First observation of the production of three massive vector bosons and search for long-lived particles using delayed photons in pp collisions at $\sqrt{s}$ = 13 TeV

Thesis by
Zhicai Zhang

In Partial Fulfillment of the Requirements for the
Degree of
Doctor of Philosophy

Caltech

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2021
Defended September 22, 2020

© 2021

Zhicai Zhang
ORCID: 0000-0002-1630-0986

ACKNOWLEDGEMENTS

This thesis and the work presented here could not have been possible without my advisor, Professor Harvey Newman. A big thank you, Harvey! It was my privilege to be your student over the past five years, and I wish I could continue to learn more from you wherever I go after my graduation. You taught me with tremendous patience how to think, ask, work, and create like a scientist. Your dedication to physics, your enthusiasm in work, and your wealth of knowledge about the field and beyond have been deeply implanted in my soul. You are and will always be my idol, and I wish one day I could be an educator and a physicist like you.

The two and half years at CERN was the core of my learning phase during my PhD, and I cannot imagine how boring it would have been without Dr. Adi Bornheim. Adi, it was great fun being your apprentice and working in labs and on numerous test beams. And the discussions with you day after day in the office were so precious. I hope I will still get a seat in that office and have another great discussion with you when I come to CERN next time after the COVID is over.

The last year at Fermilab was the most productive period during my PhD journey, and that could not have been possible without the great mentors in the lab: Dr. Sergo Jindariani, Dr. Artur Apresyan, and Dr. Si Xie. Thank you for all your support and help. You made me realize what a great place Fermilab is to do great physics.

Although I only got the chance to stay at the Caltech campus for about one and a half years, I am truly grateful for the support of many Caltech colleagues. Thank you to the Caltech CMS group members: Maria Spiropulu, Si Xie, Cristian Pena, Javier Duarte, Joosep Pata, Nan Lu, Josh Bendavid; Dustin Anderson, Thong Nguyen, Jiajing Mao, Olmo Cerri, Irene Dutta, Christina Wang. For your help, support, and discussions both on campus and remotely; for sharing the office with me; and for lunch times, afternoon cake times, morning group meeting bagels.

Special thanks to my friends who never get bored of listening to me and chatting with me about random things every day: Yongbin Feng, Thong Nguyen, Yu Yang, Christina Wang. You cheered up my day and made my every day count. I cannot imagine how depressed I would have been without you. I am also grateful for other friends who shared a lot of fun times with me over the past years: Miao, Xunwu, Jiajing, Wuming . . .

The triboson observation work presented in this thesis is a result of collaboration

with a great team: Philip, Hannsjorg, Sapta, Sergo, Michael, Si, Karri, Mia, Jonas, Yifan. Thanks to all of you. It was a fantastic journey with you guys.

The delayed photon search presented in this thesis is also a result of a work from a great team that I am grateful to: Kevin, Si, Livia, Adi. Especially to Kevin: you were a great teammate, and the times working hard together with you day and night, both remotely and in the Wilson Hall offices, are unforgettable.

The writing of this thesis is also an effort of several people. Thank you Harvey for your tremendous amount of time spent on polishing the draft and making sure every sentence, every number, table, and plot in the thesis makes sense. Also thanks to Christina for being the first reader of the draft and being the language editor.

This thesis is dedicated to my dad and mom. You are awesome parents. Thank you for your love, and I love you!

ABSTRACT

This thesis presents the first observation of the production of three massive vector bosons (VVV with $V = W, Z$) in proton-proton collisions at the LHC at $\sqrt{s} = 13$ TeV. The search was performed in final states with two same-sign charged leptons (electrons or muons) plus one or two jets, and three, four, five, or six leptons from WWW, WWZ, WZZ and ZZZ decays, with a data sample corresponding to an integrated luminosity of $137\,\mathrm{fb}^{-1}$ collected by the CMS experiment during 2016-2018. The observed (expected) significance of the combined VVV production is 5.7 (5.9) standard deviations, and the production cross section is measured to be $1010^{+210}_{-200}\,(\mathrm{stat})^{+150}_{-120}(\mathrm{syst})$ fb, corresponding to a signal strength of $1.02^{+0.26}_{-0.23}$. We also found evidence for WWW and for WWZ production, with observed significances of 3.3 and 3.4 standard deviations, respectively. The measured production cross sections for individual VVV processes are also reported. The establishment of VVV production opens a new program of many standard model studies (such as gauge-gauge and Higgs-gauge couplings), and provides a new tool for many new physics searches (such as anomalous gauge coupling searches, new resonance searches). This thesis also reports a search for long-lived supersymmetry particles decaying to photons in proton-proton collisions at the LHC at $\sqrt{s} = 13$ TeV, with a data sample corresponding to an integrated luminosity of $77.4\,\mathrm{fb}^{-1}$ collected by CMS during 2016-2017. Results are interpreted with a gauge-mediated supersymmetry breaking model, where the long-lived particle is a neutralino and the lightest supersymmetry particle is a gravitino. For neutralino proper decay lengths of 0.1, 1, 10, and 100 m, masses up to 320, 525, 360, and 215 GeV are excluded at 95% confidence level, respectively. This result extends the limits from previous searches by one order of magnitude for the neutralino proper decay length and up to 100 GeV more for the neutralino mass. Motivated by the need for precision timing measurements for long-lived particle searches, as well as for improvements in general object reconstruction performance, the timing performance of two types of sensors are studied in this thesis: one with a Cerium doped Lutetium Yttrium Orthosilicate (LYSO:Ce) crystal as the scintillator and a silicon photomultiplier (SiPM) as the photodetector, another with a Cadmium-Telluride sensor as the active material for a sampling calorimeter. Both setups have been demonstrated with test beams to be able to provide timing measurements of particles with a resolution below 30 ps.

PUBLISHED CONTENT AND CONTRIBUTIONS

[1] CMS Collaboration. *Observation of the production of three massive gauge bosons at $\sqrt{s}$ = 13 TeV*. 2020. arXiv: 2006.11191 [hep-ex].
Z.Z. was one of the lead analysts for this search. Z.Z. initiated, developed, and optimized the BDT method for all final states; Z.Z. was in charge of the combination of all the final states and the statistical interpretation of the results; Z.Z. also participated in other parts of the analysis, such as background estimations, trigger and object ID studies, etc.

[2] CMS Collaboration. "Search for long-lived particles using delayed photons in proton-proton collisions at $\sqrt{s}$ = 13 TeV." In: *Physical Review D* 100.11 (Dec. 2019). ISSN: 2470-0029. DOI: 10.1103/physrevd.100.112003. arXiv: 1909.06166 [hep-ex]. URL: http://dx.doi.org/10.1103/PhysRevD.100.112003.
Z.Z. was the lead data analyst for the 2016 data in this paper. This paper was a combination of 2016 data and 2017 data. Z.Z. also lead the combination of results from two years and participated in preparing the manuscript for publication.

[3] A. Bornheim et al. "Precision timing detectors with cadmium-telluride sensor." In: *NIM-A* 867 (2017), pp. 32–39. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2017.04.024. URL: http://www.sciencedirect.com/science/article/pii/S0168900217304746.
Z.Z. participated in the test beam measurements, analyzed the data, and assisted in preparing the manuscript for publication.

[4] CMS Collaboration. "Search for supersymmetry with Higgs boson to diphoton decays using the razor variables at $\sqrt{s}$ = 13 TeV." In: *Physics Letters B* 779 (Apr. 2017), pp. 166–190. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2017.12.069. arXiv: 1709.00384 [hep-ex]. URL: http://dx.doi.org/10.1016/j.physletb.2017.12.069.
Z.Z. performed the Standard Model background modelling tests and determined the background functions.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

Part I

Introduction

Chapter 1

THE STANDARD MODEL OF PARTICLE PHYSICS

The standard model (SM) Lagrangian density is based on the $SU(3)_c \times SU(2)_L \times U(1)_Y$ (subscript c for color, L for the left-handed fermions, Y for the hypercharge operator) gauge theory:

$$\mathcal{L} = \bar{\psi}^i (i\gamma^\mu)(\mathcal{D}_\mu)_{ij}\psi^j - \frac{1}{4}G^a_{\mu\nu}G^{a\mu\nu} - \frac{1}{4}W^a_{\mu\nu}W^{a\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}$$
$$- m_f\bar{\psi}^i_f\psi_{fi} + \frac{1}{2}m^2_G G_\mu G^\mu + \frac{1}{2}m^2_W W_\mu W^\mu + \frac{1}{2}m^2_B B_\mu B^\mu, \qquad (1.1)$$

where:

- ψ is the fermion field;

- $G^a_{\mu\nu}, W^a_{\mu\nu}, B_{\mu\nu}$ are the field strength tensors for $SU(3)_c$, $SU(2)_L$, $U(1)_Y$, respectively.

- Index a is the index of the generators, which runs from 1 to 8 for the gluon, and 1 to 3 for W boson;

- m_f, m_G, m_W, m_B are the masses of the fermion, gluon, W boson, and neutral gauge boson;

- μ, ν are the Lorentz vector indices.

The first term $\bar{\psi}^i(i\gamma^\mu)(D_\mu)_{ij}\psi^j$ is the fermion kinetic term, where the covariant derivative is:

$$\mathcal{D}_\mu = \partial_\mu - ig'B_\mu Y - igW^a_\mu T^a - ig_s G^a_\mu t^a, \qquad (1.2)$$

and where g', g, g_s are the coupling strengths of the hypercharge interaction, weak interaction, and strong interaction ($g^2_s = 4\pi\alpha_s$); Y, T^a, t^a are the hypercharge operator, SU(2) generator, and SU(3) generator, respectively. T^a and t^a are proportional to the Pauli matrices σ^a and Gell-Mann matrices λ^a, respectively: $T^a = \sigma^a/2$, $t^a = \lambda^a/2$.

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad (1.3)$$

$$\lambda^1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda^2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda^4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

$$\lambda^5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \lambda^6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \lambda^7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \lambda^8 = \begin{pmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & \frac{-2}{\sqrt{3}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{3}} \end{pmatrix},$$

$$(1.4)$$

The field strength tensors in Eq. 1.1 are:

$$\begin{aligned}
B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu, \\
W^a_{\mu\nu} &= \partial_\mu W^a_\nu - \partial_\nu W^a_\mu + g\epsilon^{abc}W^b_\mu W^c_\nu, \\
G^a_{\mu\nu} &= \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + g_s f^{abc}G^b_\mu G^c_\nu,
\end{aligned}$$

$$(1.5)$$

where f^{abc} (a,b,c run from 1 to 8) and ϵ^{abc} (a,b,c run from 1 to 3) are the structure constants of SU(3) and SU(2), and they are related to the generators by:

$$[t^a, t^b] = if^{abc}t^c, \quad [T^a, T^b] = i\epsilon^{abc}T^c. \tag{1.6}$$

Expanding the $G^a_{\mu\nu}G^{a\mu\nu}$ and $W^a_{\mu\nu}W^{a\mu\nu}$ terms, we can see that there are three-gluon, four-gluon, three-gauge-boson, and four-gauge-boson vertices that involve the structure constants. The four-body vertices depend on the square of the structure constants (f^{abc} or ϵ^{abc}) and the square of the coupling strengths (g or g_s), due to the non-Abelian terms ($g\epsilon^{abc}W^b_\mu W^c_\nu$ and $g_s f^{abc}G^b_\mu G^c_\nu$) in the field strength tensors of the gluon and W bosons. Note that the neutral gauge bosons do not self-interact (through three-boson gauge or four-boson gauge couplings of the Z or the photon) in the SM. The experimental study of three- and four-boson gauge couplings is one of the main topics of this thesis (see Chapter 7).

The second line of Eq. 1.1 gives the mass terms for the fermions, gluons, and gauge bosons. Experimentally, it has been confirmed that gluons and photons are massless, while the fermions, W and Z bosons have precisely determined, nonzero masses. However, as illustrated below, it can be shown that the mass terms in the second line of Eq. 1.1 are not invariant under local gauge transformations, and thus they must be further considered before they can be included in the Lagrangian.

First of all, for the gauge sector, in order to preserve the gauge invariance of the first line of the Lagrangian in Eq. 1.1, the gauge transformations of the gauge fields need

to be written as:

$$B_\mu \rightarrow B_\mu + \frac{1}{g'}\partial_\mu \lambda_Y(x)$$

$$W_\mu^a \rightarrow W_\mu^a + \frac{1}{g}\partial_\mu \lambda_L^a(x) + \epsilon^{abc}W_\mu^b \lambda_L^c(x)$$

$$G_\mu^a \rightarrow G_\mu^a + \frac{1}{g_s}\partial_\mu \lambda_c^a(x) + f^{abc}G_\mu^b \lambda_c^c(x). \tag{1.7}$$

Under such gauge transformations, the mass terms $\frac{1}{2}m_G^2 G_\mu G^\mu$, $\frac{1}{2}m_W^2 W_\mu W^\mu$, and $\frac{1}{2}m_B^2 B_\mu B^\mu$ are no longer invariant.

For the electroweak interactions of fermions, the fermions are grouped into three families: (ν_e, e, u, d), (ν_μ, μ, c, s), (ν_τ, τ, t, b). For each of the three families, the left-handed fields are $SU(2)_L$ doublets, while the right-handed partners are $SU(2)_L$ singlets, and each of them has different hypercharges: 1/6 for $Q_L \equiv \begin{pmatrix} u_L \\ d_L \end{pmatrix}$, 2/3 for u_R, -1/3 for d_R, -1/2 for $L_L \equiv \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}$, and -1 for e_R. Under gauge transformations, each of them will be transformed by $\psi \rightarrow e^{i\lambda_Y(x)Y}\psi$, where Y is the hypercharge. The mass term of the fermion can be written as:

$$-m_f \bar{\psi}\psi = -m\bar{\psi}_R \psi_L - m\bar{\psi}_L \psi_R. \tag{1.8}$$

Under gauge transformation, the sum of the additional phases from transformed $\bar{\psi}_R$ and ψ_L will not be zero because the hypercharge Y for $\bar{\psi}_R$ and ψ_L is different, and therefore the mass terms given in Eq. 1.1 are not invariant under gauge transformations, and are therefore not allowed.

The inclusion of the mass terms without violating gauge invariance is accomplished through the use of spontaneous symmetry breaking, and the so-called "Higgs mechanism" as outlined below. We start with the Lagrangian density without the mass terms:

$$\mathcal{L} = \bar{\psi}^i(i\gamma^\mu)(\mathcal{D}_\mu)_{ij}\psi^j - \frac{1}{4}G_{\mu\nu}^a G^{a\mu\nu} - \frac{1}{4}W_{\mu\nu}^a W^{a\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}, \tag{1.9}$$

and to accommodate the nonzero observed masses of the Z, W, and fermions, we now introduce a new ingredient: a complex scalar field Φ which is an $SU(2)_L$ doublet and color singlet, with hypercharge $Y = 1/2$:

$$\Phi = \frac{1}{\sqrt{2}}\begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}, \tag{1.10}$$

where ϕ_1 and ϕ_2 are complex scalar fields. The potential of the field Φ is given by:

$$V(\Phi) = -\mu^2\Phi^\dagger\Phi + \lambda(\Phi^\dagger\Phi)^2, \tag{1.11}$$

where λ needs to be positive so that $V(\Phi)$ has minimum value. If $-\mu^2 > 0$, then $V(\Phi)$ will have a minimum of zero when $|\Phi| = 0$, which means the vacuum expectation value of Φ is zero, and in this case the gauge transformation is still invariant and there is still no mass term required for the W, Z and fermions. So $-\mu^2$ has to be negative, and in that case the potential $V(\Phi)$ has a minimum at $\Phi^\dagger\Phi = \frac{\mu^2}{2\lambda}$, and therefore the vacuum expectation value of the field Φ is not zero. Since the gauge transformation of Φ only changes the direction of the four component field Φ without changing the potential, we can choose a particular gauge transformation such that $\phi_1 = 0$ and ϕ_2 is a real scalar field (denoted as ϕ). In such gauge choice, the potential $V(\Phi)$ can be plotted in the real plane of Φ, as shown in Figure 1.1,

Under gauge transformation, Φ gets transformed to $e^{i\lambda_Y(x)Y}\Phi$ (where $Y = 1/2$ is the hypercharge), so the vacuum is not gauge invariant (even though the Lagrangian is still gauge invariant); in other words the gauge symmetry is *spontaneously broken* in the vacuum. Since the potential has a minimum at $|\Phi| = \frac{|\phi|}{\sqrt{2}} = \sqrt{\frac{\mu^2}{2\lambda}}$, we can rewrite the real scalar field ϕ as $\phi = v + h$, where $v = \frac{\mu^2}{\lambda}$, and h is a real scalar field with zero vacuum expectation value $\langle h \rangle = 0$. With this definition, the Lagrangian from the Φ field potential part is then given by:

$$\mathcal{L}_V = -V(\Phi) = \mu^2\Phi^\dagger\Phi - \lambda(\Phi^\dagger\Phi)^2 = -\lambda v^2 h^2 - \lambda v h^3 - \frac{\lambda}{4}h^4 + \text{const.} \tag{1.12}$$

where the first term is the mass term of the Higgs, $-\lambda v^2 h^2 = -m_h h^2/2$, and the second and third terms are the triple Higgs and quartic Higgs interaction vertices. Based on the first term of the equation:

$$m_h = \sqrt{2\lambda}v \tag{1.13}$$

where λ is an unknown parameter, and $v = 246\,\text{GeV}$ based on the value of Fermi coupling constant $v = 2m_W/g = 1/\sqrt{\sqrt{2}G_F^0}$ (we will see the relationship between W boson mass and v, g later in this section). Therefore, the mass of the Higgs boson is not predicted by this model and needs to be measured experimentally. The Higgs boson was first observed by CMS and ATLAS Collaborations in 2012 with a mass around 125 GeV [1, 2], and the latest and most precise measurement of m_h is $125.38 \pm 0.14\,\text{GeV}$ from the CMS Collaboration [3]. From the m_h measurement, the value of $\lambda = 0.129$ is inferred.
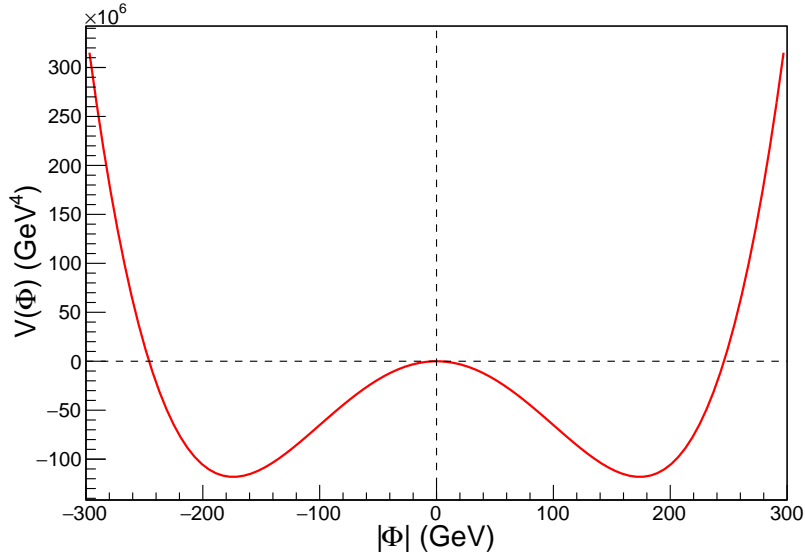
Figure 1.1: The potential of the Higgs field $V(\Phi) = -\mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2$ as a function of $|\Phi|$. The SM measured value of $-\mu^2 = (88.4\,\text{GeV})^2, \lambda = 0.129$ are used to make the plot. The minimum of the potential is at $|\Phi| = \pm \frac{\mu^2}{2\lambda} = \pm \frac{v}{\sqrt{2}} = \pm \frac{246}{\sqrt{2}}$ GeV.

Taking into account the Higgs-Higgs, Higgs-gauge and Higgs-fermion interactions, the standard model Lagrangian density now can be written as:

$$
\begin{aligned}
\mathcal{L} = &-\frac{1}{4} G^a_{\mu\nu} G^{a\mu\nu} - \frac{1}{4} W^a_{\mu\nu} W^{a\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} && \text{(gauge terms)} \\
&+ \bar{\psi}^i (i\gamma^\mu)(\mathcal{D}_\mu)_{ij} \psi^j && \text{(fermion kinetic)} \\
&+ \mu^2 \Phi^\dagger \Phi - \lambda (\Phi^\dagger \Phi)^2 && \text{(Higgs-Higgs)} \\
&+ (\mathcal{D}_\mu \Phi)^\dagger (\mathcal{D}^\mu \Phi) && \text{(Higgs-gauge)} \\
&- y_e \bar{L}_L \Phi e_R - y_u \bar{Q}_L \tilde{\Phi} u_R - y_d \bar{Q}_L \Phi d_R + \text{(h.c.)} && \text{(Higgs-fermion)}
\end{aligned}
\tag{1.14}
$$

For the Higgs-gauge term, since the Higgs field Φ is a color singlet and an SU(2) doublet, and has hypercharge $Y = 1/2$, the covariant derivative can be written as:

$$
\mathcal{D}_\mu = \partial_\mu - i\frac{g'}{2} B_\mu - i\frac{g}{2} W^a_\mu \sigma^a,
\tag{1.15}
$$

and the Higgs-gauge term can be written as:

$$(\mathcal{D}_\mu\Phi)^\dagger(\mathcal{D}^\mu\Phi) = \frac{1}{2} \left(\begin{array}{c} -\frac{i}{2}g(W_\mu^1 - iW_\mu^2)(v+h) \\ \partial_\mu h + \frac{i}{2}(gW_\mu^3 - g'B_\mu)(v+h) \end{array} \right)^\dagger \left(\begin{array}{c} -\frac{i}{2}g(W_\mu^1 - iW_\mu^2)(v+h) \\ \partial_\mu h + \frac{i}{2}(gW_\mu^3 - g'B_\mu)(v+h) \end{array} \right)$$

$$= \frac{1}{2}(\partial_\mu h)(\partial^\mu h) + \frac{1}{8}g^2(v+h)^2(W_\mu^1 - iW_\mu^2)(W^{1\mu} + iW^{2\mu})$$

$$+ \frac{1}{8}(v+h)^2\left(-g'B_\mu + gW_\mu^3\right)^2, \tag{1.16}$$

If we define the following four combinations (corresponding to the W^+, W^-, Z and photon fields):

$$W_\mu^+ = \frac{W_\mu^1 - iW_\mu^2}{\sqrt{2}},$$

$$W_\mu^- = \frac{W_\mu^1 + iW_\mu^2}{\sqrt{2}},$$

$$Z_\mu = \cos\theta_W W_\mu^3 - \sin\theta_W B_\mu,$$

$$A_\mu = \sin\theta_W W_\mu^3 + \cos\theta_W B_\mu \tag{1.17}$$

where θ_W is the weak mixing angle ($\tan\theta_W = \frac{g'}{g}$). We can rewrite the Higgs-gauge term as the following:

$$(\mathcal{D}_\mu\Phi)^\dagger(\mathcal{D}^\mu\Phi) = \frac{1}{2}(\partial_\mu h)(\partial^\mu h) \qquad \text{(Higgs kinetic)}$$

$$+ \frac{g^2v^2}{4}W_\mu^+W_\mu^- + \frac{(g^2 + g'^2)v^2}{8}Z_\mu Z^\mu \qquad \text{(W and Z mass term)}$$

$$+ \frac{g^2v}{2}hW_\mu^+W_\mu^- + \frac{g^2}{2}hhW_\mu^+W_\mu^- \qquad \text{(hWW and hhWW vertex)}$$

$$+ \frac{(g^2 + g'^2)v}{4}hZ_\mu Z^\mu + \frac{g^2 + g'^2}{8}hhZ_\mu Z^\mu \quad \text{(hZZ and hhZZ vertex)}$$

$$\tag{1.18}$$

from which we know that the masses of the W and Z bosons are:

$$m_W = \frac{gv}{2}, \qquad m_Z = \frac{v}{2}\sqrt{g^2 + g'^2} = m_W/\cos\theta_W. \tag{1.19}$$

The experimental study of the Higgs-gauge couplings by the production of three massive gauge bosons (in which two of them are produced from a Higgs boson decay) is also one of the main topics of this thesis (see Chapter 7).

The Higgs-fermion terms (Yukawa terms) in Eq. 1.14 include the fermion mass terms and Higgs-fermion coupling terms:

$$
\begin{aligned}
\mathcal{L}_{\text{Yukawa}} = &- \left(\frac{y_e v}{\sqrt{2}} \right) \bar{e}e - \frac{y_e}{\sqrt{2}} h\bar{e}e && \text{(lepton)} \\
&- \left(\frac{y_u v}{\sqrt{2}} \right) \bar{u}u - \frac{y_u}{\sqrt{2}} h\bar{u}u && \text{(up-type quark)} \\
&- \left(\frac{y_d v}{\sqrt{2}} \right) \bar{d}d - \frac{y_d}{\sqrt{2}} h\bar{d}d && \text{(down-type quark)}
\end{aligned}
\tag{1.20}
$$

from which we know that the mass of the leptons are:

$$
m_e = \frac{y_e v}{\sqrt{2}}, \qquad m_u = \frac{y_u v}{\sqrt{2}}, \qquad m_d = \frac{y_d v}{\sqrt{2}}.
\tag{1.21}
$$

In summary, the standard model (SM) of particle physics defines and classifies the elementary particles based on their quantum properties, and describes the weak, strong, and electromagnetic interactions among them. Figure 1.2 summarizes all the elementary particles and their classifications in the SM. So far, all the elementary particles in the SM have been experimentally observed and their properties and interactions have been well-tested and shown to agree within errors with the SM predictions. Figure 1.3 shows the current SM constraints on the Higgs boson mass (M_{H}) from various electroweak measurements, together with the direct measurement of M_{H}. We see that the direct measurement of M_{H} is compatible with the constraints derived from other electroweak measurements. More precise measurements and tests of the SM are continuously being conducted by particle physics experiments, in the quest for the first hints of the existence and nature of new physics beyond the standard model.
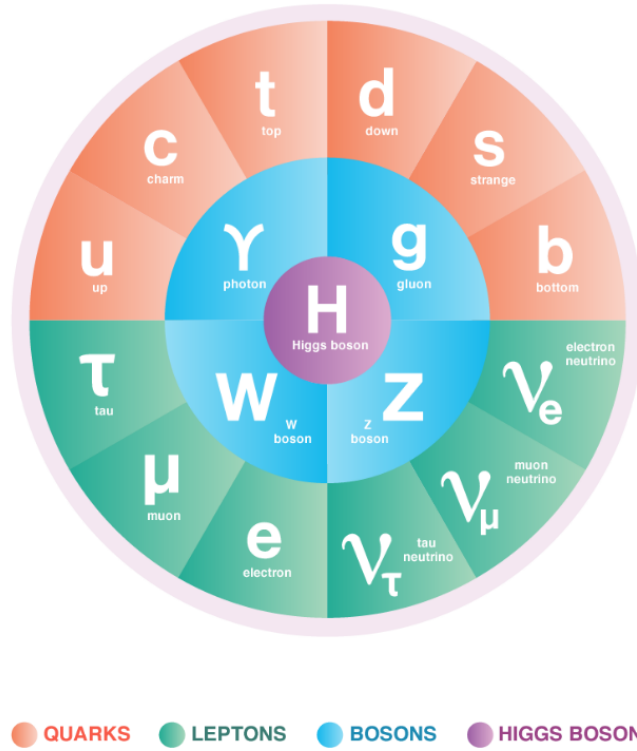
Figure 1.2: Elementary particles in the standard model: 12 fermions (quarks and leptons), 5 bosons (Higgs boson and gauge bosons). The plot is taken from [4].
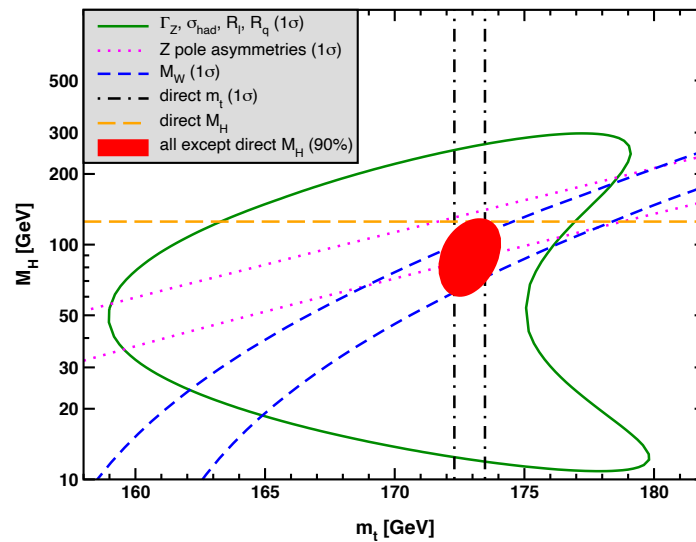
Figure 1.3: The direct measurements of Higgs boson mass (horizontal orange line, whose uncertainty is not visible in the current range) and top quark mass (vertical black lines), as well as the fit results of the Higgs boson mass as a function of top quark mass. The plot is taken from [5].

Chapter 2

SUPERSYMMETRY AND SEARCHES AT THE LHC

The Higgs boson part of the SM Lagrangian is:

$$\mathcal{L}_{\text{Higgs}} = \mu^2 \Phi^\dagger \Phi - \lambda (\Phi^\dagger \Phi)^2 + (\mathcal{D}_\mu \Phi)^\dagger (\mathcal{D}^\mu \Phi) + \mathcal{L}_{\text{Yukawa}}. \qquad (2.1)$$

The Higgs-fermion, Higgs-Higgs and Higgs-gauge interactions induce fermion/Higgs/gauge loop corrections to the Higgs free field propagator Φ and this gives radiative corrections to the Higgs mass (for the fermion loop, we only consider the top quark loop which is dominant):
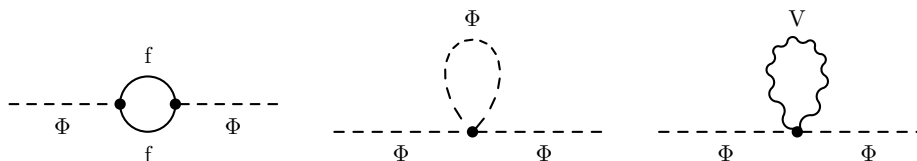
$$\delta m_h^2 \approx \frac{1}{8\pi^2} \left(-y_t^2 + 3\lambda + \frac{g^2}{4} + \frac{g^2}{4\cos^2\theta_W} \right) \Lambda_{\text{UV}}^2 \qquad (2.2)$$

where Λ_{UV} is the ultraviolet (UV) momentum cutoff of the divergent loop integrals, which is the energy scale above which the SM is no longer valid as an effective field theory. The fact that the measured Higgs boson mass is about 125 GeV and the coefficients (y_t, λ, g) in Eq. 2.2 are on the order of 1 implies that $\Lambda_{\text{UV}} \sim O(1\text{TeV})$ to make sure that $|\delta m_h^2| < m_h^2$. On the other hand, if the Λ_{UV} cutoff for the SM as an effective theory is on the order of the Planck scale, $M_P \sim 10^{19}\text{GeV}$, then $|\delta m_h^2|$ would be about 10^{36} times larger than the measured $m_h^2 \sim (125\,\text{GeV})^2$. This puzzle, known as the hierarchy problem, inevitably requires that new physics should exist at the TeV scale to alter the propagators in the loop correction and cut off the loop integral. Note that this problem only exists in the Higgs mass correction, as for the fermion and gauge boson masses, the radiative corrections do not have the direct quadratic dependence on Λ_{UV}^2, but instead they are logarithmic: $\delta m_{\text{f/W/Z}} \propto \log(\frac{\Lambda}{m_{\text{f/W/Z}}})$.

