Z boson cross section measurements and neutral Higgs boson search in the di-muon channel with the CMS detector

by

Xinmei Niu

B.S. in Physics, University of Science and Technology of China (2013)

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2019

Signature redacted

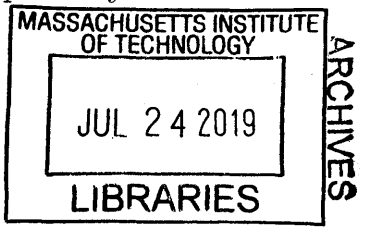
Author .....
Department of Physics
May 24, 2019

Signature redacted

Certified by .....
Markus Klute
Associate Professor of Physics
Thesis Supervisor

Signature redacted

Accepted by .....
Nergis Mavalvala
Associate Department Head of Physics



Z boson cross section measurements and neutral Higgs boson search in the di-muon channel with the CMS detector

by

Xinmei Niu

Submitted to the Department of Physics
on May 24, 2019, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

In this thesis, measurements of inclusive and differential Z boson production cross sections in proton-proton collisions at $\sqrt{s}$ =13 TeV with the CMS detector at Large Hadron Collider are performed with the di-muon channel. The measured total inclusive cross section times branching ratio is $\sigma(pp \to ZX) \times \mathcal{B}(Z \to \mu\mu) = 1870 \pm 2(stat) \pm 35(syst) \pm 51(lumi)$ pb for the di-muon invariant mass in the range of 60 to 120 GeV, which is in good agreement with the next-to-next-to leading order QCD predictions. The spectra of Z boson transverse momentum, $\phi^{\star}_{\eta}$ variable, rapidity, and the muon transverse momentum are also measured and compared with theoretical predictions. The large production cross section of Z boson and good experimental accessibility of final-state muons permit a real-time monitoring of luminosities using the counts of reconstructed $Z \to \mu\mu$ events. Preliminary results of counting of Z bosons as a luminometer are shown using the entire 2018 data-taking period. A search for beyond Standard Model neutral Higgs bosons decaying to two muons is also presented. No significant excess is observed. A 95% confidence level upper limit is set on $\sigma(pp \to \phi) \times \mathcal{B}(\phi \to \mu\mu)$ for a generic boson ϕ. The exclusion contour is also determined on the parameter phase space in the context of Minimal Supersymmetric Standard Mode representative benchmark scenarios.

Thesis Supervisor: Markus Klute
Title: Associate Professor of Physics

Acknowledgments

First and foremost, I would like to express my sincere gratitude and deepest admiration to my supervisor Prof. Markus Klute, not only for his continuous support and guidance of every research project and writing of this thesis, but also for his patience and encouragement which keep me motivated in the past six years.

Besides my supervisor, my sincere thanks also goes to the rest of my thesis committee, Prof. Philip Harris and Prof. Boleslaw Wyslouch, for their insightful comments on my thesis.

I would also like to thank Dr. Guillelmo Gomez-Ceballos, Dr. Andrea Marini, Dr. Mariarosari D'Alfonso, Dr. Jakob Salfeld-Nebgen and Dr. Aram Apyan, for sharing his/her expertise and proving precious advice during our close collaboration on research projects, as well as for being a caring and invaluable friend to me for all those years.

I thank my fellow office mates Stephanie Brandt and Miao Hu, for the days and nights we worked together and for all the fun we had together.

I want to thank all my friends and those more than friends that I met in Boston and Geneva, as well as in Handan and Hefei, for always being there, via Wechat or in person, especially in my sentimental moments (well there were a lot, I know :)).

Last of all, I would like to thank my parents. They have no idea what a boson is, but they offer me 100% freedom, 100% trust, and 100% spiritual support to do what I want since school age. Without them, I would never have had a chance to come overseas, earn a doctorate degree in physics, and become who I am now. Mom and Dad, I love you.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

On December 3rd 2018, The Large Hadron Collider (LHC) [26] successfully completed the second data-taking period that began in 2015. About 160 fb^{-1} of proton-proton collision data are taken at the center-of-mass energy $\sqrt{s} = 13$ TeV. There were another $\sim$30 fb^{-1} of data taken at $\sqrt{s} = 7$ and 8 TeV during the first period. Those data are devoted to test the Standard Model (SM) of particle physics and search for new physics beyond the SM (BSM) at the new energy regime.

The number of SM measurements performed at LHC is huge, covering more than ten orders of magnitude in the production cross sections and probing many different aspects in physics. A summary of the SM cross section measurement results obtained by the ATLAS and CMS collaboration are illustrated in Figure 1-1. Among those measurements, the inclusive jet measurements [27, 28] probe the energy and distance frontier, where the jet transverse momentum reaches up to 2 TeV and corresponds to a distance scale of 10^{-19} m. While the inclusive vector boson production measurements outperform the jet studies in terms of the precision reached, and provide the most precisely determined observables at hadron colliders. The high precision measurements contribute testing grounds for the high-order theoretical calculations and scrutinize their approximations and implementations in simulation tools. With this motivation, the precision measurements of the inclusive and differential Z boson production cross sections in the proton-proton collisions are performed using the di-muon final states, and the first topic of this thesis is dedicated to discuss these

measurements.

In addition to the physics motivation, as a standard candle, the Z boson productions also practically serve as detector calibration purpose. During the second period of data-taking, the CMS collaboration realized a complementary luminometer using the Z boson counts in defined time intervals with automated workflows. This work is also documented as the second subjects in this thesis.

The discovery of a scalar boson with a mass of 125.09 ± 0.24 GeV [29–31] completes the SM. However, the SM is certainly not a "theory of everything" as it does not explain phenomena such as the hierarchy problem [32] and evidence for the dark matter [33–35]. This has led to the development of many BSM theories that extend the Higgs sector and predict more physical Higgs bosons with potentially dramatically different properties from the SM Higgs boson. The Higgs boson decays to a pair of muons could be enhanced at specific parameter phase space and provide sensitivity for observation of the Yukawa coupling [36] to second generation fermions, which is known to be very unlikely with existing dataset in the case of SM Higgs [37]. To that end, the search for heavy neutral Higgs bosons in the di-muon channel are conducted and reported as the third topic of this thesis.

The seven chapters of this thesis is organized as follows. After the introduction, Chapter 2 reviews the fundamental theories behind the Z boson production and the BSM Higgs boson production. The experimental apparatus of the accelerators and detectors are introduced in Chapter 3 in a general manner, while the procedures to reconstruct high-level physics objects from raw detector readout are described in Chapter 4. The subsequent three chapters serve the three main subjects of this thesis: Chapter 5 for the Z boson production cross section measurements, Chapter 6 for the Z boson counting as a luminosity monitor, and Chapter 7 for the search of neutral BSM Higgs bosons. Each chapter delineates the motivation, how the measurements (searches) are performed, and the corresponding results and interpretation.

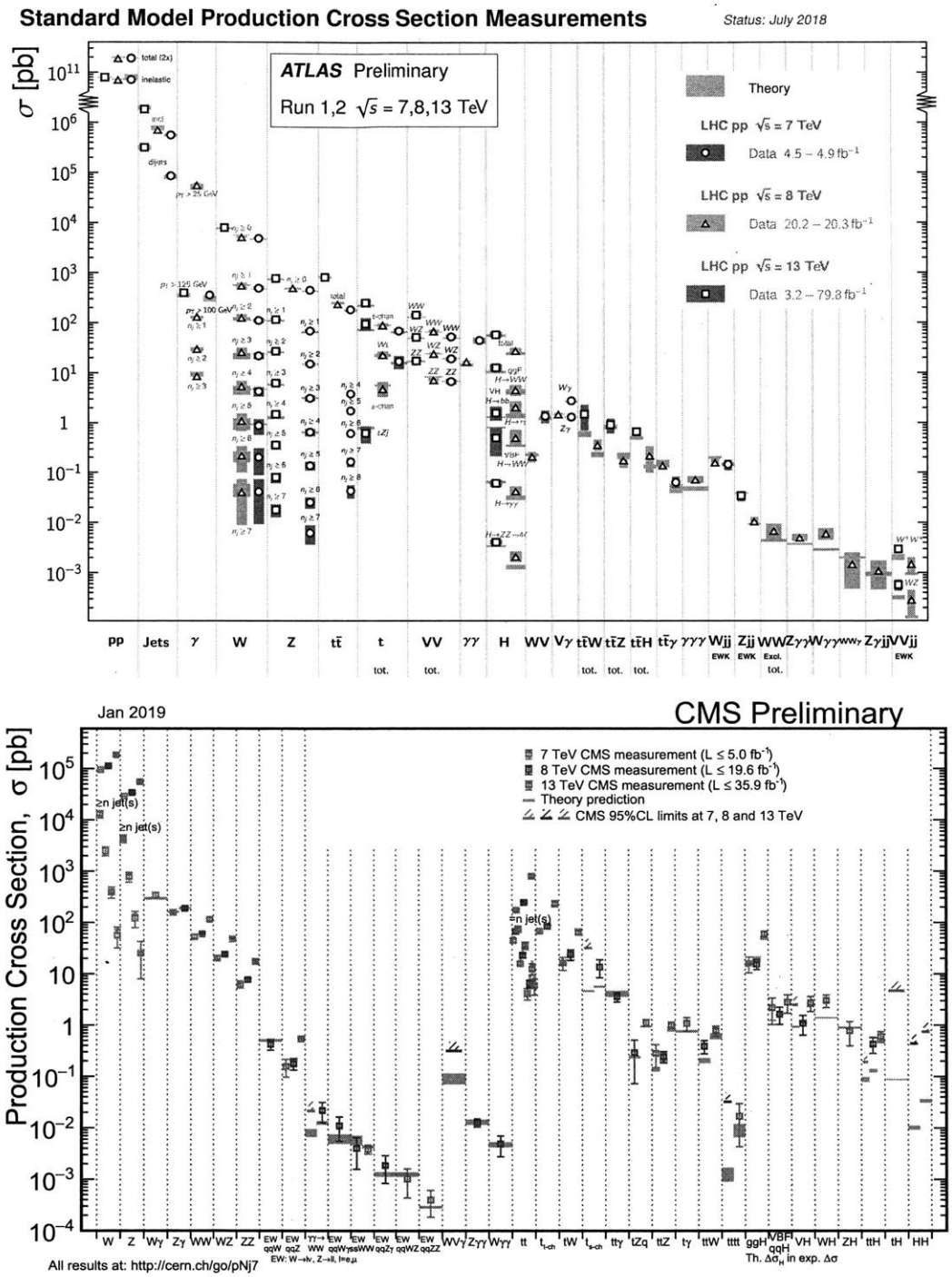


Figure 1-1: Summary of several SM cross section measurements at the ATLAS (top) and CMS (bottom) experiments [1, 2].

Chapter 2

Theoretical Approach

The Standard Model (SM) is one of the most beautiful and successful theories in modern physics. It exploits the elegant mathematics of group representations and gauge symmetries, and more importantly, describes and predicts the experimental results with very high accuracy.

The SM is built based on $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ symmetry group. It categorizes the elementary fermionic particles that make up matter: quarks and leptons, into three generations, and describes three fundamental interactions that govern the universe: strong, electromagnetic and weak forces, by exchange of gauge bosons. See Figure 2-1 for illustration. Quantum chromodynamics [38–42] is the theory describing the strong interaction with the $SU(3)_C$ component, while the electromagnetic and weak forces are nicely incorporated together into the electroweak unified theory [43–46] using the $SU(2)_L \otimes U(1)_Y$ sector. In order to acquire masses for massive gauge bosons and fermions, spontaneous symmetry breaking mechanism [47–51] is invoked. The simplest approach exploits a scalar doublet and leaves one physical Higgs boson after breaking, i.e, the SM Higgs, which completes the last piece of SM particles.

However, several other questions remain concerning and hierarchy problem [32] is one of those. To account for the diverging radiation corrections of the Higgs boson mass at the high energy scale, supersymmetry (SUSY) [52, 53] is proposed, which assumes a set of SUSY particles and a Lagrangian that cancels the divergence us-

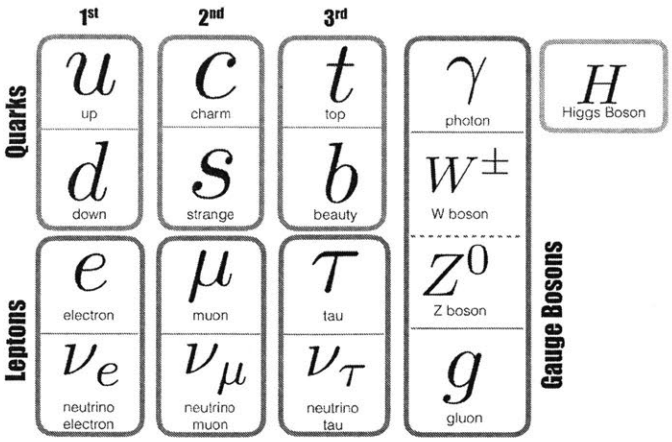


Figure 2-1: The SM particles [3].

ing those superpartners. The Higgs sector extends to a more complex structure as well. In the mostly discussed SUSY model, Minimal Supersymmetric Standard Model (MSSM) [54–56], the Higgs sectors consists of two complex doublets, to provide masses for up and down type fermions.

To provide relevant theoretical considerations for the measurements and search covered in this thesis, this chapter is organized as follows. Section 2.1 discusses the essentials in quantum chromodynamics predictions, using the Z boson production as an example. Section 2.2 describes the fundamentals of the electroweak theory and the Higgs boson physics, with emphasis on the MSSM Higgs sector.

2.1 Quantum chromodynamics

Quantum chromodynamics (QCD) describes the strong forces of quarks and gluons. The QCD Lagrangian is given in Section 2.1.1. The different aspects of the QCD prediction for experimental results are specified in Section 2.1.2, using the Z boson production in proton-proton collisions for demonstration. This example is chosen since the dedicated measurements of Z boson production will be discussed in Chapter 5 of this thesis.

2.1.1 QCD Lagrangian

Define the quark field of color α and flavor f as q_f^α and $q_f \equiv (q_f^1, q_f^2, q_f^3)$. Considering the invariance under the local $SU(3)_C$ gauge transformation, eight gauge parameters G_a^μ, which give rise to eight gauge bosons, gluons, are introduced in the covariant derivative:

$$D_\mu \equiv \partial^\mu + ig_s \frac{\lambda^a}{2} G_a^\mu, \tag{2.1}$$

and the invariant QCD Lagrangian:

$$\mathcal{L}_{QCD} = -\frac{1}{4} G_a^{\mu\nu} G_{\mu\nu}^a + \sum_f \bar{q}_f (i\gamma^\mu D_\mu - m_f) q_f. \tag{2.2}$$

Here, $G_a^{\mu\nu} = \partial^\mu G_a^\nu - \partial^\nu G_a^\mu - g_s f^{abc} G_b^\mu G_c^\mu$ is used to describe the kinematic term for the gluons, γ^μ are the Dirac γ matrices, and λ^a the generators of $SU(3)_C$ group. g_s denotes the coupling strength and m_f the quark masses.

Running coupling

The coupling strength is usually expressed as running coupling $\alpha_s = \frac{g_s^2}{4\pi}$. When α_s is small, pertubative expansions of α_s can be applied for QCD calculations. In the pertubative QCD (pQCD) framework, α_s is factorized as a function of the renormalization scale μ_R and satisfies the renormalization group equation (RGE):

$$\mu_R^2 \frac{d\alpha_s}{d\mu_R^2} = \beta(\alpha_s), \tag{2.3}$$

where $\beta(\alpha_s)$ is a polynomial series of α_s starting from second order [4]. For a specific process, the running coupling strength is usually estimated by taking the μ_R to be the momentum transfer Q. The corresponding dependence is shown in Figure 2-2 from different measurements and in fact implies one of the essential properties of QCD: asymptotic freedom [41, 42]. As can be seen, $\alpha_s \sim 0.1$ for $Q > 100$ GeV, which indicates pQCD is appropriate for high-Q (hard) processes.

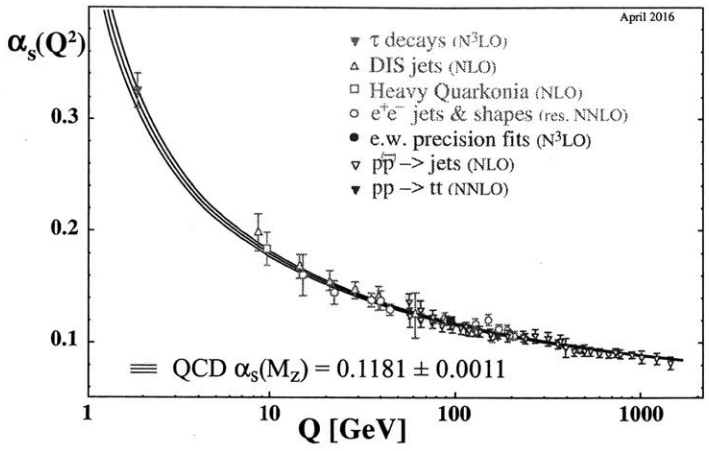


Figure 2-2: Summary of measurements of α_s as a function of the energy scale Q [4].

2.1.2 QCD predictions of the Z boson production

As Figure 2-3 suggests, it is protons that collide in particle experiments, while pQCD calculates hard process of partons (quarks and gluons). To predict experimental results, one has to weight sub-process partonic cross sections with the probability to have corresponding partons in the colliding protons for all orders. See Equation 2.4 for the full expression of the inclusive Z boson production cross section in proton-proton collisions with pQCD calculations [4].

$$\sigma(pp \to Z{+}X) = \sum_{n=0}^{\infty} \alpha_s^n(\mu_R^2) \sum_{a,b} \int dx_a dx_b f_{a/A}(x_a, \mu_F^2) f_{b/B}(x_b, \mu_F^2) \hat{\sigma}^{(n)}_{ab\to Z+X}(x_a x_b s, \mu_R^2, \mu_F^2). \tag{2.4}$$

Here, the assigned weights, $f_{a/A}$ and $f_{b/B}$, are defined as parton distribution functions, which are dependent on the factorization scale μ_F and momentum fractions, x_a and x_b, carried by relevant partons. $\hat{\sigma}^{(n)}_{ab\to Z+X}$ denotes the fixed-order calculation of the sub-process partonic cross sections (also called matrix elements).

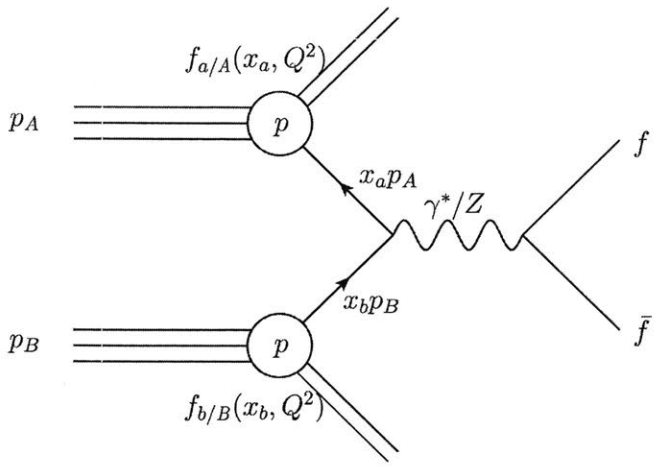


Figure 2-3: Feynman diagram of the Z boson production process in proton-proton collisions [5].

Parton distribution functions

Parton distribution functions (PDFs) $f_{q/p}(x)$ describe the number density of parton q inside proton p with a momentum fraction x. The parton momenta can be modified via gluon exchanges and radiations, which are usually soft and collinear to the partons and can cause divergence in integrals. Therefore, the factorization scale μ_F is introduced so that the collinear divergences from radiation momenta below μ_F are factorized into re-defined non-perturbative PDFs $f_{q/p}(x, \mu_F^2)$ and then determined by experimental data. The corresponding dependence on μ_F is characterized by Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) evolution equations [57–59]. For scale choices of $\mu_F^2 = 10\,GeV^2$ and $10^4\,GeV^2$, $xf(x, \mu_F^2)$ are illustrated in Figure 2-4.

Fixed-order matrix element calculations

As mentioned briefly, soft and collinear radiation of partons can cause divergence. The fixed-order calculations are finite only for the soft and collinear safe quantities, for which the divergent terms can cancel with those in loop corrections from the similar source. The inclusive cross section is one of the safe quantities, as the fully inclusive variables do not probe the topological or kinematic structure of the final

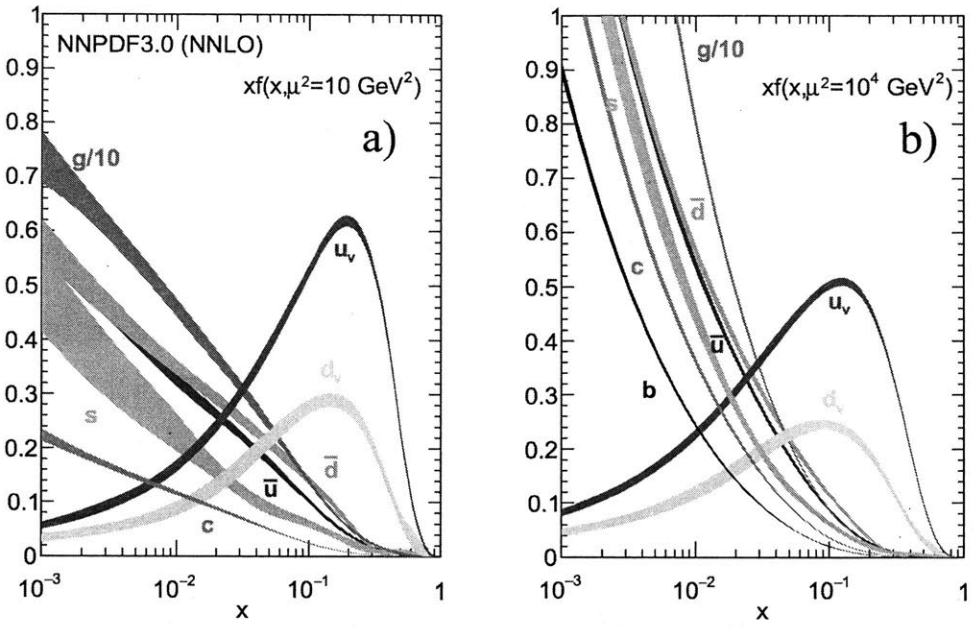


Figure 2-4: $xf(x, \mu_F^2)$ as a function of x, obtained from NNPDF3.0 global analysis at scales $\mu_F^2 = 10\,GeV^2$ (left) and $\mu_F^2 = 10^4\,GeV^2$ (right) [4].

state partons. Such inclusive hard cross sections are now calculable with pQCD up to α_s^2, the next-to-next-leading order (NNLO) [60]. For instance, simulation codes FEWZ [61] provides NNLO calculations for hadronic Z boson production process, which is used as reference for the measurements discussed in Chapter 5.

Resummation

On the other hand, not fully inclusive cross sections, such as Z boson production associated with jets, are also of great interests to allow study of the final state structures. Direct experimental measurements are susceptible to a sizable jet energy uncertainty, but it can be alternatively approached indirectly by studying the Z boson transverse momentum distributions.

The transverse momentum of Z boson (p_T^Z) appears to balance the real parton radiations in the Z boson production process. When this radiation is hard, which is corresponding to a p_T^Z of tens to hundreds of GeV, pQCD is still applicable. While for p_T^Z at the order of a few GeV, fixed-order calculations diverge. This is because the real radiations are subject to the p_T constraints but not the 1-loop corrections, thus the

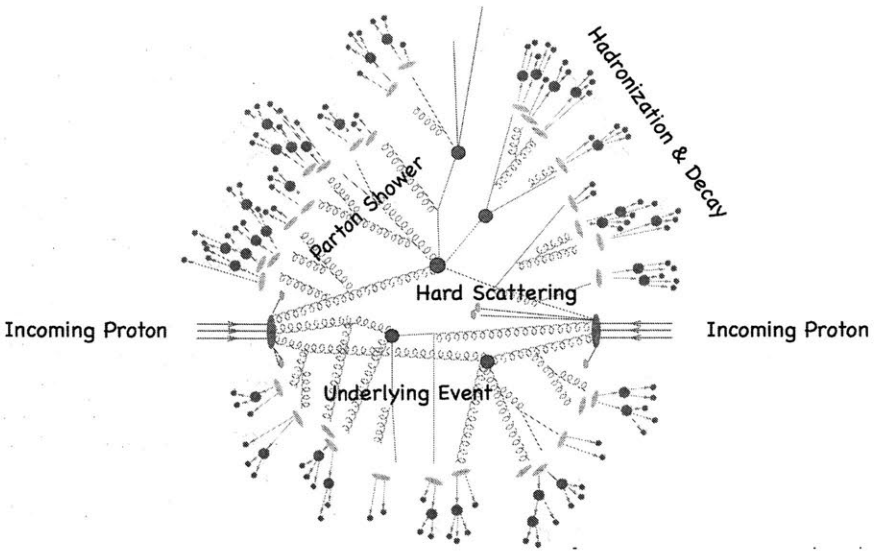


Figure 2-5: Illustration of a full picture of various processes in a proton-proton collision [6].

divergent terms cannot fully cancel and end up with large logarithmic terms $ln\frac{M_Z}{p_T^Z}$ for each order of α_s. In those cases, a different technique, which is called "resummation", sums up most significant logarithmic terms in all orders in order to obtain reasonable predictions for the differential measurements [62–64]. Currently, such resummation calculations are available up to next-to-next-leading logarithm (NNLL) in specific programs such as ResBos [65] and DYRes [66].

Parton shower and hadronization

Final state partons produced in the hard processes ultimately go through:

- parton shower, to split into multiple partons,
- hadronization, to form hadrons,
- and decay of the unstable hadrons,

which are illustrated pictorially in Figure 2-5 for a proton-proton collision.

Parton showers are a series of parton splittings, starting from partons from the hard processes and evolve into soft radiations. The factorization between pQCD and non-perturbative calculations is handled by parton splitting functions with decreasing scales to describe the different probabilities of various parton splitting processes, see Figure 2-6, which are characterized by DGLAP evolution equations [57–59] as well.

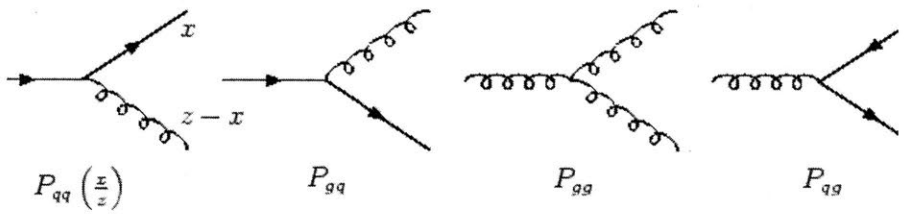


Figure 2-6: Process for lowest order splitting functions [7].

Parton showers stop and start to hadronize at a scale of ~1 GeV. The hadronization process is complex and non-perturbative and relies on phenomenological models. The two prevailing models are the string model [67,68] and cluster model [69]. String model exploits linear color confinement by building and breaking color strings between partons, while cluster model groups color singlet clusters along the color flows.

In addition, as Figure 2-5 indicates, there are multiple parton interactions (MPI) in a single proton-proton collision, which mostly lead to additional soft jets and need to be considered.

Event generators

All aspects of the proton-proton collisions discussed previously are implemented in event generators for a wide variety of processes for simulation. Those generators produce exclusive events and can be interfaced to different PDF sets and to detailed description of detector geometry for detector response, such as GEANT4 [70] for the CMS detector.

PYTHIA [71] provides leading-order (LO) matrix element calculations for an extensive list of SM and BSM processes, and also approximates high-order corrections for vector boson productions by matching to next-to-leading-order (NLO) cross sections. Parton showering is implemented with leading logarithms (LL) resummations and hadronization is realized via string model. As some aspects cannot be derived from first principles, some parameters have to be tuned, especially for hadronization and MPI simulation. The tune `CUETP8M1` [72] is applied by default for the simulated samples used in the scope of this thesis.

Madgraph [73] also calculates LO matrix elements and is useful particularly for

vector boson production associated with higher jet multiplicities. It adds real emissions of additional partons to matrix elements and virtual corrections later in parton showering, after which the processes are simulated by PYTHIA. As can be inferred, for the multi-parton final states, the parton showers can cause overlap with the matrix element calculation, as both generate associated jets. To remove the overlaps, matching and merging methods such as Catani-Krauss-Kuler-Webber (CKKW) [74,75] and Mangano (MLM) [76] scheme have been developed.

Madgraph5_aMC@NLO [77, 78] (aMC@NLO) and POWHEG [79–82] combines NLO matrix element calculations and parton showering. Similarly, to deal with the overlap, aMC@NLO removes all terms of matrix elements that appear in parton showering (FxFx scheme) [83], while POWHEG generates the hardest emission first and passes soft radiation to parton showering.

2.2 Electroweak theory and the Higgs boson physics

This section aims to provide theoretical considerations for the second topic covered in this thesis, i.e. search for a heavy neutral Higgs boson. After the electroweak unification and SM Higgs boson are briefly discussed in Section 2.2.1 and 2.2.2, Section 2.2.3 focuses on the MSSM Higgs sector. The masses, couplings, production mechanisms and decay modes of the neutral MSSM Higgs bosons are elaborated. Two representative MSSM benchmark scenarios are also introduced.

2.2.1 Electroweak unification

Denote ψ_i as the i-th generation of fermions (i =1, 2, 3). Considering the invariance under the local $SU(2)_L \otimes U(1)_Y$ gauge transformation, four gauge fields, W^i_μ and B_μ, are introduced in the covariant derivative:

$$D_\mu \equiv \partial^\mu + ig\frac{\sigma_i}{2}W^i_\mu + ig'YB_\mu, \tag{2.5}$$

and the invariant Lagrangian:

$$\mathcal{L}_{EW} = -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}W^i_{\mu\nu}W_i^{\mu\nu} + \sum_{j=1}^{3} i\bar{\psi}^{\mu}_j D_\mu \psi_j. \tag{2.6}$$

Here, $B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$ and $W^i_{\mu\nu} = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu - g\epsilon^{ijk}W^j_\mu W^k_\nu$ are defined to build the kinematic term for the gauge fields. g, g' are the coupling strengths and Y is the hypercharge. To give rise to the four physical gauge bosons, $W^{\pm}$, Z and A, the original gauge fields can be rotated as:

$$\begin{aligned} W^{(\mp)}_\mu &= (W^1_\mu \pm iW^2_\mu)/\sqrt{2}, \\ \begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} &= \begin{pmatrix} cos\theta_W & -sin\theta_W \\ sin\theta_W & cos\theta_W \end{pmatrix} \begin{pmatrix} W^3_\mu \\ B_\mu \end{pmatrix}, \end{aligned} \tag{2.7}$$

where the θ_W is denoted as the weak mixing angle.

The gauge symmetry forbids the mass terms of those gauge fields, which however would be in obvious contradiction to the experimental measurements. To solve this dilemma, a new scalar field is introduced by the spontaneous symmetry breaking mechanism.

2.2.2 Spontaneous symmetry breaking and the SM Higgs boson

To break the gauge symmetry while maintaining a symmetric Lagrangian, a $SU(2)_L$ doublet of complex scalar field:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \tag{2.8}$$

and its invariant Lagrangian with infinite degenerated vacuum states:

$$\mathcal{L}_S = (D_\mu\phi)^\dagger(D^\mu\phi) - \mu^2\phi^\dagger\phi - h(\phi^\dagger\phi)^2, (h > 0, \mu^2 < 0), \tag{2.9}$$

are considered. The symmetry is broken when one vacuum state is chosen, which satisfies $|\langle 0|\phi^0|0\rangle| = \sqrt{\frac{-\mu^2}{2h}} \equiv \frac{v}{2}$. Equation 2.8 can be rewritten into:

$$\phi \rightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v+H \end{pmatrix}, \tag{2.10}$$

and accordingly, the kinematic term in Equation 2.9 can be expanded into:

$$(D_\mu \phi)^\dagger (D^\mu \phi) \rightarrow \frac{1}{2}\partial_\mu H \partial^\mu H + (v+H)^2 \{\frac{g^2}{4} W_\mu^\dagger W^\mu + \frac{g^2}{8cos^2\theta_W} Z_\mu Z^\mu\}. \tag{2.11}$$

From Equation 2.11, one can restore mass terms for $W^\pm$ and Z bosons:

$$\begin{aligned} m_W &= \frac{1}{2} vg, \\ m_Z &= \frac{1}{2} \frac{vg}{cos\theta_W}. \end{aligned} \tag{2.12}$$

In addition, fermions acquire masses by adding the Yukawa coupling [36, 44] to the Lagrangian:

$$\mathcal{L}_Y = -(1 + \frac{H}{v})\{m_f \bar{f} f\}. \tag{2.13}$$

Given those couplings, the dominant production mechanisms at the LHC are gluon fusion, vector boson fusion, associated production with a gauge boson and with a pair of top or bottom quarks. The Feynman graphs of those processes are illustrated in Figure 2-7 and the corresponding production cross sections as a function of $\sqrt{s}$ are shown in Figure 2-8. The branching ratios of the main decay modes near $m_H = 125$ GeV are also presented.

2.2.3 MSSM Higgs sector

In contrast to the SM, the MSSM Higgs sector consists of two scalar doublets. After symmetry breaks, the two doublets condensate into five physical Higgs bosons: two neutral CP-even scalar h and H[1], with $m_h < m_H$, one neutral CP-odd pseudoscalar

[1]For the rest of this chapter, h and H represents the light and heavy MSSM neutral Higgs bosons, unless explicitly specified.

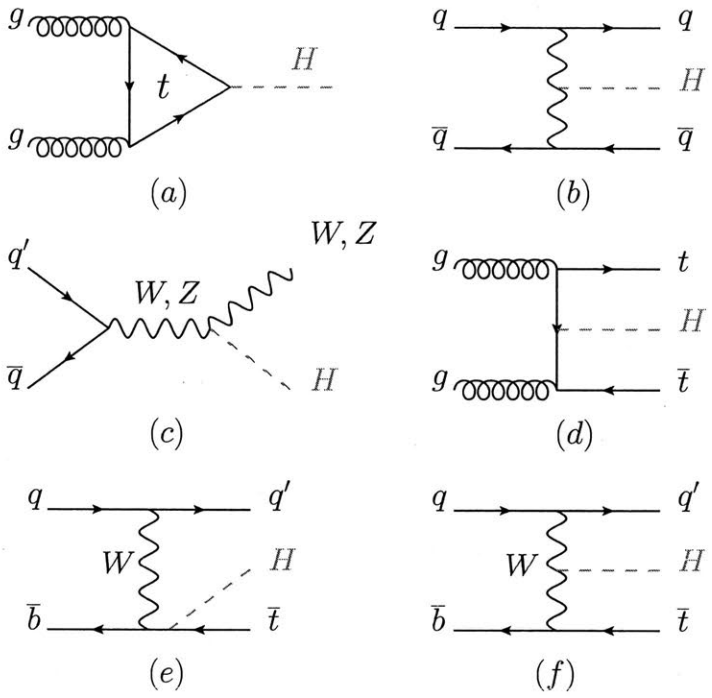


Figure 2-7: Feynman graphs for the SM Higgs boson production mechanisms at LHC: (a) gluon fusion, (b) vector boson fusion, (c) associated production with a gauge boson, (d) associated production with a pair of top (or bottom) quarks, (e-f) associated production with a single top quark [4].

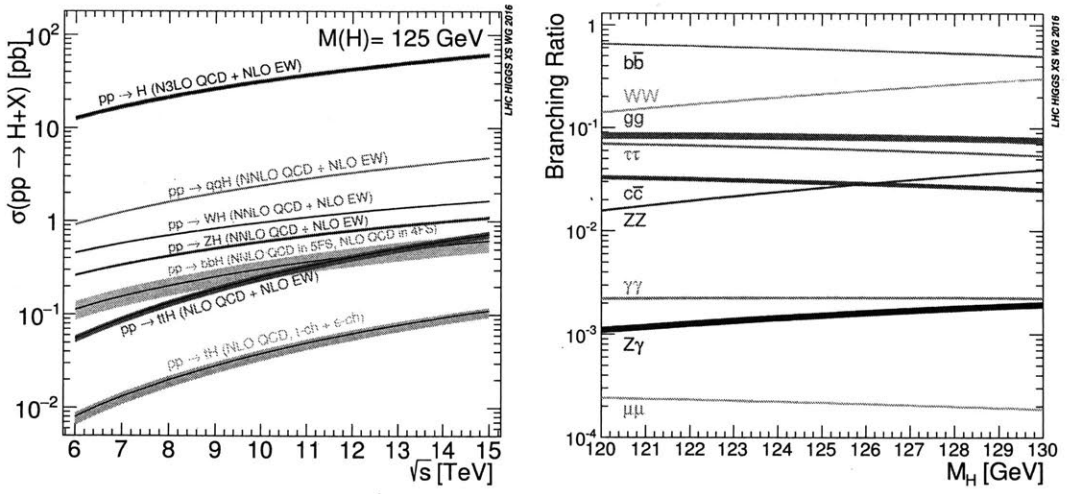


Figure 2-8: The SM Higgs boson production cross sections as a function of $\sqrt{s}$ (left) and decay branching ratios near $m_H = 125$ GeV (right) [4].

	g_{VV}	g_{uu}	g_{dd}
A	0	$cot\beta$	$tan\beta$
H	$cos(\beta-\alpha)$	$sin\alpha/sin\beta$	$cos\alpha/cos\beta$
h	$sin(\beta-\alpha)$	$cos\alpha/sin\beta$	$-sin\alpha/cos\beta$

Table 2.1: The coupling strength of MSSM neutral Higgs bosons relative to the corresponding SM counterparts.

A and a pair of charged bosons $H^{\pm}$.

At tree level, the MSSM Higgs sector can be specified by m_W, m_Z, and two additional parameters, which are chosen by convention as m_A, the mass of the A boson, and $tan\beta = \frac{v_2}{v_1}$, the ratio of the vacuum expectation values of the two Higgs doublets. The tree-level mass terms can be expresses as [9]:

$$\begin{aligned} m^2_{h,H} &= \frac{1}{2}(m^2_A + m^2_Z \mp \sqrt{(m^2_A + m^2_Z)^2 - 4m^2_A m^2_Z cos^2 2\beta}), \\ m^2_{H^{\pm}} &= m^2_A + m^2_W. \end{aligned} \tag{2.14}$$

In terms of the couplings to SM particles, analogous to the SM Higgs bosons, the neutral MSSM Higgs bosons also couple to the massive gauge bosons (VV) and up and down type fermions (uu, dd). The relative coupling strengths to corresponding SM counterparts are listed in Table 2.1. α is the angle that diagonalizes the mass matrix, which at tree level is:

$$tan\alpha = \frac{-(m^2_A + m^2_Z)sin2\beta}{(m^2_Z - m^2_A)cos2\beta + \sqrt{(m^2_A + m^2_Z)^2 - 4m^2_A m^2_Z cos^2 2\beta}}. \tag{2.15}$$

In the decoupling limits ($m_A \gg m_Z$), which seems to be more and more favored by the experimental data [84], Table 2.1 would simplify to Table 2.2. This indicates the h couplings become SM-like, A and H reduce to a heavy and mass-degenerate state that does not couple to vector bosons while the couplings to up(down) type fermions are suppressed(enhanced) with large $tan\beta$.

However, in this limit, the tree-level m_h has an infamous upper bound: $m_h \approx m_Z|cos2\beta| \leq m_Z$. To lift m_h to the observed SM Higgs mass of 125.09 $\pm$ 0.24

	g_{VV}	g_{uu}	g_{dd}
A	0	$cot\beta$	$tan\beta$
H	0	-$cot\beta$	$tan\beta$
h	1	1	1

Table 2.2: The coupling strength of MSSM neutral Higgs boson relative to the corresponding SM counterparts, in the decoupling limits ($m_A \gg m_Z$).

GeV [29–31], substantial radiative corrections from loops of quarks and superpartners are essential, in which case more model-dependent parameters have to be introduced besides m_A and $tan\beta$. When interpreting experimental results under MSSM, complete scans of the entire parameter space would be highly impractical. Thus, several different benchmark scenarios [85] has been proposed by imposing the constraints from experimental data, including the observation of an scalar boson with a mass of 125.09 $\pm$ 0.24 GeV and couplings compatible with the SM Higgs [86–88], and no evidence for additional particles beyond SM. All parameters other than m_A and $tan\beta$ are fixed in those scenarios. Two representative benchmark scenarios of the MSSM, the m_h^{mod+} [85] and hMSSM [89–91], are mostly exploited nowadays to interpret the data. Within both scenarios, the observed particle is interpreted as h. The m_h^{mod+} scenario constraints $m_h = 125 \pm 3$ GeV, where 3 GeV is the theoretical uncertainty on m_h. While in hMSSM scenario the m_h is fixed to the experimentally known value.

Production channels and decay modes

Given the couplings described above, the productions and decays of neutral MSSM Higgs bosons are very different from those of the SM Higgs. While the vector boson fusion and the associated production with a gauge boson are rather dominant in the SM Higgs production, the couplings of neutral MSSM Higgs bosons to massive gauge bosons are either suppressed in the case of H, for large values of $tan\beta$ and m_A, or even forbidden in the case of A. In addition, the associated production with bottom quarks is suppressed in the SM by two orders of magnitude with respect to the gluon fusion (see Figure 2-8), but in the MSSM case, it can be significantly enhanced for large $tan\beta$ and hence become the leading production channel. As a result, the MSSM

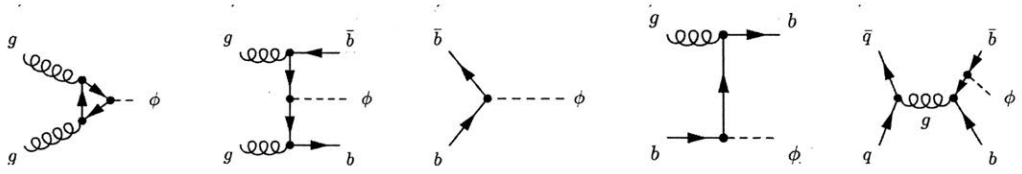


Figure 2-9: Feynman diagrams for different MSSM neutral Higgs boson production mechanisms: gluon-gluon fusion and b-quark association production [8].

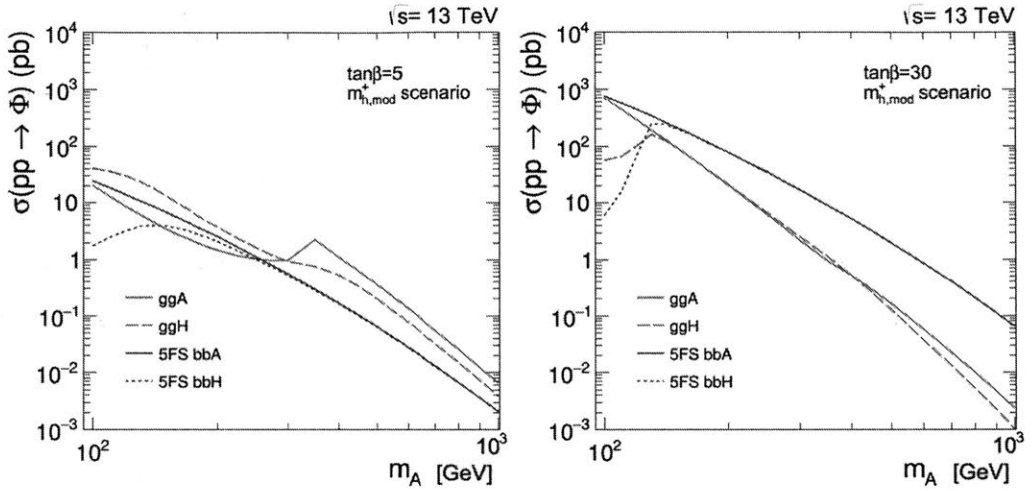


Figure 2-10: Higgs boson production cross sections in the $m^{+}_{h,mod}$ scenario for $tan\beta = 5$ (left) and $tan\beta = 30$ (right) [9].

neutral Higgs bosons are mainly produced via two mechanisms: gluon fusion (denoted as $gg\phi$, $\phi = H, A$) and associated production with bottom quarks (denoted as $bb\phi$). The corresponding Feynman diagrams are illustrated in Figure 2-9 and the cross sections predicted with the m_h^{mod+} scenario are shown in Figure 2-10 as a function of m_A, for $tan\beta = 5$ and 30, respectively. As can be observed, $gg\phi$ is dominant at low $tan\beta$ region while $bb\phi$ is enhanced at high $tan\beta$ values.

For the very same reason, while the decays to massive gauge bosons played important roles for the SM Higgs searches, they are either suppressed or forbidden in the MSSM. Instead, the couplings to third and second generation fermions are strongly enhanced in large regions of the MSSM parameter space, and the couplings to bottom type fermions are further magnified at large $tan\beta$ values. This makes the sensitivity of the Higgs search in the $\mu\mu$ channel much more significant compared to the SM case (see Figure 2-8) and increases the potential to find evidence of the Yukawa coupling

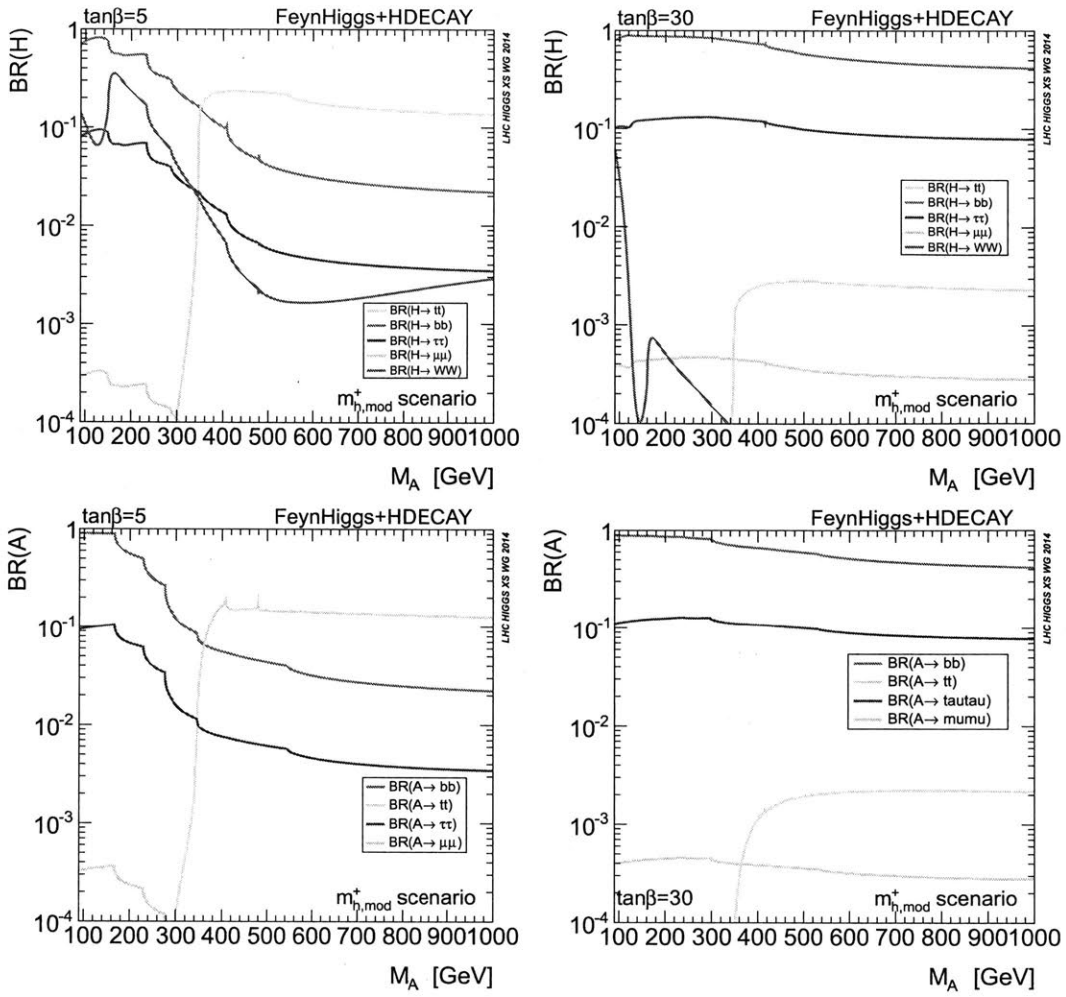


Figure 2-11: MSSM neutral boson decay branching ratios as a function of m_A, for H boson (top) and A boson (bottom), $tan\beta = 5$ (left) and $tan\beta = 30$ (right), respectively [9].

to the second generation fermions. Figure 2-11 shows the branching ratios of the H and A decay modes under the m_h^{mod+} scenario, as a function of m_A, for $tan\beta = 5$ and 30, respectively.

Chapter 3

Experimental Apparatus

To perform the measurements and search covered in this thesis, the data taken from the Compact Muon Solenoid (CMS) detector on the Large Hadron Collider (LHC) machine are exploited. Up to the present, the CMS and LHC have successfully accomplished two periods of data-taking: Run I era from 2010 to 2012, and Run II era from 2015 to 2018. Figure 3-1 reflects the status of data taking during the two periods, by showing the integrated luminosity delivered to the CMS detector. The data analyzed for physics results covered in this thesis are taken from Run II.

This chapter is organized as follows. Section 3.1 discusses the LHC machine and its two key parameters: the beam energy and instantaneous luminosity. The CMS detector and its main subsystems are described in Section 3.2.

