobulinimo, kalibracijos procedūrų veikimo kokybės tikrinimo ir dokumentavimo. Didžioji CMS eksperimento vidinėms reikmėms skirto Ph2-ACF naudotojo vadovo dalis taip pat parengta autoriaus [160]. Autorius sukūrė ir Ph2-ACF esantį duomenų perdavimo linijų diagnostinį įrankį (vadinamą *Data Transmission Test*). Jis leidžia kiekybiškai įvertinti duomenų perdavimo linijos kokybę skenuojant išėjimo signalo amplitudės nustatymą ir nustatant pereinamosios srities padėtį tarp nesėkmingo bei be klaidų vykstančio duomenų perdavimo.

Toliau pristatomi pagrindiniai CROCv1 prototipų testavimo rezultatai. Dalis matavimų palyginimui buvo pakartoti naudojant RD53A lustą. Testai atlikti laboratorijose CERN ir Vilniaus universiteto Branduolių ir elementariųjų dalelių fizikos centre. Pagrindiniai rezultatai susiję su signalo aptikimo slenksčio kalibravimu, signalo stiprio charakteristikų matavimu ir krūvio injekavimo grandinės kalibravimu. Šie rezultatai buvo pristatyti keliose publikacijose [A2,

B2, B3]. Kitus publikuotus rezultatus, prie kurių autorius prisidėjo kurdamas, tobulindamas, tikrindamas ir dokumentuodamas kalibracijos procedūras Ph2-ACF programinėje įrangoje, galima rasti nuorodose [A3, A4].

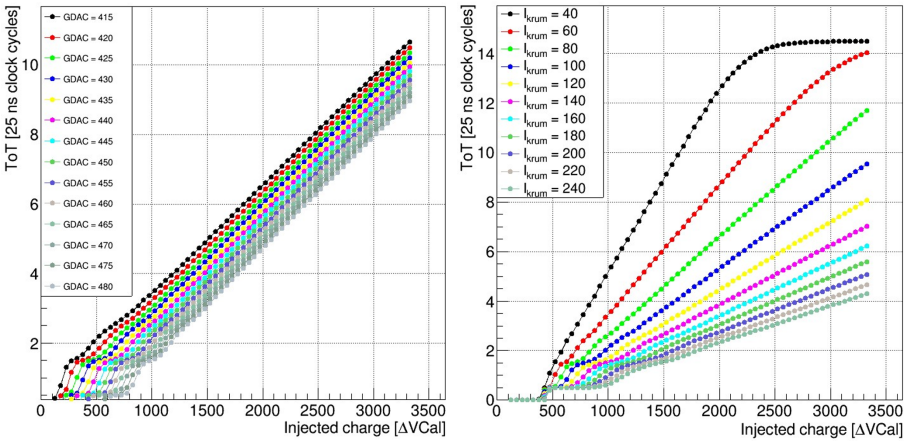
Signalo aptikimo slenksčio kalibravimo tyrimai parodė, kad CROCv1 lustas po TDAC registro kalibravimo gali pasiekti slenksčių pasiskirstymą, kurio standartinis nuokrypis siekia maždaug 40 elektronų. Tai yra apie du kartus geresnis rezultatas nei RD53A luste. Ši savybė svarbi, nes mažas slenksčių išsibarstymas leidžia dirbti naudojant žemesnį slenkstį ir išlaikyti aukštą efektyvumą bei kontroliuojamą triukšmą. Tipinis RD53A ir CROCv1 slenksčių pasiskirstymų standartinių nuokrypių palyginimas, esant skirtingoms LDAC registro reikšmėms, pateiktas 6 pav.



6 pav.: Tipinis RD53A ir CROCv1 lustų signalo aptikimo slenksčių pasiskirstymų standartinių nuokrypių palyginimas esant skirtingoms LDAC registro reikšmėms [A2, B2, B3].