The search for such new physics at the TeV scale is (and will be) one of the main goals/motivations of multi-TeV colliders such as Large Hadron Collider and

future colliders with even higher energy. Besides the hierarchy problem, many other puzzles from experimental observations can only be resolved by considering theories beyond the SM; puzzles which include:

- The vacuum instability problem, which is related to the running of the Higgs quartic coupling λ to negative values. Given the current measured values of the Higgs mass and top quark mass [6], this occurs at a renormalization group energy scale well below the Planck scale M_P.

- The nature of dark matter, as observed in galactic rotation curves, the large scale structure of the universe, weak lensing, and other obervations.

- The origin of the neutrino masses observed in long baseline and reactor-based neutrino oscillation experiments.

- A quantum theory of gravity.

Now, suppose there is a heavy scalar particle S with the Higgs-S coupling term $-\frac{\lambda_S}{2} hhSS$ in the Lagrangian. Then the sum of the fermion loop and particle S loop contributions to the Higgs mass correction would be:
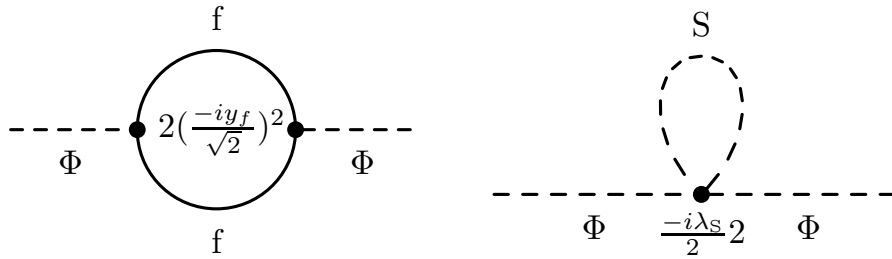
$$\delta m_h^2 \approx \frac{1}{16\pi^2} \left(-2y_f^2 + \lambda_S \right) \Lambda_{UV}^2. \tag{2.3}$$

If each of the SM fermions has two companion scalars each with $\lambda_S = y_f^2$, then the sum of the Λ_{UV}^2 term contribution in δm_h^2 from the fermion and the two scalars will be exactly zero. And similarly for gauge bosons and Higgs boson, we can cancel out the Λ_{UV}^2 term in δm_h^2 by their counterterms from their fermion companions (if they exist). To make such a theory a success, a new symmetry relating fermions and bosons, called *supersymmetry* (SUSY) [7], needs to be assumed. A supersymmetry

transformation generator Q that transforms a bosonic state into a fermionic state, and vice versa, must exist:

$$Q|\text{Boson}\rangle = |\text{Fermion}\rangle, \qquad Q|\text{Fermion}\rangle = |\text{Boson}\rangle. \tag{2.4}$$

The representations of the supersymmetry algebra are called supermultiplets. Each supermultiplet contains both fermion and boson states, which are known as the superpartners to each other, and each supermultiplet contains an equal number of fermion and boson degrees of freedom ($n_B = n_F$). The supermultiplets are classified based on the spin combinations of the fermion and scalar in the supermultiplet: a spin 0 and spin 1/2 combination is called a chiral supermultiplet, and a spin 1/2 and spin 1 combination is called a gauge supermultiplet. The possible supermultiplets which contain SM particles and their superpartner (called *sparticle*) are:

- One spin 1/2 fermion ($n_F = 2$), two spin 0 scalars ($n_B = 1 + 1 = 2$) which are called *sfermions* (including *squarks*, *sleptons*). This is a chiral supermultiplet.

- One spin 1 vector boson ($n_B = 2$), spin 1/2 fermions ($n_F = 2$) which are called *gauginos*. This is a gauge supermultiplet.

- For the Higgs boson, there are two supermultiplets, one with hypercharge $Y = 1/2$ (denoted as H_u) and one with hypercharge $Y = -1/2$ (denoted as H_d), and the third component of weak isospin $T_3 = (1/2, -1/2)$, which makes the charged and neutral components of H_u and H_d ($Q = T_3 + Y = -1, 0, 1$): $H_u = (H_u^+, H_u^0)$, $H_d = (H_d^-, H_d^0)$. The physical SM Higgs boson is a linear combination of H_u^0 and H_d^0. And the spin 1/2 superpartners of H_u and H_d are called *Higgsinos* ($\tilde{H}_u, \tilde{H}_d$). They are also chiral supermultiplets.

Among the gauginos, the superpartner of the gluon is called a *gluino* ($\tilde{g}$). The mixing of the gauginos $\tilde{W}^1, \tilde{W}^2, \tilde{W}^3, \tilde{B}$ give the mass eigenstates known as *winos* ($\tilde{W}^+, \tilde{W}^-$), the *zino* ($\tilde{Z}$) and the *photino* ($\tilde{A}$), just like the mixing among their superpartners (the SM particles) as shown in Eq. 1.17. All the supermultiplets mentioned above, which are composed of the SM particles and their superpartners, are the content of the most simple extension of the SM, called the Minimal Supersymmetric Standard Model (MSSM).

The MSSM also conserves the R-parity, which is defined based on the baryon number B, lepton number L, and spin of the particle:

$$P_R = (-1)^{3(B-L)+2s}. \tag{2.5}$$

With this definition, all SM particles have even R-parity ($P_R = 1$), and all the superpartners of the SM particles (sparticles) have odd R-parity ($P_R = -1$). The requirement of conserving R-parity implies that for any vertex that involves particles and sparticles, there can only be an even number of sparticles connecting to the vertex (even-in, even-out; or odd-in, odd-out), which means:

- From SM partons in a pp collision, the sparticles are always pair produced, for example pair-produced gluinos or squarks from a gluon.

- The sparticle must be either absolutely stable (no decay) or it must decay to SM particles plus an odd number of sparticles (usually just one).

- The lightest sparticle, called the lightest supersymmetric particle, or LSP, must be absolutely stable (no decay).

Just like electroweak symmetry, which has to be broken in the SM, supersymmetry also needs to be a broken symmetry in the vacuum state. The reason is that if supersymmetry were unbroken, we would have already easily observed the superpartners of SM particles which have the same mass as the SM particles, for example 0.511 MeV selectrons $\tilde{e}_L$ and $\tilde{e}_R$ (the superpartners of $L_L \equiv \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}$ and e_R), the massless gluino and photino. The fact that the unbroken supersymmetry already cancels the Λ_{UV}^2 terms in δm_h^2 through appropriate choices of the dimensionless coupling parameters (y_t, λ, g) requires that the Lagrangian term that violates supersymmetry can no longer have dimensionless coupling parameters. In other words, the effective Lagrangian of the MSSM can then be written as:

$$\mathcal{L} = \mathcal{L}_{\mathrm{SUSY}} + \mathcal{L}_{\mathrm{soft}} \tag{2.6}$$

where $\mathcal{L}_{\mathrm{SUSY}}$ contains all the gauge, Yukawa, and dimensionless scalar couplings, and does not break supersymmetry; and $\mathcal{L}_{\mathrm{soft}}$ violates supersymmetry and only contains couplings parameters and parameters in mass terms that are dimensionful with positive dimensions. This requirement is called soft supersymmetry breaking (for a review see [8]).

Similarly to the SM Lagrangian in Eq. 1.14, the Lagrangian density of a supersymmetric theory can be written as the following terms with all the free supermultiplets

and their interactions included:

$$\begin{aligned}
\mathcal{L}_{\text{SUSY}} = & -\frac{1}{4} F^a_{\mu\nu} F^{a\mu\nu} + i\lambda^{\dagger a} \bar{\sigma}^\mu \nabla_\mu \lambda^a + \frac{1}{2} D^a D^a && \text{(gauge term)} \\
& - \partial^\mu \phi^{*i} \partial_\mu \phi_i + i\psi^{\dagger i} \bar{\sigma}^\mu \partial_\mu \psi_i && \text{(chiral term)} \\
& - \frac{1}{2} y^{ijk} \phi_i \psi_j \psi_k - \frac{1}{2} y^*_{ijk} \phi^{*i} \psi^{\dagger j} \psi^{\dagger k} && \text{(chiral interaction)} \\
& - \frac{1}{2} M^{ij} \psi_i \psi_j - \frac{1}{2} M^*_{ij} \psi^{\dagger i} \psi^{\dagger j} - V(\phi, \phi^*) && \text{(chiral interaction)} \\
& - \sqrt{2} g (\phi^* T^a \psi) \lambda^a - \sqrt{2} g \lambda^{\dagger a} (\psi^\dagger T^a \phi) + g(\phi^* T^a \phi) D^a && \text{(gauge interaction)}
\end{aligned}$$

(2.7)

where:

- $F^a_{\mu\nu}$: the field strength tensors for the gauge bosons, as defined in Eq. 1.5.

- λ^a: the two-component fermion gaugino (superpartners of the gauge bosons).

- D^a: the real bosonic auxiliary field for the gauge boson to match the degree of freedom of the gauge boson and the gaugino.

- ϕ: the scalar field in a chiral supermultiplet, such as Higgs scalars, sleptons.

- ψ: the fermion field in a chiral supermultiplet, such as Higgsino fermion, leptons. The chiral supermultiplet term (second line) in Eq. 2.7 describes the Lagrangian of the free chiral supermultiplets.

- y^{ijk}: the Yukawa coupling of a scalar ϕ_i and two fermions $\psi_j \psi_k$.

- M_{ij}: symmetric mass matrix for the fermions ψ.

- $V(\phi, \phi^*)$: potential of the scalar ϕ, $V(\phi, \phi^*) = F^{*k} F_k$, where F is an auxiliary scalar field, $F_i = -M^*_{ij} \phi^{*j} - \frac{1}{2} y^*_{ijk} \phi^{*j} \phi^{*k}$.

- The last line of Eq. 2.7 contains the interaction of scalar, fermion, gaugino, and auxiliary fields between gauge supermultiplet and chiral supermultiplet.

The supersymmetry breaking part of the Lagrangian, $\mathcal{L}_{\text{soft}}$, in the effective MSSM Lagrangian, can be written as:

$$\mathcal{L}_{\text{soft}} = -\left(\frac{1}{2} M_a \lambda^a \lambda^a + \frac{1}{6} a^{ijk} \phi_i \phi_j \phi_k + \frac{1}{2} b^{ij} \phi_i \phi_j + t^i \phi_i \right) + \text{h.c.} - (m^2)^i_j \phi^{j*} \phi_i,$$

(2.8)

where all the parameters have positive dimensions:

- M_a: the gaugino masses, dimension 1.

- $(m^2)^i_j$: scalar mass terms, dimension 2.

- a^{ijk}: three-scalar couplings, dimension 1.

- b^{ij}: two-scalar couplings, dimension 2.

- t^i: "tadpole" couplings for scalar[1], dimension 3.

Different mechanisms have been proposed to try to explain the spontaneous super-symmetry breaking by constructing non-zero vacuum expectation values from the MSSM supermultiplets, for example the Fayet-Iliopoulos mechanism [10] (which uses the auxiliary D^a field in the gauge supermultiplet, called a D-term), or the O'Raifeartaigh models [11] (which use the auxiliary F_i field in the chiral supermul-tiplet, called an F-term). However, all of them have proven to be insufficient to be responsible for the dominant source for supersymmetry breaking [7]. As a result, an extension of the MSSM with a separate supersymmetry breaking sector (called the hidden sector) is needed, in which the spontaneous supersymmetry breaking occurs in the hidden sector, and the supersymmetry breaking effects are mediated from the hidden sector to the visible sector (supermultiplets of MSSM) such that the masses of the sparticles in the visible sector are large enough that we still have not observed them. There are two main models of such mediating interactions, one is gravity-mediated or Planck-scale mediated supersymmetry breaking (PMSB) [12–15], and the other one is gauge-mediated supersymmetry breaking (GMSB) [16–25].

For PMSB, the supersymmetry breaking is mediated from the hidden sector to the MSSM by new interactions (including gravity) that happen at the Planck scale M_P. When taking gravity into account, we get another supermultiplet which has the spin 2 graviton and a spin 3/2 fermion called the *gravitino* ($\widetilde{G}$). If supersymmetry is broken in the hidden sector by a vacuum expectation value $\langle F \rangle$, then the soft terms in MSSM are about:

$$m_{\text{soft}} \sim \frac{\langle F \rangle}{M_P} \tag{2.9}$$

---

[1]For a phenomenology of models where the electroweak symmetry breaking is driven by a tadpole, and how the Higgs potential can be destabilized by a tadpole, see [9]

and in an effective field theory, the parameters in the soft term of the Lagrangian in Eq. 2.8 are:

$$M_a = \frac{f^a \langle F \rangle}{M_P}$$

$$(m^2)^i_j = \frac{k^j_i |\langle F \rangle|^2}{M_P^2}$$

$$a^{ijk} = \frac{\alpha^{ijk} \langle F \rangle}{M_P}$$

$$b^{ij} = \frac{\beta^{ij} \langle F \rangle}{M_P} \tag{2.10}$$

where f^a, k^j_i, α^{ijk} are dimensionless couplings, and β^{ij} has dimension 1. In PMSB, the gravitino mass $(m_{3/2})$ is also proportional to $\frac{\langle F \rangle}{M_P}$, so it is comparable to the masses of the MSSM sparticles (of order 100 GeV or so).

For GMSB, the ordinary gauge interactions are responsible for the origin of supersymmetry breaking in the hidden sector, and another sector called the messenger sector is responsible for the coupling to both the hidden sector and the visible sector. The messenger sector has chiral supermultiplets called messengers, which contain messenger fermion quarks ($\psi_q, \psi_{\bar{q}}$), scalar quarks ($q, \bar{q}$), fermion leptons ($\psi_l, \psi_{\bar{l}}$), scalar leptons ($l, \bar{l}$), and all of them have very large masses (M_{mess}), at the so-called messenger scale. Scalar quark messenger loops and scalar lepton messenger loops give masses to the gauginos, which are proportional to $\langle F \rangle / M_{\text{mess}}$. The masses of the scalars in the MSSM are obtained by two-loop corrections from the messenger fermions and messenger scalars, and the scalar masses are also proportional to $\langle F \rangle / M_{\text{mess}}$. So the masses of the gauginos and the scalars in the MSSM are close to each other.

In GMSB, the gravitino mass is much lighter than the sparticles in the MSSM (order of eV compared to GeV) and it is the LSP in GMSB, and all MSSM sparticles will eventually decay into a gravitino plus SM particles. The rate of a sparticle $\tilde{X}$ decaying to an SM particle X plus a gravitino is:

$$\Gamma(\tilde{X} \rightarrow X\tilde{G}) = \frac{m_{\tilde{X}}^5}{16\pi \langle F \rangle^2} \left(1 - m_X^2 / m_{\tilde{X}}^2 \right)^4 . \tag{2.11}$$

Depending on the mass spectrum involved in this decay and $\langle F \rangle$, the decay rate could be very small which leads to long-lived sparticles $\tilde{X}$. The experimental search for such long-lived sparticles within the context of a GMSB model is one of the main topics of this thesis (see Chapters 8, 9, and 10).

After supersymmetry breaking, the superpartners in the supermultiplets discussed above might no longer be mass eigenstates. For example, the mixing now can happen between gauginos and higgsinos that have the same electric charge due to both electroweak symmetry breaking and supersymmetry breaking effects. The mix of neutral gauginos and neutral higgsinos form the *neutralinos* ($\widetilde{\chi}_1^0, \widetilde{\chi}_2^0, \widetilde{\chi}_3^0, \widetilde{\chi}_4^0$); the mix of charged gauginos and charged higgsinos form the *charginos* ($\widetilde{\chi}_1^+, \widetilde{\chi}_1^-, \widetilde{\chi}_2^+, \widetilde{\chi}_2^-$).

At the Large Hadron Collider, the dominant SUSY production modes are gluino-squark, gluino-gluino, and squark-squark pair productions from gluon-gluon and gluon-quark fusions. The production cross section of different SUSY particle pairs at the LHC can be seen in Figure 2.1. As we can see, the production cross sections for different SUSY particles vary by several orders of magnitude, which makes the production of gluinos and squarks the most experimentally sensitive channels at the LHC.
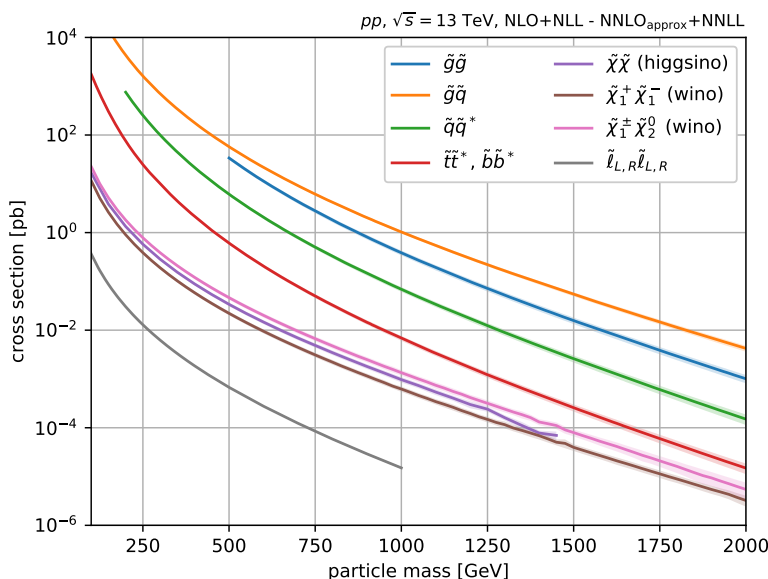
Figure 2.1: Production cross section of SUSY particles at the LHC for proton-proton collisions of $\sqrt{s} = 13$ TeV as a function of the mass of SUSY particles at the next-to-leading logarithmic accuracy (NLL) matched to next-to-leading order (NLO) predictions for weak production, and next-to-next-to-leading logarithmic accuracy (NNLL) matched to approximate next-to-next-to-leading order (NNLO$_{\text{approx}}$) predictions for strong production. The plot is taken from [26].

Different models (PMSB, GMSB, etc.) have different predictions for the sparticle masses and decay modes, depending on the values of the free parameters in such

models that have not yet been measured experimentally. When supersymmetry searches are performed, different searches usually use benchmark points in a specific model with some of the free parameters fixed and some free parameters floating to be constrained experimentally. With the data collected in 2015-2018 by the CMS and ATLAS experiments during the LHC Run2, the mass of the gluino has been excluded up to about 2.2 TeV for most decay channels, and the mass of squarks has been excluded up to about 1.7 TeV [27, 28], corresponding to an upper limit in the cross sections on the order of 1 fb, as shown in Figure 2.1.

Part II

The LHC and CMS

Chapter 3

THE LARGE HADRON COLLIDER

The Large Hadron Collider (LHC) [29] is a machine which accelerates, circulates, and focuses two counter-rotating bunched hadron (proton or heavy ion) beams in a 26.659 km circular main tunnel 50 to 150 meters below the ground spanning the border between Switzerland and France, and collides the bunches at four interaction points (IP) where the four major LHC experiments are installed: ATLAS (IP1), ALICE (IP2), CMS (IP5), and LHCb (IP8).

The LHC was built by re-using the Large Electron-Positron Collider (LEP) tunnel, which was dug in 1984-88 and which has an internal diameter of 3.7 meters. The LEP tunnel was originally foreseen to have a 30 km circumference, but civil engineering test borings showed that the tunnel would have to go too deep into the granite below the Jura mountains, where there are pockets of water under high pressure. For this reason, the circumference was reduced, the tunnel was moved to minimize the part of the circumference (less than 2%) in the granite, and it was tilted by 1.5% to fit between the foot of the mountains and the Geneva airport.

The main goal of the LHC is to search for new physics beyond the standard model (BSM) with proton-proton (pp) or heavy ion (PbPb, pPb, or XeXe) collisions of center-of-mass energies ($\sqrt{s}$) of up to 14 TeV for the proton-proton runs and typical energies of around 5 TeV during heavy ion runs. A second goal is to precisely measure many of the strong and electroweak standard model processes so as to detect signs of BSM processes, or further constrain the yet-to-be-discovered new physics that may exist. The heavy ion runs have provided new information on states of nuclear matter including the quark gluon plasma under extreme pressure and temperature.

Because the two proton beams have the same electric charge (anti-proton beams are not possible given the LHC's high beam intensity requirement), separate beam pipes and sets of coils are needed in the tunnel in order to produce the magnetic field that bends the two beams in opposite directions. Due to the limited space in the tunnel, it would have been extremely difficult to install two completely separate rings of dipole magnets to guide the counter-rotating beams. Instead, a twin-bore magnet design [30] was adopted, in which the two beam pipes and two sets of coils

are installed within the same mechanical structure and share the same cryostat, and thus the two beam channels are magnetically and mechanically coupled.

There are two main types of magnets in the LHC: 1232 dipole magnets for bending the beams and 392 quadrupole magnets for focusing the beams. There are additional pole windings to correct for higher order multipoles in the magnet lattice and specialized sextupole and octupole magnets to help control the beam orbit and focusing. The superconducting magnets are made of Nb-Ti cables, cooled by superfluid helium at a temperature of 1.9K, which provides a nominal magnetic field of 8.33 T.

Figure 3.1 shows the schematic layout of the LHC. The LHC tunnel is not a perfect circle: it has eight 528-m-long straight sections and eight 2.45-km-long arcs. A straight section plus two transition regions (dispersion suppressors) is defined as one insertion. The region between two insertion points is defined as a sector. The region between the middle of the arc to the middle of the next arc is defined as an octant. Insertion points 1, 2, 5, and 8 are where the four main experiments are located: ATLAS, ALICE, CMS, and LHCb. The two beams, Beam 1 and Beam 2, are injected from insertion point 2 and insertion point 8, with Beam 1 rotating clockwise and Beam 2 rotating anticlockwise.

The injected 450 GeV-proton beams are produced by stripping the electrons off of hydrogen gas atoms, and then accelerated in steps in the various accelerator systems of CERN [31]: Linac2 (50 MeV; after Run 2 Linac2 has been replaced by Linac4, which will accelerate the protons to 160 MeV), the Proton Synchrotron Booster (PSB, 11.4 GeV), the Proton Synchrotron (PS, 25 GeV), and the Super Proton Synchrotron (SPS, 450 GeV). Finally, the 450 GeV proton beams get injected into the LHC at the two insertion points. Upon injection into the LHC, the beams are captured and accelerated by the superconducting radio-frequency cavity system.

Each proton gains 485 keV energy during one turn through the LHC lattice, which means that 1.35×10^7 turns are needed to ramp up the energy from 450 GeV to 7 TeV, corresponding to a typical ramp up time of 20 minutes. The collisions between the bunches in the two beams happen at the four IPs, in which the two beam pipes merge into one common beam pipe of about 140 m (126 m) long for IP1/5 (IP2/8). To avoid bunch crossings at multiple points in the same beam pipe (in total 37 possible crossing points in a 140 m long beam pipe), the two beams are collided with each other at the designed point with a small crossing angle (typically a few hundred μrad). After a certain number of collisions, the two beams are dumped by

the beam dump system in insertion point 6.
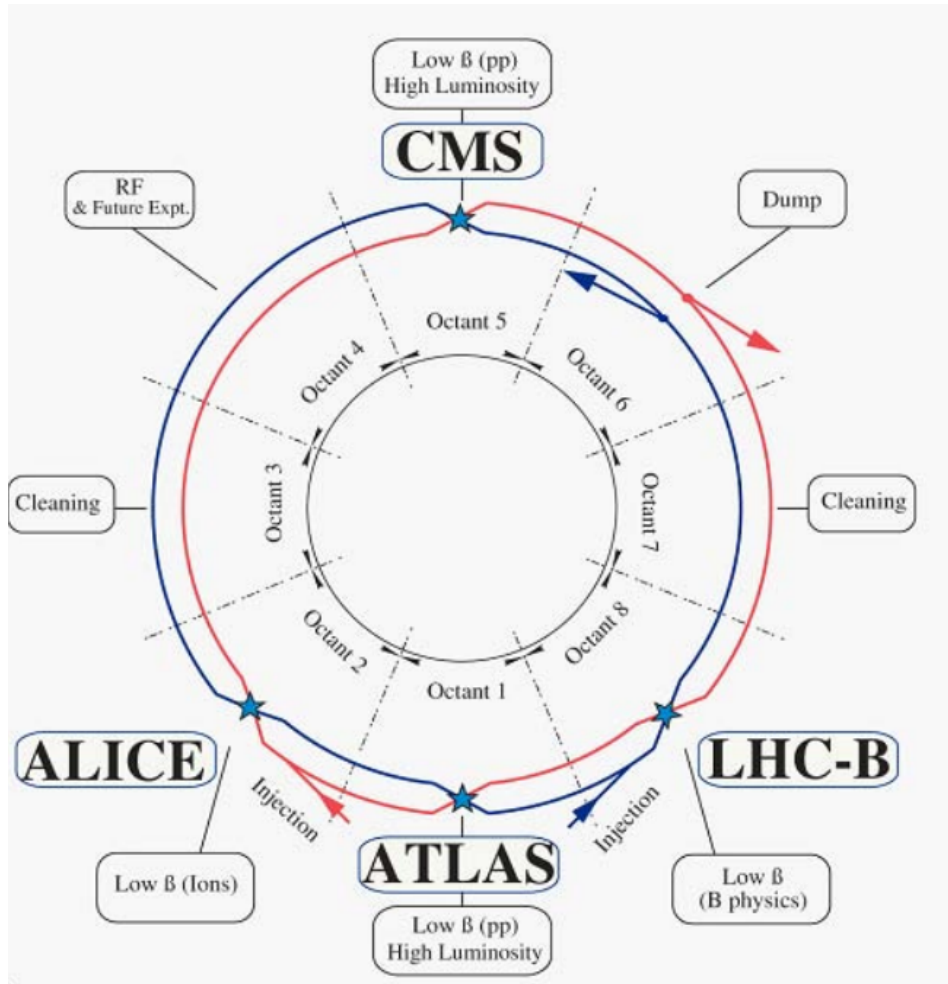
Figure 3.1: The schematic layout of the LHC; Beam 1 rotates clockwise and Beam 2 rotates anticlockwise [29].

The event production rate from the collisions depends on the machine luminosity and the production cross section: rate $= L\sigma_{xsec}$, where σ_{xsec} is the production cross section of the physics process under study, which depends on the center of mass energy $\sqrt{s}$. The luminosity L can be written (approximately) in terms of the main beam parameters as follows:

$$L = \frac{N_b^2}{4\pi\sigma^2 t} \tag{3.1}$$

where (from the design reports [29, 32]):

- N_b: nominal number of protons per bunch: 1.15×10^{11}.

- t: time between bunches: 25 ns.

- σ: nominal transverse size of the bunch at IP: 16.7 μm.

Using the numbers above, we get $L \approx 1.5 \times 10^{34} \, \text{cm}^{-2} \, \text{s}^{-1}$.

More precisely, with some corrections and expressed in more general parameters, the luminosity can be also written as:

$$L = \frac{N_b^2 n_b f_{\text{rev}} \gamma_r}{4\pi \epsilon_n \beta^*} F; \quad F = \frac{1}{\sqrt{1 + (\frac{\theta_c \sigma_z}{2\sigma})^2}} \tag{3.2}$$

where the LHC design values of the parameters in this formula are (from [29, 32]):

- N_b: number of protons per bunch: 1.15×10^{11}.

- n_b: number of bunches per beam: the maximum possible number is determined by the length of the LHC tunnel and the distance between two neighboring bunches: 26.659km/7.5 m = 3554; the design number is 2808 (allowing for some empty timeslots among the bunches).

- f_{rev}: revolution frequency: $f_{\text{rev}} = 1/T = \frac{299792 \text{km/s}}{26.659 \text{km}} = 11245$ Hz (rings/second). The product of $n_b f_{\text{rev}} = 31.6$MHz is the average bunch crossing rate.

- γ_r: relativistic gamma factor: 7461 for 7 TeV protons.

- ϵ_n: normalized (w.r.t the beam energy) transverse beam emittance: 3.75 μm.

- β^*: beta function at the collision point: 0.55 m.

- F: geometric luminosity reduction factor due to non-zero crossing angle: depends on the crossing angle at the IP, θ_c (285μrad), the bunch length, σ_z (7.55 cm), and the transverse beam size, σ (16.7μm). This parameter is about 0.836.

Using the numbers above, we get $L = 1.0 \times 10^{34} \, \text{cm}^{-2} \, \text{s}^{-1}$, which is the design luminosity of the LHC.

During actual operation of the LHC in Run 2, several operational parameters have been significantly improved relative to the design values, as a result of the large effective aperture, quality and remarkable reproducibility of the LHC magnet lattice, the growing operational experience and expertise of the LHC team, and the discovery

of new injection and focusing schemes. This led to greater maximum bunch intensity in 2016, and smaller emittance and overall a peak luminosity that reached $L = 2 \times 10^{34} \, \text{cm}^{-2} \, \text{s}^{-1}$ repeatedly at the beginning of a fill during 2017 and 2018. As discussed further below, the maximum luminosity is set by the experiments' ability to deal with a certain maximum number of "pileup" events accompanying an event of interest during the same crossing of the proton bunches. A comparison of the nominal LHC design parameters and the typical actual parameters during Run 2 operation is given in Table 3.1.

Following the next data taking run, Run 3 scheduled for 2022-24, the LHC accelerator and detectors will complete the installation and commissioning of major upgrades in 2025-26. A new phase of the program called the High-Luminosity LHC (HL-LHC) is scheduled to start in 2028 and continue for approximately 10 years, through 2038. The HL-LHC accelerator has a design luminosity of approximately $L = 6.5 \times 10^{34} \, \text{cm}^{-2} \, \text{s}^{-1}$. These upgrades will significantly boost the discovery reach and rate of scientific progress, and reduce the time needed to reach the ultimate integrated luminosity goal (about 3000-4000 fb^{-1}). The comparison of the machine parameters in LHC and HL-LHC can also be found in Table 3.1.

The HL-LHC program will enable the study of a wide variety of production and decay channels, including rare decays of the Higgs boson, vector boson scattering, di-Higgs production to provide key information on the Higgs boson self-coupling, and rare b-meson decays with great precision, enabling both direct and indirect searches for BSM physics across a vast landscape.