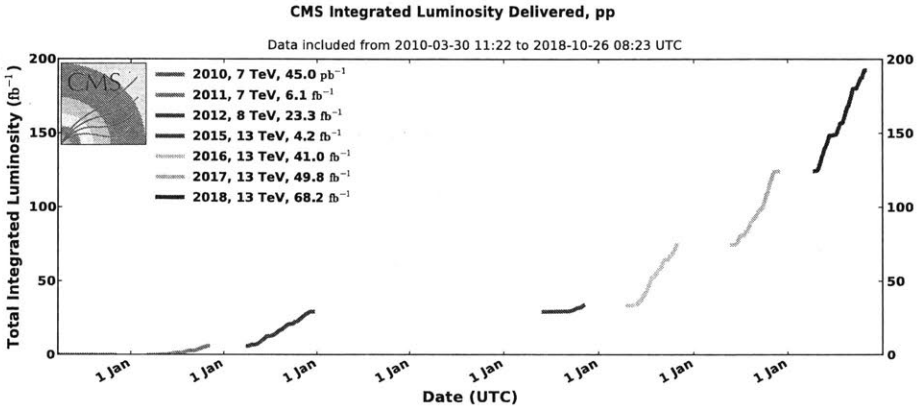


Figure 3-1: Integrated luminosity delivered to CMS by LHC from 2010 to 2018 [10].

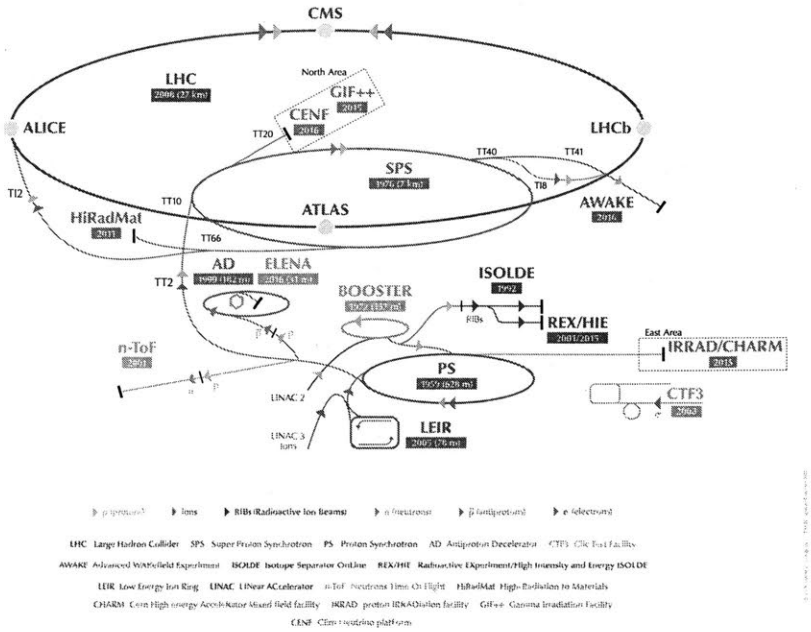


Figure 3-2: The CERN accelerator complex [11].

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) machine [26] is the world's largest superconducting circular proton accelerator and collider. It has a 27-km circumference and is installed in a 4-m-diameter tunnel ~100 m underground. It consists of two rings with counter-revolving proton beams, which are accelerated and made to collide at different interaction points (IP) for various particle detectors, including ATLAS [92] and CMS for studying the high-p_T physics, LHCb [93] for flavor physics and the ALICE [94] for heavy ion physics. Figure 3-2 illustrates the CERN accelerator complex and the largest ring is the LHC. As can be seen, the proton beams go through a series of different accelerator rings and increase energy stepwise before finally injecting into the LHC rings.

The collision energy and luminosity are the two key parameters for the LHC machine, since they decide the distance regime between particles that can be probed and data statistics to perform analyses. The LHC is designed to deliver proton-proton collisions at a center-of-mass energy of 14 TeV and an instantaneous luminosity of

	Design	Run I	2015	2016	2017	2018
Energy (TeV)	14	7/8	13	13	13	13
Bunch spacing (ns)	25	50	50/25	25	25	25
N ($\times 10^{11}$)	1.15	1.6	1.2	1.1	1.25	1.15
n	2800	1400	2200	2200	1900	2500
ϵ (μm)	3.5	2.2	3.5	2.5	2.0	2.2
β^* (cm)	55	80	80	40	30	30→25
crossing angle (μrad)	285	-	290	280	300→240	300→260
Peak $\mathscr{L}$ ($\times 10^{34} cm^{-2} s^{-1}$)	1.0	0.8	0.5	1.5	1.5 (leveled)	2.0
Peak pile-up	25	45	25	45	65	60

Table 3.1: Key LHC parameters and comparison during different data-taking periods [25].

$10^{34} cm^{-2} s^{-1}$. To achieve the desired energy and luminosity, LHC exploits sophisticated and advanced technology of magnets to deviate, squeeze and target the proton beams. Along the 27-km circumference, 1232 main steering dipoles, 858 main focusing quadrupoles and 6000 corrector magnets are utilized. The dipole superconducting coils operates at a current of 12000 A in the superfluid helium at 1.9 K and are capable to reach a magnetic field strength of 8.33 Tesla. This magnetic field strength determines the maximum beam energy that can be achieved, i.e, 7 TeV per beam. In addition, the instantaneous luminosity is decided by a few parameters according to

$$\mathscr{L} = f \frac{n N_1 N_2 \gamma_r}{4\pi \epsilon_n \beta^*} \times F, \tag{3.1}$$

where the f is revolution frequency, n the number of proton bunches per beam, N the number of protons per bunch, γ_r the relativistic factor, ϵ_n the normalized transverse beam emittance, β^* the beta function, and F the geometric reduction factor to account for the crossing angle at IPs. Table 3.1 lists and compares the designed and practical key LHC parameters during different data-taking periods.

3.2 The Compact Muon Solenoid detector

The Compact Muon Solenoid (CMS) detector [18] is named because of its compact size given the material it contains, the high momentum resolution for muon detection and the use of the most powerful solenoid ever made. The overall CMS detector

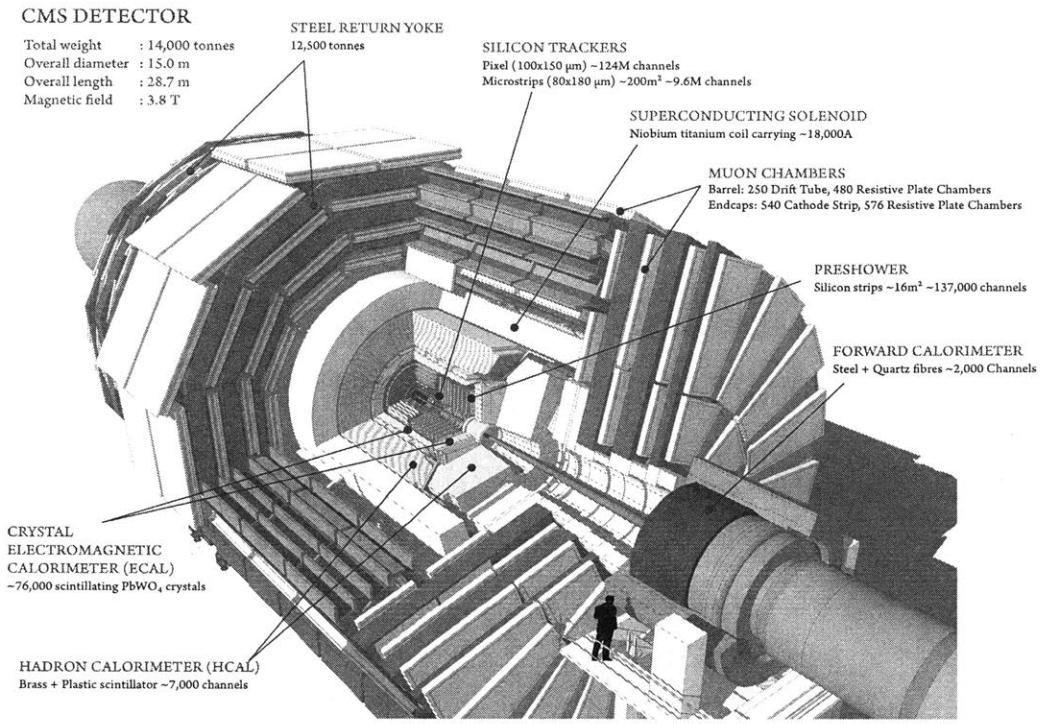


Figure 3-3: A Cutaway view of the CMS detector [12].

is 28.7 m in length and 15.0 m in diameter. In order to provide strong magnetic field to bend high energy particles and measure corresponding momenta, the CMS magnet system consists of a NbTi superconducting solenoid of 12.5-m long and 6-m in diameter, operating at 3.8 Tesla, and the steel returning yokes, which are segmented into 5 wheels in barrel and 3 disks for each endcap. See Figure 3-3 for a cutaway view of the CMS detector. The all-silicon inner tracking detector, lead tungsten crystal homogeneous electromagnetic calorimeter and the brass-scintillator sampling hadronic calorimeter are inside the solenoid, while the muon detectors are interleaved with returning yokes. Subdetectors will be described in details in the rest of this section.

The CMS coordinate system follows the right-handed rule, see Figure 3-4. The origin is located at the interaction point, the x axis points to the center of the LHC ring and the y axis points upward. The longitudinal direction is along z axis and the transverse direction lies on the $x-y$ plane[1]. In the transverse plane, the radial coordi-

[1]The transverse component of variables are usually denoted with T as subscript. For example the transverse momentum is denoted as p_T and the transverse missing momentum p_T^{miss}.

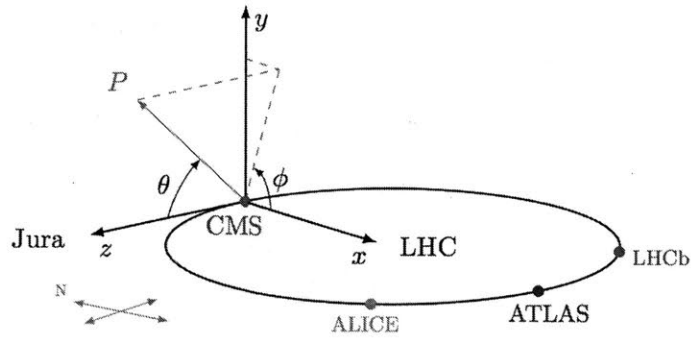


Figure 3-4: The CMS coordinate system [13].

nate is denoted as r and the azimuthal angle started from x axis is denoted as ϕ. θ is then defined in $r-z$ plane to be the angle with respect to z axis. The pseudorapidity is defined as $\eta = -ln(tan\frac{\theta}{2})$. As a result, $\eta-\phi$ corresponds to a orthogonal coordinate system, where the direction of an outgoing particle is represented by a point. Thus, the separation of two particles can be defined as $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

3.2.1 Inner tracking detectors

The inner tracking detectors are located at the most inner part of the CMS detector. At current experimental setup with an instantaneous luminosity of $\sim 10^{34} cm^{-2} s^{-1}$, more than 1000 charged particles are produced per every bunch crossing (i.e 25 ns) from over 20 overlapping proton-proton collisions. To be able to precisely measure the trajectories of charged particles and reconstruct the primary and secondary vertices in such challenging environment, the inner tracking system needs to have fine granularity, fast time response, as well as radiation hardness, which make silicon detectors to be the suitable technology.

The basic module of a silicon strip is illustrated in Figure 3-5. The energy of the charged particle passing by knocks off electrons and creates electron-hole pairs, which are then collected to form electric signals. The pixel modules function in a similar manner.

A simplified sketch of the top half of the inner tracking system is shown in Figure 3-6. As can be seen, the inner tracking system consists of a small region of pixel layers (shown in red), which provides 3-dimensional position measurements, and a large area

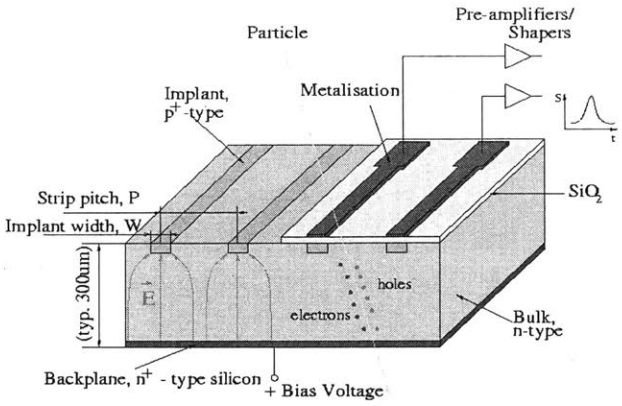


Figure 3-5: A silicon strip module and its operation principle [14].

Tracker subsystem	Layers	Cell size	Cell thickness
Pixel barrel	3 cylindrical	100 μm×150 μm	285 μm
Strip inner barrel (TIB)	4 cylindrical	80–120 μm×10 cm	320 μm
Strip outer barrel (TOB)	6 cylindrical	122–183 μm×25 cm	500 μm
Pixel endcap	2 disks	100 μm×150 μm	285 μm
Strip inner disks (TID)	3 disks	100–141 μm×10 cm	320 μm
Strip endcap (TEC)	9 disks	97–184 μm×25 cm	500 μm

Table 3.2: Summary of the main characteristics of the tracker subsystems. The number of disks are for a single endcap [14].

of silicon strip layers, which provide position measurements in 2 dimensions (shown in black) as well as 3 dimensions (shown in blue), which is achieved by placing two strip modules back-to-back and rotating a 'stereo' angle of 100 mrad. Within a layer, adjacent modules are shifted slightly in r (in barrel) or z (in endcap) to avoid gaps. The silicon strips are further divided into four subsystems: inner barrel (TIB), outer barrel (TOB), inner disk (TID) and endcap (TEC). The main characteristics for each subsystem are summarized in Table 3.2.

3.2.2 Electromagnetic calorimeter

The CMS electromagnetic calorimeter (ECAL) is a homogeneous calorimeter made of lead tungstate ($PbWO_4$) crystals, with barrel (EB) covering $|\eta|<1.479$, two endcaps

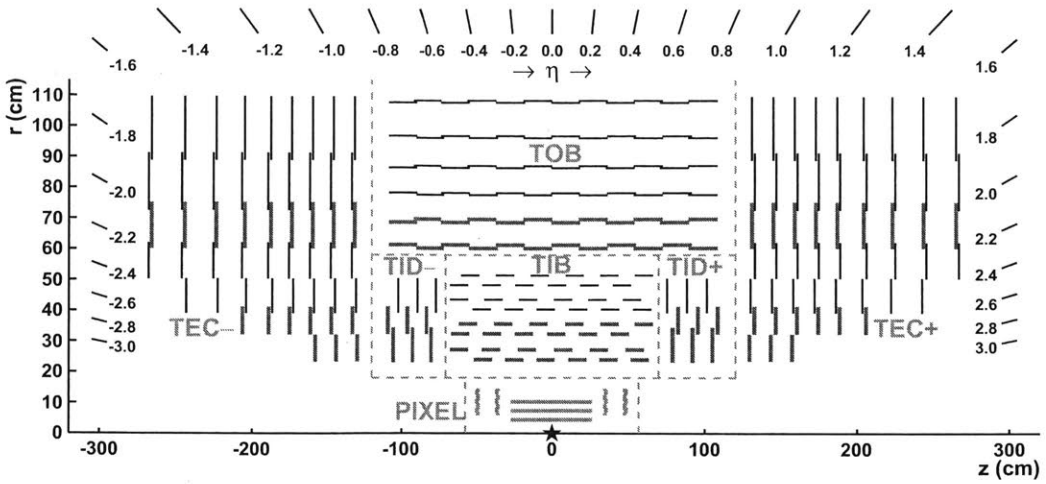


Figure 3-6: A r-z cross section view of the top half of the CMS inner tracker detector. The interaction point is indicated by a star [15].

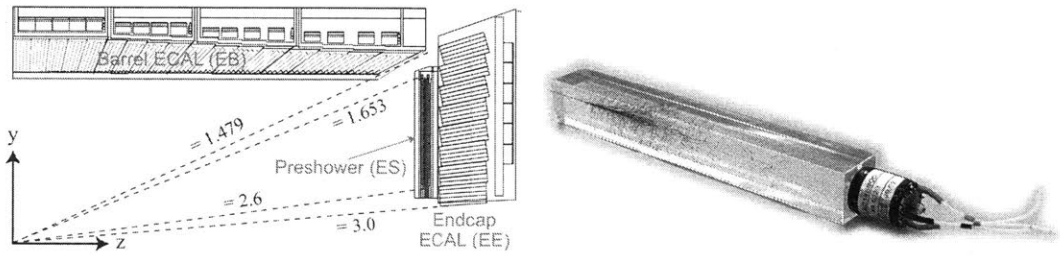


Figure 3-7: A y-z cross section view of a quadrant of the CMS ECAL (left) [16] and a $PBWO_4$ crystal with vaccum phototriodes (right) [17].

(EE) covering up to $|\eta|<3.0$ and gap regions around $|\eta|=1.5$, as shown in Figure 3-7.

The transversing electrons and photons produce secondary photons and electrons by Bremsstrahlung and pair production, which give rise to a cascade of particles, i.e, electromagnetic showers. The length and cross section of $PbWO_4$ crystals are designed to be comparable with the longitudinal and transverse size of such electromagnetic showers, which can be found in Table 3.3, where the radiation length X_0 is defined as the average distance that an electron must travel in the material to reduce its energy to $\frac{1}{e}$ of its original energy. Then the energy deposited is converted in the form of light and further amplified by avalanche photodiodes (APDs) in the barrel and vacuum phototriodes (VPTs) in the endcaps.

Additionally, the preshower detectors (ES), composed of two layers of lead and silicon strips, are installed in front of the ECAL endcaps, covering $1.65 <|\eta|< 2.6$. It

Crystal location	Length	Front-end cross section
ECAL barrel	23 cm ($\sim 26X_0$)	2.2×2.2 cm^2 ($2.5X_0 \times 2.5X_0$)
ECAL endcap	22 cm ($\sim 25X_0$)	2.9×2.9 cm^2 ($3.2X_0 \times 3.2X_0$)

Table 3.3: Summary of the length and cross section size of ECAL barrel and endcap crystals.

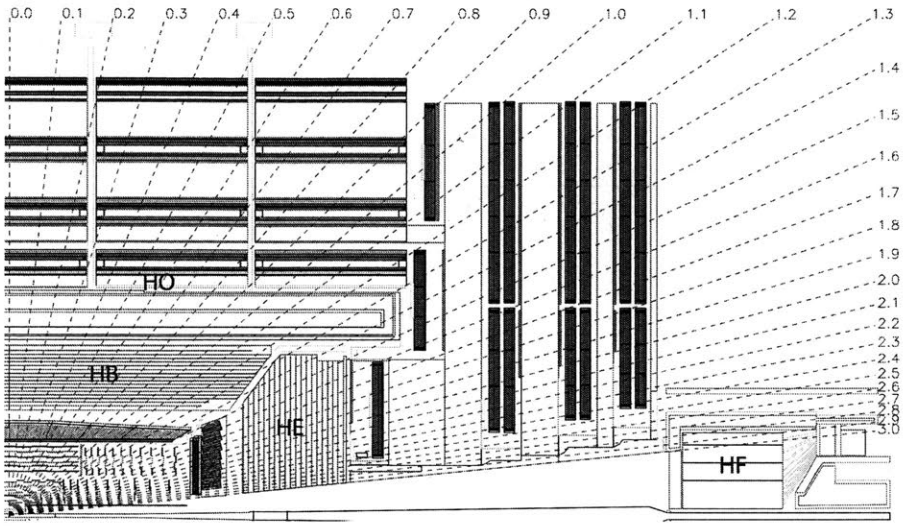


Figure 3-8: A r-z cross section view of a quadrant of the CMS HCAL [18].

was originally aimed to distinguish prompt photons and those decayed from π^0, but actually not fully exploited due to large amount of π^0 from interactions with inner tracker materials.

3.2.3 Hadron calorimeter

The CMS Hadron calorimeter (HCAL) surrounds the ECAL, consisting of barrel (HB) and two endcaps (HE) which cover up to $|\eta|<3.0$, as shown in Figure 3-8. Unlike ECAL, the HCAL exploits sampling calorimeter technique, which consists of alternating layers of brass, as absorbers to degrade energy, and scintillators, as active sensors to detector signal.

The interacting hadrons also generate a cascade of secondary particles (including neutrons, protons, neutral and charged pions, etc) through electromagnetic and strong interactions. The interaction length λ, defined as the mean free path between two inelastic nuclear collisions, is conventionally used to characterize the size of hadronic

showers. It needs about 10-11 λ to absorb showers with energy up to $\sim$1 TeV. However, limited by the solenoid, the HCAL barrel only has thickness ranging from 5 to 10 λ with increasing $|\eta|$. Therefore, a tail catcher called HCAL outer (HO) is installed to complement the hadronic shower absorption. The optical signals detected by scintillators are read out via Hybrid Photodiodes (HPDs).

In addition, HCAL forward calorimeters (HF) extend the coverage to $|\eta|<5.0$, and are essential for forward jet physics, trigger studies and luminosity determination, which will be discussed in Section 6.1 with more details. HF utilizes steel absorbers and quartz fibers as active material and functions as a Cerenkov light detector.

3.2.4 Muon detectors

The muon detectors are located outside the solenoid. To be compatible with the shape of the solenoid, the muon detectors also have cylindrical barrel (MB) covering $|\eta|<1.2$, and planar endcaps (ME) covering $0.9<|\eta|<2.4$. A simplified sketch of a quadrant of the CMS detector with emphasis on the muon detectors is illustrated in Figure 3-9. As can be seen, the 4 muon stations [2] are intersected with three layers of magnet flux-return yokes. Each station is further segmented into 12 ϕ segments in barrel and several rings in endcap. Since the volumes of those stations are large, the applied technique must be relatively cheap. Therefore, three different types of gas ionization chambers are exploited in muon systems, drift tubes (DTs) and cathode strip chambers (CSCs) as tracking detectors and resistive plate chambers (RPCs) as triggering detectors.

DT and CSC

The DTs and CSCs are used to determine the trajectories of muons by measuring (r,ϕ) or (r,z) coordinates in different layers of chambers. The CSCs, which are easier to finely segment and have a faster response compared to the DTs, are utilized in the endcap region, where the particle rates are higher and the magnetic fields are

[2]The station is defined as a chamber union with same r-coordinate in barrel and same z-coordinate in endcap.

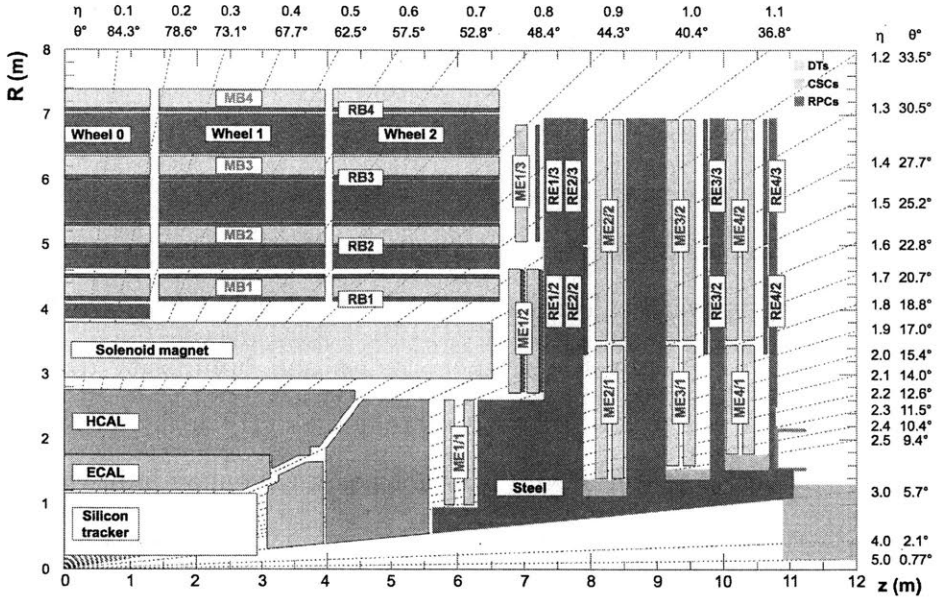


Figure 3-9: An R-z cross section of a quadrant of the CMS detector [19].

non-uniform.

The building block of the DT is a cuboid drift cell, as illustrated in Figure 3-10. It is composed by an anode wire in the central and four cathode strips, one above, one below and two on sides. Muons and other charged particles passing by ionize the gas and produce electrons, which then drift to the anode wire and the corresponding drift times are used to determine the position. As shown in 3-10, 4 layers of such parallel drift cells are stacked to make a Superlayer (SL). Neighbor layers are shifted by half cell to avoid detection gaps. Then, 2 parallel SLs, allowing (r,ϕ) measurements, and 1 perpendicular SL, allowing (r,z) measurements, are grouped to construct an arc-shaped DT chamber. As can be seen in Figure 3-9, the longitudinal size of DT chambers are same, 2.5 m as the length of a return yoke wheel. The transverse sizes vary from 1.9 to 4.1 m given different stations.

The multi-wire proportional counter is the prototype for CSC. The ionization mechanism of a multi-wire proportional counter is displayed in Figure 3-11. Each layer of counters consists of 80 radial cathode strips, each covering about 3 mrad, and thus provide a very accurate measurement of ϕ, which could describe accurately the bending trajectories of endcap muons. To reduce the number of readout channels, the

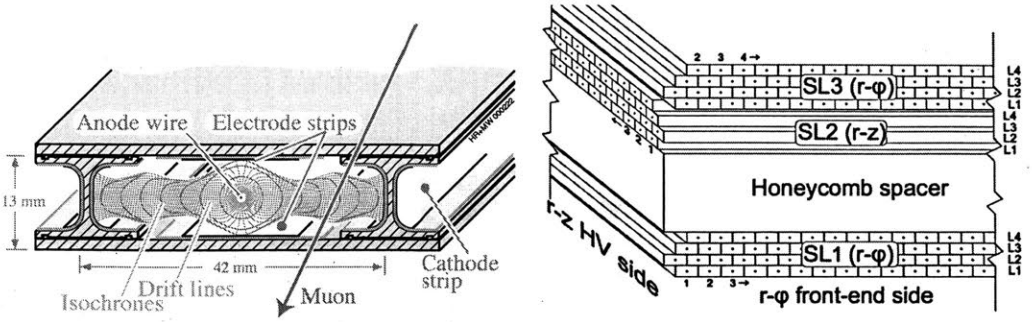


Figure 3-10: A drift cell (left) and a DT chamber (right) [20]

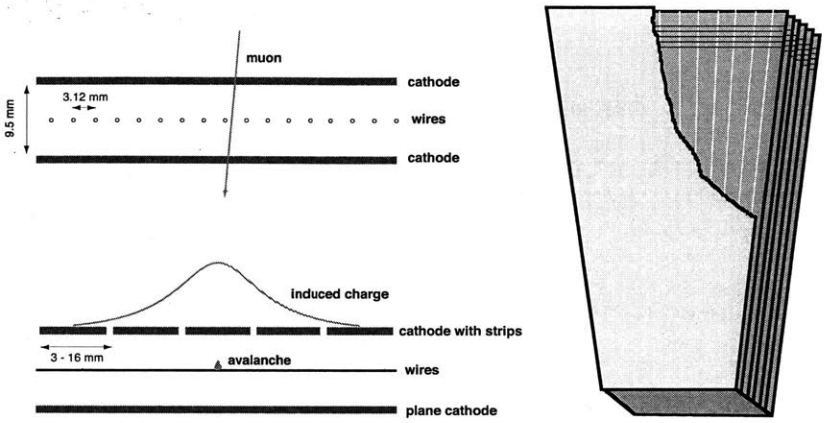


Figure 3-11: A multi-wire proportional counter and corresponding ionization mechanism (left) and a CSC chamber (right) [20]

anode wires are grouped every 5 to 16 channels, which results a coarser measurement of r. 6 layers of such finely segmented strips and wires are placed together to create a sector-like CSC chamber. The length of a CSC chamber is 1.7 to 3.4 m in radius direction, while the azimuthal coverage is 10 to 20 degrees.

RPC

Figure 3-12 illustrates a double-gap RPC. Each gap is constituted of two Bakelite plates and a gas gap, and the readout strip is positioned between the two gaps. Crossing muons and other charged particles produce avalanche electrons in the gas gap and induce signals on readout strips. It has excellent time resolution (a few ns) to distinguish ionizations in two consecutive bunch crossings. Thus, RPC is exploited as a complementary fast muon triggering detector.

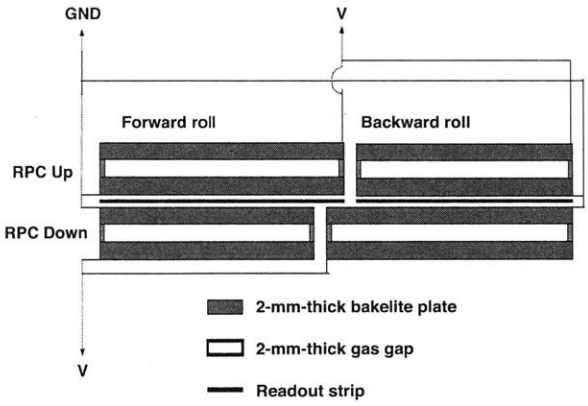


Figure 3-12: A barrel RPC [20].

Chamber type	Unit size	Gas composition	Operating voltage
DT	Drift cells 42 × 13 mm^2	85% Ar+15% CO_2	Anode: +3600 V Cathodes on sides: -1200 V Cathodes above and below: +1800 V
CSC	Cathode strips: 2.2-4.7 $mrad$ Anode wires spacing: 3.12/3.16 mm	50% CO_2+40% Ar+10% CF_4	3600 V
RPC	double gas gap: 2 mm	95.2% $C_2H_2F_4$+4.5% C_4H_10+0.3% SF_6	9600 V

Table 3.4: Technical details for each simple operation unit of DTs, CSCs and RPCs, including size, gas composition and operation voltage.

As shown in 3-9, 6 layers of RPCs (RB) are assembled in barrel region, 2 for first and second station, attaching on the innermost and outermost layer of DT, and 1 for third and last station, attaching on the innermost layer of DT only, to ensure at least 4 hit measurements even for low-p_T muons, which could be stopped before reaching far. 4 layers of RPCs (RE) are installed in the endcap region.

Table 3.4 summarizes the detailed descriptions of the DTs, CSCs and RPCs, including size of units, gas compositions and operation voltages.

3.2.5 Trigger and data acquisition system

The LHC collides proton bunches at an extremely high rate of 400 MHz. Hard scattering processes of physics interests only exist in a small fraction of those collisions, and the limited storage resources only allow to record a small fraction as well. The CMS trigger [95] and data acquisition (DAQ) system [96] is designed to make online

decisions to select interesting events and handle unprecedented event rate and data size.

The trigger system has two levels: hardware-based level 1 (L1) and software-based high-level trigger (HLT). At L1, detector signal patterns that are compatible with interesting objects, such as energy clusters in calorimeters and ionization deposits in muon detectors, create trigger primitives, which are then processed in the corresponding regional and the global triggers to make a L1 decision. The L1 selects events at a rate of 100 kHz at most, which is limited by the readout electronics. After L1 accept, the subdetector readouts are sent to the front-end driver (FED) and interfaced with DAQ system and HLT farm, where objects are further identified with the reconstruction algorithms. The HLT selects events at a rate of 400 Hz. Data of accepted events is then compressed and wrote to disk by storage manager.

Chapter 4

Event Reconstruction

Event reconstruction is the procedure to construct physics objects and determine relevant kinematic quantities from the raw signals collected by the detector. The CMS experiment exploits particle-flow (PF) algorithm [21], which first reconstructs simple PF elements such as tracks and calorimeter clusters within an individual subdetector, and then combines all subdetector information to identify particles including muons, electrons, photons, neutral and charged hadrons. Finally, on top of the reconstructed particles, high-level physics objects used in physics analyses such as jets and missing transverse momenta are further built.

According to the previous logic, this chapter is organized as follows. Reconstruction of the PF elements is first discussed in Section 4.1 to 4.3, for tracks from inner tracker detectors, muon detectors, and clusters from calorimeters, respectively. After the particle identification using PF method is described in Section 4.4, the jet and missing transverse momenta reconstruction and identification are elaborated in Section 4.5 and 4.6.

4.1 Tracks and primary vertices

The track reconstruction [15] in the CMS detector is aimed to estimate the momenta and positions of the charged particles which transverse the inner tracker detectors. It starts with local hit position reconstruction in silicon layers and construct global

tracks from hits. Based on those global tracks, the reconstruction of the primary vertices [97] is then performed to determine the positions where protons collide.

4.1.1 Hit reconstruction

As described in Section 3.2.1, charged particles crossing the inner tracking detector create electric signals. Those clustered signals after zero suppression are defined as hits. Zero suppression is achieved by requiring a minimum charge in pixel clusters and a minimum signal-to-noise ratio in strips.

Two algorithms are exploited for pixel clusters to determine the positions of the hits. The first-pass algorithm is faster and applied in seed generation and track finding steps in track reconstruction (Section 4.1.2), which uses the geometrical center of the cluster and makes correction according to the relative charge in pixels at the edges and Lorentz shift. The other algorithm [98] builds a template from simulated cluster distributions and performs template fit to determine the hit position, which is more precise and deployed in the track fitting step. Hit positions of strip clusters are determined by charge-weighted average of strip positions and corrected for Lorentz shifts.

4.1.2 Track reconstruction

Track reconstruction targets on the precise estimation of the momentum and position parameters $\{\frac{q}{p_T},\ \eta,\ \phi,\ d_z,\ d_0\}$[1] of the charged particles associated with the reconstructed hits, including not only high-p_T, prompt tracks but also low-p_T, displaced tracks. To reduce the combinatorial complexity, the *iterative tracking* method is applied. The idea is to do track reconstruction several times, and for each iteration, hits used by previously reconstructed tracks are masked. Iterative tracking usually starts with prompt, high-p_T tracks by requiring tighter d_z, d_0 and p_T cut in the first iterations, and then loose selections to reconstruct low-p_T and non-prompt ones. The exact selection requirements can be found in Ref. [15].

[1] d_z, d_0 are the longitudinal and transverse impact parameters, which is defined as the distance between the primary vertex to the point of closest approach.

Track reconstruction in each iteration consists of four steps: seed generation, track finding, fitting and selection.

Seed generation provides an initial coarse estimation of aforementioned parameters. To extract five parameters, either three 3-dimensional hits or two 3-dimensional hits plus a constraint with respect to the primary vertices[2] is needed, depending on the iteration and corresponding reconstruction targets. Seed generation starts from most inner layer outwards concerning the channel occupancy and efficiency reasons.

Track finding starts from existing seeds and adds hits based on Kalman filter technique [99] from other layers and update the parameters. It first determines the next layers that the trajectory can transverse and searches for compatible modules and hits, then updates parameters for each compatible hits. This procedure is repeated until the track reaches the last tracker layer or other termination condition [15] is satisfied.

Track fitting involves refitting the known trajectories from the track finding step, using Kalman filter and smoother, to extract the best possible estimation of tracking parameters.

Track selection is used to filter the fake tracks, which are defined to be not associated with charged particles, and assign tracks with quality labels: *loose*, *tight* and *high-purity*. The selection variables include the number of layers in which the track has hits and d_z, d_0 with respect to pixel vertices, etc. High-purity tracks are usually the ones deployed in physics analysis.

4.1.3 Primary vertex reconstruction and selection

The reconstruction of primary vertices is based on the reconstructed track collections and aimed at determining the positions of all proton-proton interactions for each bunch crossing, including the one containing hard processes of physics interests, as well as other pile-up interactions.

[2]Those primary vertices are reconstructed by a very fast algorithm purely based on pixel, which are usually called pixel vertices.

Primary vertex reconstruction starts with the selection of the reconstructed tracks, according to three conditions: their compatibility with the beam spot, number of hits involved and fit quality. Then, the tracks are clustered into different primary vertex candidates using *deterministic annealing* (DA) algorithm [100], which assigns a number for the possibility that one track belongs to one vertex and optimizes a free energy like function to determine the primary vertex candidates. Finally, a 3-dimensional fit is performed to compute the best estimation of the vertex position parameters.

In physics analyses, to select the primary vertex of hard scattering process, reconstructed vertices are required to be

- the longitudinal distance to interaction point: $|z|<24$ cm,
- the transverse distance to interaction point: $r<2$ cm,
- the number of degree of freedom of the vertex fit: $n_{dof}>4$,

Out of the selected vertices, the one with the largest $\sum_{tracks} p_T^2$ is chosen as *hard-scatter vertex* in the event. The vertices other than hard-scatter vertex are considered as *pile-up vertices.*

4.2 Calorimeter clusters

The goal of the calorimeter cluster reconstruction is to measure the energies and positions of the energy deposits in the CMS calorimeters. Local hit reconstruction determines hit energy per cell by performing template fits to the pulse shapes [101]. Then the clustering algorithms are performed separately for each subsystems and involve three steps. First, the cells with energy larger than adjacent cells and location-dependent thresholds [21] are selected to be seeds. Then nearby seeds are integrated to define a topological cluster. Finally, the positions and energies of those topological clusters are determined by an iterative expectation maximization algorithm based on Gaussian-mixture models and maximum likelihood fits [21].

Cluster energies determined in this way need to be calibrated in order to account for the energy loss due to thresholds. The calibration is performed using simulation

and assuming specific analytical functions between true energy and measured energy in each subsystem. The ECAL cluster calibration is done first, using the simulated single photon samples without conversion in trackers. With ECAL calibrated, the HCAL cluster calibration is then performed in a similar way using the simulated single neutral hadron samples without interaction in tracks.

4.3 Muons

The muon track reconstruction [19] in the CMS detector targets on the identification of muons and precise measurements of muon momenta. It starts from local reconstruction of hits and segments in a single DT, CSC and RPC chamber, then builds tracks independently in the muon systems and further combines the inner tracker information (Section 4.1) to complete muon track collections.

This subsection will also discuss the high-level muon objects deployed in physics analyses, which are usually built on top of the reconstructed muon tracks by applying additional criteria, in order to enhance the desired prompt muons. Those types of criteria are known as the muon identification and isolation requirements.

4.3.1 Hit and segment reconstruction

Hit reconstruction

As described in Section 3.2.4, muons passing by an active unit in the muon detector ionize gas and create electronic signals in the corresponding readout channel, associated to a precisely measured position of (r,ϕ) or (r,z). These digitized electronic signals are defined as hits in muon detectors. According to detection technology, DT, CSC and RPC exploit different methods [102] for the hit position reconstruction, which are summarized as follows.

- **DT drift cell**. Hit position = $(T_{TDC} - T_{ped}) \times v_{drift}$, where T_{TDC} is the arrival time registered by time-to-digital converter (TDC), T_{ped} is the correction for the

pedestal time, from bunch crossing to the arrival time of the trigger decision (L1Accept), and v_{drift} is the drift velocity of electrons.

- **CSC layer**. Hit position is reconstructed at the intersection points of active strip and wire groups.

- **RPC layer**. Hit position is reconstructed at the strip cluster centroid, as the induced charge can be shared by more than one strip.

Segment reconstruction

As illustrated in Figure 3-10 and 3-11, a DT chamber has 12 layers and a CSC has 6 layers. Thus, groups of the reconstructed hits from each layer in a single chamber create straight-line track stubs, which are defined as segments. To reconstruct segments from hits, pattern recognition algorithms and fit algorithms [103] are exploited to associate hits into groups, by matching straight-line parameters for a candidate muon producing those hits.

4.3.2 Muon track reconstruction

Three collections of muon tracks are commonly used as basic muon objects in analyses: standalone muon, global muon and tracker muon tracks. Standalone muon tracks are first reconstructed with muon detector information only.

- **Standalone muon tracks**. Based on Kalman filter technique [99], reconstructed segments are utilized as seeds. The muon track is then built using all DT, CSC, RPC segments and hits along the trajectory as input with an iterative algorithm which updates the trajectory parameters at each step.

Then, together with the inner tracker tracks as input, the other two high-level muon objects can be further reconstructed.

- **Global muon tracks**. For each "standalone muon track - tracker track" pair, a combined fit using all hits in both tracks is performed based on Kalman filter

technique. The best-matching tracker track is then selected for each standalone muon track, following the "outside-in" approach.

- **Tracker muon tracks**. This collection is designed to complement the global muon collection. The "inside-out" approach is applied in order to consider all tracker tracks to be potential muon candidates and "tracker track - muon segment" pair matching is performed. The efficiency of muon reconstruction is very high, about 99% muons inside the muon detector acceptance are reconstructed as either global muon or tracker muon.

To precisely determine the muon momentum of the reconstructed tracks, two algorithms are utilized. The Tune-P algorithm [104] selects the best p_T measurement from several refits [3] based on goodness-of-fit and resolution requirements. The PF algorithm exploits the full event information, selects refits that improve the balance of missing transverse momentum.

4.3.3 Muon identification and isolation

There are generally four sources of muons and muon-like objects after muon track reconstruction: decays of W, Z and Higgs bosons, decays of heavy-flavor hadrons, decays of light hadrons (decay in flight) and hadronic punch-through (not muons but charged hadrons). Various identification (ID) and isolation variables are studied to target on different muons with required purity and efficiency. In this section, the two sets of ID selections and two isolation variables, which will be heavily exploited in the analyses reported in Chapter 5, 6 and 7, are described.

Identification

Following the CMS convention, the two sets of IDs which are used as baselines for analyses are called Tight ID and High-p_T ID [105]. Both of them are based on

[3] Including *Inner-Track fit* which uses tracker hits only, *Tracker-Plus-First-Muon-Station fit* which uses tracker and the innermost muon station hits, etc.

global muon tracks with additional muon quality requirements, and have a muon identification efficiency of 96 - 98%, depending on detector region.

Tight ID is very widely used in CMS physics analyses, aimed to select muons from prompt decay of W, Z and Higgs bosons and disfavor muons from hadronic punch-through, decay in flight, cosmics and pile-up vertices. On top of being a global muon, the additional muon quality requirements and their logics are summarized as follows.

- *At least 6 tracker layers have hits*, to guarantee good momentum measurement as well as suppress muons from decays in flight.
- *At least 1 pixel layer has hit*, to further suppress muons from decays in flight.
- *At least 2 muon stations' segments are matched to tracker track*, to suppress muons from punch-through and accidental "tracker track - muon segment" pair matching.
- *At least 1 hit from muon system is included in global fit.*
- *Global fit quality requirement* ($\chi^2/ndof < 10$).
- *Transverse impact parameter* $|d_0| < 2$ *mm*, to suppress cosmic muons and further suppress muons from decays in flight.
- *Longitudinal impact parameter* $|d_z| < 5$ *mm*, to further suppress cosmic muons, muons from decays in flight and tracks from pile-up vertices.

High-p_T ID is designed for higher efficiency of identifying high-p_T muons with p_T > 200 GeV. Since the probability of electromagnetic radiation are enhanced when high-momentum muons transverse the iron return yokes in muon detectors. The electromagnetic showers can create additional hits and segments, resulting energy losses and causing inaccurate momentum measurements. Compared to Tight ID, the differences are:

- *Removal of global fit quality requirement*, to increase efficiency for high p_T muons, as they radiate more often and generate additional hits in muon stations.
- *Add requirement on relative* p_T *uncertainty:* $\sigma(p_T)/p_T < 30\%$, to ensure a proper momentum measurement.

Working points	Loose	Tight
PF-based isolation	0.25	0.15
Tracker-based isolation	0.10	0.05

Table 4.1: Muon isolation requirements

Isolation

Isolation requirements are implemented to further enhance prompt muons and suppress those in jets, by constraining the energy surrounding the muon. Two isolation variables are studied, one is PF-based, the other is tracker-based.

PF-based isolation sums the charged hadrons, neutral hadrons and photons in the cone size of ΔR=0.4, and corrects for the contribution from pile-up vertices:

$$I^{PF} = \sum p_T^{charged-hadrons} + max(0, \sum E_T^{neutral-hadrons} + \sum E_T^{photons} - 0.5 \times \sum p_T^{pile-up}) \tag{4.1}$$

where 0.5 is determined from simulation as the ratio of charged components in the pile-up contribution.

Tracker-based isolation sums the tracker tracks in the cone size of ΔR=0.3:

$$I^{tracker} = \sum p_T^{tracker-tracks} \tag{4.2}$$

Tight and loose working points for the two isolation variables relative to p_T are summarized in Table 4.1, which are chosen to achieve 95% and 98% selection efficiency, respectively.

4.4 Particle identification with PF algorithms

Once PF elements are reconstructed, a linking algorithm [21] based on pairs of shortest distance in (η, ϕ) plane is exploited to connect elements from all sub-detectors to form PF blocks. In each PF block, particles are identified and reconstructed in the following order: muons, electron and isolated photons, hadrons and non-isolated photons. The

four momentum of each particle is also determined based on the vectorial momentum of tracks and the energy and position of the calorimeter clusters. Once a particle is identified, the associated elements are masked for particle identification afterwards.

Figure 4-1 illustrates the idea of particle identification using inputs from PF blocks. Muons are first identified and reconstructed, with algorithms described in Section 4.3. Then, electron candidates are built from links of tracks[4] and ECAL clusters, while isolated photon candidates are seeded from ECAL clusters without linking to tracks. Lastly, hadrons and non-isolated photons are reconstructed from remaining PF blocks. For a joint ECAL-HCAL cluster not linked to any track within tracker acceptance, the non-isolated photons are first reconstructed based on the ECAL cluster, then the HCAL cluster builts neutral hadrons. Beyond tracker acceptance, such a cluster is identified simply as hadrons, either charged or neutral. Among the remaining HCAL clusters, presence of charged, neutral hadrons and photons depends on the comparison between energy determined from clusters and tracks. For example, if they are compatible, it suggests no neutral particles and presence of charged hadrons.

4.5 Jets

By the nature of QCD, quarks and gluons produced in the detector develop parton showers and undergo hadronization. The fragments from hadronization tend to travel in the same direction, thus form a cone of hadronic and other particles, which is called a jet. The purpose of jet identification and reconstruction includes constructing jet from PF particles and determining the jet four momentum (*jet clustering algorithm*), distinguishing physics jet from detector noise (*noise jet ID*), differentiating heavy-flavor and light-flavor jets (*heavy-flavor jet ID*), etc.

[4]For electron tracking, a Gaussian-sum filter (GSF) [106] is used to fit the trajectory instead of Kalman fitter (KF) in the usual inner track reconstruction, as it allows sudden and substantial energy losses.

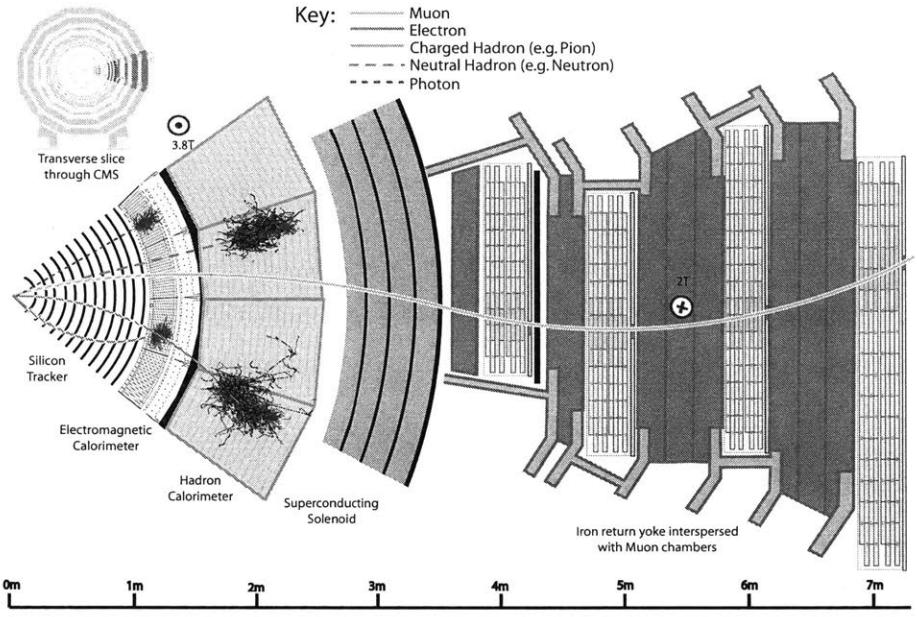


Figure 4-1: Illustration of particle interaction with detector, for muons, electrons, charged hadrons, neutral hadrons and photons, respectively [21].

4.5.1 Jet clustering algorithm

In a common jet, 65% of its energy is carried by charged hadrons, 25% by photons and 10% by neutral hadrons. Jet reconstruction starts from clustering those PF particles in the defined jet cone. The jet clustering algorithm used in analyses covered by this thesis is called *anti-k_T* algorithm [107], which is described as follows.

- Define $d_{i,j}$, distance between PF particle i and j, and $d_{i,B}$, distance between PF particle i and beam (B),

$$\begin{aligned} d_{i,j} &= min(k_{ti}^{2p}, k_{tj}^{2p})\frac{\Delta_{ij}^2}{R^2}, \\ d_{i,B} &= k_{ti}^{2p}, \end{aligned} \tag{4.3}$$

  where k_t is the transverse momentum, $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, $R = 0.4$ is the radius of the jet cone and $p = -1$ as a scale parameter.
- Identify the smallest distance among all $d_{i,j}$ and $d_{i,B}$.
- Combine particle i and j if it is a $d_{i,j}$. Otherwise, define particle i as a jet and remove it from next iterations.