Toliau tirta signalo stiprio (ToT) atsako priklausomybė nuo į pikselį injekuoto krūvio, signalo aptikimo slenksčio ir signalo stiprintuvo grandinėje esančio kondensatoriaus iškrovimo (Krummenacher) srovės. Nustatyta, kad ToT priklausomybė nuo krūvio yra tiesinė, kai surinktas krūvis yra reikšmingai didesnis už slenkstinę vertę. Taip pat pastebėta, kad tiesės posvyris nepriklauso nuo slenksčio padėties, o pagrindinis signalo stiprinimą valdantis parametras yra kondensatoriaus iškrovimo srovė. Tai reiškia, kad slenksčio ir signalo stiprio derinimą galima atlikti beveik nepriklausomai. Šios priklausomybės parodytos 7 pav.: kairėje pateiktos kreivės rodo, kad slenksčio keitimas nepakeičia ToT atsako nuolydžio, o dešinėje matyti kondensatoriaus iškrovimo srovės įtaka

signalo stiprinimui.

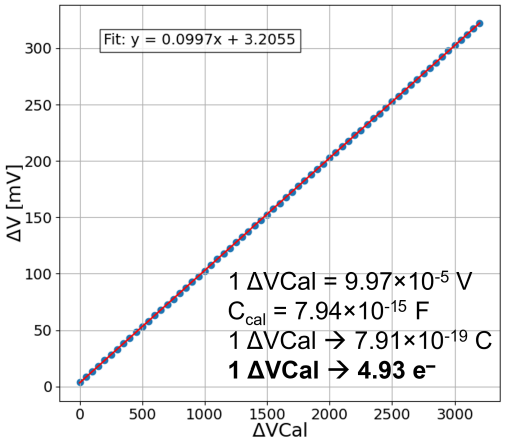


7 pav.: Vidutinė signalo stiprio priklausomybė nuo į pikselį injekuoto krūvio. Kairėje: esant skirtingoms signalo aptikimo slenksčio reikšmėms; dešinėje: esant skirtingoms signalo stiprintuvo grandinės kondensatoriaus iškrovimo srovės reikšmėms [B3].

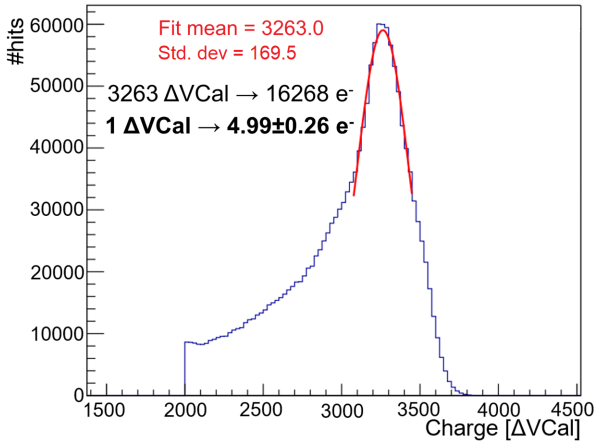
CROCv1 lustai turi savikalibracijos grandinę, kuri injekuoja nustatyto dydžio krūvį į norimą pikselį, taip sukurdama sintetinius signalus. Injekuoto krūvio dydį kontroliuoja skaitmeninis parametras Δ`VCal`. Šio parametro kalibravimui naudoti du nepriklausomi metodai.

Pirmiausia, pasinaudojant luste esančiu įtampos multipleksoriumi (VMUX), buvo išmatuota kalibracinės įtampos voltais priklausomybė nuo Δ`VCal` parametro vertės. Tuo pačiu metodu pamatavus ir kalibracinę talpą, gauta, jog 1 Δ`VCal` vienetas atitinka 4.93 elektrono. Ši tiesinė priklausomybė parodyta 8 pav.

Antrasis metodas leido įvertinti Δ`VCal` vertę panaudojant monochromatinį Rentgeno spindulių šaltinį ^{241}Am. Monochromatiniai Rentgeno fotonai detektoriaus pikselyje sukuria apytiksliai vienodo dydžio signalą. Įvertinus, kiek elektronų turėtų silicyje išlaisvinti vienas ^{241}Am fotonas ir kokio dydžio signalą jis paliko detektoriuje, buvo nustatyta, kad vienas Δ`VCal` vienetas atitinka 4.99 ± 0.26 elektrono. Išmatuotas ^{241}Am signalo stiprio pasiskirstymas pateiktas 9 pav. Tai, kad dvi nepriklausomos metodikos davė artimus įverčius, pagrindžia 5 elektronų vienam Δ`VCal` vienetui konvertavimo daugiklio naudojimą CROCv1 kalibravimui.



