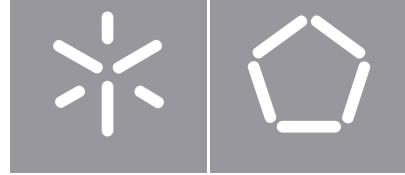




University of Minho
School of Engineering

Fábio Lucas Pereira Carneiro

**Luminosity Calibrations at the CERN
Compact Muon Solenoid Experiment**



University of Minho
School of Engineering

Fábio Lucas Pereira Carneiro

**Luminosity Calibrations at the CERN
Compact Muon Solenoid Experiment**

Master's in Physics Engineering

Specialization in Physics Information
Dissertation supervised by
Nuno Filipe Da Silva Fernandes De Castro
Andres Guillermo Delannoy Sotomayor

Copyright and Terms of Use for Third Party Work

This dissertation reports on academic work that can be used by third parties as long as the internationally accepted standards and good practices are respected concerning copyright and related rights.

This work can thereafter be used under the terms established in the license below.

Readers needing authorization conditions not provided for in the indicated licensing should contact the author through the RepositóriUM of the University of Minho.

License granted to users of this work:



CC BY

<https://creativecommons.org/licenses/by/4.0/>

Acknowledgements

First and foremost, I would like to express my gratitude to my supervisor Nuno Filipe Da Silva Fernandes De Castro. I am also profoundly grateful to my co-supervisor at CERN, Andres Guillermo Delannoy Sotomayor, for providing me with the opportunity to work at the forefront of high-energy physics. His expertise, mentorship, and constructive criticism have greatly enriched my understanding of the field and have been pivotal in shaping this work.

I would like to extend my sincere thanks to the entire BRIL team at CERN for welcoming me into the group and for their exceptional camaraderie. Your willingness to share knowledge, the countless hours of fruitful discussions, and your patient guidance on every aspect of the work have been essential in helping me navigate the complexities of this project. I have learned immensely from each of you, and this work would not have been possible without your collective support.

Lastly, I would like to thank my family and friends for their continuous encouragement and understanding throughout this journey. Their support has been a constant source of strength.

Thank you all.

Statement of Integrity

I hereby declare having conducted this academic work with integrity.

I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration.

I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

University of Minho, Braga, march 2025

Fábio Lucas Pereira Carneiro

Abstract

The quest to understand the fundamental nature of the universe drives the field of particle physics, centered around the Standard Model (SM), which describes the fundamental particles and three of the four known forces: electromagnetism, the weak force, and the strong force. Despite its success in explaining a wide range of phenomena, the SM does not address some of the most profound questions in physics, such as the nature of dark matter, the origin of mass hierarchies, and the unification of forces. Probing the SM with precision measurements is essential for uncovering potential new physics beyond the current theoretical framework.

Particle colliders like the Large Hadron Collider (LHC) at CERN are at the forefront of this exploration, where measurements of cross sections are a crucial part of the physics program that tests the SM's predictions. A cross section quantifies the likelihood of specific particle's interactions, and accurate luminosity measurements are fundamental to determining these cross sections. Luminosity, \mathcal{L} , is proportional to the rate of collisions, and thus any uncertainty in \mathcal{L} translates into uncertainties in the cross sections, affecting the reliability of the conclusions drawn from experimental data.

This thesis presents the preliminary calibration and measurement of luminosity for the 2023 proton-proton collision data-taking period at a center-of-mass energy of 13.6 TeV with the CMS experiment. The absolute luminosity scale is established using the van der Meer (vdM) scan methodology, a technique that involves scanning the particle beams across each other to measure their overlap and calibrate the visible cross sections of the luminometers. A series of detectors were employed to provide real-time luminosity measurements, with systematic corrections and stability analyses ensuring consistent and accurate results

The final integrated luminosity was measured as 32.74 fb^{-1} with a total uncertainty of 1.28%, representing the best preliminary luminosity results achieved by CMS. In addition to the experimental efforts, substantial software improvements were made to the codebases used in the analysis. These advancements enhanced data processing efficiency and streamlined the calibration workflow, enabling more rapid and reliable analysis iterations.

Keywords CERN, CMS, Luminosity, van der Meer

Resumo

A busca por compreender a natureza fundamental do universo impulsiona o campo da física de partículas, centrado no Modelo Padrão (SM), que descreve as partículas fundamentais e três das quatro forças conhecidas: o electromagnetismo, a força fraca e a força forte. Apesar do seu sucesso em explicar uma ampla gama de fenómenos, o SM não aborda algumas das questões mais profundas da física, como a natureza da matéria escura, a origem das hierarquias de massa e a unificação das forças. Testar o SM com medições de alta precisão é essencial para descobrir novas possíveis físicas para além do atual quadro teórico.

Colisores de partículas como o Grande Colisor de Partículas (LHC) no CERN estão na vanguarda desta exploração, onde as medições das secções eficazes são cruciais para testar as previsões do SM. A secção eficaz quantifica a probabilidade de interações entre específicas partículas, e medições precisas da luminosidade são fundamentais para determinar estas secções eficazes. A luminosidade, \mathcal{L} , influencia diretamente a taxa de colisões, e, portanto, qualquer incerteza em \mathcal{L} traduz-se numa incerteza nas secções eficazes, afetando a fiabilidade das conclusões extraídas dos dados experimentais.

Esta tese foca-se na calibração preliminar e medição da luminosidade para o período de 2023 que corresponde a colisões protão-protão a uma energia de centro de massa de 13,6 TeV na experiência CMS. A escala absoluta da luminosidade é estabelecida utilizando a metodologia de scans de van der Meer (vdM), uma técnica que baseada no scan dos feixes de partículas entre si para medir a sobreposição e calibrar as secções eficazes visíveis dos luminómetros. Foi empregue uma série de detectores para fornecer medições de luminosidade em tempo real, com correções sistemáticas e análises de estabilidade a garantir resultados consistentes e precisos.

A medição final da luminosidade integrada foi de 32.74 fb^{-1} com uma incerteza total de 1.28%, o melhor resultado preliminar alcançado pelo CMS. Além dos esforços experimentais, foram feitas melhorias no software utilizado nas análises. Estes avanços aumentaram a eficiência do processamento de dados e simplificaram o fluxo de trabalho da calibração, permitindo iterações de análise mais rápidas e fiáveis.

Palavras-chave CERN, CMS, Luminosidade, van der Meer

Contents

1	Introduction	1
1.1	The Large Hadron Collider	2
1.1.1	Accelerator Complex	2
1.1.2	Particle Beams	3
1.1.3	Compact Muon Solenoid Experiment	4
1.2	Luminosity	7
1.2.1	Instantaneous and Integrated Luminosity	7
1.2.2	Luminosity from Machine Parameters	7
1.2.3	Luminosity Calibration	9
1.2.4	Zero-Counting Method	10
1.3	Luminometers	10
1.3.1	Forward Calorimeter	12
1.3.2	The Pixel Luminosity Telescope	12
1.3.3	The Fast Beam Conditions Monitor	14
1.3.4	Pixel Cluster Counting	15
1.3.5	Drift Tube	16
2	Luminosity Calibration at the LHC	17
2.1	Absolute luminosity calibration	17
2.1.1	The van der Meer Scan Methodology	18
2.1.2	Experimental Conditions and Corrective Procedures	19
2.1.3	Extrapolation of vdM Calibration	20
2.1.4	Additional Scans	22
3	2023 Luminosity Calibration	24
3.1	Analysis of vdM data	24

3.1.1	Background estimation	26
3.1.2	Beam current measurement	28
3.1.3	Length scale calibration	30
3.1.4	Beam-beam effects	33
3.1.5	Orbit drift	34
3.1.6	XY factorization	39
3.1.7	Calibration results	41
3.2	Integration	45
3.2.1	vdm Consistency	46
3.2.2	Detector Stability	48
3.2.3	Detector Non-Linearity	49
3.2.4	Establishing a Reference Detector	50
3.3	Final Luminosity Results	51
4	Analysis Software	54
4.1	The vdm Framework	54
4.1.1	Flow of the Analysis	55
4.1.2	Input/Output bottleneck	56
4.1.3	Online vdm Analysis	58
4.1.4	Improving Wait Time in Merger	59
4.2	The BRIL Work Suite	60
4.2.1	Normtags and lovtags	60
4.2.2	Limitations	61
4.2.3	Enhancements to lovtags	62
4.2.4	An Alternative to brilcalc	65
5	Conclusions and future work	66
5.1	Conclusion and Future Work	66
6	Appendices	73
6.1	Listings	73

List of Figures

1	LHC Run 3 schedule	3
2	CERN Accelerator Sequence	4
3	CMS detector and coordinate system	5
4	Transverse slice of CMS detector	6
5	Bunch crossing illustration	8
6	BRIL systems at CMS	11
7	PLT detector telescopes	13
8	Triple coincidence in PLT	14
9	BCM1F sensor arrangement	15
10	Upgraded pixel detector	16
11	Example vdM scan sequence positions	19
12	Non-linearity effects on SBIL	22
13	Beam positions during fill 8999	24
14	Initial HFET visible cross section results	25
15	Initial visible cross section for other online luminometers	26
16	BCM1FUTCA hits in super-separation scan	28
17	FBCT to DCCT calibration	29
18	FBCT calibration for scan pairs	29
19	Longitudinal bunch profile with LDM	30
20	Beam positions in length scale scans	31
21	Length scale factor determination	32
22	Length scale calibration results	32
23	Electromagnetic interaction correction	34
24	Orbit drift correction calculation	35

25	Orbit drift correction summary	36
26	Residual beam position corrections	37
27	Comparison of residual orbit fit methods	38
28	Beam positions during the first offset scan and the second diagonal scan of fill 8999.	39
29	Factorization corrections on visible cross section from 2D rate fit method	40
30	Factorization corrections on visible cross section from luminous region method.	41
31	Final HFET visible cross section results	42
32	Final visible cross section results for other online luminometers	43
33	PCC calibration results	43
34	Double Gaussian fits for BCM1FUTCA	44
35	Scan fits reduced χ^2 distribution	44
36	Overview of the measured effective beam overlap widths	45
37	Overview of the average Σ values in both transverse plans measured during all vdM and im scans, by all the online luminometers.	45
38	Instantaneous luminosity during vdM	46
39	Comparison of head-on luminosity measurements	47
40	vdM luminosity ratio distributions	47
41	Full year luminosity ratios with only vdM calibration	48
42	BCM1F efficiency analysis results	49
43	PLT relative non-linearity slopes	50
44	Final full year stability	51
45	Residual relative non-linearity distributions	52
46	Overview of the flow of the analysis in the vdMFw.	55
47	Workflow of the online analysis.	59
48	Illustration of iovtag and normtag file structure	61

List of Tables

1	Background impact on calibration	27
2	Beam-beam systematic uncertainty	35
3	HFET visible cross sections for the many residual fit methods for both arcBPM and DOROS systems.	38
4	Summary of visible cross sections results	42
5	Average and RMS of all fully-corrected independently-calibrated luminometers with respect to HFET-cross-calibrated DT.	51
6	Summary of uncertainty contributions for the 2023 luminosity calibration analysis.	53
7	Initial vdMFw profiling results	57
8	vdMFw profiling results after performance improvements	58
9	Time spent waiting for scan files to close in the Merger during the 2024 vdM fill.	60

Acronyms

4NB 4 Nibble.

BCID Bunch Crossing Identifier.

BCM1F Fast Beam Conditions Monitor.

BIB Beam-Induced Background.

BRIL Beam Radiation Instrumentation and Luminosity.

CERN European Organization for Nuclear Research.

CMS Compact Muon Solenoid.

ECAL Electromagnetic Calorimeter.

HCAL Hadronic Calorimeter.

HF Forward Calorimeter.

HFET HF Transverse Energy.

HFOC HF Occupancy Count.

IP Interaction Points.

LHC Large Hadron Collider.

LS Lumi Section.

PLT Pixel Luminosity Telescope.

RHU Realtime Histogramming Unit.

SBIL Single Bunch Instantaneous Luminosity.

SM Standard Model.

VME Versa Module Europa.

Chapter 1

Introduction

The quest to understand the universe at its most fundamental level lies at the heart of particle physics. Scientists in this field aim to explore the properties, behaviors, and interactions of elementary particles. Particle physics seeks to answer fundamental questions about the universe, such as the origin of mass, the nature of dark matter, and the unification of forces.

The **Standard Model (SM)** is a theoretical framework that describes the fundamental particles and their interactions, excluding gravity. It encompasses three of the four known fundamental forces: electromagnetism, the weak nuclear force, and the strong nuclear force. It also incorporates the Higgs mechanism, giving elementary particles their mass. The particles in the Standard Model include quarks, leptons, gauge bosons, and the Higgs boson. Quarks and leptons are the building blocks of matter, while gauge bosons mediate the interactions between these particles.

Particle colliders, key instruments in this research process, accelerate and collide particles to reveal new insights through the analysis of the byproducts of such collisions. These events or processes are characterized by their cross sections. Often denoted as σ and measured in units of area, the cross section of an event quantifies the probability of it occurring in a particle interaction. By comparing observed cross sections with those predicted by the **SM**, physicists can test the validity of the model and potentially identify discrepancies indicating new physics beyond the **SM**.

The performance of particle colliders is determined by two key parameters: the center-of-mass energy and the luminosity, L . While the center-of-mass energy defines the energy scale of the collisions, the luminosity quantifies the rate at which particles interact, directly determining the total amount of events that a collider can produce over a given time period. The instantaneous luminosity, \mathcal{L} , is a proportionality constant that relates the rate of an event, R , to its cross section, $R = \mathcal{L} \cdot \sigma$. The importance of precise cross section measurements, from which the **SM** can be probed, sparks the need for precise measurements of the luminosity and fuels the motivation behind this thesis.

1.1 The Large Hadron Collider

The **Large Hadron Collider (LHC)** at the **European Organization for Nuclear Research (CERN)** is a monumental particle accelerator [1]. It's an underground facility spanning the border of France and Switzerland. Its purpose is to accelerate particles beams, one clockwise and the other counterclockwise, and make them collide at high energies to probe the fundamental properties of matter. The collisions happen at four main **Interaction Points (IP)** that correspond to the four main experiments:

- **Compact Muon Solenoid (CMS):** A general-purpose detector designed to investigate a wide range of physics, including the study of the Higgs boson, precision probes of the **SM**, and searches for new phenomena. **CMS** is located at **IP 5**.
- A Toroidal LHC Apparatus (ATLAS): The second general-purpose detector at the **LHC**. Shares similar goals with **CMS** and is located at the opposite side of the **LHC** ring. ATLAS is located at **IP 1**.
- A Large Ion Collider Experiment (ALICE): A dedicated detector designed to study the physics of strongly interacting matter at extreme energy densities, such as quark-gluon plasma. ALICE is located at **IP 2**.
- Large Hadron Collider beauty (LHCb): A detector designed to study the differences between matter and antimatter by studying particle interactions involving the "beauty" quark. LHCb is located at **IP 8**.

The **LHC** stands as the world's largest particle accelerator, capable of accelerating protons to a maximum energy of 6.8 TeV. It started operations in 2009 and has since been a cornerstone of particle physics research with its most notable achievement being the discovery of the Higgs boson in 2012 [2].

The **LHC** operates according to a schedule, which is meticulously planned and updated, when needed, to ensure the success of its experimental programs and upgrades. The current schedule, as shown in Figure 1, outlines the activities for Run 3, which is expected to continue until 2025.

1.1.1 Accelerator Complex

The **LHC** is only the final stage of a complex series of accelerators that prepare the particles for collision. Before protons are injected into the **LHC**, they pass through a sequence of smaller accelerators that progressively increase their energy. This series of accelerators forms the comprehensive CERN accelerator complex, as shown in Figure 2.

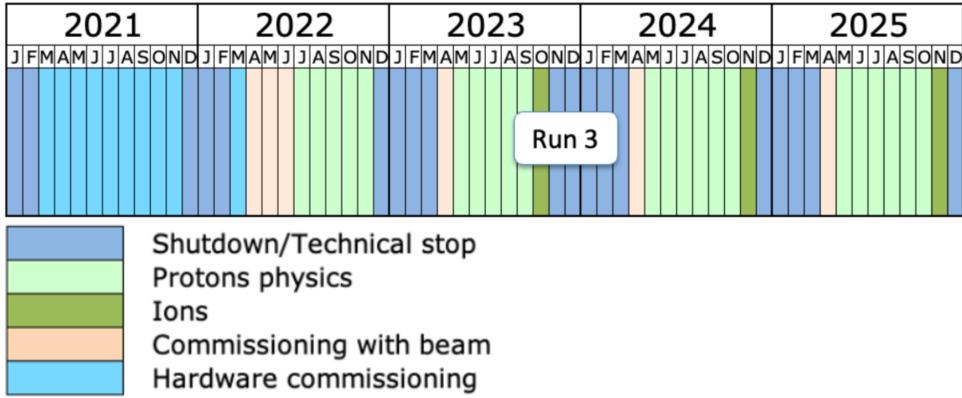


Figure 1: The LHC schedule for Run 3, detailing the operational plan until 2025 (from Ref. [3]).

The **CERN** accelerator complex begins its operations with the Linear Accelerator 4 (Linac4) [5], which, has served as the primary source for proton beams since 2020. Linac4 boosts the energy of negative hydrogen ions (H^-) — that is, hydrogen atoms with an additional electron — to 160 million MeV. Using radio frequency (RF) cavities, Linac4 charges a series of conductors in an alternating positive and negative manner, causing the particles to be accelerated. Before being pulsed through the accelerator, the H^- ions are squeezed into a tight beam using quadrupole magnets. These ions are then readied to enter the Proton Synchrotron Booster (PSB), where they lose their two extra electrons and only the protons remain. The PSB further accelerates these protons to 2 GeV, before they proceed to the Proton Synchrotron (PS). At the PS, the proton energy is increased to 26 GeV. Subsequently, the protons are directed to the Super Proton Synchrotron (SPS), where their energy reaches 450 GeV. Finally, these protons are conveyed to the **LHC**, where each beam is accelerated to an energy of 6.8 TeV.

In addition to protons, the **LHC** also accelerates heavy ions, namely, lead and xenon nuclei. These ions initially enter the Linear Accelerator 3 (Linac3). Once gathered, they are sped up in the Low Energy Ion Ring (LEIR) and eventually follow the same trajectory as the protons to reach their maximum energy.

1.1.2 Particle Beams

In the **LHC**, particles are grouped into bunches, each containing approximately 10^{11} protons at the start of a physics fill. These bunches travel around a 27 km ring at nearly the speed of light, resulting in an orbiting frequency of approximately 11245.5 Hz. The RF cavities that accelerate the bunches operate with an oscillating electric field at a frequency of 400 MHz, creating a total of 35640 RF buckets, each spaced 2.5 ns apart. This arrangement is demanding, as it requires a high degree of timing resolution for any detector system aiming to obtain measurements per bucket. Buckets are grouped into sets of 10, with only the first bucket nominally filled with particles. This grouping increases the spacing between filled particle

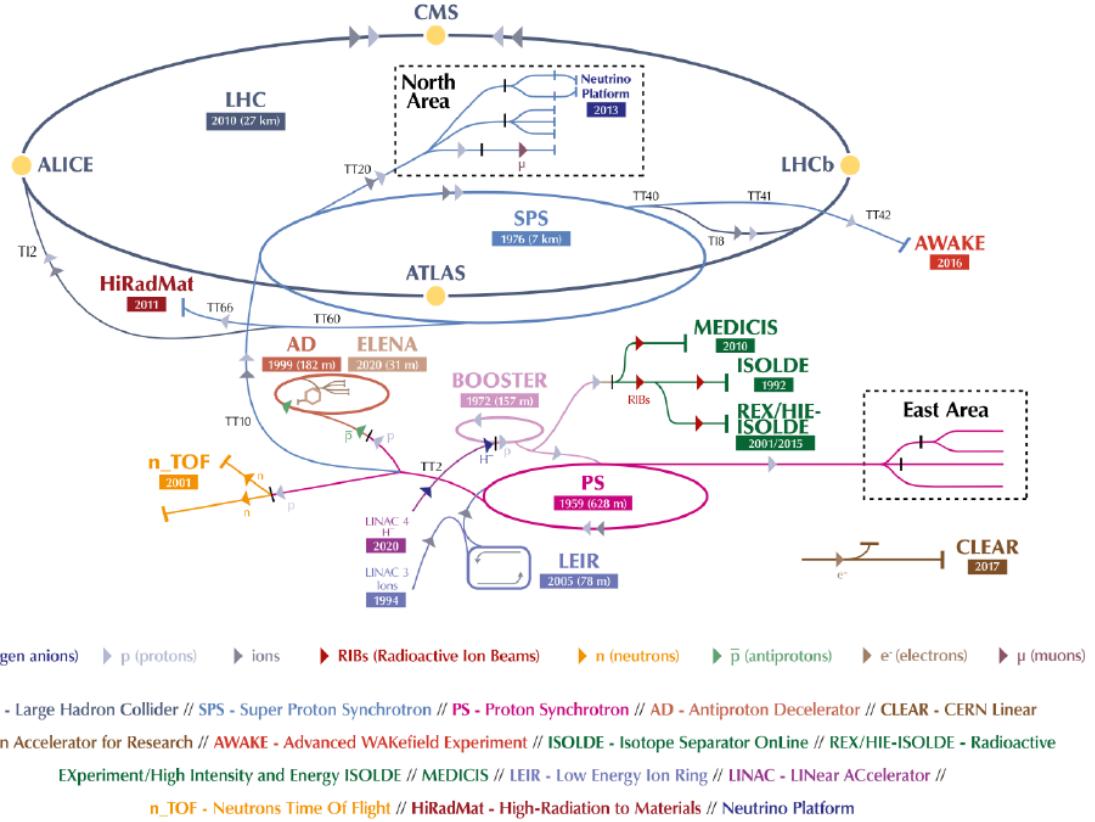


Figure 2: Diagram of the **CERN** Accelerator Sequence (from Ref. [4]): Sequence of accelerators that prepare protons for injection into the **LHC**. The path that protons follow can be traced from Linac4, to BOOSTER, to the PS, to the SPS, and finally to the LHC. Highlighted in yellow are the four main experiments at the **LHC**. The Linac3 and LEIR, whose purpose is to provide ions for the LHC, are also shown.

bunches to 25 ns. These bunch slots are referred to as bunch crossings and assigned a unique number called **Bunch Crossing Identifier (BCID)**, ranging from 1 to 3564.

Each **LHC** fill has a well-defined filling scheme. This scheme provides information on parameters such as the bunch spacing, the number of filled bunches per beam, the number of collisions at each **IP**, and the number of consecutive filled bunches, also known as a train, among others. The filling scheme heavily influences the physics conditions of the fill, dictating the experimental environment and collision dynamics. Each fill is uniquely identified by a fill number which is incremented by a counter. Special fills may be requested to study specific experimental conditions in order to achieve the desired physics goals.

1.1.3 Compact Muon Solenoid Experiment

The **Compact Muon Solenoid** [6], is a multi-purpose detector operating at the **LHC**, designed to operate under high instantaneous luminosity conditions. These conditions require radiation resistance, sub-

detector systems with high granularity, and an intricate readout and trigger system. To fulfill its requirements the **CMS** apparatus is composed of a series of subsystems. Figure 3a shows a schematic of the **CMS** detector and its subsystems. **CMS** is located underground close to the French village of Cessy, between Lake Geneva and the Jura mountains. Its coordinate system in relation to the **LHC** ring is illustrated in Figure 3b.

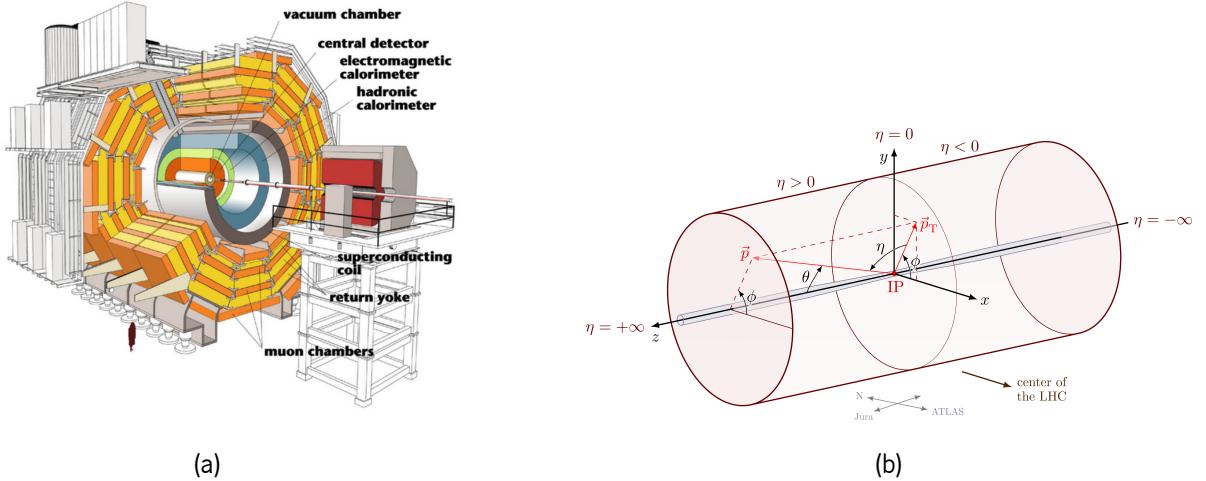


Figure 3: (a) Schematic of the **CMS** detector outlining its subsystems (adapted from Ref. [7]). (b) Diagram of the **CMS** coordinate system as seen with the **LHC** cylinder in the background. The x -axis points radially to the center of the **LHC** ring, the y -axis points upwards, and the z -axis points along the beam direction. The azimuthal angle, ϕ , is measured from the x -axis to the \vec{p}_T vector in the $x - y$ plane. The polar angle, θ , is measured from the z -axis to the \vec{p} vector. Each zone is labeled with the corresponding pseudorapidity, η , defined as $\eta = -\ln \tan(\theta/2)$.

As stated in Section 1.1, the **CMS** physics program encompasses a wide range of topics. From the study of the Higgs boson, to continuously probing the **SM** through precision measurements, to searching for new physics phenomena at high energy scales. The detector is built from the conjunction of multiple layers of subsystems, each with a dedicated focus on measuring specific products of the collision events as they traverse the detector. The subsystems, from the center outwards, are depicted in Figure 4.