Category	Neutral			Charged		
ID variable	Hadron Fraction	EM Fraction	Multiplicity	Hadron Fraction	EM Fraction	Multiplicity
Selection	< 0.99	< 0.99	> 1	> 0	< 0.99	> 0

Table 4.2: Jet noise ID selections.

- Repeat previous steps until no particles are left.

The jet four momentum is then determined as the sum of four momenta of all particles inside the jet. In order to mitigate the effect from pile-up, the *charged hadron subtraction* (CHS) algorithm [108] is exploited to remove the charged hadrons associated with pile-up vertices from the PF particle lists. An additional energy correction is applied assuming uniform energy density in the jet cone, to remove the remaining contribution from neutral and charged particles originating from the pile-up vertices.

The jet energy response, which is defined as the ratio of reconstructed jet energy to the reference jet energy, i.e the energy of all stable particles produced by the event generator inside the jet except the neutrinos, is not unity due to several sources, such as the non-linearity of the calorimeter responses to hadronic and electromagnetic (EM) particles inside the jets, energy thresholds in the calorimeters and p_T thresholds in the tracker. To account for this, jet energy corrections [109] are derived from simulation and applied to bring the jet energy response back to unity and remove any dependence on jet p_T and η.

4.5.2 Noise jet identification

To distinguish physical jets from calorimeter noise, a set of selections [110] are required on the PF particle multiplicity and jet energy fraction carried by different types of PF particles inside a jet, which are summarized in Table 4.2. Those identification (ID) variables are sensitive to different sources of noise from the ECAL and HCAL.

4.5.3 Heavy-flavor jet identification

Heavy-flavor jets are jets originating from b or c quarks. We define jets that contain at least one b hadron are b jets, those that contain no b hadron and at least one c hadron are c jets, and the remaining jets are called light-flavor jets or "udsg" jets. Heavy-flavor jets have many unique properties, compared to light-flavor jets. First, the lifetime of heavy-flavor hadrons are long, which results displaced tracks and secondary vertices. Secondly, the decay products tend to have a larger p_T as the b and c quark have larger masses. Thirdly, electron and muons are often present in the decay products. Therefore, corresponding discriminant variables [111] are exploited to identify b and c jets.

Especially, efficient identification of b jets are important for many measurements of standard model physics and searches for new physics, including the search for neutral Higgs boson generated in the b quark associated production, which will be discussed in Chapter 7. The combined secondary vertex (CSV) algorithm, which combines various discriminant variable, mainly based on displaced tracks and secondary vertex information, and outputs a single discriminator value, is deployed in the analysis. The algorithm can be applied to jets with $p_T > 20$ GeV and $|\eta| < 2.4$, with a efficiency of 66% and a mistag probability of 1% for the selected working points.

4.6 Missing transverse momentum

Missing transverse momentum (p_T^{miss}) is used to indicate the existence of particle that do not interact with detector material, most notably neutrinos, and is also expected to be a signature of new particles beyond Standard Model. In this section, the methods to reconstruct p_T^{miss} are described and filters to remove spurious energy also discussed.

4.6.1 p_T^{miss} reconstruction algorithm

The reconstruction of p_T^{miss} is based on the conservation of momentum in the transverse plane. The $\vec{p}_T^{miss}$ is defined to balance the vectorial sum of the transverse

momenta of all reconstructed objects:

$$\vec{p}_T^{miss} = -\sum_{i=1}^{N_{objects}} \vec{p}_{T,i}. \tag{4.4}$$

Similarly, to account for detector response as described in Section 4.5.1, the jet energy corrections are propagated to the p_T^{miss} calculation:

$$\vec{p}_T^{miss} = -\sum_{i=1}^{N_{objects}} \vec{p}_{T,i} - \sum_{j=1}^{N_{jets}} (\vec{p}_{T,j}^{corr} - \vec{p}_{T,j}). \tag{4.5}$$

4.6.2 p_T^{miss} filters

Along with genuine p_T^{miss} created from undetectable particles, fake p_T^{miss} is also introduced due to uninteresting reasons such as detector noise, cosmic rays and beam halo particles.[5] To reconstruct p_T^{miss} accurately, various filters [112] are applied at different reconstruction steps in order to minimize those fake p_T^{miss}.

HCAL filters are exploited to reduce the spurious energy due to noise in the hybride photodiode (HPD) and readout box (RBX) electronics. ECAL filters are used to eliminate energy from interaction with photodetector and other inactive materials. Machine-induced backgrounds including beam halo particles usually leave energy deposits along straight lines in the calorimeters, which are removed with beam halo filters.

[5]Beam protons might interact upstream of the detector and create beam halo particles.

Chapter 5

Z Boson Production Cross Section Measurements

The Z boson production process is crucial for LHC physics programs. The feature of the large cross section ($\sim$ nb, see Figure 1-1), clean detection signature of the leptonic decay products, and sufficiently understood theories (to NNLO QCD and NLO EW) allows the inclusive production cross section to be measured very accurately. The theoretical uncertainty on the detector acceptance correction can be further reduced by applying kinematic cuts and measuring the so-called fiducial cross sections. The precise measurements probe high-order pQCD and help quantify the order of magnitude of the not-yet-calculated contributions.

In addition to the inclusive production cross section, the large datasets collected at the LHC also permit for measurements of differential production cross sections with high precision. The measurement of the transverse momentum of Z boson (p_T^Z) is of particular importance, as it probes different aspects of the strong interaction. The low p_T^Z end behavior is dominated by initial-state radiation (ISR) and the intrinsic p_T distributions of the initial-state partons, which are modeled by soft-gluon resummations or parton shower models, while the high p_T^Z end performance is governed by quark-gluon scattering, which is described by pQCD.

The low p_T^Z domain can be alternatively probed with the $\phi_\eta^\star$ observable [113],

which is defined as:

$$\begin{aligned} \phi_\eta^\star &= \tan(\frac{\pi - \Delta\phi}{2}) \cdot \sin(\theta_\eta^*) \\ \cos(\theta_\eta^*) &= \tanh(\frac{\eta^- - \eta^+}{2}), \end{aligned} \tag{5.1}$$

where $\Delta\phi$ is the absolute value of the azimuthal difference of the two muons, and $\eta^\pm$ refers to the pseudo-rapidity of the positive and negative muons. The $\phi_\eta^\star$ is highly correlated with $p_T^Z/m_{\ell\ell}$, but less sensitive to the smearing effects since it is built entirely by angular quantities of the two muons, which in fact resulting in smaller experimental uncertainties than p_T^Z.

The measurement of rapidity distribution (y^Z) gives a direct determination of the momentum fraction of the interacting partons $x_{a,b}$, via [5]:

$$x_{a,b} = \frac{m_Z}{\sqrt{s}} e^{\pm y^Z}, \tag{5.2}$$

which can provide additional constraints on the PDFs. The transverse momentum distributions of muons in the final states ($p_T^{\mu^+}$, $p_T^{\mu^-}$) are also sensitive to resummations.

5.1 Measurement overview

In this chapter, measurements of the following Z boson production cross section quantities in the muon channel are presented:

- Inclusive total Z boson production cross section: $\sigma_{tot}^{Z\to\mu^+\mu^-}$,
- Inclusive fiducial Z boson production cross section: $\sigma_{fid}^{Z\to\mu^+\mu^-}$,
- Differential Z boson production cross sections with respect to observable $\mathcal{O}$: $\frac{d\sigma^{Z\to\mu^+\mu^-}}{d\mathcal{O}}$, where $\mathcal{O}$ is p_T^Z, y^Z, $\phi_\eta^\star$, $p_T^{\mu^+}$ or $p_T^{\mu^-}$.

5.1.1 Inclusive Z boson cross section measurements

By definition, the Z boson production cross section we want to measure is:

$$\sigma_{tot}^{Z\to\mu^+\mu^-} = \sigma_{tot}^{Z} \times BR(Z \to \mu^+\mu^-) = \frac{N_{tot}^{Z\to\mu^+\mu^-}}{\mathcal{L}}. \tag{5.3}$$

$\mathcal{L}$ is the recorded integrated luminosity by the CMS detector, for the data-taking periods which are used for the measurements. While $N_{tot}^{Z\to\mu^+\mu^-}$ is the total number of $Z \to \mu^+\mu^-$ events that the detector should have recorded if it has full geometric and kinematic coverage and 100% detection efficiency.

However, as the detector has limited coverage and imperfect efficiency, it is common to define *acceptance* and *efficiency* as follows:

$$A = \frac{N_{fid}^{Z\to\mu^+\mu^-}}{N_{tot}^{Z\to\mu^+\mu^-}}, \tag{5.4}$$

$$\epsilon = \frac{N_{sel}^{Z\to\mu^+\mu^-}}{N_{fid}^{Z\to\mu^+\mu^-}}. \tag{5.5}$$

$N_{fid}^{Z\to\mu^+\mu^-}$ denotes the number of $Z \to \mu^+\mu^-$ events that the detector should have recorded inside the predefined fiducial region if with 100% detection efficiency, while $N_{sel}^{Z\to\mu^+\mu^-}$ denotes the number of $Z \to \mu^+\mu^-$ events that are actually triggered, reconstructed and selected in this region.

Moreover, not only true $Z \to \mu^+\mu^-$ events, $N_{sel}^{Z\to\mu^+\mu^-}$, would pass the selection requirements, other processes that have or fake to have the same $\mu^+\mu^-$ final state can also pass the criteria and become *backgrounds*, N_{sel}^{bkg}, which have to be subtracted from the observed Z-like events, $N_{obs}^{Z\to\mu^+\mu^-}$, in order to determine the correct cross section:

$$N_{sel}^{Z\to\mu^+\mu^-} = N_{obs}^{Z\to\mu^+\mu^-} - N_{sel}^{bkg}. \tag{5.6}$$

Finally, based on Equation 5.3 - 5.6, we reach the master formula to calculate the

total and fiducial Z cross section measurements:

$$\sigma_{tot}^{Z\to\mu^+\mu^-} = \frac{N_{obs}^{Z\to\mu^+\mu^-} - N_{sel}^{bkg}}{A\cdot\epsilon\cdot\mathcal{L}}, \tag{5.7}$$

$$\sigma_{fid}^{Z\to\mu^+\mu^-} = \sigma_{tot}^{Z\to\mu^+\mu^-}\cdot A = \frac{N_{obs}^{Z\to\mu^+\mu^-} - N_{sel}^{bkg}}{\epsilon\cdot\mathcal{L}}. \tag{5.8}$$

As can be seen in Equation 5.7 and 5.8, the major components for inclusive measurements are clearly divided. In the measurements, observed Z-like events, $N_{obs}^{Z\to\mu^+\mu^-}$, are counted from data, background contributions N_{sel}^{bkg}, acceptances A and efficiencies ϵ are estimated using simulation, with data-to-simulation single muon efficiency scale factors applied to account for the muon efficiency differences between data and simulation. Measurements of integrated luminosity ($\mathcal{L}$) is an independent and established project in CMS. Its basics are described in Section 6.1 and more details can be found in Ref. [114].

5.1.2 Differential Z boson cross section measurements

Straightforwardly, for differential cross section measurements, Equation 5.8 can be applied to each differential bin of the observable under study. For instance, to measure the p_T^Z distribution, individual yields in each Δp_T^Z bin can be corrected by corresponding efficiency evaluated in that specific fiducial region, $\epsilon(\Delta)$, i.e doing a bin-by-bin correction:

$$\frac{\Delta\sigma^{Z\to\mu^+\mu^-}}{\Delta p_T^Z} = \frac{\Delta N_{obs}^{Z\to\mu^+\mu^-} - \Delta N_{sel}^{bkg}}{\epsilon(\Delta)\cdot\mathcal{L}\cdot\Delta p_T^Z}. \tag{5.9}$$

However, the bin-by-bin approach is not correct when the amount of event migrations from a bin to others is not negligible, considering there exists smearing effects due to limited detector resolution. An alternative approach, unfolding based on Bayes' theorem [115], by building a response matrix in order to unfold the detector level information to the particle level information, is deployed. In our measurements, the response matrices are estimated from simulation, with data-to-simulation single muon efficiency scale factors applied as well.

The rest of this chapter is organized as follows. First, the data and simulated sam-

ples used in the measurements are described in Section 5.2. After Z event selection is briefly discussed in Section 5.3, the fiducial region definition and acceptance results for the inclusive Z cross section measurements are presented in Section 5.4. Efficiency measurements are elaborated and corresponding scale factors are documented in Section 5.5. Section 5.7 describes the signal extraction for inclusive measurements and Section 5.8 discusses the unfolding technique exploited in differential measurements. Finally, after systematics estimations are summarized in Section 5.9, the Z cross section measurement results are presented in Section 5.10.

5.2 Data and Simulated Samples

5.2.1 Data samples and triggers

The data used in the measurements are taken from 25 ns bunch crossing conditions of proton-proton collisions at center-of-mass energy of $\sqrt{s}$=13 TeV, collected from September to November in 2015 data-taking period (see Figure 5-1), which corresponds to an integrated luminosity of 2.3 fb^{-1} [114].

The Z candidate events are selected from all events collected by the lowest p_T threshold, non pre-scaled isolated single muon trigger. For 2015 data-taking period, it is *HLT_IsoMu20*, which requires at least one muon with $p_T > 20$ GeV and $|\eta| < 2.4$, and passing loose isolation and identification requirements.

5.2.2 Simulated samples

Simulated signal samples are generated to model the Z boson production and decays into muon pairs. This is done with two NLO generators: aMC@NLO [77, 78] as the nominal one to estimated the acceptance and response matrices, and POWHEG [79–82] as an alternative to evaluate the modeling systematics. The NNPDF 3.0 [116] is used as PDF input. The simulated background samples are also used to determine the normalization of the background contamination, which is about ~0.5% and mainly comes from top pair and diboson productions such as $W^{\pm}W^{\mp}$, $W^{\pm}Z$ and

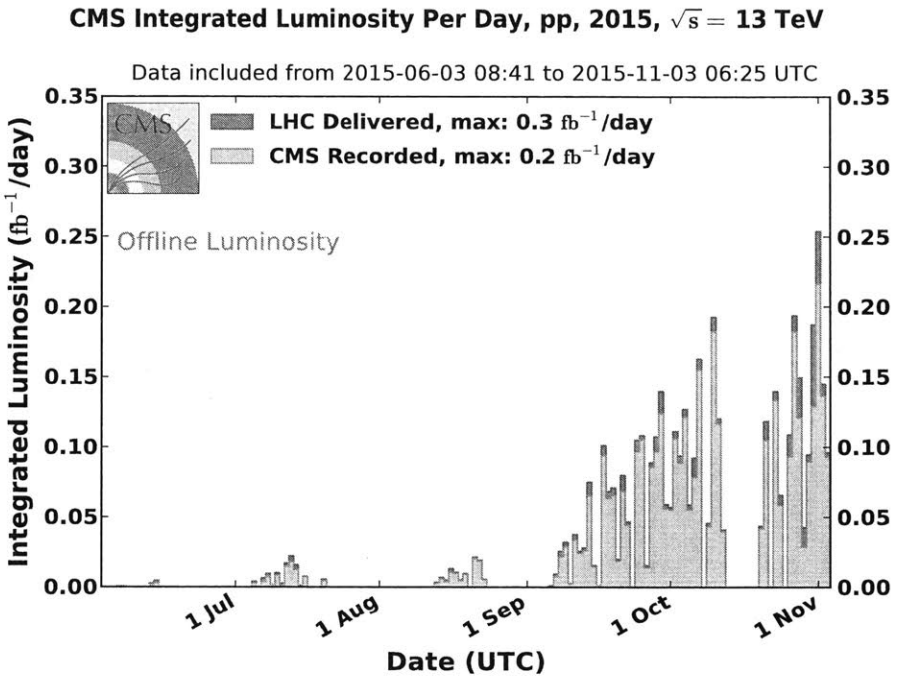


Figure 5-1: Day-by-day integrated luminosity of 2015 data-taking period [10].

ZZ. Those simulated samples are generated with POWHEG [79–82] and PYTHIA [71]. Descriptions of those generators can be found in Section 2.1.2. Additional proton-proton interactions per bunch crossing (pile-up effects) are also considered in the simulated events, by reweighting the pile-up distribution of generated events to match that observed in data.

5.3 Fiducial region and Z event selection

In the inclusive Z cross section measurements, the fiducial region is defined with cuts on the transverse momenta and pseudo-rapidities of the two muons:

- $p_T^{\mu} > 25$ GeV,
- $|\eta^{\mu}| < 2.4$.

The choice of p_T threshold is to avoid the trigger turn-on inefficiency region, while the $|\eta|$ is set to be consistent with the acceptance of the muon detectors. Depending on whether the muon momenta are evaluated before or after the final state radiation (FSR), the fiducial cross sections are measured with respect to both *dressed* muons

(pre-FSR) and *naked* muons (post-FSR). Additionally, $60 < M_Z < 120$ GeV is used to define the Z cross sections.

With above definitions, the Z event selection requires presence of at least two opposite-charged muons with $p_T > 25$ GeV and $|\eta| < 2.4$, and have an invariant mass with in the mass window of $60 < M_{\mu\mu} < 120$ GeV. The muons are selected with Tight ID and Tight PF-based isolation requirements as described in Section 4.3.3 to enhance the prompt muons from decay of Z bosons.

5.4 Acceptance

Acceptance is estimated from simulation according to the definition in Equation 5.4. The nominal signal sample, i.e aMC@NLO with NNPDF3.0, is exploited. Based on the fiducial region definition in Section 5.3, the following ratio is evaluated:

$$A_{MC} = \frac{N_{MC}^{Z\to\mu^+\mu^-}(p_T^\mu > 25GeV, |\eta^\mu| < 2.4, 60 < M_Z < 120GeV)}{N_{MC}^{Z\to\mu^+\mu^-}(60 < M_Z < 120GeV)}. \tag{5.10}$$

For *naked* level acceptance, the post-FSR particle level information is used. While for *dressed* level acceptance, the p_T^μ is calculated by adding back the four-momenta of photons that are not from hadronic decays and inside the cone $\Delta R = 0.1$ around each muon.

5.4.1 Acceptance uncertainty evaluation

The sources of uncertainty on acceptance include: the PDF uncertainty, higher-order pQCD and soft QCD corrections, and higher-order EW corrections.

PDF uncertainty mainly concerns the uncertainty on α_S and the choice of heavy-flavor quark masses. Its impact is estimated by re-calculating the acceptance using PDF replicas [1], following the NNPDF 3.0 collaboration recommendations [116].

Uncertainties due to missing higher-order QCD corrections are divided into two components. First, ResBos [65] and DYRes [66], which provides accuracy at NNLO

[1] a set of PDF created from sampling technique.

Source	PDF	NNLO+NNLL	>NNLO	FSR	EW	Total
Uncertainty	0.7%	0.9%	0.6%	0.3%	0.4%	1.4%

Table 5.1: Summary of the theoretical uncertainty on acceptance.

naked level acceptance	0.366±0.005
dressed level acceptance	0.372±0.005

Table 5.2: Acceptance results from simulation.

for matrix element calculation and NNLL for resummation, are used to compute the acceptance. The envelop of the differences with respect to the nominal value is taken as one uncertainty component. Second, the missing QCD effects beyond NNLO are estimated by scale variation (μ_F=μ_R=M_Z,$2M_Z$,$M_Z/2$) using FEWZ [61]:

$$\sigma_{scale} = \frac{1}{2} max\left[|A_{M_Z} - A_{2M_Z}|, |A_{M_Z} - A_{M_Z/2}|, |A_{2M_Z} - A_{M_Z/2}|\right], \tag{5.11}$$

where A_μ is the acceptance by setting the scales to μ.

Uncertainties from high-order EW corrections are also considered in two aspects: the FSR modeling and EW virtual corrections. To evaluate, HORACE generator [117], which enables options to include FSR correction only (*HORACE-FSR-only*) and also loop corrections (*HORACE-all*), is exploited. The uncertainty on FSR modeling by PYTHIA in the nominal sample is estimated by comparing the nominal acceptance value to that calculated from HORACE-FSR-only. The size of virtual corrections is quantified by comparing acceptances between HORACE-FSR-only and HORACE-all.

Uncertainty results are summarized in Table 5.1. As can be see, the uncertainties from all sources are small, which indicates the aMC@NLO generator already provides a good inclusive description for QCD and EW effects in the Z boson production. Therefore, the higher-order effects are treated as uncertainties instead of additional corrections to the nominal acceptance value. To conclude, the nominal acceptance values and the uncertainties are reported in Table 5.2.

5.5 Efficiency

The Z event efficiency is defined in Equation 5.5, i.e, the fraction of Z events that are actually triggered, reconstructed and selected in the fiducial region. In both inclusive and differential Z cross section measurements, the Z event efficiencies are estimated from simulation, with data-to-simulation scale factors applied to account for the trigger, reconstruction and selection efficiency difference between data and simulation. In this section, we will first describe the methodology of measurements for single muon trigger, reconstruction and selection efficiencies, present corresponding results and systematics, then discuss how to integrate them into Z event efficiency.

The total efficiency for a muon in the selected Z event can be factorized into three efficiency types, considering all potential detection inefficiencies in triggering, reconstruction (in inner tracker (*tracking*) and muon detectors (*standalone*)) and selection (via identification (*ID*) and isolation (*Iso*) requirements) steps:

$$\epsilon_{\mu} = \epsilon_{trigger} \times \epsilon_{tracking+ID+Iso} \times \epsilon_{standalone}, \tag{5.12}$$

where efficiencies in each type can be measured using the tag-and-probe method [118], which is a generic technique to measure self-defined object efficiency in a unbiased way by exploiting di-object resonances such as Z boson or J/ψ.

5.5.1 Tag-and-probe method

Based on tag-and-probe approach, the steps to measure single muon efficiencies for a specific efficiency type are summarized as follows.

1. Build di-muon resonance with one leg passing tight selection (*tag*) and the other passing a loose selection (*probe*).
2. Divide resonances into *passing* and *failing* category, depending on whether the probe is passing required criteria.
3. Fit di-muon invariant mass distribution $M_{\mu\mu}$ simultaneously for passing and failing category to extract the signal yields N^{sig}_{pass} and N^{sig}_{fail}.
4. The efficiency is then determined by $\epsilon = \frac{N^{sig}_{pass}}{N^{sig}_{pass}+N^{sig}_{fail}}$.

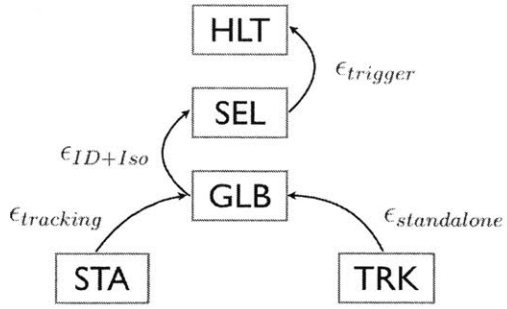


Figure 5-2: Muon object efficiency steps.

Tag and probe selection

For all efficiency types, the tag selection is to be as tight as possible, i.e, passing Tight ID and Tight PF-based isolation, and matching to trigger object, to ensure a pure and unbiased Z event sample. The probe and passing probe selections depend on the specified efficiency type. Considering the construction step of a muon object, the probe selection of current step is always the passing probe selection of previous step. To clarify, we define the following abbreviations and illustrate the efficiency steps in Figure 5-2.

- **TRK**: a tracker muon (defined in Section 4.3.2),
- **STA**: a standalone muon (defined in Section 4.3.2),
- **GLB**: a global muon (defined in Section 4.3.2),
- **SEL**: a selected muon, i.e, global muon passing ID and isolation requirements,
- **HLT**: a selected muon also matching to the trigger object.

Therefore, it is straightforward to see that the probe is TRK and passing probe is GLB for $\epsilon_{standalone}$. For $\epsilon_{tracking+ID+Iso}$, the probe is STA and passing probe is SEL. For $\epsilon_{trigger}$, the probe is SEL and passing probe is HLT.

(p_T, η) binning of probe

To account for the kinematic dependence of efficiencies and corresponding scale factors, the efficiencies are measured in defined (p_T, η) binning:

- p_T: {25 GeV, 40 GeV,∞},
- η: {-2.4, -2.1, -1.2, -0.9, -0.3, -0.2, 0, 0.2, 0.3, 0.9, 1.2, 2.1, 2.4}.

The p_T-dependence of efficiencies are small, so that only two p_T bins are applied. Instead, there are more η-dependence due to the muon detector geometry of barrel ($0<|\eta|<1.2$) and endcap ($0.9<|\eta|<2.4$), as well as the feature of less instrumentation between the central muon wheel and the two adjacent wheels at $0.2<|\eta|<0.3$ [19].

Fit model of signal and background

As described previously, simultaneous fits with "signal + background" models are performed to the invariant mass distributions for passing and failing category, in order to subtract the backgrounds that are not from the Z resonance. This is done when the background contribution is not negligible, which is the case for $\epsilon_{tracking+ID+Iso}$ and $\epsilon_{standalone}$ in data. Otherwise, for $\epsilon_{trigger}$ and all efficiencies measured in simulation (using DY signal samples only), simple counting is performed.

The two signal models exploited in efficiency measurements are denoted as *MC*Gaus* and *BW*CB*, which are both physically well motivated. *MC*Gaus* is the convolution of the invariant mass template from simulation and a Gaussian function:

$$f_{MC*Gaus}(M_{\mu\mu}) = f_{MC}(M_{\mu\mu}) * exp(\frac{(M_{\mu\mu} - M_0)^2}{2\sigma_M^2}), \tag{5.13}$$

where the $f_{MC}(M_{\mu\mu})$ template is built by selecting the simulated tag-and-probe events with identical requirements, in order to model the FSR and detector effects. While Gaussian function takes into account the energy scale shifts (by M_0, initialized at zero) and smearing effects (by σ_M, initialized at $\sim$ 2 GeV, to be the CMS detector resolution for $M_{\mu\mu}$) in the real data relative to simulation. *MC*Gaus* is used as the nominal signal model to produce central results. Alternatively, the other signal model *BW*CB* is also considered for signal shape systematic studies, which is the convolution of the Breit-Wigner [119] and Crystal-Ball function [120]:

$$f_{BW*CB}(M_{\mu\mu}) = \frac{1}{(M_{\mu\mu} - M_Z)^2 + (\frac{\Gamma_Z}{2})^2} * f_{CB}(M_{\mu\mu}), \tag{5.14}$$

where the

$$f_{CB}(M_{\mu\mu}) = \begin{cases} exp(\frac{(M_{\mu\mu}-M_0)^2}{2\sigma_M^2}) & \text{if } \frac{M_{\mu\mu}-M_0}{\sigma_M} > -\alpha \\ \frac{a}{(b-\frac{M_{\mu\mu}-M_0}{\sigma_M})^n} & \text{otherwise} \end{cases}, \tag{5.15}$$

$$\begin{aligned} a &= \frac{n}{|\alpha|} exp(-\frac{|\alpha|^2}{2}), \\ b &= \frac{n}{|\alpha|} - |\alpha|. \end{aligned} \tag{5.16}$$

Breit-Wigner is a common function to model resonance, where $M_Z = 91.1876$ GeV and $\Gamma_Z = 2.4952$ GeV are fixed to be the current best estimation of mass and width of Z boson [4]. Crystal-Ball is a piece-wise function including a Gaussian core with parameter M_0 and σ_M to account for energy scale shift and mass resolution, as well as a power-law low-end tail with parameter n and α to simulate the FSR.

For the background modeling, three simple function forms, *exponential*, *quadratic* and *power-law* function, are deployed:

$$\begin{aligned} f_{exp}(M_{\mu\mu}) &= exp(cM_{\mu\mu}), \\ f_{quad}(M_{\mu\mu}) &= p_0 + p_1 M_{\mu\mu} + p_2 M_{\mu\mu}^2, \\ f_{PL}(M_{\mu\mu}) &= \frac{1}{M_{\mu\mu}^a}, \end{aligned} \tag{5.17}$$

where c, p_0, p_1, p_2 and a are floating parameters. To ensure good fitting performance, a per-fit procedure is usually done to (60, 80)$\bigcup$(100, 120) GeV range with background models only to constrain the parameters. f_{exp} is sufficient when background contribution is small, so that it is used as nominal model to fit for $\epsilon_{tracking+ID+Iso}$. f_{quad} is motivated by the off-peak $M_{\mu\mu}$ distributions in the failing category of $\epsilon_{standalone}$, which can be well described by the polynomial curve. f_{PL} is utilized as the alternative function for both $\epsilon_{tracking+ID+Iso}$ and $\epsilon_{standalone}$ to evaluate the background shape uncertainties. Fit examples are shown in Figure 5-3, for $\epsilon_{tracking+ID+Iso}$ and $\epsilon_{standalone}$, passing and failing category, respectively.

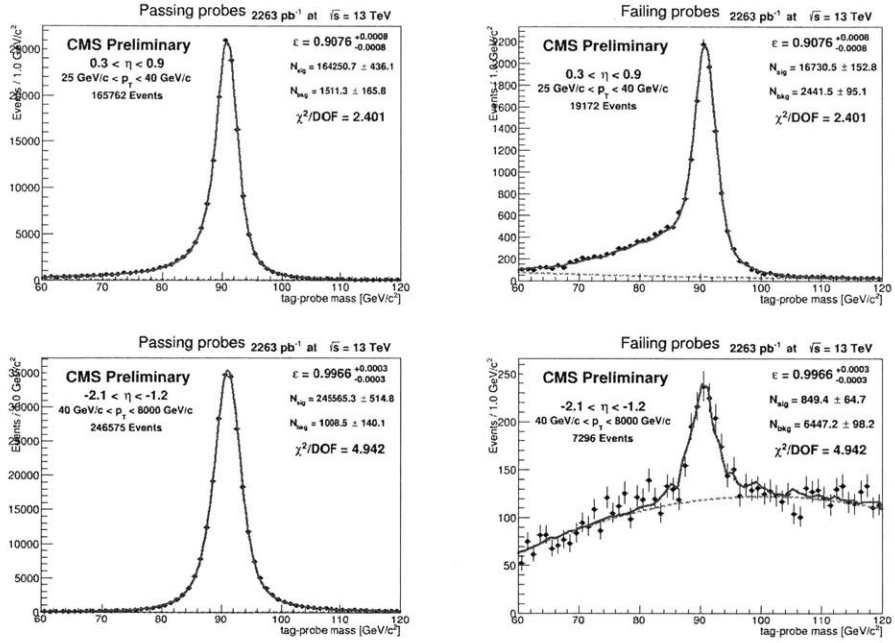


Figure 5-3: Fit examples for tag-and-probe invariant mass distributions for $\epsilon_{tracking+ID+Iso}$ (top) and $\epsilon_{standalone}$ (bottom), passing (left) and failing (right) category, respectively.

5.5.2 Single muon efficiency and scale factor results

Single muon trigger, tracking-identification-isolation and standalone efficiency scale factors are simply defined by:

$$\kappa(p_T, \eta) = \frac{\epsilon^{data}(p_T, \eta)}{\epsilon^{MC}(p_T, \eta)}, \tag{5.18}$$

where $\epsilon^{data}(p_T, \eta)$ and $\epsilon^{MC}(p_T, \eta)$ are corresponding efficiencies measured in data and simulation, respectively, based on the tag-and-probe method describe in Section 5.5.1.

Efficiency results are shown as a function of η in Figure 5-4. The trigger inefficiency is mainly due to L1 trigger and the online isolation requirements, while the variations in η are from the muon detector geometry that affects L1 trigger efficiency. The tracking-identification-isolation efficiencies vary between 90-95% for low-p_T bins and 95-98% for high-p_T bins. The standalone efficiencies are mostly consistent with unity, except for the dip in $0.2<|\eta|<0.3$, where there is a known issue of less instrumentation between the barrel wheels.

Scale factors are also reported in Table 5.3, 5.4 and 5.5. The data and simulation

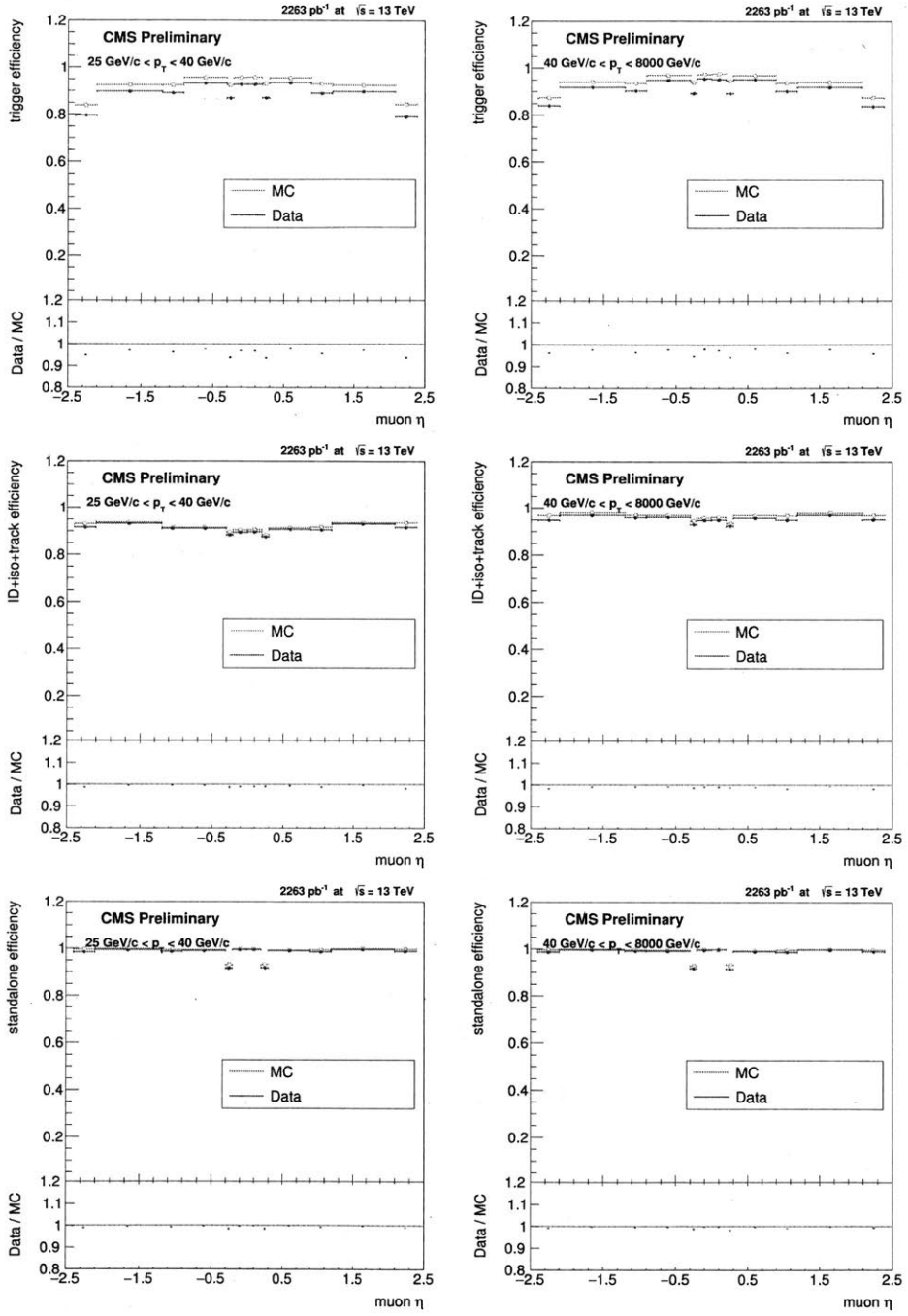


Figure 5-4: Single muon efficiency for trigger (top), tracking-identification-isolation (middle) and standalone (bottom) type as a function of pseudo-rapidity for $25 < p_T < 40$ (left) and $40 < p_T < 8000$ (right).

	$-2.4 < \eta < -2.1$	$-2.1 < \eta < -1.2$	$-1.2 < \eta < -0.9$	$-0.9 < \eta < -0.3$	$-0.3 < \eta < -0.2$	$-0.2 < \eta < 0$
$25 < p_T < 40$	0.949 ± 0.003	0.972 ± 0.001	0.964 ± 0.002	0.976 ± 0.001	0.939 ± 0.003	0.970 ± 0.001
$40 < p_T < \infty$	0.962 ± 0.002	0.976 ± 0.001	0.965 ± 0.001	0.978 ± 0.001	0.948 ± 0.002	0.981 ± 0.001
	$0 < \eta < 0.2$	$0.2 < \eta < 0.3$	$0.3 < \eta < 0.9$	$0.9 < \eta < 1.2$	$1.2 < \eta < 2.1$	$2.1 < \eta < 2.4$
$25 < p_T < 40$	0.969 ± 0.001	0.936 ± 0.003	0.979 ± 0.001	0.958 ± 0.002	0.972 ± 0.001	0.938 ± 0.003
$40 < p_T < \infty$	0.975 ± 0.001	0.943 ± 0.002	0.982 ± 0.001	0.963 ± 0.001	0.979 ± 0.001	0.958 ± 0.002

Table 5.3: Single muon trigger efficiency scale factors in (p_T, η) bins.

	$-2.4 < \eta < -2.1$	$-2.1 < \eta < -1.2$	$-1.2 < \eta < -0.9$	$-0.9 < \eta < -0.3$	$-0.3 < \eta < -0.2$	$-0.2 < \eta < 0$
$25 < p_T < 40$	0.985 ± 0.002	0.996 ± 0.001	0.996 ± 0.002	0.996 ± 0.001	0.987 ± 0.003	0.990 ± 0.002
$40 < p_T < \infty$	0.982 ± 0.001	0.990 ± 0.000	0.990 ± 0.001	0.991 ± 0.000	0.986 ± 0.002	0.990 ± 0.001
	$0 < \eta < 0.2$	$0.2 < \eta < 0.3$	$0.3 < \eta < 0.9$	$0.9 < \eta < 1.2$	$1.2 < \eta < 2.1$	$2.1 < \eta < 2.4$
$25 < p_T < 40$	0.989 ± 0.002	0.990 ± 0.003	0.993 ± 0.001	0.988 ± 0.002	0.996 ± 0.001	0.980 ± 0.002
$40 < p_T < \infty$	0.989 ± 0.001	0.987 ± 0.002	0.989 ± 0.001	0.982 ± 0.001	0.993 ± 0.000	0.982 ± 0.001

Table 5.4: Single muon tracking, identification, isolation efficiency scale factors in (p_T, η) bins.

agree within 1-2% for reconstruction and selection efficiencies. Those scale factors are then propagated to next steps.

5.5.3 Efficiency systematics evaluation

Three sources of systematics are taken into account for efficiency measurements: the choices of signal, background model and the binning. For binning uncertainty, a simplified binning is applied to re-evaluate the efficiencies and scale factors. For signal model uncertainty, a non-trivial approach by toy experiments is exploited and summarized as below.

1. Fit data with *MC*Gaus* and *exponential* to get nominal efficiency result ϵ_{true} and uncertainty σ_{true}.
2. Fit data with *CB*BW* and *exponential*.

	$-2.4 < \eta < -2.1$	$-2.1 < \eta < -1.2$	$-1.2 < \eta < -0.9$	$-0.9 < \eta < -0.3$	$-0.3 < \eta < -0.2$	$-0.2 < \eta < 0$
$25 < p_T < 40$	0.989 ± 0.001	0.996 ± 0.001	0.992 ± 0.001	0.997 ± 0.001	0.985 ± 0.003	0.998 ± 0.001
$40 < p_T < \infty$	0.991 ± 0.001	0.998 ± 0.000	0.996 ± 0.001	0.997 ± 0.000	0.989 ± 0.002	0.998 ± 0.001
	$0 < \eta < 0.2$	$0.2 < \eta < 0.3$	$0.3 < \eta < 0.9$	$0.9 < \eta < 1.2$	$1.2 < \eta < 2.1$	$2.1 < \eta < 2.4$
$25 < p_T < 40$	0.998 ± 0.001	0.984 ± 0.002	0.999 ± 0.001	0.991 ± 0.001	0.998 ± 0.001	0.991 ± 0.001
$40 < p_T < \infty$	0.999 ± 0.000	0.983 ± 0.002	0.997 ± 0.000	0.991 ± 0.001	0.998 ± 0.000	0.994 ± 0.001

Table 5.5: Single muon standalone efficiency scale factors in (p_T, η) bins.

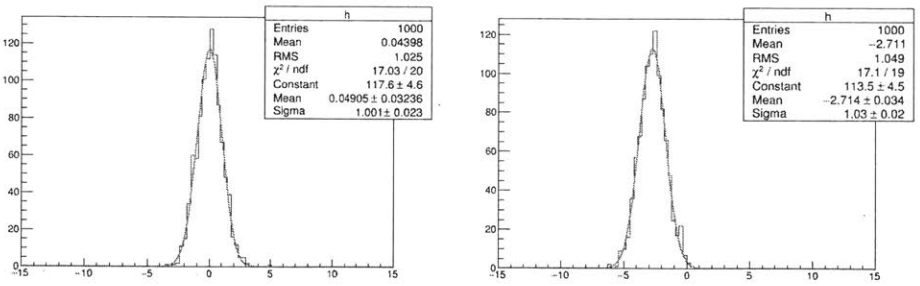


Figure 5-5: Examples of pull distribution for signal shape uncertainty study.

3. Generate 1000 toys with *CB*BW* parameters from Step 2 and *exponential* parameters from Step 1.
4. Fit each toy with *MC*Gaus* and *exponential* to get ϵ_{meas} and uncertainty σ_{meas}.
5. Construct pull distribution by: $pull = \frac{\epsilon_{meas} - \epsilon_{true}}{\sigma_{meas}}$.
6. Systematics for each (p_T, η) bin is then determined by: $\sigma_i = < pull >_{avg} \times \frac{\sigma_{true}}{\epsilon_{true}}$.

Two examples of the pull distribution when evaluating signal model uncertainty for $\epsilon_{tracking+ID+Iso}$, are shown in Figure 5-5 for illustration. It can be observed that the *MC*Gaus* and *CB*BW* behave more differently for the bin on the right than the left. Background model uncertainty follows a similar method.

Those per-bin uncertainties for single muon efficiencies are then propagated into the scale factor corrected $A \cdot \epsilon$ calculation (for inclusive measurements) or response matrix (for differential measurements) to evaluate the integrated impact on Z event efficiency.

5.5.4 Z event efficiency and systematics

As mentioned in Section 5.1.2, the Z event efficiencies in differential measurements are corrected by unfolding the response matrix, and the corresponding efficiency systematics will be presented in Section 5.9. To quantify the overall efficiency and systematic impact, the results for inclusive measurements are shown first in this section.

In the inclusive measurements, Z event efficiency is evaluated as the ratio between scale factor corrected $A \cdot \epsilon$ and A from simulation. While the uncertainties are estimated by taking the relative difference between the nominal value of $A \cdot \epsilon$ and the

	nominal	signal shape variation	background shape variation	binning variation	total
$A \cdot \epsilon$	0.3079 ± 0.0003	0.3110 ± 0.0003	0.3102 ± 0.0003	0.3083 ± 0.0003	-
Uncertainty	-	1.01%	0.73%	0.10%	1.25%

Table 5.6: Efficiency systematic uncertainty evaluations.

A	0.366±0.005
$A \cdot \epsilon$	0.308±0.004
ϵ	0.85±0.01

Table 5.7: Acceptance A, scale factor corrected $A \cdot \epsilon$ and Z event efficiency from simulation.

ones with shape and binning variations, see Table 5.6 (uncertainties quoted on $A \cdot \epsilon$ are statistical only), the total systematic uncertainty from efficiency is 1.25%. The final Z event efficiency result is reported in Table 5.7. Uncertainties quoted are theoretical systematics on A and experimental efficiency systematics on $A \cdot \epsilon$, and they are uncorrelated and add in quadrature to determine the total uncertainty on Z event efficiency ϵ.

5.6 Muon momentum scale and resolution correction

Rochester correction [121] is applied to correct muon momentum scale shifts for data, due to detector misalignment and reconstruction bias, as well as inaccurate momentum resolution for simulation, because of the imperfect modeling of the detector response and material budget.

A comparison of data/simulation agreement in di-muon invariant mass distribution before and after such a correction are shown in Figure 5-6. As can be seen, the data/simulation discrepancy around the Z mass peak is resolved by this correction. The systematic uncertainty of this correction is estimated by varying up and down by 1σ according to the calibration uncertainty of the correction, which is propagated into the final cross section results via signal yields. The impact is at per-mill level for the inclusive measurements and most of the phase space for the differential measurements, except in the high-p_T bins where the scale and resolution uncertainty increases to a few percent (see Section 5.9).

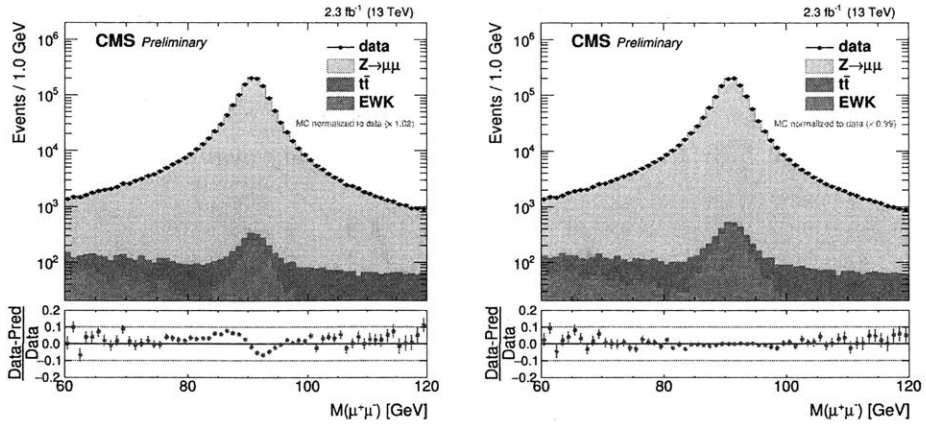


Figure 5-6: Data/simulation agreement in di-muon invariant mass ($M_{\mu\mu}$) distribution before (left) and after (right) Rochester Correction. Simulation is normalized to data yields.

Data yields	1343017 ± 1160
Background yields	7050 ± 1330
Background contribution	0.5% ± 0.1%

Table 5.8: Data and background yield in fiducial region for inclusive measurements

5.7 Signal extraction

Signal yields are simply determined by counting data events in the fiducial region and subtracting the background contribution, which is estimated using efficiency scale factor corrected simulated events. Background processes considered including $t\bar{t}$, W+jets, and di-boson productions WW, WZ and ZZ, which could either produce a same di-muon final states or fake to have one.

The inclusive yields are reported in Table 5.8, and relative contributions to fiducial region are illustrated in Figure 5-6. The data yield uncertainties are given by Poisson statistics. While the background yield uncertainties also include the systematics on $t\bar{t}$ and di-boson normalization, which is set to be 10% and 30%, respectively. Those numbers are estimated based on other early 13 TeV CMS measurements on those processes [122–124], as well cross-checked by a control region study of $e\mu$ and 3-lepton final sates. See Figure 5-7, where a Z-like object is constructed by requiring the opposite-charge, same p_T, η and $M_{\ell\ell}$ cuts as the Z event selection for the electron and muon in $e\mu$ and the same-flavor pair in $\ell\ell\ell$ control region.

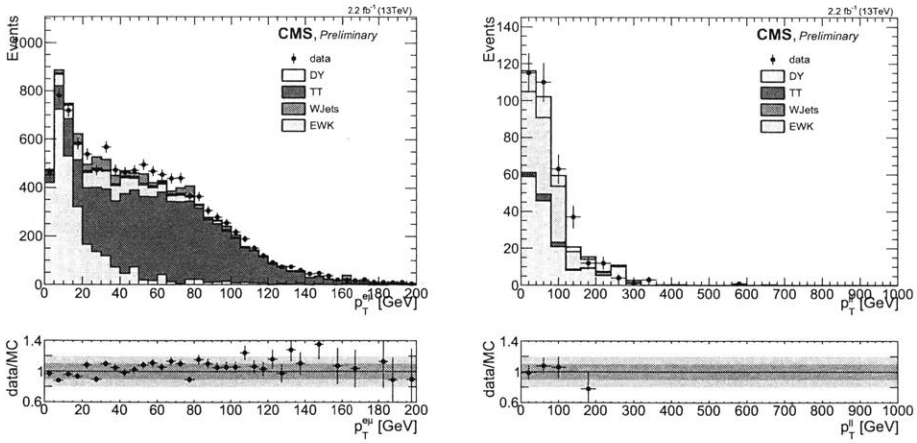


Figure 5-7: Data/simulation agreement in $p_T^{e\mu}$ in $e\mu$ control region (left) and $p_T^{\ell\ell}$ in $\ell\ell\ell$ control region (right).

The background subtraction for differential yields at the detector level is also determined in the same way[2]. The data/simulation comparisons in $p_T^{\mu^+\mu^-}$, $\phi_\eta^\star$, $|y^{\mu^+\mu^-}|$, $p_T^{\mu^+}$ and $p_T^{\mu^-}$ variables are shown in Figure 5-8. The agreement is good in general. 10% discrepancy is observed in some phase space of $p_T^{\mu^+\mu^-}$, which is because the parton shower in simulation is not well modeled. In addition, as can be see, the background contribution is small (<1%) across all phase space for differential measurements, the corresponding uncertainties are mostly at per-mill level, see Section 5.9.

5.8 Unfolding

The detector level distributions of the observables of interest have been presented in Figure 5-8. However, generally speaking, due to the parton shower, hadronization, particle decay and detector effects, the distribution of a measured observable at reconstruction (detector) level is different from the true distribution at generator (particle) level. The unfolding method we exploited in the measurements is based on Bayes' theorem [115], which is superior to other unfolding approaches [125] in the sense that it is capable to take into account the migration effects between bins, and avoid potential numerical instability of inverting matrices.