A summary of the operational experience during LHC fills during Run 2 in 2015-18 follows. Once the beams were brought into collision, the luminosity during a fill dropped as N_b was reduced due to collisions. The total cross section for pp collisions is about $\sigma_{\text{tot}} = 110$ mb for $\sqrt{s} = 13$ TeV, with small variations for different center-of-mass energies [33, 34]. This means that the typical reduction rate of the number of protons in the beam in the midst of a fill ($N_b n_b$) was roughly $\sigma_{\text{tot}} L = 110 \, \text{mb} \times 1.4 \times 10^{34} \, \text{cm}^{-2} \, \text{s}^{-1} = 1.5 \times 10^9 \, \text{s}^{-1}$ due to collisions at each IP. The sum of the loss-rates in the two high luminosity experiments CMS and ATLAS was thus about $dN/dt = 3 \times 10^9 \text{s}^{-1}$, or 1.1×10^{13} per hour. This means that the mean lifetime of the total number of protons $N_b n_b$ (or the mean lifetime of the number of protons in each bunch N_b) was about $\tau_N = \frac{N}{dN/dt} = 1.25 \times 10^{11} \times 2556/(1.1 \times 10^{13} \, \text{h}^{-1}) = 29$ hours. This gave a typical mean lifetime of the luminosity ($\tau_L \approx \tau_N/2$) of about 14 hours based on Eq. 3.2.

In practice, some of the fills were cut short due to a beam loss. On the other hand, the luminosity lifetime went up at the end of the fill, so long fills were sometimes allowed to last longer, including when any essential component in the LHC injector chain went down for some hours. In addition, the β^* was sometimes reduced on the fly during a fill, and that fill was kept longer. Without the collisions, the typical beam lifetimes were above 100 hours.

To achieve the maximum integrated luminosity, the beams were dumped when the luminosity was low enough. The optimal time length of each fill (t_{fill}) depends on the turnaround time t_T [35]: from the instant data taking is stopped during a fill, to the instant when data taking is resumed at the start of the next fill. t_T is thus the sum of the beam dump time and the preparation time (time that is needed for LHC magnet cycling and optics setup, injecting the beam, and ramping up the beam energy) to bring the beams back into collision again at full energy.

Taking an average turnaround time t_T of 6 hours and t_{fill} of 8 hours (as in 2018), the average luminosity was about $0.68 \times 10^{34} \, \text{cm}^{-2} \, \text{s}^{-1}$ (averaged over t_T and t_{fill}, $L_{ave} = \int_{t_{\text{fill}}} L dt / (t_T + t_{\text{fill}})$).Considering that the typical average availability of the LHC was about 79% (the percentage of time without hardware failures), the integrated luminosity was about $0.46 \, \text{fb}^{-1}$ per day. During 2018, there were 145 days for physics runs, which yields an integrated luminosity per year in our simplified calculation of around $70 \, \text{fb}^{-1}$. In reality, the integrated luminosity for pp collisions at $\sqrt{s} = 13$ TeV delivered by the LHC to CMS was about $41.0 \, \text{fb}^{-1}$, $49.8 \, \text{fb}^{-1}$, and $67.9 \, \text{fb}^{-1}$ during 2016, 2017 and 2018, respectively.

The integrated luminosity recorded by CMS and good for physics analysis (data collected by CMS with all essential sub-detectors working) was about $35.9 \, \text{fb}^{-1}$, $41.5 \, \text{fb}^{-1}$, and $59.7 \, \text{fb}^{-1}$ during 2016, 2017, and 2018, respectively. The LHC also delivered and CMS recorded about $4 \, \text{fb}^{-1}$ of data during 2015, but this small piece of data is not considered in the physics analysis in this thesis, as the cost of adding in the 2015 data is too high compared to the benefit. We would need separate simulations for the signal and background events, as well as dedicated calibrations, scale factors, and detector performance measurements. As a result, the computing resources and time required to analyze the 2015 data would be almost the same as for analyzing the data in other years, while the impact on the overall Run 2 data sample would not be statistically significant.

The actual day by day peak luminosity and the cumulative luminosity delivered to CMS during Run 2 can be seen in Figure 3.2.

CMS Peak Luminosity Per Day, pp, $\sqrt{s}$ = 13 TeV

Data included from 2015-06-03 08:41 to 2018-10-26 08:23 UTC

CMS Integrated Luminosity Delivered, pp, $\sqrt{s}$ = 13 TeV

Data included from 2015-06-03 08:41 to 2018-10-26 08:23 UTC
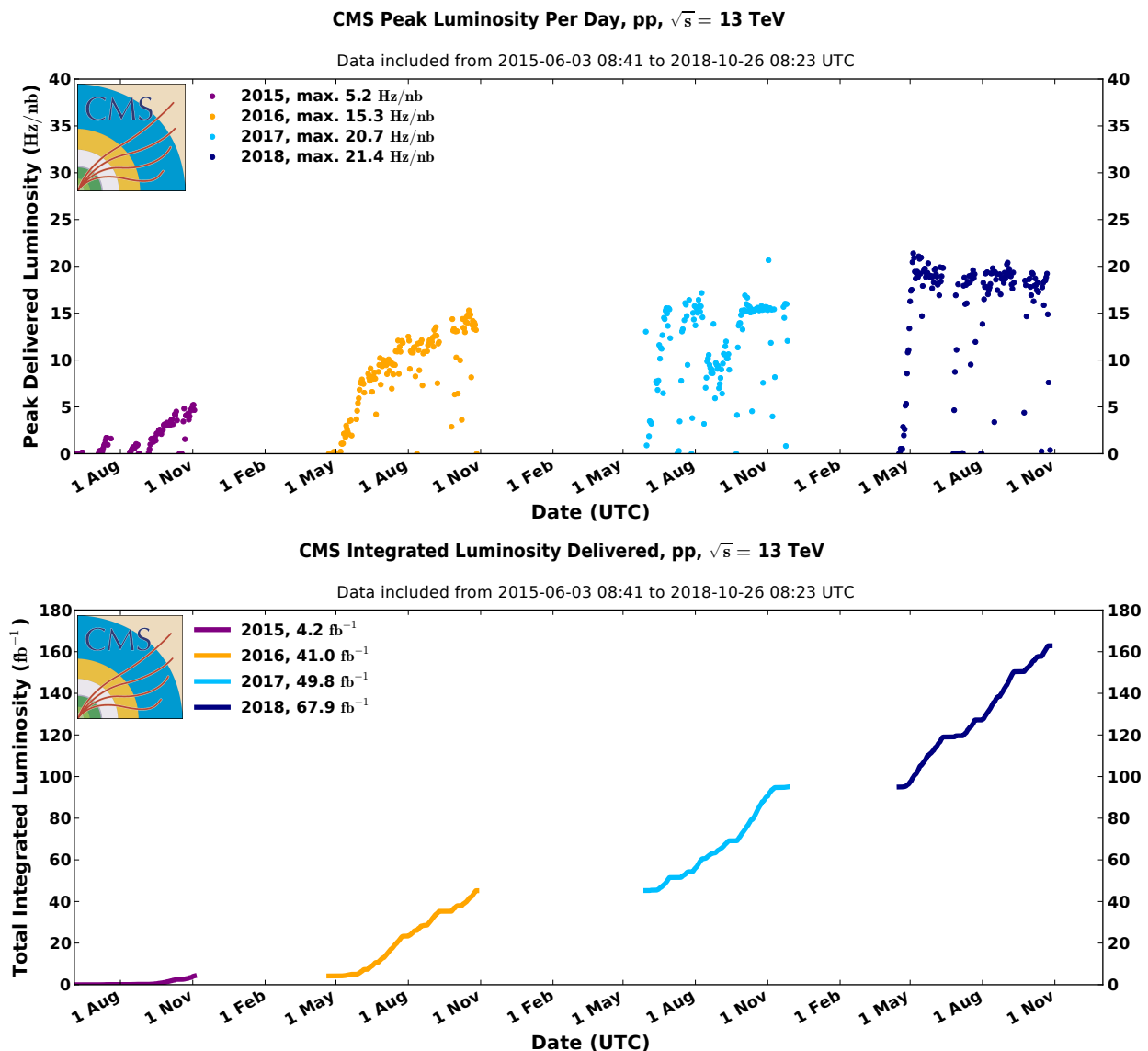
Figure 3.2: Upper (bottom) plot: peak luminosity (cumulative luminosity) versus day delivered to CMS for pp collisions during Run 2 [36].

The number of simultaneous pp interactions occurring in the same bunch crossing, defined as the pileup, can be estimated based on the inelastic pp cross section σ_{inel}, which is 79.5 mb [34] at 13 TeV. With this cross section, the number of pp interactions per second at each IP for the 2018 peak luminosity ($2.1 \times 10^{34} \, \text{cm}^{-2} \, \text{s}^{-1}$) would be about $\sigma_{\text{inel}} \, L = 79.5 \, \text{mb} \times 2.1 \times 10^{34} \, \text{cm}^{-2} \, \text{s}^{-1} = 16.7 \times 10^8 \, \text{s}^{-1}$, and the average bunch crossing rate would be approximately 28.7 MHz (for 2556 bunches

in 2018), so the number of pp interactions for each crossing would be about

$$N_{\text{pileup}} = \frac{\sigma_{\text{inel}}L}{\text{crossing rate}} = \frac{16.7 \times 10^8 \text{s}^{-1}}{28.7 \text{MHz}} = 58,$$

(3.3)

which is the estimated peak pileup for the 2018 data taking period.

Alternatively, the number of pp interactions in one crossing can also be calculated by the basic formula:

$$N_{\text{pileup}} = \frac{N_b^2 \sigma_{\text{inel}}}{4\pi \sigma_{\text{beam size}}^2} = \frac{(1.25 \times 10^{11})^2 \times 79.5 \text{mb}}{4\pi \times (11\mu m)^2} = 82.$$

(3.4)

With a correction by the geometric reduction factor F of about 0.7, we also get about 58 peak pileup for the 2018 data taking period.

In reality, the average crossing rate and luminosity changed during the different years of Run 2 running, with variations in the bunch intensity and/or the beam sizes at the interaction point, so the peak pileup also changed. The actual crossing rate, peak luminosity and also the peak pileup for different years of LHC Run 2 can be found in Table 3.1. The distribution of the pileup during each year of Run 2 running measured by CMS is shown in Figure 3.3.
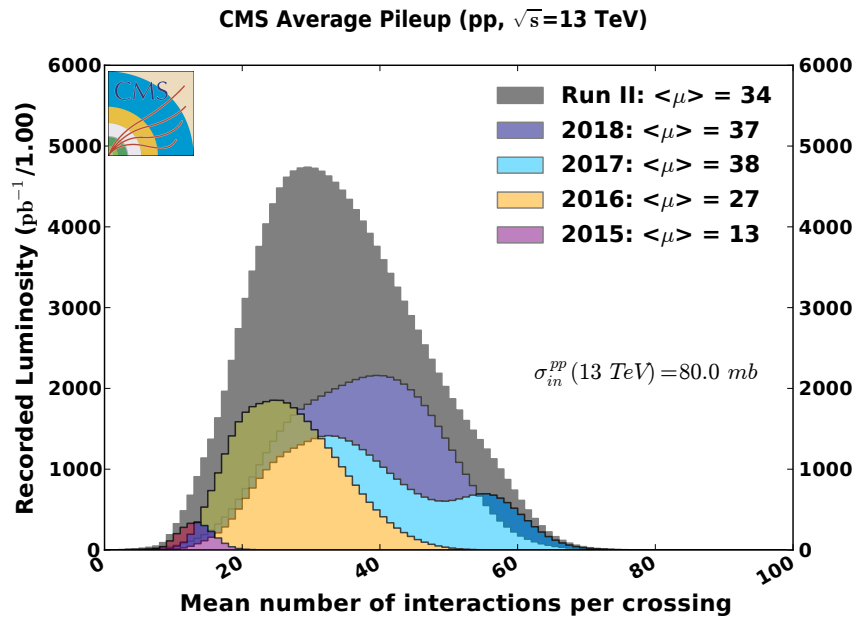
Figure 3.3: Pileup (interactions per crossing) distributions for pp collisions during different years of Run 2 running [36].

Table 3.1: Comparison of the design performance of the LHC [29, 32], the actual performance achieved during Run 2, and the design performance of the HL-LHC for pp collisions [37].

Parameter	LHC design	2016	2017/2018	HL-LHC
Peak luminosity L (10^{34} cm^{-2} s^{-1})	1.0	1.4	2.1	6.5
Integrated luminosity at IP1/5 (fb^{-1}/year)	50	41	50/68	≈ 300
Number of bunches n_b	2808	2220	2556	2748
Average crossing rate (MHz)	31.6	25.0	28.7	30.9
Peak pileup	25	45	58	157
Average pileup	20	27	38	140
Center-of-mass energy $\sqrt{s}$ (TeV)	14	13	13	14
Number of protons per bunch N_b (10^{11})	1.15	1.18	1.25	2.2
Emittance ϵ_n (μm)	3.75	2.2	2.2	2.5
β^* (m)	0.55	0.40	0.33	0.20
Crossing angle at IP1/5 σ_c (μrad)	285	280	240	590
Bunch length σ_z (cm)	7.55	≈ 10	≈ 10	7.55
Beam size at IP1/5 σ (μm)	16.7	≈ 12	≈ 11	8.2
Reduction factor F	0.836	≈ 0.7	≈ 0.7	0.369
Average turnaround time (h)	5	7	6	3
Average length of a fill (h)	14	11	8	11
Number of physics runs (days)	150	146	140/145	160
Machine availability	71%	76%	83/79%	60%

C h a p t e r 4

THE COMPACT MUON SOLENOID EXPERIMENT

The Compact Muon Solenoid (CMS) is a general purpose detector whose compact high-field design, precision silicon pixel and strip tracker, crystal electromagnetic calorimeter (ECAL), hadron calorimeter (HCAL), and extensive muon system interleaved with the steel magnetic flux return are designed to provide a wide range of precision electroweak and heavy flavor standard model (SM) measurements, and high sensitivity for many beyond standard model (BSM) searches. The high resolution of the crystal ECAL, and the combined tracker and muon system have been optimized for the discovery of the Higgs boson, with an emphasis on its two photon and four lepton decay modes. CMS also has been shown to provide exceptional tracking and good energy measurements in high density heavy ion collisions, and high sensitivity and excellent measurement capability for rare B meson decays.

CMS features a strong magnetic field (3.8 Tesla) produced by a superconducting solenoidal magnet coil, inside of which is the silicon-based pixel and strip tracker system in the innermost layer providing excellent momentum as well as precise vertex measurements for charged particles, followed by a lead tungstate crystal-based electromagnetic calorimeter with excellent energy resolution for photons and electrons, and a hadronic calorimeter measuring the energy of hadrons. Outside the coil of the solenoid sit the muon detectors for muon detection and measurements of muons with good resolution up to the TeV range.

In general, each sub-detector is composed of cylindrical-shaped layers in the central region and circular disks in the forward region. The entire system is 28.7 m long, with a radius of 7.5 m, and weighs 14,000 tonnes. A diagram of the full CMS detector and each of its main subsystems is shown in Figure 4.1.

CMS has adopted a Cartesian coordinate system following the right-hand rule: the origin is taken as the nominal interaction point, the x-axis in the horizontal plane points towards the center of the LHC circle, the y-axis in the vertical plane points upwards, and the z-axis is aligned with the beamline with +z pointing towards the Jura mountains (which is the same as the direction of Beam 2 at IP5 that travels anticlockwise in the LHC tunnel, see Figure 3.1). In addition, spherical coordinates are often used, in which the azimuthal angle ϕ is defined as the angle in the x-y
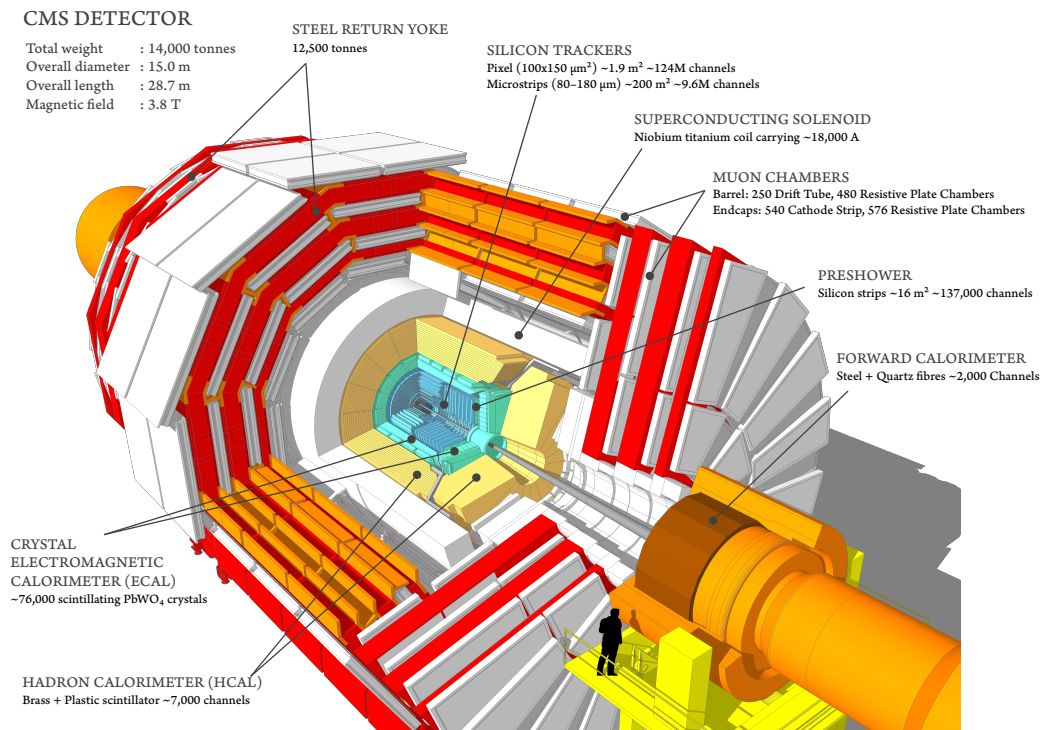
Figure 4.1: Cutaway diagram of the CMS detector and its subsystems [38].

plane measured from the +x axis, the radius *r* is defined as the distance from the origin to the point in the x-y plane, and the polar angle θ is the angle measured from the +z axis. To approximate the Lorentz invariant variable rapidity in the $m \ll |\mathbf{p}|$ limit, another coordinate variable called pseudorapidity η is defined in terms of the polar angle: $\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right]$.

The remainder of this chapter is organized as follows: Sections 4.1 to 4.5 describe the details of all the subsystems of the detector (magnets, tracker, calorimeters, muon detectors, and triggers), and Section 4.6 discusses the details of the object and event reconstruction methods and performance during Run 2.

4.1 Solenoid magnet

The CMS magnet is widely considered to be the most ambitious large bore superconducting magnet ever built, with a stored energy of 2.3 Gigajoules when in operation. The superconducting magnet coil of 6-m diameter and 12.5-m length [39–41] provides a large, intense, and homogeneous magnetic field in the inner tracking volume. The coil, composed of 2179 turns of superconducting wire braid stabilized with pure aluminum, wound in 4 layers, is operated at a current of 18,160 A. Together with

the return yoke, the solenoid generates a central magnetic field of 3.8 T along the beam line inside the solenoid (the yoke only contributes 8% to the central field). The coil is enclosed within three layers of steel flux return yoke in both the barrel and endcaps.

The mass of the iron in the return yoke is about 10000 tonnes, contributing about 71% of the total mass of the CMS detector. The map of the magnetic field in the entire CMS detector volume measured with the aid of cosmic muon events can be seen in Figure 4.2.
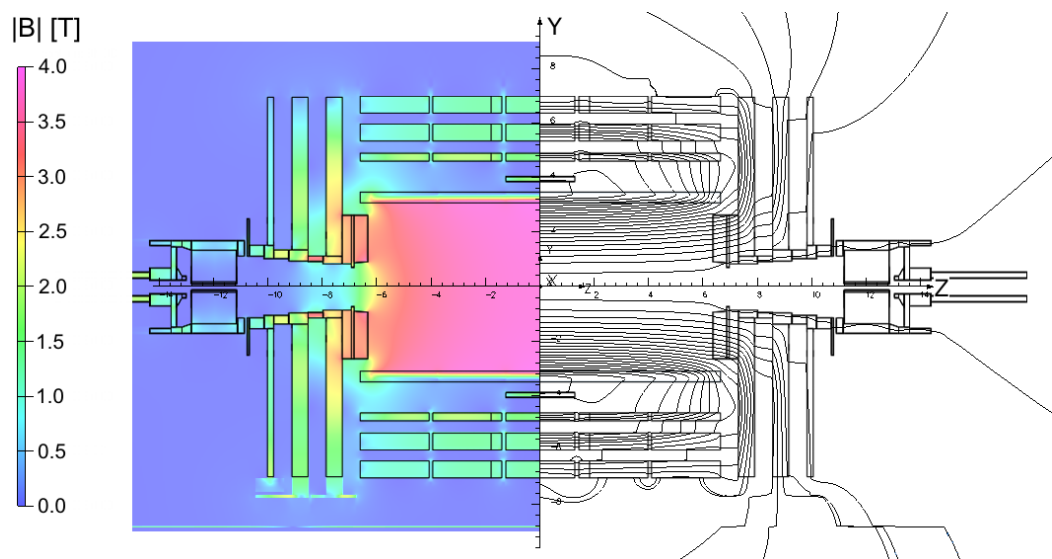
Figure 4.2: The magnetic field lines (right) and values of magnetic field $|B|$ (left) at different locations produced by the CMS magnet system. The plot is taken from [42].

4.2 Tracker

Many of the essential tasks of the CMS physics program require fast and precise position measurements of particle trajectories, yielding precise measurements of the track momenta, primary and secondary vertices, and other characteristics such as jet flavor tagging in collision events. The CMS tracker detectors which are used to reconstruct charged particle trajectories consist of an inner pixel detector and an outer strip tracker, both of which are based on solid state silicon detector technology.

The pixel detector uses n^+-in-n (n^+ pixelated implants on n-bulk) silicon sensors of 285 μm thick with a pixel size of $100 \times 150 \mu m^2$ in $r\phi \times rz$. Each pixel measures the three coordinates of a hit passing through it $(r, rz, r\phi)$. When a charged particle passes through the sensor, it usually generates signals in two or more neighboring

pixels. The amplitude of each pixel is then read out and an interpolation of the amplitudes gives a hit resolution of about 10μm (20μm) in $r\phi$ (rz), which is much better than the pixel size. In total, there are 48 million pixels in the barrel region populated in three layers at radii of 4.3, 7.3, and 10.4 cm, covering z from -27 cm to 27 cm, and 18 million pixels in the endcap populated in four disks at z of ±35.5 and ±46.5 cm, providing η coverage up to 2.5. During February-March 2017, a new pixel detector (part of the CMS Phase I upgrade [43]) was installed. The new pixel detector uses the same sensor but an optimized layout design and improved readout electronics, in order to cope efficiently with the higher luminosity (above the design value) produced by the LHC. The Phase I pixel has four barrel layers at radii of 3.0, 6.8, 10.9, and 16.0 cm, and has six endcap disks at z of ±29.1, ±39.6, and ±51.6 cm, with a total of 79 (45) million pixels in the barrel (endcaps).

The strip detector uses p-in-n sensors of 320 to 500 μm thick with distances between strips (pitch) of about 80 to 183 μm. There are 9.3 million strips in total in the CMS strip tracker. All strips run parallel to the beam line in the barrel, which gives a measurement in $r\phi$ coordinates for each hit. In addition, there are double-sided modules that have two layers of strips, with the additional layer rotated by about 100 mrad (5.7°) with respect to the nominal layer, giving two dimensional measurements of hits ($r\phi$ and rz). There are in total 10 strip layers in the barrel. Layers 1, 2, 5, 6 are double-sided and the others are single-sided. The first four layers are called the Tracker Inner Barrel (TIB), which spans a radius from 20 cm to 55 cm with a length of $|z| < 65$ cm. The last six layers are called the Tracker Outer Barrel (TOB), which spans a radius from 55 cm to 110 cm with a length of $|z| < 110$ cm. The single hit position resolution from the TIB (TOB) is about 13-38μm (18-47μm).

In addition, in the 20 cm to 55 cm radius and $65\,\text{cm} < |z| < 110\,\text{cm}$ region, disk-shaped detectors (TID) are placed. The sensors in the tracker endcaps (TECs) are trapezoidal with strips running in the radial direction (pointing towards and perpendicular to the beam line), measuring the ϕ coordinate of each hit. There are nine layers of disks in each endcap, covering $|z|$ from 120 cm to 280 cm and radius from 20 cm to 110 cm. A sketch of the CMS pixel and strip tracker layout is shown in Figure 4.3.

Although the silicon sensors in the pixel and strip detectors are very thin and thus usually cause almost no energy loss for particles passing through them, the large number of pixel/strip channels requires a dedicated mechanical support system and a large amount of readout electronics and cables to read out the signal, along
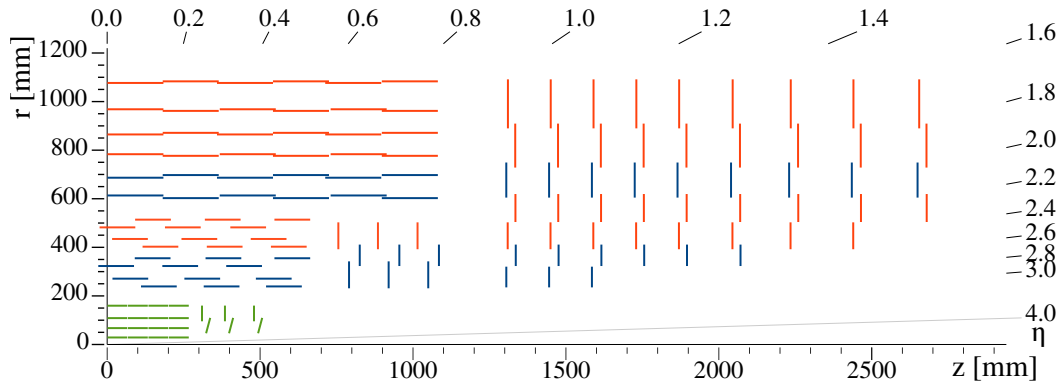
Figure 4.3: The layout of the CMS tracker system (after the Phase I upgrade) in one r-z quadrant. The pixel detector is shown in green; the red and blue segments are the single-sided and double-sided strip modules, respectively. The plot is taken from [44].

with services needed to cool the tracker. This results in a significant amount of material in the detector. These materials lead to energy loss and scattering of the charged particles and to conversion of neutral particles, and occasionally to nuclear interactions, before they reach the calorimeter systems designed to measure their energy. Prior to the pixel Phase I upgrade, the estimated total amount of material in the tracker systems as well as the beam pipe and support tube was about 0.4 radiation lengths (X_0) in the central region and up to about $1.8X_0$ at around $|\eta| = 1.5$, as shown in Figure 4.4. After the pixel Phase I upgrade, the amount of material was reduced by about $0.3X_0$ in the endcap region ($1.5 < |\eta| < 2.5$) due to a reduction in the mass of systems such as the cooling system [45].

More details on the CMS tracker system can be found in [46]. Details of the particle track and momentum reconstruction are discussed in Section 4.6 below.

4.3 Calorimeters

The CMS tracker can only measure charged particles with high resolution over a limited momentum range in a relatively clean environment as a result of the number of coordinates along each track and the limited bending of the trajectory of high momentum particles, and in a limited pseudorapidity region due to the partial alignment of the magnetic field direction and particle momentum direction in the forward region. For the momentum (energy) measurement of neutral particles (such as photons, neutrons, or neutral mesons which do not decay in the tracker volume), or charged particles in the very forward region, calorimeters are the only sources of information, and for charged particles in crowded environments (such as high
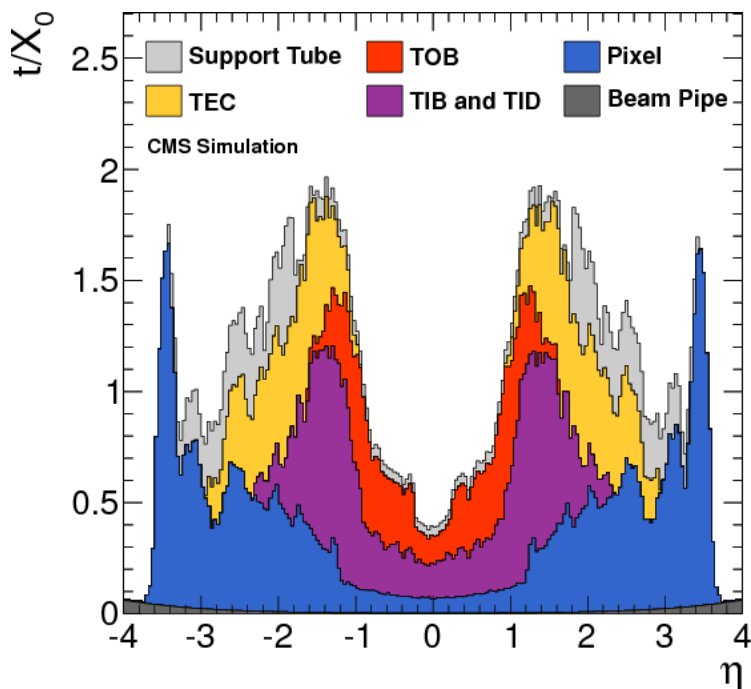
Figure 4.4: The material budget simulation of various sub-detectors of the CMS tracking system as a function of η, measured in units of radiation length (X_0). The left plot is for the detector before February 2017 (before Pixel Phase I upgrade). The plot is taken from [47].

energy jets) or charged particles with very high momentum, the calorimeters also may provide the most important energy measurement.

To achieve high energy resolution and identification of electrons and photons with high efficiency (which has been essential for the discovery and measurements of the Higgs boson in the two-photon decay channel, for example), CMS has adopted a system with two separate calorimeters: an electromagnetic calorimeter (ECAL) in the inner layer with enough number of radiation lengths (about 25 X_0) to contain and measure the energy of electromagnetic showers from high-energy electrons, positrons, photons, or neutral pions, and a hadronic calorimeter (HCAL) in the outer layer which contains and measures the energy of the rest of the hadronic showers from high-energy hadrons. The ECAL has a small enough thickness in interaction lengths (about one λ_I in the case of the CMS) such that the likelihood of absorbing all of the energy of a hadronic shower in the ECAL is small, and this aids in the separation of individual hadrons and jets from electrons and photons.

The basic unit of the CMS ECAL is a lead-tungstate (PbWO$_4$ [48]) crystal with a relatively small size both longitudinally (about 23 cm long) and transversely (about

$2.2 \times 2.2\,\mathrm{cm}^2$ at the inner end and $2.5 \times 2.5\,\mathrm{cm}^2$ at the outer end for the barrel crystals). The small crystal size is made possible by lead-tungstate's short radiation length ($X_0 = 0.89\,\mathrm{cm}$) and small Molière radius ($R_\mathrm{M} = 2.2\,\mathrm{cm}$), making the ECAL a homogeneous and hermetic calorimeter with a compact and finely granular design.