The part of the detector closest to the beam pipe is the silicon tracker. This subsystem is composed of silicon sensors that track the paths of charged particles as they ionize the silicon while moving through it. Influenced by the **CMS** magnet, the curvature of these particles' paths provides crucial information for calculating their momentum, with more curved paths indicating lower momentum. The tracker can reconstruct the paths of charged particles, including electrons, hadrons, and high-energy muons. Immediately following the tracker are the **CMS** calorimeters: the **Electromagnetic Calorimeter (ECAL)** and

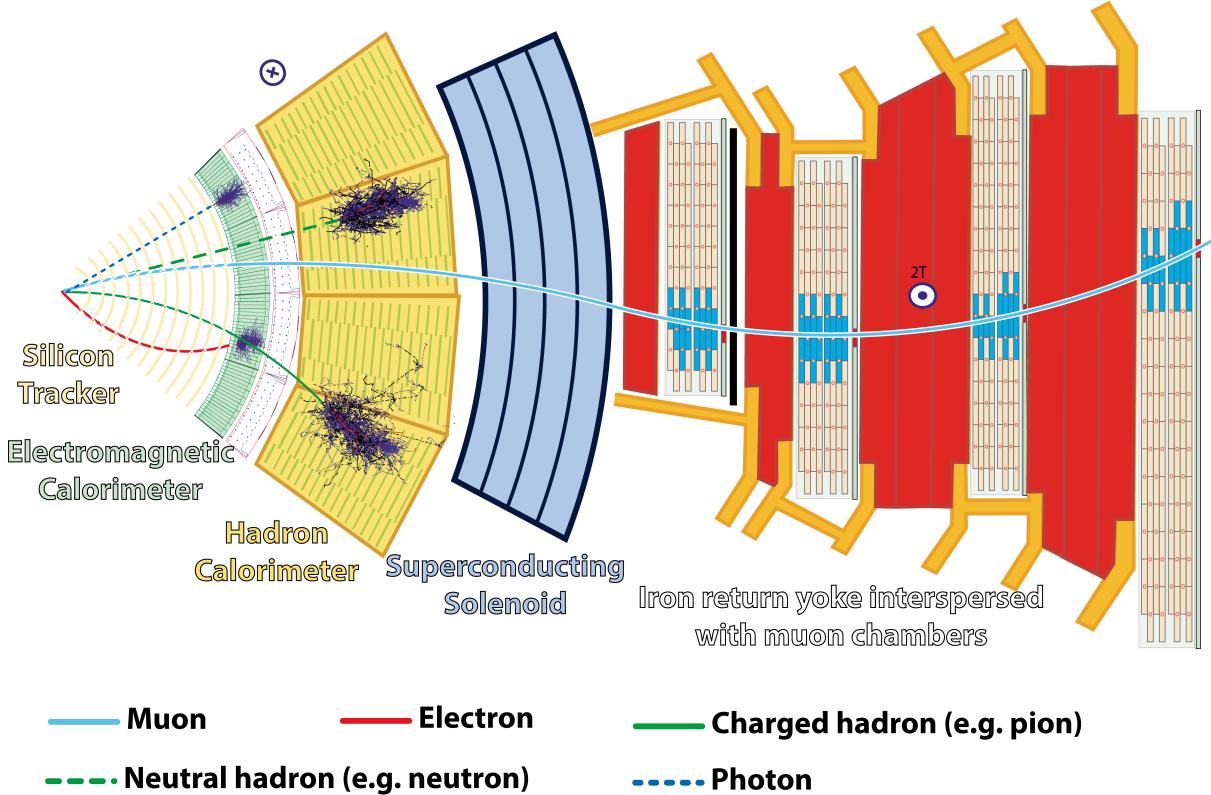


Figure 4: Transverse slice of the **CMS** detector (from Ref. [8]). The subsystems are shown from the left (center) to the right (outwards): the silicon tracker, the electromagnetic calorimeter, the hadronic calorimeter, and the muon system.

the **Hadronic Calorimeter (HCAL)**. These calorimeters measure the energy of particles by stopping them and measuring the energy they deposit. The **ECAL**, composed of lead tungstate crystals, detects electrons and photons by scintillating upon interaction with these particles. Photodetectors, designed to operate within the high magnetic field, are attached to the back of each crystal to detect the scintillation light and convert it to an electrical signal that is then amplified and analyzed. The **HCAL** measures the energy of hadrons, which are particles made of quarks and gluons, such as protons and neutrons. Additionally, it allows for the indirect measurement of non-interacting particles, such as neutrinos, through missing transverse energy, which represents the imbalance in the sum of transverse momentum of the particles in the event. The **HCAL** is constructed with alternating layers of absorber and scintillating materials, producing rapid light pulses when particles pass through. It is organized into several sections: the barrel (HB), outer barrel (HO), endcap (HE), and forward (HF) sections. Muons, heavy cousins of electrons, are the particles most likely to pass through the calorimeters and reach the muon system. The muon system is the outermost layer of the **CMS** detector and is designed to detect muons. It is composed of three types of detectors: Drift Tubes (DT), Cathode Strip Chambers (CSC), and Resistive Plate Chambers (RPC).

Detailed information on **CMS** subsystems can be found in [6].

1.2 Luminosity

Luminosity is essential in particle physics research for two key reasons. First, accurate measurements of instantaneous luminosity are crucial for tracking the performance of both the accelerator and its associated detectors [9]. Second, integrated luminosity often represents the primary source of uncertainty in numerous cross section measurements [10], [11]. This section provides a detailed overview of the concept of luminosity.

1.2.1 Instantaneous and Integrated Luminosity

Luminosity is defined as the ratio between the rate of events, R , and the cross section of a given process, σ :

$$\mathcal{L} = \frac{R}{\sigma} \quad (1.1)$$

\mathcal{L} has then units of $[\text{area}]^{-1}[\text{time}]^{-1}$ where $\text{cm}^{-2}\text{s}^{-1}$ or $\text{fb}^{-1}\text{s}^{-1}$ are common units ($1 \text{ barn} = 10^{-28} \text{ m}^{-2}$). Integrating over a given time interval, $t_2 - t_1$, yields the integrated luminosity, L :

$$L = \int_{t_1}^{t_2} \mathcal{L} dt = \int_{t_1}^{t_2} \frac{R}{\sigma} dt = \frac{N}{\sigma} \quad (1.2)$$

It follows that L is proportional to the number of events, N , produced in the collider.

From Equation 1.2 it can be seen that luminosity can be measured if the cross section of a given process is known. This is the case for lepton colliders, where the “candle” process $e^+e^- \rightarrow e^+e^-$ is used to measure luminosity [12]. However, for hadron colliders, which involve collisions of composite particles made of quarks, such as protons, this approach is not feasible due to the large uncertainties in the calculation of cross sections. Another approach is to measure luminosity using machine parameters, as described in the next section.

1.2.2 Luminosity from Machine Parameters

The **Single Bunch Instantaneous Luminosity (SBIL)**, \mathcal{L}_b , can be measured using the collider’s machine parameters [12]. For colliding bunches containing N_1 and N_2 particles respectively, the expected value of \mathcal{L}_b is given by the following equation:

$$\mathcal{L}_b = \frac{N_1 N_2 f_{rev}}{A_{eff}} \quad (1.3)$$

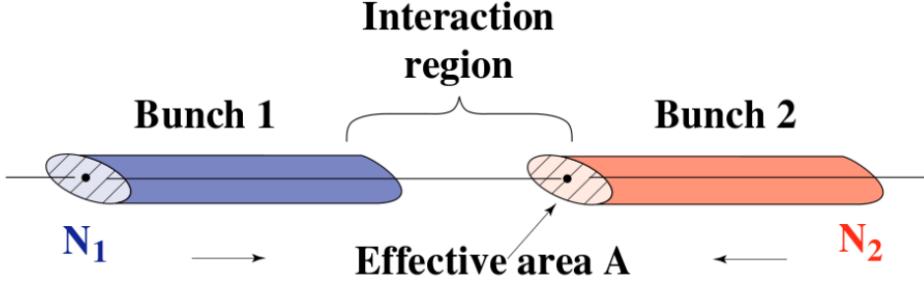


Figure 5: Bunch crossing illustration (from Ref. [12]).

In this equation, f_{rev} represents the orbiting frequency around the **LHC**, as described in subsection 1.1.2, and A_{eff} is the effective transverse area where collisions occur, also referred to as the luminous region. If the transverse distribution of particles was uniform, A_{eff} would be equivalent to the transverse bunch overlap. However, considering the transverse particle distributions ρ_1 for bunch 1 and ρ_2 for bunch 2, the effective transverse area is defined by the overlap integral:

$$\frac{1}{A_{eff}} = 2 \int \rho_1(x, y) \rho_2(x, y) dx dy \quad (1.4)$$

The factor of 2 corresponds to the Møller factor, which accounts for the fact that the two bunches are colliding at equal relativistic speeds and in opposite directions [13]. Typically, it is assumed that these transverse particle distributions are uncoupled, meaning:

$$\begin{cases} \rho_1(x, y) = \rho_{1x}(x) \rho_{1y}(y) \\ \rho_2(x, y) = \rho_{2x}(x) \rho_{2y}(y) \end{cases}. \quad (1.5)$$

Under this assumption, the effective transverse area can be rewritten as:

$$\frac{1}{A_{eff}} = 2 \int \rho_{1x}(x) \rho_{2x}(x) dx \times \int \rho_{1y}(y) \rho_{2y}(y) dy = \frac{1}{W_{eff}} \times \frac{1}{H_{eff}} \quad (1.6)$$

Here, W_{eff} and H_{eff} represent the effective transverse width and height, respectively. Substituting this back into equation 1.3, we get:

$$\mathcal{L}_b = 2N_1 N_2 f_{rev} \int \rho_{1x}(x) \rho_{2x}(x) dx \times \int \rho_{1y}(y) \rho_{2y}(y) dy \quad (1.7)$$

Further, assuming that the transverse particle distributions are Gaussian with widths σ_{x1}, σ_{x2} and heights σ_{y1}, σ_{y2} , the equation becomes:

$$\mathcal{L}_b = 2N_1 N_2 f_{rev} \int G_x(\sigma_{x1}) G_x(\sigma_{x2}) dx \times \int G_y(\sigma_{y1}) G_y(\sigma_{y2}) dy \quad (1.8)$$

$G_i(\sigma)$ in this context is the Gaussian distribution with standard deviation σ in the $i \in x, y$ direction and a mean of zero:

$$G_i(\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{i^2}{2\sigma^2}} \quad (1.9)$$

Expanding the first integral:

$$\int G_x(\sigma_{x1})G_x(\sigma_{x2})dx = \frac{1}{2\pi\sigma_{x1}\sigma_{x2}} \int e^{-\frac{x^2(\sigma_{x1}^2 + \sigma_{x2}^2)}{2\sigma_{x1}^2\sigma_{x2}^2}} dx \quad (1.10)$$

and considering

$$\int e^{-ax^2} dx = \sqrt{\frac{\pi}{2a}} \quad (1.11)$$

equation 1.10 can be simplified to:

$$\int G_x(\sigma_{x1})G_x(\sigma_{x2})dx = \frac{1}{2\pi\sigma_{x1}\sigma_{x2}} \sqrt{\frac{\pi\sigma_{x1}^2\sigma_{x2}^2}{\sigma_{x1}^2 + \sigma_{x2}^2}} = \frac{1}{2\sqrt{\pi}\sqrt{\sigma_{x1}^2 + \sigma_{x2}^2}} \quad (1.12)$$

Using the same approach for the second integral and substituting the results back into equation 1.8, we get:

$$\mathcal{L}_b = \frac{N_1 N_2 f_{rev}}{2\pi\Sigma_X\Sigma_Y} \quad (1.13)$$

where $\Sigma_X = \sqrt{2\pi}W_{eff} = \sqrt{2\pi}\sqrt{\sigma_{x1}^2 + \sigma_{x2}^2}$ and $\Sigma_Y = \sqrt{2\pi}H_{eff} = \sqrt{2\pi}\sqrt{\sigma_{y1}^2 + \sigma_{y2}^2}$. In the case of round beams, the root-mean-squared widths of the horizontal and vertical beam profiles are equivalent to $\sqrt{\epsilon_N\beta^*/\gamma}$. γ is the relativistic Lorentz factor, and β^* corresponds to the value of the optical function β at the IP. ϵ_N is the normalized transverse emittance, with a value of $3.75\mu m$ for the **LHC** [14].

1.2.3 Luminosity Calibration

Luminosity detectors, also called luminometers, are located at the **IPs** of the **LHC** and measure various observable quantities related to the rate of the collisions. Due to the differences between the luminometers, the nature of their recorded observables, and their location at the **CMS** cavern, each is assumed to only capture a fraction of the total collisions. This bias in the luminometers is accounted for by a detector-specific calibration factor, the visible cross section, σ_{vis} . Using equations 1.1 and 1.8 this calibration can be expressed as:

$$\sigma_{vis} = \frac{2\pi\Sigma_X\Sigma_Y R}{N_1 N_2 f_{rev}} \quad (1.14)$$

Once calculated for each luminometer, the visible cross section can be used to convert the measured rates to luminosity:

$$\mathcal{L} = \frac{R}{\sigma_{vis}} \quad (1.15)$$

In both these equations, R represents the measured rate of events from a particular luminometer.

1.2.4 Zero-Counting Method

Due to the constraints in the read-out electronics of some luminometers, there's a chance that simultaneous particle hits might be recorded as a single event. This necessitates a more advanced approach to counting hits accurately. It is postulated that the distribution of the number of hits can be modeled by a Poisson distribution [15] with mean, μ , that is proportional to luminosity.

$$\mathcal{L}_b = \frac{\mu f_{rev}}{\sigma} \quad (1.16)$$

The zero-counting method consists of calculating the probability that a luminometer records zero hits, P_0 , regardless of the number of collisions that occur. This probability is given by the following equation:

$$P_0 = \sum_{i=0}^{\infty} \frac{\mu^i e^{-\mu}}{i!} p^i = e^{-\mu(1-p)} \quad (1.17)$$

μ is the mean number of hits per bunch crossing, also known as pileup, and p is the probability that no observables are recorded in a bunch crossing with i interactions. It then follows that luminosity is proportional to the natural logarithm of P_0 :

$$\mathcal{L}_b = \frac{\mu f_{rev}}{\sigma} = -\frac{\log(P_0)}{1-p} \frac{f_{rev}}{\sigma} = -\log(P_0) \frac{f_{rev}}{\sigma_{vis}} \quad (1.18)$$

The value of p is here absorbed into the visible cross section [16] and does not need to be known beforehand.

The zero counting method is not without its limitations. From Equation 1.17 it can be seen that the probability of recording zero hits decreases exponentially with increasing pileup, or similarly, increasing luminosity. This leads to a phenomenon known as “zero starvation”, where the probability of recording zero hits, P_0 , becomes very small. This can lead to a situation where the luminometer underestimates the number of hits, and consequently, the luminosity.

1.3 Luminometers

The **Beam Radiation Instrumentation and Luminosity (BRIL)** group is responsible for measuring luminosity at **CMS**. In addition to this primary task, the group monitors **Beam-Induced Background (BIB)**, and dangerous beam loss events that may trigger a beam abort. They also manage beam timing, minimum bias triggers for data acquisition, and the monitoring and simulation of the radiation environment within the **CMS** cavern. Luminosity is measured with dedicated detectors called luminometers.

These specialized detectors are capable of providing luminosity measurements with a per-bunch timing granularity. At **CMS**, the number of events measured by most of these systems are integrated over

2^{14} **LHC** orbits, which corresponds to ≈ 1.458 s. This amount of time is referred to as a **4 Nibble (4NB)**. Another interval of integration often used during luminosity analysis is the **Lumi Section (LS)** and corresponds to $16 \times \text{4NB}$ (≈ 23.311 s).

Since these detectors usually operate during long periods of time, four quantities are used to identify a particular operation period. These correspond to: a fill number, synchronized across all **LHC** experiments; a run number, incremented multiple times in a fill as needed by the **CMS** data acquisition system; a **LS** number, unique to **CMS** and resetting at the start of every run; and a **4NB**, also unique to **CMS** and resetting at the start of every lumisection.

Figure 6 shows the **CMS** location of some the systems operated by the **BRIL** group.

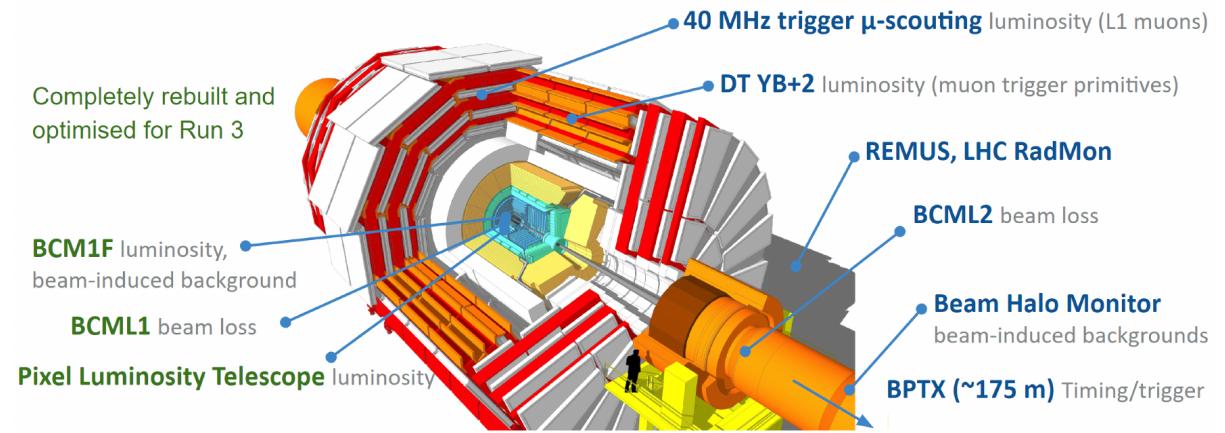


Figure 6: Systems operated by BRIL group at the Compact Muon Solenoid experiment. Green text indicates, that the instrumentation has been rebuilt for Run 3 (from Ref. [17]).

For a detector to be used as a luminometer, it must meet certain essential requirements. One of the most critical use cases for luminometers is their ability to provide online instantaneous luminosity. This capability is crucial for the operation of the **LHC** and the goals of the **CMS** experiment. Additionally, having per-bunch granularity is important, as the calibration process for the luminometers heavily depends on the available statistics, as will be addressed in subsection 2.1.1.

Two other characteristics of paramount importance are the stability of the detector over long periods of time and linear response to varying experimental conditions. The former is necessary to ensure that the luminometers can provide reliable measurements throughout the entire period of operations. A luminometer's stability is typically affected by its accumulated radiation dose, which decreases the detector's efficiency over the course of the data taking period. In order to adjust the detector's response, several adjustments are made to its operating configuration.

The linear response concerns how the detector's response changes under different experimental con-

ditions. In current LHC conditions, the pileup is at its highest near the start of the fill. As the fill progresses, the natural decay of the beam population, commonly referred to as “luminosity burn-off”, causes the pileup to decrease. A luminometer’s response to these changes in pileup conditions must be as linear as possible with respect to luminosity. Additionally, changes in the filling scheme can affect the detector’s response, as some detectors are more sensitive to “afterglow” effects, where the signal from a previous bunch crossing is still present in the immediately adjacent bunch crossings. Similar to maintaining the detector’s stability, the linearity of the luminometer’s response can be preserved through corrective procedures, which will be discussed in subsection 2.1.3 and with focus on 2023 data in chapter 3. The “online” visible cross section and linearity corrections are updated periodically to compensate for changes in the detector’s efficiency and linear behavior.

The following subsections will describe the luminometers that were considered for this work, their positions within the **CMS** detector, and how the luminosity is extracted from their observable measurements.

1.3.1 Forward Calorimeter

As stated in subsection 1.1.3, the **CMS HCAL** is composed of several sections, one of which is the **Forward Calorimeter (HF)** calorimeter [18]. The **HF** is constructed from quartz scintillators embedded in steel absorbers and is located at $z \pm 11.2\text{m}$ from the **IP**, covering a pseudorapidity range of $3 < |\eta| < 5$. This range allows for better detection of missing transverse energy, which is crucial for probing the **SM** and searching for new physics phenomena.

The **HF** calorimeter uses Photomultiplier Tubes (PMTs) to collect Cherenkov light produced by particles passing through the quartz scintillators. A dedicated readout system collects these signals and converts them into luminosity measurements using two methods. The **HFOC** method tracks the fraction of bunch crossings with no energy deposition above a threshold in specific **HF** towers (η rings 31 and 32). The average number of these counts is then converted to luminosity via the zero-counting method.

The **HFET** method, was implemented in 2016. It is based on the assumption that the measured transverse energy sum is proportional to the luminosity. This assumption is used to estimate the number of particles depositing signals and convert it to luminosity. The HFET method is used in parallel with the HFOC method to cross-check the luminosity measurements.

1.3.2 The Pixel Luminosity Telescope

One of the dedicated luminosity detectors is the **Pixel Luminosity Telescope (PLT)** [19]. This detector is based on the use of silicon pixel sensors, 48 in total, arranged in groups of 3 in a telescope-like structure.

The sensors are placed at a distance of 1.75m from both ends of the **IP** at a $|\eta| \approx 4.2$, nearly parallel to the beam line. Figure 7 shows the a diagram for the 24 sensors and the half of the detector with the mounting support.

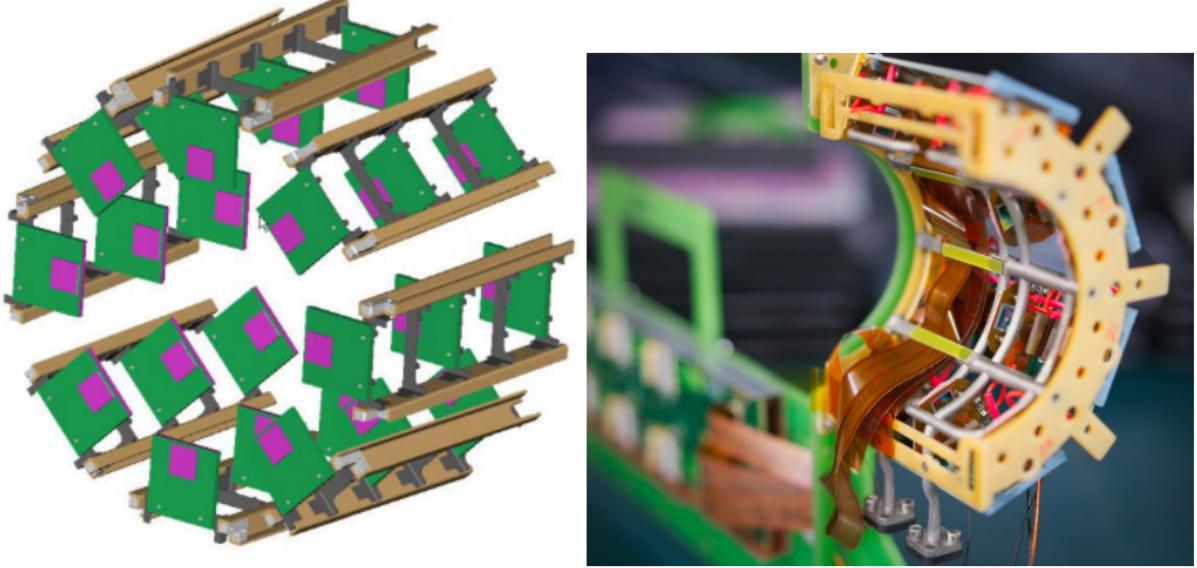


Figure 7: **Left:** Picture illustrating eight telescopes of the **PLT** detector, which correspond to half detector (from Ref. [20]). **Right:** One quarter of the full **PLT** detector with mounting support and with visible connections to the readout electronic (from Ref. [21]).

The **PLT**'s silicon sensors are bump bonded to PSI46v2 readout chips [22], which offer two readout modes: the "fast-or" mode allows for a readout rate of 40MHz, allowing per bunch measurements. A **PLT** "track" is registered when a particle passes through all 3 sensors of a telescope, referred to as a "triple coincidence". The full pixel data readout mode reads at a lower rate of 3.3 kHz being useful for other studies. This method allows for track reconstruction which helps discern between particles from the collision and background noise. Figure 8 shows an illustration of a triple coincidence.

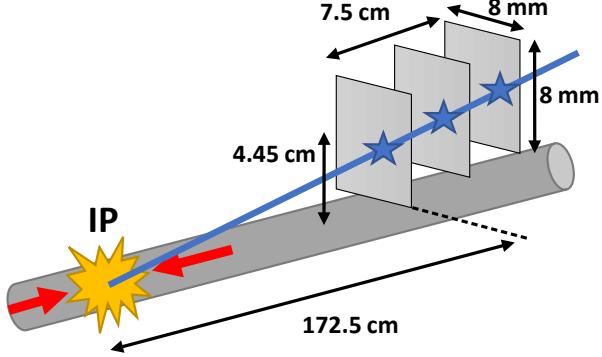


Figure 8: Illustration of a triple coincidence in the **PLT** detector. A hit is considered when a particle passes through the 3 sensors of a telescope (from Ref. [23]).

The zero-counting method, described in subsection 1.2.4, is used to convert the number of hits to luminosity. For PLT, the zero starvation effect is usually not a problem, as the typical PLT occupancy is on the order 0.1-0.2 triple coincidences per telescope per bunch crossing at high luminosity conditions.

1.3.3 The Fast Beam Conditions Monitor

The **Fast Beam Conditions Monitor (BCM1F)** [24] can measure not only the luminosity, but also the **BIB**. It consists of 24 double-pad silicon sensors arranged in four semi-circles, referred to as “C-shapes”, positioned 7.2 cm from the beam pipe and at a distance of 1.8 m on either side of the **IP**. This placement allows for the separation of incoming and outgoing particles, corresponding to a flight time of 6.25 ns for particles traveling at relativistic speeds. This setup enables the discernment of signals originating from collisions, **BIB**, and noise caused by the activation of surrounding material [25]. Figure 9 shows the arrangement of one of the C-shapes.

A particle passing through the bulk of a silicon sensor will ionize the material, creating electron-hole pairs. These pairs are then separated by an electric field, generating a current pulse proportional to the energy of the particle. The signals detected in the sensors are read by two parallel systems. The older back-end system utilizes the **Versa Module Europa (VME)** standard and was the primary system during Run 2. The newer μ TCA-based system began operation in Run 3 [27].

The **VME** back-end system discriminates signals using a constant threshold, and the hits are sent to a **Realtime Histogramming Unit (RHU)**. The hits per bunch are then converted to luminosity using the zero-counting method. The μ TCA system employs a more sophisticated approach, using a peak-finding algorithm that counts pulses based on the signal’s derivative and amplitude. This method not only allows for more accurate distinction between hits and noise but also enables the detection of simultaneous

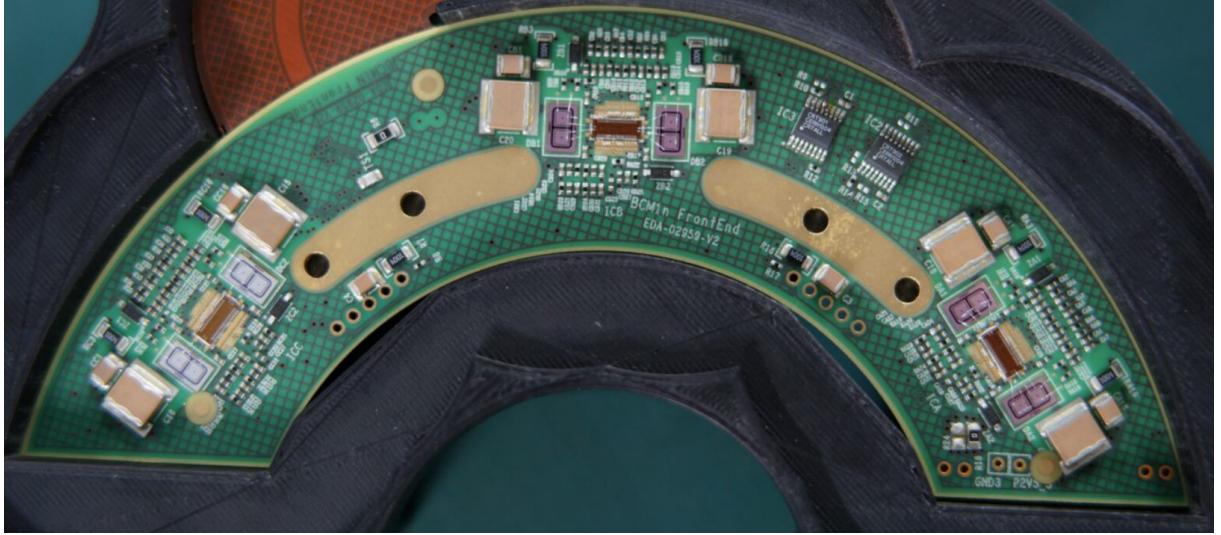


Figure 9: Arrangement of the six sensors in one of the C-shapes of the **BCM1F** detector (from Ref. [26]).

hits by analyzing the pulse shape. Both systems, BCM1F and BCM1FUTCA, are used as independent luminometers.