[2]An alternative approach for background subtraction in differential measurements is to exploit the unfolding method based on Bayer' theorem, where the background contributions from different processes can be simply treated as additional bins for truth, see 5.8.1.

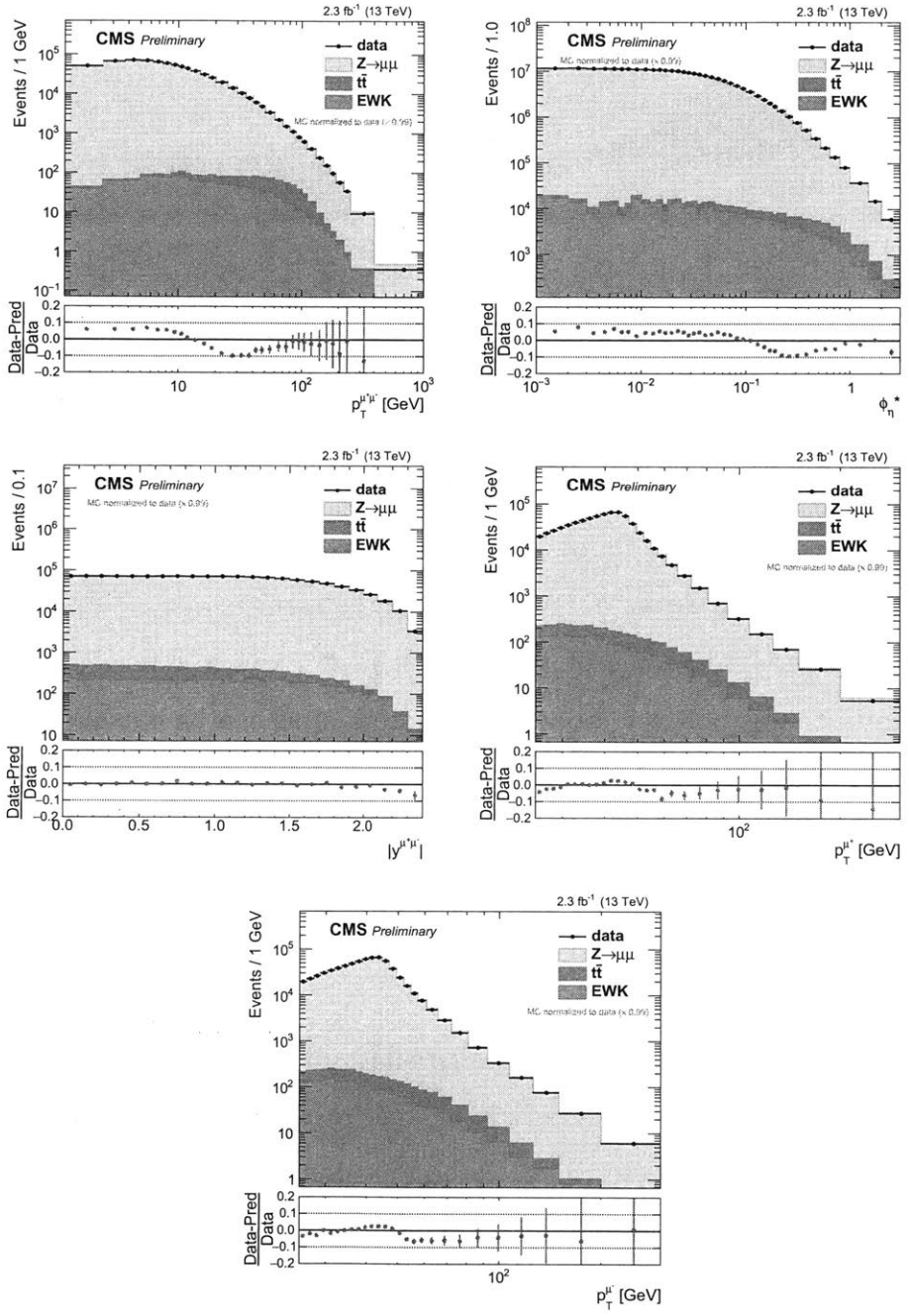


Figure 5-8: Data/simulation agreement in $p_T^{\mu^+\mu^-}$ (top-left), $\phi_\eta^\star$ (top-right), $|y^{\mu^+\mu^-}|$ (middle-left), $p_T^{\mu^+}$ (middle-right) and $p_T^{\mu^-}$ (bottom) variables at the detector level. Simulation is normalized to data yields.

5.8.1 Unfolding methodology based on Bayes' theorem

Starting with Bayes' theorem in terms of the truth in i^{th} bin, $\mathcal{O}_i^{truth}$, and measured observable in j^{th} bin, $\mathcal{O}_j^{meas}$, we have

$$\mathcal{P}(\mathcal{O}_i^{truth}|\mathcal{O}_j^{meas}) = \frac{\mathcal{P}(\mathcal{O}_j^{meas}|\mathcal{O}_i^{truth})\mathcal{P}_0(\mathcal{O}_i^{truth})}{\sum_{l=1}^{n_{truth}} \mathcal{P}(\mathcal{O}_j^{meas}|\mathcal{O}_l^{truth})\mathcal{P}_0(\mathcal{O}_l^{truth})}, \tag{5.19}$$

where $\mathcal{P}_0(\mathcal{O}_i^{truth})$ is the *priori* distribution of $\mathcal{O}_i^{truth}$, which is initialized to be same as the simulated distribution at generator level. This does not bias the result, as it will be updated iteratively in the method to approximate the true one, see Section 5.8.3 for details. $\mathcal{P}(\mathcal{O}_j^{meas}|\mathcal{O}_i^{truth})$ is defined as *response matrix (or migration matrix)*. Normalization conditions of those probabilities are:

1. $\sum_{i=1}^{n_{truth}} \mathcal{P}_0(\mathcal{O}_i^{truth})$=1,
2. $\sum_{i=1}^{n_{truth}} \mathcal{P}(\mathcal{O}_i^{truth}|\mathcal{O}_j^{meas})$=1,
3. 0<$\epsilon = \sum_{j=1}^{n_{meas}} \mathcal{P}(\mathcal{O}_j^{meas}|\mathcal{O}_l^{truth})$<1.

After the detector level distribution $\{n(\mathcal{O}_j^{meas})\}$ from data is recorded in measurements, the best estimator of particle level distribution $\{\hat{n}(\mathcal{O}_i^{truth})\}$ is then built by:

$$\begin{aligned}
\hat{n}(\mathcal{O}_i^{truth}) &= \frac{1}{\epsilon}\sum_{j=1}^{n_{meas}} \mathcal{P}(\mathcal{O}_i^{truth}|\mathcal{O}_j^{meas})n(\mathcal{O}_j^{meas}) \\
&= \sum_{j=1}^{n_{meas}} \frac{\mathcal{P}(\mathcal{O}_j^{meas}|\mathcal{O}_i^{truth})\mathcal{P}_0(\mathcal{O}_i^{truth})}{[\sum_{l=1}^{n_{meas}} \mathcal{P}(\mathcal{O}_l^{meas}|\mathcal{O}_i^{truth})][\sum_{l=1}^{n_{truth}} \mathcal{P}(\mathcal{O}_j^{meas}|\mathcal{O}_l^{truth})\mathcal{P}_0(\mathcal{O}_l^{truth})]} n(\mathcal{O}_j^{meas}) \\
&= \sum_{j=1}^{n_{meas}} \mathcal{M}_{ij} n(\mathcal{O}_j^{meas}),
\end{aligned} \tag{5.20}$$

where $\mathcal{M}_{ij}$ is the *unfolding matrix*, which is estimated from simulation.

Practically, the unfolding matrix is constructed using a common package called RooUnfold [126], which takes three histograms as inputs: the 1-dimensional truth distribution $\{n^{MC}(\mathcal{O}_i^{truth})\}$, the 1-dimentional measured distribution $\{n^{MC}(\mathcal{O}_j^{meas})\}$ and the 2-dimensional response matrix $\mathcal{P}(\mathcal{O}_j^{meas}|\mathcal{O}_i^{truth})$. The truth distribution is filled

by events in fiducial region at the generator level, while the measured distribution is filled by events in fiducial region at reconstruction level, including the data-to-simulation efficiency scale factors to account for muon efficiency difference between data and simulation. The response matrix is filled by events in fiducial region both at the generator and reconstruction level. Examples of response matrices for observables of interest are illustrated in Figure 5-9.

5.8.2 Binning

Choice of binning concerns two aspects: the migration between bins due to detector smearing effect and the number of events for each bin. In order to quantify the migration effect, the bin *purity* is defined for each bin at the reconstruction level (i.e, bins on y axis of histograms in Figure 5-9), as the fraction of events coming from the same bin at the generator level. A minimum of 60% on bin purity is required in order to have a satisfactory bin resolution $\lesssim 1\sigma$. As angular observables have higher resolution, 90% bin purity is required to prevent large number of bins. Additionally, each bin must have at least 50 events to avoid tail of distributions. Figure 5-10 presents the bin purity after the optimization procedure described above.

5.8.3 Iterative algorithm and regularization

After unfolding, the prior distribution can be recovered by:

$$\mathcal{P}(\mathcal{O}_i^{truth}) = \frac{\hat{n}(\mathcal{O}_i^{truth})}{\sum_{i=1}^{n_{truth}} \hat{n}(\mathcal{O}_i^{truth})}, \tag{5.21}$$

which can be verified[3] to lie between the initial value $\mathcal{P}_0(\mathcal{O}_i^{truth})$ and the true one. Thus, an iterative procedure is performed until the χ^2 of $\hat{n}(\mathcal{O}_i^{truth})$ between iterations are small enough. However, the ultimate true priori determined by numerical calculation might be highly unstable. Therefore, regularization is introduced by stopping iterations before the priori reaches the true value.

[3]See Section 6 of G. D'Agostini' paper [115]

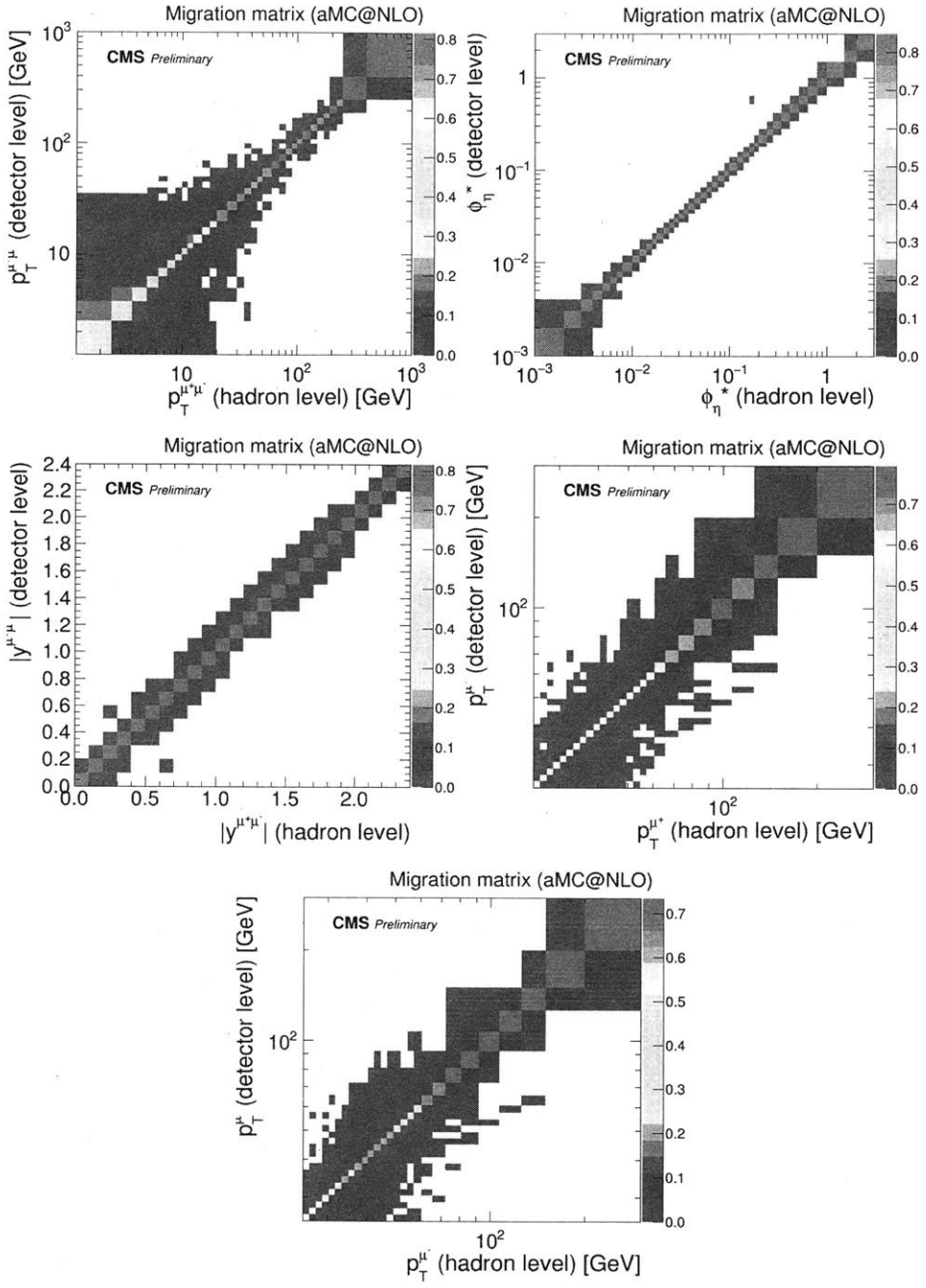


Figure 5-9: Response matrices for $p_T^{\mu^+\mu^-}$ (top-left), $\phi_\eta^\star$ (top-right), $|y^{\mu^+\mu^-}|$ (middle-left), $p_T^{\mu^+}$ (middle-right) and $p_T^{\mu^-}$ (bottom) observables. Matrices are determined from aMC@NLO simulation and normalized according to normalization condition 2 in Section 5.8.1.

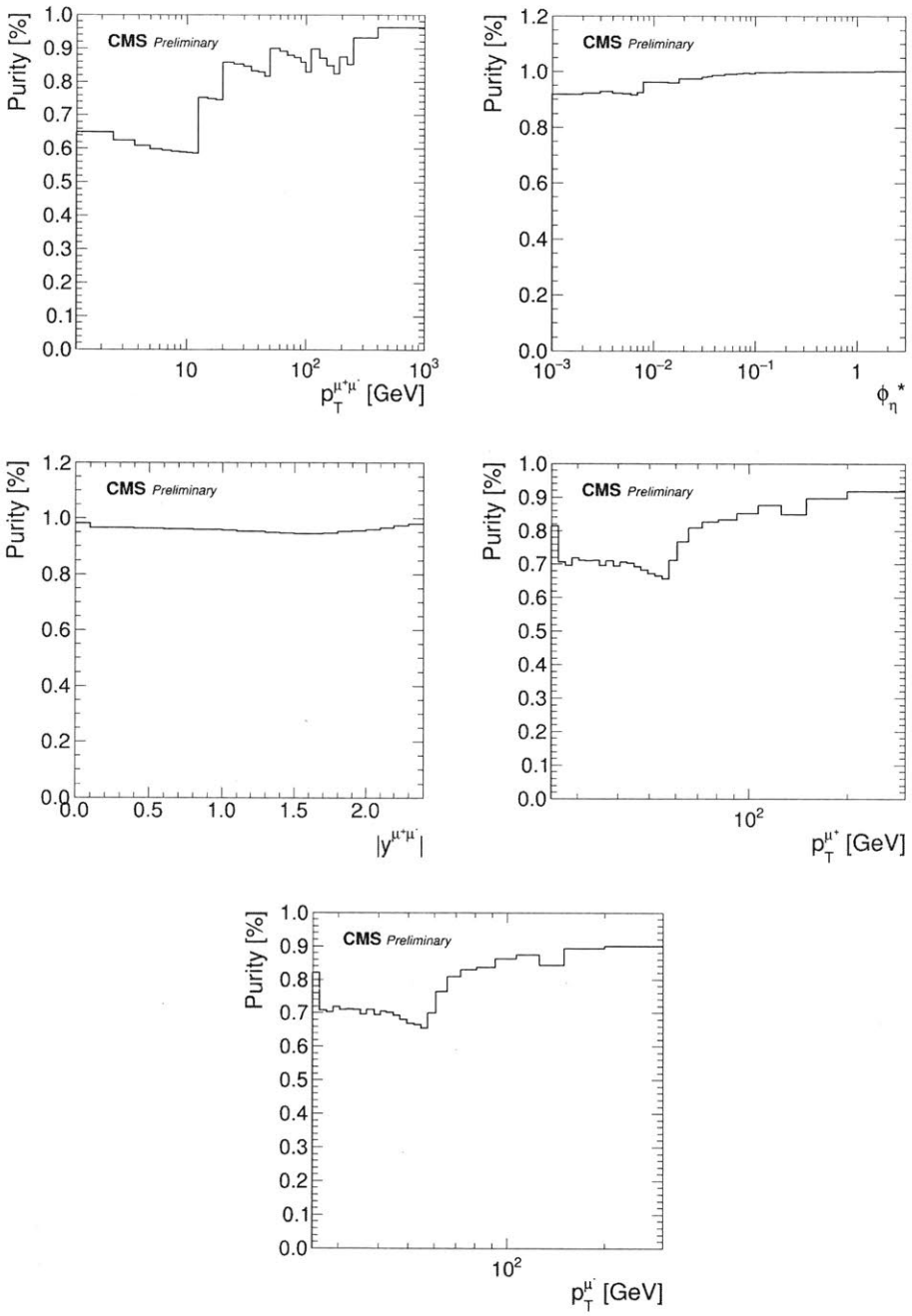


Figure 5-10: Bin purity for observable $p_T^{\mu^+\mu^-}$ (top-left), $\phi_\eta^\star$ (top-right), $|y^{\mu^+\mu^-}|$ (middle-left), $p_T^{\mu^+}$ (middle-right) and $p_T^{\mu^-}$ (bottom) binnings.

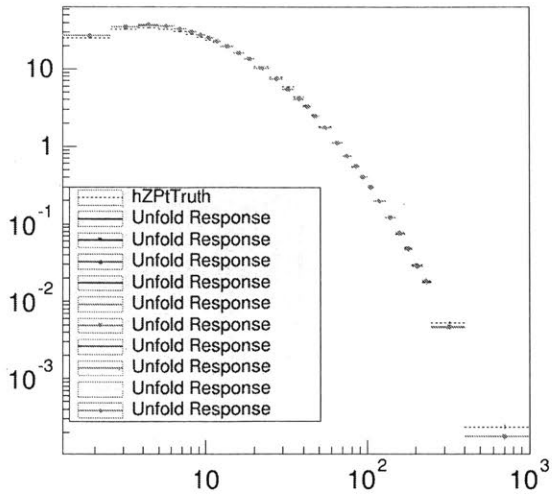


Figure 5-11: Unfolded distribution of $p_T^{\mu^+\mu^-}$ after different number of iterations.

Iteration	1	2	3	4	5	6	7	8
χ^2	6.303	1.240	0.472	0.214	0.110	0.063	0.040	0.027

Table 5.9: χ^2 of a toy $p_T^{\mu^+\mu^-}$ distribution between iterations.

Studies suggest the results are actually not very insensitive to the number of iterations. Figure 5-11 shows the unfolded distribution of $p_T^{\mu^+\mu^-}$ after different number of iterations, and Table 5.9 compares the χ^2 of a toy $p_T^{\mu^+\mu^-}$ distribution between iterations, which indicates the optimal number of iteration is 2-3.

5.8.4 Unfolding uncertainties

The uncertainties introduced specifically by the unfolding procedure include two items: model dependence systematic uncertainty of unfolding matrix and statistical uncertainties propagated from measured distributions $n(\mathcal{O}_i^{meas})$ and the unfolding matrix $\mathcal{M}_{ij}$.

Model dependence uncertainty

As described previously, the unfolding matrix is built entirely relying on the nominal simulated sample, aMC@NLO, which means that one assumes the parton shower,

hadronization and detector smearing effects in aMC@NLO are same as in data. The systematic uncertainty of this assumption is estimated by comparing the unfolded distributions using two different unfolding matrices: one from the nominal simulated sample ($\hat{n}_{nominal}$), and the other one from an alternative simulated sample, POWHEG ($\hat{n}_{alternative}$). Figure 5-12 shows corresponding response matrices using POWHEG sample and the solid black histograms in Figure 5-13 present the relative systematic shifts of model dependence uncertainty, which is defined as:

$$\mathcal{D}(\mathcal{O}_i^{truth}) = \frac{\hat{n}_{alternative}(\mathcal{O}_i^{truth}) - \hat{n}_{nominal}(\mathcal{O}_i^{truth})}{\hat{n}_{nominal}(\mathcal{O}_i^{truth})}. \tag{5.22}$$

However, as can be observed, the model dependence uncertainty determined in this way is highly affected by statistical fluctuations in some phase space of the observables, especially for high-p_T regions, where number of events in both data and simulation are limited. To have a more reliable uncertainty estimation, Gaussian kernel smoothing [127] is exploited. The smoothed uncertainty $\tilde{\mathcal{D}}(\mathcal{O}_i^{truth})$ is the weighted average of the neighborhood bins, while the weights $\mathcal{W}$ are defined in the sense that weights are larger for bins which are closer and have smaller statistical fluctuations (i.e, more events). Formally,

$$\begin{aligned}
\tilde{\mathcal{D}}(\mathcal{O}_i^{truth}) &= \frac{\sum_{j=1}^{n_{truth}} \mathcal{D}(\mathcal{O}_j^{truth}) \cdot \mathcal{W}(\mathcal{O}_i^{truth}, \mathcal{O}_j^{truth})}{\sum_{j=1}^{n_{truth}} \mathcal{W}(\mathcal{O}_i^{truth}, \mathcal{O}_j^{truth})}, \\
\mathcal{W}(\mathcal{O}_i^{truth}, \mathcal{O}_j^{truth}) &= \omega(\mathcal{O}_j^{truth}) \times exp(-\frac{\mathcal{O}_i^{truth} - \mathcal{O}_j^{truth}}{2\sigma^2_{\mathcal{O}^{truth}}}), \\
\omega(\mathcal{O}_j^{truth}) &= \frac{\frac{1}{(\delta\sigma(\mathcal{O}_j^{truth}))^2}}{\sum_{j=1}^{n_{truth}} \frac{1}{(\delta\sigma(\mathcal{O}_j^{truth}))^2}}.
\end{aligned} \tag{5.23}$$

The smoothing curves and the smoothed model dependence uncertainties can also be found in Figure 5-13.

Statistical uncertainty

To evaluate the statistical uncertainty in the unfolded distribution $\hat{n}(\mathcal{O}_i^{truth})$, one can not simply do $\sqrt{\hat{n}(\mathcal{O}_i^{truth})}$. Because it is the measured distribution $n(\mathcal{O}_j^{meas})$ follows

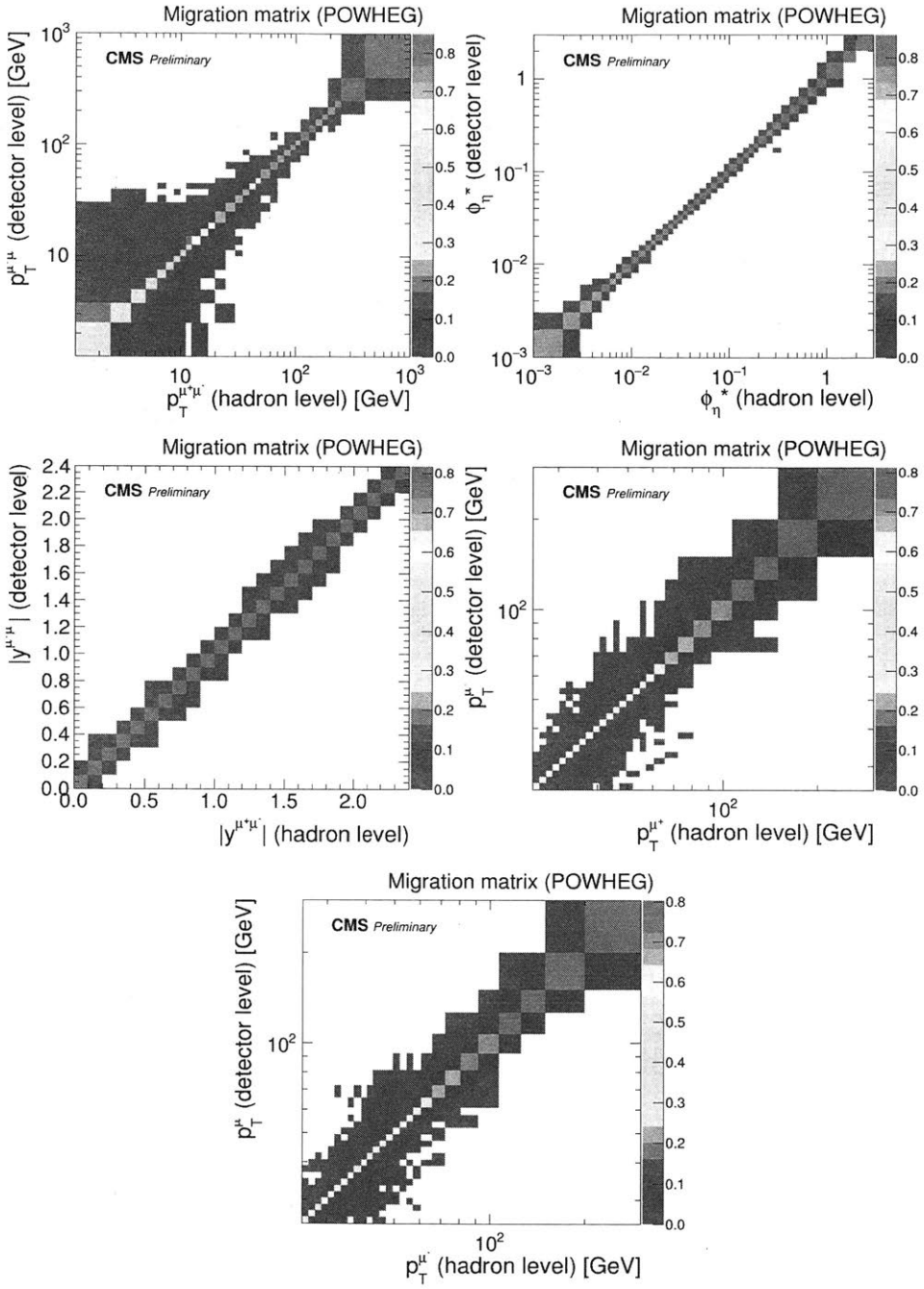


Figure 5-12: Response matrices for $p_T^{\mu^+\mu^-}$ (top-left), $\phi_\eta^\star$ (top-right), $|y^{\mu^+\mu^-}|$ (middle-left), $p_T^{\mu^+}$ (middle-right) and $p_T^{\mu^-}$ (bottom) observables. Matrices are determined from POWHEG simulation and normalized according to normalization condition 2 in Section 5.8.1.

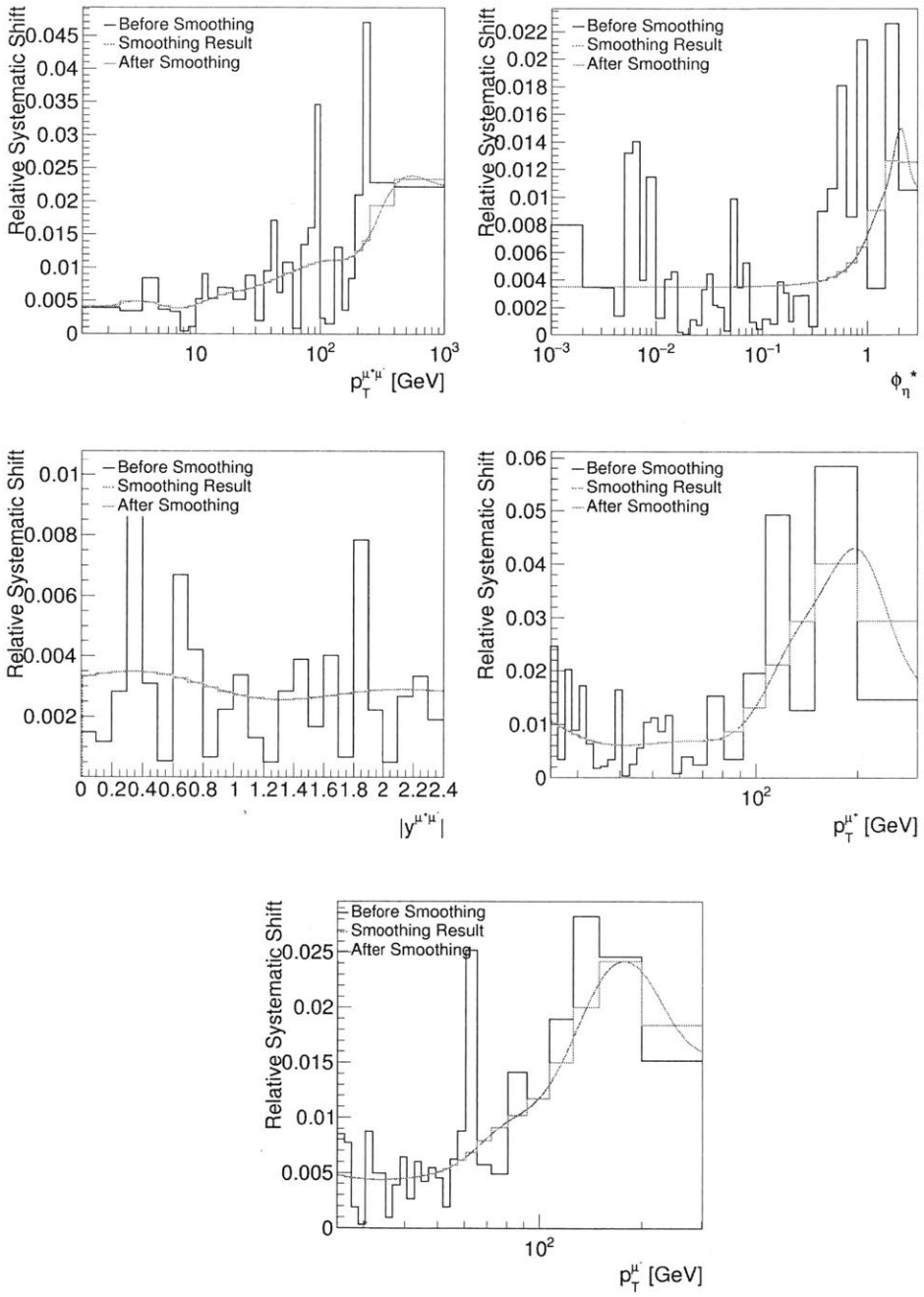


Figure 5-13: Model dependence uncertainty before and after smoothing for $p_T^{\mu^+\mu^-}$ (top-left), $\phi_\eta^\star$ (top-right), $|y^{\mu^+\mu^-}|$ (middle-left), $p_T^{\mu^+}$ (middle-right) and $p_T^{\mu^-}$ (bottom) unfolded distributions.

Poisson distribution, but not $\hat{n}(\mathcal{O}_i^{truth})$, which actually consists of contributions from different measured bins so that has correlations across bins. It is also affected by limited simulated events in unfolding matrix. Formally, according to Equation 5.20, the covariance matrix of $\hat{n}(\mathcal{O}_i^{truth})$ can be determined by uncertainty propagation from $n(\mathcal{O}_j^{meas})$:

$$\mathcal{V}(\hat{n}(\mathcal{O}_k^{truth}), \hat{n}(\mathcal{O}_l^{truth})) = \sum_{i=1}^{n_{meas}} \sum_{j=1}^{n_{meas}} \frac{\partial \hat{n}(\mathcal{O}_k^{truth})}{\partial n(\mathcal{O}_i^{meas})} \mathcal{V}(n(\mathcal{O}_i^{meas}), n(\mathcal{O}_j^{meas})) \frac{\partial \hat{n}(\mathcal{O}_l^{truth})}{\partial n(\mathcal{O}_j^{meas})}. \tag{5.24}$$

However, an alternative approach, to create 1000 toys by $n^{toy}(\mathcal{O}_j^{meas}) \sim Gaus(n(\mathcal{O}_j^{meas}), \sigma_j)$ and construct the sample covariance of the unfolded distributions, is more accurate and practical, especially for the unfolding method based on Bayes' theorem, and thus deployed in the measurements. The statistical correlation matrices determined in this way are illustrated in Figure 5-14.

5.9 Systematics

The sources of systematics considered in the analysis consist of theoretical uncertainties on the predictions, including:

- PDF choice,
- higher-order QCD correction,
- higher-order EW correction,
- FSR modeling,

as well as experimental uncertainties on measurements, which are summarized as follows:

- luminosity,
- muon efficiency scale factor correction,
- muon momentum scale and resolution correction,
- background subtraction,
- model dependence in unfolding.

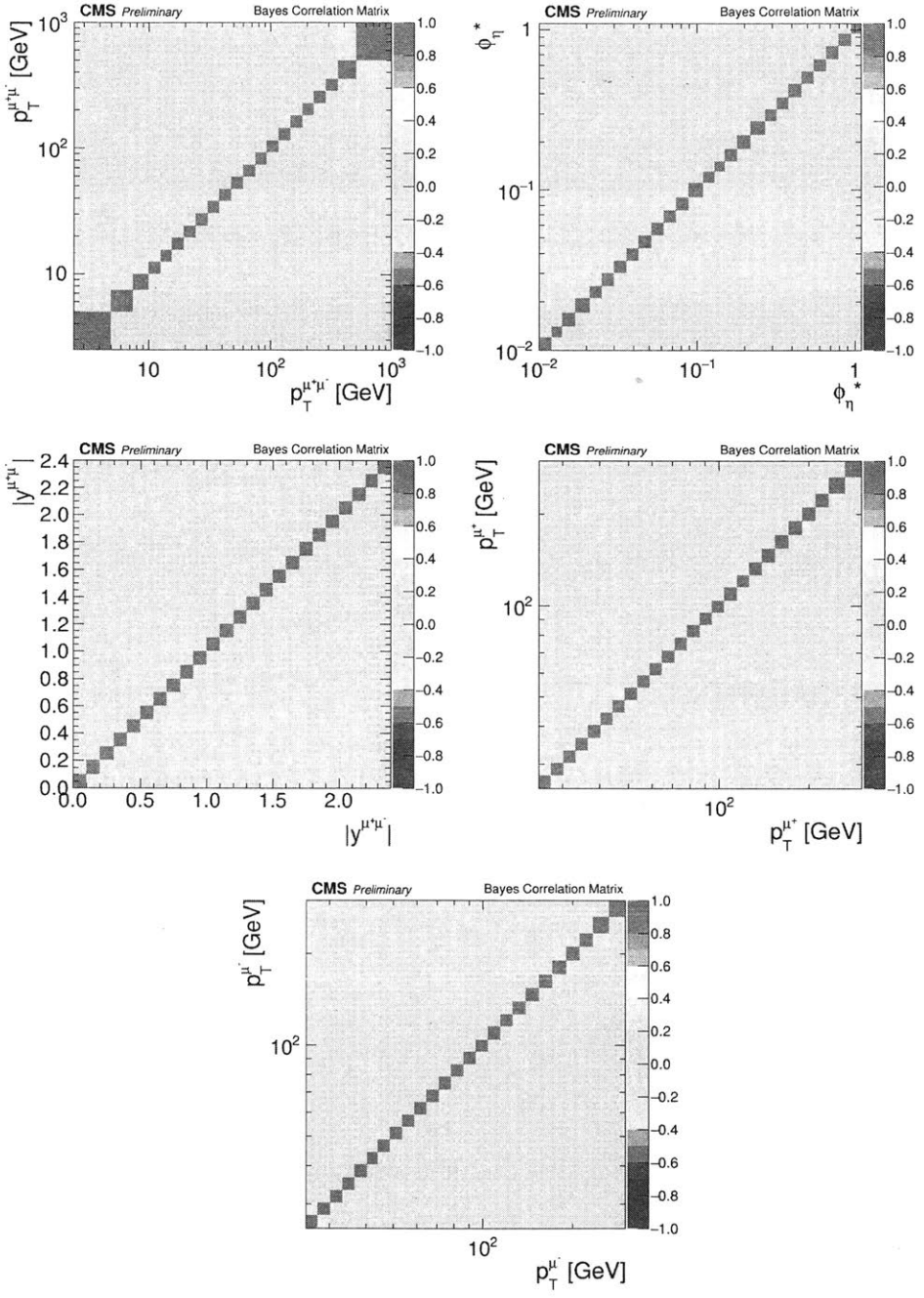


Figure 5-14: Statistical correlation matrices for $p_T^{\mu^+\mu^-}$ (top-left), $\phi_\eta^\star$ (top-right), $|y^{\mu^+\mu^-}|$ (middle-left), $p_T^{\mu^+}$ (middle-right) and $p_T^{\mu^-}$ (bottom) unfolded distributions.

Muon efficiency	1.3%
Background subtraction	0.1%
Total experimental	1.3%
PDF	0.7%
QCD correction	1.1%
EW correction	0.4%
Total theoretical	1.4%
Luminosity	2.7%
Total	3.3%

Table 5.10: Systematic uncertainties on the inclusive measurements

The dedicated methods for uncertainty evaluation have been elaborated in Section 5.4 for theoretical sources and Section 5.5 to 5.8 for experimental ones. Additionally, the luminosity uncertainty for 2015 data-taking period is 2.7% [114].

5.9.1 Inclusive systematics

The systematic results for each individual component in the inclusive measurements have been discussed in Section 5.4 to 5.7 and are now summarized in Table 5.10. As can be seen, the luminosity uncertainty is dominant, and the other experimental and theoretical systematics are comparable, which are both at percent level.

5.9.2 Differential systematics

To integrate the systematics into differential measurement results, for each source k, 1σ variation of the correction (for efficiency scale factor and momentum scale and resolution uncertainty) or the normalization (for luminosity and background uncertainty) is applied to simulated events individually. The modified unfolded distributions ($\hat{n}^k_{modified}$) are then determined using correspondingly adjusted unfolding matrices and detector level data distributions with the adjusted backgrounds subtracted. Then the systematic uncertainty for source k can be calculated by taking the relative

difference from the nominal unfolded distributions ($\hat{n}^k_{modified}$):

$$\mathcal{D}^k(\mathcal{O}_i^{truth}) = \frac{\hat{n}^k_{modified}(\mathcal{O}_i^{truth}) - \hat{n}_{nominal}(\mathcal{O}_i^{truth})}{\hat{n}_{nominal}(\mathcal{O}_i^{truth})}. \tag{5.25}$$

Figure 5-15 illustrates the relative systematic uncertainties for each specific source as well as the total contribution except for the luminosity, over entire phase space of corresponding observables of interest. As can be observed, the luminosity uncertainty of 2.7% is still the leading systematics in most regions of the phase space. While the unfolding uncertainty is also dominant in some regions, typically ranging from 1-5% and increasing to ~10% where there is limited simulated events. The efficiency scale factor systematics account for 1-2% and stay very stable over entire regions. The uncertainty for muon momentum scale and resolution correction is negliable for angular observables as expected but increases as a function of p_T. The background systematics are tiny compared to others. Finally, the different sources of systematic uncertainties are added in quadrature to determine the total uncertainty.

As might be seen, in the differential measurements, different sources of systematics are treated uncorrelated, while for a same source of systematics, different bins are treated 100% correlated. To visualize the correlation for the total systematics across phase space of interested observables, the following correlation matrices are defined:

$$\begin{aligned} \rho(i,j) &= \frac{cov(i,j)}{\sigma(\mathcal{O}_i^{truth})\sigma(\mathcal{O}_j^{truth})}, \\ cov(i,j) &= \sum_k \mathcal{D}^k(\mathcal{O}_i^{truth}) \times \mathcal{D}^k(\mathcal{O}_j^{truth}), \end{aligned} \tag{5.26}$$

where $\sigma(\mathcal{O}_i^{truth})$ is the total uncertainty for bin i. Figure 5-16 presents the results of those correlation matrices. The uncertainties across bins are still strongly correlated, and some of the off-diagonal bins lose correlation because the non-correlated systematics.

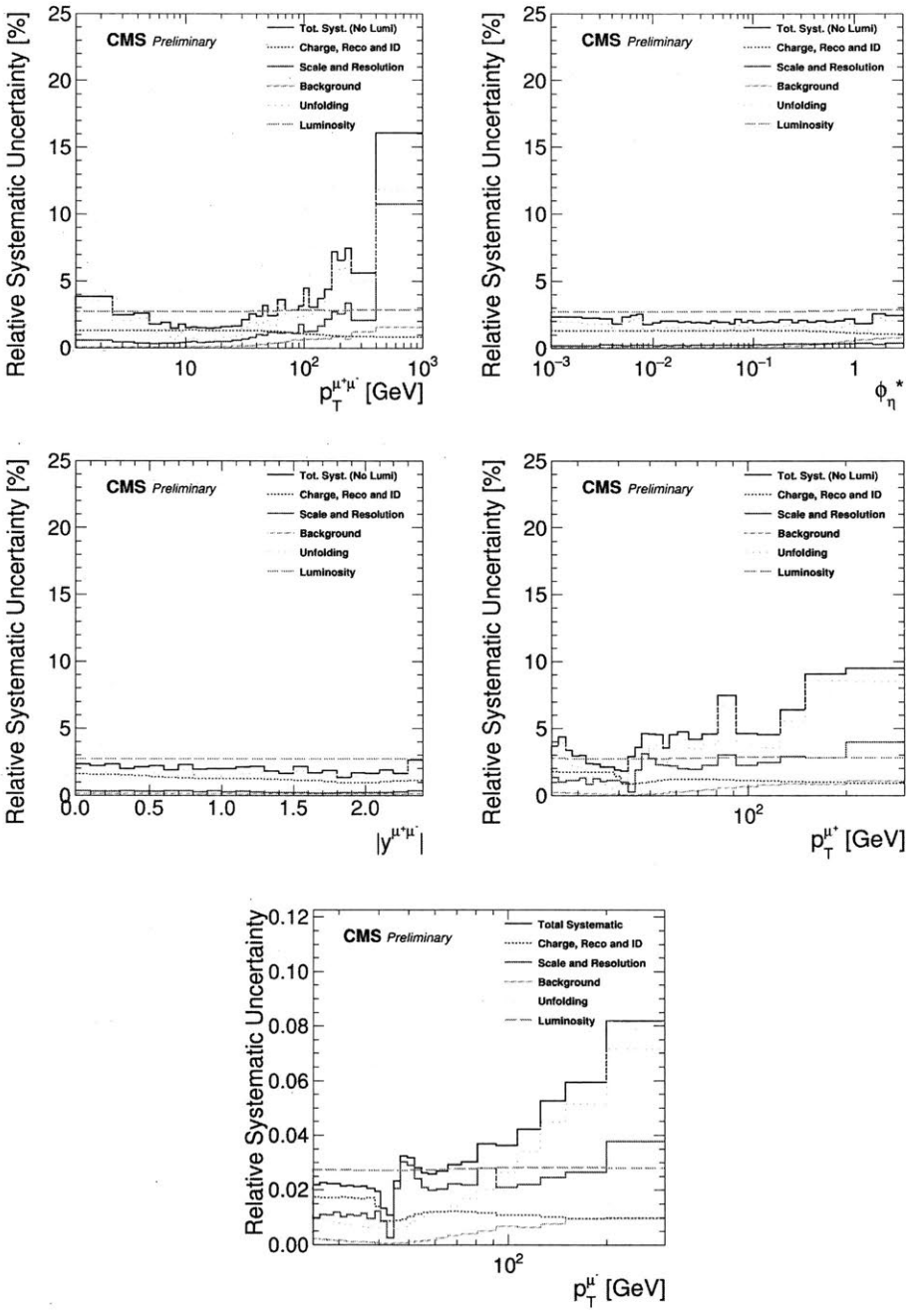


Figure 5-15: Relative systematic uncertainty for $p_T^{\mu^+\mu^-}$ (top-left), $\phi_\eta^\star$ (top-right), $|y^{\mu^+\mu^-}|$ (middle-left), $p_T^{\mu^+}$ (middle-right) and $p_T^{\mu^-}$ (bottom) unfolded distributions.

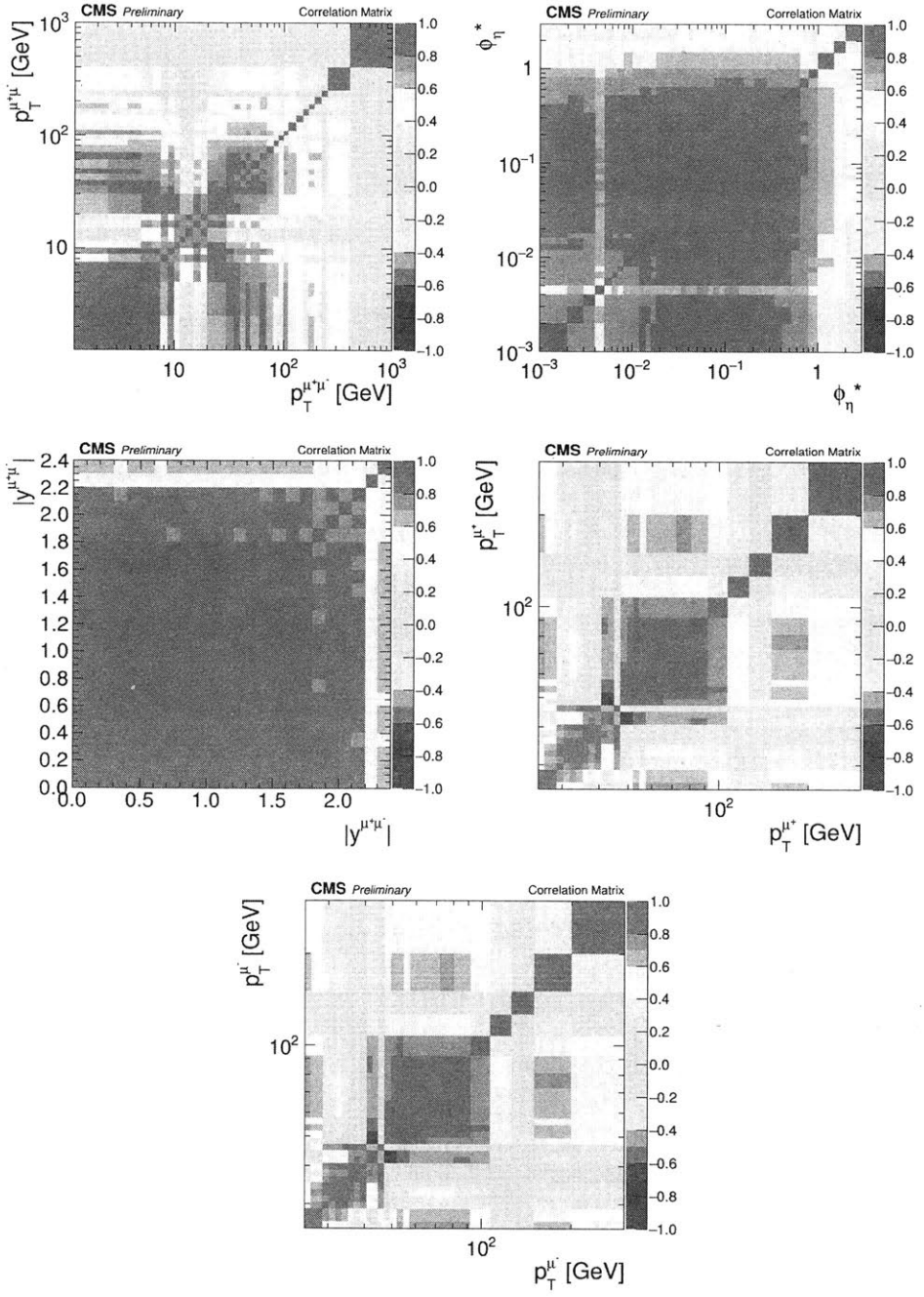


Figure 5-16: Correlation matrices of systematic uncertainty for $p_T^{\mu^+\mu^-}$ (top-left), $\phi_\eta^\star$ (top-right), $|y^{\mu^+\mu^-}|$ (middle-left), $p_T^{\mu^+}$ (middle-right) and $p_T^{\mu^-}$ (bottom) unfolded distributions.

Total data yields	1343017 ± 1160
Background yields	7050 ± 1330
Dressed acceptance	0.372 ± 0.005
Naked acceptance	0.366 ± 0.005
Efficiency	0.85 ± 0.01

Table 5.11: The total data and background yields, dressed and naked level acceptance and overall Z event efficiency for inclusive measurements.

5.10 Results

5.10.1 Inclusive results

The intermediate results for acceptances, efficiencies and yields and corresponding uncertainties have been presented separately in previous sections for inclusive measurements and are now summarized in Table 5.11.

Total Z cross section

According to Equation 5.7, the measured inclusive total Z cross section for dimuon mass in the range of 60-120 GeV is:

$$\sigma(pp \to ZX) \times \mathcal{B}(Z \to \mu\mu) = 1870 \pm 2(stat) \pm 35(syst) \pm 51(lumi)pb.$$

In order to compare with theoretical predictions, FEWZ code [61], which provides cross section calculations at NNLO QCD and NLO EW, and various PDF set as inputs, are exploited. The uncertainties in those calculations, including PDF uncertainties and scale variations, are also considered. Table 5.12 summarizes the predicted inclusive total cross sections by FEWZ, with NNPDF 3.0 [116], CT14 [128], MMHT2014 [129], ABM12LHC [130] and HERAPDF15 [131] as PDF set input. The comparison shows good agreement between measured value and theoretical prediction, which indicates NNLO QCD and NLO EW describes the inclusive cross section reasonably well.