The scintillation light produced in the crystals after being excited by particles in an electromagnetic shower has a spectrum peaked at around 420-430 nm. Avalanche photodiodes (APDs) and vacuum phototriodes (VPTs) are attached to one end of the crystal as photodetectors in the barrel and endcap, respectively. After taking into account the light transmission in the crystal, the light collection efficiency by the photodetector, and the quantum efficiency of the photodetectors, about 4.5 photoelectrons are produced for 1 MeV energy deposition in the crystal [49]. This corresponds to a stochastic term of about 2.8% in the energy resolution, for the sum of the energy measured by a group of 3×3 crystals, i.e. $\sigma_E(\,\mathrm{GeV}) = 2.8\%\sqrt{E(\,\mathrm{GeV})}$. Note that due to radiation damage, the light transmission was progressively reduced during running (see Chapter 6 for details), so the stochastic term increased accordingly. The electronic noise in each readout channel was about 60-80 MeV during Run 2, corresponding to an uncertainty of about 200 MeV in the sum of energy in a group of 3×3 crystals. In addition, effects such as the non-uniformity of the light collection can also contribute about 0.3% to the energy resolution.

In summary, the energy resolution of ECAL can be expressed as the sum of three contributions (stochastic, noise, and constant terms):

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E(\,\mathrm{GeV})}} \oplus \frac{b}{E(\,\mathrm{GeV})} \oplus c, \qquad (4.1)$$

where for the sum of energy in a group of 3×3 crystals during Run 2, the values of a, b, c are measured to be about 2.8%, 20%, and 0.3%, respectively, not counting the effects of the inter-calibration among crystals which tends to increase the constant term, and of radiation damage which reduces the light transmission and therefore increases the stochastic term.

The crystals in the ECAL barrel (EB) are placed at an inner radius of $r = 1.29\,\mathrm{m}$, corresponding to a $\Delta\eta \times \Delta\phi = 0.0174 \times 0.0174$ coverage for each crystal. A total of 360 crystals cover the full 2π in the ϕ direction in each ring, and the η coverage of EB is from -1.479 to 1.479, corresponding to 170 crystals in the η direction. The total number of crystals in EB is therefore $360 \times 170 = 61200$. Crystals with the same value of $|\eta|$ have exactly the same size, but crystals at different $|\eta|$ positions have slightly different sizes in order to have all of the crystals pointing near to the

interaction point (with a small non-pointing angle of about $3°$ to avoid particles traveling out from the interaction point along the crack between crystals). In order to reduce the number of crystal types and ease the construction and assembly, in total 17 types of crystal sizes are used instead of 85, with every 5 crystals in η grouped together and sharing the same size. A group of $2 \times 5 = 10$ crystals in $\phi \times \eta$ is called a submodule, which are contained in an alveolar supporting structure. The gaps between crystals in the same submodule are about 0.3 mm, and the gaps between two submodules are about 0.5 mm. Groups of submodules are then grouped together and sit in the same supporting basket to form a module, which are called module type 1, 2, 3, and 4 from $|\eta| = 0$ to $|\eta| = 1.479$. Each module consists of 20×25 (20×20) crystals in $\phi \times \eta$ for module 1 (2, 3, 4). The crack between two modules is about 7 mm. Every four modules ranging from $\eta = 0$ to $\eta = 1.479$ are called a supermodule, which contains 20×85 crystals in $\phi \times \eta$; in total, there are 18 supermodules in the $\eta > 0$ region and 18 supermodules in the $\eta < 0$ region.

The crystals in the ECAL endcaps (EE) are placed at an inner $z = \pm 3.2$ m, covering a pseudorapidity range from 1.479 to 3.0. Similarly to the crystals in EB, the crystals in EE are also pointing to the interaction point with a small non-pointing angle in order to contain a shower in a minimum number of crystals and reduce the effect of gaps and cracks. A rectilinear grid of 5×5 crystals in $x \times y$ is called a supercrystal. Each endcap is composed of 276 supercrystals and 36 partial supercrystals at the edges (which are each composed of less than 5×5 crystals), corresponding to a total of 7324 crystals in each endcap. This rectilinear grid geometry forms an approximately circular shape for each endcap.

In order to enhance photon identification against photon pairs from neutral pion decay, a sampling calorimeter called a preshower which has a much finer granularity is placed just in front of each ECAL endcap in the region from $1.653 < |\eta| < 2.6$. The preshower is made of two layers of orthogonal silicon strip planes, each of which is accompanied by a layer of lead absorber in the front. The first layer of lead is about $2X_0$ thick and the second layer of lead is about $1X_0$ thick. The silicon strips have a pitch size of 1.9 mm, which provides much better spatial resolution and was designed in principle to improve the separation of photons against neutral pions.

The reconstruction of the energy of an electromagnetic shower starts from the digitized signal pulses converted from the analog output signals of the APDs or VPTs. Each pulse is sampled in 10 consecutive time-samples of 25 ns, corresponding to the inter-bunch spacing of the LHC. As a result, the recorded pulse is an overlap of the

signal from the current bunch crossing (BX) with signals from neighboring bunches, referred to as the out-of-time (OOT) pileup. To extract the signal from these OOT pileup contributions, a multifit algorithm [50] is employed that performs a template fit using a set of in-time and out-of-time predefined pulse shapes. The extracted signal amplitude in each channel is then calibrated individually and converted from ADC counts into GeV. Signals from different channels are then combined and clustered into basic clusters of 5×5 crystal matrices in EE (for EB, the basic cluster is a group of $5 \times N$ channels in $\eta \times \phi$, N ranges from 1 to 35) [51]. To collect all the energy of the photons that converted into electron-positron pairs, or energy of electrons that undergo bremsstrahlung before reaching the ECAL, one or multiple basic clusters in the same η ring are combined into a supercluster (SC). The energy of the SC is then corrected to account for incomplete shower containment of the total electron/photon energy, such as shower losses before reaching the ECAL, gaps between ECAL crystals, etc. The details of the energy corrections and calibration methods are found in Chapter 6 of this thesis. The typical relative electron/photon energy resolution (σ_E/E) has been measured to be approximately 2% in the EB and about 4% in the EE during Run 2, as shown in Figure 4.5.

The overall layout of EB and EE, together with the preshower detector, is shown in Figure 4.6. More details about the CMS ECAL can be found in [53]. The details of electron and photon reconstruction and identification are discussed in Section 4.6.

The CMS HCAL is placed right behind the ECAL, and is divided into 4 subsystems: barrel (HB) for $|\eta| < 1.4$, endcap (HE) for $1.3 < |\eta| < 3.0$, forward (HF) for $3.0 < |\eta| < 5.2$, and an additional subsystem in the barrel but outside the solenoid cryostat called HO which covers $|\eta| < 1.2$. Both HB and HE are sampling calorimeters with brass as the absorber and plastic scintillator as the active layers. HO also uses plastic scintillators as the active layers, but uses iron (as well as the solenoidal coil itself) as the absorber. HF uses scintillating quartz fibres as the active material and steel as the absorber. In HB, HE and HO, the scintillation light is carried by wavelength-shifting fibers to hybrid photodiodes (HPDs) for detection, while the scintillating light in HF is read out by photomultiplier tubes (PMTs). Each segment of the calorimeter tower covers $\eta \times \phi$ of about 0.087×0.087 in HB, HE, and HO, and about 0.175×0.175 in HF.

Figure 4.7 shows the details of the layout of HB, HE, HO, and HF. Details of the hadron reconstruction are discussed in Section 4.6.
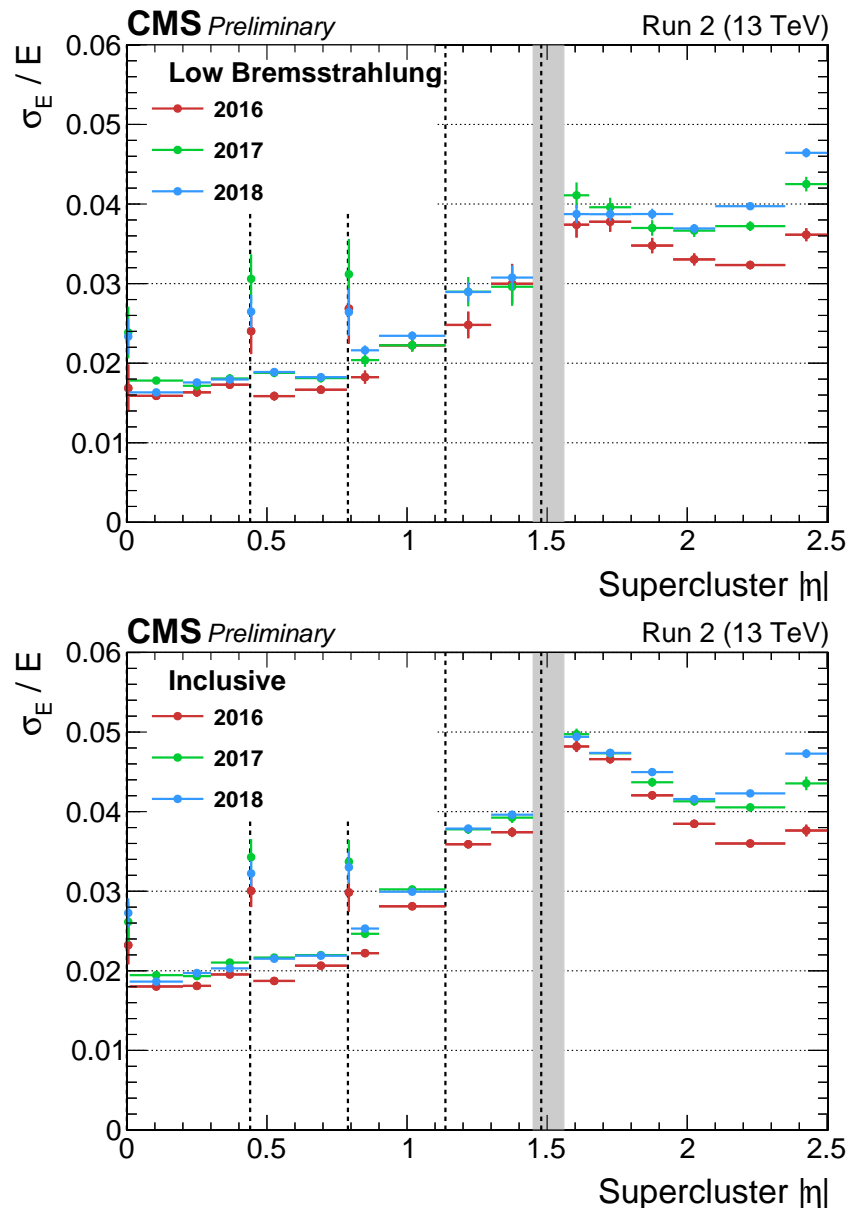
Figure 4.5: ECAL energy resolution measured with electrons from $Z \rightarrow ee$ events as a function of the supercluster η for different years during Run 2. The top plot is for electrons with low bremsstrahlung, and the bottom plot is for all electrons. The vertical lines indicate the boundaries between modules or boundary between EB and EE. The plots are from [52].

4.4 Muon detectors

Unlike electrons, muons with an energy in the GeV range predominantly lose energy through ionization in matter due to their much larger mass. The ionization energy loss (dE/dx) of a minimum ionizing particle (MIP) muon is only about 11 MeV/cm in materials such as iron, so muons can easily penetrate the material inside the
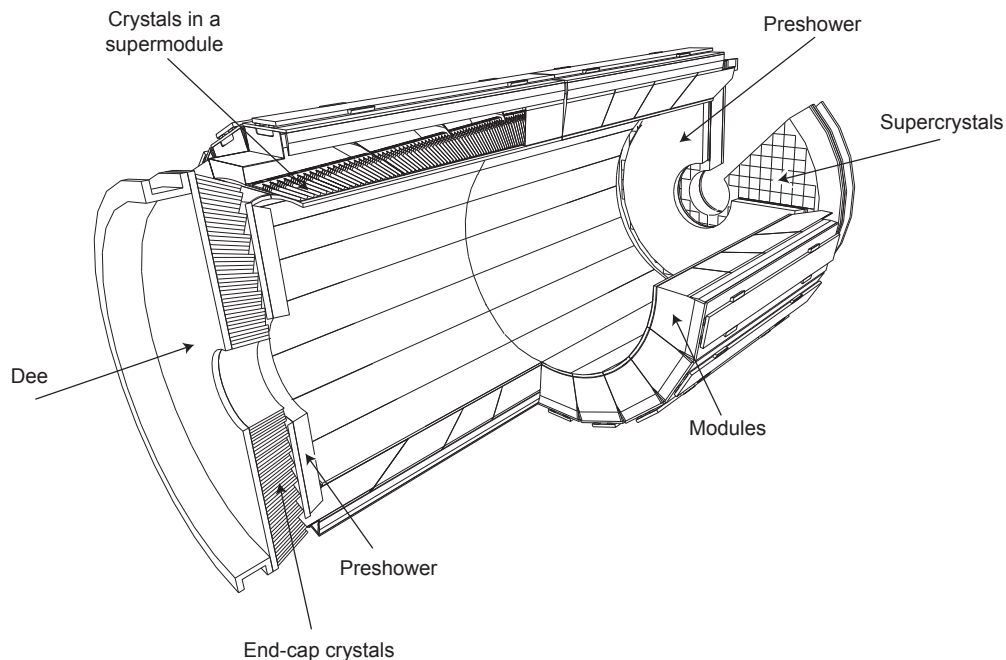
Figure 4.6: The overall layout of CMS ECAL in barrel and endcap, along with the preshower detector in the endcap [49].

solenoid magnet coil without losing much energy. This property makes the area outside the coil of the solenoid magnet a relatively clean environment to place the muon detectors.

The precise momentum measurement of muons in CMS is achieved with the combination of the inner tracker and the muon detectors. The measurement in the tracker dominates the precision for muons up to approximately 200 GeV, while the muon detector measurement which is limited by multiple scattering in the iron return yoke layers, dominates the momentum resolution for higher energy muons, and provides a resolution of better than 10% up to the TeV range [55]. The muon detectors are also essential for the identification and triggering of muon objects in CMS.

The CMS muon system is composed of three types of gas ionization chambers: the drift tubes (DTs) in the barrel ($|\eta| < 1.2$), cathode strip chambers (CSCs) in the endcaps ($0.9 < |\eta| < 2.4$), and resistive plane chambers in both the barrel and endcap regions ($|\eta| < 1.9$).

The basic unit of a DT chamber is a drift cell, which is a tube of around 2.4 meters long and with a size of 13 mm × 42 mm in the transverse plane. At the center of the tube is a positively charged wire (anode) of 50 μm diameter made with gold-plated
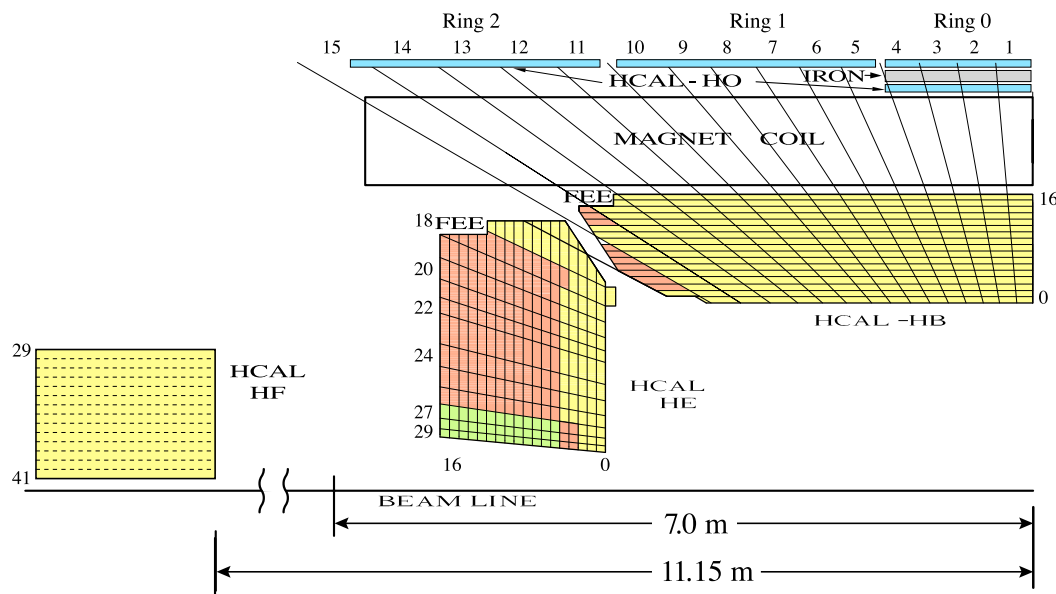
Figure 4.7: The overall layout of subsystems in the CMS HCAL, in one r-z quadrant [54].

stainless-steel, and the tube is filled with an 85% Ar + 15% CO_2 gas mixture. When a muon travels through a tube, it ionizes the gas and generates electrons and ions. A +3600 V voltage is applied to the wire (anode); +1800 V is applied to the two 42 mm wide walls of the tube (electrode); and -1200 V is applied to the two 13 mm wide walls of the tube (cathode). Under such an electric field, electrons travel to the anode with a drift velocity of about 54 μm/ns (corresponding to about a 390 ns travel time for the 21 mm distance from the tube wall to the wire), undergo an avalanche close to the wire, and generate output signals. (Ions, on the other hand, do not have such a directional movement due to the much larger energy losses resulting from collisions of the ions with gas atoms).

Figure 4.8 shows the transverse view of the position of the wires, the walls of the tube, and the drift lines (the electric field lines) of the electrons. The readout electronics measures the time of the signal with a high performance time-to-digital converter (TDC). The time measured from the TDC is then converted to the drift time of the electrons from the ionization cluster to the wire, and then the drift time is converted to the distance of the muon track in the cell to the wire, and this distance measurement (from a single wire) has a resolution of about 260 μm [56].

Four layers of drift cells with staggered centers are fixed together to form a superlayer. Each DT chamber is made of three superlayers, with the wires of the center superlayer being orthogonal to the beam line and providing the z position of the muon track,

while the wires of the two outer superlayers are parallel to the beam line and provide the $r - \phi$ measurement of the track. The combination of 8 track points from the two outer superlayers provide a global resolution of about 100 μm in $r - \phi$ and 150 μm in z.

A ring of DT chambers with the same value of the radius r makes a DT station. There are four DT stations in total, with the first station (MB1) placed just outside the solenoid, the 2nd and 3rd station (MB2, MB3) sitting in between the flux return yoke layers, and the 4th station (MB4) sitting outside the yoke. Note that in MB4, each DT chamber only has two superlayers, with the wires parallel to the beam line. A DT chamber is about 2.5 m long in the z dimension and about 1.9 m (4.1 m) long in the $r - \phi$ dimension for the chambers in MB1 (MB4). The layout of the DT stations is shown in Figure 4.9.

The CSC chamber is a multiwire proportional chamber but with two-dimensional read-out from both the anode and the cathode [57], which provides both the r position (from the anode wires) and the ϕ position (from the cathode strips) of a muon hit. A CSC chamber has 6 layers of gold-plated tungsten wires, with each wire running in the ϕ direction. The wire layers are embedded within 7 layers of cathode panels, and each cathode panel has milled copper strips which run perpendicular to the wires. The distance between two neighboring wires is about 3.1 mm and the distance between two neighboring strips is about 3 mm at the small r end and 16 mm at the large r end. The gas gap, defined as the distance between two cathode panels, is about 9.5 mm.

The chamber is filled with a 40% Ar + 50% CO_2 + 10% CF_4 gas mixture, and is operated at a high voltage of 3600 V, which produces an electric field large enough to generate an ionization cascade (an avalanche) when a muon passes through the chamber. After the avalanche is formed, the wire collects the electrons and a positive charge is induced on the cathode strips. The charge on the cathode strips depends on the distance from the strip to where the avalanche occurred. By interpolating among the charges on the different strips, one can measure the ϕ position of the muon hit. Figure 4.8 shows the position of the wires and strips and an illustration of how such an interpolation is done.

There are four CSC stations in each endcap (ME1 to ME4), and each station is a ring of chambers embedded in between two flux return yoke disks. Figure 4.9 shows the layout of the CSC stations in one endcap. With offline reconstruction, the spatial resolution measured by one CSC chamber is about 45 to 105 μm for the

ME1 chambers and around 130 μm for the other chambers [55].
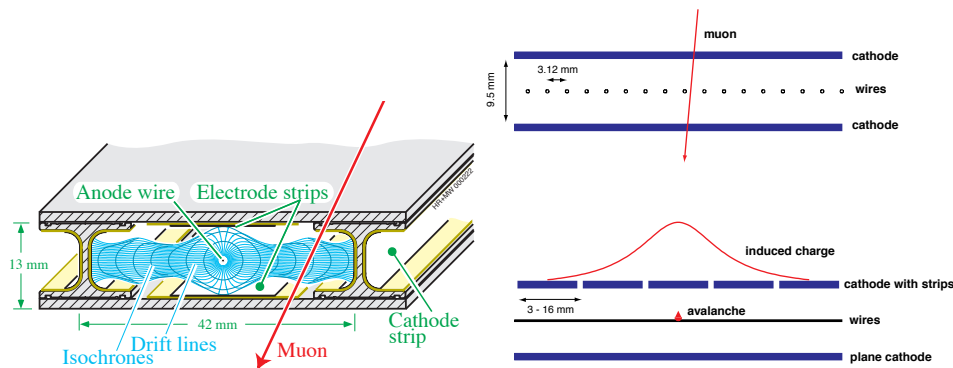
Figure 4.8: Left: transverse view of a drift cell which shows the position of the anode wires and the walls of electrode strips and cathode strips. Right: cross-sectional view of the CSC wires and cathode strips and an illustration of how the ϕ information of the hit is obtained, by interpolating the induced charge distribution on the cathode strips. The plots are taken from [58].

The main purpose of the Resistive Plate Counter (RPC) system is to provide excellent measurements of the times of muon hits together with acceptable position measurements, which are combined with the corresponding position measurements from the DT or CSC chambers in order to trigger on a muon candidate, and assign it to the correct bunch crossing with high efficiency.

A basic RPC module is made up of two gas gaps of 2 mm thickness each, and each gas gap consists of two resistive Bakelite plates. A conductive graphite layer coated on each of the surfaces of the Bakelite plates facing the gas gaps is used to provide a voltage of about 9600 V across each gap. The other side of each gap has milled copper strips at ground, protected with a mylar sheet. The gaps are filled with a gas mixture of 95.2% Freon + 4.5% isobutane + 0.3% sulphur hexafluoride (SF6), where the highly electronegative SF6 helps prevent breakdowns in the gas.

When a muon passes through a gas gap, an avalanche occurs and generates induced charges on the readout strips, which are located between and are shared by the two gas gaps. The readout electronics then measures the time of the signal with an intrinsic resolution of around 2 ns. Adding the uncertainty of the propagation time of the signal along the strip, the RPC measures the time of the hit with an overall resolution of better than 3 ns, which is much smaller than the LHC bunch spacing of 25 ns, and is precise enough to assign the muon track to the proper bunch crossing with high efficiency (above 95% during Run 2 [59]).

The layout of RPCs is similar to the DTs in the barrel and the CSCs in the endcaps. There are four RPC stations in the barrel: RB1 to RB4. RB1 and RB2 each have two layers on both sides of the DT station, while RB3 and RB4 each have only one layer placed on the inner side of the DT, facing the interaction point. The RPC strips in the barrel are parallel to the beam line and measure the $r - \phi$ coordinates of the hit. There are four RPC stations in each endcap (RE1 to RE4), with RE1 and RE3 on the outer side of the ME1 and ME3, and RE2 and RE4 on the inner side of ME2 and ME4. The RPC strips in the endcap are parallel to the CSC strips and measure the ϕ coordinate of the hit. The layout of the RPC stations is shown in Figure 4.9.

Further details on the CMS muon detectors can be found in [60]. The details of the muon triggers are discussed in Section 4.5, and details of the muon reconstruction are discussed in Section 4.6.
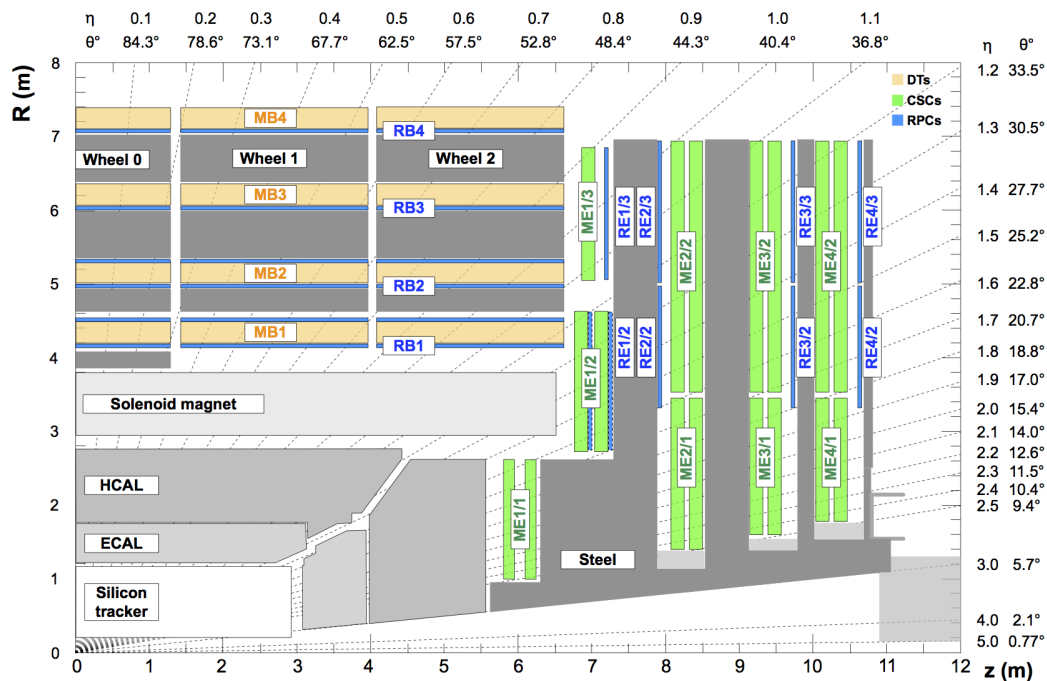
Figure 4.9: The layout of CMS subdetectors in the r-z quadrant with muon detectors (DT, CSC, RPC) marked and labeled in color. DTs are in orange color and labeled as MB1/2/3/4; CSCs are in green color and labeled as MEn/m, with n the index in the z direction and m the index in the R direction; RPCs are in blue color and labeled as RB1/2/3/4 for barrel RPCs and REn/m for endcap RPCs, with n the index in the z direction and m the index in the R direction. This plot is taken from [55].

4.5 Triggers

The storage needed to store the raw digitized signals from all CMS detector components for one bunch crossing event is about 1 megabyte, and with an average crossing rate of about 40 MHz, this means that CMS would have to transfer and store a few tens of terabytes of data per second if it were to store events from all bunch crossings. Processing and storing data at this rate would require much more computing power than is available, and would greatly exceed the CMS budget.

Therefore, an online data filtering system, the CMS Trigger and Data Acquisition (TriDAS) system [61, 62], is needed. As currently designed, the main CMS TriDAS output stream selects, digitizes, and stores events at a maximum event rate of about 1 kHz, which corresponds to processes with a total cross section of 50 nb for $L = 2.0 \times 10^{-34} \, \mathrm{cm}^{-2} \, \mathrm{s}^{-1}$. The TriDAS only stores the collision events of greatest interest, using dozens of trigger lines [62, 63] designed to maximize the coverage of all of the known and potentially new BSM channels, within the constraints imposed both by the trigger system itself and the available budgets that limit the ability of the CMS Collaboration to process, distribute and analyze the data.

In addition to the main output stream, there are specialized streams designed to enable various calibrations, including the stream dedicated to calibration of the ECAL using π^0 decays as described in Section 6.1 of this thesis, and a so-called "data scouting" stream in which only a small subset of the reconstructed event quantities is stored amounting to a few kilobytes per event. The use of the scouting stream enables CMS to explore a much larger parameter space and/or kinematic range, as in the study of a large sample of B-meson decays, for example.

The CMS trigger system is segmented into two levels: the Level-1 (L1) trigger which makes decisions to accept or reject every collision event with custom electronics at an acceptance rate of up to 100 kHz, and the High Level Trigger (HLT) which processes the L1 accepted events with software filters running in commercial CPUs and accepts events with interesting physics content at a rate of up to about 1 kHz.

The digitized output signals from the sub-detectors are continuously stored in the sub-detector front-end pipelined buffers. In addition, some sub-detectors (the calorimeters and muon detectors) also send their output data to the L1 trigger, which gives an L1 decision about 4 μs later (called the latency time) to the detector front-ends via the Timing, Trigger, and Control (TTC) system. Upon arrival of the L1 acceptance decision, the data in the detector front-end buffers are extracted and pushed into the Data Acquisition (DAQ) system by the Front-End Drivers (FEDs).

Data fragments from the FEDs are then transported from the electronics room (the Underground Service Cavern or USC, where the FEDs are located) about 90 meters deep underground to the surface building (SCX) by the FED-builder. The FED-builder also assembles data fragments into super-fragments and stores them in the buffers of the Read-out Units (RU). The RU-builder then assembles all data super-fragments that belong to the same event and sends them to a single unit of the HLT filter farm. Accepted events from the filter units are then passed through the computing services local area and site networks and are sent to the storage systems.

The CMS L1 trigger [61, 64] is composed of the calorimeter trigger and the muon trigger. A plot which illustrates the architecture and data flow of the L1 calorimeter trigger and L1 muon trigger is shown in Figure 4.10. There are two layers of Xilinx Virtex-7 Field Programmable Gate Array (FPGA)-based processing cards [65] in the calorimeter trigger. The layer 1 calorimeter trigger receives trigger primitives (TPs) produced by the ECAL and HCAL (HBHE, HF) front-end electronics and calibrates and sorts those TPs. Each TP contains the energy sum in a $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ region (corresponding to 5×5 ECAL crystals or one HCAL readout in HB). There are 18 cards in layer 1, and there are 72 TPs in the ϕ direction. Each card will process 4 out of the 72 TPs in the ϕ direction and for all the η regions. The calibrated and sorted TPs are then passed to one of the 9 cards in the layer 2 calorimeter trigger with a time multiplexed algorithm [66], in which physics objects such as electrons/photons, tau leptons, jets, and energy sums (such as scalar sum of p_T of all jets H_T) are reconstructed. Those objects are then passed to a demultiplexer (DeMux) board (also FPGA-based) where events are formatted and sent to the global trigger (μGT) for processing.