1.3.4 Pixel Cluster Counting

A large number of modules from the **CMS** Pixel detector is also used as a luminometer. Upgraded in 2017, it consists of four barrel layers, BPIX, and three disks, FPIX, on each side of the **IP** at distances of 291, 396, and 516 mm, respectively. The BPIX is made up of a total of 79 million pixels in 1184 modules, while the FPIX contains 45 million pixels in 672 modules. The layouts of the two detectors, the original and the upgraded one, are compared in Figure 10. Further details on the design and construction of the upgraded Pixel detector can be found in [28].

The Pixel Cluster Counting (PCC) method counts the number of clusters, a conglomerate amount of charged particle hits, in the pixel detector in zero-bias events (events from nominally colliding bunch crossings without any specific activity requirement). The mean number of clusters, averaged over several measurements, can be expressed as

$$\langle N_{\text{clusters}} \rangle = \langle N_{\text{clusters/interaction}} \rangle \cdot \frac{\sigma_{\text{minBias}}}{f_{\text{rev}}} \cdot \mathcal{L}_b \quad (1.19)$$

from which the PCC calibration factor can be calculated as $\sigma_{\text{vis}} = \sigma_{\text{minBias}} \cdot \langle N_{\text{clusters/interaction}} \rangle$ where σ_{minBias} is the proton-proton cross section for minimum bias events. The $\langle N_{\text{clusters}} \rangle$ is measured by averaging the number of clusters over **LS** intervals with a 0.1% statistical uncertainty for orbit integrated luminosity. The PCC method has been shown to provide stable and linear luminosity measurements and

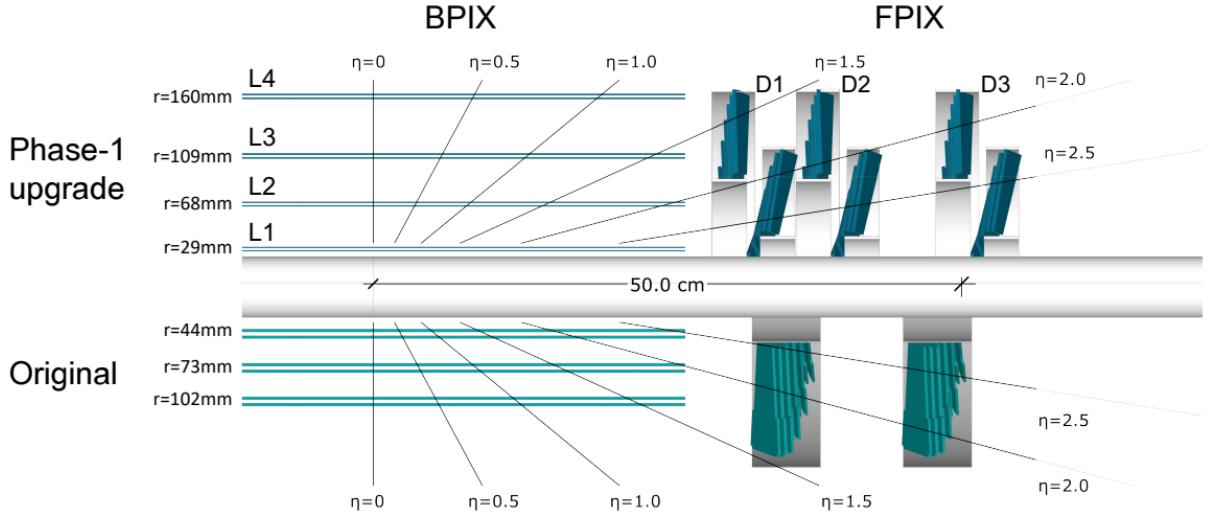


Figure 10: Longitudinal view of the Phase 1-upgraded pixel detector compared to the original detector layout (from Ref. [28]). Labels L1 to L4 indicate the barrel layers, while D1 to D3 indicate the disks.

was used as the primary luminometer for the **CMS** experiment in previous years [16]. The measurements from this detector are only used during the offline analysis.

1.3.5 Drift Tube

Located in the **CMS** muon system, the Drift Tube (DT) is a sub-system of the muon barrel (MB) that provides an estimate of the number of muon tracks passing through the detector. This estimate, done with the Barrel Muon Track Finder (BMTF) algorithm [29], is assumed to be proportional to the luminosity which allows for DT to be used as a luminometer. However, it does not provide luminosity measurements with per bunch granularity, but is rather **LS**-integrated. These limitations make the measurement of its calibration factor, σ_{vis} , dependent on the other luminometers. However, due to its good stability and linearity, it is often calibrated to match other luminometers as a way to compare for stability and linearity.

Chapter 2

Luminosity Calibration at the LHC

CERN's mission of understanding the fundamental structure of the universe is heavily reliant on particle accelerators. In this chapter, the van der Meer (vdM) scan methodology used to measure the luminometer's calibration factor is explained first. Subsequently, the corrective procedures applied to optimize the accuracy and precision of this method are discussed. Finally, the challenges associated with extrapolating the results obtained with the vdM method for the rest of the data-taking period are addressed.

2.1 Absolute luminosity calibration

As explained in subsection 1.2.3, all luminometers require a specific calibration, σ_{vis} , to convert their measured rates into an absolute measurement of luminosity. A common limitation in making precise theoretical predictions of Standard Model processes is the uncertainty in the parton distribution functions within the proton [30]. These limitations necessitate methods that do not rely on theoretical assumptions of these distributions. While data-driven methods have been proposed, they introduce correlations between low and high pileup data-taking periods [31]. A more precise and purely experimental procedure to determine this detector calibration is the van der Meer (vdM) scan methodology.

In order to measure the visible cross section, beam scans are performed, in which the two LHC beams are moved with respect to each other in the transverse ($x - y$) plane in incremental steps. This procedure was pioneered by Simon van der Meer at the Intersecting Storage Rings (ISR) [32] and extended by Carlo Rubbia to the case of a collider with bunched beams [33]. Instead of measuring the transverse bunch density functions, this method allows for the measurement of the bunch overlap integral from the rates measured at different beam separations. This method has been used by every LHC experiment [15], [16], [34], [35].

2.1.1 The van der Meer Scan Methodology

Recalling Equation 1.3 and Equation 1.6, we can rewrite the SBIL expression for beams separated by Δ_x in the horizontal plane and Δ_y in the vertical plane as:

$$\mathcal{L}_b(\Delta_x, \Delta_y) = N_1 N_2 f_{rev} \int \rho_1(x + \frac{\Delta_x}{2}, y + \frac{\Delta_y}{2}) \rho_2(x - \frac{\Delta_x}{2}, y - \frac{\Delta_y}{2}) dx dy \quad (2.1)$$

As shown in [32], when a separation scan is performed in either transverse direction, where Δ_x and/or Δ_y is varied in a step-wise manner, the effective width and height of the luminous region can be expressed as:

$$W_{eff} = \frac{\int \int \rho_1(x) \rho_2(x + \frac{\Delta_x}{2}) dx d\Delta_x}{\int \rho_1(x) \rho_2(x) dx} = \frac{\int \mathcal{L}_b(\Delta_x, 0) d\Delta_x}{\mathcal{L}_b(0, 0)} \quad (2.2)$$

$$H_{eff} = \frac{\int \int \rho_1(y) \rho_2(y + \frac{\Delta_y}{2}) dy d\Delta_y}{\int \rho_1(y) \rho_2(y) dy} = \frac{\int \mathcal{L}_b(0, \Delta_y) d\Delta_y}{\mathcal{L}_b(0, 0)}$$

where the beam populations, N_1 and N_2 , and the LHC revolution frequency have been canceled in the second step for both equations.

Assuming Gaussian-distributed bunches, the scan curves $\mathcal{L}_b(\Delta_x, 0)$ and $\mathcal{L}_b(0, \Delta_y)$ will also be Gaussian. Thus, we arrive at the equality expressed in subsection 1.2.2, where $\Sigma_X = \sqrt{2\pi} W_{eff}$ and $\Sigma_Y = \sqrt{2\pi} H_{eff}$, thereby proving the equivalence of the method. However, as mentioned at the beginning of this section, no assumptions are made about the nature of the particle bunch distribution, which means the scan curves are not guaranteed to be Gaussian.

Frequently, these curves are not well described by simple Gaussians. In this analysis, we fit these curves with Double Gaussian (DG) functions of the form:

$$f_{DG}(\chi) = \frac{r_\chi}{\sqrt{2\pi}} \left[\frac{\epsilon_\chi}{\sigma_{1_\chi}} \exp \left(-\frac{(\Delta_\chi - \mu_\chi)^2}{2\sigma_{1_\chi}^2} \right) + \frac{1 - \epsilon_\chi}{\sigma_{2_\chi}} \exp \left(-\frac{(\Delta_\chi - \mu_\chi)^2}{2\sigma_{2_\chi}^2} \right) \right] \quad (2.3)$$

where $\chi \in \{X, Y\}$ denotes the scanning plane, Δ_χ is the nominal beam separation, r_χ and μ_χ are the peak and peak position of the DG and σ_{1_χ} , σ_{2_χ} are the widths of the two individual Gaussians. The two Gaussians are weighted by ϵ_χ and $1 - \epsilon_\chi$, respectively, and Σ_X and Σ_Y are related to the individual beam widths by:

$$\Sigma_\chi = \frac{\sigma_{1_\chi} \sigma_{2_\chi}}{\epsilon_\chi \sigma_{2_\chi} + (1 - \epsilon_\chi) \sigma_{1_\chi}} \quad (2.4)$$

The visible cross section, σ_{vis} , can also be expressed as a function of the fit parameters as:

$$\sigma_{\text{vis}} = \frac{2\pi\Sigma_X\Sigma_Y R_{\text{peak}}}{N_1 N_2 f_{\text{rev}}} \quad (2.5)$$

where R_{peak} is taken as the arithmetic mean of the peak values, r_X and r_Y , of both scan curves. This fit procedure is applied for every bunch crossing. Since the visible cross section results heavily depend on how well the fit converges we only apply this method on the luminometers which have per-bunch granularity.

Figure 11 illustrates the beam separations during two vdM scans, one in each transverse direction.

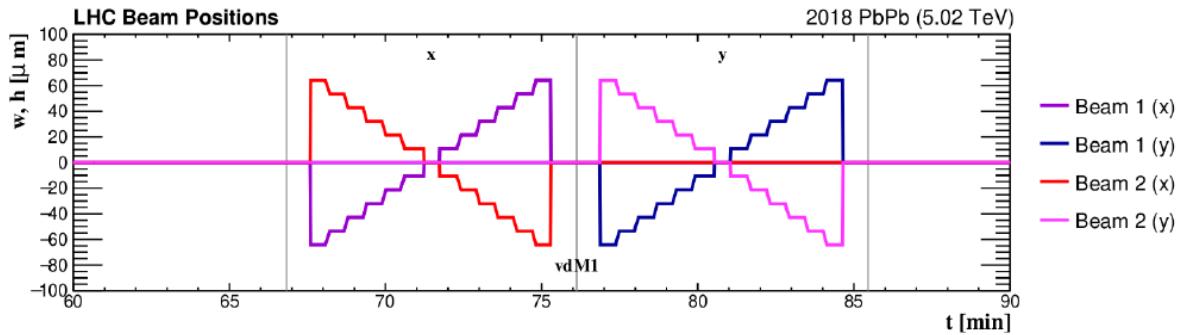


Figure 11: Nominal LHC beam positions displaying the van der Meer scan sequence first in x -axis and then in y -axis (from Ref. [17]).

2.1.2 Experimental Conditions and Corrective Procedures

A vdM scan is conducted under experimental conditions that allow for the measurement of the visible cross section to the highest accuracy. These conditions include smaller beam intensities, which minimize the effect of BIB and reduce non-linearity effects, and well-separated bunches, which minimize afterglow effects [30]. In addition to these optimized experimental conditions, a series of offline corrections are performed to ensure the required level of accuracy. These corrections include:

- **Background estimation:** Background signals, whether beam-induced, from detector noise, or machine-induced, are typically present in the raw recorded rates. These additional contributions are removed.
- **Beam current measurement:** Various systems measure and monitor the beam conditions throughout the entire fill to enable offline corrections, such as the removal of spurious charges.
- **Beam-beam effects:** The electromagnetic interaction between the two LHC beams causes disturbances in the nominal beam separation and shapes, affecting the calibration measurement [36]. These effects are accounted for and corrected in the analysis.

- **Orbit drift:** During the vdM fill, the nominal position of the LHC beams may shift. These shifts are measured by Beam Position Monitors (BPM), which provide information to correct the nominal beam separations.
- **Length scale calibration:** The operational displacement of the LHC beams is achieved using a pair of steering dipoles located on either side of the IP. These displacements come with associated uncertainties due to effects such as magnet hysteresis or lattice imperfections [37]. These effects are taken into account.
- **XY factorization:** The vdM method assumes that the transverse particle distributions are independent in each transverse direction and therefore factorizable. If this assumption is not true, the calculated value for A_{eff} will be biased. Studies are conducted to understand the magnitude of this bias and correct it.

A more detailed explanation of these corrections, as well as their impact on the measured calibration, will be provided in chapter 3.

2.1.3 Extrapolation of vdM Calibration

The experimental conditions in which a vdM fill is performed, as well as all the extra corrections applied to the collected data, ensure a visible cross section measurement with high precision and accuracy. However, the optimal experimental conditions for probing the Standard Model, known as “physics conditions”, introduce several effects that impact the calibration of the luminometers. Therefore, it is necessary to extrapolate the results obtained during a vdM fill to correctly calibrate the luminosity measured under physics conditions.

Physics conditions are tailored to obtaining the largest amount of data possible. This includes higher bunch populations compared to vdM fills and more colliding bunches, all of which increase the number of colliding particles. These conditions bring about two classes of effects on our luminometers:

- **Efficiency effects:** Prolonged exposure to ionizing radiation from highly-energetic interactions deteriorates the luminometers, making them less efficient. These effects result in a gradual decrease in the reported luminosity throughout the data-taking period.
- **Non-linearity effects:** Varying experimental conditions can affect the reported luminosity of our detectors, as mentioned in section 1.3. While some of these effects have their own corrective procedures, such as out-of-time pileup [16], others are corrected empirically.

To account for these two classes of effects, the measured luminosity undergoes a final correction described by Equation 2.6.

$$\mathcal{L} = \frac{f_{rev} \cdot \mu}{\sigma_{vis} \cdot \epsilon} - \alpha \left(\frac{f_{rev} \cdot \mu}{\sigma_{vis} \cdot \epsilon} \right)^2 \quad (2.6)$$

Two new parameters have been introduced. ϵ is the parameter that corrects for losses in detector efficiency throughout the data-taking period. To obtain these factors, mini vdM-like scans, called emittance scans (em), are performed. These scans are used to measure the visible cross section using the same vdM method, allowing for the analysis of the evolution of this quantity over the data-taking period. A figure of merit (FOM) is then calculated as:

$$FOM = \frac{\sigma_{vis}^{\text{emitt}}}{\sigma_{vis}^{\text{vdM}}} \quad (2.7)$$

where $\sigma_{vis}^{\text{vdM}}$ is the visible cross section measured under vdM conditions, and $\sigma_{vis}^{\text{emitt}}$ is the visible cross section measured during emittance scans under physics conditions. These FOM values are calculated for every emittance scan throughout the data-taking period and used to correct for variations in each luminometer's efficiency.

The second new parameter, α , corrects for any non-linearity in the luminometers and is derived empirically. Non-linearity is defined as any non-linear response to changes in experimental conditions. Figure 12 illustrates the response of a theoretical linear detector, a detector with positive non-linearity, and a detector with negative non-linearity.

The reasons for a detector to have a positive or negative non-linear response vary:

- PLT has consistently shown signs of positive non-linearity, mainly due to the increased probability of “accidental” triple coincidences at higher pileup due to background tracks, leading to over counting.
- HFOC suffers from the “zero starvation” effect, as explained in subsection 1.2.4, leading to an underestimation of the measured rates, categorized as negative non-linearity.

Correcting for non-linearity is challenging due to the lack of a perfectly linear reference detector. Instead, a reference detector, assumed to be linear, is used. In this method, a linear fit is performed on the ratio between a luminometer and this reference detector as a function of SBIL.

The slope extracted from the fit is then used as the value of α in Equation 2.6. Unlike the efficiency corrections (ϵ), the non-linearity corrections are done at the expense of introducing some degree of correlation between the corrected detector and the reference. To keep the luminometers as independent as possible, the slopes are extracted for only a few fills with a high SBIL range throughout the year.

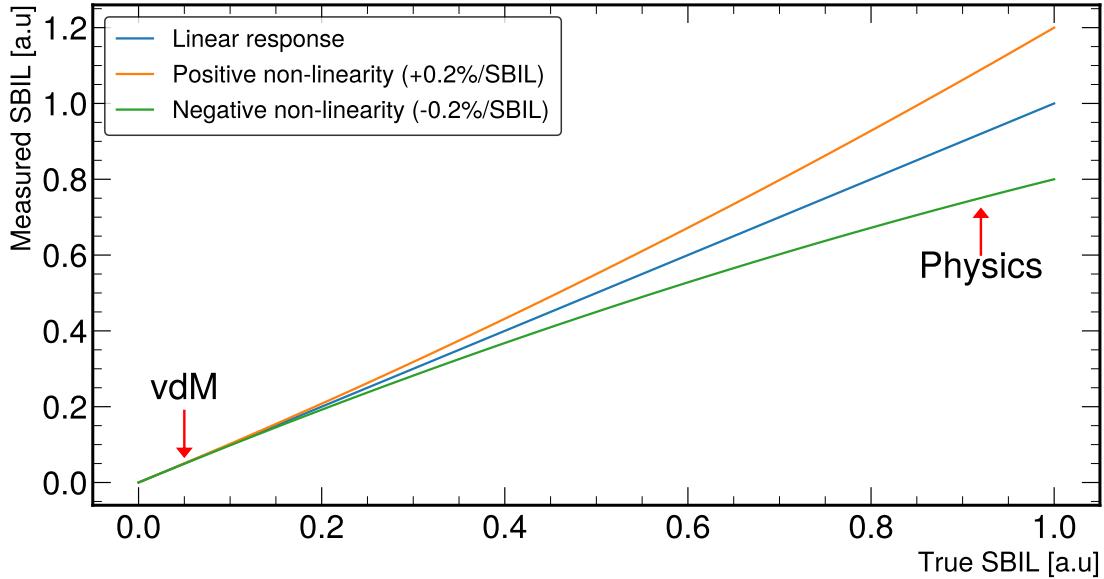


Figure 12: Illustration of non-linearity effects on measured SBIL as a function of the true SBIL. The two marked regions illustrate the effects of extrapolating from vdM to physics conditions.

The specialized vdM conditions, along with the series of corrections applied to the vdM data and the extrapolation to physics conditions, all contribute to the significant challenge of accurately measuring luminosity at hadron colliders like the LHC.

2.1.4 Additional Scans

In addition to the vdM scan, other types of scans are performed during the vdM fill to provide specific insights into the various experimental conditions:

- **Super Separation scans (ss):** These scans involve maximally separating the beams in the LHC, allowing for an accurate measurement of the background for each luminometer.
- **Beam Imaging scans (im):** In these scans, one beam is kept fixed in its nominal head-on position while the other beam is scanned. These scans are used for the beam-imaging method [38], which is employed in XY factorization analysis. A visible cross section is also extracted from these scans, similar to vdM scans.
- **Diagonal scans (diag):** These are vdM scans performed at distinct angles, such as $+45^\circ/-45^\circ$, $+30^\circ/-60^\circ$, and $+60^\circ/-30^\circ$. They sample over different regions of the beam overlap shape and are used in XY factorization studies.

- **Offset scans (off):** These scans have the same motion as vdM scans, but one of the beams is not in the nominal head-on position. They are used in beam shape studies to determine the XY factorization bias.
- **Variable/Constant length scale scans (vls, cls):** These scans involve moving the beams together at constant or variable separation and are used to calibrate the distance by which the steering magnets displace the beams.

Each scan is separately recorded and analyzed offline. After thorough analysis, the corrective procedures determined from each of these special scans are applied to the vdM scan analysis to correct the extracted detector calibration. These corrections affect every luminometer in a similar fashion.

Chapter 3

2023 Luminosity Calibration

3.1 Analysis of vdM data

CMS's vdM scan program was performed in June 2023 during LHC fill 8999. During this fill, there were 138 filled bunches per beam and a total of 136 colliding bunches at the CMS IP. The bunches were designed to be wider than usual ($\beta^* = 1920\text{cm}$) and separated by 525ns in order to minimize afterglow effects. In order to minimize non-linear effects, the pileup was significantly limited from the average during 2023, 52, to just 0.5.

The LHC beam positions throughout the entire vdM program is shown in Figure 13.

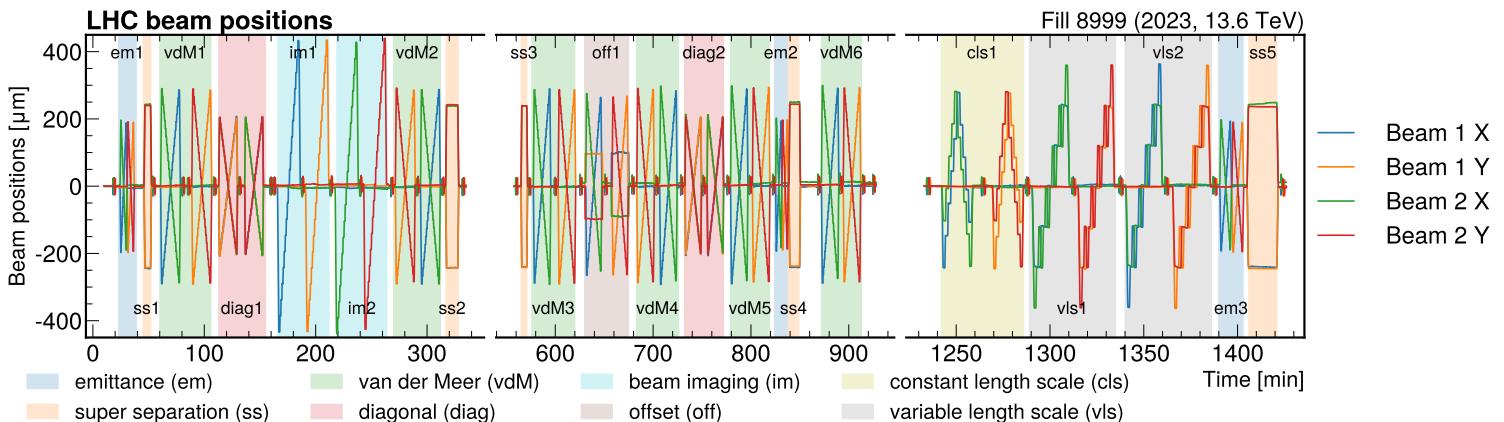


Figure 13: The vertical (X) and horizontal (Y) positions of beam 1 and beam 2 are shown as a function of time during LHC fill 8999 as measured by the LHC beam position monitors. Each kind of scan pair, which contains two orthogonal scans, is delimited by the shaded regions and labeled with the appropriate scan name. The cut out time periods correspond to when CMS was taking head-on data while the ATLAS vdM scan program was ongoing.

As can be seen in Figure 13, during the vdM scans, the beams start by being maximally separated by approximately $578\mu\text{m}$. They are then scanned across each other in 25 steps of approximately $48\mu\text{m}$,

remaining in the same position for 30 seconds for each step. In the beam imaging scans, one of the beams is fixed in the nominal head-on position while the other is moved from approximately $-433\mu\text{m}$ to $+433\mu\text{m}$ in 19 steps of 46 seconds each.

The measured luminometer rates are fitted as a function of the beam separation with the model described in Equation 2.3 and the luminometer calibrations are determined from the fit parameters as detailed in Equation 2.5. The fits were performed on the data from the 6 vdM and the 2 im scan pairs and its associated visible cross sections were measured. Figure 14 shows the visible cross section measurement results for the HFET luminometer with none of the corrections mentioned in subsection 2.1.2 having been applied.

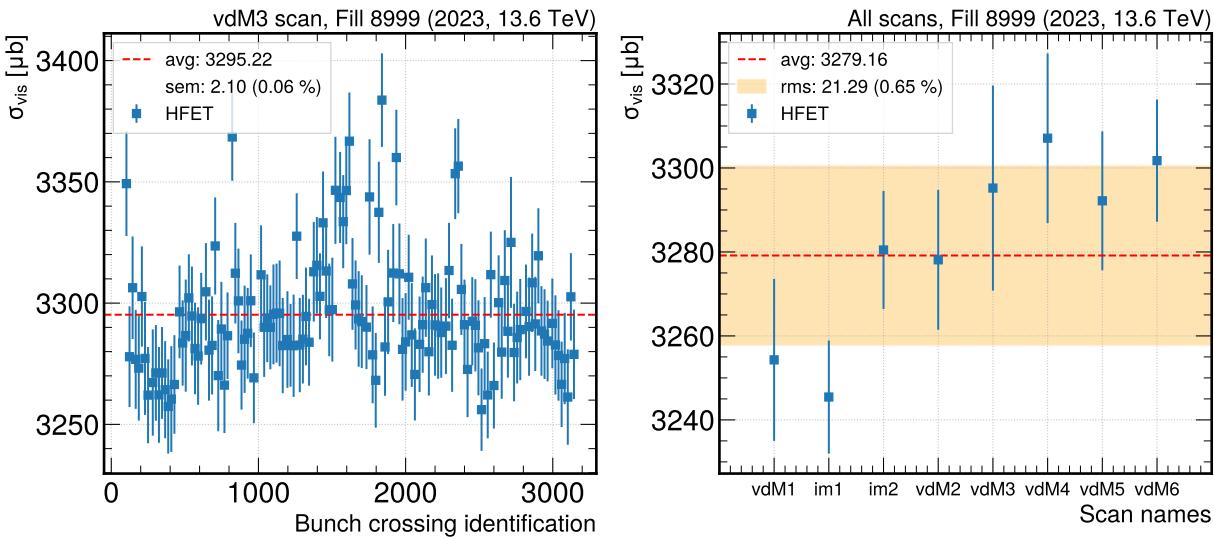


Figure 14: Initial visible cross section results obtained from uncorrected, raw data for the HFET luminometer, as an example. On the left, the measured per-bunch visible cross sections with their associated fit uncertainties are shown. The standard error on the mean (sem) is also depicted in the legend. The right image shows the measured visible cross section taking the average over all 136 BCIDs for each vdM and im scan pairs. The error bars are defined by the root mean squared (RMS) over all BCIDs. The central value and its uncertainty are determined by calculating the uncertainty-weighted average and standard deviation over the scan pairs.

There is a noticeable increasing trend in the measured visible cross sections as a function of time. Furthermore, this trend is visible in all other luminometers, as can be seen in Figure 15.

To increase the precision and accuracy of these measurements, a series of corrections have been applied to the raw data as will be described in the next subsections.