8 pav.: Lusto savikalibracijos grandinės įtampos priklausomybė nuo `ΔVCal` parametro vertės [B3].



9 pav.: ^{241}Am rentgeno spindulių detektoriuje sukelto signalo stiprio pasiskirstymas [B3].

CMS detektoriaus II fazės atnaujinimo įtaka Drell–Yan proceso matavimo tikslumui

Šiame skyriuje pateikiamas įvertinimas, kaip II fazės CMS detektorius gali paveikti būsimų DY matavimų tikslumą. Tam CMSSW aplinkoje sukurta pilna MC modeliavimo grandinė procesui $pp \to Z/\gamma^* \to \mu^+\mu^-$. Įvykiai sumodeliuoti naudojant POWHEG MiNNLO [55, 56, 68], PYTHIA8 [48] ir PHOTOS [123] modeliavimo programas, detektoriaus atsakas sumodeliuotas naudojant GEANT4 [125], o įvykių atkūrimas atliktas naudojant realistišką II fazės CMS konfigūraciją, 1000 fb^{-1} detektoriaus nusidėvėjimo modelį ir vidutinį 200 vienalaikių protonų susidūrimų skaičių. Gauta II fazės imtis palyginta su sumodeliuota CMS 2018 metų etalonine imtimi, paruošta taip, kad būtų galima tiesiogiai palyginti detektoriaus ir įvykių atkūrimo poveikį stebimiems DY dydžiams.

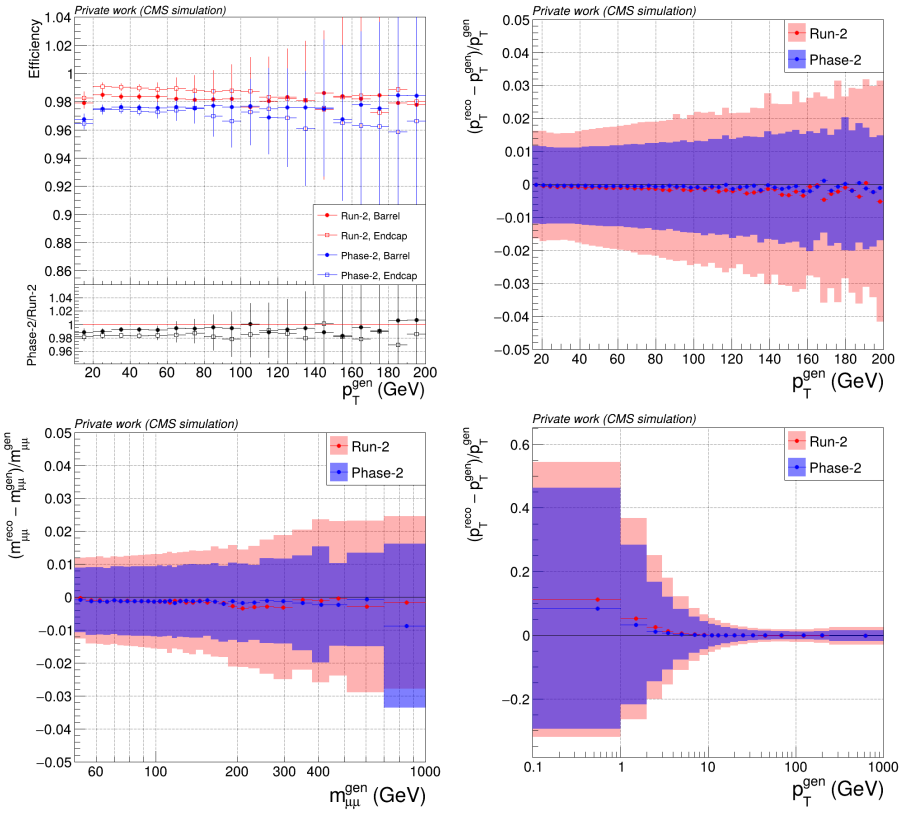
Tyrimas atliktas miuonų kanale, nes miuonų atkūrimo kokybė II fazės detektoriaus modelyje jau yra pakankamai išvystyta, kad tiktų prasmingam palyginimui. Pirmiausia įvertintas „globalaus“ miuono (užregistruoto tiek trekeryje, tiek miuonų detektoriuje) atkūrimo efektyvumas. Nustatyta, kad II fazės aplinkoje, net esant 200 vienalaikių protonų susidūrimų, efektyvumas sumažėja tik apie 1%, lyginant su 2018 m. etalonu. Tai rodo, kad modernizuotas detektorius išlaiko patikimą miuonų atkūrimą sudėtingoje HL-LHC aplinkoje.