The global L1 muon trigger is composed of groups of processing cards (also FPGA-based) that are responsible for muons in different η regions: the barrel ($|\eta| < 0.83$), the endcap ($1.24 < |\eta| < 2.4$), and the overlap region ($0.83 < |\eta| < 1.24$). The drift tube (DT) and cathode strip chambers (CSC) send TPs to L1, while the resistive plate counter (RPC) system sends its hits (with precise timing information) to the L1 trigger system. The TPs from DT consist of coordinates and quality information of identified track segments; and the TPs from CSC have the information of a local track combined from track segments from the cathode and anode readout. The TPs of DT and the hits of RPC from the same station get merged by a so-called TwinMux card, in which the spatial information from DT and timing information from the barrel RPC get combined into superprimitives. The TwinMux card then sends the

superprimitives to the Barrel Muon Track Finder (BMTF) card. The TPs from CSC are collected by the Muon Port Cards, which sort the local tracks and send up to three candidates to the Endcap Muon Track Finder (EMTF). The EMTF also takes input from TPs that are reconstructed from endcap RPC hits in adjacent strips in a CPPF (concentrator preprocessor and fan-out) card. In the overlap region, the tracks are built in the Overlap Muon Track Finder (OMTF) card with inputs from the DT, RPC barrel, RPC endcap, RPC endcap, and CSC.

The three track finders (BMTF, EMTF, OMTF) then all run the same algorithm independently and simultaneously to build muon track candidates, and each of the track finders are segmented into several sectors, with each sector responsible for building track candidates in a given range of ϕ ($30°$ for BMTF, $60°$ for OMTF and EMTF). All of the three track finders will then send the muon candidates to the Global Muon Trigger (μGMT), in which the muons get sorted and duplicate muons from different boards get cleaned, and then the μGMT will send up to eight muons of the best quality and p_T to the Global Trigger for processing.
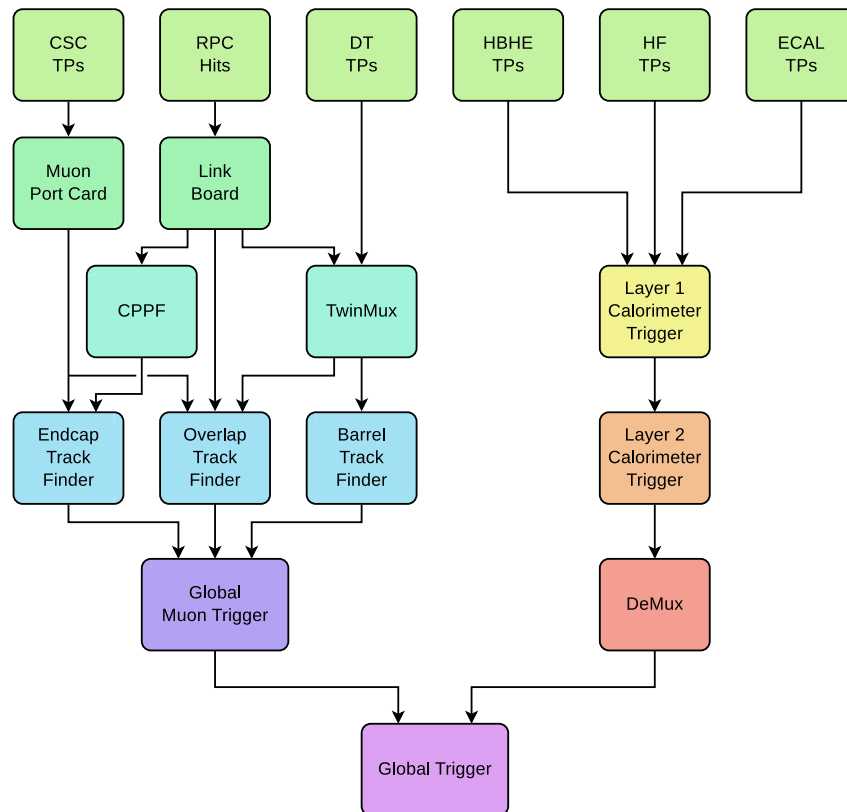
Figure 4.10: The architecture and data flow of the CMS Level-1 Trigger during LHC Run 2 [64].

The processing boards (also FPGA-based) in the Global Muon Trigger take all the objects from the L1 calorimeter trigger and the L1 muon trigger (electrons/photons, tau leptons, jets, energy sums, and muons), and then make a trigger decision to accept or reject the event, based on a list of criteria called the trigger "menu." An event that passes any criterion (called a Level 1 or L1 seed) in the menu is accepted by L1 and then passed through to the High Level Trigger (HLT). Each criterion in the menu is typically a simple requirement of the presence of at least one or more objects (of the same type or different types) with p_T or energy greater than some thresholds.

The list of the most used algorithms in the L1 menu during CMS Run 2 running can be found in [64]. Among these, the two simplest L1 trigger algorithms are the single muon trigger and the single electron/photon trigger. The L1 trigger efficiency for a muon object with p_T above the trigger threshold is typically 93%, with slightly smaller efficiency (around 90%) for muons with very high p_T [67]. The L1 trigger efficiency for electron/photon objects with energy above the trigger threshold is larger, and is very close to 100% [68]. Figure 4.11 shows the efficiency of the single e/γ and single muon triggers measured with Run 2 data as a function of the transverse energy (transverse momentum) of the e/γ (muon) objects. More specific requirements on higher level variables such as invariant mass or transverse mass can also be executed with the FPGA boards in μGMT. For example a vector boson fusion Higgs trigger has a cut on the invariant mass of the two jets [69].

The HLT takes the full information from all sub-detectors (including the tracker) of an event that passes the L1 trigger decision, and runs a faster version of the offline event reconstruction algorithm (including the most time consuming track reconstruction algorithm). Each filter algorithm in the HLT menu can have multiple requirements on different objects as well as requirements on high level variables (in principle, an arbitrary software code can be implemented and executed in the HLT CPUs). Different HLT algorithms are usually designed targeting a specific physics process. The definition of the HLT filters used in this thesis, along with the efficiency measurements, can be found in the corresponding Chapters for each physics analysis.

4.6   Event reconstruction

The reconstruction of physics objects, such as electrons, muons, taus, hadrons, and missing transverse momentum (p_T^{miss}) relies on the information from more than one
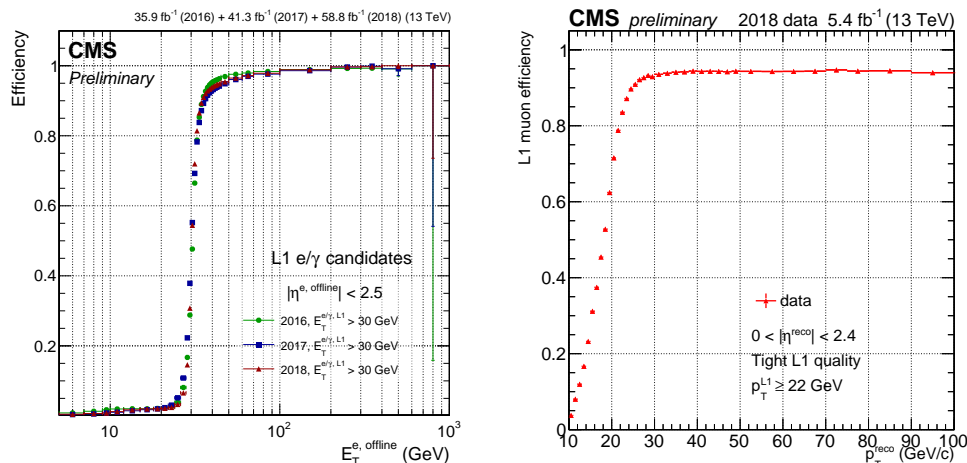
Figure 4.11: Left (right): L1 trigger efficiency for electron/photon (muon) objects as a function of the transverse energy (transverse momentum) of the object measured with Run 2 data for the single e/γ and single muon L1 triggers. The plots are from [68, 70] (see details of the measurement methods in the references).

sub-detector in CMS, as many of the objects leave signals in multiple sub-detectors. Figure 4.12 shows an illustration of the interactions of different particles with the CMS detector. To get a complete and non-overlapping list of particles in one event, we start with the basic reconstructed elements from each sub-detector, namely tracks from the tracker and muon detectors and clusters in the ECAL and HCAL, and then we topologically link those elements across sub-detectors based on the trajectory projections for charged and neutral particles. From those links, the list of particles are extracted based on their characteristic contents in the following order (once a particle is identified, all the associated elements are removed from the list): muons, electrons, photons, charged hadrons, and neutral hadrons. The properties of each particle in the output list are then calculated by combining the measurements from the different sub-detectors. The above steps of identifying and reconstructing the full list of particles in the entire event is called particle-flow (PF) reconstruction [71], which is widely used in CMS.

Tracking is the cornerstone of PF reconstruction, and plays a role in the identification of all particles: a muon is a link between a tracker track and a muon track; an electron is a link between a track (often a short or bad track due to bremsstrahlung) and ECAL cluster(s); a converted photon is a link between a track (without first hits in the first few pixel layers) and ECAL clusters; a charged hadron is a link between a track and calorimeter clusters (ECAL and/or HCAL); an unconverted photon is an ECAL
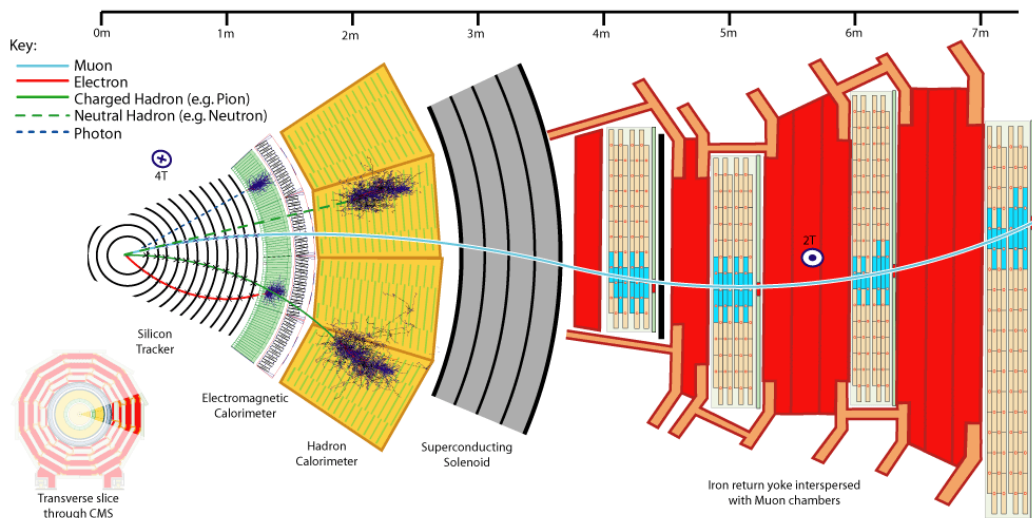
Figure 4.12: A cross sectional view of a sector of the central region of the CMS detector, together with the interactions of different types of particles with the detector [72].

cluster not linked to a track; and a neutral hadron is an HCAL cluster not linked to a track.

Track reconstruction is a process of combining hits in different layers of the tracker or muon detectors to obtain the momentum and position information of charged particles making those hits. As described in Section 4.2 and 4.4, the local reconstruction of hits can provide local coordinates of the hits with a resolution of 10-50 μm in the tracker, or 100-150 μm in the muon detectors. The measured local coordinates are then translated to global coordinates with the mapping of the location of the detector elements. From the global coordinates of the hits, tracker tracks and muon tracks are reconstructed independently using the track finder algorithm based on a combinatorial Kalman filter [73].

Tracker tracks are reconstructed in an iterative procedure, with each iteration targeting different kinds of tracks: the first iteration is for high p_T prompt tracks, the second iteration is for low p_T prompt tracks, and the third to fifth iterations are for non-prompt tracks (tracks that originate from outside the beam spot) and for any remaining tracks not found in the previous iterations. Once one iteration is finished, all the associated hits are removed from consideration for the next iteration. Each iteration begins with a list of seeds, which are hit pairs or triplets from pixel layers and/or strip layers with the p_T and the distance to the assumed origin compatible with the target of that iteration. From each seed, the track finder extrapolates the

trajectory of seed hits to the outer layers and searches for hits that are on or close to the trajectory with the Kalman filter.

A track-fitting procedure is then performed to estimate the parameters of the trajectory with a Kalman filter and smoother. The output of this procedure is a collection of track candidates with some quality variables, such as the number of layers with hits associated to the track, the number of layers in the middle of the track without hits, and the track's impact parameters. Requirements on the quality variables are then applied to the candidates, in order to select good tracks and reject fake tracks (tracks that are not associated with any charged particle). For tracks in the muon detector (called standalone-muon tracks), the reconstruction procedure uses seeds consisting of groups of DT or CSC segments (a segment is a straight-line track built from all layers in one DT or CSC chamber) and applies the Kalman-filter technique.

After the reconstruction of tracker tracks and standalone-muon tracks, a muon object can be reconstructed by combining the two of them. According to the order of this matching procedure, two types of muons can be reconstructed: tracker muons and global muons. For a tracker muon, the matching starts from a tracker track, and then one extrapolates the tracker track to the muon system and looks for DT or CSC segments. If one finds at least one matching segment in the projected trajectory, a tracker muon is then considered to be found. The global muon, on the other hand, is reconstructed by comparing a standalone-muon track with tracker tracks and selecting a tracker track whose parameters match with the standalone-muon track, and then performing a combined fit with the Kalman filter. The tracker muons have higher efficiency, since some tracker muons that are not global muons typically only have segments in the innermost muon chamber layers, and the reconstruction of segments in the outer layer chambers are sometimes corrupted by punch-through hadron showers. The global muons, however, are much cleaner and have a smaller misidentification rate compared to tracker muons.

The momentum of the muon object is measured based on the combined information from both the tracker and muon system, which improves the momentum resolution compared to tracker-only measurements, especially for high p_T muons ($> 200\,\mathrm{GeV}$). Based on the different needs for the balance between efficiency and purity, different sets of muon identifications can be defined based on the muon properties, such as the track fit χ^2 and the number of hits on the track. Muons are used in the search for the production of three massive bosons, which is part of this thesis, and the details of the selection criteria and efficiency of the muon object used in making

that observation are discussed in the corresponding chapter (see Section 7.3).

Electrons are reconstructed by combining the momentum measurement from the tracker and the energy measurement from the ECAL. Due to the material present in the tracker (see Section 4.2), the electrons traveling through the tracker lose a fraction of their energy by bremsstrahlung. When there is little or no bremsstrahlung, the electron deposits nearly all of its energy in a small group of neighboring ECAL crystals; for example, a 120 GeV electron entering the center of the front face of a crystal and traveling along the long crystal axis deposits 97% of its energy on average in a 5x5 crystal array. When bremsstrahlung occurs and photons are emitted, the electron and photons will generally end up hitting different locations in the ECAL due to the very strong magnetic field in the tracker and ECAL volume. Because of the direction of the magnetic field, which is predominantly parallel to the beam axis, the energy of the initial electron is mainly spread in the ϕ direction. In order to deal with this effect, the total energy of the initial electron is reconstructed by grouping together one or several 5×5 or 5×N crystal arrays in EE or EB (each a "basic cluster") into a "supercluster" (SC), starting from the seed crystal which has the maximum energy deposited, and then extending outward along the $\pm\phi$ direction [51].

Once the SC is identified, its position and energy can then be used to find the first few hits (seeds) of the electron in the tracker, either by matching the reconstructed tracker hits to the SC, or by matching reconstructed tracks from a general charged particle algorithm to the SC.

Once the electron seeds are found, they are then used to initiate the dedicated electron track building and fitting procedure, which is based on the Kalman filter method. The dedicated electron track building process takes into account the possibility of a change of curvature due to bremsstrahlung of the electron by applying a loose matching requirement to the predicted and found hit positions in each layer. The track parameters of the electron are then estimated based on the hits collected in the above procedure, by using a Gaussian sum filter (GSF) technique which takes into account the energy loss of the track in each layer. The momentum of the electron is determined by regression using both the measured track momentum and the energy of the SC. This method greatly improves the momentum resolution, down to about 2% for both low and high energy electrons (see Figure 4.5).

Electrons are also used in the search for three massive vector boson production, which is part of this thesis, and the details of the selection criteria and efficiency of the electron object used in making that observation are also discussed in the

corresponding chapter (see Section 7.3).

The reconstruction of photon objects is primarily based on ECAL clusters, with additional selection requirements that no hits are found in the corresponding pixel area and no (or very little) energy is deposited in the corresponding HCAL area. To reconstruct all the energy for converted photons, a supercluster is also used. A photon is one of the main objects used in the delayed photon search in this thesis, and dedicated identification criteria have been developed for photon objects that originate from a displaced vertex. The details of these criteria are described in Section 8.2.

Once the muons, electrons, and photons are reconstructed in an event, the corresponding elements in the tracker, calorimeter, and muon detectors are removed prior to the reconstruction of other particles. The two main remaining types of particles are charged hadrons and neutral hadrons. Within the tracker coverage ($|\eta| < 2.5$), a link of an ECAL cluster with an HCAL cluster without any matching track is defined as a neutral hadron, while such a link with a matching track is classified as a charged hadron. Outside of the tracker coverage ($|\eta| > 2.5$), the charged hadrons and neutral hadrons are indistinguishable, and therefore such links are assumed to come from the same (neutral or charged) hadron shower.

More complicated objects, such as jets and tau leptons decaying to hadrons (τ_{h}), can be reconstructed based on the reconstruction of the above PF objects (muons, electrons, photons, charged hadrons, neutral hadrons). Jets are reconstructed with the anti-k_{T} algorithm [74–76]. The basic elements used for jet clustering can be either PF candidates (PF jets) or calorimeter clusters (Calo jets). PF jets are used throughout this thesis. For the hadronically decaying tau leptons, the final state might include one or several charged hadrons and zero, one, or more neutral pions (π^0). Such tau objects are reconstructed from particles inside a jet object by the hadrons-plus-strips (HPS) algorithm [77–79]. The π^0 mesons from τ_{h} decay will decay to photon pairs, which will likely convert to electron-positron pairs, leaving ECAL clusters along the ϕ direction, which are called a "strip" (defined as a $\Delta\eta \times \Delta\phi = 0.05 \times 0.20$ window). The energy of a π^0 is obtained by summing up all the energy deposits in that strip. Different hadronic decay modes are then identified based on the set of charges in the charged hadrons and strips contained in a jet as well as the mass of the system of charged hadrons plus strips.

C h a p t e r   5

A MIP TIMING DETECTOR FOR THE CMS PHASE-2 UPGRADE

As mentioned in Chapter 3, the LHC will undergo a major upgrade in 2025-2027, after which a new phase of the physics program called the High Luminosity LHC (HL-LHC) will start. The luminosity of the HL-LHC is projected to reach 5 to $7.5 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, with a corresponding pileup of 140-200 interactions per bunch crossing. The reconstruction of physics objects in such a high pileup environment is very challenging, so the CMS detector will also undergo major upgrades (as part of the so-called Phase-2 Upgrade) before data taking at the HL-LHC starts. Along with the upgrades of many existing sub-detectors, CMS will install a new sub-detector, the MIP Timing Detector (MTD), placed in between the tracker and the ECAL [80].

The MTD is designed to measure the time of a minimum ionizing particle (a MIP) or other charged particle passes through an MTD detector layer with a precision of 30-40 ps at the beginning of HL-LHC operation, and 50-60 ps at the end of HL-LHC operation following irradiation over a ten year period.

With the additional requirements on the times of the tracks measured in the MTD, the number of tracks from pileup vertices that are incorrectly assigned to the hard primary vertex can be significantly reduced. This reduction of pileup tracks associated to the vertex will effectively clean away much of the pollution from pileup, and enable the particle flow reconstruction to function with significantly improved efficiency, resolution, and background rejection for many physics objects. A few examples of the improvements expected as a result of use of the MTD precise timing measurements include: increased identification efficiency for leptons and photons due to fewer tracks in the isolation cone, significantly improved resolution and more accurate angles in the reconstruction of jets and $p_{\mathrm{T}}^{\mathrm{miss}}$, and better efficiency and accuracy in the identification and reconstruction of other objects that heavily rely on primary and secondary vertex reconstruction, such as b-jets.

In addition to the physics object reconstruction improvement, the timing measurement of charged particles as well as the vertices will bring new CMS physics capabilities, as in the search for long-lived particles (LLP), for which the MTD will significantly improve the reach in mass and proper lifetime $c\tau$ for LLPs that decay

inside the tracker volume.

The MTD is composed of one layer with a cylindrical barrel layout in the central barrel region (BTL) and two disks in each endcap (ETL), as shown in Figure 5.1. The ETL is made with Low Gain Avalanche Detectors (LGAD) sensors $300\ \mu m$ thick (including a $50\ \mu m$ thick depletion region) with cell sizes of $1.3 \times 1.3\ mm^2$. The internal gain (about 10-30) of the LGAD sensors provides a large output signal and thus gives excellent timing resolution. Each endcap has two disks of sensors in order to provide full coverage of the space occupied by the readout electronics and cables in one layer. In total, 16624 LGAD sensors are used in the two endcaps to cover $|\eta|$ from 1.6 to 3.0 at a distance of $|z| = 2.98\ m$ from the Interaction Point.
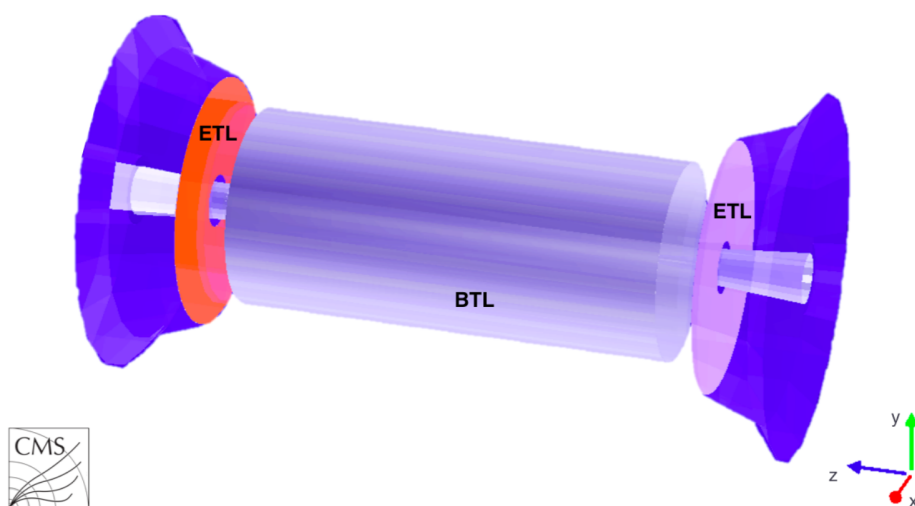
Figure 5.1: A schematic view of the MTD layout: the BTL is the grey cylinder and the orange and light violet dics are ETL disks.

In the barrel, the BTL is placed at $r = 1.148\ m$ with a length of 5 m, with a much larger total surface area than that of ETL (about 2.5 times bigger). Therefore a more cost effective sensor choice is made for BTL, with small Cerium doped Lutetium Yttrium Orthosilicate (LYSO:Ce) crystal scintillator bars as the sensitive material elements, each of which is read out by silicon photomultipliers (SiPMs) at the two ends. LYSO crystals have a very large light yield (about 32 scintillation photons per keV deposited) and a fast rise time, which makes the thin layer BTL design which fits in the space between the tracker support tube and the ECAL possible, with very good timing resolution.

Originally, there were two proposals for the size and alignment of a basic BTL sensor. The first proposal, called the tile geometry [81], is based on an LYSO

tile with a size of approximately $11 \times 11 \times 4 \, \text{mm}^3$, with one of the $11 \times 11 \, \text{mm}^2$ faces glued to a silicon photomultiplier (SiPM) of $4 \times 4 \, \text{mm}^2$ size, where a MIP particle from the interaction point enters the opposite $11 \times 11 \, \text{mm}^2$ face. The second proposal, called the bar geometry [80], uses a $57 \times 3 \times 3 \, \text{mm}^3$ LYSO bar as the basic element, with each of the two $3 \times 3 \, \text{mm}^2$ end faces attached to a $3 \times 3 \, \text{mm}^2$ SiPM, where a MIP particle from the interaction point enters one of the $57 \times 3 \, \text{mm}^2$ faces and generates scintillation light that travels along the bar in both directions and is collected by the SiPMs at the two ends.

Based on test beam data, it has been found that for the tile geometry, the measured timestamp has a strong dependence on the impact point of the MIP particle on the LYSO tile. In order to correct for that, an impact point measurement with a precision of 1 mm or better would be required, which could be achieved from the extrapolation of tracks with $p_{\text{T}} > 2 \, \text{GeV}$, but not for tracks with $p_{\text{T}} < 2 \, \text{GeV}$ (which is the p_{T} range of many of the pileup tracks). In addition, the measured time resolution is worse when the MIP passes near the edge of the LYSO tile, since less scintillation light is propagated and collected by the SiPM. In order to minimize the impact point dependence of the timestamp as well as the time resolution, it was found that a much larger area SiPM would be needed, to cover a larger fraction of the LYSO tile surface, but such a SiPM would significantly increase the cost of the MTD as a whole.

In the remainder of this chapter, the experimental study (with beam tests) of the impact point dependence of the measured timestamp and time resolution for different sizes of LYSOs and SiPMs of tile geometry is presented, where the results from the beam tests have been cross-checked by comparing to simulations of the light propagation for different LYSO and SiPM combinations.

5.1 Test beam study of timing performance for different sizes of LYSO and SiPM sensors

In order to study the LYSO+SiPM timing performance for different impact points of a MIP on the sensor and for different sensor sizes of LYSO and SiPM, the sensors have been tested with proton beams provided by the Fermilab Test Beam Facility (FTBF). The results presented in this section are based on the test data taken during June 2018 at the FTBF. The schematic diagram of the layout of the sensors under test and the auxiliary devices in the beam line can be seen in Figure 5.2. The test used MIPs from the 120 GeV proton beam provided by the FTBF, which go from right to left

in Figure 5.2. The LYSO+SiPM sensors under test were placed downstream of the pixel and strip telescope which was used for position measurement of the beam. The pixel detectors in the telescope were leftover modules from the CMS pixel detector and the strip detectors were leftover from the D0 tracker, which together gave a very precise determination of the extrapolation of the beam position downstream, where the sensors under test were placed. The resolution on this extrapolated position was better than 10 μm) [82, 83], which was sufficiently precise for the testbeam study.
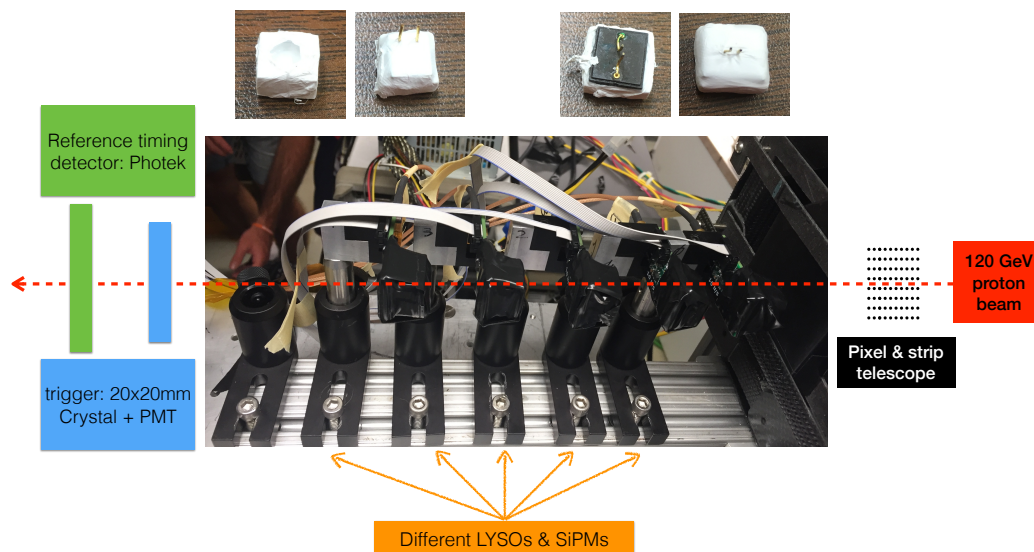
Figure 5.2: The alignment of the LYSO+SiPM sensors together with the pixel and strip telescope, the trigger and reference timing detectors in the June 2018 test with FTBF. The four pictures on the upper part show an example of the wrapping of the LYSO+SiPM sensors.

In total, 10 different LYSO+SiPM combinations from two types of SiPMs with different sizes (3×3 mm^2 and 6×6 mm^2 of the Hamamatsu S13360 series) and 5 types of LYSO crystal tiles with different sizes (5×5 mm^2, 8×8 mm^2, 11×11 mm^2, 12×12 mm^2, and 13×13 mm^2; all 4 mm thick) have been tested. The SiPM was mated to one face of the LYSO crystal with optical silicone grease for light transmission from the LYSO tile to the SiPM. The remaining surfaces of the LYSO crystal were wrapped with Teflon tape, as shown in the upper pictures in Figure 5.2, to reflect back the light that is exiting those surfaces, in order to increase the overall light collection efficiency.

The LYSO+SiPM sensors were then attached to a readout board (with an internal gain of 10), and black tape was used to wrap the entire readout board and LYSO+SiPM sensor in order to shield out light from the environment. A $20 \times 20 \times 5$ mm^3 LYSO

crystal attached to a photomultiplier tube (PMT) was placed downstream of the LYSO+SiPM sensors, to act as the trigger for the readout of the entire setup. At the very end of the setup downstream, a Photek 240 Micro Channel Plate-PMT (MCP) was placed as the reference timing detector. The resolution of the timestamp provided by the MCP was measured in situ to be 14 ps (measured by placing another MCP of the same type in the beam line), which is well below the time resolution of the LYSO+SiPM sensor.

The output signals from the MCP and the different LYSO+SiPM boards were digitized by the same VME-based Domino Sampling Chip (DRS) waveform digitizer module (CAEN V1742 [84]) with a sampling rate of 5 GHz. From the digitized output signal pulse, the timestamp was reconstructed as the time when the pulse reached 10% of the pulse maximum amplitude. This constant ratio timing method minimized the dependence of the reconstructed timestamp on the amplitude of the pulse, and any residual dependence was corrected (with a time-walk correction). From the digitized pulse, we also measured the rise time, which is defined as the time from 10% to 90% of the pulse amplitude. The time resolution of the LYSO+SiPM channel was measured by taking the time difference (Δt) between the LYSO+SiPM channel and the MCP channel, and by fitting the distribution of the Δt from many events with a Gaussian function and taking the standard deviation as the time resolution.

For a beam passing the LYSO+SiPM sensor at different locations, the output pulse had a different amplitude, rise time, mean timestamp, and time resolution. Figure 5.3 shows the 2D maps of the measured values and uncertainties in these four variables, in bins of the beam impact point X and Y coordinates measured by the pixel and strip telescope. From these 2D maps, one can clearly see that the time resolution of the LYSO+SiPM is best when the beam hits the center of the sensor, and it gets worse as the impact point approaches an edge of the sensor. This position dependence is due to the larger amount of scintillation light collected by the SiPM and the faster rise time when a particle in the beam hits the center, as can be seen from the amplitude and rise time maps. In addition, the measured mean timestamp also gets larger for a hit at the edge of the sensor (up to 240 ps larger on average for hits in the corner of the LYSO sensor relative to hits at the center), as it takes longer for the majority of the scintillation light to reach the SiPM photosensor for the impact points nearer to the edges.