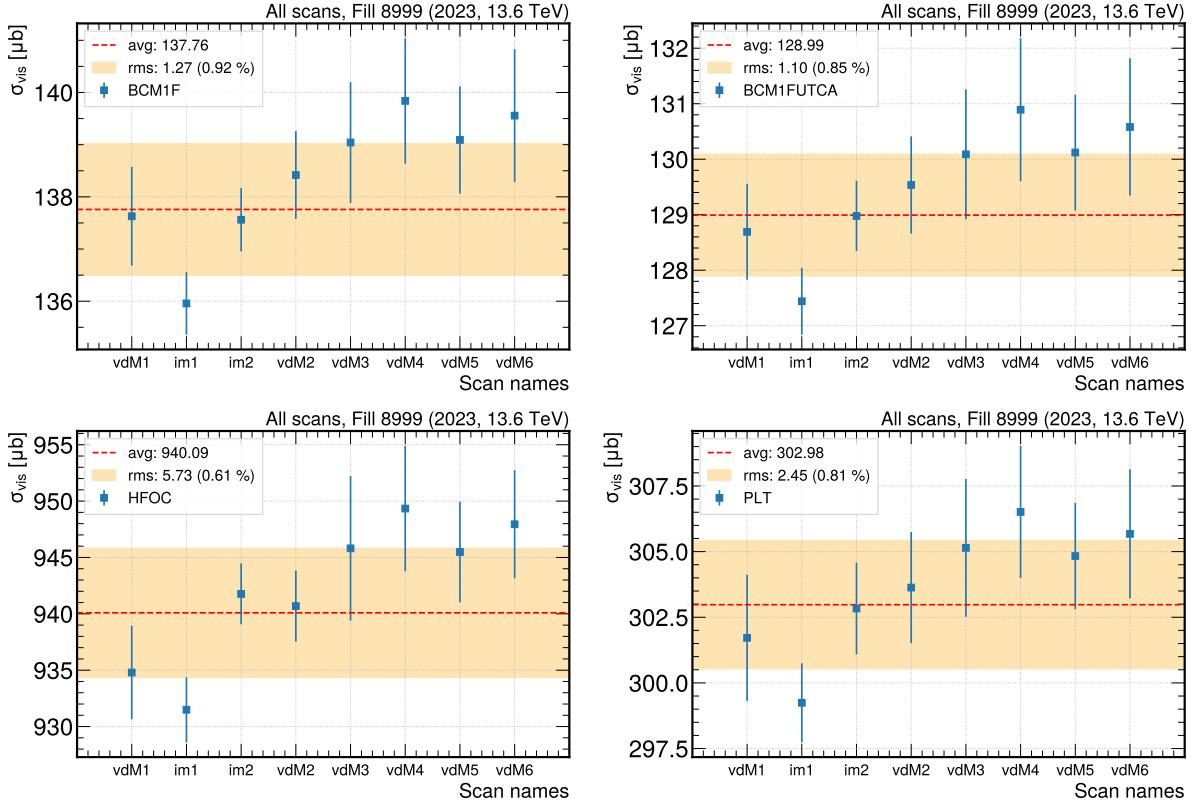


Figure 15: Initial visible cross section results obtained from uncorrected, raw data for BCM1F, BCM1FUTCA, HFOC and PLT luminometers.

3.1.1 Background estimation

Beam-induced background and detector noise are both sources of background that are taken into account and subtracted from the raw detector rates. These sources have been estimated via two distinct methods:

- **Estimation from super-separation scans:** The background contribution can be directly measured from super-separation scans (denoted as ss in Figure 13). During these scans, the beams are separated in both planes in order to render the contributions from collision products negligible. This method is most commonly used during special fills, such as the vdm fill, since it consumes beam time.
- **Estimation from non-colliding bunches:** An estimation for background can be done as long as there are filled but non-colliding bunches in the filling scheme. This indirect method assumes

the background in colliding bunches can be described as:

$$\text{BKG} = 2 \cdot \text{BIB} + \eta \quad (3.1)$$

where the beam-induced background, BIB, from the two beams is considered as well as the detector-specific noise, η . Furthermore, BIB is calculated like:

$$\text{BIB} = \text{FNCR} - \eta. \quad (3.2)$$

Here, FNCR, is the measured rate from filled non-colliding bunches. Finally, the detector noise is calculated as the average signal from empty non-colliding bunches, ENCR. Combining Equation 3.1 and Equation 3.2, the background in colliding bunches is calculated as:

$$\text{BKG} = 2 \cdot \text{FNCR} - \text{ENCR} \quad (3.3)$$

This method heavily relies on the number of filled non-colliding bunches, of which there were only two in the vdM fill, in order to accurately measure BIB. Since super separation scans are never done under physics conditions, this method is the only option available for emittance scans during physics fills.

Figure 16 shows the BCM1FUTCA average of the measured rates for every BCID using the super-separation scan method. Table 1 summarizes the measured background contributions for all the luminometers except for HFET since its background subtraction is applied online.

Table 1: Summary of measured background per luminometer and its impact on the vdM calibration. The average is computed on the measured background for all the colliding bunches across all scans. The associated uncertainty is the standard error on the average.

Luminometer	Average [μ]	Uncertainty [μ]	% to peak rate at vdM1	Effect on σ_{vis} (%)
BCM1F	1.59E-6	4.68E-9	0.18	-1.56
BCM1FUTCA	1.34E-6	4.63E-9	0.16	-1.39
HFOC	1.17E-7	6.10E-9	0.002	-0.02
PLT	2.11E-6	9.25E-9	0.11	-1.05
PCC	1.65E-2	6E-4	0.12	-1.08

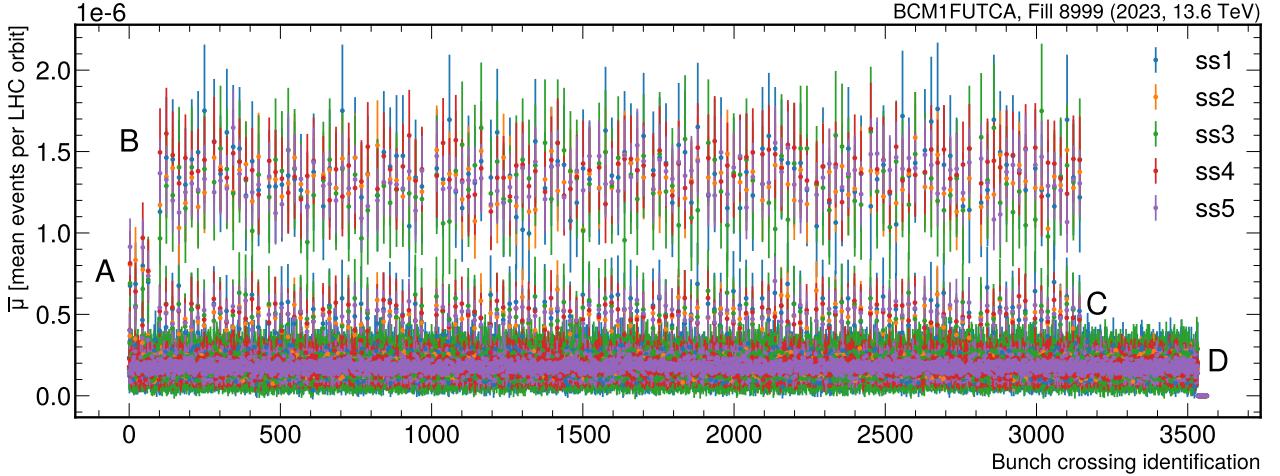


Figure 16: BCM1FUTCA average of the measured rates per BCID during super separation scans. The labels A, B, C, and D refer to the following: A represents the four filled non-colliding BCIDs at the beginning of the orbit, B represents the 136 colliding BCIDs (indicated by the top band), C corresponds to the empty BCIDs adjacent to filled BCIDs (around 0.5), and D represents the remaining empty non-colliding bunches.

3.1.2 Beam current measurement

The beam currents play a crucial role in the determination of the visible cross section, serving as a normalization factor to the detector rates, as can be seen from Equation 2.5. In the LHC, they are measured with two dedicated instruments: the Fast Beam Current Transformer (FBCT) [39] and the Direct-Current Current Transformer (DCCT) [40].

The FBCT system is able to provide current measurements with per-bunch granularity (every 25ns) while DCCT provides orbit integrated measurements with an associated uncertainty of 0.2% [41], achieving higher levels of accuracy compared to FBCT [42]. This motivates a calibration procedure in which the orbit-integrated FBCT measurements are normalized to the measurements from DCCT. Figure 17 shows the value of this calibration across the entire vdm fill for both beams.

The application of this correction is done only in the measurements that correspond to time intervals where a scan is being performed. The overall magnitude of this calibration is calculated as the average among all the scan steps. Figure 18 shows the calibration used for the first vdm scan pair, vdm1.

Both systems are also susceptible to the presence of erratic charges. As mentioned in subsection 1.1.2 only one out of the 10 RF buckets in a bunch crossing are nominally filled. Spurious charges in the nominally empty buckets, referred to as “ghost charges”, have an impact on the DCCT measurements but do not have an effect on the per-bunch FBCT measurement, since it is impervious to charges below a given threshold. Another class of additional charges, called “satellites”, emerge in the RF buckets adjacent

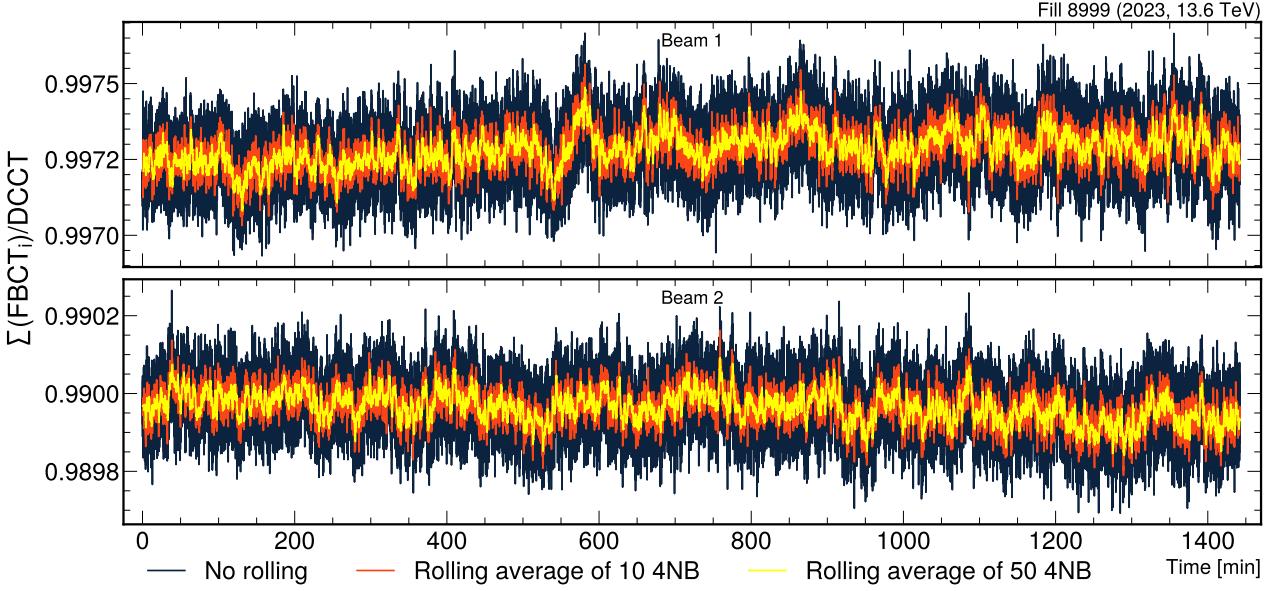


Figure 17: Value of the FBCT to DCCT calibration throughout the vdM fill 8999. The calibration factor is computed as the ratio between the orbit-integrated FBCT measurements and that of DCCT.

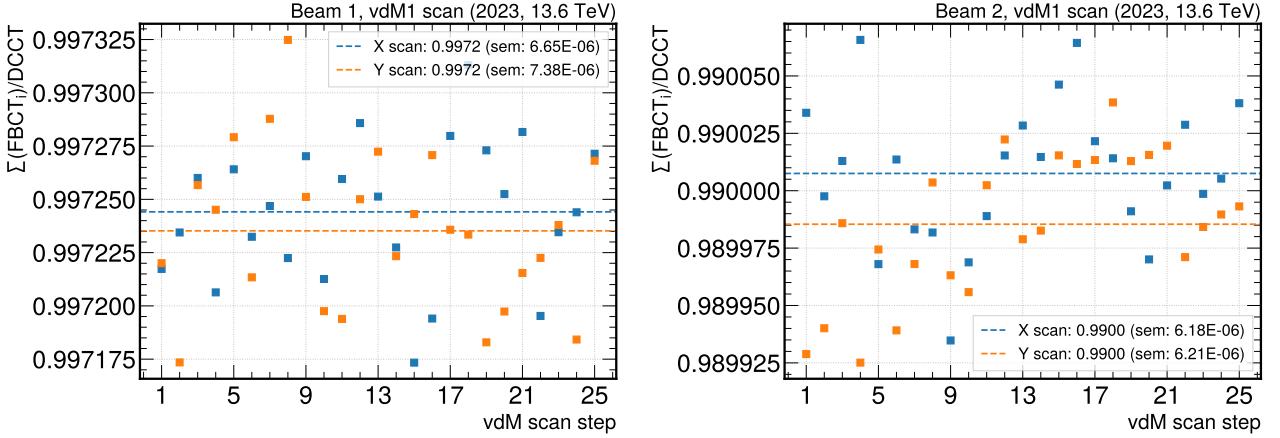


Figure 18: Calculated FBCT to DCCT calibration for the scan pairs of the first vdM scan in fill 8999. Every squared point corresponds to the ratio between the orbit-integrated FBCT beam current and the measurement from DCCT for a given scan step (25 in total for a vdM scan). The average and error on the mean are shown in the legend.

to the nominally filled one. These extra charges affect the FBCT measurements. Figure 19 illustrates the position of these charges in relation to the main bunch.

The LHC Longitudinal Density Monitor (LDM) system, measures the fraction of ghost charges in the DCCT measurement, f_{ghost} , as well as the per-bunch satellite fractions in the FBCT measurements, f_{sat_i} . Combining the calibration procedure specified above with the corrections from these LDM measurements,

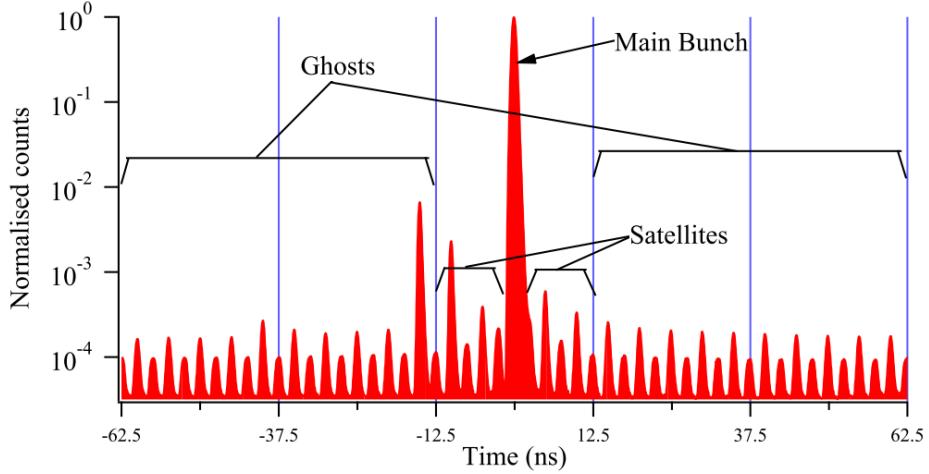


Figure 19: Longitudinal profile taken with the LDM, in logarithmic scale. The ghost charges are evenly spread while the satellites are adjacent to the main bunch (from Ref. [43]).

the total correction to the beam current can be expressed as:

$$n_i = \frac{\text{FBCT}_i (1 - f_{sat_i}) \cdot \text{DCCT} (1 - f_{ghost})}{\sum_i \text{FBCT}_i} \quad (3.4)$$

where n_i is the corrected current for BCID i .

These corrections have a similar effect on all luminometers. The FBCT to DCCT calibration decreased the visible cross section by 1.27% with an associated uncertainty equal to that of the DCCT measurement, 0.2%. The correction from ghost-satellite charges had a +0.1% effect on the detector calibration and the full correction was taken as an uncertainty.

3.1.3 Length scale calibration

During a vdM scan pair, the beams perform 25 steps in both planes. The movement of these beams is controlled with a pair of steering dipoles located on either side of the IP. However, the scale of each step may differ from the nominally intended step length. For example, an intended nominal step of, say, $1 \mu\text{m}$ may correspond to $0.99 \mu\text{m}$. This can be generalized for any nominal position, x_{nom} , and the associated actual measurement, x_{meas} , as:

$$x_{\text{meas}} - x_{\text{nom}} = (\alpha - 1) x_{\text{nom}} \quad (3.5)$$

where α is called the length scale factor. This parameter encodes the first-order linear response of the magnets to the nominal input positions.

This correction revolves around the comparison between the nominal and actual beam spot positions. The beam spot is the overall distribution of the vertices reconstructed from tracks registered in the tracking

detector and its position refers to the average of this position. The CMS tracker [16], provides the actual beam spot measurements, $\overline{x}_{\text{BS}}^{\text{meas}}$, and the nominal measurement assumes that the individual beam profiles are Gaussian and factorizable. Therefore, the beam spot is computed as the weighted average from the two beams as [44]:

$$\overline{x}_{\text{BS}}^{\text{nom}} = \frac{\overline{x}_1 \sigma_{x1}^{-2} + \overline{x}_2 \sigma_{x2}^{-2}}{\sigma_{x1}^{-2} + \sigma_{x2}^{-2}}, \quad (3.6)$$

and similarly for the y direction. The $\overline{x}_{\text{BS}}^{\text{nom}}$ is the nominal beam spot position and \overline{x}_1 and \overline{x}_2 are the means of the beam profiles.

In practice, the length scale scans, as shown in Figure 13, are used. There are two versions of these scans: constant length scale (cls) scans and variable length scale (vls) scans. The cls scans are made up of 2 stages: forward stage, when the beams are moved from negative to positive positions, and the backward stage, where the beams are moved from positive to negative positions. In the cls scan, the relative beam separations are kept constant. In the vls scans, we have a cycle repeating a total of five times. This cycle comprises the moments in which beam 1 is above beam 2 (Up), the beams are head-on and, lastly, when beam 1 is below beam 2 (Dn). Figure 20 shows the beam positions during the cls1 and vls1 scan pairs.

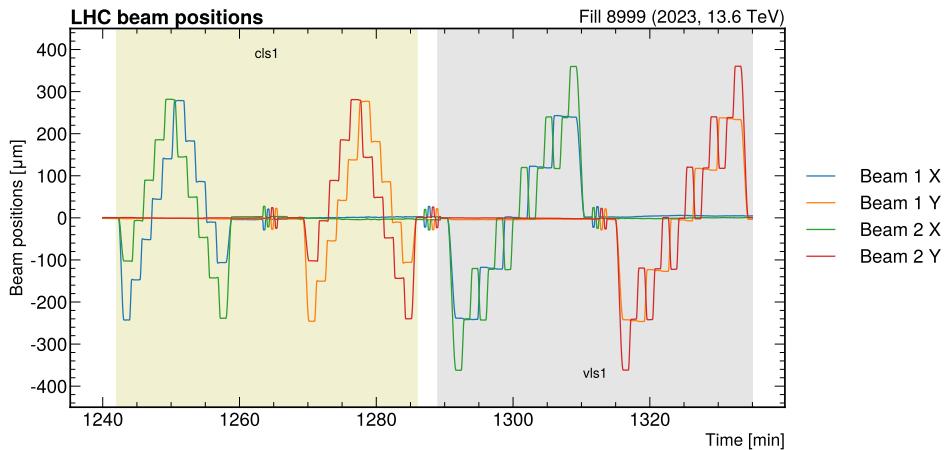


Figure 20: Beam positions during the constant and variable length scale scans.

The difference between the measured and nominal beam spot measurements are plotted as a function of the beam positions and a linear fit is performed to obtain the α factor. An example fit for the Up stage of the variable length scale scan is shown in Figure 21. The overall length scale factor was calculated from a weighted average on the factors obtained from every possible scan and scan stage. The results are summarised in Figure 22.

As the backward stage of the constant length scale for the vertical plane yielded poor quality fit results, it was not included in the weighted average. However, in order to account for this in the associated

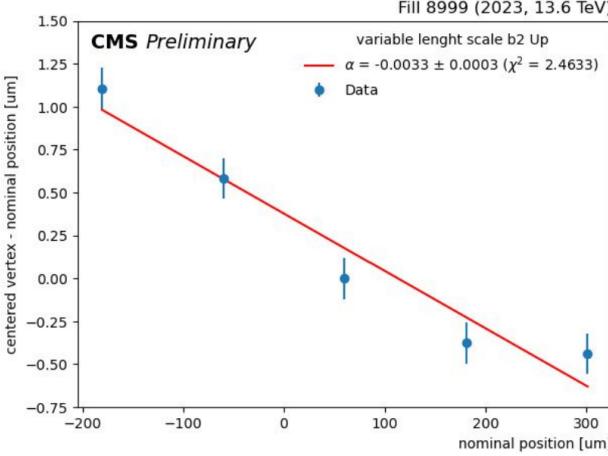


Figure 21: Determination of the length scale factor α from a variable length scale scan (from Ref. [45]).

uncertainty, a systematic uncertainty was added to the already-obtained 0.0011 RMS. This systematic was computed as the difference between the weighted average with and without this faulty scan stage and has an absolute value of 0.0012. This results in a total uncertainty for the length scale factor in the y plane of $\sqrt{0.0011^2 + 0.0012^2} \approx 0.0016$.

The final uncertainty is the quadrature sum of the uncertainties in both planes, 0.2%. The effect of this correction impacts all luminometers equally and resulted in a decrease of the σ_{vis} of 0.9%.

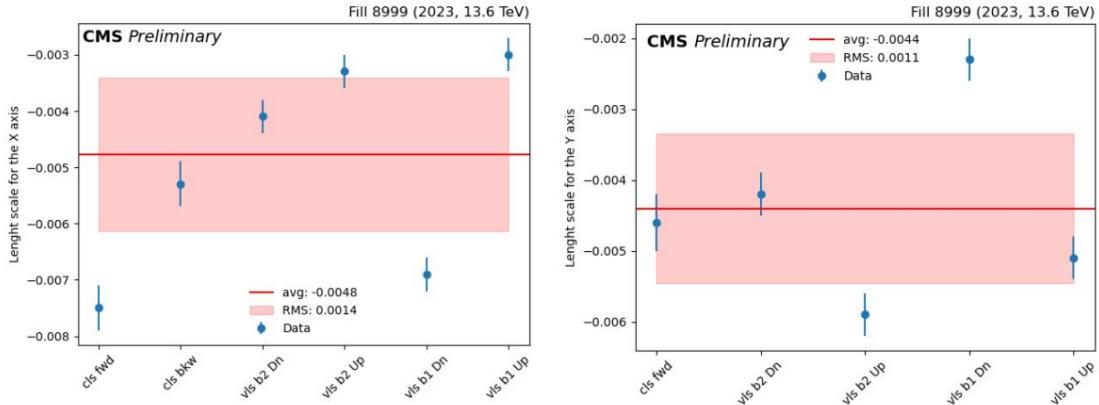


Figure 22: Results of the length scale calibration for the X (left) and Y (right) coordinates, using constant-separation length scale (cls) scans in the forward (fwd) and backward (bkw) directions, and variable-separation length scale (vls) scans (analysed as cls scans) for beam 1(2) scan with beam 2(1) (b2(b1)) being separated in the direction down (Dn) or up (Up) (from Ref. [45]). Length scale values are shown for fits whose reduced- χ^2 is below 100, which was not the case for cls bkw scan in the Y plane.

3.1.4 Beam-beam effects

When proton bunches collide at the IP, a repulsion force will be present due to their electromagnetic interaction. This interaction will have an effect on the luminometer rates, on the beam separations and, consequently, on the measured visible cross section. These effects, referred to as beam-beam effects, are corrected for in two ways.

As a consequence of the repulsive forces, the beam orbits will be deflected. The transverse electric field generated due to the interaction of the two particle bunches, $E(x, y)$, is calculated using the Bassetti-Erskine formula [46]. With the modelling of the transverse electric field, the deflection angle caused by electrostatic repulsion and the orbit shifts can be calculated with Equation 3.7 and Equation 3.8 [47], respectively.

$$\theta_\chi = \frac{2r_p}{\gamma} N_p E_\chi(x, y), \text{ with } \chi \in \{x, y\} \quad (3.7)$$

$$\delta_\chi = \frac{\theta_\chi \beta^*}{2 \tan(\pi Q_\chi)}, \text{ with } \chi \in \{x, y\} \quad (3.8)$$

In these equations, r_p is the classical proton radius, N_p is the opposing bunch population, γ is the relativistic factor and Q_χ is the betatron tune, a beam optics parameter specific to the each experiment. The calculated orbit deflections, δ_χ , are then used to correct the nominal beam separations for every bunch crossing and at each scan step.

When the bunches cross each other, the shape of the convoluted bunch distributions will also be affected by the electromagnetic interaction. This effect, commonly called dynamic- β effect, is the second beam-beam effect that is taken into account. The distorted beam overlap shape will be reflected in a β^* value different from the nominal one, which describes the envelope of the beam's transverse emittance and will affect the measured rates. This way, a corrective factor is multiplied to the measured rates in order to account for this effect. The magnitude of both corrections, beam-beam deflection and dynamic- β , are shown in Figure 23.

Applying these corrections to all vdM and im scans yields an increase in σ_{vis} of 0.7% for all luminometers. The associated uncertainty is computed from the quadrature sum of all the relevant sources of uncertainty summarised in Table 2. A more detailed explanation of the beam-effects, the corrections and their associated uncertainties can be found in [48].

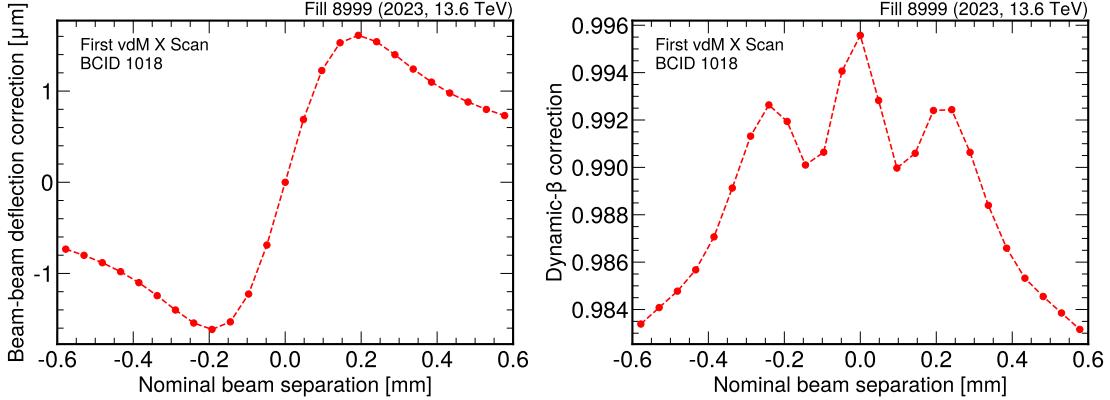


Figure 23: Correction values for the electromagnetic interaction between the 2 bunches. On the left, the orbit deflections due to electromagnetic repulsion are shown while on the right, the rate correction factors due to the distortion of the beam overlap shape are depicted.