Taip pat ištirta vieno miuono ir miuonų poros kinematinių dydžių skyra. Atsako pasiskirstymai buvo apskaičiuoti kaip sumodeliuoto ir iš detektoriaus signalų atkurto dydžių santykinis skirtumas. Atsakas kiekviename sumodeliuoto dydžio intervale aproksimuotas Gauso pasiskirstymu. Priderinto pasiskirstymo vidurkis naudotas kaip atkūrimo skalės įvertis, o plotis σ — kaip skyros įvertis. Pagrindiniai efektyvumo ir skyros palyginimai pateikti 10 pav.

II fazės imtyje globalaus miuono atkūrimo efektyvumas išlieka labai artimas 2018 m. etalonui, o pagrindiniai kinematiniai dydžiai rekonstruojami su geresne skyra. Miuono p_T skyra pagerėja maždaug 30–40%, miuonų poros invariantinės masės skyra – apie 20–40%, o miuonų poros p_T skyra – apie 15–40%, priklausomai nuo nagrinėjamos kinematinės srities. Pagerėjusi skyra mažina detektoriaus nulemtą masės ir impulso spektrų matavimų neapibrėžtį ir gerina dekonvoliucijos stabilumą. Kartu su gerokai didesne HL-LHC statistika tai turėtų reikšmingai pagerinti būsimų DY matavimų tikslumą.

Išvados

Disertacijoje pateikti rezultatai apima tiek dabartinius didelio tikslumo DY proceso matavimus CMS eksperimente, tiek pasirengimą būsimiems HL-LHC



10 pav.: 2018 m. ir II fazės palyginimas: globalaus miuono atkūrimo efektyvumas (viršuje kairėje), miuono p_{T} skyra (viršuje dešinėje), miuonų poros invariantinės masės skyra (apačioje kairėje) ir miuonų poros p_{T} skyra (apačioje dešinėje).

laikotarpio matavimams. CMS Antrojo duomenų rinkimo etapo (2016–2018 m.) DY analizėje sukurti ir pritaikyti matavimu grįsti fono įvertinimo metodai, leidžiantys kontroliuoti klaidingai atpažintų leptonų foną elektronų ir miuonų kanaluose. Parodyta, kad pateikta vienodo krūvio ženklo metodika leidžia šį foną įvertinti taip, kad sisteminis neapibrėžtumas išliktų mažesnis nei 50% srityje $M_{\ell\ell} < 900$ GeV. Didesnėse masių srityse neapibrėžtumą pradeda dominuoti kontrolinių imčių statistika. Tai pagrindžia pirmuosius du disertacijos teiginius. Taip pat įgyvendinta $e\mu$ kontrolinės srities pagrindu paremta viršūninių kvarkų fono pataisa, sumažinanti didelių masių srityje stebėtą duomenų ir simuliacijos neatitikimą.

II fazės CMS Vidinio trekerio atnaujinimo dalyje parodyta, kad CROCv1 nuskaitymo lustas leidžia pasiekti maždaug du kartus siauresnį signalo aptikimo slenksčių pasiskirstymą nei RD53A. Taip pat nustatyta, kad jo signalo stiprio priklausomybė nuo pikselyje surinkto krūvio yra tiesinė ir iš esmės nepriklauso nuo signalo aptikimo slenksčio. Vidinės krūvio injekavimo grandinės kalibravimas dviem nepriklausomais metodais patvirtino maždaug 5 elektronų per Δ`VCal` žingsnį. Šie rezultatai pagrindžia trečiąjį, ketvirtąjį ir penktąjį disertacijos teiginius.