NNPDF3.0	1870^{+50}_{-40} pb
CT14	1900^{+50}_{-50} pb
MMHT2014	1920^{+40}_{-40} pb
ABM12LHC	1920^{+20}_{-20} pb
HERAPDF15	1930^{+90}_{-40} pb

Table 5.12: The FEWZ prediction for inclusive cross sections, with various PDF sets as input.

Naked $\sigma^{total}_{FEWZ} \cdot A$	684 ± 23 pb
Dressed $\sigma^{total}_{FEWZ} \cdot A$	695 ± 23 pb

Table 5.13: The theoretical prediction for fiducial cross sections.

Fiducial Z cross section

According to Equation 5.8, the measured Z cross sections in defined fiducial regions for dimuon mass in the range of 60-120 GeV are:

$$\sigma^{fid}_{naked}(pp \to ZX) \times \mathcal{B}(Z \to \mu\mu) = 684 \pm 1(stat) \pm 9(syst) \pm 19(lumi)pb,$$
$$\sigma^{fid}_{dressed}(pp \to ZX) \times \mathcal{B}(Z \to \mu\mu) = 695 \pm 1(stat) \pm 9(syst) \pm 19(lumi) \pm 2(FSR)pb.$$

To compare, predictions for fiducial cross sections are also calculated, using the total cross section prediction by FEWZ and the acceptance from aMC@NLO simulation (Table 5.2), which are summarized in Table 5.13. As expected, the measured values also agree well with the predictions.

5.10.2 Differential results

Differential Z cross sections are determined from the unfolded distributions of the observable of interest by:

$$(\frac{d\sigma}{d\mathcal{O}^{truth}})_i = \frac{\hat{n}(\mathcal{O}^{truth}_i)}{w(\mathcal{O}^{truth}_i) \cdot \mathcal{L}}, \tag{5.28}$$

where the $w(\mathcal{O}^{truth}_i)$ is the width of bin i, while $\mathcal{L}$, the integrated luminosity.

Figure 5-17 and 5-18 present the differential cross section results and the com-

parison with theoretical predictions from aMC@NLO (NLO generator), POWHEG (NLO generator) and FEWZ (NNLO fixed-order calculation). The shaded (colored) bands around data points (prediction values) represent total uncertainties, while the solid lines are for the statistical components. For predictions, the total uncertainties include the PDF uncertainty and scale variations.

The precision of the differential cross section measurements is mostly dominated by systematic uncertainties for all phase space, except for high-p_T regions, where the results are statistically limited. Leading systematics for different observables and regions have been discussed in 5.9.2. As can be seen, in the $p_T^{\mu^+\mu^-}$ region of tens to hundreds of GeV, the NNLO FEWZ predictions agree with data very well within a few prercent, but below 10 GeV the fixed-order calculations fail and resummation is needed to restore the agreement with data. Correspondingly, the generator aMC@NLO and POWHEG, which match NLO matrix elements to LL parton shower, provide relatively good predictions at low and intermediate $p_T^{\mu^+\mu^-}$, but deviate at high $p_T^{\mu^+\mu^-}$, still covered by the theoretical uncertainties. This is also reflected in $\phi_\eta^\star$ distributions as which is highly correlated with $p_T^{\mu^+\mu^-}$.

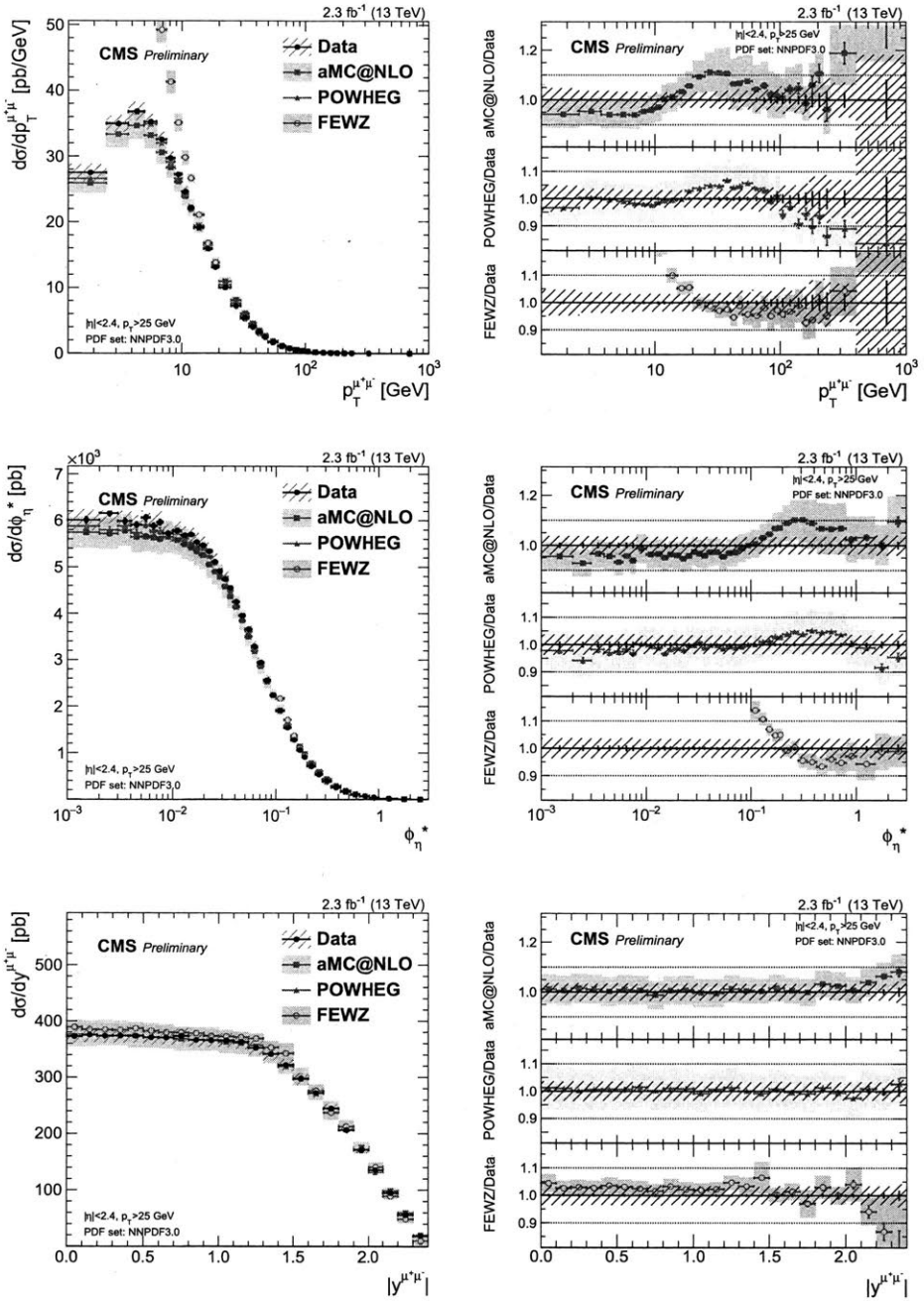


Figure 5-17: Unfolded distribution (left) and prediction-to-data ratio (right) of $p_T^{\mu^+\mu^-}$ (top), $\phi_\eta^\star$ (middle), $|y^{\mu^+\mu^-}|$ (bottom) observables.

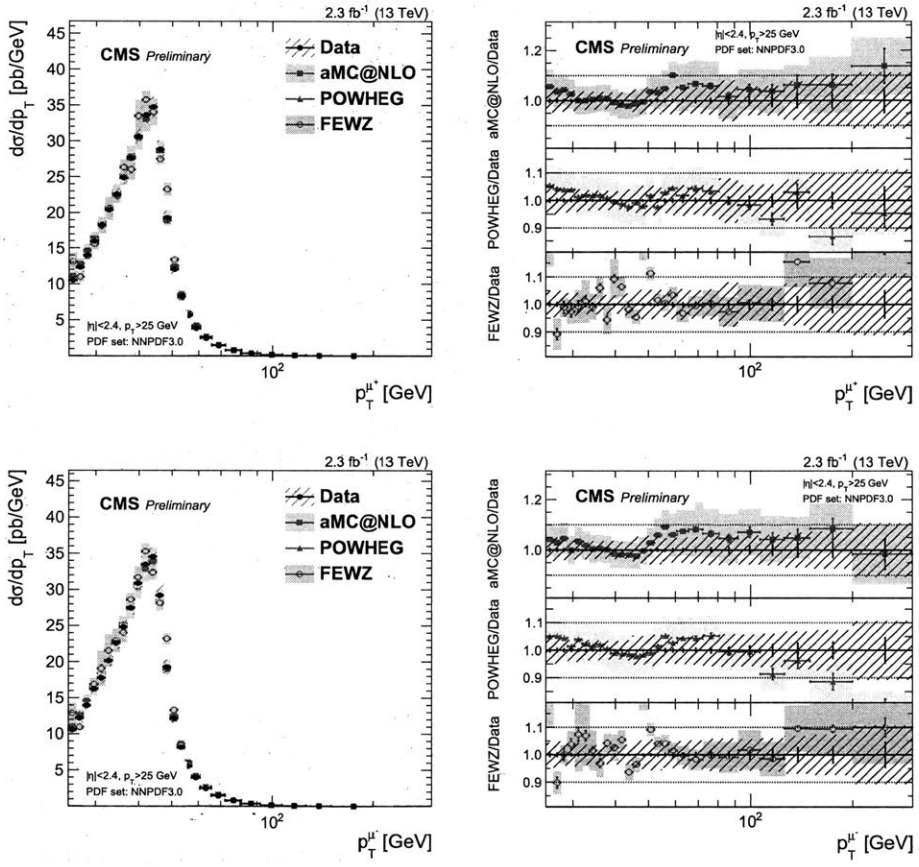


Figure 5-18: Unfolded distribution and prediction-to-data ratio (right) of $p_T^{\mu^+}$ (top) and $p_T^{\mu^-}$ (bottom) observables.

Chapter 6

Z Boson Counting as Luminosity Monitor

The precise and prompt measurement of luminosity is of great importance for the LHC. First, integrated luminosity recorded by the detector provides overall normalization of data used in all physics analyses. This is crucial, especially for precision measurements of SM parameters, such as W and Z boson and $t\bar{t}$ production cross section measurements [132, 133], where the luminosity uncertainty has become the leading experimental uncertainty. An accurate luminosity scale helps reduce uncertainties of those measurements, which as a result could help precisely quantify the high-order QCD corrections and constrain PDFs. Second, during data-taking period, the detector has responsibility to provide real-time feedback about instantaneous bunch-by-bunch luminosity delivered by LHC, in order to allow quick diagnosis and optimization of LHC performance.

To satisfy both precision and promptness requirement of luminosity measurement, CMS has developed a hybrid approach: online luminometers with preliminary calibration to provide fast response and offline luminometers with dedicated calibration for physics purpose. Five subdetectors have been exploited as luminosity monitors in CMS, taking the advantage of intrinsic linear relationship between instantaneous luminosity and the rates of corresponding observables measured on each luminometers. The principles of luminosity measurements in CMS are discussed in Section 6.1.

However, independent cross-checks of the absolute normalization as well as the linearity and stability of luminometers are strongly desired, in order to increase confidence of both online and offline luminosity measurements. Z bosons, one of the standard candles, are deployed as a new luminometers in this context. The method, which is based on the fact that Z boson production rate on LHC should be also proportional to the instantaneous luminosity, will be denoted as "Z counting" from now on. Dedicated arguments for motivation of establishing such a method can be found in Section 6.2.

Following the luminosity measurements in CMS and the motivation of Z counting, the rest of this chapter is organized as follows. First, the overall strategy of Z counting and workflow implementation are described in Section 6.3. In Section 6.4, data samples used in Z counting analyses and event selection are discussed. Measurements of the three components in Z counting analysis, acceptances, efficiencies and signal yields are then presented in details in Section 6.5, 6.6, and 6.7, respectively. Lastly, the Z counting measurement results are discussed in Section 6.8.

6.1 Luminosity measurements in CMS

By definition, the integrated luminosity $\mathcal{L}$ (time integral of the instantaneous luminosity $\mathscr{L}$) follows:

$$\mathcal{L} = \int \mathscr{L}(t)dt = \frac{N}{\sigma}, \tag{6.1}$$

where N is the number of events of interest and σ the corresponding cross section. Luminosity measurements in CMS are exactly based on the differential form of this definition:

$$\mathscr{L}(t) = \frac{R_{vis}}{\sigma_{vis}}, \tag{6.2}$$

where R_{vis} is the rate of observable measured at each luminometer, while σ_{vis} is the corresponding visible cross section, which is determined by absolute scale calibration using van der Meer scans.

The five CMS luminometers and the corresponding rate observables are described

in Section 6.1.1. The principles of absolute scale calibrations are elaborated in Section 6.1.2. Section 6.1.3 discusses the correction and uncertainties of R_{vis} and σ_{vis} measurements.

6.1.1 Luminometers and rate observables

Pixel Luminosity Telescope and Fast Beam Conditions Monitor

The Pixel Luminosity Telescope (PLT) [134] and the Fast Beam Conditions Monitor (BCM1F) [135] are two independent sub-detectors for luminosity measurements. They share the same carriages, which are located outside the pixel endcaps, about 1.8 m from the IP. The PLT consists of 8 telescopes at each end around the beam pipe, with 3 radiation-hard silicon pixel sensor layers for each telescope. For BCM1F, 24 sensors with three different technologies are exploited, including polycrystalline diamond (pCVD), single crystal diamond (sCVD) and silicon pixel.

To monitor luminosity, PLT uses the number of *triple coincidence* of the three layers as the rate observable and then applies zero-counting algorithm [1] to calculate the average rate. BCM1F follows a similar idea. Together with HF, these three systems utilize a separate data acquisition system to allow fast online bunch-by-bunch luminosity measurements. The position of PLT, BCM1F and HF luminometers with respect to the entire CMS detector are shown in Figure 6-1.

Hadronic forward calorimeter

The hadronic forward calorimeter (HF), introduced in Section 3.2.3, exploits two main approaches to measure luminosity. *HF-Occupancy* (HFOC) uses tower occupancy as the rate observable while *HF-E_T* (HFET) utilizes the linear relationship between the sum of transverse energy deposited in HF detector and the number of interactions per bunch crossing. Compared to HFOC, HFET has the advantage that it does not saturate at high instantaneous luminosity, and as a fact HFET was proved to provide the best offline measurements for most of the 2017 data-taking period, when the pixel

[1] Zero-counting algorithm assumes the triple coincidence counts follows a Poisson distribution, then the mean counts μ can be determined from the probability of zero counts $p(0)$ via $\mu = -lnp(0)$.

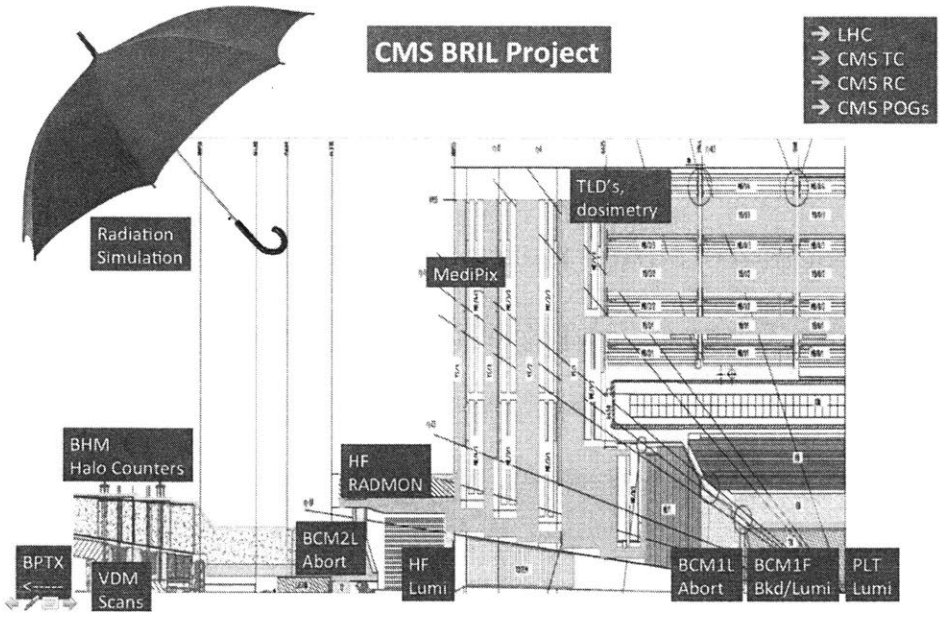


Figure 6-1: The PLT, BCM1F and HF luminometers in the CMS detector. [22]

cluster counting, which used to be the most reliable luminometer, experienced severe hardware failure [2].

Pixel cluster counting

The extremely high-granularity feature of the silicon pixel sensors ($\sim 10^7$) results in very low occupancy in the CMS pixel detector, which as a result could provide good linearity and stability for high-precision luminosity measurements.

Number of pixel clusters per event is the rate variable for this method, thus this luminometer is denoted as pixel cluster counting (PCC). PCC was the primary offline luminometer to provide the integrated luminosity for data taken in 2015 and 2016.

Drift tube

The drift tube (DT) in barrel muon system is also deployed to provide reference for offline luminosity measurements. The corresponding rate variable is the number of muon segments in muon barrel track finder (MBTF). The DT also has the low-occupancy feature, but unlike PCC, the rate is relatively low, so that DT cannot

[2]DC-DC converter issues in 2017 caused exclusion of a large number pixel modules.

provide per-bunch measurements.

6.1.2 Absolute scale calibration and Van der Meer scans

To interpret the measured rates from different luminometers into luminosity values, the corresponding absolute scale calibrations, i.e. visible cross sections σ_{vis} in Equation 6.2, are determined for each rate observable through an indispensable technique that Simon van der Meer developed during his work at the Intersecting Storage Rings (ISR), van der Meer (VdM) scans [136].

The principle of calibration with VdM scan methodology is illustrated as follows [23]. Given number of protons (N_1, N_2) and proton density ($\rho_1(x,y)$, $\rho_2(x,y)$) in two colliding bunches, and orbit frequency f, if assuming no crossing angle, the instantaneous luminosity of colliding two bunches separated by (Δx, Δy) can be expressed by:

$$\mathscr{L}(\Delta x, \Delta y) = N_1 N_2 f \int_{-\infty}^{+\infty} \rho_1(x,y)\rho_2(x+\Delta x, y+\Delta y)dxdy. \tag{6.3}$$

If further assuming there is no x-y correlation in the proton density functions[3], i.e, $\rho_i(x,y) = \rho_{i,x}(x)\rho_{i,y}(y)$, where i = 1,2 for colliding bunches:

$$\mathscr{L}(\Delta x, \Delta y) = N_1 N_2 f [\int_{-\infty}^{+\infty} \rho_{1,x}(x)\rho_{2,x}(x+\Delta x)dx][\int_{-\infty}^{+\infty} \rho_{1,y}(y)\rho_{2,y}(y+\Delta y)dy]. \tag{6.4}$$

The integral of proton density can be replaced [4]by the ratio of luminosity quantities and then scaled into measurable rate quantities, and further rewritten into convolved

[3]This is not necessarily true in reality. As a fact, the existence of x-y correlation introduces one of the leading systematics in luminosity measurements. See Section 6.1.3 for details.

widths (beam overlap widths) Σ_x and Σ_y:

$$\begin{aligned}\mathscr{L}(\Delta x, \Delta y) &= N_1 N_2 f \frac{\mathscr{L}(\Delta x_0, \Delta y_0)}{\int_{-\infty}^{+\infty} \mathscr{L}(\Delta x_0, \Delta y) d(\Delta y)} \frac{\mathscr{L}(\Delta x_0, \Delta y_0)}{\int_{-\infty}^{+\infty} \mathscr{L}(\Delta x, \Delta y_0) d(\Delta x)} \\ &= N_1 N_2 f \frac{R(\Delta x_0, \Delta y_0)}{\int_{-\infty}^{+\infty} R(\Delta x_0, \Delta y) d(\Delta y)} \frac{R(\Delta x_0, \Delta y_0)}{\int_{-\infty}^{+\infty} R(\Delta x, \Delta y_0) d(\Delta x)} \\ &= \frac{N_1 N_2 f}{2\pi \Sigma_x \Sigma_y}.\end{aligned} \tag{6.5}$$

Then with Equation 6.2, the visible cross section can be then calculated with:

$$\sigma_{vis} = \frac{2\pi \Sigma_x \Sigma_y R(\Delta x_0, \Delta y_0)}{N_1 N_2 f}. \tag{6.8}$$

Experimentally, two separate scans over x and y direction are performed. Rates on luminometers are measured at certain intervals of beam separation and scan curves are fitted with a double Gaussian after subtracting backgrounds, which are due to machine introduced background (MIB) or detector noise, or alternatively, fitted with a double Gaussian after subtracted backgrounds. Example of such a fit to scan curves can be found in Figure 6-2.

Then the beam overlap widths Σ_x and Σ_y can be derived from the fitting parameters of the two Gaussian functions (amplitudes A_1,A_2 and widths σ_1, σ_2) by:

$$\Sigma = \frac{A_1 \sigma_1 + A_2 \sigma_2}{A_1 + A_2}, \tag{6.9}$$

while the rest of Equation 6.8 can be derived from amplitudes of the normalized scan curves. Since the beam overlap widths are characteristic features of the beam, they do not depend on either luminometers or rate observables. In practice, VdM scans are

[4]Integrate Equation 6.4 over Δx, while $\Delta y \equiv \Delta y_0$ is fixed:

$$\int_{-\infty}^{+\infty} \mathscr{L}(\Delta x, \Delta y_0) d(\Delta x) = N_1 N_2 f [\int_{-\infty}^{+\infty} \rho_{1,y}(y) \rho_{2,y}(y + \Delta y) dy]. \tag{6.6}$$

Substitute Equation 6.6 into Equation 6.4, the integral of proton density can be expressed by:

$$\int_{-\infty}^{+\infty} \rho_{1,x}(x) \rho_{2,x}(x + \Delta x_0) dx = \frac{\mathscr{L}(\Delta x_0, \Delta y_0)}{\int_{-\infty}^{+\infty} \mathscr{L}(\Delta x_0, \Delta y) d(\Delta y)}. \tag{6.7}$$

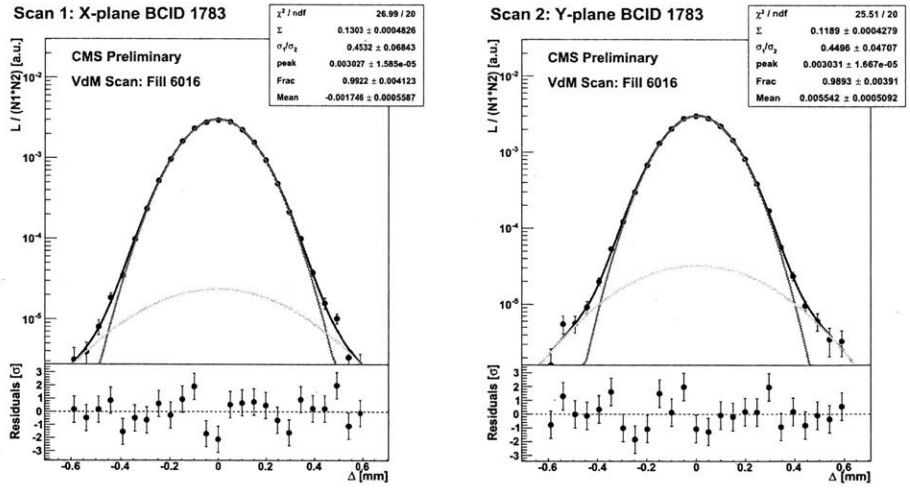


Figure 6-2: An example of double Gaussian fit to normalized scan curves for a single bunch recorded by PLT for a scan in the x (left) and y (right) direction. [23]

Luminometer	Measured σ_{vis}
BCM1F	$225.6 \pm 1.3\ \mu b$
HFET	$2644.9 \pm 24.1\ \mu b$
HFOC	$838.5 \pm 5.7\ \mu b$
PLT	$297.7 \pm 1.8\ \mu b$
PCC	$4.719 \pm 0.035\ b$

Table 6.1: Summary of measured visible cross sections for different luminometers. The uncertainty quoted includes variation among scans. [23]

done multiple times and with different bunch pairs. The overall visible cross section is then determined by taking the weighted average over bunch crossings and scans. A summary of the measured visible cross sections for different luminometers in 2017 data-taking period can be found in Table 6.1.

6.1.3 Systematic uncertanties

The systematic uncertainties of luminosity measurements can be grouped into two categories: normalization uncertainties on the absolute scale calibration, which is introduced by performing VdM scans and independent of luminometers, and the in-

tegration uncertainties on luminometer-specific rate measurements.

The sources of uncertainties considered for both categories are listed as follows and the corresponding correction and systematic results of luminosity measurement for 2017 data-taking period are summarized in Table 6.2.

Sources of normalization systematics

- Length scale calibration. The actual beam separation might be different from the nominal beam separation determined by beam steering magnets. Beam spot positions that are reconstructed by using information from the inner tracker detectors are used to calibrate the nominal beam separations.

- Orbit drift. Beam positions are measured right before, at head-on, and right after each scan to monitor the orbit drift during scans. This is done by Diode ORbit and OScillation (DOROS) Beam Position Monitor (BPM) system and arc BPMs.

- X-y correlation. Proton density functions in colliding bunches are not x-y factorizable. Beam imaging scans [137] are taken to reconstruct proton density using measured vertex distributions.

- Beam-beam effect and dynamic-$\beta^{\star}$. Deflection and de-focusing effect due to electromagnetic forces are parameterized as a function of beam separation.

- Variations among different scans, colliding bunch pairs and luminometers.

Sources of integration systematics

- Luminometer-specific effect.

  - For instance, HFET was the primary offline luminometer for 2017. Afterglow effects (out-of-time response due to detector material activation) were estimated by modeling the in-time bunch response leaking into next bunch and forming long tail over bunches.

- Stability and linearity are evaluated from luminosity ratio between luminometers as a function of time/instantaneous luminosity.

- CMS DAQ deadtime: the time that CMS does not record data due to busy triggers.

	Systematic	Correction (%)	Uncertainty (%)
Normalization	Length scale	-0.9	0.3
	Orbit drift	-	0.2
	x-y correlations	+0.8	0.8
	Beam-beam deflection	+1.6	0.4
	Dynamic-$\beta^{\star}$	-	0.5
	Beam current calibration	-	0.3
	Ghosts and satellites	-	0.1
	Scan to scan variation	-	0.9
	Bunch to bunch variation	-	0.1
	Cross-detector consistency	0.4–0.6	0.6
Integration	Afterglow (HF)	-	0.2⊕0.3
	Cross-detector stability	-	0.5
	Linearity	-	1.5
	CMS deadtime	-	0.5
	Total		2.3

Table 6.2: Summary of the corrections and systematic uncertainties of CMS luminosity measurements for 2017 data-taking period. [23]

6.2 Motivation of introducing Z counting

Despite the existence of aforementioned various luminometers in CMS, other luminosity determination methods are still required for both the LHC and CMS detector. First, during LHC running, significant luminosity difference have had been observed unexpectedly between CMS and ATLAS. See Figure 6-3 for example. However, we didn't know the discrepancy was real (i.e. the LHC delivered different amount of luminosities to CMS and ATLAS, which could cause different physics reach for the two detectors in long run) or "artificial" (i.e. mismeasurement by one or both experiments), since those measurements are based on online luminometers, which are

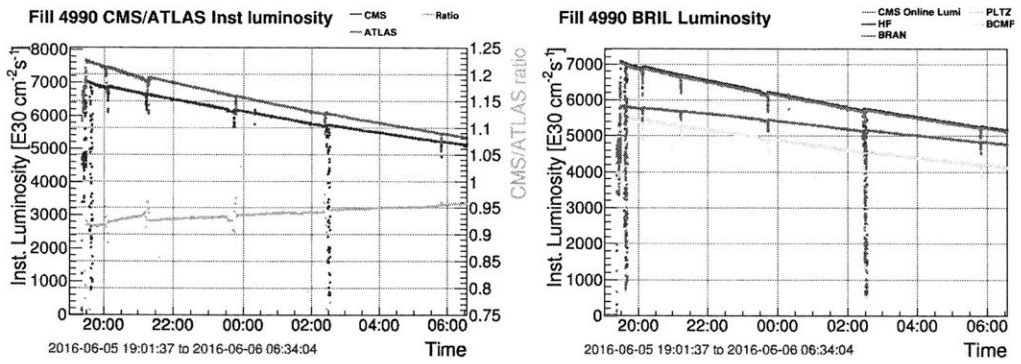


Figure 6-3: Instantaneous luminosity discrepancy between CMS and ATLAS online measurements (left) and CMS online luminometers (right) in Fill 4990. [24]

potentially affected by mis-calibration and could have large uncertainty as high as ~10%. Precise (~3%) calibration from VdM scans could be achieved but take months to finalize and not acceptable for prompt LHC diagnosis and optimization. Therefore, luminosity information with intermediate precision and timescale of 1-2 weeks after data taken is at request to help monitor online measurements.

Second, a new luminometer which has comparable precision as offline luminometers and does not depend on VdM scan is desired for CMS luminosity measurements as well. Before linearity was improved for HFET in 2017, PCC was the only luminometer which could provide luminosity measurements at 2-3% precision level. It is necessary to find another approach with similar accuracy to cross check PCC results. All standard luminometers take absolute scale calibration from VdM scans. Although dedicated studies have been performed to estimate the uncertainty, the existence of an unknown but non-negligible source of systematics is still possible.

Given the above, $Z \rightarrow \mu^+\mu^-$ events are ideal for being a rate observable as luminosity monitor. First of all, as a standard candle, Z boson production process on LHC is theoretically well understood. NNLO predictions of Z boson production cross sections with percent level uncertainty are available and in good agreement with experimental results [132]. Secondly, large Z boson production cross sections on LHC and high, stable muon detection efficiencies make it possible to accumulate enough event yields to fit in the time budget required for luminosity monitoring, e.g. ~20 minutes. Thirdly, clean experimental signature of di-muon final state and small

background contamination further reduce the experimental uncertainty, in order to achieve the desired precision level.

The Z counting method has two unique features. As Z boson production cross sections are same at IP1 and IP5, the Z boson production rate itself is a comparable observable between ATLAS and CMS. In addition, as muon reconstruction, selection and trigger efficiencies are intermediate products in Z counting analyses, the efficiency results in such fine time interval provides complementary information for data quality monitoring of muon system as well.

6.3 Analysis strategy and workflow overview

The overall analysis strategy exploited in the inclusive Z boson cross section measurements in Chapter 5 also applies to Z counting analysis. Compared to the Z boson cross section measurement, the most different and challenging point is that Z counting analysis has to be done continuously per very short time interval and promptly to provide results after data-taking. This requires not only a well-established analysis strategy but also a reliable framework for automation.

In the Z counting analysis, instead of measuring total Z boson cross section $\sigma_{tot}^{Z\to\ell^+\ell^-}$ or fiducial Z cross section $\sigma_{fid}^{Z\to\ell^+\ell^-}$ to provide indirect luminosity information, it is more convenient to define the following variables as observables for luminosity monitoring purpose.

6.3.1 Z counting observables

Recorded and delivered luminosity measured by Z counting

Given Equation 5.7, a straightforward approach to determine the *recorded luminosity measured by Z counting*, which is a function of time t, $\mathcal{L}_{rec}^{Z-counting}(t)$, is to simply twist the equation:

$$\mathcal{L}_{rec}^{Z-counting}(t) = \frac{N_{obs}^{Z\to\ell^+\ell^-}(t) - N_{sel}^{bkg}(t)}{A \cdot \epsilon(t) \cdot \sigma_{tot}^{Z\to\ell^+\ell^-}}, \tag{6.10}$$

where $N_{obs}^{Z\to\ell^+\ell^-}$, N_{sel}^{bkg}, A and ϵ are defined as same as in Chapter 5, but $N_{obs}^{Z\to\ell^+\ell^-}(t)$, $N_{sel}^{bkg}(t)$ and $\epsilon(t)$ can vary, as the detector condition changes from time to time, and thus have to be measured repeatedly for every time interval. While the $\sigma_{tot}^{Z\to\ell^+\ell^-}$ can either be taken from theoretical predictions or existing experimental measurements. Then, $\mathcal{L}_{rec}^{Z-counting}(t)$ can be compared with the corresponding recorded luminosity by other luminometers for the same data-taking period, i.e, $\mathcal{L}_{rec}^{PLT}(t)$, $\mathcal{L}_{rec}^{BCM1F}(t)$, $\mathcal{L}_{rec}^{HF}(t)$, $\mathcal{L}_{rec}^{DT}(t)$ or $\mathcal{L}_{rec}^{PCC}(t)$, if available, to help monitor CMS online and offline luminosity values.

In addition, to provide luminosity information for LHC diagnosis, the *delivered luminosity* is further determined by taking correction of the deadtime of the detector:

$$\mathcal{L}_{del}^{Z-counting}(t) = \frac{N_{obs}^{Z\to\ell^+\ell^-}(t) - N_{sel}^{bkg}(t)}{A\cdot\epsilon(t)\cdot\sigma_{tot}^{Z\to\ell^+\ell^-}} \times \frac{1}{1 - deadtime[\%](t)}. \tag{6.11}$$

Z counts and Z rates

As CMS and ATLAS agreed on a same fiducial region for Z counting studies, both A and $\sigma_{tot}^{Z\to\ell^+\ell^-}$ are constant at both detectors, the rest of the equation is *delivered Z counts*, or *Z rates* if further dividing by time T, in the defined fiducial region:

$$N_{fid,del}^{Z\to\ell^+\ell^-}(t) = \frac{N_{obs}^{Z\to\ell^+\ell^-}(t) - N_{sel}^{bkg}(t)}{\epsilon(t)} \times \frac{1}{1 - deadtime[\%](t)} \tag{6.12}$$
$$(= \mathcal{L}_{del}^{Z-counting}(t)\cdot A\cdot\sigma_{tot}^{Z\to\ell^+\ell^-})$$

$$R_{fid,del}^{Z\to\ell^+\ell^-}(t) = \frac{N_{obs}^{Z\to\ell^+\ell^-}(t) - N_{sel}^{bkg}(t)}{\epsilon(t)} \times \frac{1}{T} \times \frac{1}{1 - deadtime[\%](t)}, \tag{6.13}$$
$$(= \mathscr{L}_{del}^{Z-counting}(t)\cdot A\cdot\sigma_{tot}^{Z\to\ell^+\ell^-}).$$

These two observables are in fact used to compare delivered luminosity to assist LHC machine diagnosis in the joint CMS-ATLAS synchronization exercise. Since the theoretical and other experimental uncertainties on A and $\sigma_{tot}^{Z\to\ell^+\ell^-}$ do not enter the measurements, Z counts and Z rates can be determined at high precision.

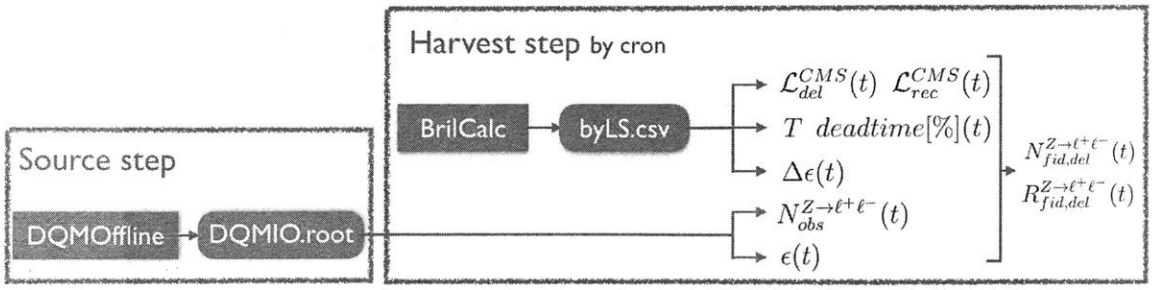


Figure 6-4: A sketch of Z counting workflow.

6.3.2 Workflow automation

Automation of the Z counting in muon channel has been realized with the aid of the CMS DQMOffline framework [138] and Cron job scheduler. The implementation includes two steps: source step and harvest step. A flow diagram is sketched in Figure 6-4 to illustrate the steps and main components involved in the workflow.

Let us first introduce some terms to better illustrate the workflow. A *LHC fill* is defined as the period from beam injection to the beam dump. Within a same LHC fill, there are one or more *CMS runs*, which consist of a number of *lumisections*. Each lumisection lasts 23.3 seconds.

Source step

Source step is executed in DQMOffline framework. The Z counting module has been implemented in the DQM sequence for SingleMuon primary dataset[5], where histograms for calculating observed Z yields $N_{obs}^{Z\to\mu^+\mu^-}(t)$ and efficiencies $\epsilon(t)$ are booked and filled. The output file containing desired histograms, DQMIO.root, is produced for every CMS run and usually ready 3-5 days after data taking, which is the time needed for dataset preparation.

[5]Data recorded via triggers that select events with at least one muon object.

```
#Data tag : v2 , Norm tag: None
#run:fill,ls,time,hltpath,delivered(/ub),recorded(/ub),avgpu,source
321305:7056,18:18,08/15/18 02:01:36,HLT_IsoMu27_v16,390641.216,381945.521,53.4,HFET
321305:7056,19:19,08/15/18 02:01:59,HLT_IsoMu27_v16,434793.288,327176.345,41.7,HFET
321305:7056,20:20,08/15/18 02:02:22,HLT_IsoMu27_v16,389555.261,302480.396,53.8,HFET
321305:7056,21:21,08/15/18 02:02:46,HLT_IsoMu27_v16,439728.572,422538.329,48.8,HFET
321305:7056,22:22,08/15/18 02:03:09,HLT_IsoMu27_v16,450370.398,438665.109,52.8,HFET
321305:7056,23:23,08/15/18 02:03:32,HLT_IsoMu27_v16,458726.538,446312.319,55.1,HFET
321305:7056,24:24,08/15/18 02:03:56,HLT_IsoMu27_v16,457533.500,432634.310,54.9,HFET
321305:7056,25:25,08/15/18 02:04:19,HLT_IsoMu27_v16,135350.385,133003.998,54.9,HFET
321305:7056,26:26,08/15/18 02:04:42,HLT_IsoMu27_v16,160575.220,158504.363,12.9,HFET
321305:7056,27:27,08/15/18 02:05:05,HLT_IsoMu27_v16,320044.886,313916.587,30.8,HFET
321305:7056,28:28,08/15/18 02:05:29,HLT_IsoMu27_v16,438739.730,427660.554,46.4,HFET
321305:7056,29:29,08/15/18 02:05:52,HLT_IsoMu27_v16,384926.751,376542.907,52.0,HFET
321305:7056,30:30,08/15/18 02:06:15,HLT_IsoMu27_v16,222675.671,219370.138,34.7,HFET
321305:7056,31:31,08/15/18 02:06:39,HLT_IsoMu27_v16,86691.251,85807.331,18.5,HFET
321305:7056,32:32,08/15/18 02:07:02,HLT_IsoMu27_v16,280150.900,273625.977,18.9,HFET
321305:7056,33:33,08/15/18 02:07:25,HLT_IsoMu27_v16,58422.355,57877.563,3.6,HFET
321305:7056,34:34,08/15/18 02:07:49,HLT_IsoMu27_v16,186646.232,184007.425,9.9,HFET
321305:7056,35:35,08/15/18 02:08:12,HLT_IsoMu27_v16,391571.963,382669.294,42.8,HFET
321305:7056,36:36,08/15/18 02:08:35,HLT_IsoMu27_v16,416927.898,407149.208,54.4,HFET
321305:7056,37:37,08/15/18 02:08:59,HLT_IsoMu27_v16,228360.746,224790.987,47.0,HFET
321305:7056,38:38,08/15/18 02:09:22,HLT_IsoMu27_v16,73847.433,73128.532,12.3,HFET
321305:7056,39:39,08/15/18 02:09:45,HLT_IsoMu27_v16,304306.058,296740.253,4.5,HFET
321305:7056,40:40,08/15/18 02:10:08,HLT_IsoMu27_v16,453147.340,435053.304,54.4,HFET
321305:7056,41:41,08/15/18 02:10:32,HLT_IsoMu27_v16,452900.983,434774.940,54.4,HFET
321305:7056,42:42,08/15/18 02:10:55,HLT_IsoMu27_v16,451645.798,439957.628,54.3,HFET
321305:7056,43:43,08/15/18 02:11:18,HLT_IsoMu27_v16,451426.167,439673.573,54.1,HFET
321305:7056,44:44,08/15/18 02:11:42,HLT_IsoMu27_v16,452120.209,434140.086,54.2,HFET
321305:7056,45:45,08/15/18 02:12:05,HLT_IsoMu27_v16,451925.482,440269.318,54.3,HFET
321305:7056,46:46,08/15/18 02:12:28,HLT_IsoMu27_v16,451004.296,439398.260,54.1,HFET
321305:7056,47:47,08/15/18 02:12:52,HLT_IsoMu27_v16,450614.068,439033.118,54.1,HFET
321305:7056,48:48,08/15/18 02:13:15,HLT_IsoMu27_v16,450106.922,438521.101,54.1,HFET
321305:7056,49:49,08/15/18 02:13:38,HLT_IsoMu27_v16,449563.217,431753.372,54.0,HFET
321305:7056,50:50,08/15/18 02:14:02,HLT_IsoMu27_v16,448788.315,437322.280,53.9,HFET
```

Figure 6-5: An example of byLS.csv from BRIL tools, exhibiting CMS run number, LHC fill number, lumisection number, time, HLT path, delivered luminosity, recored luminosity, average number of pile-up and the source of luminosity measurements for each lumisection per line.

Harvest step

In addition to $N_{obs}^{Z\to\mu^+\mu^-}(t)$ and $\epsilon(t)$, time stamps and deadtime information are also required to determine Z counts, $N_{fid,del}^{Z\to\mu^+\mu^-}(t)$, and Z rates, $R_{fid,del}^{Z\to\mu^+\mu^-}(t)$. The efficiency correction $\Delta\epsilon(t)$ needs to be applied and is parameterized as a function of the mean number of pile-up. Online luminosity measurements are also desired to check statistics when merging lumisections. A tool developed under the CMS Beam Radiation Instrumentation and Luminosity (BRIL) project [139, 140] is exploited to provide all those information per lumisection in the format of byLS.csv, which is the other input to the harvest step. An example of a byLS.csv file can be found in Figure 6-5, which illustrates CMS run number, LHC fill number, lumisection number, time, HLT path, delivered luminosity, recored luminosity, average number of pile-up and the source of luminosity measurements for each lumisection per line.

Once the histograms from source step are accessible, they will be merged according to desired size of the data chunk. The size parameters are chosen in order to ensure the statistical uncertainty for each measurement is smaller than 1%. Currently, the size is set to be every 50 lumisections, with an integrated luminosity of 10 pb^{-1} as threshold to ensure enough number of events. That is, if a data chunk has integrated luminosity less than 10 pb^{-1}, it will be merged with next chunk until the entire merged chunk has reached 10 pb^{-1}. This usually happens towards the end of a long fill, where the instantaneous luminosity has dropped dramatically before the fill is dumped. Then, for every data chunk, one Z counting measurement is performed and corresponding results are posted to webpages to enable cross checks of the CMS online luminometers and synchronization with ATLAS for further detector and machine diagnosis.

Accessing DQM histograms, creating BRIL byLS files and running Z counting analyses to post result plots are all automatically done via Cron, a time-base job scheduler. Harvest step usually takes several minutes to hours, depending on the data size.

6.4 Data samples and event selection

6.4.1 Data samples and triggers

The Z counting studies analyzed data from 2015 and 2016 with the offline framework. Since the beginning of the 2017 data-taking period, the Z counting automation implementation has been deployed in production and effective. The Z candidate events are selected from all events collected by the lowest p_T threshold, non pre-scaled isolated single muon trigger. For 2017 and 2018 data-taking period, it is *IsoMu27*, which requires at least one muon with $p_T > 27$ GeV and $|\eta| < 2.4$, and passing loose isolation and identification requirements.

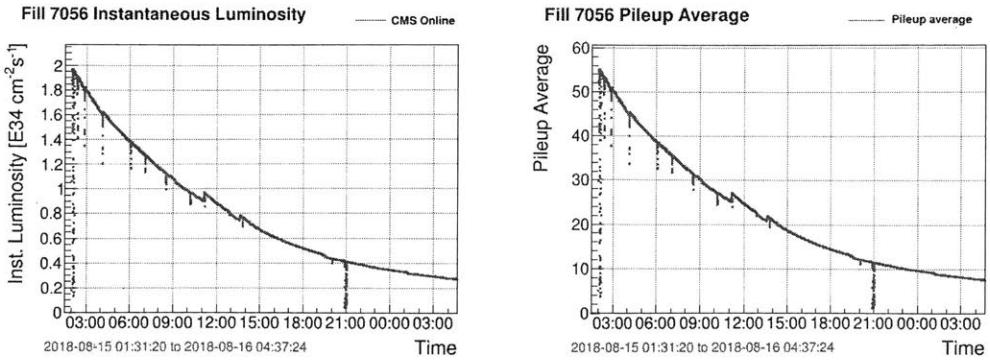


Figure 6-6: Instantaneous luminosity and average pile-up as a function of time in Fill 7056. [24]

CMS Run	Begin time	End time	$\mathcal{L}_{rec}^{CMS}[pb^{-1}]$	nMeas
321305	2018.08.15 01:53:49	2018.08.15 19:07:39	612.38	41
321310	2018.08.15 19:37:36	2018.08.15 20:44:49	16.85	2
321311	2018.08.15 20:47:33	2018.08.15 20:52:10	0.48	1
321312	2018.08.15 20:54:30	2018.08.16 01:50:33	61.57	6
321313	2018.08.16 01:54:20	2018.08.16 04:38:07	27.74	3

Table 6.3: LHC Fill 7056 report. [24]

6.4.2 Example LHC fill

To illustrate, LHC Fill 7056 is chosen from now on to explain the intermediate steps and results. It is a relatively long fill (26.7 hours) and thus spans a wide range of instantaneous luminosity values and pile-up conditions, see Figure 6-6. More details for this fill including begin and end time, recorded luminosity by CMS online luminometers, $\mathcal{L}_{rec}^{CMS}$, and number of Z counting measurements, $nMeas$, for each CMS run can be found in Table 6.3.

6.4.3 Event selection

The Z event selection in Z counting analysis is based the Z boson cross section measurements [132], described in Section 5.3. That is, two opposite-charge high-p_T muons come from a good primary vertex, and the reconstructed di-muon invariant mass is in a mass window around Z boson mass.

However, there are several differences from the nominal selection requirements in

Section 5.3.

- No further isolation requirements on muon objects. Pile-up condition varies dramatically from the beginning to the end of a typical fill (Figure 6-6). To avoid the pitfall that isolation variables and thus results would be affected by pile-up, additional cuts on those variables were removed.

- $p_T > 30$ GeV, to adapt to the 2017 and 2018 trigger p_T threshold and avoid the turn-on inefficiency.

- $|\eta| < 0.9$ is added for barrel measurements only, in addition to $|\eta| < 2.4$ for inclusive measurements, in order to allow a cross check and reduce systematic uncertainties.

- $66 < M_{\mu^+\mu^-} < 116$ GeV is used to define the Z cross section. In order to allow direct comparison of Z yield and Z rate for the joint CMS-ATLAS synchronization exercise, the Z cross section is defined to adapt to ATLAS convention, instead of the usual CMS convention of 60 to 120 GeV.

6.4.4 Implementation and $N_{obs}^{Z\to\mu^+\mu^-}(t)$ results

Three one-dimensional histograms for barrel-barrel, endcap-endcap, and inclusive di-muon final states, are booked for each CMS run and filled with selected events to determine $N_{obs}^{Z\to\mu^+\mu^-}(t)$. An example of such a histogram, showing the inclusive $N_{obs}^{Z\to\mu^+\mu^-}(t)$ as a function of lumisections in CMS Run 321305, LHC Fill 7056, can be found in Figure 6-7.

6.5 Acceptance

As described in 5.4, acceptances are determined with simulation. The simulated sample used for 2017 and 2018 Z counting analysis is generated using MadGraph [73] generator with MLM merging scheme [76].

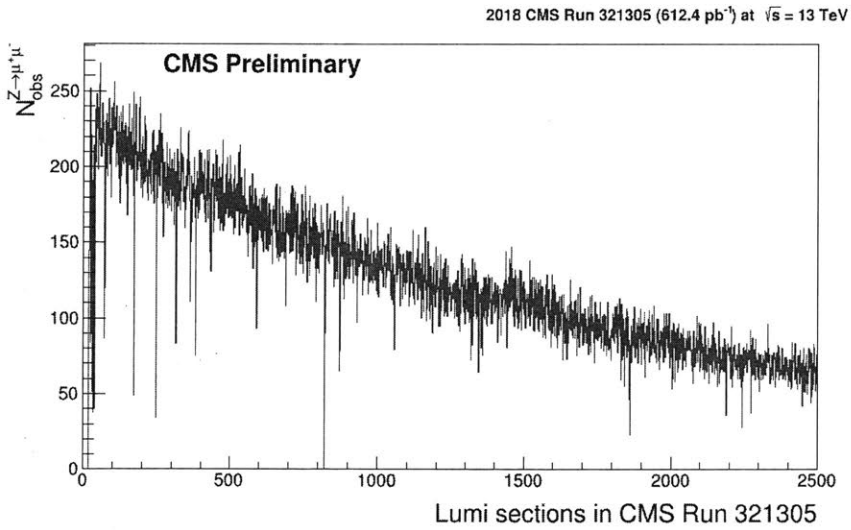


Figure 6-7: An example of $N_{obs}^{Z\to\mu^+\mu^-}(t)$ histogram for CMS Run 321305, LHC Fill 7056.