3.1.5 Orbit drift

Any drift in the beam positions, be it random or systematic, will result in a discrepancy from the nominal positions and, consequently, affect the visible cross section measurements. To correct for this effect, two beam position monitor systems (BPM) are used: the Diode Orbit and Oscillation System (DOROS) [49] and the BPMs located in the LHC arcs (arcBPM) [16].

The correction applied to the nominal beam positions, using the measurements from both BPMs, is done in two steps.

- **Linear orbit drift:** In this correction, BPM measurements that are taken immediately before and after the scans, when the beam are in their nominal head-on positions, are used to determine the linear orbit drift at each step of a scan. To this effect, a linear interpolation is performed between these two points for every scan and the nominal positions are corrected accordingly. Figure 24 shows this procedure applied to the first vdm scan done in the horizontal transverse direction and Figure 25 illustrates the orbit drift for all vdm and im scan pairs.

After correcting for the orbit drift in the nominal beam separations, we follow up with a correction to the measured rates in the non-scanning direction. To understand why, let us recall our assumption that the particle bunch distributions are XY-factorizable. Under this assumption, the rates measured in some arbitrary (x, y) point would be:

$$R(x, y) = r_x(x) \cdot r_y(y). \quad (3.9)$$

Taking the vertical plane as non-scanning, for example, an orbit drift (Δ_y) would mean that our

Table 2: Contributions to total beam-beam systematic uncertainty. Detailed description of these sources can be found in [48]

Source	Uncertainty (%)
Nominal tunes	0.15
β^* at IP5	0.11
Non-Gaussian tails	0.16
Beam size imbalance	0.01
Reference ambiguity	0.20
Multi-IP tune shift	0.04
Ellipticity	0.03
Polynomial	0.10
Total	0.34

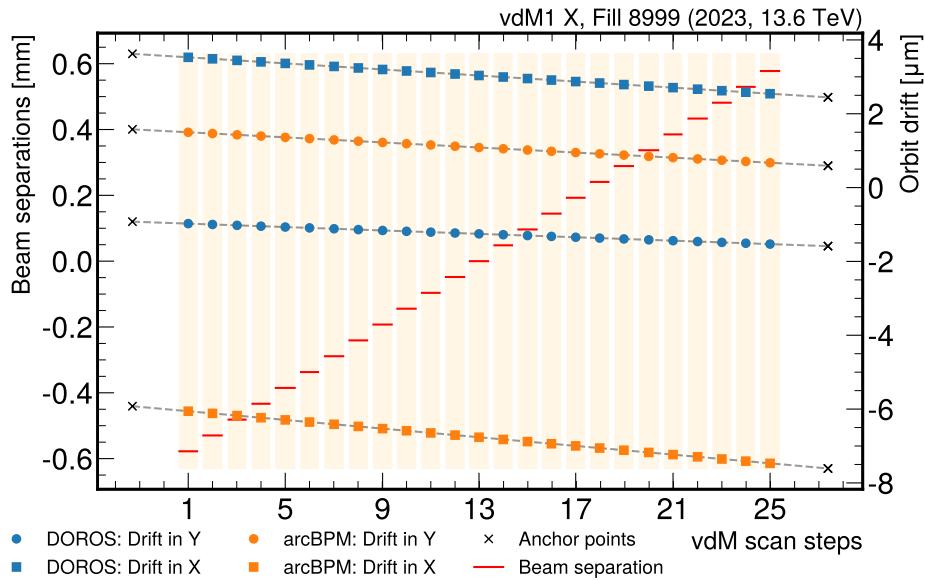


Figure 24: Orbit drift correction calculated from the interpolation of the BPM measurements done before and after the vdM1 scan in the horizontal plane. The nominal beam separations (left axis), at every scan step, are represented by the solid red lines, the anchor points can be seen as black crosses. The linear interpolation was conducted for both BPM systems, DOROS (blue) and arcBPM (orange), and both planes, horizontal (squares) and vertical (circles).

measured rates would be given as:

$$R(x, y + \Delta_y) = r_x(x) \cdot r_y(y + \Delta_y). \quad (3.10)$$

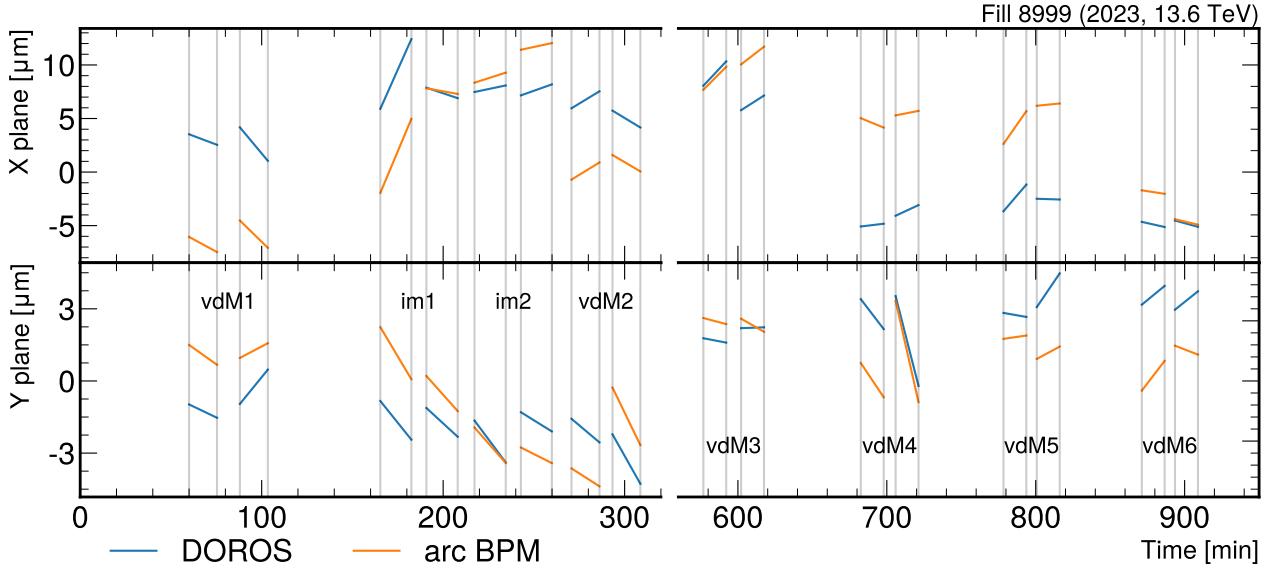


Figure 25: Orbit drift correction for all vdM and im scan pairs.

We can then combine both equations to arrive at the rate correction:

$$R(x, y) = R(x, y + \Delta_y) \cdot \frac{r_y(y)}{r_y(y + \Delta_y)}. \quad (3.11)$$

- **Residual orbit drift:** After the linear orbit drift (linOD) correction is applied, some residual differences remain between the nominal (NOM) and the BPM measured beam positions. The length scale and the beam-beam (BB) deflection corrections are part of this residual difference and are accounted for as described in the previous sections. To correct for the remaining unpredictable deviations the residual differences are fitted as follows:

$$\text{BPM}_{\text{scanning}} - \text{linOD}_{\text{scanning}} = \alpha \text{NOM}_{\text{scanning}} + \beta \text{BB}(\Delta \text{NOM}_{\text{scanning}}) + c_{\text{scanning}} \quad (3.12)$$

$$\text{BPM}_{\text{stationary}} - \text{linOD}_{\text{stationary}} = \gamma \text{NOM}_{\text{scanning}} + \frac{\gamma}{\alpha} \beta \text{BB}(\Delta \text{NOM}_{\text{scanning}}) + c_{\text{stationary}} \quad (3.13)$$

α encodes the length scale factor difference between the nominal positions and the BPM measurements, γ , referred to as the leakage parameter, deals with the apparent misalignment between the measured and the nominal coordinate system, β is a multiplicative factor to the beam-beam deflection at the IP and c is simply a constant factor. After extracting the fit parameters, the residual orbit drifts are calculated as:

$$\Delta x_{\text{scanning}} = \text{BPM}_{\text{scanning}} - \text{linOD}_{\text{scanning}} - \alpha \text{NOM}_{\text{scanning}} - \beta \text{BB}(\Delta \text{NOM}_{\text{scanning}}) \quad (3.14)$$

$$\Delta x_{\text{stationary}} = \text{BPM}_{\text{stationary}} - \text{linOD}_{\text{stationary}} - \gamma \text{NOM}_{\text{scanning}} - \frac{\gamma}{\alpha} \beta \text{BB}(\Delta \text{NOM}_{\text{scanning}}) \quad (3.15)$$

The fittings were performed simultaneously on both equations and with different relaxation levels for the α and β parameters. The methods were:

- **Separate fit (Sep)**: Every parameter was fitted separately for every scan in the vdM program.
- **Average BB (AvgBB)**: The β parameter used in the residual calculation is the average among all vdM scans.
- **Common BB (ComBB)**: The β parameter is fitted globally across the scans.
- **Common LS (ComLS)**: The α parameter is fitted globally across the scans.
- **Common LSBB (ComLSBB)**: ComBB and ComLS methods combined.

In all methods, the leakage parameter γ and the constant c are fitted separately. The per-scan step residual correction calculated with the fit parameters obtained from the CommonLSBB fit method are shown in Figure 26.

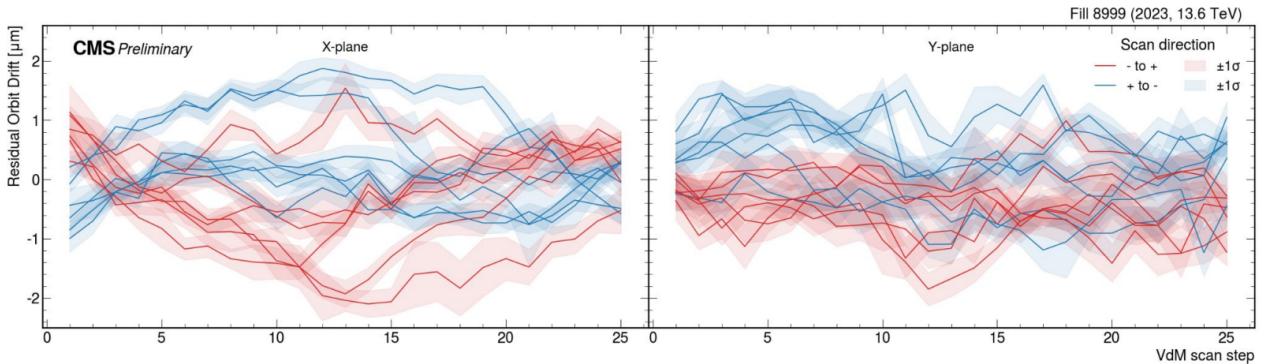


Figure 26: Residual beam position corrections as a function of the scan step number for each vdM scan estimated using the arc BPM data and performing a global fit to all scans, which assumes a common BPM length scale and a common scale factor for the beam-beam deflection amplitude after correcting for the slow linear orbit drift. The red (blue) lines represent the residuals for the beam moving from the negative (positive) to the positive (negative) direction with respect to the reference CMS coordinate system. The band is the uncertainty that is estimated from the standard deviation of the position measurements (performed every second) within a scan step (from Ref. [45]).

Upon correcting the beam separations and the non-scanning beam rates for the linear orbit drift, the visible cross section results increased by 0.22% and 0.26% for the arcBPM and DOROS measurement, respectively. The uncertainty of this correction was taken as half the difference of the two implementations,

0.02%. For the residual orbit drift corrections, the impact of all fit methods were compared. Figure 27 displays the comparison for the HFET luminometer as an example. The numerical results are shown in Table 3.

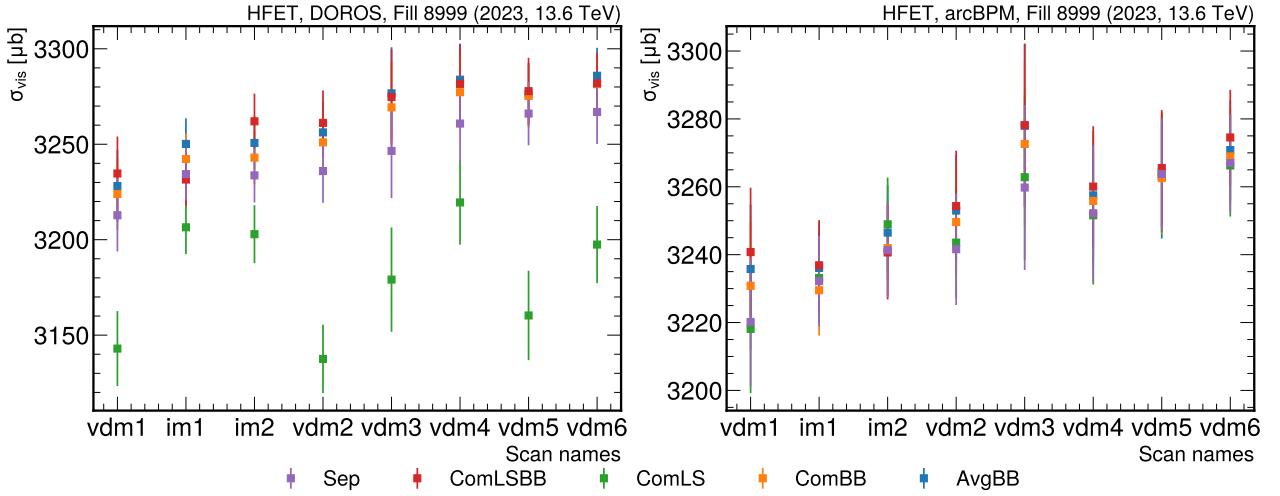


Figure 27: HFET visible cross section estimates for the vdM and im scan pairs after applying the orbit drift corrections. The central values are obtained by calculating the weighted average of the scan results. The scan to scan consistency is represented by the equivalent weighted standard deviation.

Table 3: HFET visible cross sections for the many residual fit methods for both arcBPM and DOROS systems.

Method	DOROS		arcBPM	
	$\sigma_{\text{vis}} [\mu\text{b}]$	Uncertainty (%)	$\sigma_{\text{vis}} [\mu\text{b}]$	Uncertainty (%)
Sep	3243.48	0.53	3246.49	0.47
AvgBB	3262.38	0.56	3252.75	0.42
ComLS	3183.52	0.91	3247.55	0.46
ComBB	3255.48	0.58	3249.46	0.47
ComLSBB	3260.45	0.59	3254.09	0.46

The results obtained from the separately-fit length scale parameter methods (Sep, AvgBB and ComBB) have not been considered since the length scale parameter as computed with the length scale calibration is constant throughout the vdM program. The ComLSBB method with the measurements from the arcBPM system was chosen for its smaller scan-to-scan variation (0.46%). This correction increased the visible cross sections of all luminometers by 0.2% with an associated uncertainty of 0.16%, estimated from the standard deviation of results from all fits methods.

3.1.6 XY factorization

The assumption that the particle bunch distributions are factorizable in the transverse plane has been maintained throughout all the previously mentioned corrections. Since this assumption is not guaranteed, the visible cross section differs from the one measured with Equation 2.5 by what is referred to as the factorization bias. This bias is calculated by reconstructing the bunch overlap densities of the two beams during the vdm fill and comparing the results to the ones obtained from the one-dimensional vdm scans.

Different methods are used for this purpose such as the beam-imaging method described in [38], [50], [51]. For this analysis two other methods were used and their results were compared.

One of these methods is the 2D rate fit method, which has been carried out several times in previous CMS luminosity analyses [44], [52]. In this method, on-axis scans, like the previously mentioned vdm and im scans, are combined with off-axis scans, such as diagonal and offset scans, that were performed close in time. In diagonal scans, the scanning beam is moving in both the horizontal and vertical directions in such a way that the scanning plane is at a 45° angle. Offset scans are like normal vdm scans with an additional separation in the non-scanning plane. These off-axis scans are illustrated in Figure 28.

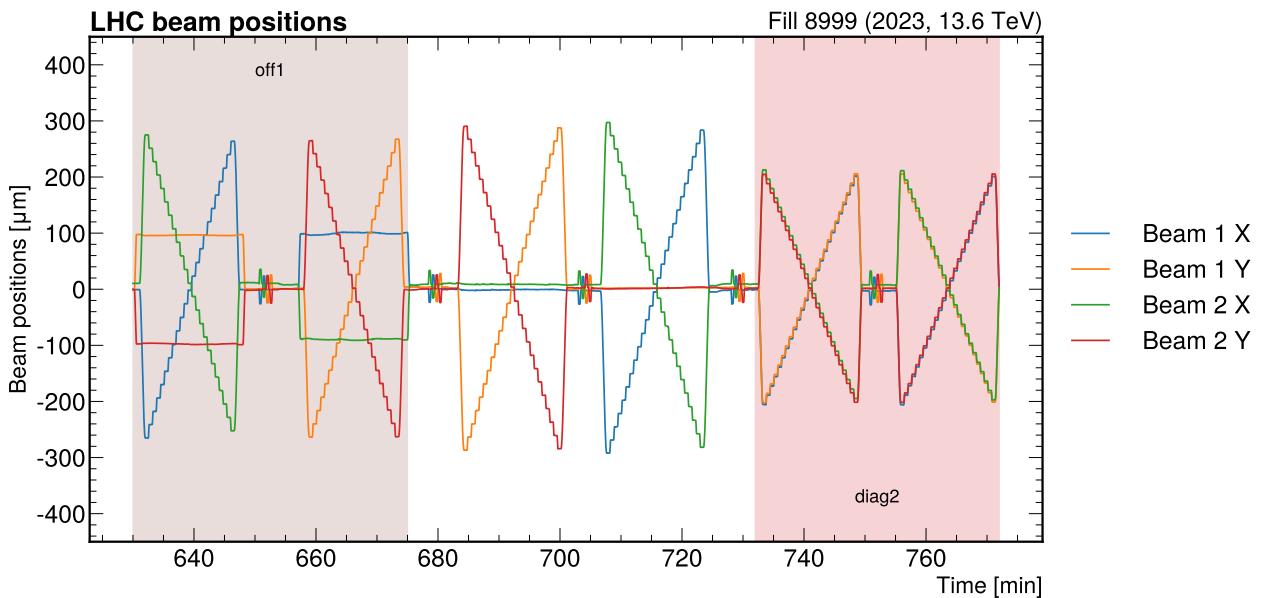


Figure 28: Beam positions during the first offset scan and the second diagonal scan of fill 8999.

The combination of on-axis with off-axis scans allows for the probing of the overlap bunch distribution along different paths. This provides more information about the non-factorizable nature of this distribution. For 2023 vdm program, the scan combinations were vdm1+diag1, diag1+Bl1, vdm3+off1, off1+vdM4, vdm4+diag2 and diag2+vdM5. The visible cross section corrections for all these combinations are shown in Figure 29.

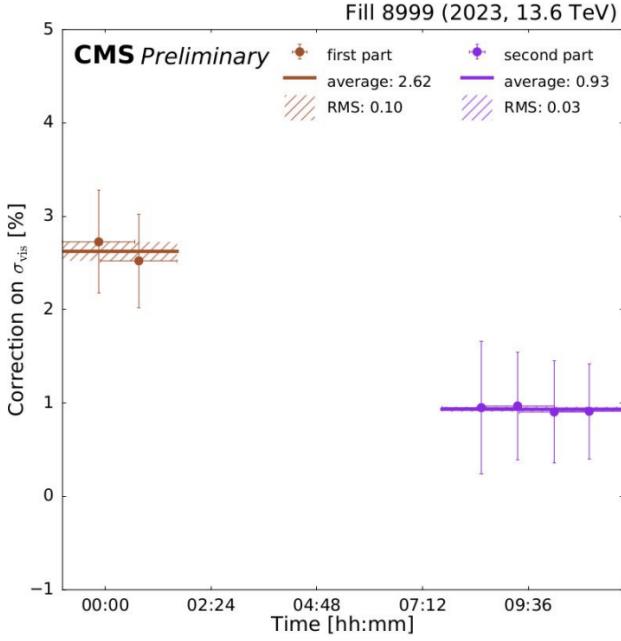


Figure 29: Factorization corrections on the visible cross section from several on-axis scan and off-axis scan pairs as a function of time. The points from left to right correspond to vdM1+diag1, diag1+BI1 combinations taken during the first part of the vdM program (in brown), and to vdM3+off, off+vdM4, vdM4+diag2, diag2+vdM5 combinations taken during the second part of the vdM program (in purple). Over the nine hours of the data taking, the factorization correction changed from about 2.62% to approximately 0.93%, indicating a significant time dependence (from Ref. [45]).

Lastly, a 3D fit of the luminous region is performed to luminometer rates using parameters of the 3D ellipsoid of the reconstructed beamspot and the bunch length data from the LHC LDM. One of the main advantages of this method, in comparison to the 2D rate method, is the ability to obtain a cross section measurement for every vdM and im scan separately. For this reason, the corrections to the detectors calibration were derived from this method. A comparison to the 2D rate fit method is shown in Figure 30, which shows good agreement between both methods. The large time dependence in the correction values demonstrates the change in beam overlap shape throughout the vdM fill.

The XY-factorization bias was the main responsible for reducing the scan-to-scan variation seen in Figure 14 and Figure 15. The total uncertainty for this correction is 0.67%, which was estimated from the uncertainty associated with model- and BCID-dependence and the maximum discrepancy with the 2D rate fit method, among other sources [44].

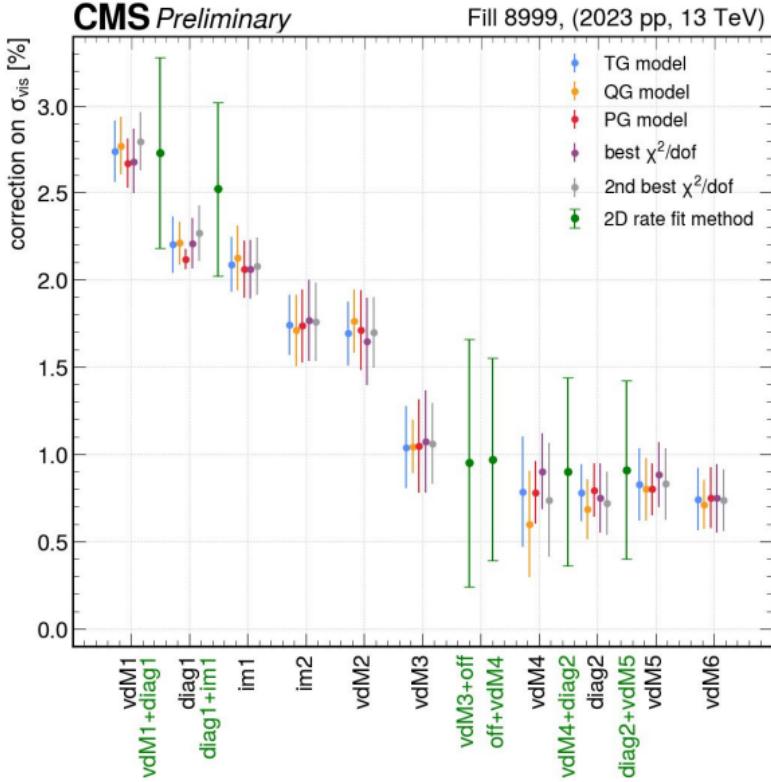


Figure 30: Average correction for the visible cross section over nine BCIDs using five different models of the luminous region (LR) method. The scans are ordered in time. The averages include only BCIDs for which high-statistics zero bias samples are available, and the uncertainties indicate the RMS over these BCIDs. A small number of bad quality fits with anomalously large reduced $\chi^2 > 5$ were discarded. Results for three models of the bunch proton density are shown: weighted sum of three (TG), four (QG) or five (PG) 3-dimensional Gaussians. The results – choosing the best and the second best (i.e. smallest) reduced χ^2 fits from the twenty seven tested models – are also given (from Ref. [45]).

3.1.7 Calibration results

The visible cross section results were determined by performing a weighted average on the calibrations measured in the eight analyzed scans, incorporating all prior corrections. Table 4 provides a summary of the results for all independently calibrated luminometers, along with their associated uncertainties. The associated σ_{vis} evolution plots are in Figure 31 for the HFET detector and in Figure 32 and Figure 33 for the remaining luminometers.

The bunch-to-bunch variation in the visible cross section measurements is attributed as a source of uncertainty and has a magnitude of 0.6%, the standard deviation of the mean in HFET. Similarly, the calibration variation across scans was also used as another source of uncertainty. HFET's weighted RMS

Table 4: Summary of the measured visible cross sections. The results have been computed through an average weighted on the error of each scan and are accompanied by their standard error on the mean. The bunch-to-bunch uncertainty range across all scans is shown in the "Bunches RMS" column. The "Scans RMS" depicts the overall scan-to-scan variation.

Luminometer	Measured σ_{vis}	SEM	Bunches RMS range	Scans RMS
BCM1F	136.36 μb	0.17 μb	0.58-1.03 μb (0.43-0.76%)	0.41 μb (0.30%)
BCM1FUTCA	127.55 μb	0.15 μb	0.61-1.09 μb (0.48-0.85%)	0.40 μb (0.31%)
HFET	3294.69 μb	3.86 μb	12.82-21.78 μb (0.39-0.66%)	9.28 μb (0.28%)
HFOC	946.05 μb	1.12 μb	2.49-5.92 μb (0.26-0.63%)	2.63 μb (0.28%)
PLT	301.39 μb	0.37 μb	1.52-2.38 μb (0.50-0.79%)	0.90 μb (0.30%)
PCC	1305.23 mb	1.21 mb	3.31-8.85 mb (0.25-0.68%)	3.27 mb (0.25%)

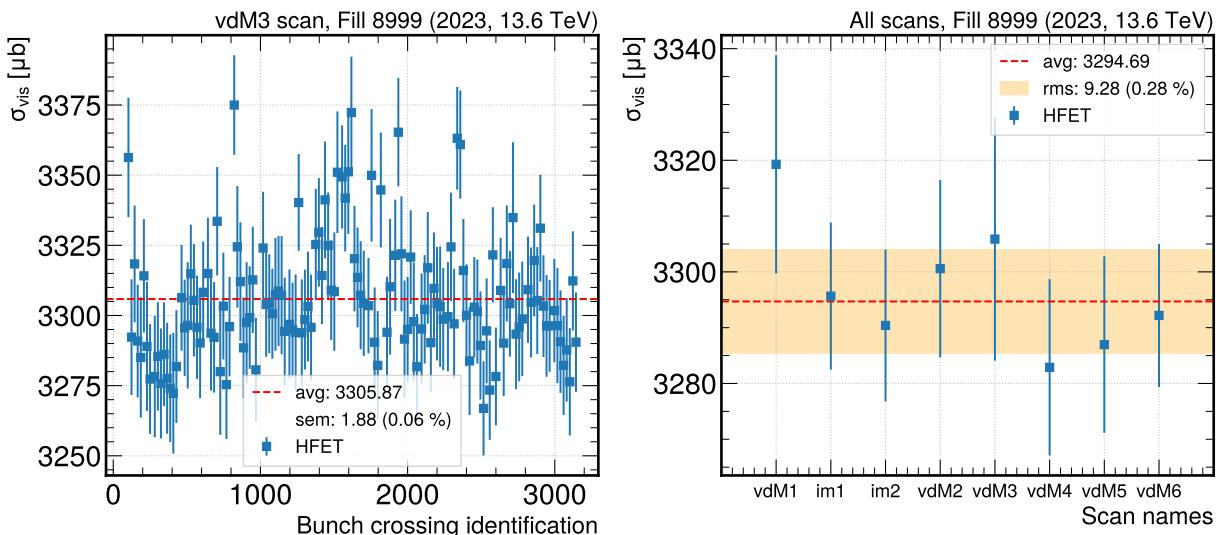


Figure 31: Final visible cross section results obtained after all data corrections for the HFET luminometer, as an example. On the left, the measured per-bunch visible cross sections with its associated fit uncertainty are shown. The right image shows the measured visible cross section taking the average over all 136 colliding BCIDs for each vdM and im scan pairs. All statistics are determined as in Figure 14.

across all scans, 0.28%, was the attributed uncertainty.