Galiausiai, II fazės DY MC modeliavimo tyrimas parodė, kad modernizuotas CMS detektorius turėtų išlaikyti labai aukštą miuonų atkūrimo efektyvumą ir tuo pačiu ženkliai pagerinti miuono bei miuonų poros kinematinių dydžių skyrą. Miuono skersinio impulso skyra pagerėja maždaug 30–40%, o miuonų poros invariantinės masės skyra – apie 20–40%, priklausomai nuo nagrinėjamos kinematinės srities. Šie rezultatai pagrindžia šeštąjį disertacijos teiginį. Tai reiškia, kad HL-LHC laikotarpiu, kartu su daug didesne statistika, DY miuonų kanalas išliks itin svarbus ir patikimas įrankis tiksliems SM matavimams bei galimų nukrypimų nuo SM paieškoms.

Disertacijoje parodytas nuoseklus ryšys tarp dabartinių CMS tikslumo matavimų, II fazės detektoriaus technologinio pasirengimo ir būsimų HL-LHC fizikos galimybių. Sukurti fono vertinimo metodai pagerina dabartinių DY matavimų patikimumą, o atlikti pikselinių lustų ir II fazės detektoriaus tyrimai prisideda prie būsimos CMS matavimų programos, kurioje DY procesas išliks svarbiu Standartinio modelio tikslumo patikrinimo įrankiu.

CURRICULUM VITAE

Marijus Ambrozas
Email: marijus.ambrozas@cern.ch

Education

Faculty of Physics, Vilnius University

PhD in Physics	2020 – 2026
Master of Physical Sciences	2018 – 2020
Bachelor of Physics	2014 – 2018

Faculty of Philosophy, Vilnius University

Physics teacher qualification	2018 – 2019

Tauragė "Versmės" gymnasium

High school diploma	2010 – 2014

Conference presentations and posters

CERN Baltic Conference

Plenary talk "Prospects for Drell-Yan Precision Measurements with the CMS Phase-2 Detector" at the HL-LHC"	Oct. 2025
Plenary talk "CMS Phase-2 Inner Tracker Pixel Chip Prototype Tests"	Oct. 2023
Plenary talk "Characterization and Calibration of the RD53B CMS Chip"	Oct. 2022
Plenary talk "Estimating the Fake Lepton Backgrounds for the Drell-Yan Differential Cross Section Measurement Using 2016 CERN CMS Data"	Oct. 2021

International Conference of High-Energy Physics (ICHEP)

Poster "Data-driven Dilepton Background Estimation Improvements for the Drell-Yan Analysis at CMS"	Jul. 2024

Lithuanian National Physics Conference

Parallel talk "Pixel Detector Chip Calibration for the CMS Phase-2 Upgrade"	Oct. 2023
Parallel talk "Non-Prompt Lepton Background Estimation for Drell-Yan Differential Cross-Section Measurement"	Oct. 2021
Parallel talk "Drell-Yan Process Background Estimation Using eμ Method"	Oct. 2019

Parallel talk "Data-Driven Methods to Estimate the Number of Drell-Yan Process Background Events" — Oct. 2017

EPS Conference on High Energy Physics (EPS-HEP)

Poster "CROC_v1 Pixel Chip Calibration for the Phase-2 CMS Inner Tracker" — Aug. 2023

International students' conference "Open Readings"

Poster "Calibration of Pixel Detector Prototypes for the Phase-2 CMS Inner Tracker Upgrade" — Mar. 2022

Parallel talk "Improved Fake Lepton Background Estimation for the Drell-Yan Differential Cross-Section Measurement Using 2016 CERN CMS Data" — Mar. 2021

Poster "Drell-Yan process analysis using 2016 CERN CMS proton-proton collision data" — Mar. 2019

Poster "Drell-Yan Process Background Estimation Using Data-Driven Methods" — Mar. 2018

Schools

International School of Trigger and Data Acquisition
Vilnius, Lithuania — Jun. 2025
* Part of the organizing committee