Category	Total	BB	BE	EE	Inclusive
Number of events	8724920	679707	1022560	920835	2623102
Acceptance	1	0.077904	0.117200	0.105541	0.300644

Table 6.4: Pre-FSR generator level acceptance.

According to the definition of acceptance in Equation 5.4 and fiducial region definition in Section 6.4.3, we can define the following categories of $Z \to \mu^+\mu^-$ events:

- Total: $66<M_{\mu^+\mu^-}<116$ GeV, no p_T, η cuts,
- Barrel-Barrel (BB): $66<M_{\mu^+\mu^-}<116$ GeV, $p_{T,\mu^+}, p_{T,\mu^-}>30$ GeV, $|\eta_{\mu^+}|, |\eta_{\mu^-}|<0.9$,
- Endcap-Endcap (EE): $66<M_{\mu^+\mu^-}<116$ GeV, $p_{T,\mu^+}, p_{T,\mu^-}>30$ GeV, $0.9<|\eta_{\mu^+}|, |\eta_{\mu^-}|<2.4$,
- Barrel-Endcap (BE): $66<M_{\mu^+\mu^-}<116$ GeV, $p_{T,\mu^+}, p_{T,\mu^-}>30$ GeV, $|\eta_{\mu^+}|<0.9$, $0.9<|\eta_{\mu^-}|<2.4$ or $|\eta_{\mu^-}|<0.9$, $0.9<|\eta_{\mu^+}|<2.4$,
- Inclusive: $66<M_{\mu^+\mu^-}<116$ GeV, $p_{T,\mu^+}, p_{T,\mu^-}>30$ GeV, $|\eta_{\mu^+}|, |\eta_{\mu^-}|<2.4$.

Simple counting in number of events in each category is performed using simulated sample, and corresponding ratio to the number of events in Category Total is the acceptance by definition. Results are summarized in Table 6.4.

6.6 Efficiency

Efficiency measurements are essential to the Z counting analyses. The overall strategy exploited here is same as the Z boson cross section measurements described in Section 5.5. That is, the event efficiencies are obtained from simulation and corrections are applied via single muon trigger, selection, and reconstruction efficiency data-to-simulation scale factors. Single muon efficiencies are measured with tag-and-probe technique (see Section 5.5.1) as well. Although the number of efficiency measurements performed in Z counting is huge (for instance, 53 measurements are done for Fill 7056), no shortcut on methodology is taken in order to remain the precision. Technical differences are summarized as follows:

- 2 $|\eta|$ bins (Barrel and Forward) are used to measure single muon efficiencies, instead of multiple (p_T, η) bins, in order to trade off small event yields ($\sim 10pb^{-1}$) per measurements.
- Efficiency corrections are parameterized offline using simulation before data taking as a function of pile-up, in order to fit in automation. More details can be found in Section 6.6.2.
- Dedicated pre-fit procedure and fit χ^2 monitor are enabled in order to ensure efficiency fit quality and robustness in automation.

6.6.1 Single muon trigger, selection, reconstruction efficiency

The total single muon efficiency is factorized in the same way as in Equation 5.12. As described in Section 6.4.3, no isolation requirement is applied in the Z counting selection. Therefore, the factorization formula can be revised into:

$$\epsilon_\mu = \epsilon_{trigger} \times \epsilon_{tracking+selection} \times \epsilon_{standalone}, \tag{6.14}$$

where the selection includes muon identification criteria only. The corresponding tag-and-probe definitions are also revised accordingly.

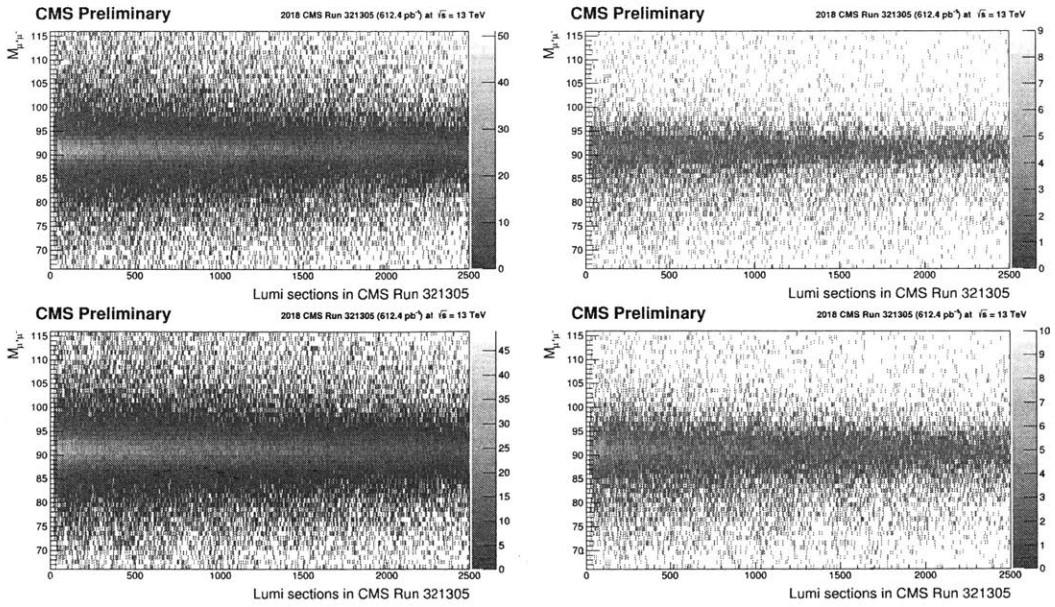


Figure 6-8: Example of $\epsilon_{trigger}$ histograms, for barrel passing (top-left), barrel failing (top-right), endcap passing (bottom-left) and endcap failing (bottom-right) probes, in CMS Run 321305, LHC Fill 7056.

Tag-and-probe implementation

In the Z counting automation, tag-and-probe invariant mass distributions are filled into 12 ($= 3 \times 2 \times 2$) two-dimensional histograms for each CMS run:

- 3 efficiency types: $\epsilon_{trigger}$, $\epsilon_{tracking+selection}$ and $\epsilon_{standalone}$,
- 2 $|\eta|$ regions: barrel and endcap,
- 2 event categories: passing and failing.

Examples of such tag-and-probe histograms are shown in Figure 6-8, 6-9 and 6-10 for $\epsilon_{trigger}$, $\epsilon_{tracking+selection}$ and $\epsilon_{standalone}$ efficiency type, respectively.

At harvest step, those histograms are merged and projected into Y axis to recover the traditional tag-and-probe invariant mass distributions, then standard countings/fittings are performed to extract the efficiencies. The functions used to fit backgrounds are the same as in Z boson cross section measurements, *exponential* for $\epsilon_{tracking+selection}$ and *quadratic* for $\epsilon_{standalone}$. The function used to fit signals for $\epsilon_{tracking+selection}$, however, is *BW*CB* instead of *MC*Gaus*. This is because the two models have comparable fitting performance for $\epsilon_{tracking+selection}$, but *BW*CB* does

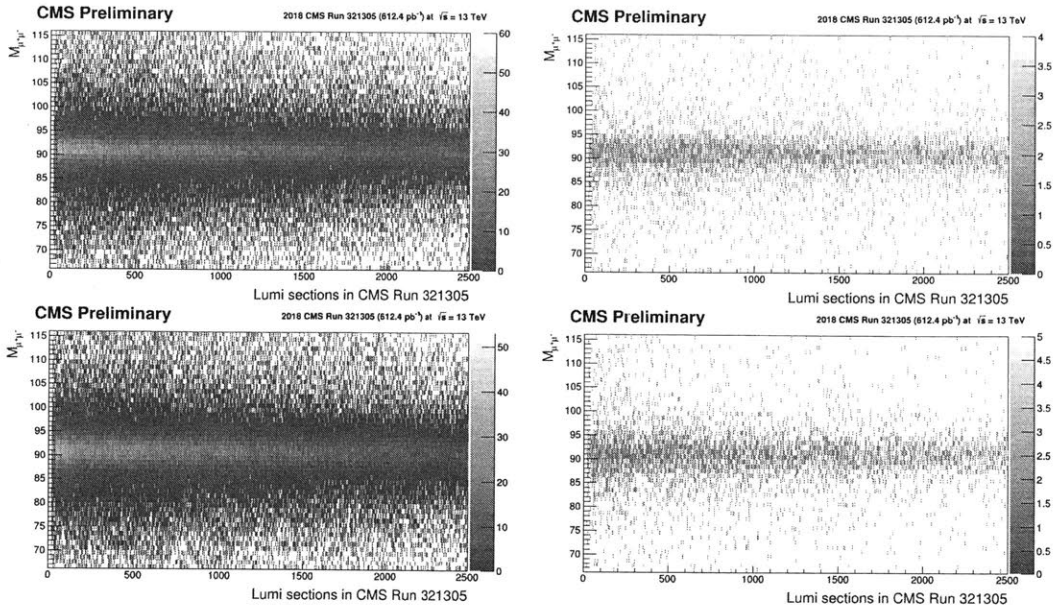


Figure 6-9: Example of $\epsilon_{tracking+selection}$ histograms, for barrel passing (top-left), barrel failing (top-right), endcap passing (bottom-left) and endcap failing (bottom-right) probes, in CMS Run 321305, LHC Fill 7056.

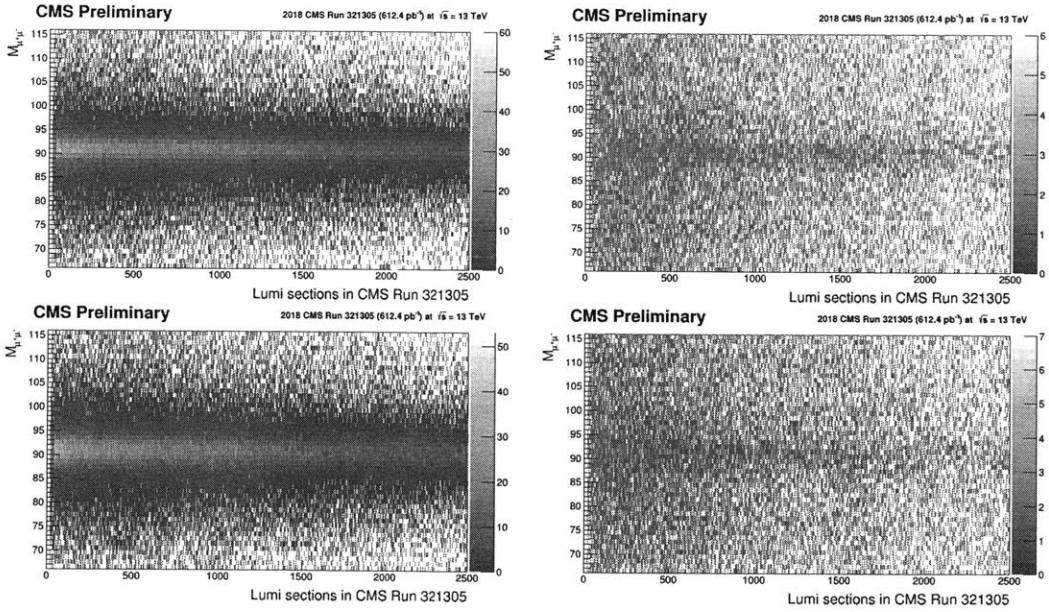


Figure 6-10: Example of $\epsilon_{standalone}$ histograms, for barrel passing (top-left), barrel failing (top-right), endcap passing (bottom-left) and endcap failing (bottom-right) probes, in CMS Run 321305, LHC Fill 7056.

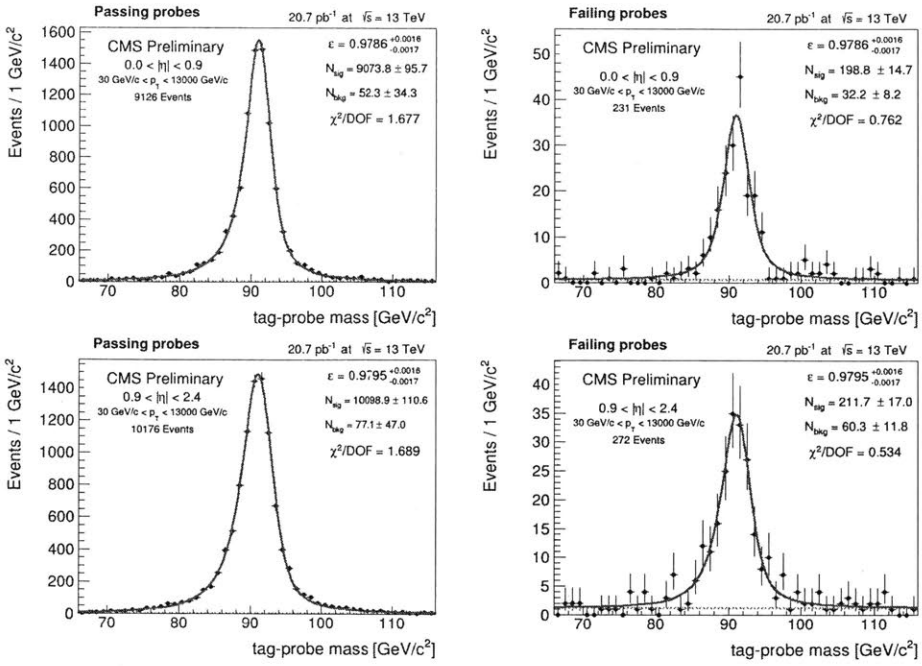


Figure 6-11: $\epsilon_{tracking+selection}$ tag-and-probe invariant mass distributions and fits for barrel passing (top-left), barrel failing (top-right), endcap passing (bottom-left) and endcap failing (bottom-right) probes, in the first 50 lumisections of CMS Run 321305, LHC Fill 7056.

not require creating templates and therefore save CPU time in automation. On the other hand, it is not the case for $\epsilon_{standalone}$ mass distributions, where $BW*CB$ behaves much worse than $MC*Gaus$. Thus, for $\epsilon_{standalone}$, the $MC*Gaus$ is still exploited and MC templates are prepared offline before data-taking periods for different pile-up conditions.

Together with the pre-fit procedure for backgrounds in $\epsilon_{standalone}$ failing category and χ^2 monitor mentioned previously, the automatic efficiency fits are good and robust over time and continuously provide reliable single muon efficiency values. For example, the fit performance for the first measurements (first 50 lumisections) in CMS Run 321305, LHC Fill 7056 can be found in Figure 6-11 and 6-12 for $\epsilon_{tracking+selection}$ and $\epsilon_{standalone}$ efficiency type, respectively.

Results

The tag-and-probe methods are performed for all measurements for each LHC Fill. For instance, LHC Fill 7056 are condensed into 53 measurements and corresponding

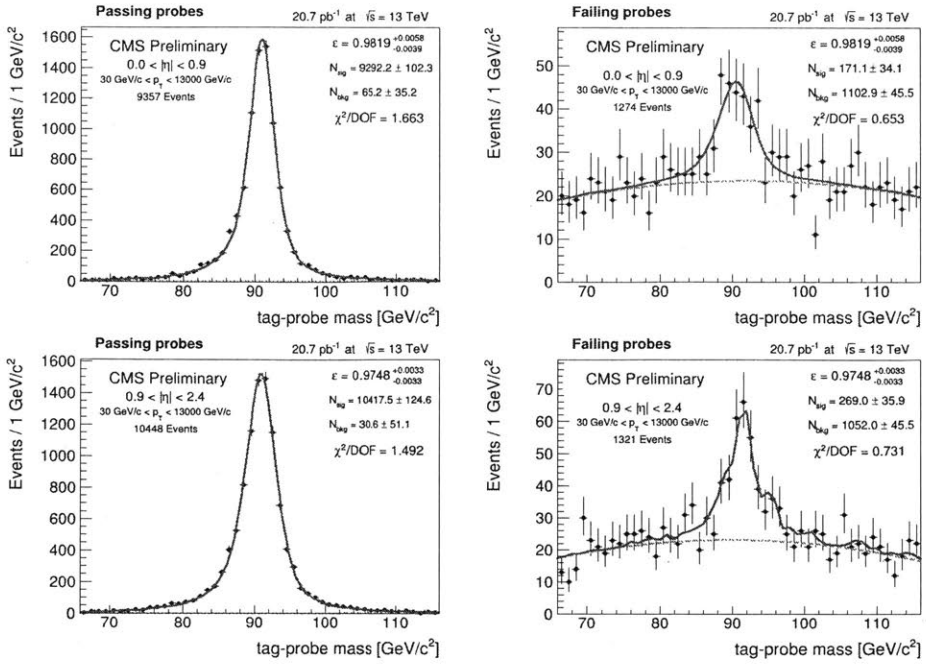


Figure 6-12: $\epsilon_{standalone}$ tag-and-probe invariant mass distributions and fits for barrel passing (top-left), barrel failing (top-right), endcap passing (bottom-left) and endcap failing (bottom-right) probes, in the first 50 lumisections of CMS Run 321305, LHC Fill 7056.

single muon efficiency measurement results as a function of time are illustrated in Figure 6-13.

As can be seen, $\epsilon_{tracking+selection}$ and $\epsilon_{standalone}$ are ~98% and stable over the entire fill, for both barrel and endcap region. However, slopes are observed for $\epsilon_{trigger}$, which is more pronounced in forward region, ranging from 85% at the beginning to 90% at the end of the fill. This is expected since there is a loose isolation selection in trigger requirement and the isolation variables are affected by pile-up, which is large at the beginning of the fill (Figure 6-6).

One straightforward solution is to use a non-isolated trigger. However, the lowest p_T threshold for a non pre-scaled non-isolated trigger for 2017 and 2018 data-taking period is 50 GeV, which is too high to accumulate event yields in the desired time interval for luminosity monitoring purpose. Introducing a non-isolated trigger might add more backgrounds too.

Because we require only one of the two muons to fire the trigger in selection, as described in Section 6.4.1, the event trigger efficiency is $1-(1-\epsilon_{\mu_1,trigger})(1-\epsilon_{\mu_2,trigger})$.

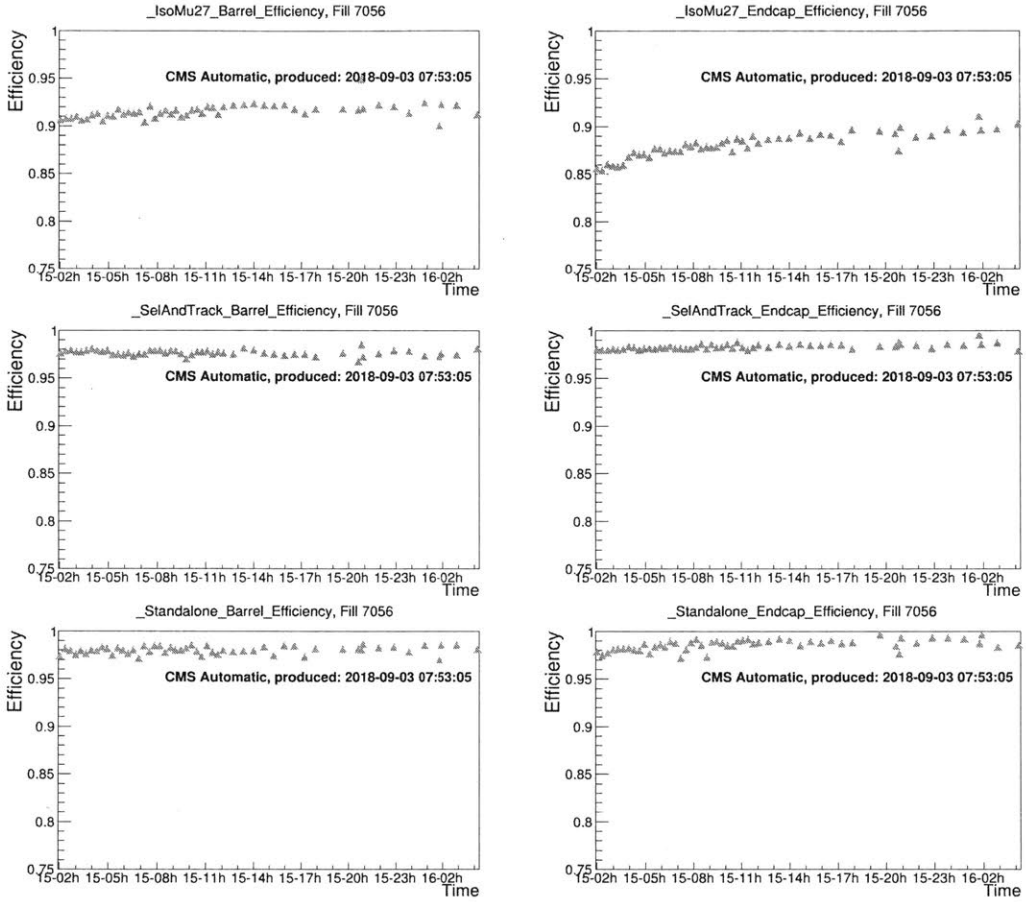


Figure 6-13: Barrel (top-left) and endcap (top-right) $\epsilon_{trigger}$, barrel (mid-left) and endcap (mid-right) $\epsilon_{tracking+selection}$, barrel (botoom-left) and endcap (bottom-right) $\epsilon_{standalone}$ results as a function of time in Fill 7056.

To estimate, in the worst case that both muons are from endcap region, we have:

$$\frac{1-(1-0.9)(1-0.9)}{1-(1-0.85)(1-0.85)} = 1.013, \tag{6.15}$$

which means the impact from the 5% difference in $\epsilon_{trigger}$ to the final result is actually $\sim 1\%$.

6.6.2 Z event efficiency

As mentioned earlier, the Z event efficiency is still determined from simulation with muon efficiency data-to-simulation scale factors applied. In the Z counting machinery, this is done in two steps. First, simple combinatorics of single muon trigger, reconstruction and selection efficiencies purely from data is calculated on-the-fly. This is correct to the first order. Further, a second-order correction due to correlation between two muons, which is prepared offline before data-taking using simulation, is added.

Z event efficiency purely from data

To construct Z event efficiency from on data, the following formula is considered, assuming there is no correlation between the two muons:

$$\begin{aligned}\epsilon_{Z\to\mu\mu}^{no-correlation} = &[1-(1-\epsilon_{\mu_1,trigger})(1-\epsilon_{\mu_2,trigger})] \times \epsilon_{\mu_1,tracking+selection} \times \epsilon_{\mu_2,tracking+selection} \\ &\times \epsilon_{\mu_1,standalone} \times \epsilon_{\mu_2,standalone}.\end{aligned} \tag{6.16}$$

Results of pure-data Z event efficiencies as a function of time are shown in blue curves in Figure 6-14 for Fill 7056.

Z event efficiency correction from simulation

The simulation samples used to derive correction is described in 6.5. Given the Z event efficiency definition in Equation 5.5, the correction due to correlation between

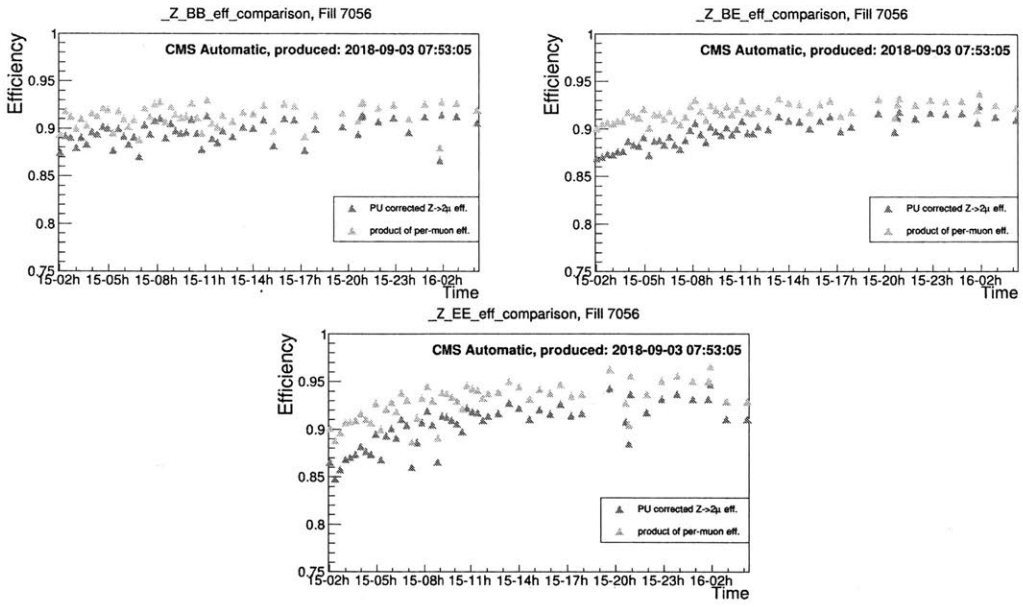


Figure 6-14: Z event efficiency results from pure data calculation and after correction for Fill 7056.

two muons are determined by:

$$\begin{aligned}\Delta\epsilon =& \epsilon_{Z\to\mu\mu}^{with-correlation} - \epsilon_{Z\to\mu\mu}^{no-correlation} \\ =& \frac{N_{sel}^{Z\to\mu^+\mu^-}}{N_{fid}^{Z\to\mu^+\mu^-}} - [1-(1-\epsilon_{\mu_1,trigger})(1-\epsilon_{\mu_2,trigger})] \\ & \times \epsilon_{\mu_1,tracking+selection} \times \epsilon_{\mu_2,tracking+selection} \times \epsilon_{\mu_1,standalone} \times \epsilon_{\mu_2,standalone}.\end{aligned} \tag{6.17}$$

All quantities in Equation 6.17 are derived from simulation. Fiducial regions and selections are defined in Section 6.4.3 and 6.5, and same tag-and-probe techniques are applied for efficiency calculations. The corrections are determined for BB, BE, EE separately and parameterized as functions of average pile-up, to account for the pile-up dependence of this correlation. This step is performed by generating pseudo-data from Poisson distribution with potential average pile-up coverage and reweighting simulated events accordingly. Examples of selected possible pile-up scenario and reweighting processes are shown in Figure 6-15.

Parameterizations are illustrated in Figure 6-16, where the blue curves "combinatorics" denote $\epsilon_{Z\to\mu\mu}^{no-correlation}$ and red curves denote $\epsilon_{Z\to\mu\mu}^{with-correlation}$ on the left hand side and $|\Delta\epsilon|$ curves are shown on right hand side. $|\Delta\epsilon|$ is then fitted with piecewise linear

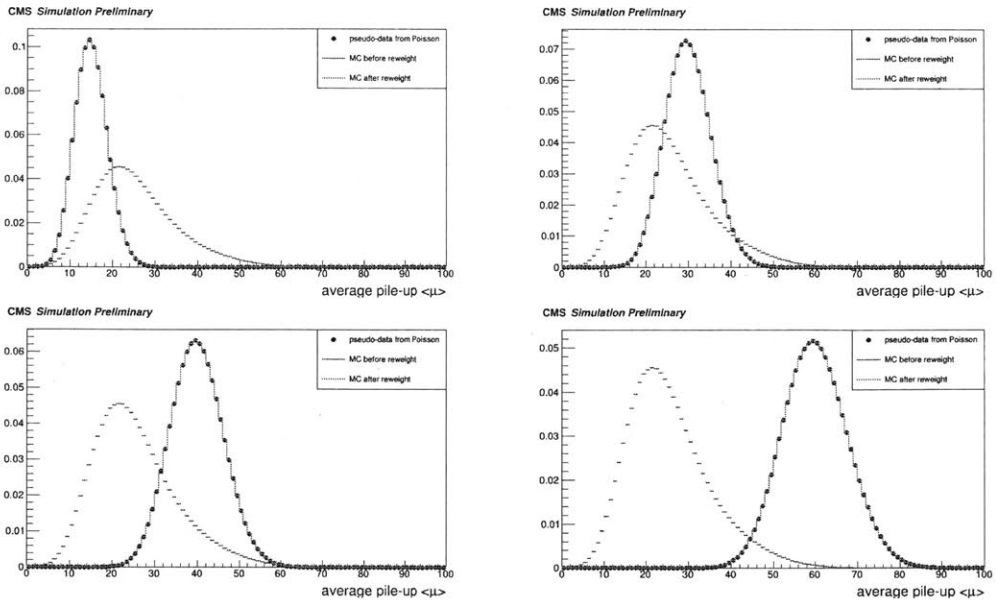


Figure 6-15: Pile-up re-weighting of the simulation sample, according to the pseudo-data distribution with an average pile-up of 15 (top-left), 30 (top-right), 40 (bottom-left) and 60 (bottom-right), respectively.

Parameters	BB	BB	BE	BE	EE	EE
pile-up range	(15,50)	(50,60)	(15,55)	(55,60)	(15,40)	(40, 60)
p_0	0.0123 ± 3e-4	-0.0878 ± 2e-4	0.0088 ± 6e-4	-0.0601 ± 4e-4	0.0161 ± 7e-4	0.0013 ± 2e-4
p_1	0.00016 ± 1e-5	0.002187 ± 3e-06	0.00049 ± 2e-05	0.001790 ± 7-06	0.00030 ± 3e-05	0.00074 ± 5e-06

Table 6.5: Z event efficiency correction parameterization as functions of pile-up using simulation.

functions to describe the different feature for medium and high pile-up conditions, for BB, BE, and EE, respectively:

$$|\Delta\epsilon| = p_0 + p_1 \times < \mu > . \tag{6.18}$$

Average pile-up ($< \mu >$) range (15,60) is exploited for fitting, which was expected to cover most realistic pile-up conditions in 2017 and 2018 data-taking period. Beyond this range, extrapolation is applied. Fitting parameter results are recorded in Table 6.5.

After the correlation correction is applied, the Z event efficiency is what finally enters the Z rate and other Z counting observable calculation. For Fill 7056, the red curves in Figure 6-14 denote the Z event efficiency results after correction. It is a

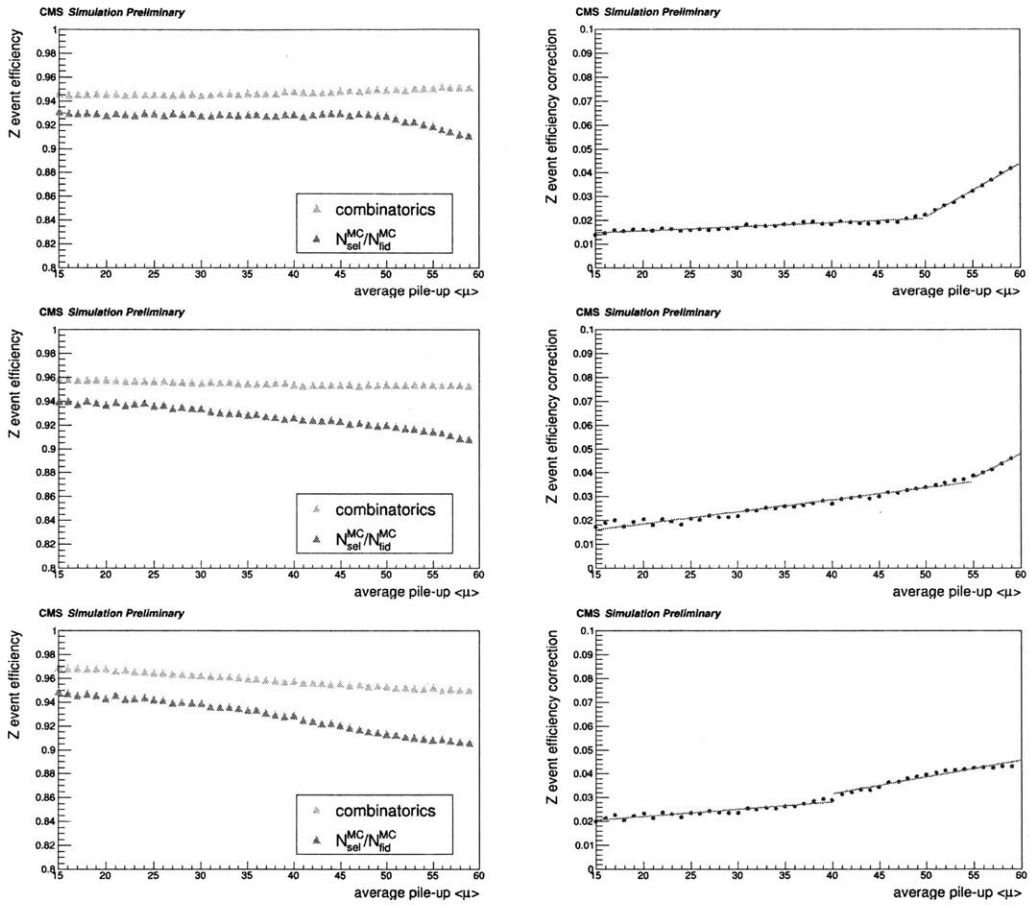


Figure 6-16: Z event efficiency correction parameterization as a function of average pile-up using simulation. Combinatorics denotes Equation 6.16. The pile-up distributions in simulated samples are re-weighted according to the pseudo-data distribution with hypothesized pile-up scenarios.

Process	Yields
$Z \to \mu\mu$	325777 ± 291
$t\bar{t}$	1523 ± 8
WW, WZ, ZZ	816 ± 35

Table 6.6: Signal and background yields passing event selection from simulation.

1-3% correction depending on pile-up, as observed.

6.6.3 Monitor tool for muon efficiencies

As shown in Section 6.6.1, $\epsilon_{trigger}$, $\epsilon_{tracking+selection}$ and $\epsilon_{standalone}$ are intermediate products in the Z counting analyses. They are constantly monitored for every Z counting time interval ($\sim$ 20 minutes). In-fill stability plots have been seen previously in Figure 6-13. Multi-fill stability plots are illustrated in Figure 6-17, where average efficiencies are plotted for single LHC fill for the entire 2018 data-taking period. As can be seen, all efficiencies are stable overall, except that the trigger efficiencies in forward region are relatively lower in earlier 2018, due to calorimeter upgrades. Those types of plots can be exploited for muon efficiency monitoring and complementary to other muon efficiency studies using a larger dataset.

6.7 Signal extraction

As described in Section 5.7, since the background contributions are small in signal region, the $Z \to \mu\mu$ yields, $N_{sel}^{Z\to\mu^+\mu^-}(t)$, are extracted by simply counting events passing selection $N_{obs}^{Z\to\mu^+\mu^-}(t)$, and then applying a correction to subtract background contribution, which is determined beforehand from simulation.

Simulated samples used for background contribution calculation are the same ones in Section 5.2.2. The signal and background yields passing event selection are recorded in Table 6.6, from which the background contribution can be determined to be 1%. Additionally, this correction is taken as a constant in all Z counting measurements, since dedicated studies show no obvious dependence on pile-up conditions.

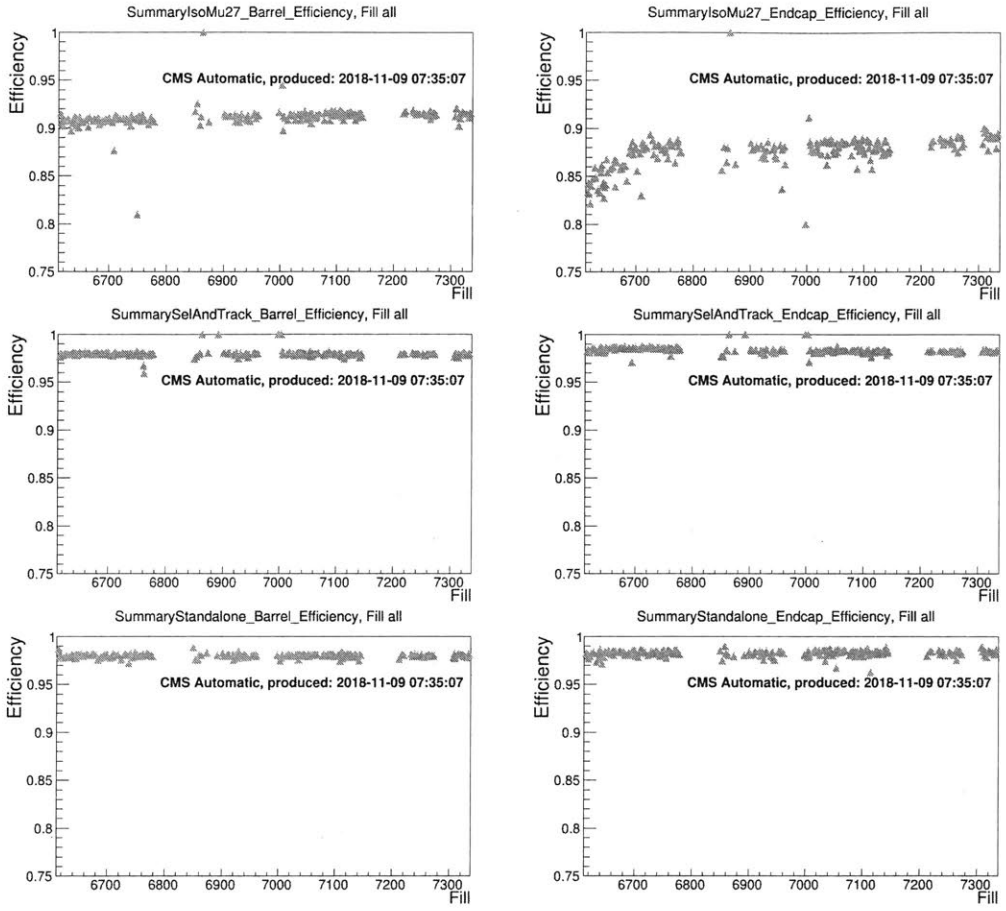


Figure 6-17: $\epsilon_{trigger}$, $\epsilon_{tracking+selection}$ and $\epsilon_{standalone}$ results for all LHC fills for the entire 2018 data-taking period.

6.8 Results

In this section, results are presented for measurements of the various Z counting observables described in Section 6.3.1, including Z rates, Z counts, Z fiducial cross sections as well as measured luminosity by Z counting and its ratio to measurements by luminometers.

6.8.1 Z rates, Z counts and synchronization with ATLAS

Once we have $N_{obs}^{Z\to\mu^+\mu^-}(t)$, $\epsilon(t)$ and time information, Z counts and Z rates can be determined by Equation 6.12 and 6.13.

For the joint CMS-ATLAS synchronization exercise, a mutually agreed format of csv files is written for each Z counting measurements, with following contents: LHC fill number, fill start time, end time, Z rates, average instantaneous luminosity, integrated luminosity and Z yields which are delivered by LHC. An example of such a file for the 53 measurements performed with LHC Fill 7056 is illustrated in Figure 6-18.

In automation workflow, an inclusive csv file is updated daily and results from new measurements are added by Cron. Z counts, Z rates and other Z counting observable plots are also posted automatically for luminosity monitoring purpose.

In-fill results

To present in-fill results, z rate observable and two specific LHC fills are deployed: Fill 7056, the long one (26.6 hours) used for illustration throughout this chapter, and Fill 6620, which is much shorter (5.2 hours) but has some unusual features we could exploit to cross check the performance of Z counting and online luminometers.

Z rate measurements and instantaneous luminosity measured by luminometers from both experiments are shown in Figure 6-19. For each measurements, the total uncertainty is conservatively estimated to be 3%, which includes ~1% statistical uncertainty on yields, ~ 2% systematic uncertainty on efficiencies and per-mill level uncertainty on time and deadtime determination. In Fill 7056, CMS and ATLAS

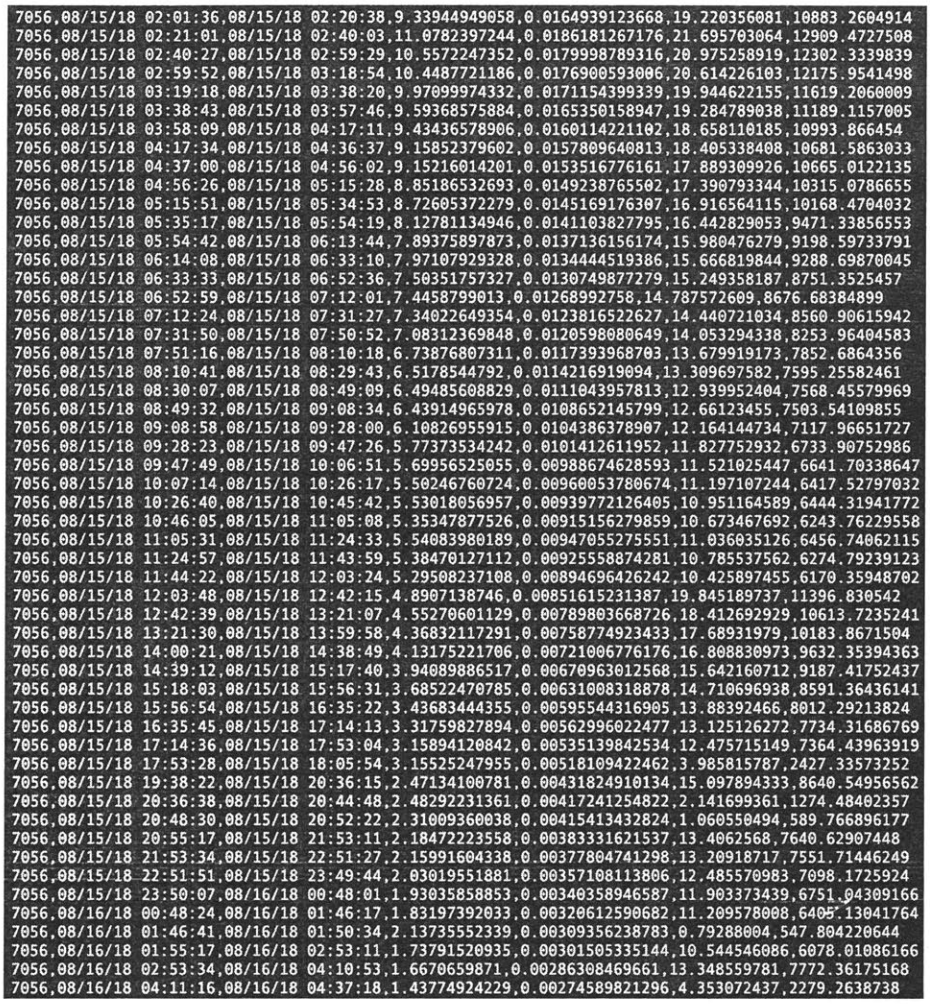


Figure 6-18: An example of output csv file for Z counting synchronization. Each line exhibits the LHC fill number, fill start time, end time, Z rates, average instantaneous luminosity, integrated luminosity and Z yields for each measurements.

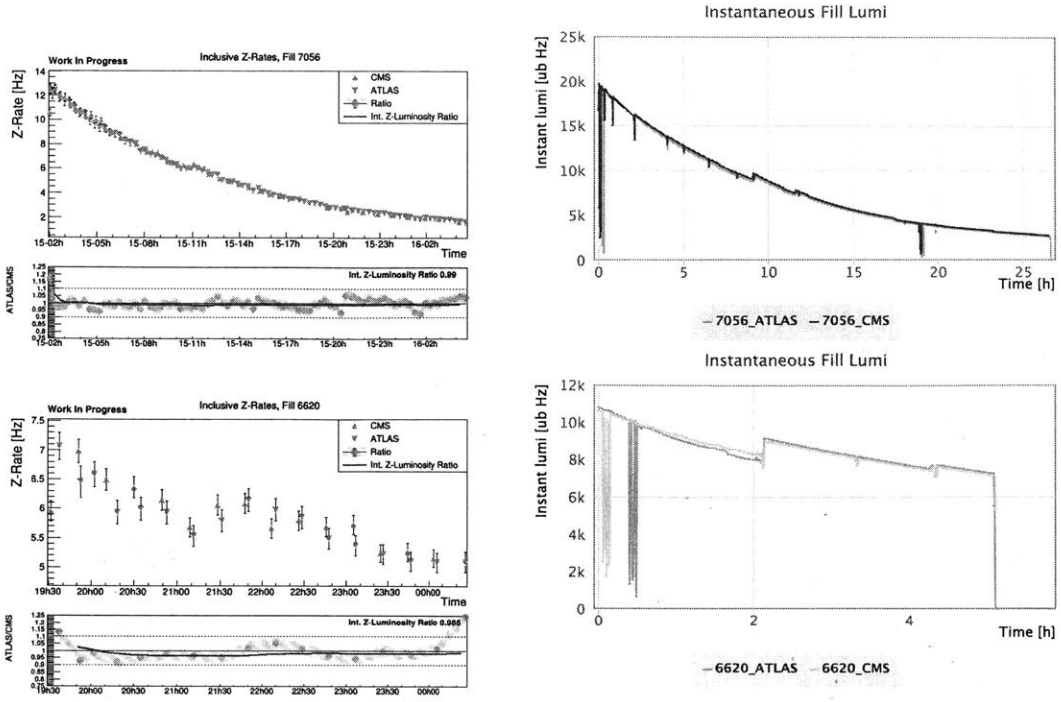


Figure 6-19: Z rate measured by CMS and ATLAS Z counting workflow (left) and instantaneous luminosity measured by CMS and ATLAS luminometers (right) for LHC Fill 7056 (top) and 6620 (bottom).

instantaneous luminosities are almost indistinguishable, which is also verified in Z rate curves. The kick at 9 hours after fill starts is also manifested. In Fill 6620, three features are suggested from online luminosity measurements: (1) LHC delivered more luminosity for CMS than ATLAS from 1 to 2 hours after fill stars, (2) luminosity drops at very beginning of the fill for CMS and 0.5 hour after fill starts for ATLAS, (3) luminosity jumps for both experiments $\sim$ 2 hours after fill starts. All those features are well captured by Z rate measurements. As can be seen, Z counting provides reliable and complementary information for in-fill luminosities.

Multi-fill results

Two types of monitoring plots are illustrated here for multi-fill results. First, ATLAS-to-CMS average Z rate ratios per fill are shown in Figure 6-20. Further, Figure 6-21 shows ATLAS-to-CMS Z count ratios per fill as well as comparison with online luminosity ratios. As marked in Figure 6-21, no data is taken during machine development (MD) and technical stop (TS), which is accounted for the blank region in the Z count-

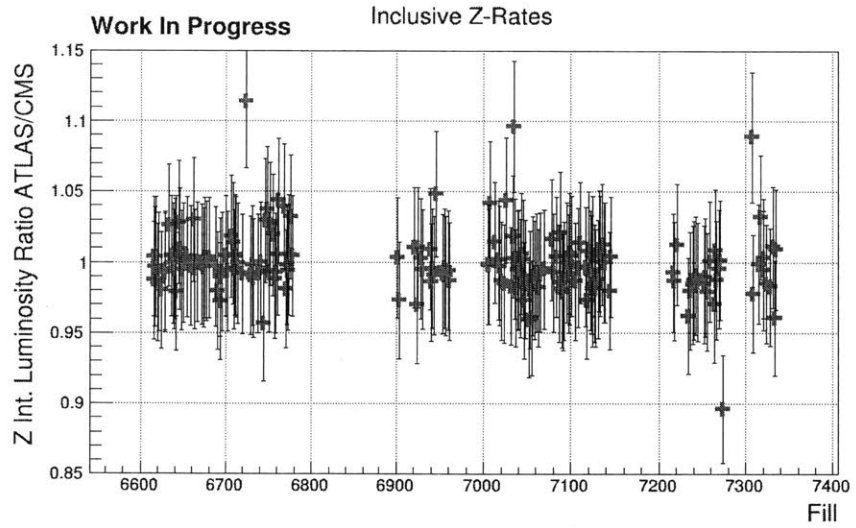


Figure 6-20: ATLAS-to-CMS average Z rate ratios per fill for the 2018 data-taking period.

ing measurements.

In general, both Z rate and Z count ratios are pretty stable over entire 2018 data-taking periods, and consistent with unity, which indicates good and even LHC performance at IP1 and IP5. On the other hand, online luminosity ratios suggest a 3-5% difference between CMS and ATLAS, which might imply potential mis-calibration of online luminometers. Moreover, in Figure 6-21, three outliers beyond 1.1 are observed for Z count ratio, which is also confirmed by luminosity ratios. This is then found to be affected by gruffalo strikes(beam losses). Given the above, the Z counting is an precise and complementary tool to check online luminometer calibrations for both detectors and identify potential problems for LHC machine.

6.8.2 Z fiducial cross sections

If using online luminosity measurements, Z cross sections can be determined and exhibit another representation of the Z counting results. Independent determination from theoretical prediction could be used as reference to cross check the results.