The fit results were compared among the different luminometers and their individual reduced χ^2 per scanning plane were plotted for all vdM scans. Overall, the results indicate good fit quality for both axis, with fits performed on HFOC reporting to have slightly higher reduced χ^2 in the vertical plane. An example of the fits for the BCM1FUTCA detector is shown in Figure 34 and the distributions for all online

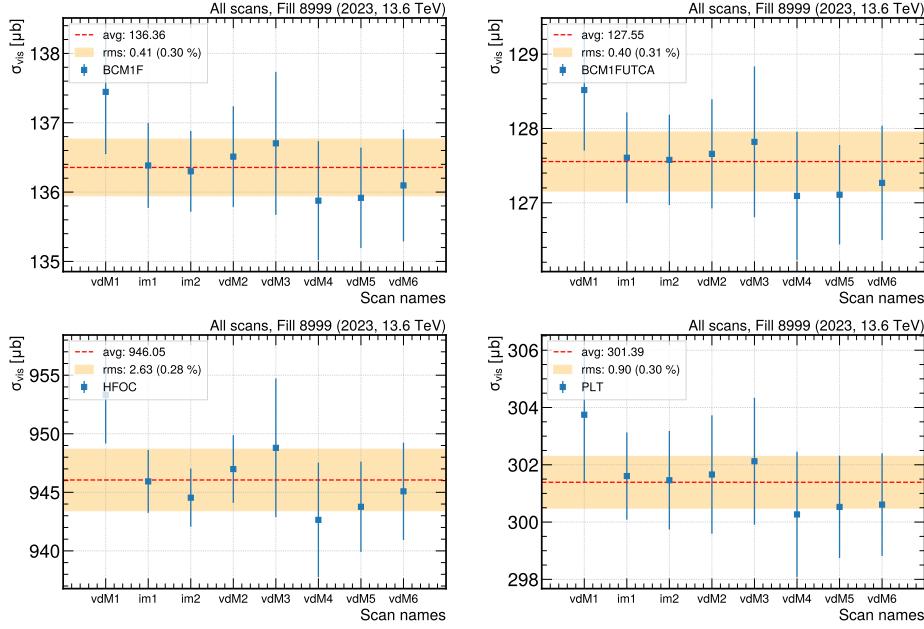


Figure 32: Final visible cross section results obtained after data corrections for BCM1F, BCM1FUTCA, HFOC and PLT luminometers.

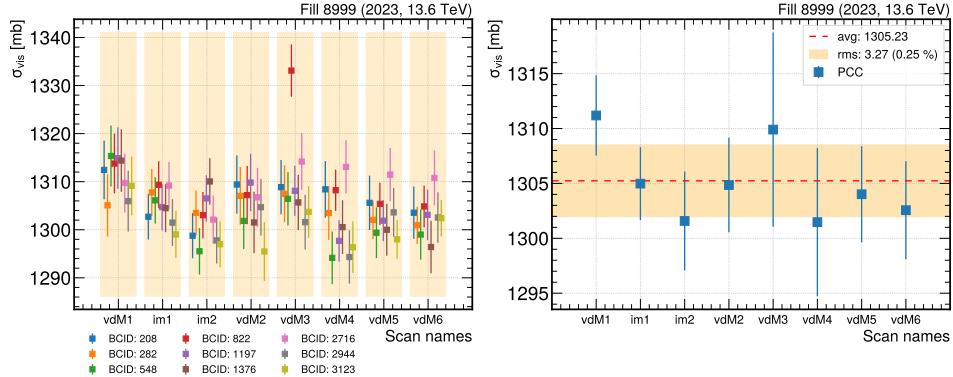


Figure 33: Left: PCC calibration results for all the vdm fill scans and recorded BCIDs (208, 282, 548, 822, 1197, 1376, 2716, 2944, and 3123). Right: PCC scan-to-scan variation across all vdm and im scans.

luminometers can be seen in Figure 35.

Using the resulting fits, we extract the measured effective width and height of the luminous area, Σ_X and Σ_Y , for all of the luminometers. As it is a property of each bunch-pair, determined by their emittances and the constant β^* at CMS, it should be consistent for all luminometers. The comparison is shown in Figure 36 and Figure 37, and good agreement is observed. In the latter figure, the time-evolution of beam emittances can be observed.

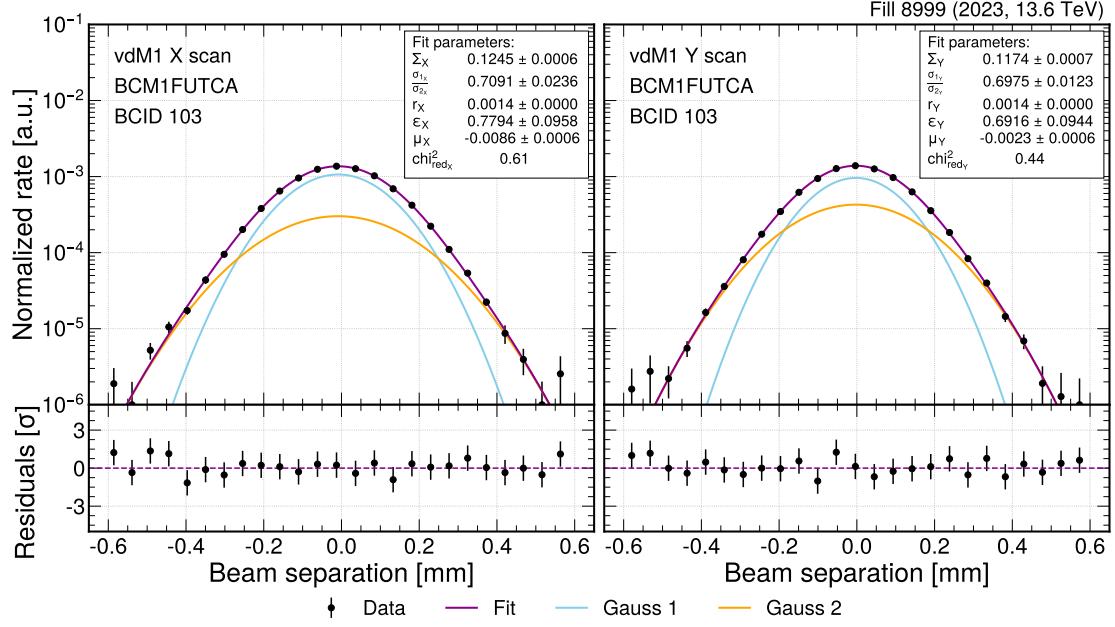


Figure 34: Double Gaussian fits to the BCM1FUTCA data recorded during the first vdM scan pair (vdM1), shown for the x (left) and y (right) scan. In the bottom panels, the statistical pull for each data point is shown. The boxes indicate the fit parameters for the fit in question. From top to bottom these contain the bunch overlap width, the ratio between both Gaussians, the rate at the peak, the weight factor for the first Gaussian curve, the mean of the overall distribution, and the reduced χ^2 .

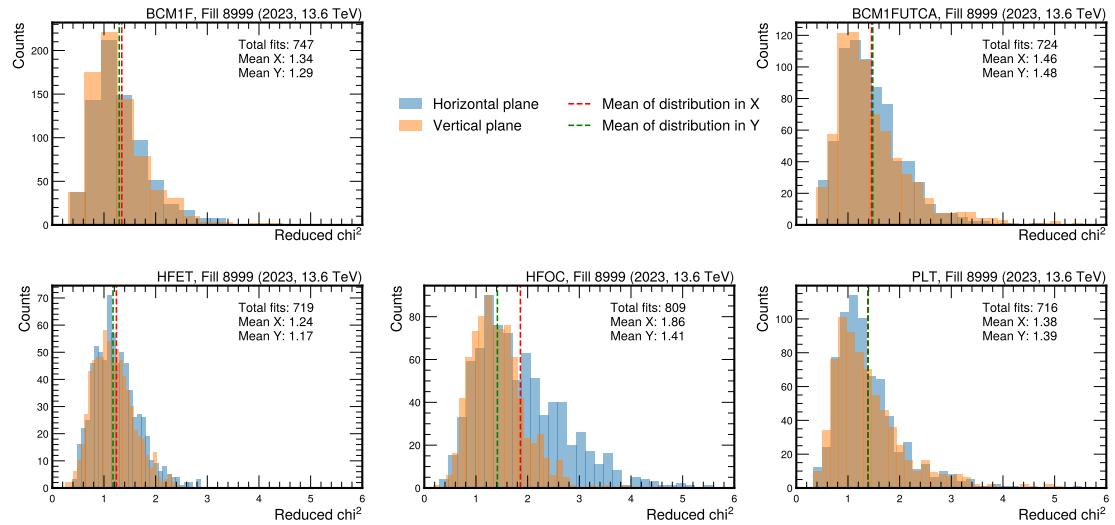


Figure 35: Overview of the reduced χ^2 distributions for all online luminometers for both scanning planes and all vdM scans. Only fits which converge are considered which leads to a slightly different total number of fits per detector.

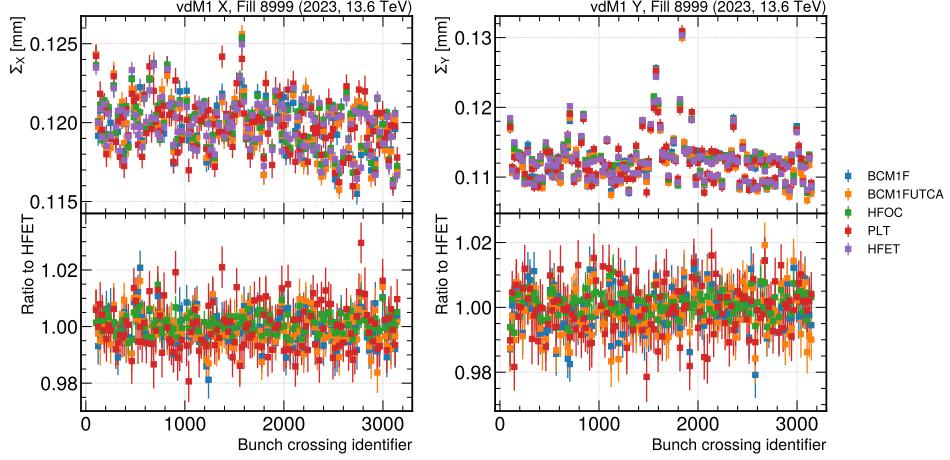


Figure 36: Overview of the effective width and height of the luminous area, Σ_X and Σ_Y , measured during the vdM1 scan pair, by all the online luminometers (upper row). Ratios of the measured values by different systems with respect to HFET are also shown (bottom row).

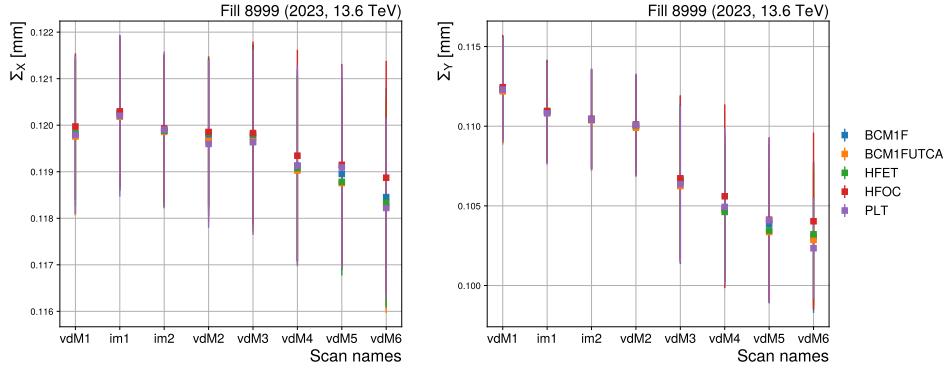


Figure 37: Overview of the average Σ values in both transverse plans measured during all vdM and im scans, by all the online luminometers.

3.2 Integration

In the previous section, the process by which the detector calibration factors were obtained was discussed. In this section, the performance of these calibrations in maximizing the agreement between luminometers during the vdM fill and, more importantly, under physics conditions is evaluated. A discussion on the year-long detector consistency will be presented, along with details on the specific stability corrections applied to the BCM1F. Additionally, the issue of detector non-linearity will be addressed, focusing on the PLT corrections. Finally, this section will present the preliminary integrated luminosity for 2023, along with the associated relative uncertainty.

3.2.1 vdM Consistency

Using the calibration values shown in Table 4, the measured luminosity of each independently calibrated detector can be compared to visualize the detector agreement under vdM conditions. Figure 38 shows the instantaneous luminosity throughout the entire vdM program.

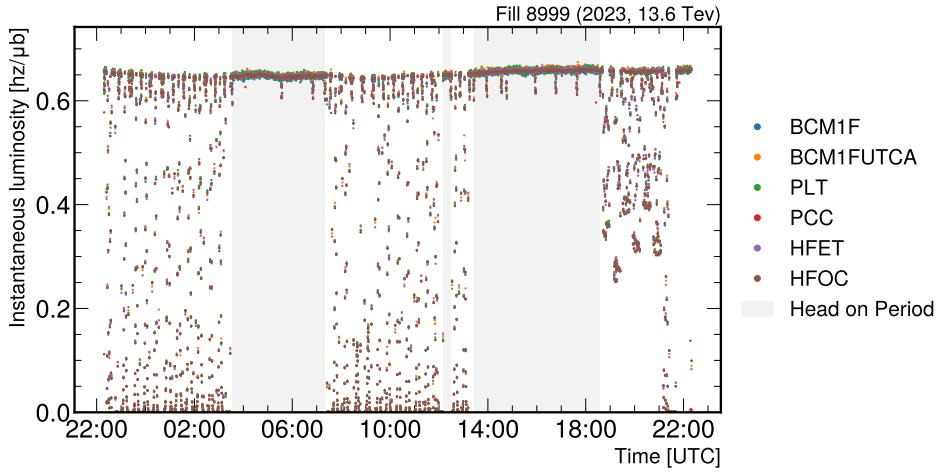


Figure 38: Measured instantaneous luminosity as a function of time during the vdM program. All independently-calibrated luminometers show good agreement, especially during the head-on periods marked by the light grey areas.

The luminometers show very good agreement with each other. The cross-detector consistency during the vdM fill was evaluated from the head-on ratios of all the luminometers with respect to the overall average. Figure 39 shows the agreement of the luminometers during these periods in the vdM fill. The overall spread of the detectors was taken as the standard deviation of the means, with a magnitude of 0.16% (see Figure 40).

The data in Figure 39 corresponds to the periods when the LHC beams were colliding head-on, with all six luminometers taking high-quality data, and for the sum of the nine BCIDs for which PCC data was recorded in zero bias events. The ratio to the average measured luminosity is shown on the right and depicts an agreement well below 1%. The statistical uncertainty associated with each point is found to be negligible and is not shown in the plot.

With the vdM calibrations validated, they can be used to calibrate the rates of all luminometers for the entire running year in accordance with Equation 1.1. The results of this calibration procedure are shown in Figure 41.

Three main observations can be made when analyzing the ratios in Figure 41. The most obvious is the high discrepancy between different luminometers. PLT stands out as a clear example of this, but

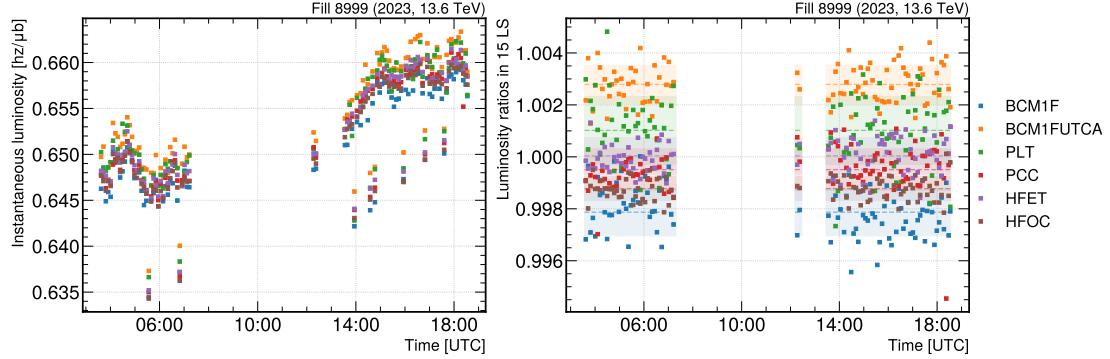


Figure 39: Head-on instantaneous luminosity measurements provided by the individual independently calibrated luminometers BCM1F, BCM1FUTCA, PLT, PCC, HFOC, and HFET during the LHC fill 8999 as a function of time. Each point corresponds to a time window of 15 lumisections, with each lumisection spanning about 23 seconds.

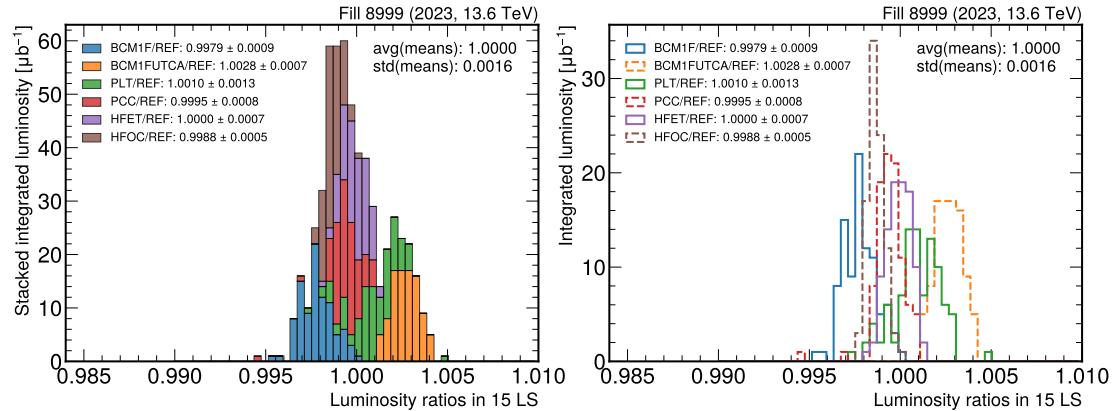


Figure 40: Distributions of the ratio of the measured luminosities between different luminometers and the reference luminosity. The left plot shows a stacked distribution while the right plot shows individual distributions. The means and standard deviations of the individual distributions are shown. The reference luminosity (REF) corresponds to the average luminosity of all the individual luminometers.

in general, the luminometers do not show good agreement. Secondly, some detectors exhibit significant steps throughout the year. Notable examples include BCM1F at the beginning and HFET from 20.5 fb^{-1} to 22 fb^{-1} and from 27 fb^{-1} to 27.5 fb^{-1} . Lastly, the ratios show some non-linear responses, as is clearly seen in PLT. These observations highlight the difficulty in extrapolating the calibrations outside the vdM fill, as explained in subsection 2.1.3.

In the next sections, the issues of detector stability and non-linearity will be addressed, with examples of these corrections for the BCM1F and PLT detectors.

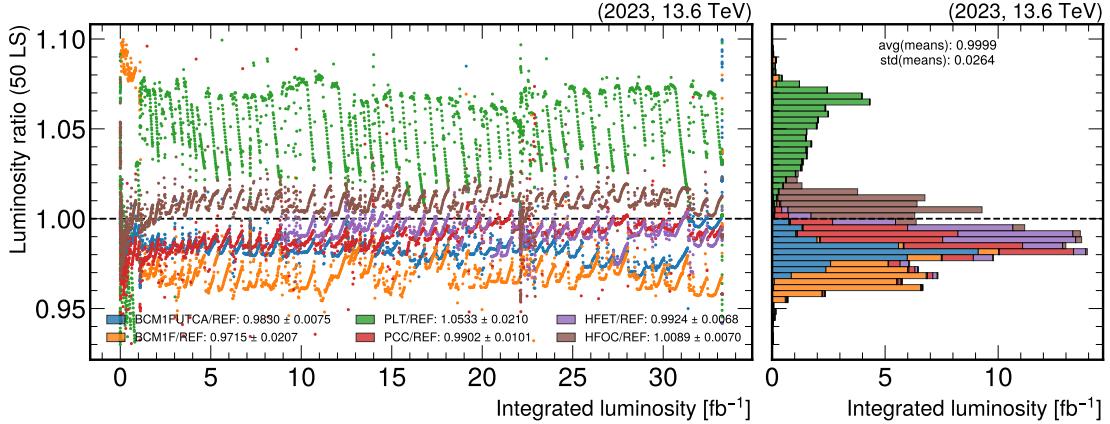


Figure 41: The left plot shows the ratios of the luminosity measured in time windows of 50 LS (about 20 minutes) between the different luminometers and the average luminosity among them as a function of the integrated luminosity. The ratio distributions are shown in a histogram on the right, with the associated averages and standard deviations of the means indicated.

3.2.2 Detector Stability

The loss of efficiency that a detector exhibits throughout the year is corrected empirically by multiplying the detector's visible cross section by a factor ϵ , as explained in Equation 2.6. The method for obtaining these factors is detailed below.

Many fills conducted during the year 2023 had designated beam time to perform emittance scans. From these scans, the detectors' visible cross sections are continuously measured, allowing for the tracking of the detectors' efficiency throughout the year. The loss of efficiency experienced by a detector is indicated by gradually decreasing σ_{vis} measurements. When subsequent measurements do not follow the expected trend, it typically indicates that the detector is operating under non-optimized conditions. In such cases, the detector's running configurations are adjusted to restore normal operation conditions.

When extrapolating the vdM calibration to other fills, a linear regression is performed on the FOM values as a function of integrated luminosity. The fitted line is then normalized to the vdM since this represents our best calibration measurement. Figure 42 shows the fit performed for the BCM1F measurements. The significant step at the beginning of Figure 42 corresponds to a period when the BCM1F current thresholds were set so high that it was counting noise as hits, leading to the apparent increase in efficiency.

Once the fits are performed, the fit values at each point are used to extrapolate the vdM calibrations for the corresponding fills.

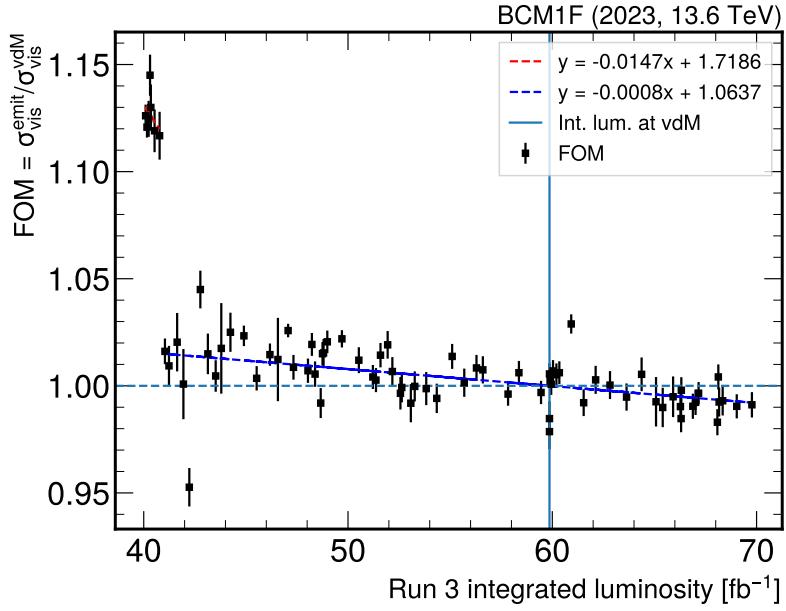


Figure 42: The results of emittance scan analysis for the BCM1F detector. The FOM values are shown as a function of integrated luminosity throughout Run 3. There are two distinct regions that are fitted independently, with the fit parameters shown in the legend.

3.2.3 Detector Non-Linearity

At the time of writing, non-linearity corrections were only carried out for BCM1F and PLT. As explained in subsection 2.1.3 a detector assumed to have a good linear response to varying experimental conditions is used as the reference for which other detectors will be corrected for. In this analysis the DT luminometer, referenced in subsection 1.3.5, will be used as the reference assumed to be linear.

As in the previous section, this correction is also done at the fill level. However, in order to minimize the correlation between the luminosity of the detectors to the reference detector, only a small number of fills with a large SBIL range are considered. The linearity correction for PLT is detailed below.

A total of 10 fills, spread across the year, were chosen to determine the slope quantifying the amount of non-linearity. For each fill, the luminosity ratios with respect to DT as a function of SBIL were computed. The fills before and after the vdM fill were placed in different groups since PLT's running configurations were changed in order to improve its efficiency for the vdM fill. The results of this analysis are shown in Figure 43.

The extracted slopes have to be 1%/SBIL since this was the non-linearity configuration assumed during online data taking. Averaging the 2 regions yielding 1.06%/SBIL and 0.98%/SBIL for before and after vdM, respectively.

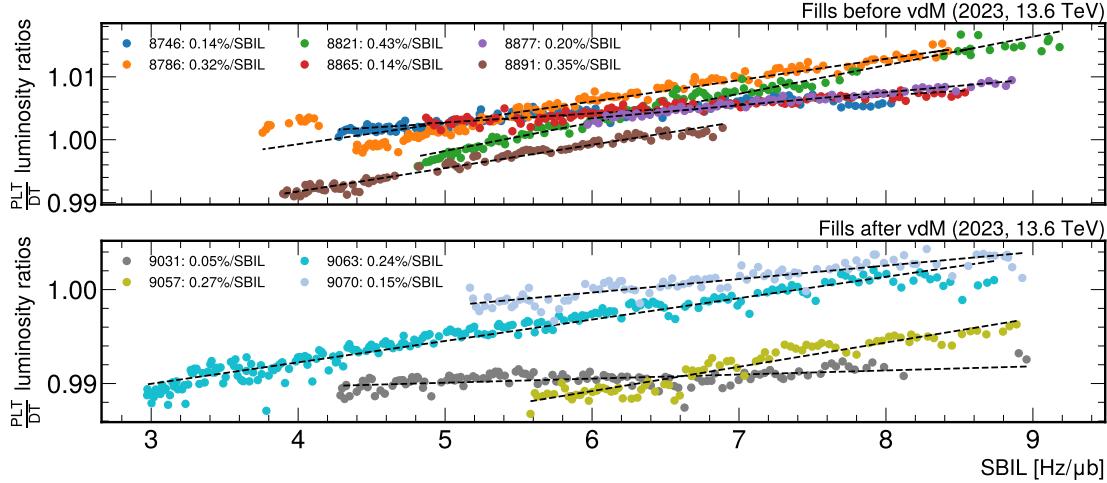


Figure 43: PLT non-linearity slopes extracted using DT as the reference luminometer for selected fills. The dashed lines correspond to the linear fits and the slopes are shown in % per SBIL in the legend. The upper plot contains the fills before the vdm program and the bottom plot contains the fills after.