MCnet Summer School
CERN, Switzerland — Jun. 2024

Terascale Statistics School
DESY, Germany — Apr. 2024

European School of High Energy Physics
Grenaa, Denmark — Sep. 2023

Baltic School of High-Energy Physics and Accelerator Technologies
Klapkalnciems, Latvia — Aug. 2021

Computer Algebra and Particle Physics
Remote — Apr. 2021

4th Tracker Upgrade DAQ School
Remote — Mar. 2021

CMS Data Analysis School
DESY, Germany — Oct. 2018

CERN Spring Campus
RTU, Latvia — Apr. 2018

Training

Workshops and visits at IIHE, Belgium

Work on the Drell-Yan differential cross section measurement

One-week visit	Jan. 2026
Three-day workshop	Nov. 2025
Four-day workshop	Sep. 2024

Training visits at PSI, Switzerland

Work on module tests for the Tracker Phase-2 Upgrade

One-week training	Nov. 2024
Two-day visit	Jan. 2024

Training visits at CERN, Switzerland

Work on system tests for the Tracker Phase-2 Upgrade

Three-week training	Nov. 2023
Three-week training	Aug. 2022

Student research project grants

Vilnius University, Lithuania

Grant agreements with the Research Council of Lithuania

Project No 09.3.3-LMT-K-712-16-0131	2019 – 2020
Project No 09.3.3-LMT-K-712-10-0128	2018 – 2019

Erasmus+ internship

CERN Experimental Physics Department, Switzerland

Work on CMS Outer Tracker Phase-2 upgrade	Summer
Two-month internship	2019

Teaching

Introduction to Programming

Exercise sessions for 1st-year BSc students	2021 – now

Numerical Methods

Coordinating teacher and lecturer for 1st-year BSc students	2023 – now

Numerical Optimization Methods in Physics

Exercise sessions for 1st-year MSc students	2021 – 2023

Supervision of students' bachelor's theses

3 students successfully defended their theses	2025 – now

Outreach

IPPOG International CMS Masterclass

Organizer, lecturer and coordinator in a yearly event	2018 – now

Science festival "Spaceship Earth"

Yearly outreach lectures on particle physics	2020 – now

IPPOG International Particle Therapy Masterclass

Organizer and coordinator in a yearly event	2023 – now

Seminars for high-school physics teachers

Bi-monthly seminars on experimental particle physics	2024 – now

Students' camps

Lecture on particle physics in "Fotonas" students' camp	Jul. 2025
Lectures on particle physics in Lithuania's National Students' Academy camps	Feb. 2025, July 2021

Conference "Let's Know Beyond"

Outreach lecture on particle physics	October 2021

ACKNOWLEDGEMENTS/PADĖKOS

This thesis wraps up the most difficult six years of my life (yet). I would have never managed to get through this path on my own.

I would like to express my sincere gratitude to my supervisor, Andrius Juodagalvis, for teaching me how to think like a scientist and a leader, for his trust, patience, and, very importantly, for keeping my financial status stable. I am also grateful to the members of the CMS Collaboration and the SMP-V Drell–Yan working group for the opportunity to contribute to a large international research effort and to learn from many excellent physicists. I am especially thankful to Kyeongpil Lee and Laurent Favart of the Drell–Yan working group for leading the analysis and making me feel like an important member of the group.

I would like to thank Stefano Mersi for initially accepting me as a summer student and then inviting me to join the CMS Tracker Phase–2 upgrade efforts as a PhD student, supervising my work at CERN, and teaching me everything I know about the specifics of the CMS Tracker. I would also like to thank Mauro Dinardo for his guidance and support in developing and using the Ph2-ACF framework.

Additional thanks go to my TFAI work colleagues: Thomas, Vytautas, Darius, and Ugnius. Thank you for the nice discussions at work and during lunch time, and for helping me with the outreach activities. Special thanks go to Aušra for initiating the outreach activities that I later needed help with.

I am also thankful to the previous and current (at the time of writing) vice-deans for studies, Olga and Justinas, for always being supportive, helpful, and caring. Entering your office has never felt stressful or uncomfortable.

I would like to thank my undergraduate students for keeping me busy. Supervising and guiding you has been a learning experience for me, too, and I hope that at least some of my passion for physics was passed down to you.