The same measurements have been performed with all combinations of the different
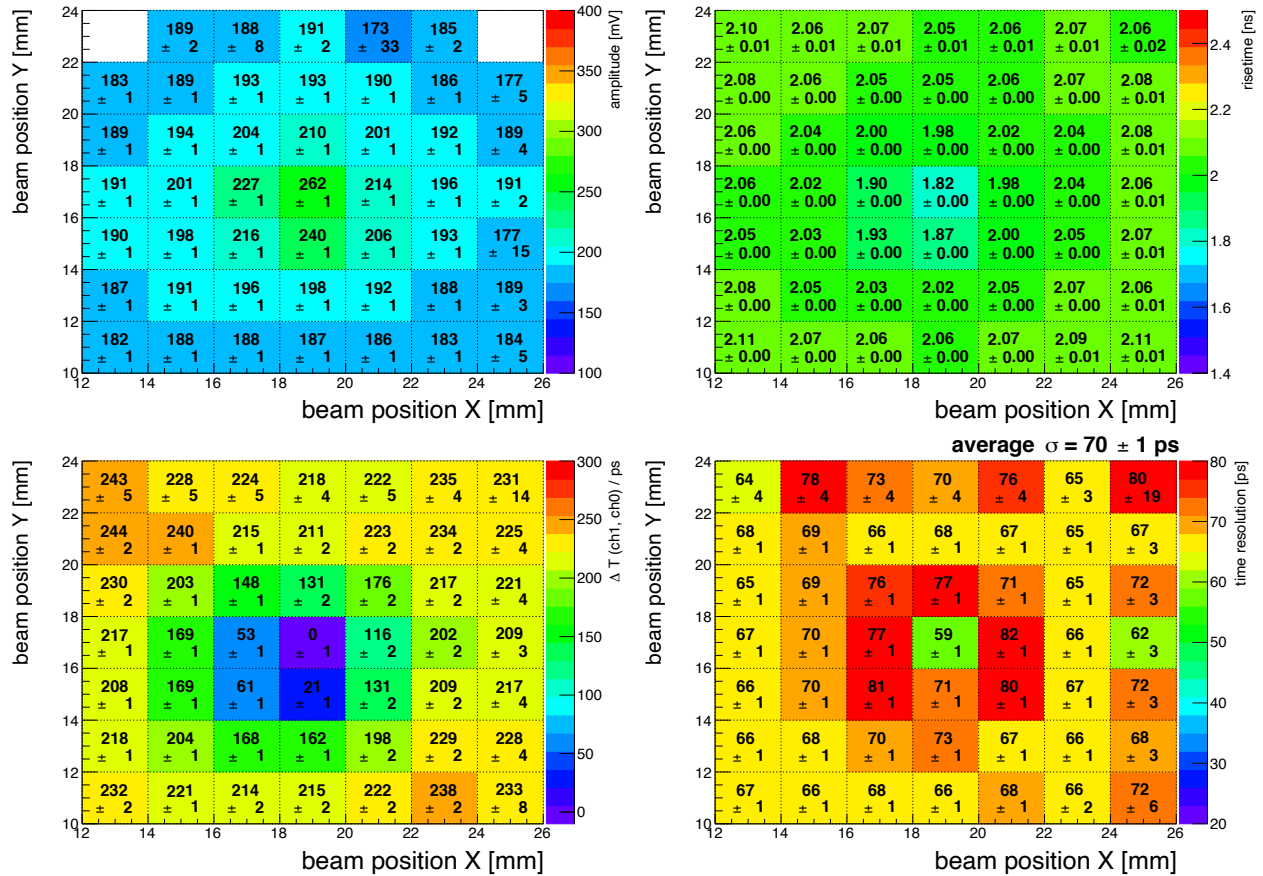
Figure 5.3: Pulse amplitude (top-left), risetime (top-right), time difference of the LYSO+SiPM signal and MCP signal (bottom-left), and time resolution of the LYSO+SiPM sensor (bottom-right) in 2D bins of the beam impact point position X and Y coordinates, for a $3 \times 3 \, \text{mm}^2$ SiPM attached to a $13 \times 13 \times 4 \, \text{mm}^3$ LYSO.

SiPM sizes and LYSO sizes. It was found that the bigger the SiPM relative to the LYSO tile size, the better the time resolution. Figure 5.4 shows the measured time resolution as a function of the SiPM and LYSO sensor aspect ratio, which is defined as the surface area of the SiPM divided by surface area of the LYSO tile (for example, a $3 \times 3 \, \text{mm}^2$ SiPM attached to a $13 \times 13 \times 4 \, \text{mm}^3$ LYSO tile has an aspect ratio of $3 \times 3 / (13 \times 13) = 0.053$). A clear improvement of the time resolution is seen as the aspect ratio increases.

In order to achieve an average time resolution below 30 ps for a MIP hitting anywhere across the LYSO tile face, an aspect ratio greater than 0.7 is required, as shown in Figure 5.4. For an LYSO tile of size $11 \times 11 \times 4 \, \text{mm}^3$ (the size in the tile geometry proposal [81]), that means that a SiPM with an area of at least $9 \times 9 \, \text{mm}^2$ is required.

To better understand the aspect ratio dependence of the time resolution, we measured the pulse amplitude and rise time for different aspect ratios, as shown in Figure 5.5. It can be seen that a larger aspect ratio results in a somewhat larger amplitude and faster rise time, which helps to explain why the time resolution is better for configurations with a larger aspect ratio.
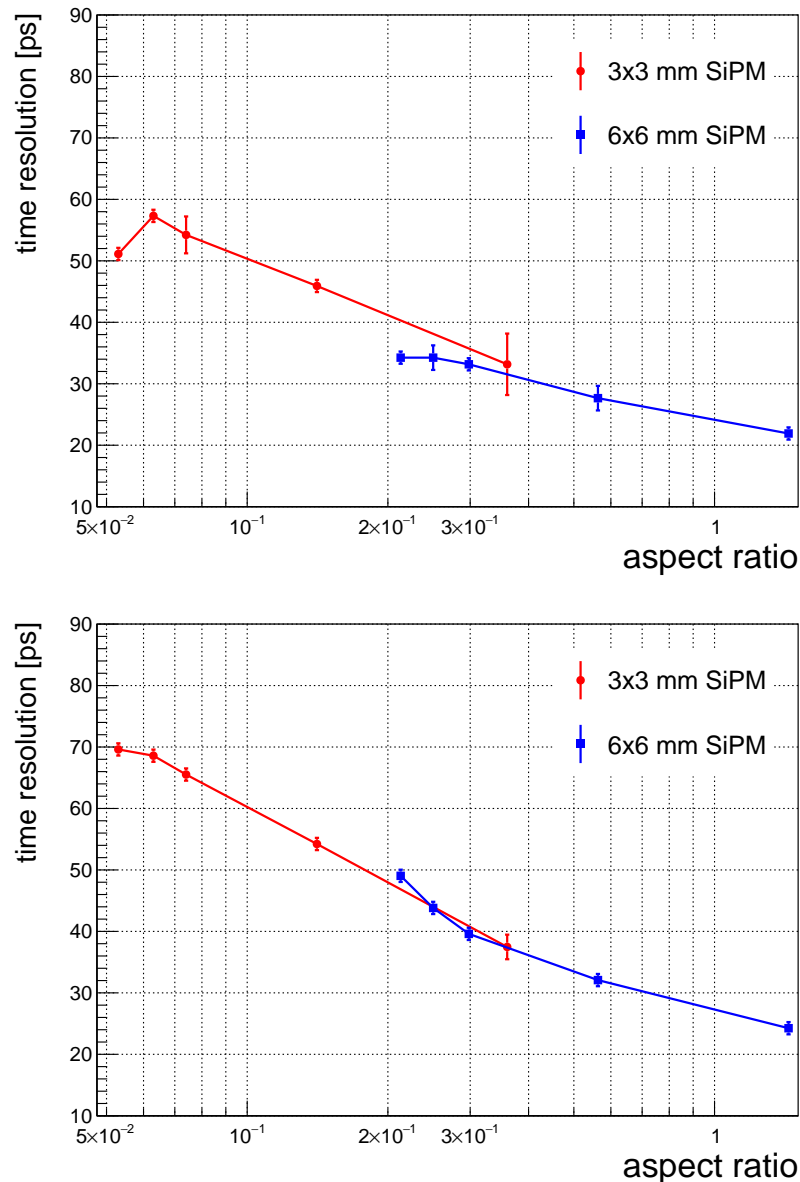
Figure 5.4: Time resolution of an LYSO+SiPM sensor as a function of the aspect ratio of the SiPM and LYSO sensors (defined as the surface area of the SiPM divided by surface area of the LYSO tile). The top plot is the time resolution for a beam particle hitting the center of the LYSO tile, and the bottom plot is the average time resolution for all impact points across the entire LYSO tile.
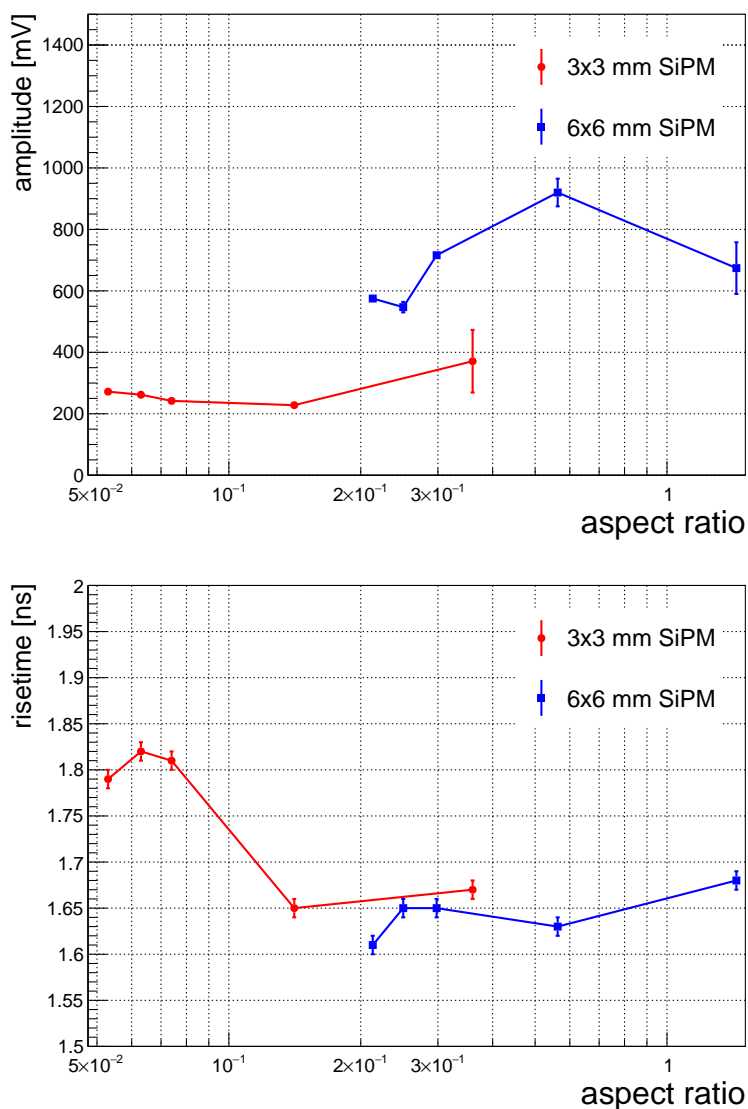
Figure 5.5: Pulse amplitude (top plot) and risetime (bottom plot) as a function of the aspect ratio of the SiPM and LYSO sensors for a MIP hitting the center of the LYSO tile. For the last point (with an aspect ratio = 1.44) in the amplitude plot, the amplitude needs to be scaled up by a factor of about 2.9 as that point was measured with a lower overvoltage on the SiPM compared to the other points (2.9 is estimated according to the gain vs. overvoltage relationship of the SiPM being used).

5.2 Geant4 simulation of light propagation and collection for different sizes of LYSO and SiPM sensors

In order to understand the timing performance for the different LYSO and SiPM size combinations from the scintillation light propagation point of view, a simulation of MIPs hitting the LYSO+SiPM sensor and of the resulting scintillation light

emission, propagation, and collection by the SiPM has been performed using the GEANT4 simulation toolkit [85].

In the simulation, the scintillator LYSO:Ce is defined as having a 99.81% mass fraction of $Lu_{1.8}Y_{0.2}SiO_5$ and 0.19% of Ce, with a density of 7.4 g/cm^3. The refractive index of the LYSO:Ce is set to be 1.81, and the scintillation light yield is set to be 32 photons per keV of deposited energy, with the spectrum of the fast component of the scintillation light emission taken from Figure 5.6. The absorption length of the crystal is set to be 50 cm (zero) for light with wavelength larger (smaller) than 400 nm. The scintillation decay time of the fast component is set to be 41 ns. A 0.1 mm thick layer of optical grease is put between the LYSO:Ce and the SiPM. The grease material is defined as silicone (C_2H_6SiO), with a refractive index of 1.50. The window of the SiPM is simulated by a 1 mm thick layer of epoxy resin ($C_{21}H_{25}ClO_5$), The simulation process ends at the resin layer, and the photons that enter that layer, weighted by the quantum efficiency of the SiPM, are counted as photoelectrons (PE), and the corresponding PE current is calculated.

The absorption length of the resin layer is set to be very small such that no light can escape once it enters that layer. The LYSO:Ce surface is modeled to be not perfectly polished, in which the normal vector at each point ($\vec{m}$) is the nominal one ($\vec{n}$, with size = 1.0 and perpendicular to the macrosurface of the LYSO crystal) plus a small vector ($\vec{r}$) with size r and a random direction: $\vec{m} = \vec{n} + \vec{r}$. The value of r is taken to be a small value (0.01) to simulate the relatively well polished LYSO crystal used in the test beam.

Figure 5.7 shows a sketch of the simulated system. The MIP particle entering the LYSO:Ce sensor is a 1 GeV electron. Standard electromagnetic processes (scattering, ionisation, bremsstrahlung, Cherenkov radiation) are simulated for the electron. Standard optical photon processes (boundary processes, absorption, Rayleigh scattering) are simulated for optical photons generated by scintillation or Cherenkov light emission. Inside the LYSO:Ce crystal, the generated optical photons bounce around and eventually arrive at the SiPM area and get absorbed by the SiPM, or escape the detector, or get absorbed by the LYSO:Ce crystal. Figure 5.8 shows the photon paths from some example simulated events.

The readout electronics chain of the SiPM is not simulated in this study. Instead, the photon current, defined as the number of photons arrived at the SiPM area in every time interval, is recorded for each event and used to approximate the output signal from test beam. In order to mimic the photon detection efficiency of the
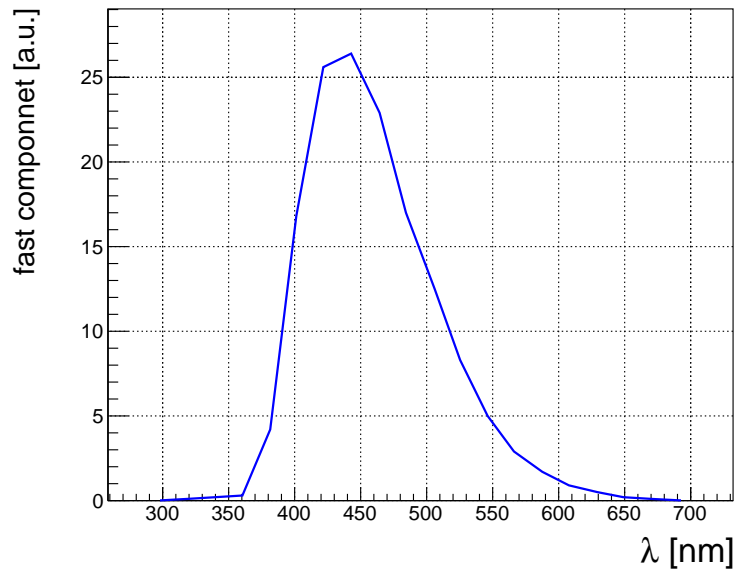
Figure 5.6: Fast component spectrum of the LYSO:Ce scintillator used in the simulation [86].

SiPM, we randomly pick only 20% of the photons arrived at the SiPM in the above photon current reconstruction procedure (the 20% is tuned such that the total number of photoelectrons obtained in simulation matches the measurement from the test beam with the same LYSO and SiPM size). Figure 5.9 shows the photon current pulse (averaged over many pulses) for a 1 GeV electron entering the center of the $8 \times 8 \times 4\,\mathrm{mm}^3$ LYSO:Ce and collected by a $3 \times 3\,\mathrm{mm}^2$ SiPM size. The plot shows two beam impact situations: one when the beam particle goes through the center (X = 0, Y = 0) of the LYSO:Ce surface, and the other one when the beam particle impact point is at the edge (X = 4 mm, Y = 0 mm) of the LYSO:Ce surface. It can be seen that the amplitude of the pulse for edge impact is only about 70% of the center impact case, since for the edge impact case, the optical photons need to go through more bounces inside the LYSO:Ce crystal in order to reach the SiPM. Therefore, the chance of the photon being absorbed by the crystal or escaping the detector is larger in the edge impact case relative to the center impact case.

In order to simulate the timing performance for different aspect ratios of the LYSO:Ce crystal and the SiPM sensor, a scan of a $4 \times 4\,\mathrm{mm}^2$ SiPM with various sizes of LYSO:Ce ($4 \times 4\,\mathrm{mm}^2$, $6 \times 6\,\mathrm{mm}^2$, $8 \times 8\,\mathrm{mm}^2$, $10 \times 10\,\mathrm{mm}^2$, $12 \times 12\,\mathrm{mm}^2$, all 4 mm thick) has been performed. Figure 5.10 shows the photon current pulse amplitude as a function of the aspect ratio, which shows a very clear increase of the amplitude
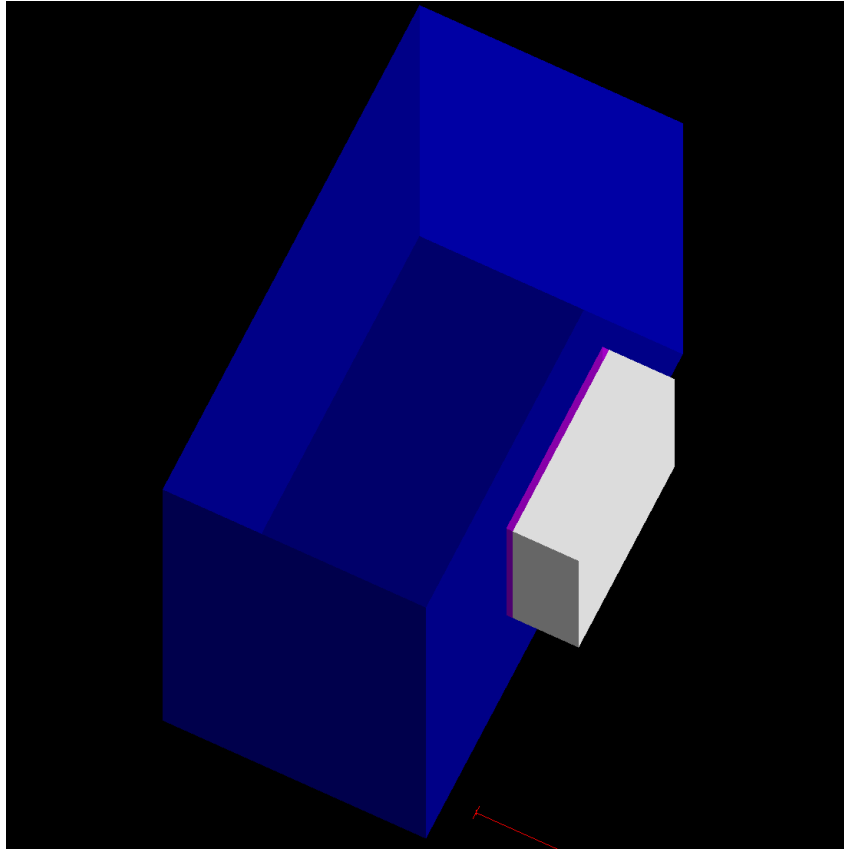
Figure 5.7: Overview of the simulated detector system, which contains a LYSO:Ce scintillator (blue), a layer of optical grease (purple), and a resin SiPM window layer to collect photons (white). An $8 \times 8 \times 4 \, mm^3$ LYSO:Ce crystal and a $3 \times 3 \, mm^2$ SiPM is shown in this sketch, with the SiPM glued to the $8 \times 8 \, mm^2$ face of the LYSO:Ce crystal. The direction of the entering MIP (the Z direction) is perpendicular to the $8 \times 8 \, mm^2$ face of the LYSO:Ce crystal (which is in the X-Y plane).

as the aspect ratio increases, similar to the results obtained from the test beam (Figure 5.5).

Figure 5.11 shows the photon current pulse rise time as a function of the aspect ratio, from which we can see that the predicted rise time continues to decrease from about 1.8 ns for a 0.1 aspect ratio, to only 0.3 ns as the aspect ratio goes to 1.0. However, from the test beam data as shown in Figure 5.5, when the aspect ratio is large enough, we see that the rise time no longer drops. This difference is due to the fact that the simulation only includes the rise time due to the scintillation process and photon propagation time, but in the test beam case, the SiPM itself also contributes to the output signal rise time (for example due to the avalanche formation process).

The time offset, defined as the difference between the time the pulse reaches 10%
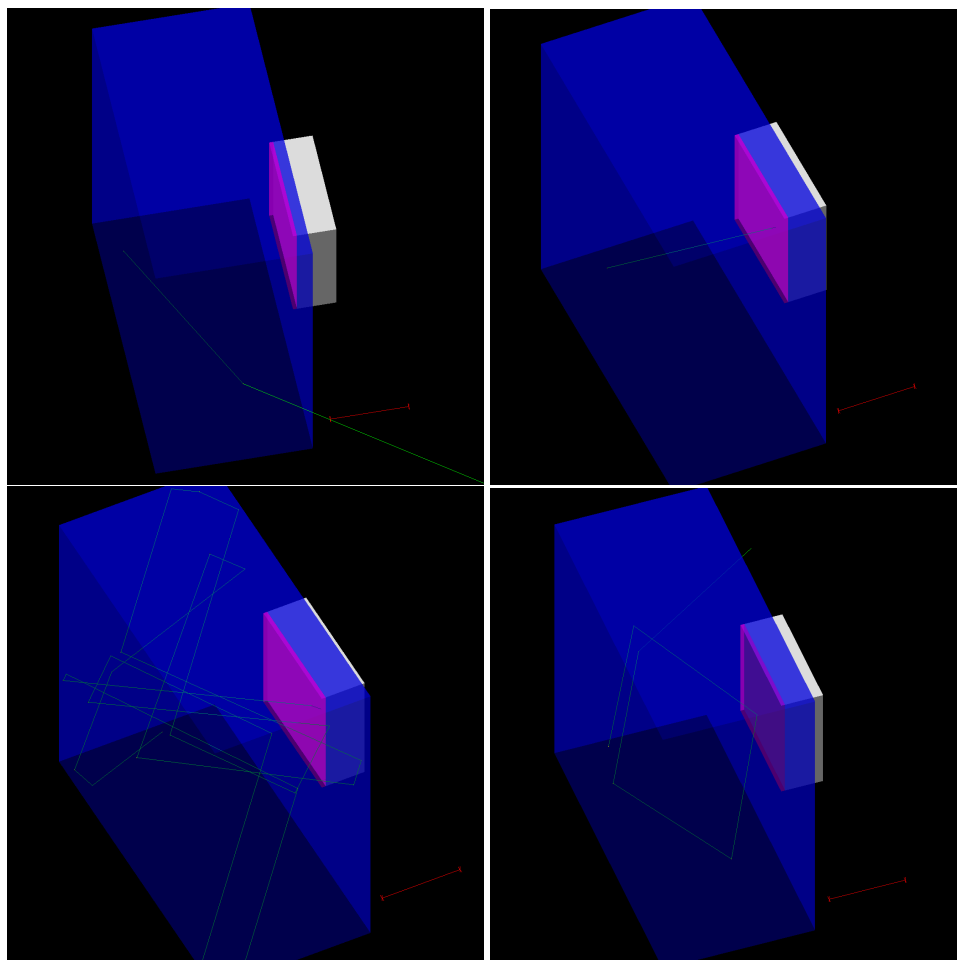
Figure 5.8: Some examples of possible paths of an optical photon (green line) inside the LYSO:Ce crystal: top left: the photon escapes the detector; top right: the photon gets absorbed by the SiPM; bottom left: the photon goes through multiple bounces and finally reaches the SiPM area and gets absorbed by the SiPM; bottom right: the photon goes through multiple bounces and eventually escapes the detector.

of its maximum amplitude for a MIP hitting the edge of the sensor, and the corresponding time for a MIP hitting the center of the sensor, has also been measured as a function of the aspect ratio, as shown in Figure 5.12. The plot shows that, in the case of an aspect ratio near 1.0, the time offset is at the 10 ps level, which is well below the time resolution of the sensor.

5.3 Conclusion

In summary, the timing performance of the LYSO tile+SiPM sensor has been studied with test beams for different sizes of LYSO and SiPM sensors. The results show that, in order to get a good time resolution, the size of SiPM needs to be large
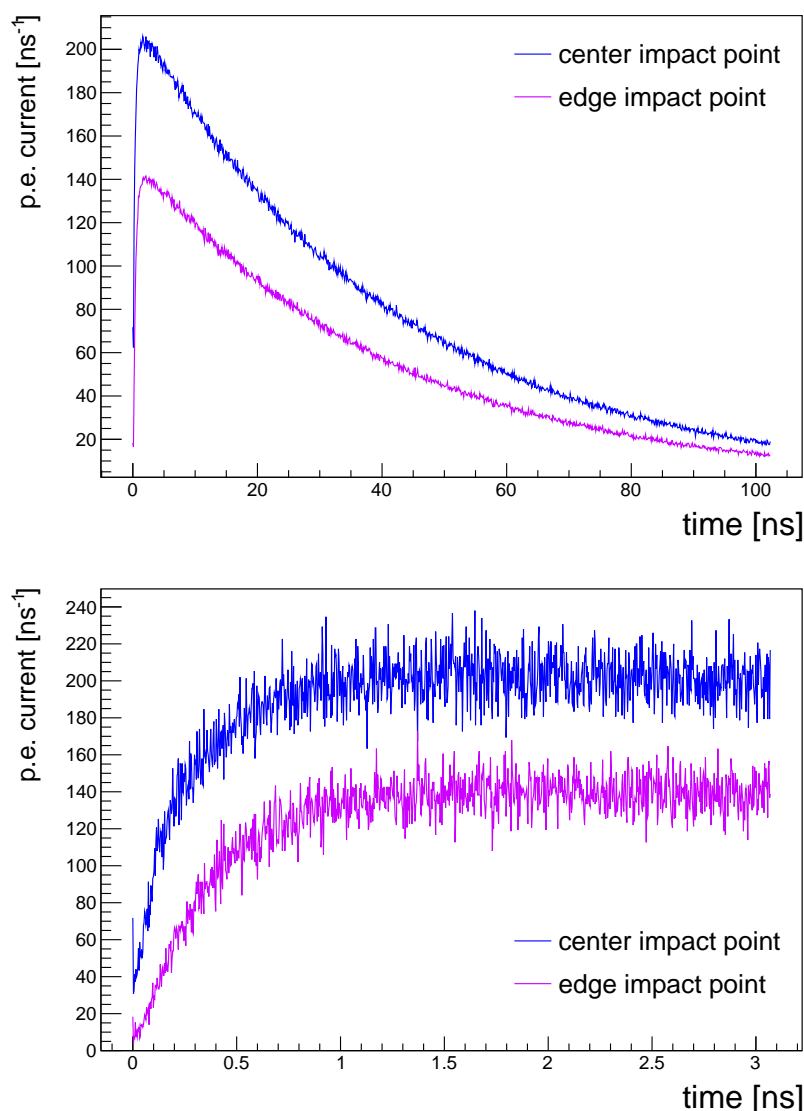
Figure 5.9: An example of the photon current pulse for a 1 GeV electron entering the center of an $8 \times 8 \times 4\,\mathrm{mm}^3$ LYSO:Ce, compared to the case of an electron entering the edge, where the light is collected by a $3 \times 3\,\mathrm{mm}^2$ size SiPM, as a function of time. The plot on the bottom only shows the first 3 ns of the pulse in each case.

enough compared to the LYSO size (a large aspect ratio). The reason is that a large aspect ratio configuration gives a larger amplitude and smaller rise time of the output signal. The same trend has also been seen in the light propagation simulation with GEANT4. In addition, in the small aspect ratio case, the time offset between edge impact and center impact of the MIP on the LYSO sensor becomes large (of the order of a few hundred ps) such that a dedicated impact point correction with a precision measurement of the MIP position would be required.
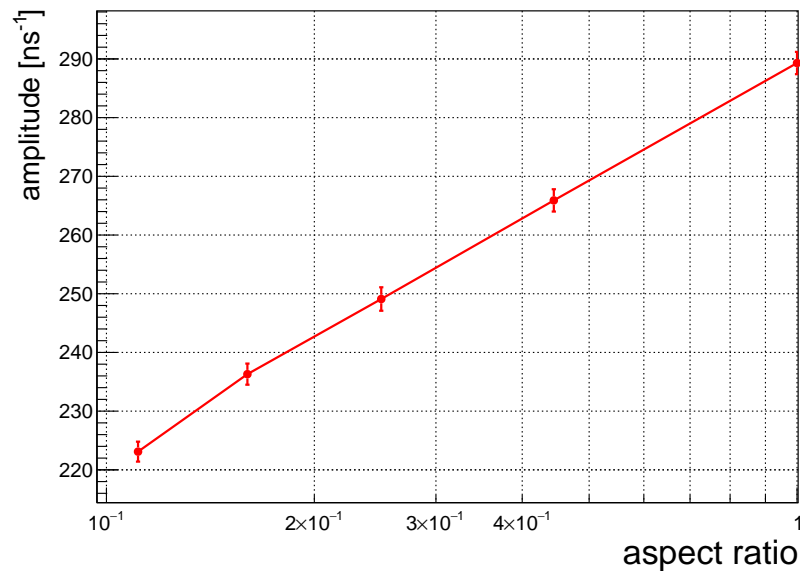
Figure 5.10: Amplitude of the photon current pulse as a function of the aspect ratio of the SiPM and LYSO:Ce sensor.

In particular, the aspect ratio in the proposed BTL tile geometry, with an LYSO tile size of $11 \times 11 \times 4\,\text{mm}^3$ and a SiPM size of $4 \times 4\,\text{mm}^2$ [81] has been demonstrated to be too small to achieve the 30 ps resolution required across the entire tile surface. In order to achieve the target resolution, one would have to either reduce the LYSO size (and thus increase the number of readout channels) or increase the SiPM size to an area larger than those which are currently mass produced. As both options to modify the tile design have been deemed not to be sufficiently cost effective, the alternate design with the bar geometry has been adopted instead.