In-fill (for LHC Fill 7056) and multi-fill results for Z fiducial cross sections and stability are illustrated in Figure 6-22. Stability from ATLAS are also presented for multi-fill plots. Direct fiducial cross section comparisons between two experiments

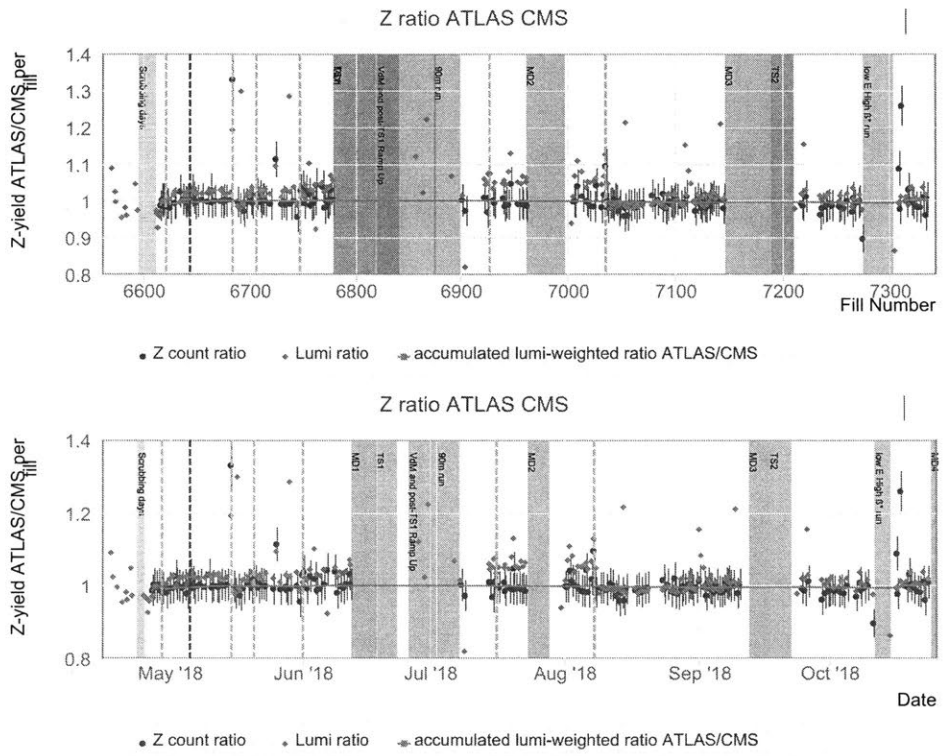


Figure 6-21: ATLAS-to-CMS Z count ratios versus luminosity ratios per fill for the 2018 data-taking period, with X axis labeled as fill number (top) and date (bottom).

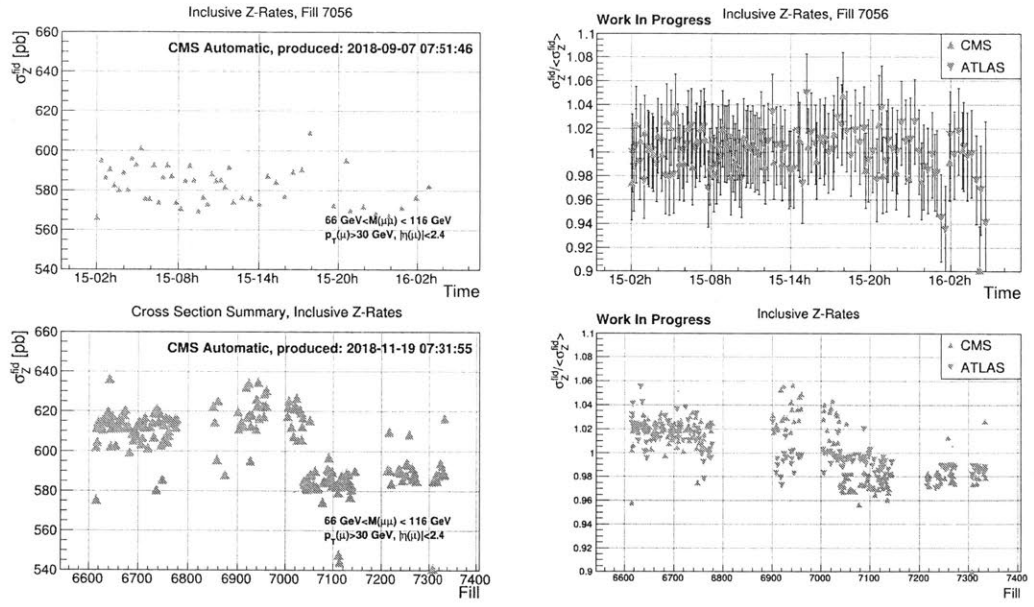


Figure 6-22: In-fill (top) and multi-fill (bottom) Z fiducial cross sections (left) and stability (right).

are avoided here to follow the analysis strategy.

Theoretical prediction calculated by FEWZ [61] using NNPDF 3.1 [116] as PDF input suggests $\sigma_{tot}^{Z\to\ell\ell} = 1923 \pm 15pb$ for the total Z cross section, which is corresponding to a fiducial cross section $\sigma_{fid}^{Z\to\ell\ell} = 578 \pm 5pb$ (acceptance taken from simulation). If taking as reference, it is observed that prior to LHC Fill 7040, the measured Z fiducial cross sections are 5-6% higher than the predicted value, and stay consistent afterwards. This drop is corresponding to the drop in ATLAS-to-CMS luminosity ratios in Figure 6-21, which could indicate that CMS online luminometers were potentially mis-calibrated before Fill 7040 and thus resulted in lower luminosity values, which is fixed afterwards. Otherwise the overall Z cross section is stable for both in-fill and multi-fill, no pile-up dependent slope is found for in-fill measurements.

6.8.3 Measured luminosity by Z counting

As mentioned in Section 6.3.1, if taking theoretical prediction as reference and utilize Equation 6.10, the measured luminosity by Z counting allows a straightforward comparison with measurements by other luminometers, see Figure 6-23. Comparison against PCC, DT, PLT and the best values used for physics analyses are shown

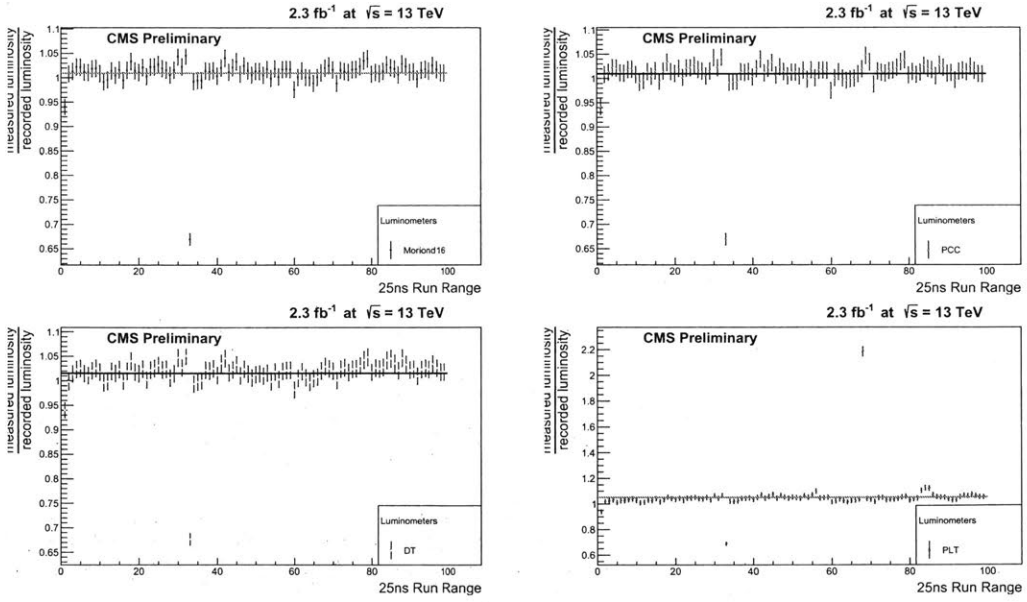


Figure 6-23: Luminosity ratio between Z counting measurements and offline (top-left), PCC (top-right), DT (bottom-left) and PLT (bottom-right)

for 2015 data-taking period. In general, good agreements are observed in, difference between different luminometers are within $\sim 1\%$.

6.9 Summary and outlook

The Z counting method as a luminosity monitor in muon channel has been well established and fully automatized during Run II. Z rates, Z counts, Z fiducial cross sections and other Z counting observables are measured at high precision. This is of great importance for online and offline luminometer calibration for CMS detector as well as LHC diagnosis. In addition, the muon trigger, selection, and reconstruction efficiencies, as by-product of the Z counting study, are also useful for efficiency and muon-related object monitoring.

Several relevant prospects are being investigated. First, towards the precision end, as indicated in Section 6.3.1 and 6.8.3, 1-2% theoretical uncertainty has to be included on the absolute luminosity determination by Z counting, if using theoretical predictions as input. To avoid the unnecessary uncertainties, a new approach [141] is proposed by taking advantage of the special low pile-up runs. The idea is to measure

luminosity with CMS luminometers at low pile-up conditions, $\mathcal{L}_{low}^{CMS}$, and scale by efficiency-corrected Z yield ratio measured in the special low pile-up, $N_{low}^{\prime Z}$, and the realistic pile-up condition when taking data, $N_{high}^{\prime Z}(t)$. Then, the luminosity can be determined by:

$$\mathcal{L}_{high}^{Z-counting}(t) = \frac{N_{high}^{\prime Z}(t)}{N_{low}^{\prime Z}} \times \mathcal{L}_{low}^{CMS}. \tag{6.19}$$

The luminosity measurements are very likely to be improved, as the uncertainty on $\mathcal{L}_{low}^{CMS}$ is mainly the normalization one (see Section 6.1.3) and small ~1-1.5%, and correlated systematics on $N^{\prime Z}$ can be cancelled. There is also possibility by exploiting the high precision luminosity measurements at LHCb experiments [142], $\mathcal{L}_{LHCb}$, and also scale by efficiency-corrected Z yield ratio measured at CMS, $N_{CMS}^{\prime Z}(t)$, and LHCb, $N_{LHCb}^{\prime Z}(t)$, i.e:

$$\mathcal{L}_{CMS}^{Z-counting}(t) = \frac{N_{CMS}^{\prime Z}(t)}{N_{LHCb}^{\prime Z}(t)} \times \mathcal{L}_{LHCb}(t). \tag{6.20}$$

However, dedicated investigation on correlation of Z yields is required to establish both methods.

Second, towards the promptness end, the Z counting at scouting stage is also proposed and currently being investigated. Lastly, towards the information end, the Z counting in electron channel is being finalized for automation, which can provide independent self cross-check of Z rates as well as electron efficiency monitoring. W and $t\bar{t}$ counting points to a possibility for tracking missing energy performance. J/ψ counting could potentially allow comparison among CMS, ATLAS and LHCb experiments.

To end the chapter, accumulated number of Z bosons by Z counting over time for 2018 data-taking period is exhibited in Figure 6-24. Integrated luminosity measured by online luminometers are also shown to allow comparison.

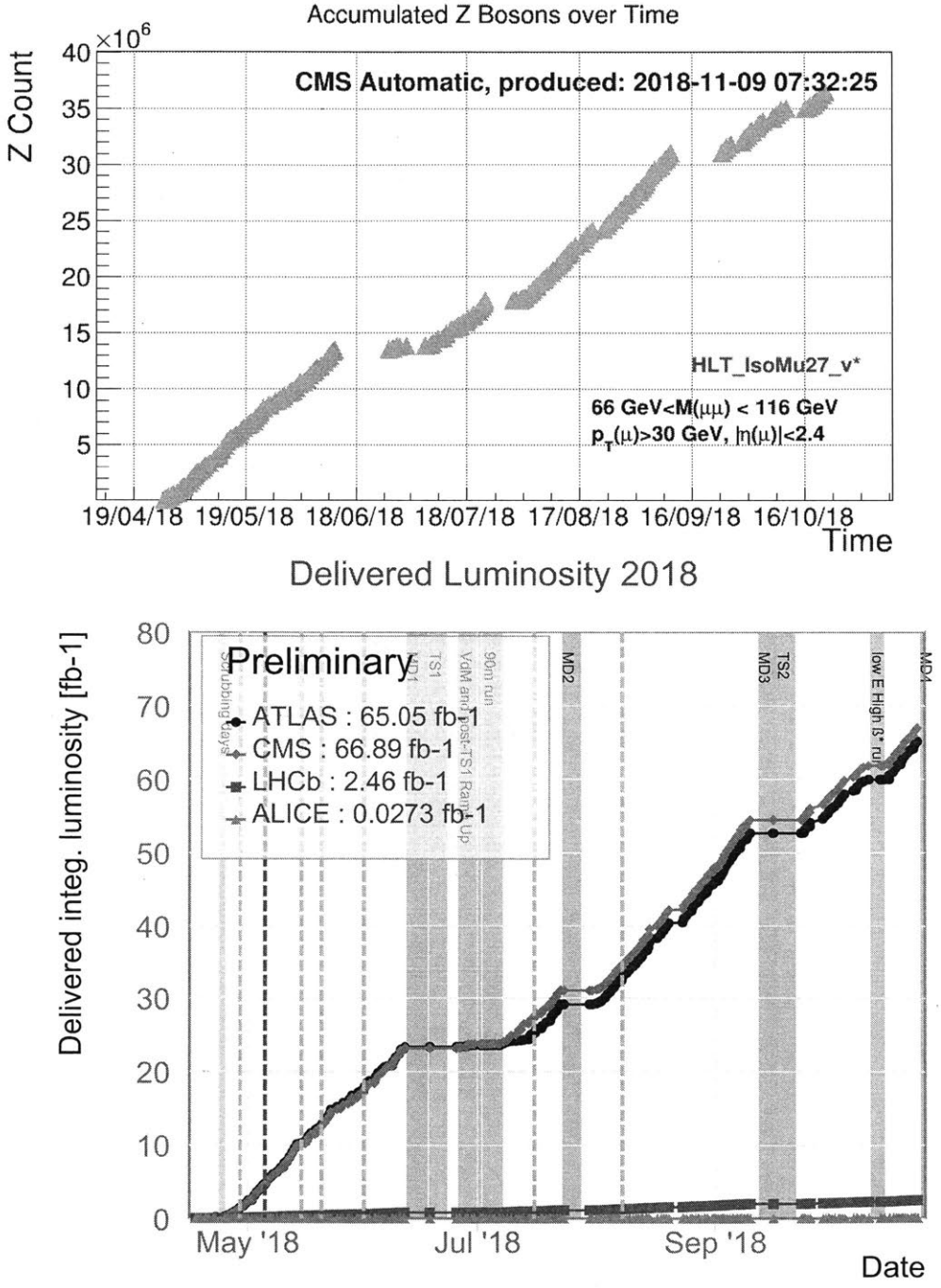


Figure 6-24: Accumulated number of Z bosons (top) and integrated luminosity (bottom) as a function of time/fill number in 2018 data-taking period

Chapter 7

Search for the neutral Higgs bosons in the di-muon channel

The light scalar boson with a mass of 125.09 $\pm$ 0.24 GeV [31], discovered [29, 30] at LHC in 2012, has properties [86–88] that are well consistent with the predicted SM Higgs boson (denoted as H^{SM} in this chapter). 2018 is a "Yukawa Year". After the observation of the $t\bar{t}H^{SM}$ production [143, 144], $H^{SM} \to \tau\tau$ [145, 146], and $H^{SM} \to b\bar{b}$ [147, 148], the Yukawa coupling mechanism of SM Higgs to the third generation fermions are fully confirmed. But the journey is not over, and the Higgs coupling to second generation fermions are subject to investigation. The next accessible channel would be $H^{SM} \to \mu\mu$. However, the potential to observe it is very unlikely using existing data [37], since the branching ratio is tiny and the other SM processes, such as the Drell-Yan and leptonic $t\bar{t}$ decays, contribute a large backgrounds for the search.

Nevertheless, there are beyond SM models that predict an enhancement of decays of the neutral Higgs bosons (generally denoted as ϕ) into the di-muon channel, which motivate this search. For instance, as discussed in Section 2.2.3, in the mostly studied SUSY model, MSSM, the decays of heavy neutral Higgs bosons, H and A, into a pair of muons are significantly enhanced for large $tan\beta$ values, which provides the potential for a discovery or exclusion of the parameter spaces for specific scenarios.

In the context of the MSSM, the two dominant production channels at LHC for the neutral Higgs bosons are gluon-gluon fusion ($gg\phi$) and b-quark associated production

($bb\phi$). The $bb\phi$ is dominant for high $tan\beta$ values, while the $gg\phi$ is more relevant at low $tan\beta$ region (see Figure 2-10). Although the branching ratios decaying to $\mu\mu$ are almost 3–4 orders of magnitude smaller than that of $\tau\tau$ and $b\bar{b}$ (see Figure 2-11), the di-muon final state has better experimental accessibility as well as provides complementary probes, as it can be fully reconstructed and the CMS detector has an excellent resolution on muon momentum resolution.

7.1 Analysis overview

The expected experimental signature of signal events contains two oppositely charged, high-p_T isolated muons and small p_T^{miss}. The b jets are also present in $bb\phi$ production channel while this activity is supposed to be rather limited for $gg\phi$ production. The SM processes of Drell-Yan, $t\bar{t}$, and diboson productions also produce signal-like signature and thus contribute as backgrounds.

To enhance the sensitivity, events are categorized and selections are optimized individually to better target on both production mechanisms. This procedure is performed with simulation, and data-to-simulation scale factors are applied to account for the difference in muon triggering, reconstruction, selection efficiencies and b jet identification efficiencies. Those scale factors are determined in a similar way as described in Section 5.5.

The presence of signals would be characterized by an excess of events over the SM backgrounds in the di-muon invariant mass ($M_{\mu\mu}$) distributions. To quantify the number of signal, simultaneous maximum likelihood fits are performed to the $M_{\mu\mu}$ distributions in all event categories for statistical analysis. The results are interpreted based on the likelihood ratios for the background-only and signal-plus-background hypotheses, both in a model independent way and under the framework of MSSM. For each interpretation, the background is entirely modeled based on data. In the model independent interpretation, a single resonance, either narrow or wide [1], is assumed to

[1]Whether a resonance is narrow or wide depends on how the intrinsic width compares with the experimental resolution.

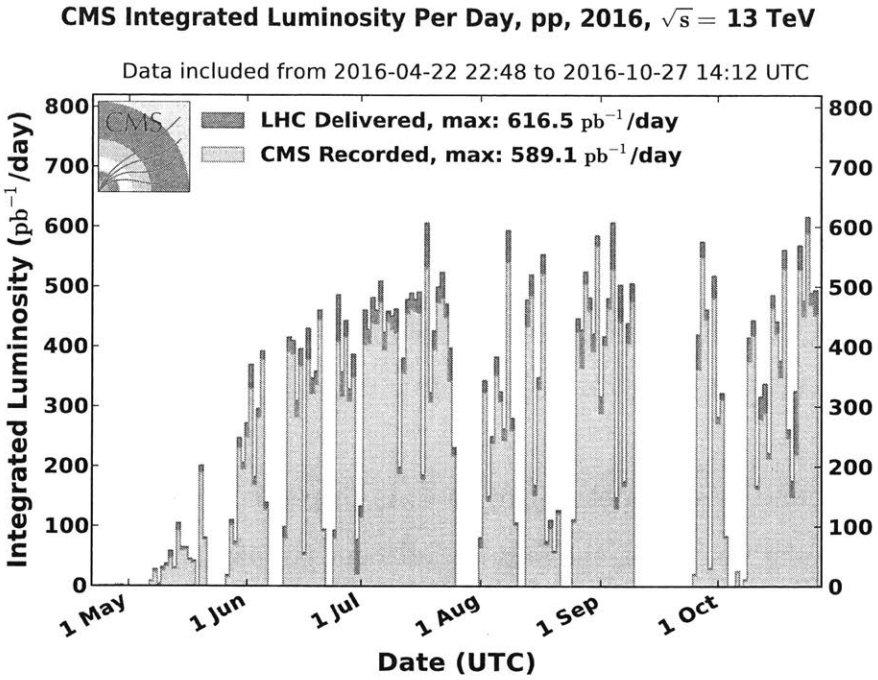


Figure 7-1: Day-by-day integrated luminosity of 2016 data-taking period. [10]

be produced entirely via either $gg\phi$ or $bb\phi$ production. For this purpose, simulated samples of $bb\phi$ and $gg\phi$ with mass between 130 GeV and 1 TeV are used to model the signal. In the MSSM interpretation, the simulated samples are combined into the multi-resonance signal structure according the tested MSSM benchmark scenarios.

The rest of this chapter is organized as follows. First, the data and simulated samples used in this search are described in Section 7.2. Then, the event selection and categorization are discussed in Section 7.3. Section 7.4 presents the signal acceptance and efficiencies and Section 7.5 elaborates the signal and background modeling. Finally, after systematic uncertainties are discussed in Section 7.6, the search results are presented in Section 7.7.

7.2 Data and simulated samples

This search is performed using the data of proton-proton collisions at center-of-mass energy of $\sqrt{s}$=13 TeV, collected in 2016 data-taking period (see Figure 7-1), which corresponds to an integrated luminosity of 35.9 fb^{-1}.

Simulated signal samples are generated with PYTHIA [71], to model the characteristics of the Higgs boson productions and decays into muon pairs. Those samples are produced for a large number of m_A and $tan\beta$ combinations, where m_A goes from 130 to 1000 GeV and $tan\beta$ from 5 to 60. For the model specific predictions of the Higgs boson masses, cross sections and branching ratios that relevant for this search, the LHC cross section working group recommendations [9] are followed. Especially, FEYNHIGGS codes [149] are exploited to correct the Higgs boson mass approximation in PYTHIA.

The backgrounds are modeled using a data-driven approach and do not rely on background simulation. The simulated background samples are only used to optimize the selection criteria. The two dominant backgrounds are the Drell-Yan process and $t\bar{t}$ production. Other less dominant contributions come from single top and diboson production processes such as $W^{\pm}W^{\mp}$, $W^{\pm}Z$ and ZZ. Those simulated samples are generated with aMC@NLO [77, 78] and POWHEG [79–82]. Descriptions of those generators can be found in Section 2.1.2. Additional proton-proton interactions per bunch crossing (pile-up effects) are also considered in the simulated events, which are reweighted to match that observed in data.

7.3 Event selection and categorization

Dedicated selections and categorizations are designed in order to improve signal sensitivity for both production channels in this search. As the di-muon decay channel is exploited, the pairs of energetic and isolated muons form the baseline selection. Then, to target on the two Higgs production mechanisms with different b jet activities and different background contributions, events passing the baseline selection are classified into two categories for individual optimization. This procedure is performed with simulation, as mentioned previously.

ID	Isolation	Triggers	Offline p_T, η cut
Tight	PF-based, loose	IsoMu24, IsoTkMu24	26 GeV, $\|\eta\| < 2.4$
High-p_T	Tracker-based, loose	Mu50, TkMu50	53 GeV, $\|\eta\| < 2.4$

Table 7.1: Summary of online and offline muon selections for two different muon objects.

7.3.1 Muon pair selection

Since the search is performed for the Higgs bosons with a large range for mass assumptions from 130 to 1000 GeV, the muon pairs decaying from the Higgs bosons are expected to distribute in a wide p_T range. To ensure high muon detection efficiency over the wide p_T range, two different IDs are exploited: PF-based Tight ID and non PF-based High-p_T ID, which aims to complement efficiency for muons with $p_T >$ 200 GeV, as described in Section 4.3.3. Two isolation variables are used accordingly: PF-based isolation if identified by Tight ID and tracker-based isolation if identified by High-p_T ID. Loose working points are applied for both isolation variables.

The offline kinematic cuts are determined to adapt to the corresponding lowest-threshold unprescaled triggers. Tight ID muons are selected from events collected by an isolated trigger with p_T threshold of 24 GeV. Since the online isolation variable is PF-based, High-p_T ID muons utilize a non-isolated trigger with p_T threshold of 50 GeV. Therefore, in order to be fully efficient, the offline p_T cuts are chosen to be 26 and 53 GeV, respectively, and $|\eta| < 2.4$. Above online and offline muon selections have been summarized in Table 7.1.

The events are required to contain at least two oppositely charged tight or high-p_T muons. Most likely, both muons in the selected pairs pass both Tight and High-p_T ID. In those cases, the priority is given to Tight ID, in the sense of choosing the isolation variables and matching to the trigger objects. Number of signal events increases about $\sim$5% when adding the High-p_T ID. The contribution from cases that one muon passes Tight ID and the other passes High-p_T ID is $\precsim$ 0.05% and not included in selection to avoid trigger object matching complication. If more than one oppositely charged muon pair are found, the pair with the highest p_T are considered.

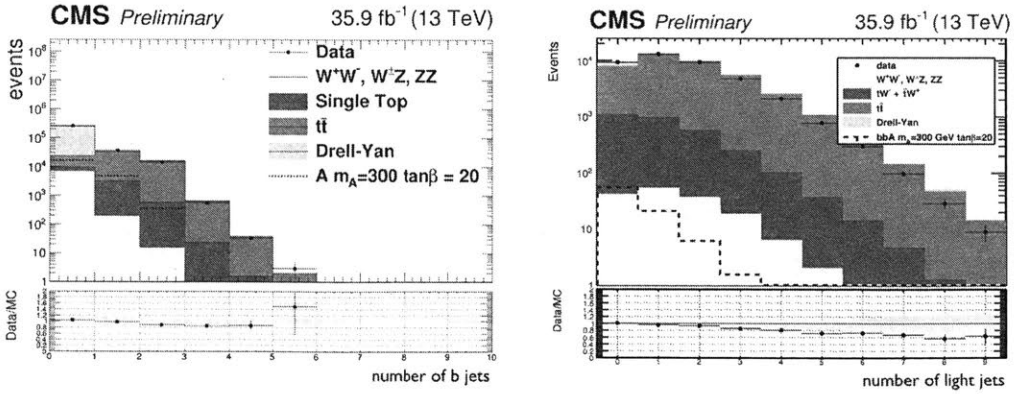


Figure 7-2: Left: b jet multiplicity in events with $M_{\mu\mu} > 130$ GeV. Right: light-flavor jet multiplicity in events with b jets.

7.3.2 Event categorization

B jet multiplicity is exploited to split the events into two categories. The distribution of number of b jets for events passing the di-muon selection with $M_{\mu\mu} > 130$ GeV are shown in Figure 7-2, for data, backgrounds, and $bb\phi$ signals. As can be seen, the presence of at least one b jet is able to heavily suppress the dominate Drell-Yan backgrounds and thus provide better sensitivity for $bb\phi$ production channel. In addition, the shape of b jet multiplicity in $bb\phi$ signals falls more sharply than in $t\bar{t}$ backgrounds. Hence, to achieve the best sensitivity, the *b-tagged* category is defined by requiring the events to have exactly one b jet. Correspondingly, the events without b jets provide better sensitivity for $gg\phi$ signals, which are defined as *no-b-tagged* category.

Light-flavor jet multiplicity is further used to optimize the signal sensitivity in b-tagged category. Distribution of number of light-flavor jets in events with b jets are illustrated in Figure 7-2. Similarly, signals tend to have a lower light-flavor jet multiplicity compared to $t\bar{t}$ backgrounds, thus the background rejection power is further improved by rejecting events with more than one light-flavor jets, without affecting significantly the signal efficiency.

Furthermore, signal events produced via both $gg\phi$ and $bb\phi$ are characterized by a rather small amount of p_T^{miss}, while the $t\bar{t}$ backgrounds contribute noticeable tails in

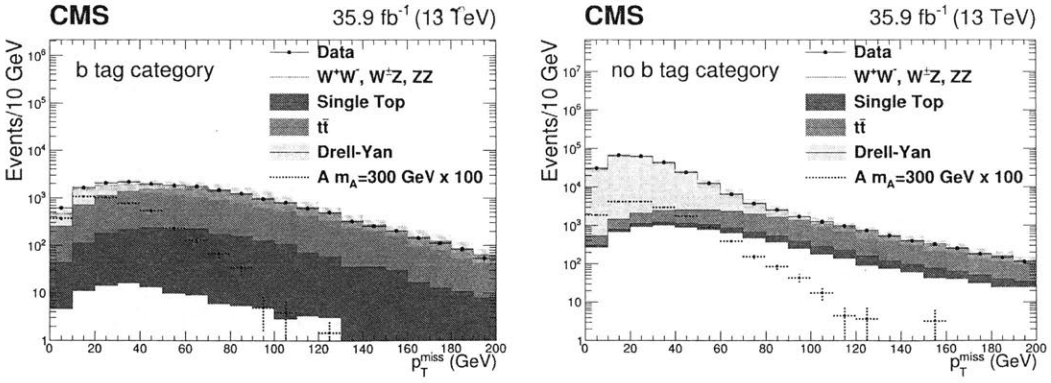


Figure 7-3: p_T^{miss} distributions in *b-tagged* and *no-b-tagged* category for events with $M_{\mu\mu} >$ 130 GeV. The shaded grey bands represent the systematic uncertainty in simulation. The expected signal is scaled by a factor of 100 for illustration.

the p_T^{miss} distributions, see Figure 7-3. For this reason, p_T^{miss} is exploited to maximize the expected significance separately for the two categories, using Punzi formula [150]:

$$\frac{\epsilon}{5/2 + \sqrt{B}}, \tag{7.1}$$

where the ϵ is the selection efficiency and B is the number of backgrounds passing corresponding p_T^{miss} cut inside the defined $M_{\mu\mu}$ window[2] around the Higgs boson mass assumptions. Figure 7-4 illustrates this optimization procedure, and the expected significances from Punzi formula are shown as a function of the cut on p_T^{miss} for various m_A and $tan\beta$ combinations. As can be observed, the significance has a peak around $p_T^{miss} = 40$ GeV for $bb\phi$ signals in the b-tagged category, while reaches a plateau of maximum around $p_T^{miss} = 80$ GeV in the no-b-tagged category. Consequently, additional selection of $p_T^{miss} < 40$ (80) GeV are applied in the b-tagged (no-b-tagged) category.

The selection criteria in the two event categories are summarized in Table 7.2.

[2] $\pm 3\Gamma_\phi$, where Γ_ϕ is the observed width of the Higgs boson resonance.

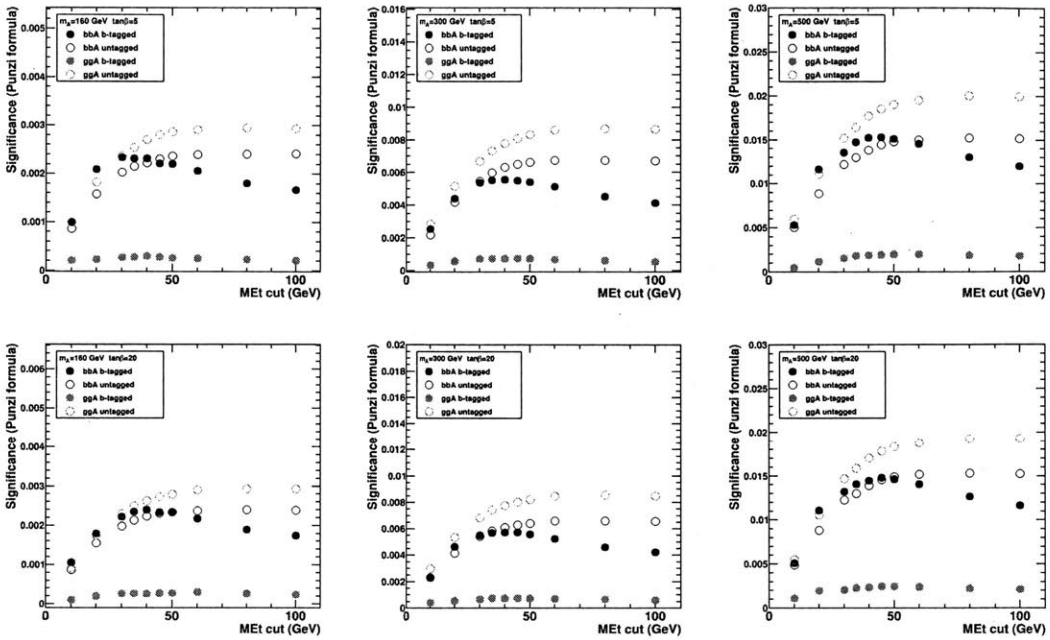


Figure 7-4: Expected significance using Punzi formula as a function of p_T^{miss} selection for various m_A and $tan\beta$ combinations.

Selection criteria	*b-tag* category	*no-b-tag* category
single muon trigger	($p_T > 24$ GeV, $\|\eta\| < 2.4$, isolation) OR ($p_T > 50$ GeV, $\|\eta\| < 2.4$)	
two oppositely charge muons	muon ID, $p_T > 26$ (53) GeV, $\|\eta\| < 2.4$, isolation	
b jets	1 with $p_T > 20$ GeV, $\|\eta\| < 2.4$	veto
light-flavor jets	0,1 with $p_T > 20$ GeV, $\|\eta\| < 2.4$	
p_T^{miss}	< 40 GeV	< 80 GeV

Table 7.2: Summary of the selection criteria in the two event categories.

7.4 Signal acceptance and efficiencies

Signal acceptance A and efficiencies ϵ are calculated using simulated signal samples to determine the signal normalization, which follows:

$$N_{sig} = \mathcal{L} \times \sigma(pp \to \phi) \times \mathcal{B}(\phi \to \mu\mu) \times A \times \epsilon, \tag{7.2}$$

where N^{sig} is the expected signal yields, $\sigma(pp \to \phi)$ the Higgs production cross section, and $\mathcal{B}(\phi \to \mu\mu)$ the branching ratio decaying to two muons. The Higgs boson events generated within the defined mass window[3] of the Higgs boson mass assumptions are considered as signals, and the number of which are normalized according to the theoretical prediction of the Higgs boson production cross sections. The acceptance is defined as the fraction of those signal events with at least two muon objects passing the p_T and η cuts. For each specific later step in the selections, the efficiency is defined with respect to the events passing previous selection step.

The corresponding results for the pseudoscalar A are shown in Figure 7-5 and 7-6, for the two production channels, respectively. Since the variation of $tan\beta$ mostly affects the Higgs boson width and leaves the event kinematics unchanged, the acceptances and efficiencies for a given Higgs boson mass are very weakly dependent on $tan\beta$. As a result, the signal acceptance and efficiencies are illustrated as a function of m_A by taking the average over different $tan\beta$ as the central value and the variations as error bands.

Based on Figure 7-5 and 7-6, the acceptance increases as a function of m_A as expected, since the decaying muons get harder. The muon identification and isolation efficiency is ~95% and the inefficiency at lower mass region is mainly due to the relative isolation cuts. The trigger requirement is fully efficient and mass-independent. b-tagged category efficiency is as low as ~20% even for bbA signals. This is because the b jets are soft and largely emitted towards the forward region, which makes them beyond the inner tracker detector acceptance and outside the applicable p_T range for b-tagging algorithms. The overall acceptance times efficiency for A is also illustrated

[3] $\pm 3\Gamma$, where Γ is the intrinsic width of the Higgs boson resonance.

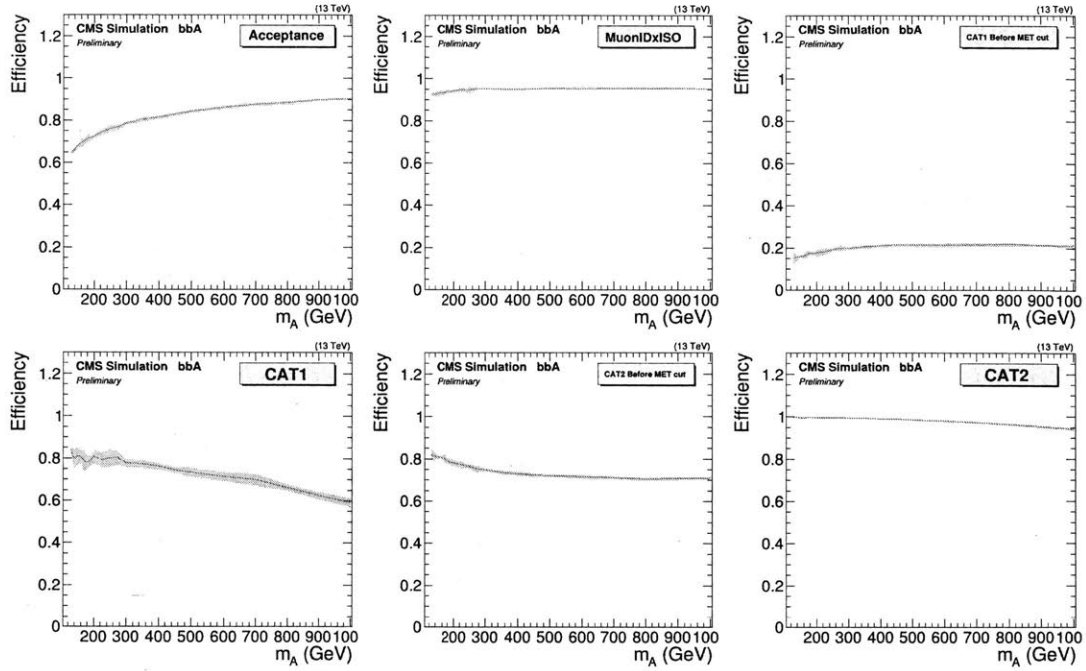


Figure 7-5: The signal acceptance (top-left) and efficiency of muon identification and isolation selection (top-middle), 1 b jet and 0/1 light-flavor jet requirements (top-right) and p_T^{miss} cuts (bottom-left) in b-tagged category, 0 b jet requirements (bottom-middle) and p_T^{miss} cuts (bottom-right) in no-b-tagged category, as a function of m_A for bbA signals.

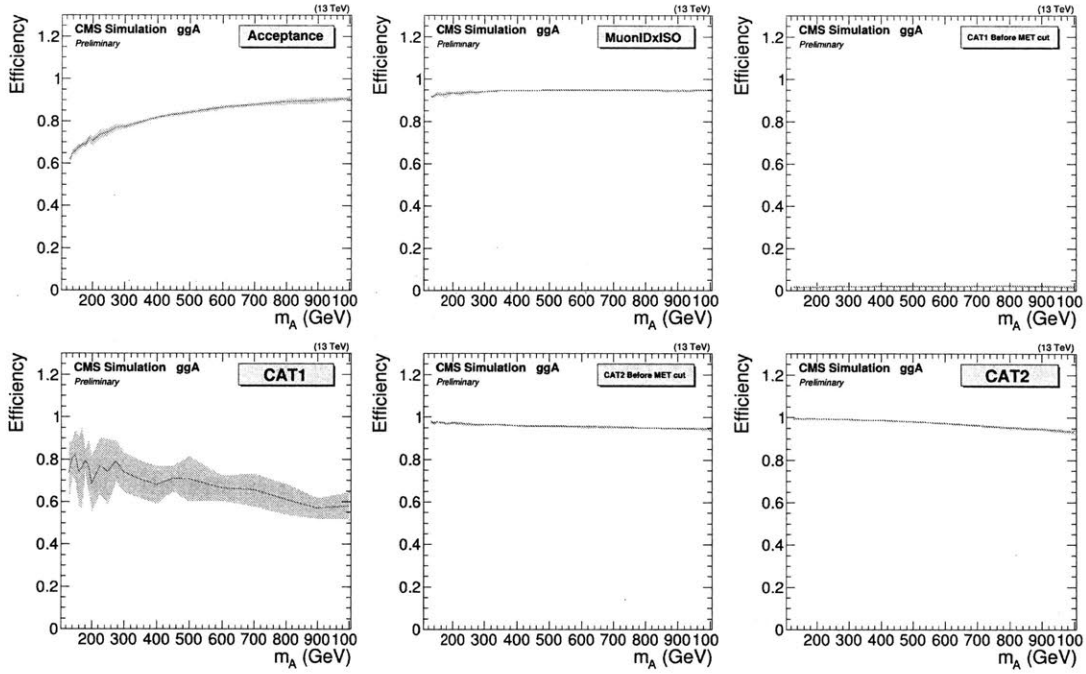


Figure 7-6: The signal acceptance (top-left) and efficiency of muon identification and isolation selection (top-middle), 1 b jet and 0/1 light-flavor jet requirements (top-right) and p_T^{miss} cuts (bottom-left) in b-tagged category, 0 b jet requirements (bottom-middle) and p_T^{miss} cuts (bottom-right) in no-b-tagged category, as a function of m_A for ggA signals.

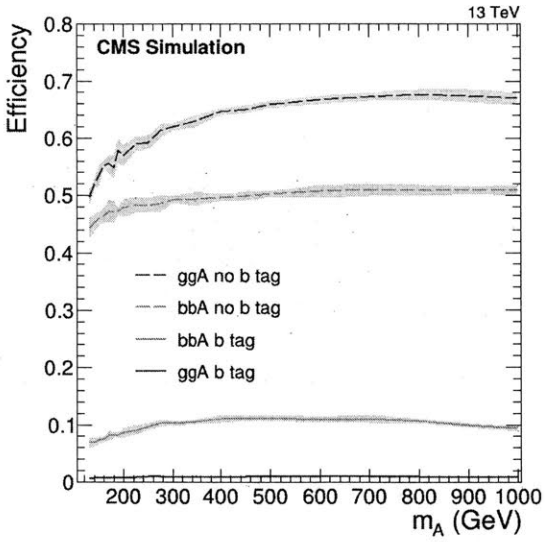


Figure 7-7: Overall acceptance times efficiency of the ggA and bbA production in b-tagged and no-b-tagged category.

in 7-7, for the two production channels and in b-tagged and no-b-tagged category, separately. The corresponding results for the heavy scalar H are consistent with what are shown for A.

7.5 Signal and background modeling

As described previously, the di-muon invariant mass ($M_{\mu\mu}$) distributions of the selected events in the two categories are used to quantify the number of signal through the maximum likelihood (ML) fits.

7.5.1 Signal modeling

Simulated events are exploited to model the shape of the Higgs boson signals. In the model independent interpretation, the signal modeling is done by fitting the single resonance of the simulated A signals of at $tan\beta = 5$ as a template, using the parametric function:

$$f_{sig}^{BW*Gaus}(M_{\mu\mu}) = \frac{1}{(M_{\mu\mu} - M_{\phi})^2 + (\frac{\Gamma_{\phi}}{2})^2} * exp(\frac{M_{\mu\mu}^2}{2\sigma_M^2}), \tag{7.3}$$

which is a convolution of Breit-Wigner, to describe the Higgs boson resonance, and Gaussian function, to account for the detector resolution [4]. The parameters, including the Higgs mass M_ϕ, width Γ_ϕ, and resolution σ_M, are free to float. Once they are determined from the fits, the signal intrinsic width Γ_ϕ is manually set to desired values in the final ML fit, according to the narrow or width resonance assumption. This will be further discussed in Section 7.7.1. This template is also exploited to determine the signal acceptance and efficiency of a generic neutral Higgs boson ϕ decaying to a muon pair, as which are mostly independent from $tan\beta$, as discussed in Figure 7-7.

In the framework of MSSM, the signal shape is modeled by fitting the h, H and A multi-resonances, for various combinations of m_A and $tan\beta$ points. The signal function is defined as

$$f_{sig}(M_{\mu\mu}) = w_h \cdot f_h(M_{\mu\mu}) + w_H \cdot f_H(M_{\mu\mu}) + w_A \cdot f_A(M_{\mu\mu}), \tag{7.4}$$

which consists of a linear combination of $f_{MC}(M_{\mu\mu})$, $f_H(M_{\mu\mu})$ and $f_A(M_{\mu\mu})$, describing the mass shape of the h, H and A, respectively. Each function is a $f^{BW*Gaus}$ as described in Equation 7.3. The parameter M_ϕ, Γ_ϕ, and σ_M are left free to float in the fits while corresponding weights, w_h, w_H, and w_A, are determined from the number of expected events of each Higgs bosons passing the event selection in the specific MSSM scenario. Fit examples are illustrated in Figure 7-8, for the case of $m_A = 140$ GeV and $tan\beta = 11$. For the m_A-$tan\beta$ combinations where the simulated signal events are not generated, those parameters are interpolated from nearby points. Additionally, as the Higgs boson masses in PYTHIA generator are calculated approximately, there are differences of the order of few GeV from the values from FEYNHIGGS predictions, especially below 200 GeV. Therefore, the $M_{\mu\mu}$ distributions are shifted by the corresponding amount.

[4] A Crystal ball model was also evaluated but didn't work well in the sense that it describes the tail but not the peak

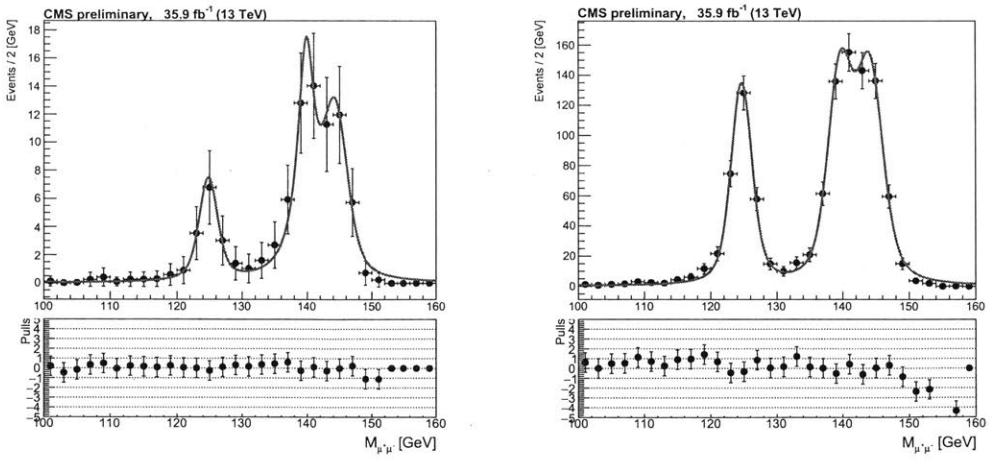


Figure 7-8: Fit example of signal modeling, for $m_A = 140$ GeV and $tan\beta = 11$, and for b-tagged category (left) and no-b-tagged category (right), respectively.

7.5.2 Background modeling

The background contributions are modeled using a data-driven approach and not estimated from simulation, as which has rather limited number of events and high theoretical uncertainties in high-mass region. For this reason, as well as given the smooth dependence on the $M_{\mu\mu}$ distributions, the backgrounds are modeled by fitting the $M_{\mu\mu}$ distributions with a parametric function. Various function families have been investigated, including *Bernstein*, *Laurent*, *Exponential*, *Polynomial*, *Power law* series as well as other physically motivated functions based on the Z peak and decaying tail shape. f-tests have been exploited to decide the order of a specific function series and χ^2-tests and bias studies (see Section 7.6.2 for details) are further utilized to determine the nominal function form for the backgrounds. The final one deployed in the ML fit is:

$$f_{bkg}^{BWZ\gamma}(M_{\mu\mu}) = exp(\lambda M_{\mu\mu}) \cdot [w \cdot \frac{1}{(M_{\mu\mu} - M_Z)^2 + (\frac{\Gamma_Z}{2})^2} + (1-w) \cdot \frac{1}{M_{\mu\mu}^2}], \quad (7.5)$$

which is factorized into the exponential part, the Z boson Breit-Wigner resonance and the photon-exchange contribution. M_Z and Γ_Z are determined in advance by fitting the $M_{\mu\mu}$ around Z boson mass and then fixed in the fits, while λ and w are free to float.

7.6 Systematic uncertainty

The normalization and shape of the signals and backgrounds determined using the procedure described in Section 7.4 and 7.5 are subject to multiple uncertainties. Those uncertainties are handled by introducing a set of nuisance parameters, θ, so that the signal and background expectation become functions of θ in the ML fits: $N_{sig}(\theta)$, $N_{bkg}(\theta)$, $f_{sig}(\theta)$, and $f_{bkg}(\theta)$. Section 7.6.1 and 7.6.2 discuss the estimation of these uncertainties prior to the scrutiny of the observed data entering the statistical analysis, and the corresponding impacts on the analysis are presented in Section 7.6.3.

7.6.1 Uncertainty on signal modeling

Various sources of systematics are considered to account for potential mismodeling of the signal shape and normalization.

The shape of the reconstructed distribution of the Higgs boson mass is affected by the muon momentum scale and resolution, and corresponding uncertainties are propagated to the $M_{\mu\mu}$ shape assuming a Gaussian distribution, yielding maximal variation of 0.05% in the mass position and 10% in the width. These uncertainties are then taken as signal shape variations in the exclusion limit calculation.

The sources of systematics that impact signal acceptance and efficiency are summarized in Table 7.3. The limited number of simulated signal events introduces an uncertainty at the order of 0.2 – 6%. To correct for the differences between data and simulation on muon trigger, identification and isolation efficiencies as well as b tagging efficiencies, corresponding data-to-simulation scale factors are applied to reweight simulated events and thus affect signal efficiencies. The uncertainties associated to those scale factors are propagated to the final signal efficiency and have an impact of 1 – 2%. The uncertainties on the weights to account for mismodeling of pile-up distributions in simulated events are determined in a similar way and introduce systematics of $\leq$1%. Variations that correspond to jet energy correction and unclustered energy are propagated p_T^{miss} calculation to estimate their impacts. The uncertainty on luminosity calibration is 2.5% [151].

Source	Systematic uncertainty(%)	
	b tag	no b tag
MC statistics	0.5–6	0.2–2
Trigger efficiency	0.9	0.9
Muon reconstruction	2	2
Muon Isolation	1	2
Pile-up	0.8	0.9
Jet energy scale	1.6	0.4
Unclustered energy	4.1	0.3
Luminosity	2.5	2.5
PDF	3	3
Higgs boson p_T	1–4	1–4
B tag	2	0.6
B jet multiplicity	20–30	7–20
light-flavor jet multiplicity	7–25	–

Table 7.3: Systematics on the signal normalization for the two event categories.