I would also like to thank my friends, who made this long path much more bearable and meaningful. I am grateful to my friends who have been there since school days: Simas, Rimvydas, Linas, Enrika, and Ingrida. I am equally grateful to my friends whom I met in Vilnius: Simonas, Šarūnė, Domantas, Gintarė, Matas, Greta, Vytautas, Arnoldas, Ignas, Justas, Aurelija, the other Ignas, the other Greta, Ieva, Vilius, Gerda, Birutė, Fjora, Vincentas, Julija, Tomas, Rasa, and Darius. Your conversations, humor, encouragement, and shared moments helped me keep perspective outside the world of physics.

A special part of my life during these years has been the Vilnius University wind orchestra “Oktava”. I would like to thank the orchestra community for

giving me a place to grow musically, rest mentally, and feel part of something joyful and creative. I am especially grateful to Albinas Kučinskas and Donatas Viaževičius for their leadership, dedication, and inspiration.

I would like to thank my psychotherapist, Akvilė, for helping me pick myself up when I was at my lowest point.

Finally, I am most grateful to my family. I thank my parents for their unconditional support, encouragement, and belief in me throughout all stages of my life and studies. I thank my brother Eugenijus and my sisters Vijolė and Julija for always making me feel part of the family. I am especially thankful to my wife, Aneta, for her love, patience, understanding, and support during the most difficult periods of this work. I couldn't have done it without you.

That said, I hope Lithuania improves its education system to better support PhD students working in large collaborations such as CMS. The current scientific publication requirements for PhD candidates do not fit the specifics of work at CERN and can lead to difficult situations that take a serious psychological toll. I hope that the PhD students who come after me will have an easier time.

Anyway, this has not discouraged me from continuing my work in science. It has only made me more resilient. See you around in big collaboration papers.

LIST OF WoS PUBLICATIONS

[A1] CMS Collaboration, "Measurement of the Mass Dependence of the Transverse Momentum of Lepton Pairs in Drell-Yan Production in Proton-Proton Collisions at $\sqrt{s} = 13$ TeV", *European Physical Journal C*, vol. 83, no. 7, p. 628, 2023, ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-023-11631-7.

[A2] G. Alimonti *et al.*, "RD53 Pixel Readout Integrated Circuits for ATLAS and CMS HL-LHC Upgrades", *Journal of Instrumentation*, vol. 20, no. 03, P03024, 2025, ISSN: 1748-0221. DOI: 10.1088/1748-0221/20/03/P03024, Corrected in G. Alimonti, M. Ambrozas *et al.*, "Erratum: RD53 Pixel Readout Integrated Circuits for ATLAS and CMS HL-LHC Upgrades", *Journal of Instrumentation*, vol. 21, no. 02, E02002, 2026, ISSN: 1748-0221.
DOI: 10.1088/1748-0221/21/02/E02002.

[A3] The Tracker Group of the CMS Collaboration, "Measurement of the Fractional Radiation Length of a Pixel Module for the CMS Phase-2 Upgrade via the Multiple Scattering of Positrons", *Journal of Instrumentation*, vol. 19, no. 10, P10023, 2024.
DOI: 10.1088/1748-0221/19/10/P10023.

[A4] The Tracker Group of the CMS Collaboration, "Evaluation of 3D Pixel Silicon Sensors for the CMS Phase-2 Inner Tracker", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 1088, p. 171 359, 2026, ISSN: 0168-9002. DOI: 10.1016/j.nima.2026.171359.

LIST OF OTHER PUBLICATIONS

[B1] M. Ambrozas *et al.*, "Measurement of the Differential Drell-Yan Mass Cross Section with the Full Run2 Data", CMS Analysis Note, CMS AN-22-177, *non-public document*, 2025.

[B2] L. Gaioni, M. Ambrozas *et al.*, "Design and Test of the CROCv1 Analog Front-End Processor for the CMS Pixel Readout at the HL-LHC", in *2022 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, 2022, pp. 1–3.
DOI: 10.1109/NSS/MIC44845.2022.10398954.

[B3] M. Ambrozas, “Calibration of the CROC_v1 Pixel Chip for the CMS Phase-2 Inner Tracker”, *Proceedings of Science*, vol. EPS-HEP2023, p. 581, 2024. DOI: 10.22323/1.449.0581.