As shown in Figure 5.13, the final BTL design, based on the crystal bar option, has since been proven to be able to meet the 30 ps precision specification without increasing the number of channels or increasing the SiPM size, and this design shows no (or very little) dependence on the impact point of the MIP on the bar [80].
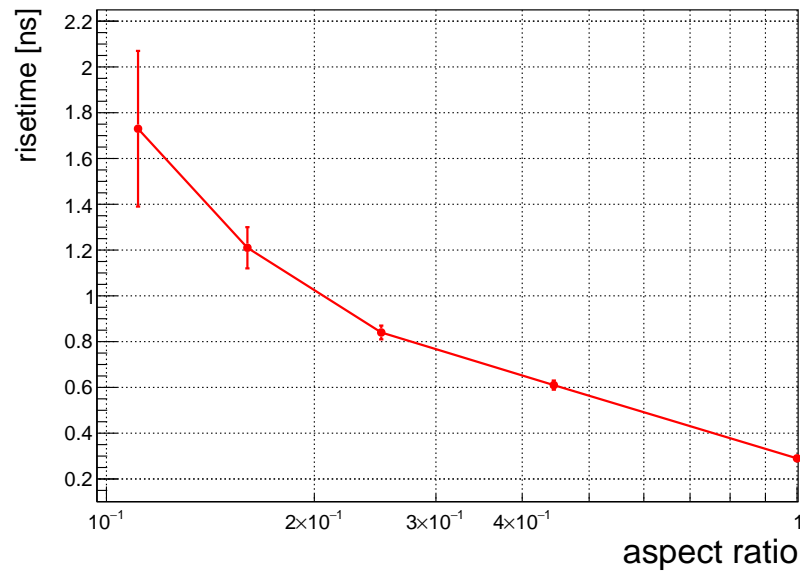
Figure 5.11: Risetime of the photon current pulse as a function of the aspect ratio of the SiPM and LYSO:Ce sensor.
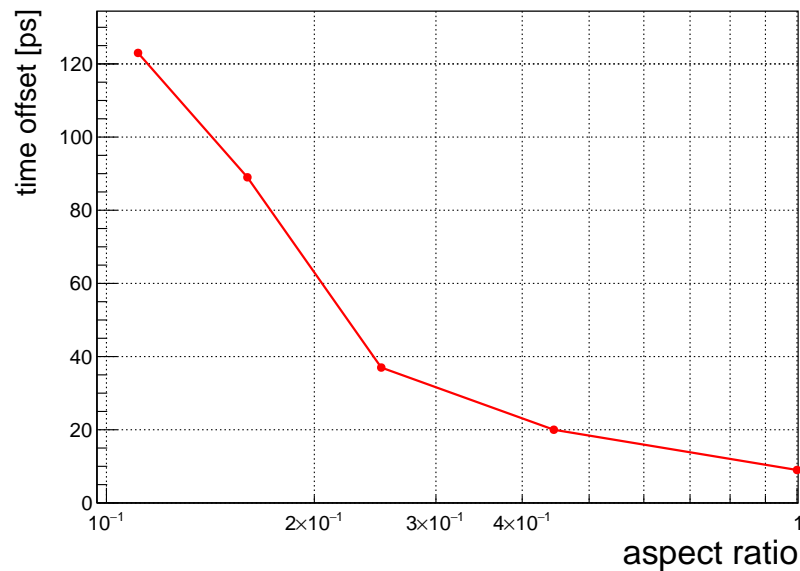
Figure 5.12: Time offset (difference of photon arrival time for edge impact and center impact) as a function of the aspect ratio of the SiPM and LYSO:Ce sensors.
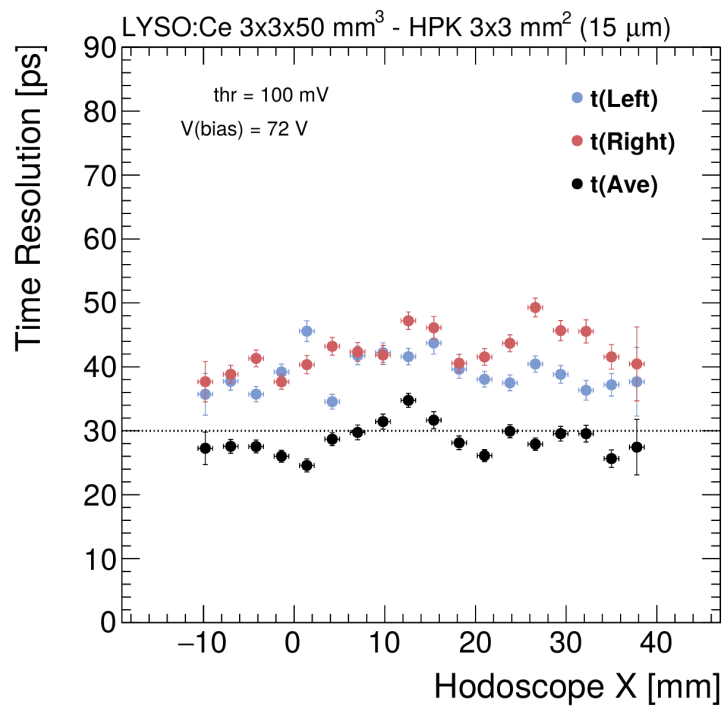
Figure 5.13: Time resolution (measured in a test beam) of a CMS BTL LYSO crystal bar with two SiPMs attached at each end of the bar, measured as a function of the beam impact position along the bar. The plot shows the time resolution from the individual SiPM measurements (blue and red dots), as well as from the combined measurement of the two SiPMs (black dots) based on the average timestamp of the two SiPMs [80].

C h a p t e r   6

ECAL ENERGY CALIBRATION WITH $\pi^0 \rightarrow \gamma\gamma$ EVENTS

The energy of an electron or photon supercluster (SC) can be expressed as a sum of the energy deposited in each ECAL channel of the SC and the energy deposited in the preshower detector, with some corrections for each channel and an overall correction of the total SC energy:

$$E_{e/\gamma} = F_{e/\gamma} \left[G \sum_i S_i(t) C_i A_i + E_{\text{ES}} \right],$$ (6.1)

where:

- E_{ES} is the estimate of the energy deposited in the preshower (for the endcap only), which is a weighted sum of the energies in the two preshower planes.

- A_i is the pulse amplitude of channel i in the cluster, measured in units of ADC counts. The pulse amplitude reconstruction is based on the multifit algorithm, as discussed in Section 4.3.

- $S_i(t)$ is the correction for the change in the channel response over time due to radiation effects (such as color center formation [87]) in the crystal that leads to a drop in the crystal's transparency, counterbalanced by thermal annealing at room temperature which results in partial recovery. The response changes continuously as the luminosity during a data taking run evolves, as well as between runs. Therefore, the correction is derived continually for all channels every 40 minutes with a laser monitoring system which injects light into each crystal and measures the response of each channel. Figure 6.1 shows the change in the response of the crystals over time as the luminosity changes, in several pseudorapidity ranges. For the pseudorapidity range covered by the tracker ($|\eta| < 2.5$), which is most relevant for e/γ physics in CMS, the radiation damage by the end of Run 2 in the Fall of 2018 had caused the average response in a ring of crystals to drop by 12 to 60%, followed by some recovery due to annealing during the shutdown which is now underway.

  The response corrections are then validated with collision data, including a fast validation procedure that uses the invariant mass of $\pi^0 \rightarrow \gamma\gamma$ events (that

provides one validation point in just a few minutes), with an example shown in Figure 6.2; and a slower validation (but at higher energy) that uses the ratio of the energy measurement from the ECAL to the momentum measurement from the tracker, for electrons from W or Z bosons.

- G is the ADC-to-GeV conversion factor (absolute energy calibration). This calibration correction is derived in a region where other impacts such as upstream tracker material or crack between ECAL modules are minimal. For Monte Carlo (MC) events, G is defined such that the reconstructed energy from a 5×5 crystal matrix is equal to the true energy of an unconverted photon. For data, G is defined such that the invariant mass of $Z \rightarrow ee$ from the data matches the mass from the MC.

- $F_{e/\gamma}$ is the supercluster energy containment correction. This correction is used to account for incomplete e/γ energy containment in the supercluster, including losses due to showering in the material in front of the ECAL, energy leaking from gaps and cracks between crystals, or in-time pileup energy pollution in the cluster. This correction varies event by event and is evaluated on the fly during the reconstruction of each SC with a Boosted Decision Tree (BDT) regression. The BDT is trained using simulated photon SCs, with the ratio of the true photon energy to the reconstructed SC energy as the target. The input variables to this correction include the position of the SC (η and ϕ, which is related to the material budget in front of the ECAL), the distance from the SC to the module boundaries (which is related to the gaps and cracks), and the shower shape variables.

- C_i is the inter-calibration constant for channel i. This set of constants is used to correct for any residual response differences among the different channels that are not already corrected by $S_i(t)$ (which only corrects the short term effects), such as the aging of photodetectors due to accumulated charges, and the differences among each channel's relationship between transparency for scintillation light and transparency for laser light.

  Four methods have been developed to derive C_i, and the weighted average of the four sets of C_i are used as the final set of constants: the ϕ-symmetry method, which is based on the fact that the energy flux around each ϕ ring should be uniform, where the C_i's should correct for any non-uniformity; the electron E/p method, which is based on the fact that the energy measured

from the ECAL and the momentum measurement from tracker should be equal, where the C_i's should correct for any discrepancies; the π^0 method, which is based on the fact that the measured diphoton invariant mass from $\pi^0 \rightarrow \gamma\gamma$ events should equal to the true mass of π^0 meson, where the C_i should correct for any differences; and finally the $Z \rightarrow ee$ method, which is similar to the $\pi^0 \rightarrow \gamma\gamma$ method but at a higher energy and with much less statistics.
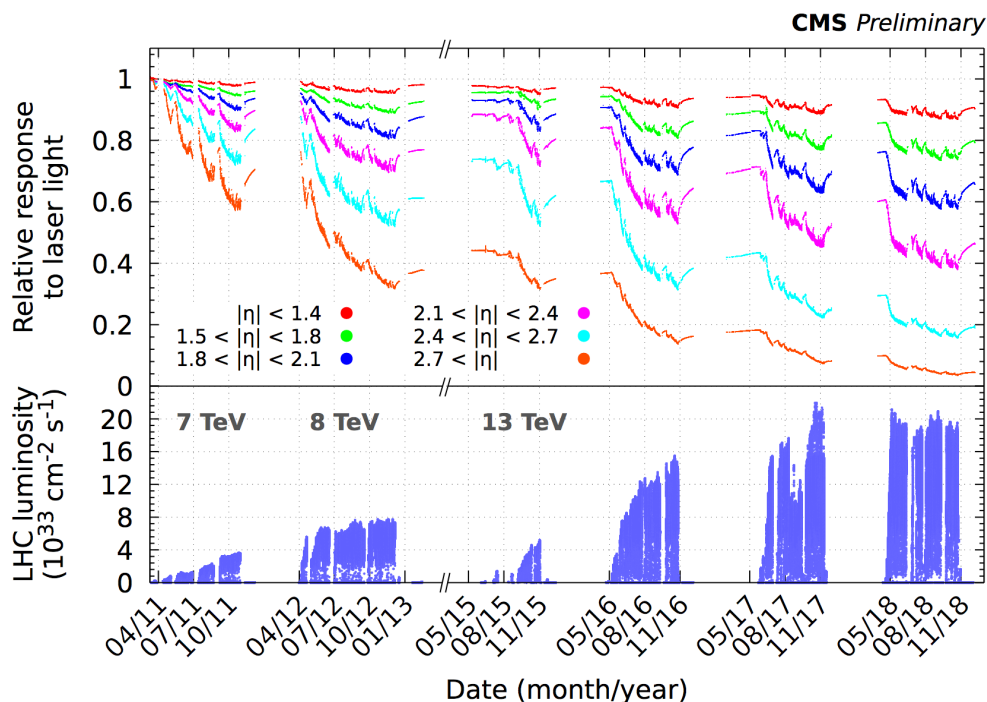
Figure 6.1: (Top) The relative response to laser light measured in bins of $|\eta|$ of the ECAL crystals for the entire Run 1 and Run 2 data taking period from 2011 through 2018, and (Bottom) the corresponding luminosity delivered to CMS [88].

In order to maintain the constant term of the energy resolution (Eq. 4.1) to a sub-percent level, it is critical to derive a set of inter-calibration constants C_i with sub-percent precision. During Run 1, when there were still not enough statistics for inter-calibrations with electrons (using the E/p and $Z \rightarrow ee$ methods), the π^0 method dominated the precision and was the only method which could give an inter-calibration with 0.5% precision (in EB). In Run 2, as the statistics for electrons became larger, and the π^0 event reconstruction became more challenging for reasons that will be discussed later, the precision of the electron methods became similar to the precision of the π^0 method (both at around the 0.5% level in EB), and therefore a
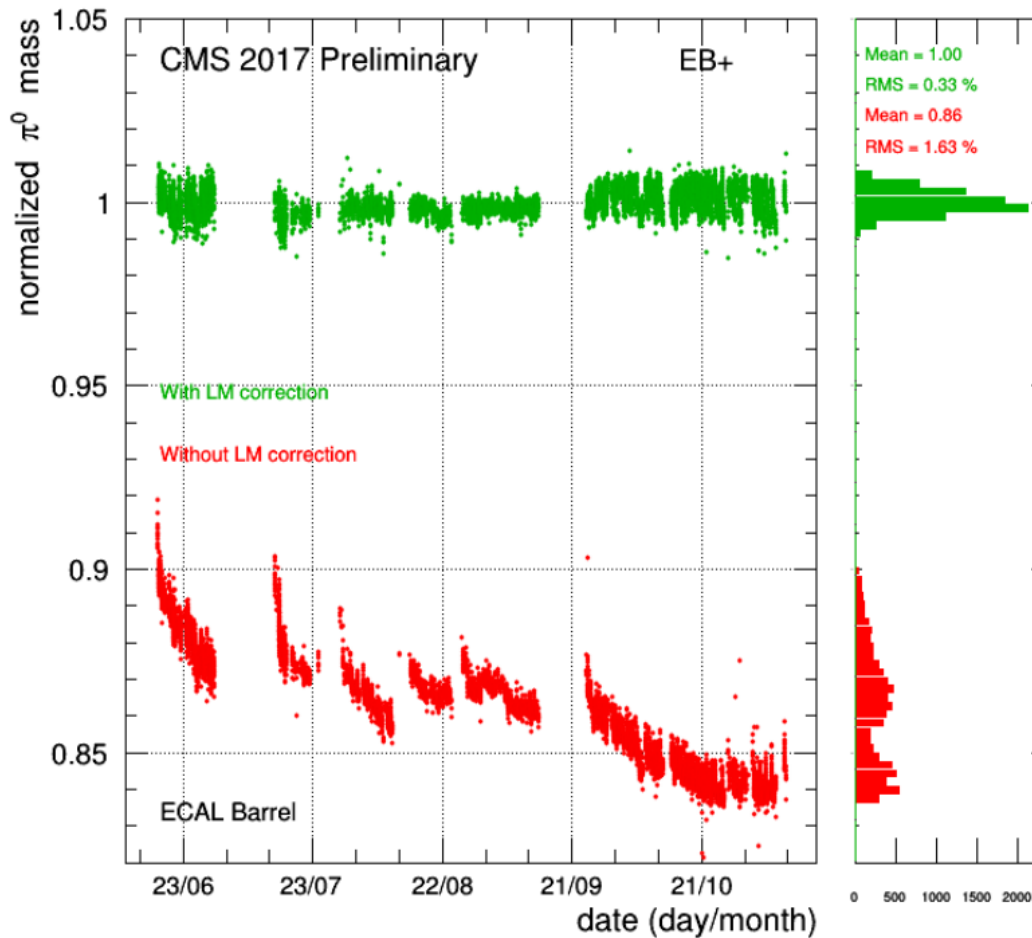
Figure 6.2: Evolution of the invariant mass of the reconstructed $\pi^0 \to \gamma\gamma$ events (normalized by the π^0 meson mass 135 MeV) before and after applying the laser correction to the response channel-by-channel during 2017 data taking [89].

combination of all methods is used for the determination of the final inter-calibration constants. Comparisons of the precision of different inter-calibration methods in the Run 1 and Run 2 data taking periods can be seen in Figure 6.3.

The remainder of this Chapter focuses on the inter-calibration with the $\pi^0 \to \gamma\gamma$ method, including the details about the data stream used, the event reconstruction, the procedure to derive the constants C_i, the results with Run 2 data, and the prospects for the future of this calibration method.

6.1 The π^0 trigger stream

The inter-calibration of the CMS ECAL with $\pi^0 \to \gamma\gamma$ decays relies on the reconstruction of a diphoton invariant mass ($m_{\gamma\gamma}$) peak around the π^0 mass for
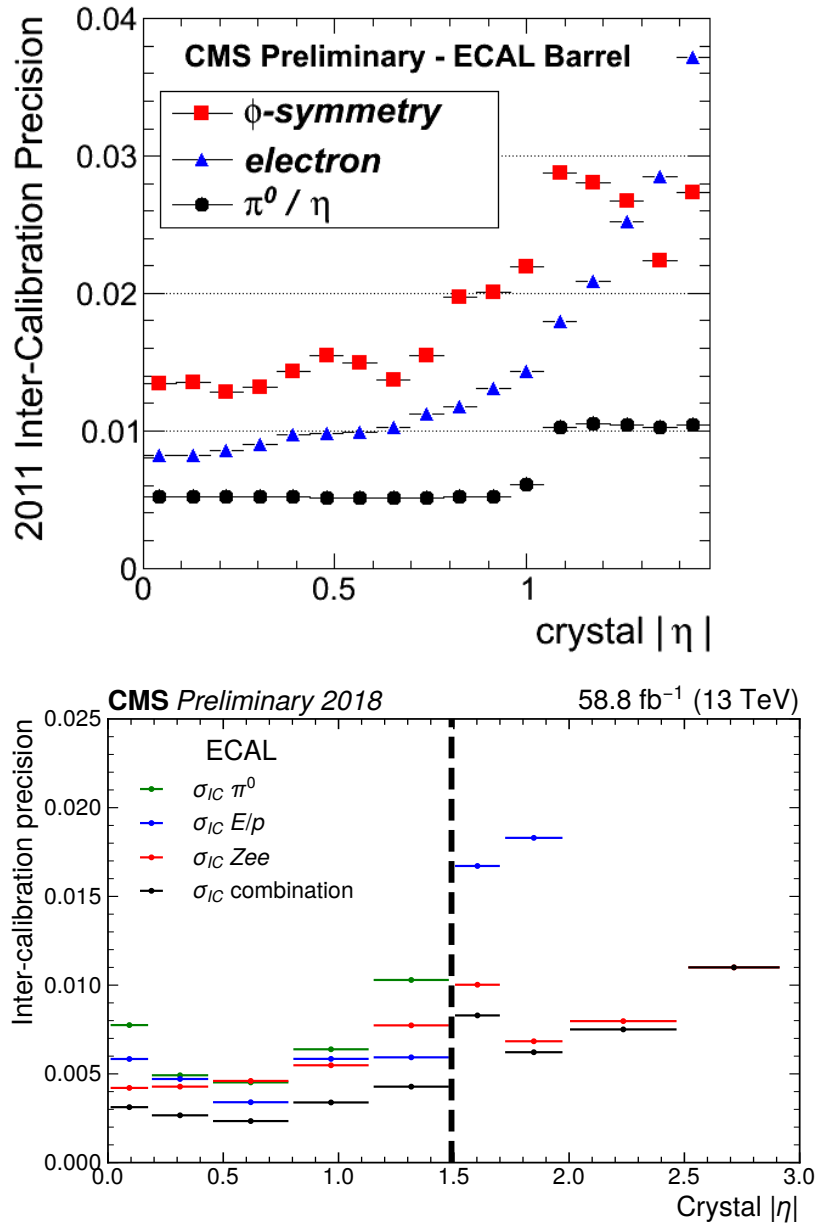
Figure 6.3: Inter-calibration precision obtained from the various methods discussed in the text and their combination (evaluated by comparing the C_i differences from the different methods), as a function of the η of the crystal, for the data taken during Run 1 (top plot) [90] and Run 2 (bottom plot) [91].

each crystal in the ECAL, with enough statistics for each crystal to precisely fit the reconstructed π^0 mass. In order to collect such a huge amount of events, a dedicated HLT stream called the π^0 stream is implemented in addition to the main output stream for physics. The π^0 stream uses only about 2% of the total bandwidth of

writing to storage (about 40 MB/s out of 2 GB/s) at an event rate about 20 times higher than the main output stream (about 20 kHz compared to 1 kHz for the main stream).

The high trigger rate is achieved by looking for π^0 objects from an OR of many L1 trigger algorithms (involving jet, e/γ signatures), as the π^0 is very often produced from the decay chain of hadrons, and it is easy to find a π^0 near an L1 jet seed or from an event with a high p_T electron or photon which often has associated jets.

In fact the rate was so large in 2016 that only a fraction of the events that passed the π^0 trigger were saved (this is called a trigger pre-scale). During 2016, the pre-scale was about a factor of 5 (which means only 1 out of 5 events passing the trigger selection were saved). In order to reduce the time between calibration points with sufficient statistics for all regions of the ECAL, the total 20 kHz bandwidth allocation was further divided into four trigger paths: two paths for $\pi^0 \to \gamma\gamma$ events in EB and EE and two paths for $\eta^0 \to \gamma\gamma$ events in EB and EE, where the η^0 decay paths have a lower rate but produce photons with a higher energy on average leading to a mass peak at 549 MeV, compared to the 135 MeV peak for the π^0 decays. The selection criteria for the EB paths were tuned to have a rate of roughly 7 kHz each for π^0 and η^0 decays (with a π^0 rate a bit higher than the η^0 rate), while the criteria for the EE paths were tuned to have a rate of about 3 kHz each during 2016.

For the rest of this Chapter only π^0 events are discussed, as the η^0 events were not used for the actual calibration during Run 2. This was due to the lower rate and more complicated decay modes of the η^0, which led to a mass peak that was not as well reconstructed as the π^0 peak.

Starting from 2017, the selection criteria in the trigger were optimized (as discussed in the next section) with higher selection efficiency but lower rate, so the pre-scale was removed.

The π^0 trigger reconstructs small clusters of 3×3 crystals, and then selects pairs of such clusters that have an invariant mass consistent with the π^0 meson. The reconstruction and selection of the clusters as well as the selection of the pairs of clusters are similar to what is done in the offline reconstruction and selection (to be discussed in the next section). Events with at least one pair of clusters passing the selection criteria have the raw digitized pulses for the ECAL and preshower channels saved to storage, while information of all other subdetectors is discarded. As a result, the event size in the π^0 trigger is only about $1/1000^{th}$ of the event size

in the main HLT stream.

It is also important to note that, due to the limited ECAL readout bandwidth and storage, there is a data filtering process performed with a Selective Readout Processor (SRP) to make sure that the amplitude of each ECAL channel's signal is large enough to be sent from the on-detector electronics to the upper level stream processing and storage (L1, HLT, etc.) subsystems in the trigger. As a simple calculation, the pulse of each channel is digitized into 10 time slices and saved in 24 bytes, so that the ECAL event size would be about 1.8 MB if all 76,000 ECAL channels were read out. In reality the average ECAL event size gets reduced to less than 2 kB.

The threshold (called the zero-suppression or ZS threshold) was 4.5 (6.5) ADC counts for all EB (EE) channels in 2016, where one ADC count corresponds to about 40 (60) MeV in EB (EE). The zero-suppression could cause incomplete clusters for some of the e/γ objects that we are interested in, and this would affect the energy reconstruction precision. In order to deal with this issue, a so-called "high interest threshold" (HTH) was included in the trigger that reads out the entire 15×15 matrix of channels surrounding the 5×5 trigger tower of the cluster if the energy sum of the 5×5 tower is higher than the HTH. If the 5×5 energy is smaller than the HTH but higher than another threshold called the "medium interest threshold" (MTH), then only the 5×5 channels are fully readout and the surrounding channels are zero-suppressed. For an even lower 5×5 energy, all channels are zero-suppressed.

During 2016, the HTH and MTH were 2.5 GeV and 1.5 GeV, respectively, and the HTH, MTH as well as the ZS threshold were gradually increased as the data taking progressed, in order to deal with increasing luminosity. As we will see in the next section, the energy of the photon clusters from π^0 decay often fell below the HTH or MTH threshold, and as a result the clusters saved for offline reconstruction usually had some channels not read out, and this presented one of the main challenges in the π^0 reconstruction.

6.2 π^0 event reconstruction

The mass of the π^0 meson is rather low (135 MeV), and in order to survive the ZS threshold from ECAL, the energy of the photons from $\pi^0 \to \gamma\gamma$ needs to be at the GeV scale. As a result the two photons from the π^0 usually have a very small opening angle. In the meantime the π^0 can not be too boosted, since the two photons are completely merged in this case, and indistinguishable with the ECAL granularity. As such, a special clustering algorithm is used to reconstruct the photons from π^0

decay, which is different from the clustering algorithm used for high p_T photons and electrons.

To minimize the electronic noise contribution to the reconstructed photon energy, and also considering that the two photons from π^0 decay have a small opening angle, the reconstruction of π^0 photons is based on 3×3 clusters, rather than 5×5 as is used for high p_T photon clusters. The clustering begins by sorting all ECAL channels based on their energy, and the channel with the highest energy is used as the seed for the leading photon (γ_1). A 3×3 matrix of channels centered on the seed channel is then defined as the cluster for γ_1. Once the cluster for γ_1 is formed, all of the channels belonging to it are removed from further consideration. The procedure then continues with the highest energy channel from among the leftover channels being chosen as the seed for the 2nd (3rd, 4th, ...) 3×3 cluster for γ_2 (γ_3, γ_4, ...). The procedure stops when the energy of the seed from the leftover channels falls below 0.5 GeV.

The outcome of this procedure is a list of photon clusters which in principle should each have the size of a 3×3 crystal matrix. However, for a photon cluster that gets clustered later (for example γ_2, γ_3, ...), some of its 3×3 channels might have already been claimed by photon clusters that were clustered earlier, and thus the number of channels in a photon cluster is not necessarily 9. In an extreme case, a photon cluster might have only one channel when all the 8 channels (except the seed channel) have been claimed by four other photon clusters around that cluster. Figure 6.4 shows the four possible results of clustering when there are two photons in the event. In addition, the zero-suppression will also cause zero energy being read out for some of the channels, which further reduces the number of channels in a photon cluster.

Once the list of photon clusters is reconstructed, the energy of each cluster is reconstructed in the same way as in Eq. 6.1, with all the corrections that are available in the existing database (which are derived from the corresponding calibrations). The two exceptions are the containment correction $F_{e/\gamma}$ and the inter-calibration constants C_i.

The purpose of π^0 reconstruction itself is to update the C_i, and this is achieved in an iterative way: for the starting iteration, the process takes the C_i from the database, and after the reconstruction of π^0 events, the procedure updates the C_i based on the reconstructed π^0 mass, which is then used as the input for the next iteration (further details of this procedure are discussed in the next section).
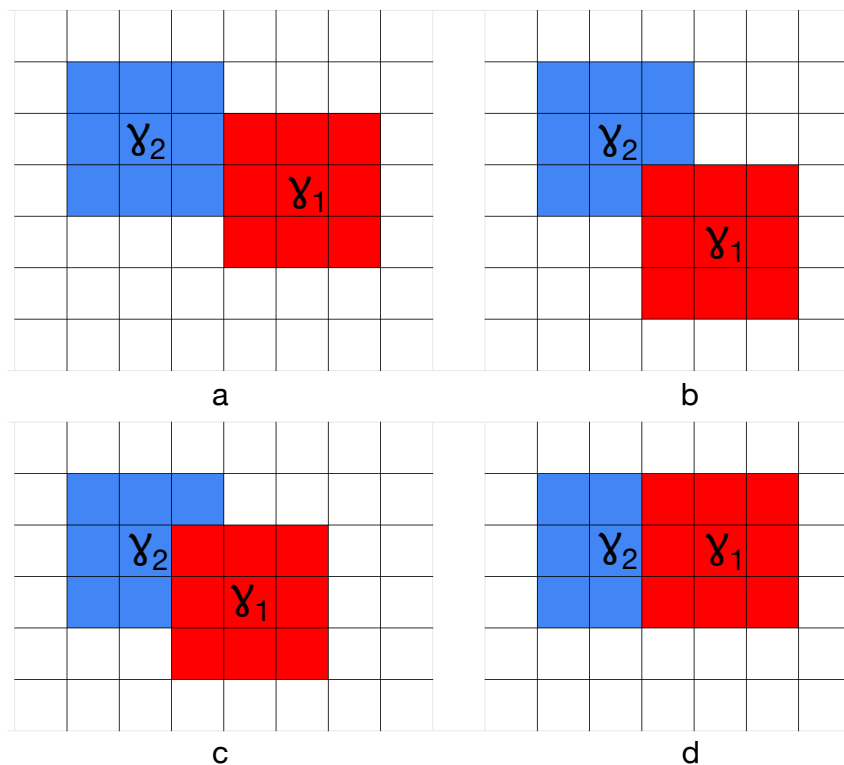
Figure 6.4: Illustration of two reconstructed photon clusters from a π^0 decay. Each cell represents an ECAL channel in this sketch. γ_1 (the red cells) is the leading photon and gets clustered first, and γ_2 (the blue cells) is the subleading photon and is clustered from the leftover channels. Figure (a) shows a non-overlapping case and the other three figures show overlapping cases. The energy from the overlapping channels is assigned to γ_1.

The cluster containment correction $F_{e/\gamma}$ used for this special 3×3 clustering algorithm is fundamentally different compared to the containment correction used for high p_T photon or electron clusters. The reason is that for this low energy, small cluster case, the containment correction needs to correct for the additional factors that cause incomplete clusters, such as the zero-suppressed energy (which is not a problem for high p_T clusters), the missing channels being claimed by previously clustered clusters, etc.