Theoretical uncertainty is also assessed on signal acceptance and efficiencies. Uncertainties on the Higgs production cross sections are quoted following the LHC cross section working group recommendations [9], when interpreting results under MSSM. The PDF set used is NNPDF3.0 [116] and 3% uncertainty from that is considered on signal efficiency. Furthermore, as the simulated signal events are generated with PYTHIA at LO, signal acceptance and efficiencies are corrected as a function of m_A to match the NLO results, in order to account for the higher order effects on muon acceptance, jet multiplicity and event migration between the two categories. The uncertainty of the correction is set to be the size of the correction itself. The corresponding results are shown in Figure 7-9. The NLO values are determined using signal events produced by POWHEG for the $gg\phi$ and aMC@NLO for the $bb\phi$ signals. As can be seen, the correction on di-muon acceptance is 1 – 4% and that on b jet multiplicity is 20 – 30% in the b-tagged category and 7 – 20% in the no-b-tagged category. An additional correction is applied to b-tagged category on the light-flavor jet multiplicity, which is 7 – 25%. Those uncertainties are summarized in the last three lines of Table 7.3, which only apply to the $bb\phi$ signals in the model independent analysis, and are weighted by the relative contribution of $bb\phi$ signals in MSSM

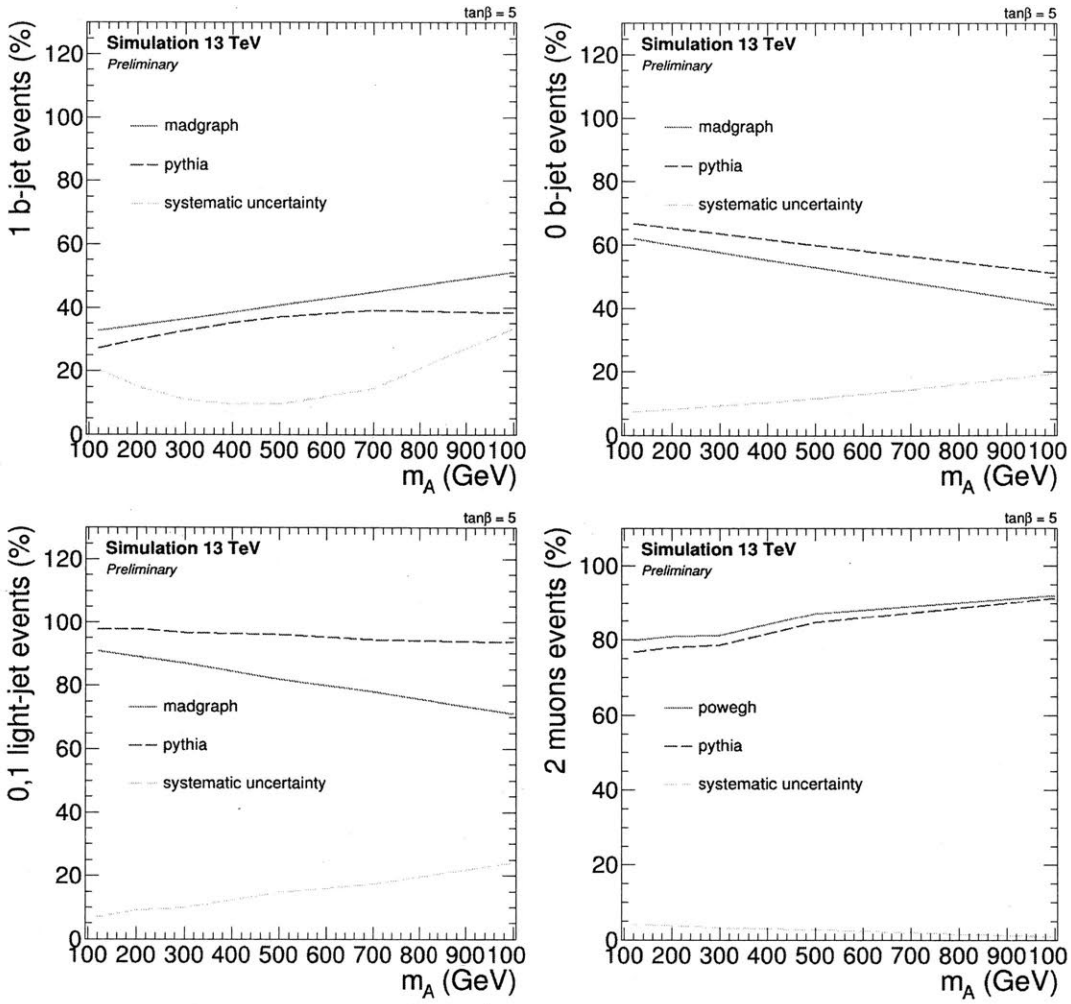


Figure 7-9: Signal acceptance and efficiency result from LO, NLO samples and their differences as a function of m_A, including one b jet (top-left), zero b jet (top-right), zero or one light-flavor jet (bottom-left) and di-muon selection (bottom-right).

interpretation.

7.6.2 Uncertainty on background modeling

The uncertainty on background modeling comes from the choice of the parametric functions. Alternative function forms, including:

$$f_{bkg}^{BWZRedux}(M_{\mu\mu}) = \frac{exp(\alpha_a M_{\mu\mu} + \alpha_b M_{\mu\mu}^2)}{(M_{\mu\mu} - M_Z)^{\alpha_c} + (\frac{\Gamma_Z}{2})^{\alpha_c}},$$
$$f_{bkg}^{SumExp}(M_{\mu\mu}) = \sum_{i=1}^{n} \beta exp(\alpha_i M_{\mu\mu}), \tag{7.6}$$

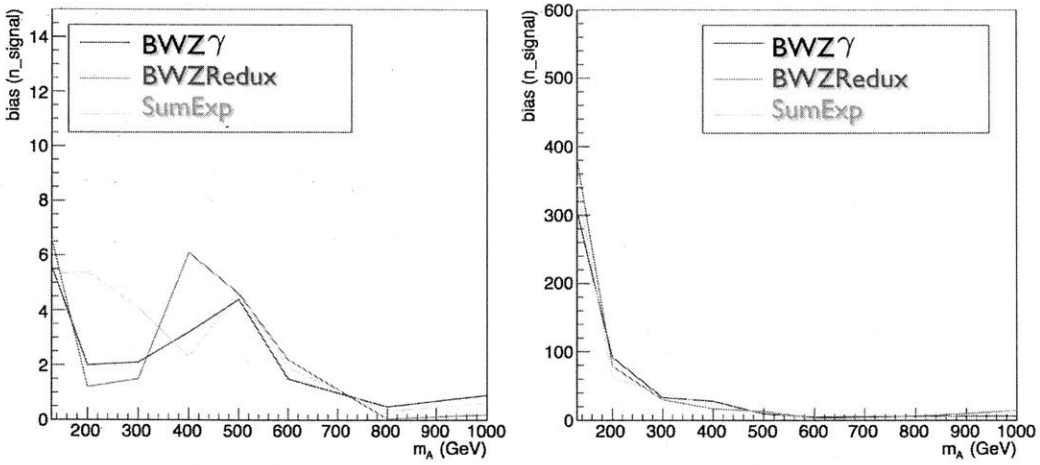


Figure 7-10: Number of bias events in *b-tagged* (left) and *no-b-tagged* (right) category as a function of the Higgs boson mass, for different reference function forms.

are exploited in order to determine the spurious signal yields that are introduced by the choice of functions. The following procedure is repeated for each of the Higgs boson mass assumptions, and among different pairs of reference (f_{bkg}^{ref}) – alternative (f_{bkg}^{alt}) function forms:

1. Fit $M_{\mu\mu}$ in data with f_{bkg}^{ref} and generated 2000 toys from obtained parameters.
2. Fit $M_{\mu\mu}$ in toys with $f_{sig} + f_{bkg}^{ref}$ to determine spurious signal yields N_{sig}^{ref}.
3. Fit $M_{\mu\mu}$ in toys with $f_{sig} + f_{bkg}^{alt}$ to determine spurious signal yields N_{sig}^{alt}.
4. The median of $N_{sig}^{ref} - N_{sig}^{alt}$ distribution is defined as bias.

The largest bias, N_{bias}, is taken to be uncertainty on the signal yields due to the choice of the reference function, for the corresponding mass points. The results for different reference functions are shown in Figure 7-10, which suggests the three function forms lead to similar biases over the entire mass range. The induced spurious signal is then treated as an additional background in the ML fit, whose shape is same as signal and normalization is subjected to Gaussian $\mathcal{N}(0, N_{bias})$.

7.6.3 Impacts of systematic uncertainties

Figure 7-11 shows the impact on the signal strength of the aforementioned systematic uncertainties and the constraint power on them from data. The impact is determined

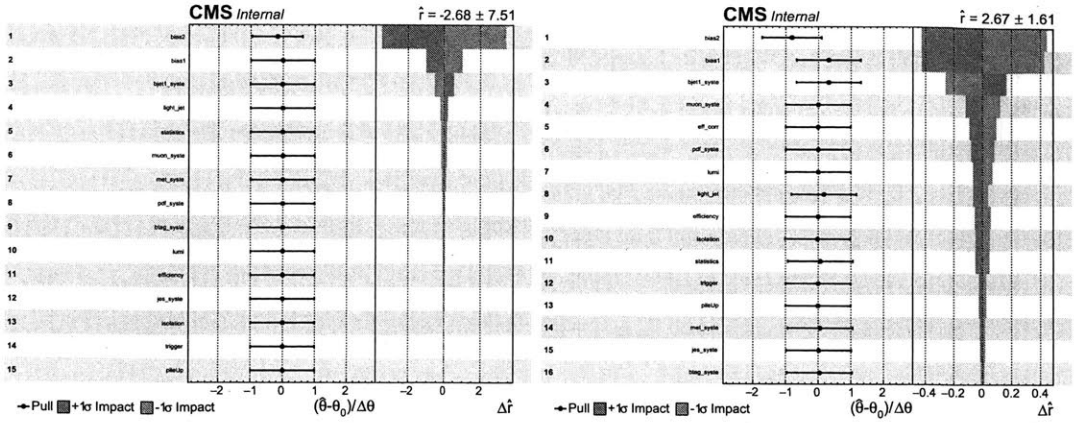


Figure 7-11: Impact of systematics for the search of a generic Higgs boson produced associated with b quark at a mass of 140 (left) and 400 (right) GeV.

through varying the corresponding nuisance parameters by $\pm 1\sigma$. As can be observed, the uncertainty due to background modeling has the leading impact, compared to other uncertainty sources from signal modeling.

7.7 Results

As described previously, to extract the number of signal, simultaneous maximum likelihood fits are performed to the $M_{\mu\mu}$ distributions in all event categories for statistical analysis. The likelihood function is defined as:

$$\mathcal{L}(data|\mu,\theta) = Poisson(data|\mu N_{sig} f_{sig} + N_{bkg} f_{bkg}) \cdot p(\tilde{\theta}|\theta). \tag{7.7}$$

Here, N_{sig}, N_{bkg}, f_{sig}, and f_{bkg} are the normalization and shape of the signal and background contribution, which are described in Section 7.4 and 7.5. The parameter μ denotes the signal strength and $p(\tilde{\theta}|\theta)$ defines the systematic uncertainty probability density functions (p.d.f), where θ represents the entire set of nuisance parameters and the $\tilde{\theta}$ is the estimated values, as determined in Section 7.6. For each of the Higgs boson mass assumptions, the fit is performed in a dedicated mass region, which is chosen to compromise between small coverage to ensure fit stability, small fit bias and large coverage to account for signal width and detector resolution. The mass windows

m_ϕ mass assumption (GeV)	$M_{\mu\mu}$ fit mass window (GeV)
m_ϕ <300	$m_\phi \pm$ 50
300$\leq m_\phi$ <400	$m_\phi \pm$ 75
400$\leq m_\phi$ <500	$m_\phi \pm$ 100
$m_\phi \geq$500	400< m_ϕ <1200

Table 7.4: Fit mass window for different Higgs boson mass assumptions.

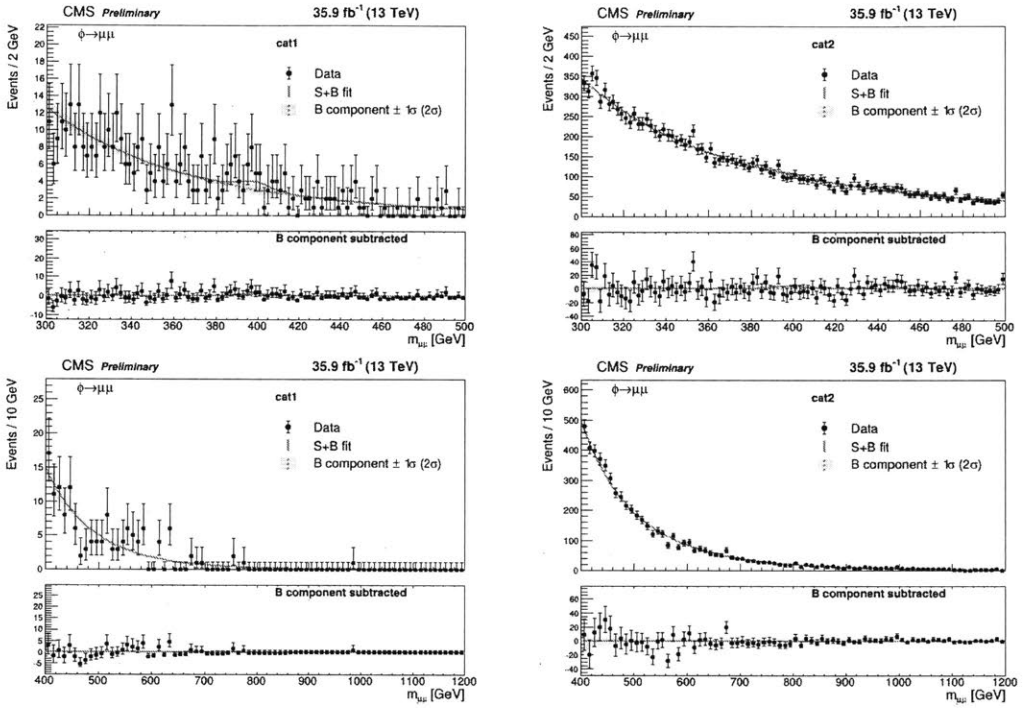


Figure 7-12: Fit examples for signal plus background modeling for a narrow width resonance with mass assumptions of 400 GeV (top) and 980 GeV (bottom), and for b-tagged category (left) and no-b-tagged category (right), respectively.

deployed in the fits are shown in Table 7.4. Fit examples are shown in Figure 7-12 to illustrate the signal-plus-background fits to data and the background components for two different mass assumptions of a narrow width resonance.

No significant evidence is observed for neutral Higgs bosons beyond the SM production in the mass range from 130 to 1000 GeV. Therefore, the results of this search are presented by setting the upper limits on the presence of a signal under the model independent and model dependent interpretation of data. The limits are computed using the Confidence Level (CL) criterion [152, 153] and the hybrid frequentist-bayesian

approach [154], where the test statistic $\tilde{q}_\mu$ is constructed based on the likelihood ratio:

$$\tilde{q}_\mu = -2ln\frac{\mathcal{L}(data|\mu,\hat{\theta}_\mu)}{\mathcal{L}(data|\hat{\mu},\hat{\theta})}, 0 \leq \hat{\mu} \leq \mu. \tag{7.8}$$

Here, $\hat{\mu}$, $\hat{\theta}$ correspond to the global maximum of the likelihood, while $\hat{\theta}_\mu$ represents the conditional maximum given μ. The distributions of the test-statistic are derived from pseudo-experiments.

7.7.1 Model independent limits

A single generic neutral Higgs boson ϕ is searched with the mass assumption from 130 GeV to 1 TeV, and the 95% CL upper limits are set on the Higgs production cross section times branching ratio decaying to two muons: $\sigma(pp \to \phi) \times \mathcal{B}(\phi \to \mu\mu)$. The model independent interpretation is performed both under the narrow width resonance assumption, in which case the intrinsic width is smaller than the $M_{\mu\mu}$ resolution, and a wide width assumption of 10% of the ϕ boson mass m_ϕ, in which case the resonance width is larger than $M_{\mu\mu}$ resolution even for high mass region near 1 TeV. For both width assumptions, the ϕ boson is assumed to be produced entirely either associated with b quarks ($bb\phi$) or by gluon-gluon fusion ($gg\phi$), and the limits are computed separately for the two production mechanisms. The corresponding results of 95% CL expected upper limits, including the one and two standard deviation bands, and the observed upper limits on $\sigma(pp \to \phi) \times \mathcal{B}(\phi \to \mu\mu)$ are shown as a function of m_ϕ in Figure 7-13 and 7-14, for narrow and wide width resonance assumptions, respectively. For $bb\phi$ production, the expected limits for a narrow width resonance range between 20 fb at m_ϕ= 130 GeV and 0.7 fb at m_ϕ= 1 TeV. For $gg\phi$ case, they range between 35 fb at m_ϕ= 130 GeV and 0.7 fb at m_ϕ= 1 TeV. No significant deviation from the expectation is observed. As expected, the results under the wide width assumption are worse by a factor of 3, compared to the narrow width assumption. Additionally, it is no longer possible to distinguish the fine structure of the limits, as observed in the narrow width case.

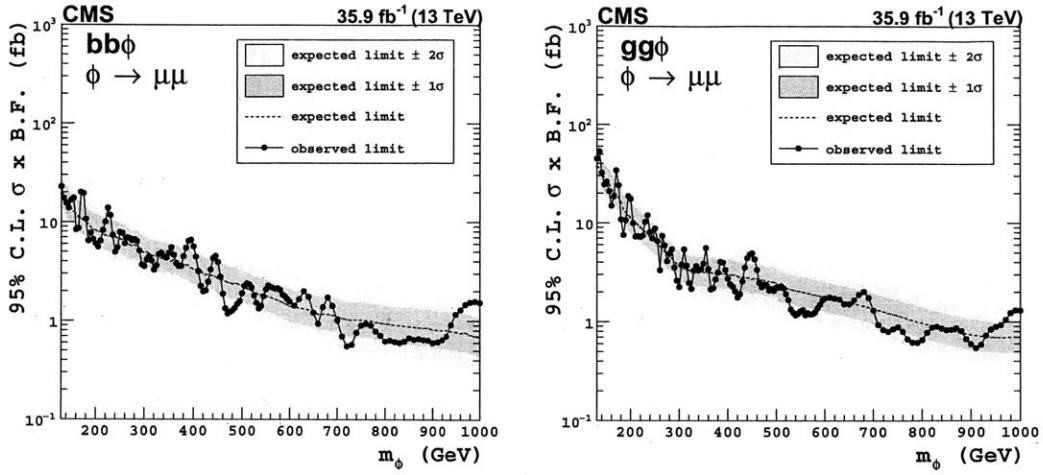


Figure 7-13: The 95% CL expected, including the one and two standard deviations, and observed upper limits on $\sigma(pp \to \phi) \times \mathcal{B}(\phi \to \mu\mu)$ as a function of m_ϕ, for $bb\phi$ (left) and $gg\phi$ (right) using the narrow width assumption.

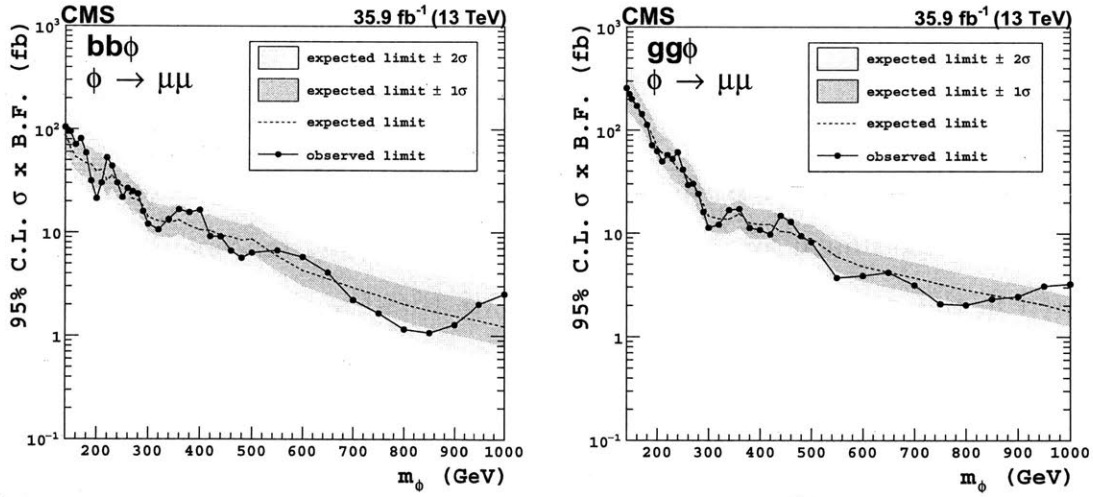


Figure 7-14: The 95% CL expected, including the one and two standard deviations, and observed upper limits on $\sigma(pp \to \phi) \times \mathcal{B}(\phi \to \mu\mu)$ as a function of m_ϕ, for $bb\phi$ (left) and $gg\phi$ (right) using a wide width assumption of 10% of the m_ϕ.

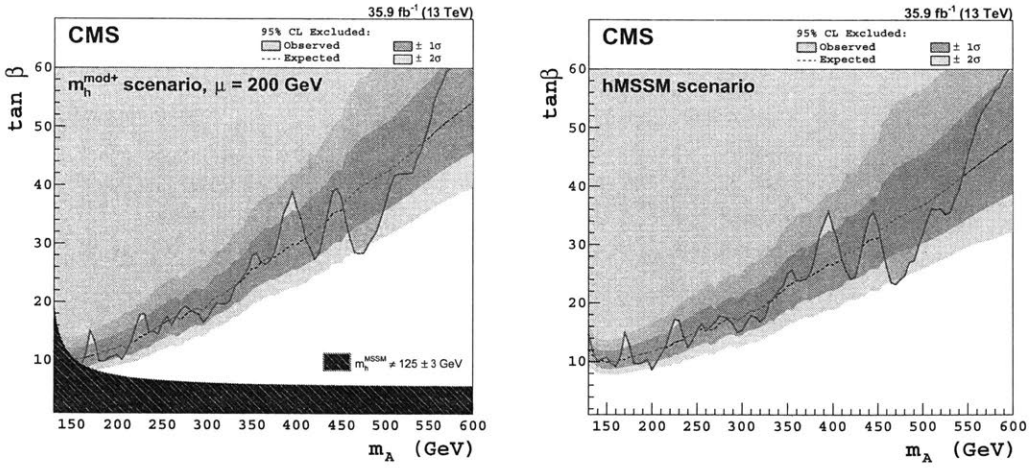


Figure 7-15: The 95% CL expected, including the one and two standard deviations, and observed exclusion contours on m_A-$tan\beta$ plane for the MSSM m_h^{mod+} (left) and hMSSM (right) benchmark scenarios.

7.7.2 Model dependent limits

In the model dependent interpretation of data, exclusion contours are determined in the m_A-$tan\beta$ plane for representative benchmark scenarios of the MSSM. Figure 7-15 shows the observed and expected 95% CL exclusion contours, including the one and two standard deviation bands, for the m_h^{mod+} and the hMSSM scenario. The exclusion contours reach up to 600 GeV, where the excluded $tan\beta$ value exceeds 60. For higher $tan\beta$ values, the MSSM predictions are no longer reliable. The contours are very similar between the two scenarios, as the h boson masses predicted by m_h^{mod+} are well consistent with the SM Higgs boson mass, and the production cross sections of A and H are also similar for the two scenarios inside the m_A-$tan\beta$ parameter space considered in this analysis. Compared with CMS Run-I analysis [155], the results reported here exclude a larger parameter space and extend to a larger range of m_A. The contours determined in the $\tau\tau$ final state search [156] exclude a much larger parameter space, reaching m_A= 1.6 TeV for $tan\beta$= 60, as the $\mathcal{B}(\phi \to \tau\tau)$ is about 300 times larger than $\mathcal{B}(\phi \to \mu\mu)$. The contours obtained in the $b\bar{b}$ final state, instead, exclude a smaller m_A-$tan\beta$ region, even if $\mathcal{B}(\phi \to b\bar{b})$ is larger.

Chapter 8

Conclusion and Outlook

The LHC has successfully completed the second data-taking period from 2015 to 2018. Proton-proton collisions at the center-of-mass energy $\sqrt{s}$ =13 TeV have been recorded by the CMS detector, corresponding to a data sample with an integrated luminosity of 150 fb^{-1}. This thesis described three analyses of di-muon final-state events exploiting the collected data.

Measurements of the inclusive and differential Z boson production cross sections are performed using 2015 dataset, corresponding to an integrated luminosity of 2.3 fb^{-1}. The measured total inclusive cross section times branching ratio is $\sigma(pp \rightarrow ZX) \times \mathcal{B}(Z \rightarrow \mu\mu) = 1870 \pm 2(stat) \pm 35(syst) \pm 51(lumi)$ pb for the di-muon invariant mass in the range of 60 to 120 GeV. This result is in good agreement with the NNLO QCD calculations. The differential cross sections are also measured as functions of Z boson transverse momentum, $\phi^{\star}_{\eta}$ variable, rapidity, and the muon transverse momentum, and compared with NLO and NNLO theoretical predictions.

The large production cross section of Z boson and good experimental accessibility of final-state muons enable a real-time monitoring of luminosities using the counts of reconstructed $Z \rightarrow \mu\mu$ events. The basic principal, workflow automation, and preliminary results using 2018 data, are documented in this thesis. This Z boson counting approach provided a valuable, reliable, and independent reference to assess and optimize the LHC machine and CMS detector performance during past data-taking period. It also points to a potential high-precision luminosity determination

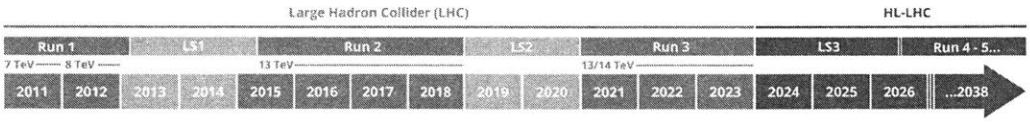


Figure 8-1: LHC and HL-LHC timeline.

method for next era, High-Luminosity LHC (HL-LHC).

A search for beyond the Standard Model neutral Higgs bosons is presented using 2016 dataset, corresponding to an integrated luminosity of 36.9 fb^{-1}. No significant excess is observed. 95% confidence level upper limits are set on $\sigma(pp \to \phi) \times \mathcal{B}(\phi \to \mu\mu)$ for a generic boson ϕ with mass hypotheses ranging from 130 GeV to 1 TeV, produced via gluon fusion and b-quark association, respectively. The results are also interpreted in the context of two MSSM representative benchmark scenarios, which excluded a much larger parameter phase space than previous search in di-muon channel.

No deviations from the Standard Model predictions have been found in above results. Neither were signatures for exciting new physics observed so far in other measurements and searches performed with LHC data. The journey will be continued in the next period and HL-LHC era (see Figure 8-1), to find answer to fundamental questions such as:

- What prevents the Higgs boson mass from getting quantum corrections at the Planck scale?
- What is the nature of dark matter and dark energy?
- Is supersymmetry the solution for above questions?
- ...

A quote from New Scientist [157].

It's an old joke. A woman returning home finds a neighbour searching for his keys beneath a street lamp. "Is that where you dropped them?" she asks. "No," he replies, "but it's where the light is."

Bibliography

[1] CMS Collaboration. Summaries of CMS cross section measurements. http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/SM/. Accessed: 2019-01-12.

[2] CMS Collaboration. Summaries of CMS cross section measurements. https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsCombined. Accessed: 2019-01-12.

[3] University of Zurich. The elementary particles of the Standard Model. https://www.physik.uzh.ch/en/researcharea/lhcb/outreach/StandardModel.html. Accessed: 2019-01-12.

[4] Particle Data Group. The Review of Particle Physics. *Phys. Rev. D 98, 030001 (2018).*

[5] M. Schott and M. Dunford. Review of single vector boson production in pp collisions at $\sqrt{s}$=7 TeV. *Eur. Phys. J. C (2014) 74.*

[6] S. Catani et al. SHERPA 1.α, a proof-of-concept version. *JHEP02 (2004) 056.*

[7] W. Verkerke. The QCD improved parton model. https://www.nikhef.nl/pub/experiments/zeus/theses/wouter_verkerke/latex2html/node9.html. Accessed: 2019-01-12.

[8] ATLAS Collaboration. Expected Performance of the ATLAS Experiment Detector, Trigger and Physics. *CERN-OPEN-2008-020 (2008).*

[9] D. de Florian et al. Handbook of LHC Higgs cross sections: 4. deciphering the nature of the Higgs sector. *CERN-2017-002-M (2017).*

[10] CMS Collaboration. CMS luminosity: public results. `https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults`. Accessed: 2019-01-12.

[11] CERN. The CERN accelerator complex. *OPEN-PHO-ACCEL-2013-056 (2013).*

[12] CMS Collaboration. 3D models of the CMS detector and events in SketchUp. `https://twiki.cern.ch/twiki/bin/view/CMSPublic/SketchUpCMS`. Accessed: 2019-01-12.

[13] University of Zurich. Simple example of 3D axes with spherical coordinates. `https://wiki.physik.uzh.ch/cms/latex:example_spherical_coordinates`. Accessed: 2019-01-12.

[14] D. Bortoletto. How and why silicon sensors are becoming more and more intelligent? *JINST 10 C08016 (2015).*

[15] CMS Collaboration. Description and performance of the track and primary-vertex reconstruction with the CMS tracker. *JINST 9 P10009 (2014).*

[16] A. Benaglia. The CMS ECAL performance with examples. *JINST 9 C02008 (2014).*

[17] CMS Collaboration. Materials for CMS ECAL poster. `https://cms-docdb.cern.ch/cgi-bin/PublicDocDB/ShowDocument?docid=12030`. Accessed: 2019-01-12.

[18] CMS Collaboration. The CMS experiment at the CERN LHC. *JINST 3 S08004 (2008).*

[19] CMS Collaboration. Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s}$=13 TeV. *JINST 13 P06015 (2018).*

[20] CMS Collaboration. The performance of the CMS muon detector in proton-proton collisions at $\sqrt{s}$=7 TeV at the LHC. *JINST 8 P11002 (2013).*

[21] CMS Collaboration. Particle-flow reconstruction and global event description with the CMS detector. *JINST 12 P1003 (2017).*

[22] P. Lujan. Luminosity at cms and plans for hl-lhc. In *32nd FCAL workshop*, 2018.

[23] CMS Collaboration. CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s}$ = 13 TeV. *CMS-PAS-LUM-17-004 (2017).*

[24] W. Badgett et al. Web based monitoring in the CMS experiment at CERN. *J. Phys.: Conf. Ser. 331 022025 (2011).*

[25] J.Boyd. The LHC machine in run-2. In *international school of subnuclear physics*, 2018.

[26] L. Evans and P. Bryant. LHC Machine. *JINST 3 S08003 (2008).*

[27] ATLAS Collaboration. Measurement of the inclusive jet cross-sections in proton-proton collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector. *JHEP20 (2017) 2017.*

[28] CMS Collaboration. Measurement of the double-differential inclusive jet cross section in protonâĂŞproton collisions at $\sqrt{s}$ = 13 TeV. *EPJC 76 (2016) 451.*

[29] ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B716 (2012) 1.*

[30] CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett. B716 (2012) 30.*

[31] ATLAS and CMS Collaborations. Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s}$ = 7 and 8 TeV with the ATLAS and CMS Experiments. *Phys. Rev. Lett. 114 (2015) 191803.*

[32] F. Jegerlehner. The hierarchy problem of the electroweak Standard Model revisited. *DESY-13-093 (2013).*

[33] G. Bertone, D. Hooper and J. Silk. Particle dark matter: Evidence, candidates and constraints. *Phys. Rept. 405 (2005) 279.*

[34] J.L. Feng. Dark Matter Candidates from Particle Physics and Methods of Detection. *Ann. Rev. Astron. Astrophys. 48 (2010) 495.*

[35] T.A. Porter, R.P. Johnson and P.W. Graham. Dark Matter Searches with Astroparticle Data. *Ann. Rev. Astron. Astrophys. 49 (2011) 155.*

[36] Y. Nambu and G. Jona-Lasinio. Dynamical model of elementary particles based on an analogy with superconductivity. I. *Phys. Rev. 122 (1961) 345.*

[37] CMS Collaboration. Search for the Higgs Boson Decaying to Two Muons in Proton-Proton Collisions at $\sqrt{s} = 13$ TeV. *Phys. Rev. Lett. 122, 021801 (2018).*

[38] M. Gell-Mann. A Schematic Model of Baryons and Mesons. *Phys. Lett. 8 (1964) 214.*

[39] M. Gell-Mann. Quarks. *Acta Phys. Austriaca Suppl. 9 (1972) 733.*

[40] M. Han and Y. Nambu. Three-Triplet Model with Double SU(3) Symmetry. *Phys. Rev. 139B (1965) 1006.*

[41] D.J. Gross and F. Wilczek. Ultraviolet Behavior of Non-Abelian Gauge Theories. *Phys. Rev. Lett. 30, 1343 (1973).*

[42] H.D. Politzer. Reliable Perturbative Results for Strong Interactions? *Phys. Rev. Lett. 30, 1346 (1973).*

[43] S.L. Glashow. Partial Symmetries of Weak Interactions. *Nucl. Phys. 22 (1961) 579.*

[44] S. Weinberg. A Model of Leptons. *Phys. Rev. Lett. 19 (1967) 1264.*

[45] A. Salam. Weak and Electromagnetic Interactions. *Conf. Proc. C680519 (1968) 367.*

[46] S.L. Glashow, J. Iliopoulos and L. Maiani. Weak Interactions with Lepton-Hadron Symmetry. *Phys. Rev. D2 (1970) 1285.*

[47] F. Englert and R. Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Phys. Rev. Lett. 13 (1964) 321.*

[48] P.W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett. 13 (1964) 508.*

[49] P.W. Higgs. Spontaneous Symmetry Breakdown without Massless Bosons. *Phys. Rev. 145 (1966) 1156.*

[50] G.S. Guralnik, C.R. Hagen and T.W.B. Kibble. Global Conservation Laws and Massless Particles. *Phys. Rev. Lett. 13 (1964) 585.*

[51] T.W.B. Kibble. Symmetry Breaking in Non-Abelian Gauge Theories. *Phys. Rev. 155 (1967) 1554.*

[52] Yu.A. Golfand and E.P. Likhtm. Extension of the algebra of Poincare group generators and violation of p invariance. *JETP Lett. 13 (1971) 323.*

[53] J. Wess and B. Zumino. Supergauge transformations in four-dimensions. *Nucl. Phys. B 70 (1974) 39.*

[54] P. Fayet. Supersymmetry and Weak, Electromagnetic and Strong Interactions. *Phys. Lett. B64, 159 (1976).*

[55] G.R. Farrar and P. Fayet. Phenomenology of the Production, Decay, and Detection of New Hadronic States Associated with Supersymmetry. *Phys. Lett. B76, 575 (1978).*

[56] S. Dimopoulos and H. Georgi. Softly Broken Supersymmetry and SU(5). *Nucl. Phys. B193, 150 (1981).*

[57] V.N. Gribov and L.N. Lipatov. Deep inelastic e p scattering in perturbation theory. *Sov. J. Nucl. Phys. 15 (1972) 438.*

[58] Yu. L. Dokshitzer. Calculation of structure functions of deep-inelastic scattering and e^+e^- annihilation by perturbation theory in quantum chromodynamics. *Sov. Phys. JETP 46 (1977) 641.*

[59] G. Altarelli and G. Parisi. Asymptotic Freedom in Parton Language. *Nucl. Phys. B126 (1977) 298.*

[60] R. Hamberg, W.L. van Neerven and T. Matsuura. A complete calculation of the order α_s^2 correction to the Drell-Yan K factor. *Nucl. Phys. B359, 343 (1991).*

[61] K. Melnikov and F. Petriello. Electroweak gauge boson production at hadron colliders through $\mathcal{O}(\alpha_s^2)$. *Phys. Rev. D74, 114017 (2006).*

[62] Y.L. Dokshitzer, D.Diakonov and S.I. Troian. On the transverse momentum distribution of massive lepton pairs. *Phys. Lett. 79B (1978) 269.*

[63] G. Parisi and R. Petronzio. Small transverse momentum distributions in hard processes. *Nucl. Phys. B154, 427 (1979).*

[64] J.C. Collins, D.E. Soper and G. Sterman. Transverse momentum distribution in Drell-Yan pair and W and Z boson production. *Nucl. Phys. B250, 199 (1985).*

[65] C. Balazs and C.P. Yuan. Soft gluon effects on lepton pairs at hadron colliders. *Phys. Rev. D56, 5558 (1997).*

[66] S. Catani et al. Vector boson production at hadron colliders: transverse-momentum resummation and leptonic decay. *JHEP12 (2015) 047.*

[67] B. Andersson et al. Parton fragmentation and string dynamics. *Phys. Reports 97, 31 (1983).*

[68] T. Sjöstrand. The Lund Monte Carlo for Jet Fragmentation and e^+e^- Physics: Jetset Version 6.2. *Nucl. Phys. B248, 469 (1984).*

[69] B.R. Webber. A QCD Model for Jet Fragmentation Including Soft Gluon Interference. *Nucl. Phys. B238, 492 (1984).*

[70] GEANT4 Collaboration. GEANT4 âĂŞ a simulation toolkit. *Nucl. Instrum. Meth. A 506 (2003) 250.*

[71] T. Sjöstrand et al. An introduction to PYTHIA 8.2. *Comp. Phys. Comm. 191, 159 (2015).*

[72] CMS Collaboration. Underlying Event Tunes and Double Parton Scattering. *EPJC (2016) 76: 155.*

[73] T. Sjöstrand et al. An introduction to PYTHIA 8.2. *Comp. Phys. Comm. 191, 159 (2015).*

[74] S. Catani et al. QCD matrix elements + parton showers. *JHEP 11, 063 (2001).*

[75] F. Krauss. Matrix elements and parton showers in hadronic interactions. *JHEP 08, 015 (2002).*

[76] M.L. Mangano et al. matrix elements and shower evolution in hadronic collisions: $Wb\bar{b}$ + n jets as a case study. *Nucl.Phys. B632:343âĂŞ362 (2002).*

[77] S. Frixione and B. Webber. Matching NLO QCD computations and parton shower simulations. *JHEP 06, 029 (2002).*

[78] J. Alwall et al. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP 07, 079 (2014).*

[79] P. Nason. A New Method for Combining NLO QCD with Shower Monte Carlo Algorithms. *JHEP 11, 040 (2004).*

[80] S. Frixione et al. Matching NLO QCD computations with parton shower simulations: the POWHEG method. *JHEP 11, 070 (2007).*

[81] S. Alioli et al. A General Framework for Implementing NLO Calculations in Shower Monte Carlo Programs: the POWHEG BOX. *JHEP 06, 043 (2010).*

[82] S. Alioli et al. NLO Vector Boson Production Matched with Shower in POWHEG. *JHEP 07, 060 (2008).*

[83] R. Frederix and S. Frixione. Merging meets matching in MC@NLO. *JHEP 12, 061 (2012).*

[84] CMS Collaboration. Summary results of high mass BSM Higgs searches using CMS run-I data. *CMS-PAS-HIG-16-007 (2016).*

[85] M. Carena et al. MSSM Higgs Boson Searches at the LHC: Benchmark Scenarios after the Discovery of a Higgs-like Particle. *EPJC 73 (2013) 2552.*

[86] ATLAS and CMS Collaboration. Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV. *JHEP 08 (2016) 045.*

[87] ATLAS Collaboration. Measurements of the Higgs boson production and decay rates and coupling strengths using pp collision data at $\sqrt{s} = 7$ and 8 TeV in the ATLAS experiment. *EPJC 76 (2016) 6.*

[88] CMS Collaboration. Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV. *EPJC 75 (2015) 212.*

[89] L. Maiani, A.D. Polosa and V. Riquer. Bounds to the Higgs sector masses in minimal supersymmetry from LHC data. *Phys. Lett. B 724 (2013) 274.*

[90] A. Djouadi et al. The post-Higgs MSSM scenario: Habemus MSSM? *EPJC 73 (2013) 2650.*

[91] A. Djouadi et al. Fully covering the MSSM Higgs sector at the LHC. *JHEP 06 (2015) 168.*

[92] ATLAS Collabration. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST 3 S08003 (2008).*

[93] LHCb Collabration. The LHCb Detector at the LHC. *JINST 3 S08005 (2008).*

[94] ALICE Collabration. The ALICE experiment at the CERN LHC . *JINST 3 S08002 (2008).*

[95] CMS Collaboration. The CMS trigger system. *JINST 12 P01020 (2017).*

[96] G. Bauer et al. The data-acquisition system of the CMS experiment at the LHC. *J. Phys.: Conf. Ser. 331 022021 (2011).*

[97] T. Speer et al. Vertex Fitting in the CMS Tracker. *CMS Note 2006-032 (2006).*

[98] M. Swartz et al. A new technique for the reconstruction, validation, and simulation of hits in the CMS Pixel Detector. *PoS VERTEX2007 (2007) 035.*

[99] R. Fruhwirth. Application of Kalman filtering to track and vertex fitting. *Nucl. Instrum. Meth. A 262 (1987) 444.*

[100] K. Rose. Deterministic Annealing for Clustering, Compression, Classification, Regression and related Optimisation Problems. *Proc. IEEE 86 (1998) 2210.*

[101] CMS collaboration. New method of out-of-time energy subtraction for the CMS hadronic calorimeter. *CMS-CR-2018-140 (2018).*

[102] CMS collaboration. The performance of the CMS muon detector in proton-proton collisions at $\sqrt{s}$=7 TeV at the LHC. *JINST 8 P11002 (2013).*

[103] CMS collaboration. Validation of the mean-timer algorithm for DT local reconstruction and muon time measurement, using 2012 data. *CMS-DP-2015-026 (2015).*

[104] CMS collaboration. Performance of CMS muon reconstruction in pp collision events at $\sqrt{s}$=7 TeV. *JINST 7 P10002 (2012).*

[105] CMS Collaboration. Baseline muon selections for Run-II. `https://twiki.cern.ch/twiki/bin/viewauth/CMS/SWGuideMuonIdRun2`. Accessed: 2019-01-12.

[106] W. Adam et all. Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC. *J. Phys. G 31 (2005) N9.*

[107] M. Cacciari, G. Salam and G. Soyez. The anti-k_t jet clustering algorithm. *JHEP04 (2008) 063.*

[108] CMS Collaboration. Pileup Removal Algorithms. *CMS-PAS-JME-14-001 (2014).*

[109] CMS Collaboration. Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV. *JINST 12 P02014 (2017).*

[110] CMS Collaboration. Jet algorithms performance in 13 TeV data. *CMS-PAS-JME-16-003 (2016).*

[111] CMS Collaboration. Identification of heavy-flavor jets with the CMS detector in pp collisions at 13 TeV. *JINST 13 P05011 (2018).*

[112] CMS Collaboration. Performance of missing transverse momentum in proton-proton collisions at $\sqrt{s}$=13 TeV using the CMS detector. *CMS-PAS-JME-17-001 (2017).*

[113] A. Banfi et all. Probing the low transverse momentum domain of Z production with novel variables. *JHEP01 (2012) 044.*

[114] CMS Collaboration. CMS luminosity measurement for the 2015 data-taking period. *CMS PAS LUM-15-001 (2015).*

[115] G. D'Agostini. A multidimensional unfolding method based on Bayes' theorem. *Nucl. Instrum. Meth. A 362 (1995) 487.*

[116] NNPDF Collaboration. Parton distributions for the LHC Run II. *JHEP 04 (2015) 040.*

[117] C. Calame et all. Precision electroweak calculation of the production of a high transverse-momentum lepton pair at hadron collider. *JHEP 10 (2007) 109.*

[118] CMS Collaboration. Measurement of the inclusive W and Z production cross sections in pp collisions at $\sqrt{s}$= 7 TeV with the CMS experiment. *JHEP 10 (2011) 132.*

[119] G. Breit and E. Wigner. Capture of Slow Neutrons. *Phys. Rev. 49, 519 (1936).*

[120] J. Gaiser. Charmonium Spectroscopy from Radiative Decays of the J/Psi and Psi-Prime. *PhD thesis, Stanford University (1982).*

[121] A. Bodek et al. Extracting muon momentum scale corrections for hadron collider experiments. *EPJC (2012) 72: 2194.*

[122] CMS Collaboration. Measurement of the Top Quark Pair Production Cross Section in Proton-Proton Collisions at $\sqrt{s}$=13 TeV. *Phys. Rev. Lett. 116, 052002 (2016).*

[123] CMS Collaboration. Measurement of the WZ production cross section in pp collisions at $\sqrt{s}$=13 TeV. *CMS-PAS-SMP-15-006 (2015).*

[124] CMS Collaboration. Measurement of the ZZ production cross section in $\ell\ell\ell'\ell'$ decays in pp collisions at $\sqrt{s}$=13 TeV. *CMS-PAS-SMP-15-005 (2015).*

[125] G. Cowan. A survey of unfolding methods for particle physics. *Conf.Proc. C0203181 (2002) 248-257.*

[126] T. Adye. Unfolding algorithms and tests using RooUnfold. *PHYSTAT Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding (2011).*

[127] T. Hastie, R. Tibshirani and J. Friedman. *The Elements of Statistical Learning*, chapter 6. Springer, 2001.

[128] S. Dulat et al. New parton distribution functions from a global analysis of quantum chromodynamics. *Phys. Rev. D 93, 033006 (2016).*

[129] L. Harland et al. Parton distributions in the LHC era: MMHT 2014 PDFs. *EPJC (2015) 75: 204.*

[130] S. Alekhin, J. BlÃijmlein and S. Moch. The ABM parton distributions tuned to LHC data. *Phys. Rev. D 89, 054028.*

[131] A. Cooper. PDF Fits at HERA. *Europhysics Conference on High Energy Physics (2011).*

[132] CMS Collaboration. Measurement of inclusive W and Z boson production cross sections in pp collisions at $\sqrt{s}$= 13 TeV. *CMS-PAS-SMP-15-004 (2015).*

[133] CMS Collaboration. Measurement of the top quark pair production cross section in proton-proton collisions at $\sqrt{s}$= 13 TeV. *Phys. Rev. Lett. 116 (2016) 052002.*

[134] P. Lujan. Performance of the Pixel Luminosity Telescope for luminosity measurement at CMS during Run 2. *PoS 314 (2017) 504.*

[135] M. Hempel. Development of a Novel Diamond Based Detector for Machine Induced Background and Luminosity Measurements. *PhD thesis, DESY, Hamburg (2017).*

[136] S. van der Meer. Calibration of the effective beam height in the ISR. *CERN-ISR-PO-68-31 (1968).*

[137] M. Klute, C. Medlock and J. Salfeld-Nebgen. Beam imaging and luminosity calibration. *JINST 12 (2017) P03018.*

[138] M. Rovere. The Data Quality Monitoring Software for the CMS experiment at the LHC. *J. Phys.: Conf. Ser. 664 072039 (2015).*

[139] A. Dabrowski. The performance of the beam conditions and radiation monitoring system of CMS. *Proceedings of the IEEE Nuclear Science Symposium and Medical Imaging Conference (2001).*

[140] CMS Collaboration. BRIL Work Suite. `https://cms-service-lumi.web.cern.ch/cms-service-lumi/brilwsdoc.html`. Accessed: 2019-01-12.

[141] J. Salfeld-Nebgen and D. Marlow. Data-driven precision luminosity measurements with Z bosons at the LHC and HL-LHC. *JINST 13 P12016 (2018).*

[142] LHCb Collaboration. Precision luminosity measurements at LHCb. *JINST 9 P12005 (2014).*

[143] CMS Collaboration. Observation of $t\bar{t}H$ Production. *Phys. Rev. Lett. 120, 231801 (2018).*

[144] ATLAS Collaboration. Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector. *Phys. Lett. B 784, 173 (2018).*

[145] CMS Collaboration. Observation of the Higgs boson decay to a pair of τ leptons with the CMS detector. *Phys. Lett. B 779, 283 (2018).*

[146] ATLAS Collaboration. Cross-section measurements of the Higgs boson decaying into a pair of ÏĎ-leptons in protonâĂŞproton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *CERN-EP-2018-232 (2018).*

[147] CMS Collaboration. Observation of Higgs Boson Decay to Bottom Quarks. *Phys. Rev. Lett. 121, 121801 (2018).*

[148] ATLAS Collaboration. Observation of $H \rightarrow b\bar{b}$ decays and VH production with the ATLAS detector. *Phys. Lett. B 786, 59 (2018).*

[149] S. Heinemeyer, W. Hollik and G. Weiglein. FeynHiggs: A program for the calculation of the masses of the neutral CP-even Higgs bosons in the MSSM. *Comput. Phys. Commun. 550 124 (2000) 76.*

[150] G. Punzi. Sensitivity of searches for new signals and its optimization. *eConf C030908 (2003).*

[151] CMS Collaboration. CMS luminosity measurement for the 2016 data-taking period. *CMS PAS LUM-17-001 (2017).*

[152] T. Junk. Confidence level computation for combining searches with small statistics. *Nucl. Instrum. Meth. A 434 (1999) 435.*

[153] A. L. Read. Presentation of search results: The CLs technique. *J. Phys. G 28 (2002) 2693.*

[154] ATLAS and CMS Collaborations. Procedure for the LHC Higgs boson search combination in summer 2011. *CMS NOTE 2011/005 (2011).*

[155] CMS Collaboration. Search for neutral MSSM Higgs bosons decaying to $\mu^+\mu^-$ in pp collisions at $\sqrt{s}$= 7 and 8 TeV. *Phys. Lett. B 752 (2016) 221.*

[156] CMS Collaborations. Search for additional neutral MSSM Higgs bosons in the $\tau\tau$ final state in proton-proton collisions at $\sqrt{s}$= 13 TeV. *JHEP 08 (2018) 007.*

[157] Justin Eure. Evidence of new physics could have been under our noses all along. ://www.newscientist.com/article/mg24132170-200-evidence-of-new-physics-could-have-been-under-our-noses-all-along/.