LIST OF PRESENTATIONS

∗ marks international conferences

[C1] M. Ambrozas and A. Juodagalvis, “Improved Fake Lepton Background Estimation for the Drell-Yan Differential Cross Section Measurement Using 2016 CERN CMS Data”, available: https://openreadings.eu/wp-content/abstracts/abstract-2021.pdf, p. 56, talk, presented at the Open Readings 2021 (Vilnius, Lithuania (remote)), Mar. 17, 2021.

[C2] ∗ M. Ambrozas and A. Juodagalvis, “Estimating the Fake Lepton Backgrounds for the Drell-Yan Differential Cross Section Measurement Using 2016 CERN CMS Data”, available: https://indico.cern.ch/event/970609/contributions/4415898/, talk, presented at the 1st CERN Baltic Conference (Tartu, Estonia (remote)), Jun. 29, 2021.

[C3] M. Ambrozas and A. Juodagalvis, “Foninių procesų su netikrais leptonais įvertinimas Drell-Yan proceso diferencialinio reakcijos skerspjūvio matavime”, available: http://lnfk.ftmc.lt/wp-content/uploads/2021/11/LNFK-44-pranesimu-medziaga-2021-spalio-6-8-Vilnius.pdf, p. 80, talk, presented at the 44-oji Lietuvos nacionalinė fizikos konferencija (Vilnius, Lithuania), Oct. 7, 2021.

[C4] M. Ambrozas and S. Mersi, “Calibration of Pixel Detector Prototypes for the Phase-2 CMS Inner Tracker Upgrade”, available: https://openreadings.eu/wp-content/abstracts/abstract-2022.pdf, p. 140, poster, presented at the Open Readings 2022 (Vilnius, Lithuania (remote)), Mar. 16, 2022.

[C5] ∗ M. Ambrozas and S. Mersi, “Characterization and Calibration of the RD53B CMS Chip”, available: https://indico.cern.ch/event/1147717/contributions/5082290/, talk, presented at the 2nd CERN Baltic Conference (Vilnius, Lithuania), Oct. 12, 2022.

[C6] ∗ M. Ambrozas, “CROC_v1 Pixel Chip Calibration for the Phase-2 CMS Inner Tracker”, available: https://indico.desy.de/event/34916/contributions/150119/, poster, presented at the European Physical Society Conference on High Energy Physics 2023 (Hamburg, Germany), Aug. 22, 2023.

[C7] ∗ M. Ambrozas, “CMS Phase-2 Inner Tracker Pixel Chip Prototype Tests”, available: https://indico.cern.ch/event/1288731/contributions/5585494/, talk, presented at the 3rd CERN Baltic Conference (Riga, Latvia), Oct. 10, 2023.

[C8] M. Ambrozas, “Pikselinių detektorių lustų prototipų, skirtų CMS II fazės atnaujinimui, kalibracijos tyrimas”, available: https://www.lietuvos-fizikai.lt/chessidr/2026/04/LNFK45_teziu_knyga-opt.pdf p. 71, talk, presented at the 45-oji Lietuvos nacionalinė fizikos konferencija (Vilnius, Lithuania), Oct. 27, 2023.

[C9] ∗ M. Ambrozas, “Data-driven Dilepton Background Estimation Improvements for the Drell-Yan Analysis at CMS”, available: https://indico.cern.ch/event/1291157/contributions/5878682/, poster, presented at the 42nd International Conference of High-Energy Physics (Prague, Czech Republic), Jul. 22, 2024.

[C10] ∗ M. Ambrozas, “Prospects for Drell-Yan Precision Measurements with the CMS Phase-2 Detector at the HL-LHC”, available: https://indico.cern.ch/event/1510460/contributions/6702887/, talk, presented at the 5rd CERN Baltic Conference (Kaunas, Lithuania), Oct. 15, 2025.

Vilniaus universiteto leidykla
Saulėtekio al. 9, III rūmai, LT-10222 Vilnius
El. p. info@leidykla.vu.lt, www.leidykla.vu.lt
bookshop.vu.lt, journals.vu.lt
Tiražas 15 egz.