For BCM1F, the non-linearity was extracted from a μ -scan fill. A μ -scan is a vdm-like scan that also spans a much wider SBIL range. The slopes were extracted in the same way as in the 2022 analysis [44] and a $-0.12\%/\text{SBIL}$ correction was applied.

3.2.4 Establishing a Reference Detector

In Figure 41, the ratios were taken with respect to the average luminosity of all luminometers. However, the luminosity results have to be available via the *brilcalc* database, which places constraints on the data format. For this reason, at the time of writing, DT was chosen as the reference detector. Since DT is not an independently-calibrated luminometer due to lack of available statistics, it was calibrated to the average of the other luminometers, effectively representing the average luminosity among them.

During the 2023 analysis, DT had been cross-calibrated to HFET in fill 9036 which yielded a calibration of $\sigma_{\text{vis}}^{\text{DT}} = 179725 \mu\text{b}$. To scale this cross-calibrated version of DT to the luminosity average, the average ratio of every other luminometer with respect to this DT was taken. The results are shown in Table 5.

Weighting the average of the values in Table 5 by the inverse square of the individual RMS yields a corrective factor of 0.991. This final cross-calibrated version of DT is used as a reference for the stability and non-linearity uncertainties discussed in the following section.

Table 5: Average and RMS of all fully-corrected independently-calibrated luminometers with respect to HFET-cross-calibrated DT.

Luminometer	Average ratio to cross-calibrate DT	RMS
BCM1F	0.985	0.005
BCM1FUTCA	0.985	0.005
HFET	0.998	0.005
HFOC	1.002	0.007
PLT	0.986	0.007
PCC	0.997	0.006

3.3 Final Luminosity Results

As in subsection 3.2.1, the year-long detector stability is illustrated. The fully corrected visible cross sections are utilized, with stability and non-linearity corrections applied.

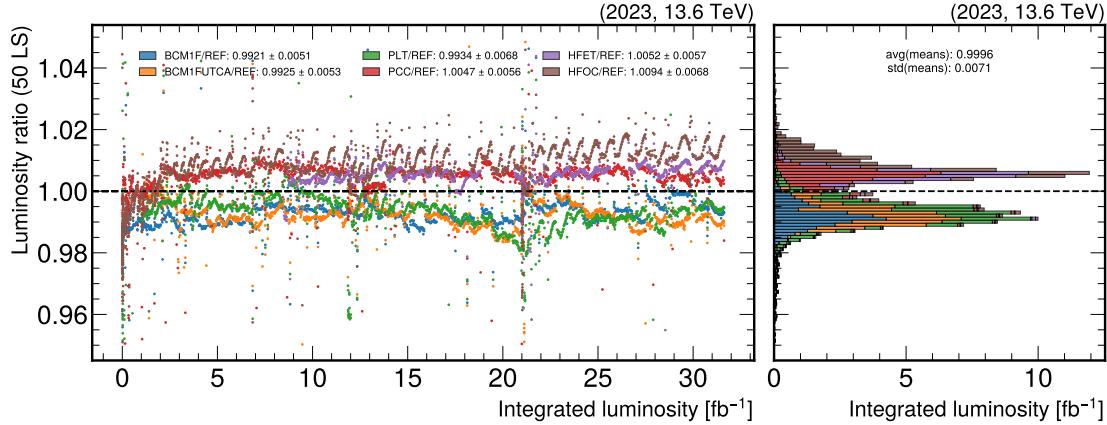


Figure 44: The left plot shows the ratios of the luminosity measured in time windows of 50 LS (about 20 minutes) between different luminometers and the reference luminometer (the cross-calibrated DT) as a function of the integrated luminosity over the 2023 pp data-taking period.

A comparison of Figure 41 and Figure 44 demonstrates a significant improvement in detector agreement. The standard deviation of the means decreased from 2.6% to 0.71%, which corresponds to the uncertainty associated with the stability corrections.

To assign an uncertainty to any residual non-linearity in the detectors, the slopes are extracted for all fills, as explained in subsection 3.2.3. The resulting distributions are analyzed, as shown in Figure 45.

The detector with the maximum deviation, HFET, is selected to assign the uncertainty for this step.

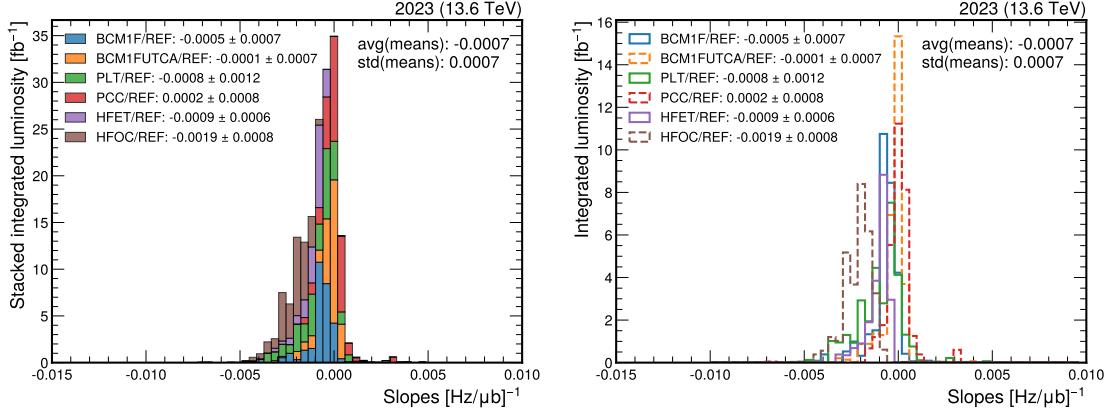


Figure 45: Distributions of relative co-linearity slope values between different luminometers and the DT luminometer taken as reference (REF). Each histogram entry is weighted by the integrated luminosity of the fill.

HFOC is excluded due to its complex non-linearity effects, which were not corrected for. The value of 0.09%/SBIL is multiplied by the average SBIL in 2023 of 6.6 $\text{Hz}/\mu\text{b}$, resulting in an overall uncertainty of 0.59%.

Certain luminometers are unavailable for some periods throughout the year. Hence, all available luminometers are prioritized based on their measured luminosity contributions to the integrated luminosity. For this analysis, the priority order is DT, PCC, HFET, PLT, and BCM1FUTCA. The total delivered luminosity was 32.74 fb^{-1} , and the CMS recorded luminosity was 30.12 fb^{-1} , with an associated uncertainty of 1.28%, as summarized in Table 6.

Table 6: Summary of uncertainty contributions for the 2023 luminosity calibration analysis.

Source	Uncertainty (%)
Calibration	
Beam current	0.20
Ghosts & satellites	0.10
Orbit drift	0.02
Residual beam positions	0.16
Beam-beam effects	0.34
Length scale	0.20
Factorization bias	0.67
Scan-to-scan variation	0.28
Bunch-to-bunch variation	0.06
Cross-detector consistency	0.16
Integration	
Cross-detector stability	0.71
Cross-detector linearity	0.59
Calibration	0.89
Integration	0.92
Total	1.28

Chapter 4

Analysis Software

The analysis of vdM data, along with the application of stability and linearity corrections, is carried out using two distinct frameworks, each playing a crucial role in ensuring the accuracy and reliability of the results.

The vdM Framework [53] is employed to execute all the corrections and fits detailed in chapter 3. This tool is vital for managing the complex data processing required during the vdM scans, enabling precise calibration and adjustment of the luminometers. My contributions to the development and enhancement of this framework are discussed in section 4.1.

For integration corrections, the BRIL work suite, (*brilws*) [54], is utilized. This software suite takes on the critical task of applying the final layer of luminosity corrections, as well as making the results available to other CMS groups. The specific program responsible for these corrections is called *brilcalc*. The inner workings, limitations, developed alternatives, and my contributions to the improvement of this program are detailed in section 4.2.

4.1 The vdM Framework

The vdM Framework (vdMFw) is the analysis software used by CMS to perform luminosity calibration measurements for all luminometers with per-bunch granularity. First commissioned in 2015, it has been continuously improved in line with the collaboration's growing understanding of the systematic uncertainties discussed in chapter 3. Today, it serves as the primary framework for analyzing vdM scans, applying data corrections, and tracking luminometer operational conditions through emittance scan analysis throughout the year.

4.1.1 Flow of the Analysis

The flow of analysis in the vdMFw, outlined in Figure 46, begins with user-provided input. This input can be categorized into two types: the data to be analyzed and the user specifications that dictate how the analysis should be conducted.

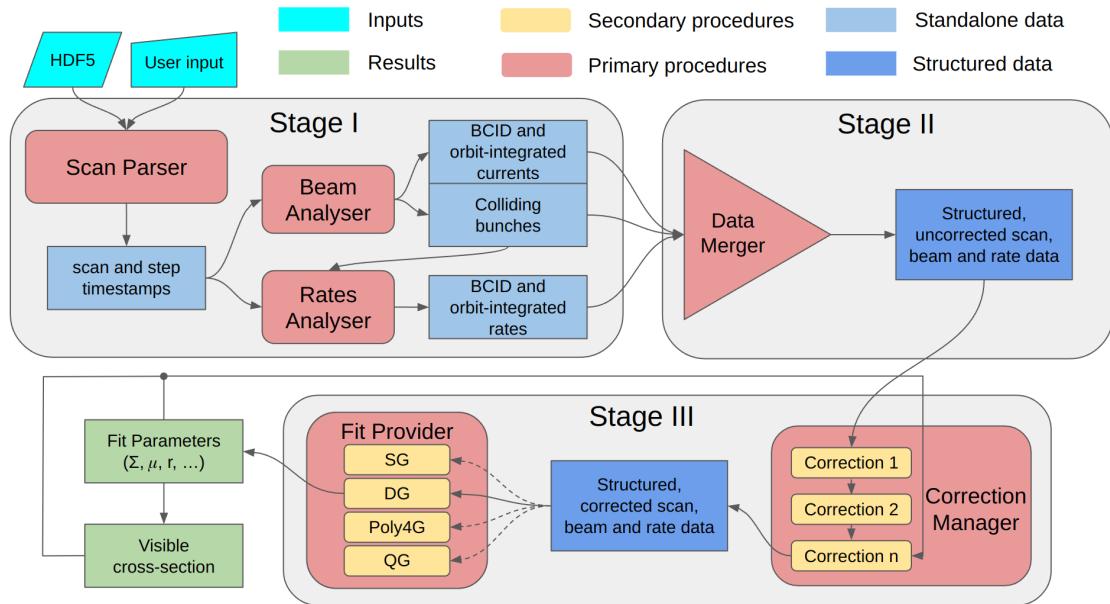


Figure 46: Overview of the flow of the analysis in the vdMFw.

The data is provided in Hierarchical Data Format 5 (HDF5) [55], which includes scan, luminometer, and beam information for a particular period of time:

- **Scan data:** Includes beam separation information, which is used to extract the beginning and end timestamps of each scan step.
- **Beam data:** Contains the beam currents, both per-orbit and per-bunch, saved in the beam table of the input file.
- **Luminometer data:** Stores the rates for every luminometer in operation during the time period associated with the input file.

The user can specify the following input parameters:

- Which luminometer data to analyse. The results of the analysis will correspond to the calibration constant for this luminometer.
- The fit function to be used for the bunch overlap shapes.

- The corrections to be applied. Some corrections are derived from the input data file, while others require external inputs from parallel analyses.
- Optionally, the user may provide a *ratefile*, also in HDF5 format, which contains only luminometer rates. This is useful when the original luminometer data is corrupted, non-existent, or when some offline correction has been applied to the original data.

Additionally, there is a user configuration file where further details can be adjusted.

The first stage of the analysis involves the preparation of the individual inputs that go into the fitting procedure, such as the beam separations and the rates normalized to the beam currents. The *Scan Parser* extracts the timestamps associated with each scan step and feeds them into the subsequent blocks. The *Beam Analyser* identifies the colliding bunches and computes the beam currents. At this stage, the beam current corrections explained in subsection 3.1.2 are applied. The colliding bunch indices are then passed to the *Rates Analyser*, where the luminometer rates are calculated.

The data processed in the previous stage is passed to Stage II. This stage consolidates all the pieces of information into a structured format to facilitate the application of any corrections and fits that follow.

Lastly, Stage III applies all the requested corrective procedures, and culminates in a fit to the corrected data, which provides the fit parameters used in calculating the detector calibration constants. While all corrections are done sequentially, some require outputs from previous corrections, hence the backward connector from the results to the beginning of this stage. The analysis concludes with all the results, both intermediate and final, being saved to files for later inspection.

4.1.2 Input/Output bottleneck

The vdMFw is executed multiple times throughout the year, particularly during emittance scans and vdM fills, where numerous corrections must be applied. Additionally, the process of adjusting the calibrations and ensuring detector stability is highly iterative, requiring the analysis to be run repeatedly until the results achieve the desired precision. Consequently, efforts have been made to enhance the software's performance.

To identify the primary bottlenecks within the framework, an initial profiling was conducted on input HDF5 files corresponding to vdM scans. These files were chosen because they undergo the complete correction pipeline, providing a comprehensive diagnosis of performance issues.

The profiling was performed on a system with the following specifications: x86_64 architecture, equipped with 32 CPUs¹, each with 2 threads per core, distributed across 8 cores per socket, and 2

¹ The Central Processing Unit (CPU) is the primary component of a computer that performs most of the processing inside a computer. It executes instructions

sockets in total. The CPUs operate at a base frequency of 1.0 GHz, with a maximum frequency of 3.5 GHz, supported by 64 GB of RAM and an L3 cache of 11 MB.

The results of the profiling are presented in Table 7, which highlights the time spent in each stage of the analysis process.

Table 7: Initial profiling results shown separately for the procedures in Stage I and combined for Stages II and III. The times correspond to an average of 50 runs where the full analysis pipeline was executed.

Procedure	Time [s]	% of total program
Scan Parser	0.34 ± 0.01	0.27
Beam Analyser	37.11 ± 0.65	29.60
Rate Analyser	57.85 ± 1.18	46.15
Stage II + III	30.06 ± 0.57	23.98

As shown in Table 7, the most significant bottlenecks were identified in the Beam Analyser and Rate Analyser procedures, which together account for over 75% of the total execution time. This result was unexpected, as the bulk of the framework’s workload was anticipated to revolve around data corrections and fitting procedures.

Upon closer inspection of the code within these two time-consuming procedures, the reason for their substantial execution time became apparent. To generate the expected outputs, these procedures need to read relevant data from the input HDF5 files. In the original implementation, the reading process occurred for every scan step, as follows:

1. Read the entire rate (or beam) table² from the HDF5 file into memory.
2. Filter for the timestamps corresponding to the current scan step.

Reading the entire table content each time is computationally expensive, as the amount of data to be read can consume a significant portion of the execution time. The profiling mentioned above was conducted using data files already filtered by the time period during which the scan was performed, resulting in smaller data files (approximately 23 MB). However, when the same profiling was run with an optional *ratefile* of 1631 MB in memory, the execution time of the Rate Analyser procedure increased dramatically from 57.85 seconds to 446.34 seconds, clearly indicating an Input/Output-related bottleneck.

To improve these execution times, the following optimizations were implemented:

from software programs by performing arithmetic, logic, control, and input/output operations.

² HDF5 files are organized in folder-like structures. The *tables* are analogous to files in this structure, and in BRIL’s case, each table contains specific data related to the LHC.

- The corresponding rate (or beam) table is now read only once before computing the per-step results.
- Instead of loading the entire table content into memory, which could lead to crashes if the data size exceeds the available RAM + SWAP³ space, a first pass is performed to gather all the entries that fall within any scan step.
- Only the relevant entries identified in the previous step are then read.

These changes led to the improved profiling results shown in Table 8:

Table 8: Profiling results after implementing the optimizations in the analyser procedures. The results show significant improvements in both procedures. The bulk of the framework (Stages II and III) now accounts for 95% of the total program execution time, and the overall program speedup was approximately 4 times.

Procedure	Time [s]	% of total program
Scan Parser	0.35 ± 0.02	1.08
Beam Analyser	0.90 ± 0.11	2.80
Rate Analyser	0.36 ± 0.02	1.11
Stage II + III	30.46 ± 0.54	95.01

When running the profiling on the same larger *ratefile* as before, the average execution time was reduced to 7.9 ± 0.22 seconds, representing a speedup of approximately 56 times compared to the original implementation.

4.1.3 Online vdM Analysis

An important aspect of analyzing LHC scans is the ability to process data immediately after a scan is completed, a process referred to as online analysis. Quick feedback from the framework allows for the possibility of repeating scans if any issues are identified during their execution. This capability is especially critical during vdM fills, motivating the enhancements made to this workflow.

The workflow, illustrated in Figure 47, begins with the Listener stage. This component constantly queries the xDAQ data acquisition system [56] for the *Active* variable, which is set to true while a scan is ongoing at the CMS IP. During these periods, two independent processes are at work:

³ SWAP space is a portion of the hard drive designated to act as virtual memory when the physical RAM is fully utilized. It helps prevent crashes by providing additional memory space, though it is much slower than actual RAM.

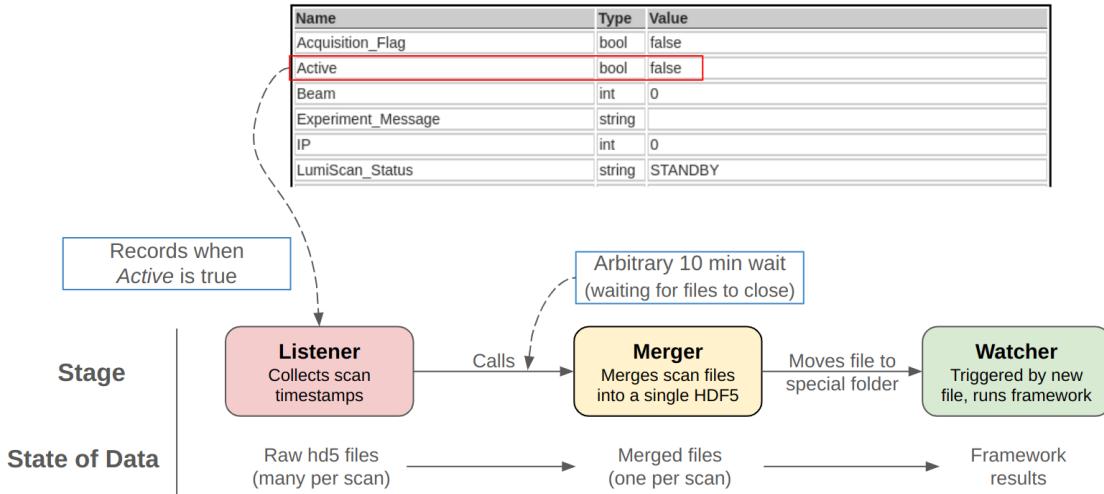


Figure 47: Workflow of the online analysis.

- **BRIL Data Acquisition System (BRILDAQ):** BRILDAQ, built on top of xDAQ, is the data acquisition and processing system used by all BRIL systems. It is configured to save extra HDF5 files to a specific destination while *Active* is set to true. These files are smaller and each contains data for a portion of the entire scan.
- **Scan Listener:** The Listener, a separate process also built using xDAQ, continuously monitors the *Active* variable and records the timestamps corresponding to the beginning and end of a scan (when *Active* changes from false to true and from true to false).

Once a scan is completed, the Listener passes the timestamp information to the Merger, which consolidates all files created within those timestamps. Since the two processes mentioned earlier operate independently, a safeguard in the form of a 10-minute wait time was implemented to protect against potential I/O errors. The output of the Merger is a single file containing data for the entire scan. This file is then moved to a designated folder monitored by the Watcher.

The Watcher is a filesystem event manager designed to trigger whenever a new file is created in a specific directory. Each trigger initiates the vdMFw to analyze that particular scan, applying multiple fits, corrections, and processing all online luminometers.

4.1.4 Improving Wait Time in Merger

The fixed waiting time in the Merger block is sub optimal. If the wait time is too short, it could lead to I/O problems, whereas if it is too long, it could delay the ability to repeat a scan, which is particularly critical during vdM fills.

To address this, an alternative approach was introduced in the Merger step that continuously polls the intermediate HDF5 files to check if they have been closed. This was implemented using the *lsof* Linux utility, which provides a convenient API to identify files that are open by a specific process. The implementation of this polling mechanism can be seen in Listing 6.1.

During the 2024 vdM program, which started on May 16th, the time spent waiting for the files to close for all vdM and im scans was significantly shorter than the fixed 10-minute interval, as shown in Table 9.

Table 9: Time spent waiting for scan files to close in the Merger during the 2024 vdM fill.

Scan	vdm1	vdm2	vdm3	vdm4	vdm5	im1	im2	vdm6
Time [s]	27.97	24.53	25.04	26.69	32.79	18.29	19.19	27.25

Since the wait time in the Merger and running the vdMFw were the most time consuming stages of this workflow, these improvements now allow for a feedback time of approximately 1 min after a particular scan has finished; a significant improvement over the 10+ minutes it used to take previously.

4.2 The BRIL Work Suite

The *brilcalc* tool, part of the BRIL work suite, plays a critical role in providing other CMS groups with access to the most accurate and up-to-date luminosity data. This section will first explain how *brilcalc* operates, including the necessary special input files, before discussing the enhancements made to its usability and maintenance. Finally, a highly requested alternative to *brilcalc* will be introduced, along with an analysis of its advantages and disadvantages.

4.2.1 Normtags and Iovtags

Given the high data demands from both BRIL and other CMS groups, *brilcalc* needs to efficiently retrieve luminosity results for any requested period. To meet this requirement, data stored in HDF5 files is loaded into the brilws database. *Brilcalc* relies on two specific input files for querying the database: a *normtag* and an *iovtag* file. The *normtag* file defines the LHC periods to be queried, while the *iovtag* file specifies how to apply corrections to the data from those periods. The structure of these files is illustrated in Figure 48.

To interpret these files:

- The first entry in the normtag specifies: For all lumisections in [72, 213] of run 366403, use the iovtag hfet23v09.

```

[ "hfet23v09", {"366403": [[72, 213]]}], name: hfet23v09
[ "hfet23v09", {"366406": [[1, 34]]}], applyto: lumi
[ "hfet23v09", {"366409": [[1, 35]]}], datasource: hfet
[ "hfet23v09", {"366410": [[1, 30]]}], since:
[ "hfet23v09", {"366413": [[1, 56]]}], - 366396:
[ "hfet23v09", {"366419": [[1, 31]]}], comments: A comment
[ "hfet23v09", {"366422": [[1, 30]]}], func: poly1d
[ "hfet23v09", {"366424": [[1, 29]]}], payload: {'coefs': '-0.0,3.318323758526599,0.0'}
[ "hfet23v09", {"366426": [[1, 31]]}], - 366448:
[ "hfet23v09", {"366427": [[1, 29]]}], comments: A comment
[ "hfet23v09", {"366429": [[1, 37]]}], func: poly1d
[ "hfet23v09", {"366432": [[1, 31]]}], payload: {'coefs': '-0.0,3.318323758526599,0.0'}
[ "hfet23v09", {"366433": [[1, 8]]}], - 366463:
[ "hfet23v09", {"366434": [[1, 5]]}], comments: A comment
[ "hfet23v09", {"366436": [[1, 45]]}], func: poly1d
[ "hfet23v09", {"366436": [[1, 45]]}], payload: {'coefs': '-0.0,3.3183287666912515,0.0'}

```

a) normtag

b) iovtag

Figure 48: a) Example of a Normtag file specifying LHC periods for queries in the brilws’ database. Each period is defined by a run number and a range of valid lumisections for that run. b) Example of an iovtag file corresponding to the HFET luminometer, specifying correction functions and payload arguments for each run number at the beginning of a fill.

- The first entry in the iovtag specifies: For all runs in [366396, 366447], correct HFET’s luminosity by applying the function *poly1d* with the parameters *{'coefs': '-0.0,3.318323758526599,0.0'}*.

The correction applied by the *poly1d* function corresponds to the procedure explained in subsection 2.1.3. The *coefs* field contains three comma-separated values that represent the coefficients of a second-degree polynomial in descending order. Since *brilcalc* operates on rate data from the HDF5 files, the coefficients are derived using:

$$\mathcal{L} = \underbrace{C_0}_{\text{0th order}} + \underbrace{\frac{f_{rev}}{\sigma_{vis} \cdot \epsilon}}_{\text{1st order}} \mu + \underbrace{-\alpha \left(\frac{f_{rev}}{\sigma_{vis} \cdot \epsilon} \right)^2}_{\text{2nd order}} \mu^2 \quad (4.1)$$

4.2.2 Limitations

Despite being the primary tool for obtaining the latest luminosity results, *brilcalc* was not originally designed for iterative analysis processes. Each time a new version of HDF5 data for a specific luminometer is created, it must be uploaded to the database. This process is not only time-consuming, as only authorized maintainers can update the database, but it also poses a risk of contaminating the database with incorrect data since uploads are not automatically validated for accuracy.

There are also issues with the construction of iovtags. As shown in Figure 48b, the *payload* parameter of iovtags is often not descriptive, requiring creators to add meaningful comments that help users interpret the values of σ_{vis} , ϵ , and α that generated the *coefs*.

Another significant problem is the lack of thorough validation for iovtags submitted to the database. The original system only checks the iovtag schema, not its content. Combined with minimal commenting, this can complicate the management of multiple iovtags across different systems and increase the difficulty of ensuring reproducibility of results.

4.2.3 Enhancements to iovtags

The software responsible for applying corrections specified by *iovtags* is built around a Python class containing various methods for implementing different types of luminosity corrections. For each correction, the specific method to be called is determined by the *func* parameter within the *iovtag*, and the method uses the corresponding *payload* to compute the corrected luminosity. Below is a simplified version of the existing code:

```

1 import numpy as np
2 class LuminosityCorrector:
3     def poly1d(self, database_input, payload):
4         coefs_array = np.fromstring(payload["coefs"], dtype=np.float, sep=',')
5
6         # Logic to fetch the rates and number of colliding bunches from the
7         # database
8         rates = ...
9         ncol = ...
10
11        # Divide by ncol to get the SBIL
12        return ncol * np.poly1d(rates / ncol, coefs_array[::-1])

```

The first enhancement to the iovtag workflow was to introduce a new set of payload parameters. A new function, *poly1d2*, where “2” denotes version 2, was implemented as follows:

```

1 def poly1d2(self, database_input, payload):
2     # Logic to fetch the rates and number of colliding bunches from the
3     # database
4     rates = ...
5     ncol = ...
6
7     lumi = (payload["frev"] * rates / ncol) / (payload["sigvis"] * payload["eff"])
8
9     return payload["constant"] + lumi - payload["alpha"] * lumi**2

```

This alternative implementation offers several advantages. Firstly, when using *poly1d2*, the payload values become transparent, allowing anyone with access to the *iovtag* to see the individual components rather than just the final coefficients from Equation 4.1. This is evident in the following example:

```
1 - 366396 :  
2     comments: A comment  
3     func: poly1d2  
4     payload:  
5         freq: 11245.5  
6         sigvis: 3394.33  
7         eff: 1.02  
8         slope: 0.0  
9         constant: 0.0
```

This new format simplifies comparisons between different versions of *iovtags* for the same system, as only the *comments* and *payload* values differ, whereas in the previous format, users could only speculate about the reasons behind differences in coefficients. Moreover, the payload arguments effectively document the *iovtag*, reducing reliance on comments, which can often be incorrect or outdated. This self-documentation makes *iovtags* more robust, embedding essential details directly within the parameters themselves. Additionally, including the calculation components directly within the *iovtag* centralizes the code, reducing the risk of errors when applying the formula described in Equation 4.1.

Although this new approach improves the readability of *iovtags*, it does not fully address the lack of stable, up-to-date documentation for the parameters of each method within the corrector class, nor does it validate *iovtags* upon upload. To address both issues while maintaining tight coupling with the source code, Python decorators were employed [57].

A decorator named *register_sanitizer* was used on all correction methods to document each *payload* parameter. An example of decorating the *poly1d2* method is shown below:

```
1 @register_sanitizer(  
2     required_params = (  
3         ("freq", "LHC's revolution frequency."),  
4         ("sigvis", "Luminometer visible cross-section.")),
```

```

5     ("eff", "Efficiency factor applied to 'sigvis.'"),
6     ("alpha", "Slope extracted from a linear fit to sbil_1 / sbil_2 = a *
7         sbil_2 + b."),
8     ("constant", "0th degree coefficient of the 2nd-degree polynomial."),
9 )
10 def poly1d2(self, database_input, payload):
11     ...

```

This decorator ensures that the `poly1d2` function checks for the presence of all required parameters in the `iovtag` before executing. For methods requiring constraints on parameters, the decorator also accepts a `validators` parameter, as shown below:

```

1 @register_sanitizer(
2     required_params=(
3         ("frev", "LHC's revolution frequency."),
4         ("sigvis", "Luminometer visible cross-section."),
5         ("eff", "Efficiency factor applied to 'sigvis.'"),
6         ("alpha", "Slope extracted from a linear fit to sbil_1 / sbil_2 = a *
7             sbil_2 + b."),
8         ("intercept", "Intercept extracted from the linear fit."),
9         ("constant", "0th degree coefficient of the 2nd-degree polynomial."),
10    ),
11     validators=(
12         (lambda payload: not np.isclose(payload["alpha"], 0.0), "'alpha' "
13          parameter cannot be 0."),
14    )
15 )
16 def another_method(self, database_input, payload):
17     ...

```

In the implementation of this decorator (see Listing 6.2), the attributes `_required_params` and `_validators` are added to the function objects to differentiate them from other class attributes. A class decorator named `collect_sanitizers` is used as a metaclass [58], which gathers all registered sanitizers into class attributes upon loading, as illustrated in Listing 6.3.

This setup allows the sanitizers to be used, not only on already loaded `iovtags`, but also on those

requested for upload. The `validate_ivtag_data` function, shown in Listing 6.5, implements the validation logic and is called each time a new ivtag is to be uploaded.

All these modifications allow the users of *brilws* to be more certain of the contents of the *ivtags* they use as well as preventing polluting the database with erroneous correctors.

4.2.4 An Alternative to *brilcalc*

One of the most technically challenging aspects of luminosity analysis is managing the different versions of data. Any experimental effect not addressed during operations must be corrected in offline analysis, a process commonly referred to as reprocessing. Each reprocessing step takes as input a specific version of the HDF5 data for a given system, or set of systems, and applies corrections for one or more systematic effects, resulting in a new set of HDF5 files with the added corrections.

Ideally, each new set of HDF5 files would be thoroughly checked to ensure that the corrections were properly applied. However, verifying the results required that the files first be uploaded to the *brilws* database, which not only risked contaminating the database with incorrect data but also unnecessarily increased the number of data versions stored. Consequently, there was a high demand for an alternative tool capable of producing the same output as *brilcalc* directly from HDF5 files without the need for database integration.

The proposed software solution, while separate from *brilws*, was designed to be compatible with the *brilws* workflow. It requires users to provide a *normtag* file and specify the location of the HDF5 files to be analyzed. The *ivtags* referenced in the *normtag* must exist on disk but do not need to be loaded into any database. The final output of this tool matches the format produced by *brilcalc*, enabling rigorous testing of any reprocessing results before deciding to share the data with the rest of the CMS.

Although both tools produce the same output, each has distinct advantages. The alternative tool provides absolute certainty regarding the version of the data being analyzed since users must specify the exact path to the data files. However, this alternative is significantly slower than the *brilws* implementation, because reading data directly from disk is much slower than querying a database. Despite this slower speed, it is still fast enough to allow multiple iterations of analysis. This made it particularly useful for 2023, where the stability and non-linearity corrections for the BCM1F system benefited greatly from this approach, as multiple reprocessing steps were required.

In conclusion, while this alternative tool proves invaluable for conducting detailed analyses, *brilcalc* remains the preferred option for production-level results, offering a more streamlined and faster approach once the data is finalized.

Chapter 5

Conclusions and future work

5.1 Conclusion and Future Work

This thesis has presented the comprehensive calibration and measurement of luminosity for the 2023 data-taking period of the CMS experiment at the LHC. The luminosity calibration was conducted using the van der Meer (vdM) scan methodology, which included meticulous corrections for several systematic effects. Among these, the beam-beam interaction, XY factorization bias, and scan-to-scan vdM variation were identified as the most significant sources of uncertainty. Despite these challenges, the final calibration achieved an impressive consistency across all independently calibrated luminometers within a 0.16% uncertainty during the head-on period of the vdM fill.

The extrapolation of the vdM calibration to physics conditions required additional stability and linearity corrections and detector specific corrections like afterglow and out-of-time pileup. These corrections significantly reduced the stability uncertainty from an initial 2.6% to 0.71%, with residual non-linearity contributing an additional 0.51% uncertainty. As a result, the total delivered integrated luminosity for the 2023 data-taking year was measured to be 32.74 fb^{-1} , with an associated overall uncertainty of 1.28%. This represents the best preliminary luminosity results ever achieved by CMS, marking a significant milestone in the precision of luminosity measurements.

A key part of my contributions focused on enhancing the performance of the vdM scan data analysis conducted by the VdMFramework. After identifying a critical input/output (I/O) bottleneck, which accounted for approximately 75% of the total execution time, the implementation of caching techniques led to a fourfold increase in efficiency, significantly speeding up the analysis. Further improvements were made to the automatic online scan analysis workflow with the introduction of a file polling mechanism, which reduced the dead time by a factor of 20, enabling the entire scan analysis pipeline to complete in approximately one minute.

In addition to these performance enhancements, substantial contributions were made to the BRIL

Work Suite, particularly in improving the usability and maintainability of the *iovtag* structure. By making the payload parameters more transparent and integral to the calibration process, the updates allowed for a more intuitive and error-resistant workflow. Moreover, an alternative tool was developed to work alongside the BRIL Work Suite, providing a means to circumvent lengthy database upload times. This tool allowed for quicker iterations during the analysis process and enabled cross-verification of the data by comparing database content with raw HDF5 files.

While significant progress has been made, there remains much work to be done. The largest contributions to the overall luminosity uncertainty stem from stability and linearity corrections, marking them as the primary focus for future improvements. On the software side, optimizing the alternative to *brilcalc* promises the most immediate gains, as it will streamline the analysis workflow further. Additionally, optimization efforts will continue for Stages II and III of the VdMFramework, where data corrections and fits are applied. Ultimately, these efforts aim to reduce the overall uncertainty of the 2023 luminosity measurement, bringing it closer to the long-term goal of reaching the challenging 1% uncertainty target.

In conclusion, this work has not only advanced the precision of luminosity measurements at CMS but also set a new benchmark as the best preliminary luminosity results achieved by the experiment. By continuing to refine both the calibration techniques and the supporting software infrastructure, the goal of achieving a more precise and accurate luminosity measurement is ever closer.

Bibliography

- [1] L. Evans and P. Bryant, “LHC Machine,” *Journal of Instrumentation*, vol. 3, no. 08, S08001–S08001, Aug. 2008, ISSN: 1748-0221. DOI: 10.1088/1748-0221/3/08/S08001. [Online]. Available: <https://iopscience.iop.org/article/10.1088/1748-0221/3/08/S08001> (visited on 11/27/2024).
- [2] CMS Collaboration, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC,” en, *Physics Letters B*, vol. 716, no. 1, pp. 30–61, Sep. 2012, ISSN: 03702693. DOI: 10.1016/j.physletb.2012.08.021. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0370269312008581> (visited on 11/27/2024).
- [3] *LHC long term schedule*. [Online]. Available: <https://lhc-commissioning.web.cern.ch/schedule/LHC-long-term.htm> (visited on 11/27/2024).
- [4] *The accelerator complex / CERN*. [Online]. Available: <https://home.cern/science/accelerators/accelerator-complex> (visited on 11/27/2024).
- [5] CERN Yellow Reports: Monographs, *CERN Yellow Reports: Monographs*, Vol. 6 (2020): *Linac4 design report*, en, Sep. 2020. DOI: 10.23731/CYRM-2020-006. [Online]. Available: <https://epublishing.cern.ch/index.php/CYRM/issue/view/121> (visited on 11/27/2024).
- [6] CMS Collaboration, “The CMS experiment at the CERN LHC,” *Journal of Instrumentation*, vol. 3, no. 08, S08004–S08004, Aug. 2008, ISSN: 1748-0221. DOI: 10.1088/1748-0221/3/08/S08004. [Online]. Available: <https://iopscience.iop.org/article/10.1088/1748-0221/3/08/S08004> (visited on 11/27/2024).
- [7] CERN AC, *Layout of CMS*, 1998. [Online]. Available: <https://cds.cern.ch/record/39040> (visited on 11/27/2024).
- [8] D. Barney, *CMS Detector Slice*, 2016. [Online]. Available: <https://cds.cern.ch/record/2120661> (visited on 11/27/2024).
- [9] M. Hostettler *et al.*, “Luminosity scans for beam diagnostics,” en, *Physical Review Accelerators and Beams*, vol. 21, no. 10, p. 102 801, Oct. 2018, ISSN: 2469-9888. DOI: 10.1103/PhysRevAccelBeams.21.102801. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevAccelBeams.21.102801> (visited on 11/27/2024).
- [10] CMS Collaboration, “Measurement of differential cross sections for the production of a Z boson in association with jets in proton-proton collisions at $\sqrt{s} = 13$ TeV,” en, *Physical Review D*, vol. 108, no. 5, p. 052 004, Sep. 2023, ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.108.052004. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevD.108.052004> (visited on 11/27/2024).

- [11] CMS Collaboration, “Measurement of the $t\bar{t}$ production cross section, the top quark mass, and the strong coupling constant using dilepton events in pp collisions at $\sqrt{s} = 13$ TeV,” en, *The European Physical Journal C*, vol. 79, no. 5, p. 368, May 2019, ISSN: 1434-6044, 1434-6052. DOI: 10.1140/epjc/s10052-019-6863-8. [Online]. Available: <http://link.springer.com/10.1140/epjc/s10052-019-6863-8> (visited on 11/27/2024).
- [12] H. Burkhardt and P. Grafström, *Absolute Luminosity from Machine Parameters*, 2007. [Online]. Available: <https://cds.cern.ch/record/1056691> (visited on 11/27/2024).
- [13] M. Furman, “The Moeller luminosity factor,” en, Tech. Rep. LBNL–53553, CBP Note - 543, 836235, Aug. 2003, LBNL–53553, CBP Note –543, 836235. DOI: 10.2172/836235. [Online]. Available: <http://www.osti.gov/servlets/purl/836235-38uG55/native/> (visited on 11/27/2024).
- [14] K. Schindl *et al.*, *LHC Design Report*, en, 2004. DOI: 10.5170/CERN-2004-003-V-3. [Online]. Available: <http://cds.cern.ch/record/823808> (visited on 11/27/2024).
- [15] LHCb Collaboration, “Precision luminosity measurements at LHCb,” *Journal of Instrumentation*, vol. 9, no. 12, P12005–P12005, Dec. 2014, ISSN: 1748-0221. DOI: 10.1088/1748-0221/9/12/P12005. [Online]. Available: <https://iopscience.iop.org/article/10.1088/1748-0221/9/12/P12005> (visited on 11/27/2024).
- [16] CMS Collaboration, “Precision luminosity measurement in proton–proton collisions at $\sqrt{s} = 13$ TeV in 2015 and 2016 at CMS,” en, *The European Physical Journal C*, vol. 81, no. 9, p. 800, Sep. 2021, ISSN: 1434-6044, 1434-6052. DOI: 10.1140/epjc/s10052-021-09538-2. [Online]. Available: <https://link.springer.com/10.1140/epjc/s10052-021-09538-2> (visited on 11/27/2024).
- [17] S. Saariokari, “Calibration of luminometers at Compact Muon Solenoid experiment. Luminometrien kalibointi Compact Muon Solenoid -kokeessa,” Ph.D. dissertation, Tampere U., 2022. [Online]. Available: <https://cds.cern.ch/record/2826125> (visited on 11/27/2024).
- [18] J. Mans *et al.*, “CMS Technical Design Report for the Phase 1 Upgrade of the Hadron Calorimeter,” Tech. Rep., 2012. [Online]. Available: <https://cds.cern.ch/record/1481837> (visited on 11/27/2024).
- [19] CMS BRIL Collaboration, “The Pixel Luminosity Telescope: A detector for luminosity measurement at CMS using silicon pixel sensors,” en, *The European Physical Journal C*, vol. 83, no. 7, p. 673, Jul. 2023, ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-023-11713-6. [Online]. Available: <https://link.springer.com/10.1140/epjc/s10052-023-11713-6> (visited on 11/27/2024).
- [20] F. Romeo, *Performance of the Pixel Luminosity Telescope for Luminosity Measurement at CMS during Run2*, 2019. [Online]. Available: <https://cds.cern.ch/record/2797807> (visited on 11/27/2024).
- [21] A. G. Delannoy Sotomayor and J. Wanczyk, *BRIL LS2 Activity: PLT & BCM1f Integration*, 2021. [Online]. Available: <https://cds.cern.ch/record/2765247> (visited on 11/27/2024).
- [22] H. Kästli *et al.*, “Design and performance of the CMS pixel detector readout chip,” en, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 565, no. 1, pp. 188–194, Sep. 2006, ISSN: 01689002. DOI: 10.1016/j.nima.2006.05.038. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0168900206007674> (visited on 11/27/2024).

- [23] P. J. Lujan, *The Pixel Luminosity Telescope: A silicon sensor detector for luminosity measurement at CMS*, 2022. [Online]. Available: <https://cds.cern.ch/record/2797692> (visited on 11/27/2024).
- [24] J. Wańczyk, “Upgraded CMS Fast Beam Condition Monitor for LHC Run 3 Online Luminosity and Beam Induced Background Measurements,” en, 5 pages, 8.278 MB, Dec. 2022, ISSN: 2673-5350. DOI: 10.18429/JACOW-IBIC2022-TH2C2. [Online]. Available: <https://jacow.org/ibic2022/doi/JACoW-IBIC2022-TH2C2.html> (visited on 11/27/2024).
- [25] A. Zagódzinska *et al.*, “New Fast Beam Conditions Monitoring (BCM1F) system for CMS,” *Journal of Instrumentation*, vol. 11, no. 01, pp. C01088–C01088, Jan. 2016, ISSN: 1748-0221. DOI: 10.1088/1748-0221/11/01/C01088. [Online]. Available: <https://iopscience.iop.org/article/10.1088/1748-0221/11/01/C01088> (visited on 11/27/2024).
- [26] A. G. Delannoy Sotomayor and J. Wanczyk, *CMS BRIL BCM1F c-shape*, 2022. [Online]. Available: <https://cds.cern.ch/record/2809025> (visited on 11/27/2024).
- [27] O. Karacheban, “Luminosity measurement at CMS,” Ph.D. dissertation, DESY, 2017. [Online]. Available: <https://cds.cern.ch/record/2294183> (visited on 11/27/2024).
- [28] The Tracker Group of the CMS Collaboration, *The CMS Phase-1 Pixel Detector Upgrade*, 2020. [Online]. Available: <https://cds.cern.ch/record/2745805> (visited on 11/27/2024).
- [29] A. Triossi *et al.*, “The CMS Barrel Muon trigger upgrade,” *Journal of Instrumentation*, vol. 12, no. 01, pp. C01095–C01095, Jan. 2017, ISSN: 1748-0221. DOI: 10.1088/1748-0221/12/01/C01095. [Online]. Available: <https://iopscience.iop.org/article/10.1088/1748-0221/12/01/C01095> (visited on 11/27/2024).
- [30] P. Grafström and W. Kozanecki, “Luminosity determination at proton colliders,” en, *Progress in Particle and Nuclear Physics*, vol. 81, pp. 97–148, Mar. 2015, ISSN: 01466410. DOI: 10.1016/j.ppnp.2014.11.002. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0146641014000878> (visited on 11/27/2024).
- [31] J. Salfeld-Nebgen and D. Marlow, “Data-driven precision luminosity measurements with Z bosons at the LHC and HL-LHC,” *Journal of Instrumentation*, vol. 13, no. 12, P12016–P12016, Dec. 2018, ISSN: 1748-0221. DOI: 10.1088/1748-0221/13/12/P12016. [Online]. Available: <https://iopscience.iop.org/article/10.1088/1748-0221/13/12/P12016> (visited on 11/27/2024).
- [32] S. van der Meer, *Calibration of the effective beam height in the ISR*, Geneva, 1968. [Online]. Available: <https://cds.cern.ch/record/296752> (visited on 11/27/2024).
- [33] C. Rubbia, *Measurement of the luminosity of p-pbar collider with a (generalized) Van der Meer Method*, Geneva, 1977. [Online]. Available: <https://cds.cern.ch/record/1025746> (visited on 11/27/2024).
- [34] S. Acharya *et al.*, “ALICE luminosity determination for Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02 \text{ TeV}$,” *Journal of Instrumentation*, vol. 19, no. 02, P02039, Feb. 2024, ISSN: 1748-0221. DOI: 10.1088/1748-0221/19/02/P02039. [Online]. Available: <https://iopscience.iop.org/article/10.1088/1748-0221/19/02/P02039> (visited on 11/27/2024).
- [35] S. Maettig, “Luminosity Measurements with the ATLAS Detector,” Ph.D. dissertation, Hamburg U., 2012. [Online]. Available: <https://cds.cern.ch/record/1513982> (visited on 11/27/2024).

- [36] A. Babaev *et al.*, “Impact of beam-beam effects on absolute luminosity calibrations at the CERN Large Hadron Collider,” en, *The European Physical Journal C*, vol. 84, no. 1, p. 17, Jan. 2024, ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-023-12192-5. [Online]. Available: <https://link.springer.com/10.1140/epjc/s10052-023-12192-5> (visited on 11/27/2024).
- [37] T. Persson *et al.*, *LHC Optics Corrections in Run 2*, 2019. [Online]. Available: <https://cds.cern.ch/record/2750277> (visited on 11/27/2024).
- [38] M. Klute, C. Medlock, and J. Salfeld-Nebgen, “Beam imaging and luminosity calibration,” en, *arXiv*, 2017. [Online]. Available: <https://dspace.mit.edu/handle/1721.1/134474> (visited on 11/27/2024).
- [39] D. Belohrad *et al.*, *The LHC Fast BCT system: A comparison of Design Parameters with Initial Performance*, Geneva, 2010. [Online]. Available: <https://cds.cern.ch/record/1267400> (visited on 11/27/2024).
- [40] P. Odier, M. Ludwig, and S. Thoulet, *The DCCT for the LHC Beam Intensity Measurement*, Geneva, 2009. [Online]. Available: <https://cds.cern.ch/record/1183400> (visited on 11/27/2024).
- [41] C. Barschel, *Results of the LHC DCCT Calibration Studies*, 2012. [Online]. Available: <https://cds.cern.ch/record/2649533> (visited on 11/27/2024).
- [42] J. J. Gras *et al.*, “Optimization of the LHC beam current transformers for accurate luminosity determination,” 2011. [Online]. Available: <https://cds.cern.ch/record/1379466> (visited on 11/27/2024).
- [43] A. Alici *et al.*, *Study of the LHC ghost charge and satellite bunches for luminosity calibration*. 2012. [Online]. Available: <https://cds.cern.ch/record/1427728> (visited on 11/27/2024).
- [44] *Luminosity measurement in proton-proton collisions at 13.6 TeV in 2022 at CMS*, 2024. [Online]. Available: <https://cds.cern.ch/record/2890833> (visited on 11/27/2024).
- [45] *Measurement of the offline integrated luminosity for the CMS proton-proton collision dataset recorded in 2023, 2024*. [Online]. Available: <https://cds.cern.ch/record/2904808> (visited on 11/27/2024).
- [46] M. Bassetti and G. A. Erskine, *Closed expression for the electrical field of a two-dimensional Gaussian charge*, Geneva, 1980. [Online]. Available: <https://cds.cern.ch/record/122227> (visited on 11/27/2024).
- [47] P. Bambade *et al.*, “Observation of beam-beam deflections at the interaction point of the SLAC Linear Collider,” en, *Physical Review Letters*, vol. 62, no. 25, pp. 2949–2952, Jun. 1989, ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.62.2949. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.62.2949> (visited on 11/27/2024).
- [48] A. Babaev, *Coherent deflection of elliptic bunches colliding at crossing angle*, 2021. DOI: 10.48550/ARXIV.2104.02595. [Online]. Available: <https://arxiv.org/abs/2104.02595> (visited on 11/27/2024).
- [49] M. Gąsior *et al.*, *First Operational Experience with the LHC Diode ORbit and OScillation (DOROS) System*, 2017. [Online]. Available: <https://cds.cern.ch/record/2313935> (visited on 11/27/2024).
- [50] J. Knolle, *The factorization bias in the Van der Meer method: Run 2 experiences at the CMS experiment*, 2019. DOI: 10.48550/ARXIV.1910.14066. [Online]. Available: <https://arxiv.org/abs/1910.14066> (visited on 11/27/2024).

- [51] J. Knolle, “Measuring luminosity and the $\mathit{\bar{Z}}$ production cross section with the CMS experiment,” Ph.D. dissertation, Hamburg U., 2020. [Online]. Available: <https://cds.cern.ch/record/2744549> (visited on 11/27/2024).
- [52] CMS Collaboration, *CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV*, 2019. [Online]. Available: <https://cds.cern.ch/record/2676164> (visited on 11/27/2024).
- [53] BRIL Collaboration, *VdM Framework*. [Online]. Available: <https://gitlab.cern.ch/bril/VdMFramework>.
- [54] Z. Xie, *BRIL Work Suite*. [Online]. Available: <https://cms-service-lumi.web.cern.ch/cms-service-lumi;brilwsdoc.html>.
- [55] S. Koranne, “Hierarchical Data Format 5 : HDF5,” en, Boston, MA: Springer US, 2011, pp. 191–200, ISBN: 9781441977182 9781441977199. DOI: 10.1007/978-1-4419-7719-9_10. [Online]. Available: https://link.springer.com/10.1007/978-1-4419-7719-9_10 (visited on 11/27/2024).
- [56] V. Brigljevic *et al.*, “Using XDAQ in Application Scenarios of the CMS Experiment. COMPACT MUON SOLENOID,” *eConf C*, vol. 0303241, MOGT008, 2003. [Online]. Available: <https://cds.cern.ch/record/845273> (visited on 11/27/2024).
- [57] *PEP 318 – Decorators for Functions and Methods* / [peps.python.org](https://peps.python.org/pep-0318/), en. [Online]. Available: <https://peps.python.org/pep-0318/> (visited on 11/27/2024).
- [58] *PEP 3115 – Metaclasses in Python 3000* / [peps.python.org](https://peps.python.org/pep-3115/), en. [Online]. Available: <https://peps.python.org/pep-3115/> (visited on 11/27/2024).

Chapter 6

Appendices

6.1 Listings

```
1 def wait_until_files_closed(files: list[str], wait: float) -> bool:
2     """Blocks until all the provided files are closed or until the provided
3         timeout ends.
4
5     Parameters
6     -----
7     files: List[str]
8         The list of files to watch.
9     wait: float
10        The maximum time allowed to block in seconds.
11
12    Returns
13    -----
14    bool:
15        True if exited after timeout expired. False otherwise.
16    """
17    timeout_seconds = wait
18    start_time = time.time()
19
20    while time.time() - start_time < timeout_seconds:
21        all_closed = True
22        for file_path in files:
23            # Use grep to check if the specific file is still open
24            result = subprocess.run(['lsof', '|', 'grep', file_path],
25                                   stdout=subprocess.PIPE, stderr=subprocess.PIPE, text=True)
26            if result.stdout:
27                logging.debug(f"The file {file_path} is still open.")
28                all_closed = False
29                break # If any file is still open, break and retry after a
30                      # pause
31            else:
32                logging.debug(f"The file {file_path} is now closed.")
33
34        if all_closed:
35            return True
36
37        time.sleep(1) # Wait for 1 second before trying again
38
39    return False
```

Listing 6.1: Polling implementation in Merger. Once all the files to be merged are full the function stops blocking.

```
1 def register_sanitizer(required_params, validators=None):
2     def decorator(func):
3         func._required_params = required_params
4         func._validators = validators if validators else []
5         @wraps(func)
6         def wrapper(instance, arg1, kwds):
7             cls = instance.__class__
8             cls.sanitize_arguments(kwds, func.__name__)
9             return func(instance, arg1, kwds)
10        return wrapper
11    return decorator
```

Listing 6.2: Decorator function that registers the sanitizers associated with a specific corrector function.

```
1 def collect_sanitizers(cls):
2     cls.required_params = {}
3     cls.validators = {}
4     for name, method in inspect.getmembers(cls, predicate=inspect.
5         isfunction):
5         if hasattr(method, '_required_params') and hasattr(method, '_validators'):
6             cls.required_params[name] = method._required_params
7             cls.validators[name] = method._validators
8     return cls
```

Listing 6.3: Class decorator that aggregates all registered sanitizers into class attributes for subsequent validation.

```
1 def validate_sanitizers(params, func_name):
2     if not hasattr(LuminosityCorrector, func_name):
3         raise ValueError(f"Function '{func_name}' does not exist in class
4             '{LuminosityCorrector.__name__}'.")
5     if func_name not in LuminosityCorrector.required_params:
6         raise ValueError(f"Function '{func_name}' cannot be used for
7             corrections.")
8
9     # At this stage params will only be None if The iovtag
10    # provided no arguments under 'payload' This may be intended,
11    # if the func_name actually does not receive arguments.
12    # Otherwise it will be an improper iovtag.
13    if params is None:
14        params = {}
15
16    param_desc_list = LuminosityCorrector.required_params[func_name]
17    missing_params = []
18    for param, desc in param_desc_list:
19        if param not in params:
20            missing_params.append(f"- {param}: {desc}")
21    if missing_params:
22        err_str = (
23            f"The configuration for '{func_name}' is missing required
24                arguments:\n"
25            + "\n".join(missing_params)
26        )
27        raise ValueError(err_str)
```

```

25
26     validators_list = LuminosityCorrector.validators[func_name]
27     validator_errs = []
28     for validator, err_msg in validators_list:
29         if not validator(params):
30             validator_errs.append(f"- {err_msg}")
31     if validator_errs:
32         err_str = (
33             f"The configuration for '{func_name}' has invalid parameters:\n"
34             ""
35         ) + "\n".join(validator_errs)
36     raise ValueError(err_str)

```

Listing 6.4: Function implementing the validation logic. It uses the information set by *register_sanitizer* and collected by *collect_sanitizers*.

```

1 def validate_iovdata(iovdata):
2     for entry in iovdata:
3         for run, data in entry.items():
4             func_name = data["func"]
5             payload = data["payload"]
6             try:
7                 validate_required_arguments(payload, func_name)
8             except ValueError as err:
9                 raise ValueError(f"Invalid config at run {run}.") from err

```

Listing 6.5: Function to validate *iovtags* before uploading them into the database, preventing contamination.