Therefore, the containment correction gets re-derived specifically for this usage with a similar BDT regression method. In order to take overlapping clusters into account, the containment correction for the leading photon and sub-leading photon are trained separately. Furthermore, the number of channels in the photon cluster is also used as an additional input for the regression training, in order to take into account zero-suppression effects as well as overlapping cluster effects. Figure 6.5 shows the

energy response (the ratio of the reconstructed energy to the MC true energy) before and after applying the containment correction. We see that after the containment correction, the reconstructed energy gets closer to the true energy and the energy resolution also improves. However, the very low energy tail, which includes some of the extreme cases where many of the channels in the cluster get zero-suppressed or claimed by other clusters, could not be recovered by the containment correction derived with this regression method.
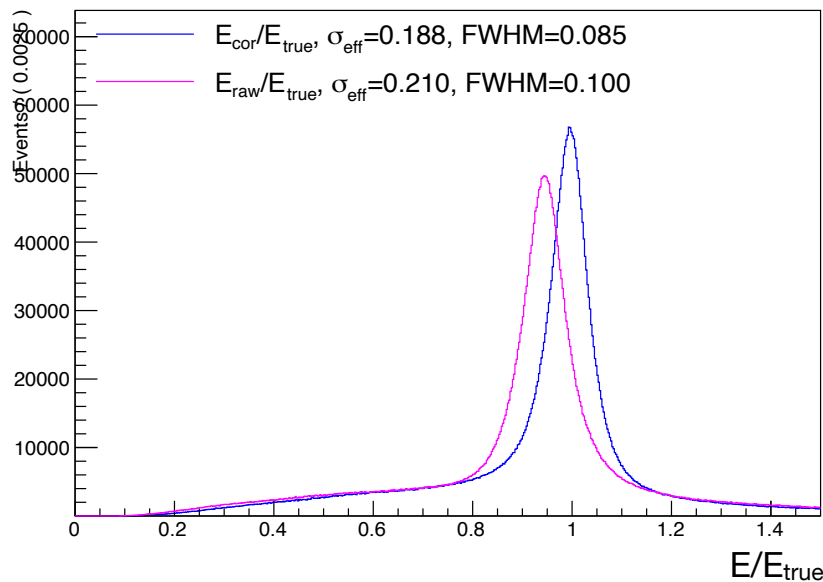
Figure 6.5: Distribution of the ratio of the reconstructed 3×3 photon cluster energy to the MC true energy before (E_{raw}/E_{true}, purple) and after (E_{cor}/E_{true}, blue) applying the containment correction to the cluster. Also shown in the legend are the σ_{eff} and the full width at half maximum (FWHM), where σ_{eff} is defined as the smallest x interval that contains 68.3% of the distribution.

In order to reconstruct a good π^0 mass peak, a set of requirements is applied to the photon clusters before they can be paired to form a π^0. Figure 6.6 shows the 2D distribution of the reconstructed photon p_T and the true p_T of the generated photons from simulated π^0 events. We see that for a large fraction of the photons, the reconstructed p_T is very small (close to zero) regardless of the true p_T. Therefore, a cut of $p_T > 0.9$ GeV is applied to all photon clusters. A basic shower shape selection is also applied to the cluster, which requires the highest energy among all possible 2×2 matrices in the cluster (called E_4) to be greater than 90% of the total energy of the cluster (E_9). In addition, a cut on the number of crystals (channels) in the photon cluster is also applied, in order to remove extreme cases of cluster

overlap or zero-suppression that cannot be recovered by the containment correction.

Figure 6.7 shows the number of crystals (Nxtal) distribution for the leading and sub-leading photon cluster in EB and EE. We can see that a large fraction of the photon clusters have less than 9 crystals, and the situation is much worse for EE clusters. A cut of Nxtal > 6 is applied to all reconstructed photon clusters. Also shown in Figure 6.7 are the diphoton mass spectrum before and after applying this Nxtal cut. It can be clearly seen that without this Nxtal cut, no π^0 peak can even be seen in the endcap. In the barrel, the mass peak resolution as well as the signal over background ratio in the spectrum is significantly improved after the Nxtal cut.

One thing to note is that in the π^0 trigger stream of 2016, there was no such Nxtal cut applied to the stream. As a result, as shown in Figure 6.7, a large fraction of events fail the offline Nxtal > 6 selection, causing the trigger efficiency to be very low in 2016.

In order to improve the trigger efficiency, the same Nxtal > 6 cut was applied to the trigger stream starting from 2017 on, and since the rate got reduced significantly after applying this cut, the trigger prescale, which was set to about 5 in 2016, was completely removed starting from 2017. As a result, the statistics of usable π^0 events was significantly increased starting from 2017 without increasing the rate of the π^0 trigger.
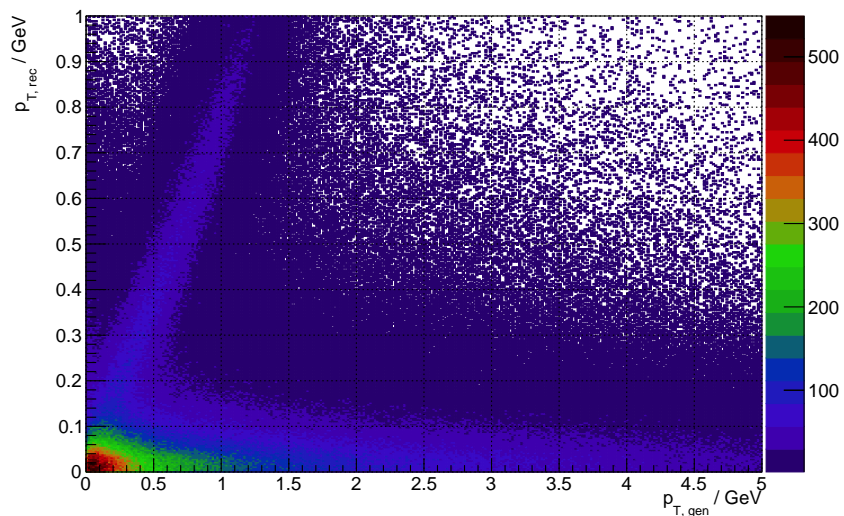
Figure 6.6: 2D distribution of the reconstructed photon cluster p_T (Y-axis) and the MC truth of the generated photon (X-axis) for photons from π^0 decay.
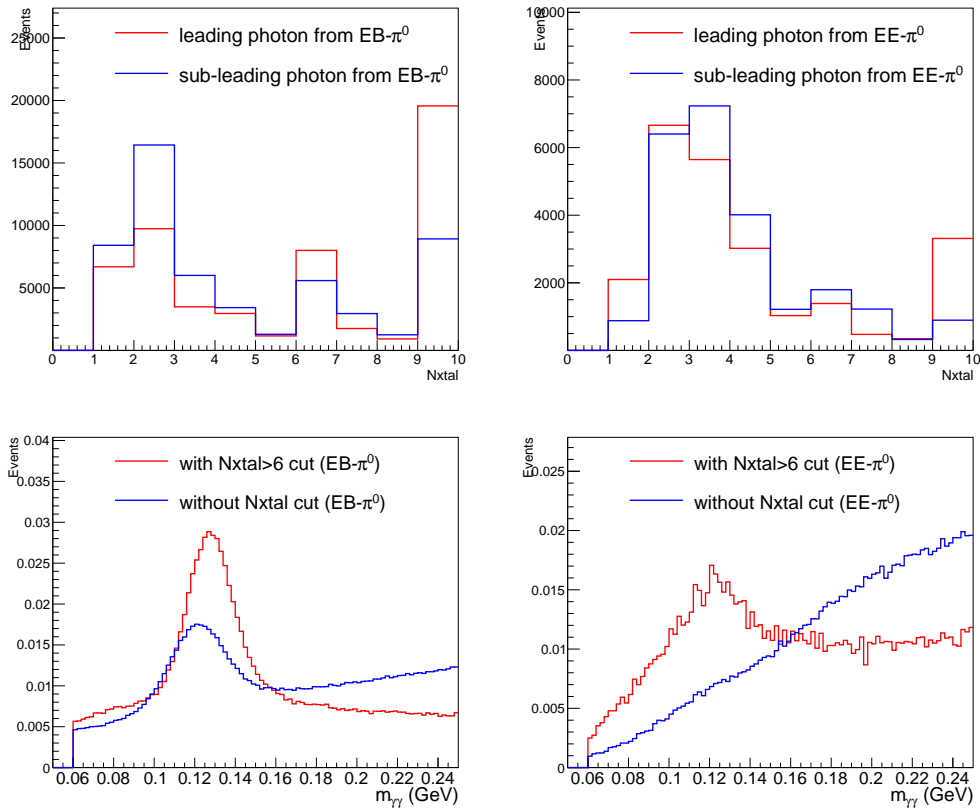
Figure 6.7: Top plots: distributions of the number of crystals (Nxtal) in a photon cluster for the leading and sub-leading photons in EB (top-left) and EE (top-right). Bottom plots: the diphoton mass distribution before and after applying the Nxtal > 6 cut for each of the photon clusters in EB (bottom-left) and EE (bottom-right). The plots are made from QCD multijet MC samples.

The photon clusters in the list passing the above selection are then paired to form π^0 events. First they are sorted based on their seed energy (the same order as γ_1, γ_2, ... in the clustering process). Then, the pairing starts with the first photon, by looking for the next photon that can pair with the first photon which can satisfy some selection to make a π^0 event. The process continues until all photons in the list have been used and no more diphoton pairs can be formed. The selection applied on the diphoton pair includes a selection on the diphoton mass ($m_{\gamma\gamma}$), which needs to be inside the π^0 mass window: $60\,\text{MeV} < m_{\gamma\gamma} < 250\,\text{MeV}$. The π^0 object also needs to be well isolated, for which the scalar sum of the p_T of all other clusters inside a $\Delta R < 0.2$ cone centered on the π^0 object has to be smaller than 50% of the p_T of the π^0 itself. In addition, the p_T of the π^0 object has to be greater than 2 GeV in order to minimize the mis-reconstruction for low p_T objects. This is shown in Figure 6.8

which compares the reconstructed p_T of the π^0 object to the MC truth.
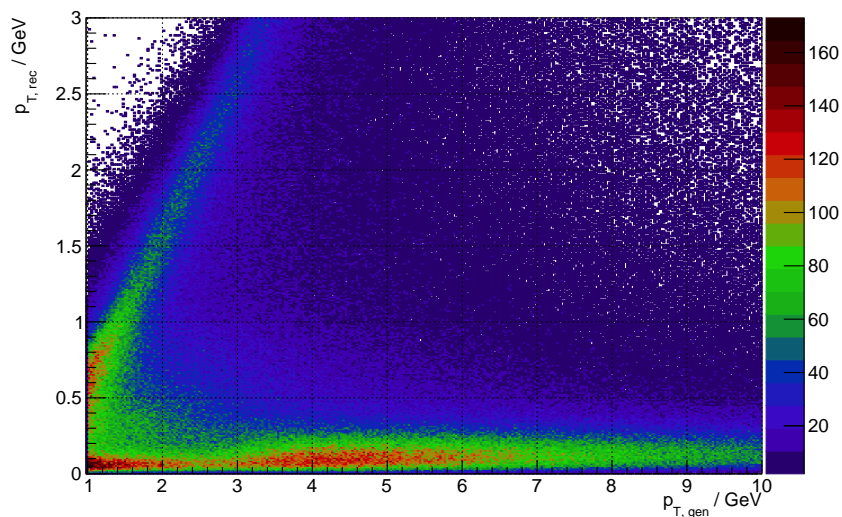
Figure 6.8: 2D distribution of the p_T of the reconstructed π^0 (Y-axis) and the p_T of the generated π^0 (X-axis).

Once the list of π^0 objects has been reconstructed, the invariant mass of the π^0 is then used to derive the inter-calibration constants, for which the procedure is discussed in the next section. As an example, Figure 6.9 shows two example mass spectra of the two photons from the π^0 candidate with 0.1 fb^{-1} of data taken during 2016. The spectrum is fitted with a Gaussian peak (signal) and a second order Chebychev polynomial (background), and the best fit mass of the π^0 is extracted from the fitted value of the mean of the Gaussian function.

6.3 Procedure to derive the inter-calibration constants and results

As mentioned above, the energies of the photon clusters used to reconstruct the π^0 mass are calculated based on Eq. 6.1, which relies on the inter-calibration constants C_i, where the reconstructed π^0 mass peak is again used to update the constants C_i. Therefore, the derivation of the inter-calibration constants C_i is an iterative procedure.

For beginning the first iteration, one could in principle just set the C_i to 1.0 for all channels. However, in order to reduce the number of iterations, we usually take C_i from the database where the values stored usually have been derived during a previous calibration campaign. Furthermore, to update the C_i for a particular crystal i, we record the diphoton mass spectrum ($m_{\gamma\gamma}$) of all π^0 events for which one of
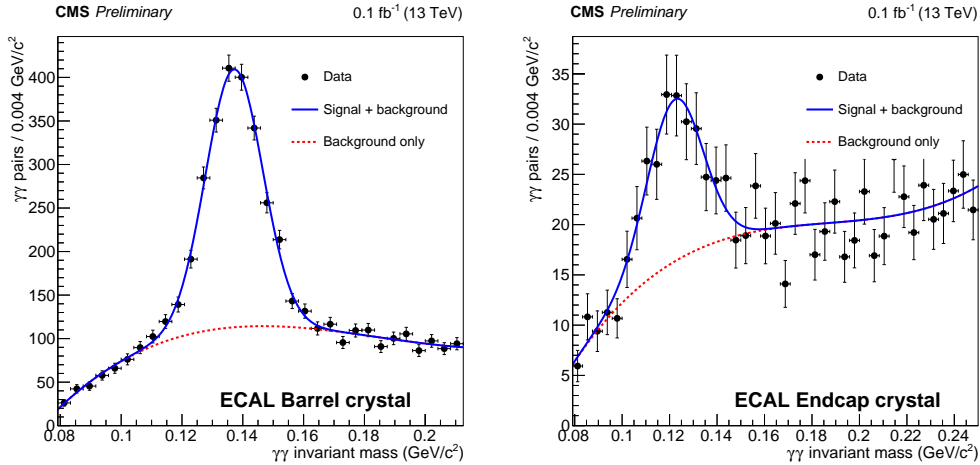
Figure 6.9: Invariant mass of the two photons from a π^0 candidate recorded by one particular ECAL crystal (that crystal belongs to one of the two photon clusters), from 0.1 fb^{-1} of data taken during 2016. The left plot is for an EB crystal and the right plot is for an EE crystal.

the photon cluster's seeds is on that particular crystal. Usually one can get an $m_{\gamma\gamma}$ spectrum with very good statistical precision for every ECAL crystal, with a few fb^{-1} of data. For each crystal i, we then fit the $m_{\gamma\gamma}$ spectrum to obtain the best fit π^0 mass (m_{fit}), and the calibration constant C_i for that crystal is then updated as:

$$C_i = \frac{1}{1+r}, \qquad \text{where} \qquad r = \frac{1}{2}\left[\left(\frac{m_{\text{fit}}}{m_{\text{PDG}}}\right)^2 - 1\right], \qquad (6.2)$$

where m_{PDG} is the true π^0 mass (134.9768 MeV according to the Particle Data Group (PDG) [5]). After one iteration, the C_i of all crystals are updated, and are then used as the input for the next iteration.

The process continues until the C_i of two consecutive iterations converge. The convergence is evaluated by looking at the distribution of the C_i differences between two consecutive iterations for each of many crystals belonging to the same block in an η region, and calculating the RMS of the distribution of differences. Once the RMS of the distribution no longer decreases with more iterations, we conclude that convergence has been reached, and the C_i from the final iteration are then used as the output of the π^0 calibration. Typically about 10 iterations are needed in order to converge.

The statistical uncertainty of the C_i derived by using this procedure with about 10 fb^{-1} data is found to be below 0.2%, which is much smaller than the systematic

uncertainty. To evaluate the total uncertainty of the C_i with the π^0 method, the C_i obtained from this method is compared to the C_i from the other methods (in this case, the other two methods are the E/p and $Z \to ee$ methods as mentioned above). Assuming that the inter-calibration constants obtained from the three methods for a particular crystal are C_1, C_2, and C_3, with uncertainties (to be derived) σ_1, σ_2, and σ_3, and assuming their uncertainties are uncorrelated, then we have:

$$\sigma_{1-2}^2 = \sigma_1^2 + \sigma_2^2$$
$$\sigma_{1-3}^2 = \sigma_1^2 + \sigma_3^2$$
$$\sigma_{2-3}^2 = \sigma_2^2 + \sigma_3^2 \tag{6.3}$$

where σ_{1-2}, σ_{1-3}, and σ_{2-3} are the uncertainty of the $C_1 - C_2$, $C_1 - C_3$, and $C_2 - C_3$, respectively, which are obtained by calculating the RMS of the distributions of $C_1 - C_2$, $C_1 - C_3$, and $C_2 - C_3$ in a given η ring. Solving Eq. 6.3, we get:

$$\sigma_1^2 = \frac{1}{2}(\sigma_{1-2}^2 + \sigma_{1-3}^2 - \sigma_{2-3}^2)$$
$$\sigma_2^2 = \frac{1}{2}(\sigma_{1-2}^2 + \sigma_{2-3}^2 - \sigma_{1-3}^2)$$
$$\sigma_3^2 = \frac{1}{2}(\sigma_{1-3}^2 + \sigma_{2-3}^2 - \sigma_{1-2}^2) \tag{6.4}$$

Figure 6.10 shows the precision of the three methods estimated using the above procedures. As we can see, the three methods give a similar precision in EB (about 0.5%), but for EE, the $Z \to ee$ method gives the best precision.

6.4 Outlook

To further improve the π^0 calibration precision, it will be important to improve the photon cluster and therefore the π^0 event reconstruction. Currently the energy loss from zero-suppression is one of the big challenges for this calibration method, and in order to improve that, a more dedicated containment correction will be needed to recover the zero-suppressed energy.

Fortunately for the HL-LHC in the future, the ECAL endcaps will be replaced by a high granularity electromagnetic and hadronic calorimeter (HGCAL) [92]. In the ECAL barrel crystal calorimeter, the zero-suppression and selective readout will not be implemented and all ECAL crystal channels in the barrel will be read out using a faster readout chain that will better-sample the pulse and provide fast timing measurements for electrons and photons [93].
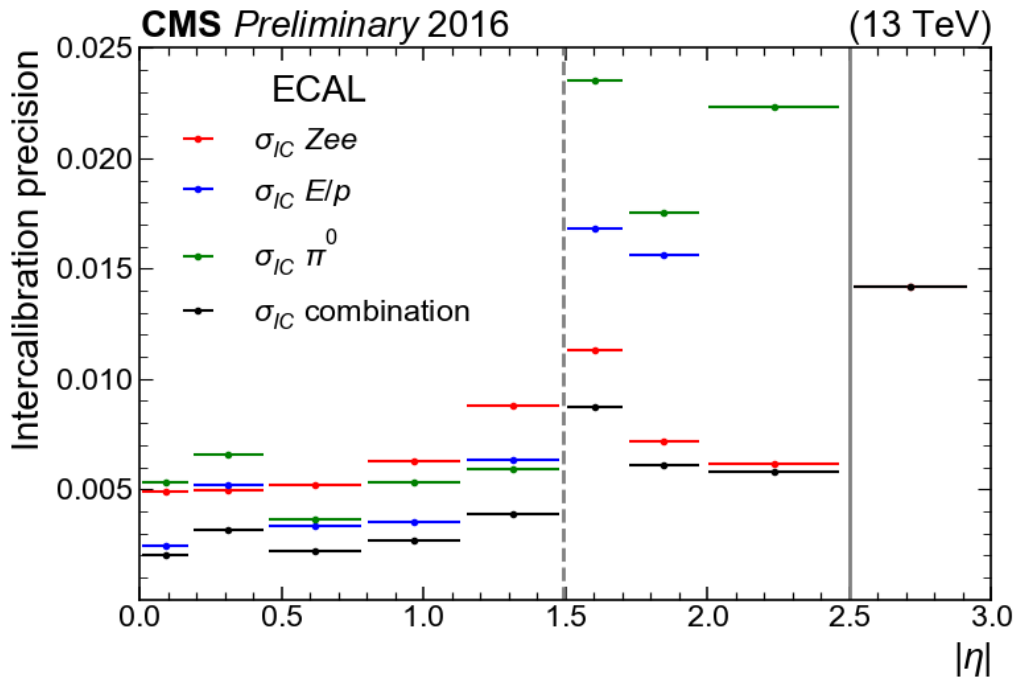
Figure 6.10: Inter-calibration constant precision obtained with the different methods and their combination, as a function of the pseudorapidity $|\eta|$ of the crystal for the 2016 data: red, blue, and green are from the $Z \rightarrow ee$, E/p, and $\pi^0 \rightarrow \gamma\gamma$ methods, respectively, and the black is the precision of the combination from the three methods (the weighted average of the constants obtained by using the three methods) [52].

But in the meantime, the electronic noise will continue to grow as the irradiation of the APD photodetectors [93] will continue, including during LHC Run 3 that is scheduled for 2022-24. The noise subtraction from the small photon clusters will likely be an increased challenge, for these reasons, during Run 3.

In addition to the zero-suppression challenge, the photon clustering algorithm could also be improved such that the energy from the overlapped channels gets shared between two clusters instead of being assigned to the leading photon cluster. This further development could be important to maintain and possibly improve the sub-percent level precision of the inter-calibration, which is important for the study of many SM and BSM physics processes that require precise measurements of electrons and/or photons, such as the measurements of Higgs boson properties and the search for deviations from SM predictions using the diphoton and four lepton channels.

Part III

Search for triboson production at the LHC

Chapter 7

FIRST OBSERVATION OF THE PRODUCTION OF THREE MASSIVE GAUGE BOSONS

In the standard model (SM), the weakly interacting gauge bosons W and Z interact with each other through gauge-gauge couplings, the triple gauge coupling (TGC) and quartic gauge coupling (QGC), and they also interact with the Higgs boson via a Higgs-gauge coupling (HGC). Together with the direct radiation from quarks, these interactions contribute to the simultaneous production of multiple massive gauge bosons, as shown in Figures 7.1 and 7.2. As a result, the measurement of multiboson production is a key tool in probing the TGC, QGC, HGC, and any anomalous couplings beyond the standard model.
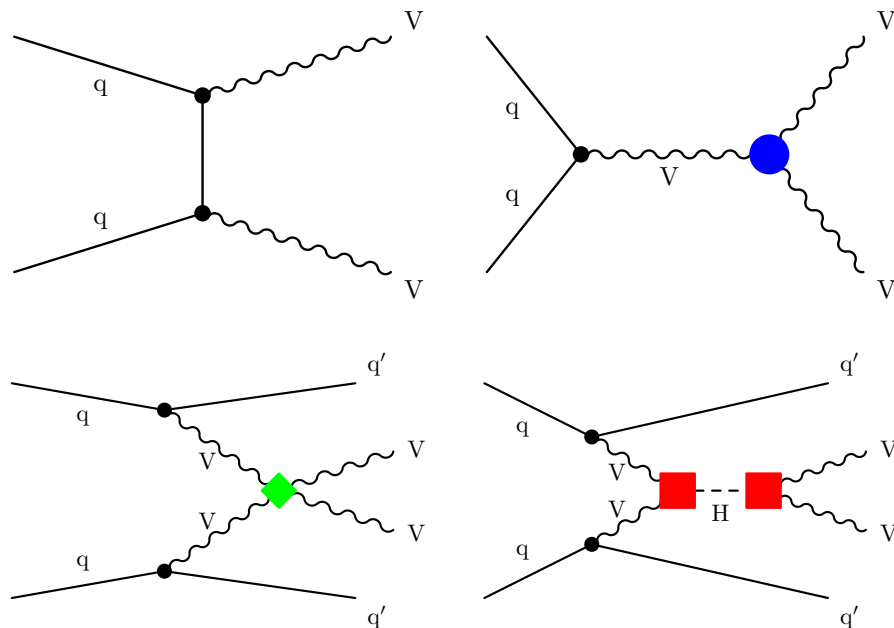
Figure 7.1: Tree-level Feynman diagrams of di-boson productions ($V = W, Z$), where the triple gauge-boson coupling, quartic gauge-boson coupling, and Higgs-gauge couplings are marked by ●, ◆, and ■, respectively.

The electroweak production of two weakly interacting gauge bosons (VV, $V = W, Z$), including $W^{\pm}W^{\pm}$ [94, 95], WZ [96, 97], and ZZ [98, 99], have all been observed individually at the LHC with Run 2 data. However, the production of
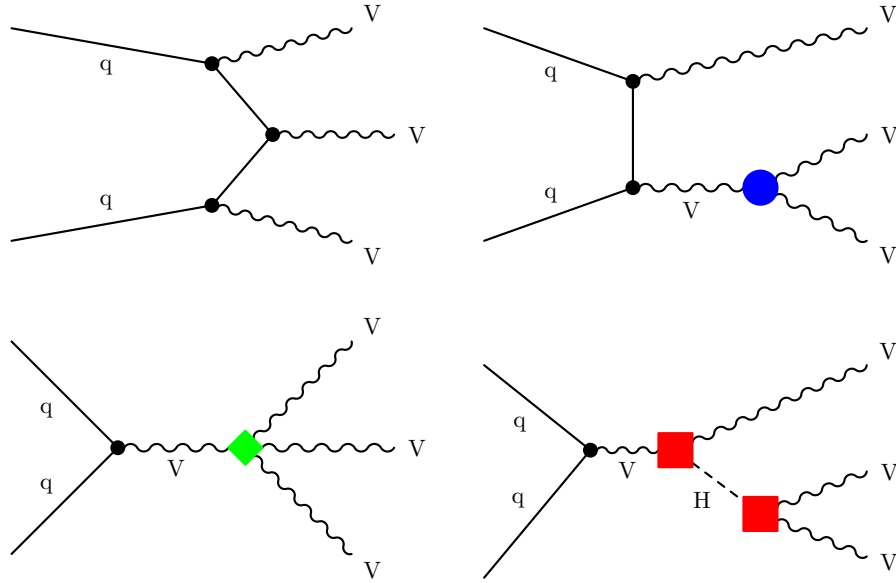
Figure 7.2: Tree-level Feynman diagrams of triple boson productions (V = W, Z), where the triple gauge-boson coupling, quartic gauge-boson coupling, and Higgs-gauge couplings are marked by ●, ◆, and ■, respectively.

three weakly interacting gauge bosons (VVV, V = W, Z) has not been observed so far (prior to the analysis described in this thesis chapter, that is), due to the much lower cross section. Previously, ATLAS found the first evidence of combined VVV production with $4.0(3.0)\sigma$ observed (expected) significance by using a data sample corresponding to an integrated luminosity of $80.0\,\mathrm{fb}^{-1}$ data [100], and CMS also published a WWW search result with a $35.9\,\mathrm{fb}^{-1}$ data sample, which resulted in an observed (expected) significance of WWW production of $0.60(1.78)\sigma$ [101].

In this chapter, we report on the first observation of combined VVV production (V = W, Z) with $137\,\mathrm{fb}^{-1}$ of data collected by CMS [102]. The search is performed in multiple final states with two same-sign leptons plus one or two jets (for WWW), three leptons (for WWW), four leptons (for WWZ), five leptons (for WZZ), and six leptons (for ZZZ). The observed (expected) significance of the combined VVV production is $5.7(5.9)\sigma$, and the corresponding measured cross section is 1010^{+210}_{-200} (statistical)$^{+150}_{-120}$ (systematic) fb, corresponding to $1.02^{+0.26}_{-0.23}$ times the SM prediction. The significance and cross section measurement results for the individual WWW, WWZ, WZZ, and ZZZ production channels are also reported. The VVV signals considered in this search include production from both gauge-gauge coupling (on-shell VVV) and Higgs-gauge couplings (VH(VV)). Fitting results

without including VH(VV) as part of the signal (and thus treating it as a background predicted by SM) are also presented. The results of a simultaneous fit of the on-shell VVV and VH(VV) signal strengths are also discussed.

The remaining sections of this chapter are organized as follows: Section 7.1 gives an overview of the analysis and the overall strategy used to perform the search. Section 7.2 lists the data and simulated event samples used in this search and the High-Level Triggers (HLT) used to collect the signal events. Section 7.3 discusses the details of the reconstruction and identification of the related "physics objects" (leptons, jets, p_T^{miss}). The details of the event selection are presented in Section 7.4. Background estimates and the systematic uncertainties on the backgrounds and the signal predictions are presented in Section 7.5 and Section 7.6; Section 7.7 presents the interpretation of the search as well as the results. Finally, a summary is given in Section 7.8.

7.1   Analysis strategy

To suppress the enormous QCD multijet backgrounds, only final states containing at least one same-sign charged lepton pair resulting from decays of the VVV process are considered, namely: $W^\pm W^\pm W^\mp \to \ell^\pm \ell^\pm \nu^n q\overline{q}'$, $W^\pm W^\pm W^\mp \to \ell^\pm \ell^\pm \ell^\mp \nu^n$, $W^\pm W^\mp Z \to \ell^\pm \ell^\mp \nu^n \ell^\pm \ell^\mp$, $W^\pm ZZ \to \ell^\pm \nu^n 2(\ell^\pm \ell^\mp)$, and $ZZZ \to 3(\ell^\pm \ell^\mp)\nu^n$, where $\ell = $ e or μ and the number of neutrinos $n \geq 0$ depending on number of $\tau \to \ell\nu\nu$ decays in the chain.

The total cross section of VVV production is approximately 1 pb, where WWW has the largest cross section and ZZZ has the smallest cross section. Table 7.1 summarizes the production cross sections for the different VVV processes and the cross section of the corresponding leptonic decay processes, along with the number of expected signal events produced in the Run 2 dataset ($137\,\mathrm{fb}^{-1}$). Signal cross sections are calculated from MADGRAPH5_AMC@NLO interfaced to PYTHIA 8, including NNLO QCD + NLO EW corrections [103–111] with the branching ratios taken from [111]. In general, the VH(VV) contribution is approximately 55% of the total VVV production cross section. However, as we will see later in the event selection and results sections, the VH(VV) contribution to the total event yields after the full event selection is less, especially in some of the most sensitive channels, since the selections favor the kinematics of non-VH VVV processes.

To have full coverage of the leptonic channels of VVV decays, as specified above, events are categorized based on the number of leptons in the final state, ranging from

Table 7.1: Summary of signal process cross sections (at $\sqrt{s}$ =13 TeV) and the expected total number of events produced in the Run 2 data set (137 fb^{-1}).

Process ($\ell = \mathrm{e}/\mu, n \geq 0$)	Cross section (fb)	Events (137 fb^{-1})
$\mathrm{pp} \rightarrow \mathrm{WWW}_{\text{non-WH}}$	216.0	29592
$\mathrm{pp} \rightarrow \mathrm{WH} \rightarrow \mathrm{WWW}$	293.4	40196
$\mathrm{pp} \rightarrow \mathrm{WWZ}_{\text{non-ZH}}$	165.1	22619
$\mathrm{pp} \rightarrow \mathrm{ZH} \rightarrow \mathrm{WWZ}$	188.9	25879
$\mathrm{pp} \rightarrow \mathrm{WZZ}_{\text{non-WH}}$	55.7	7631
$\mathrm{pp} \rightarrow \mathrm{WH} \rightarrow \mathrm{WZZ}$	36.0	4932
$\mathrm{pp} \rightarrow \mathrm{ZZZ}_{\text{non-ZH}}$	14.0	1918
$\mathrm{pp} \rightarrow \mathrm{ZH} \rightarrow \mathrm{ZZZ}$	23.1	3165
sum of $\mathrm{pp} \rightarrow \mathrm{VVV}_{\text{non-VH}}$	450.8	61759
sum of $\mathrm{pp} \rightarrow \mathrm{VH} \rightarrow \mathrm{VVV}$	541.4	74171
$\mathrm{pp} \rightarrow \mathrm{WWW} \rightarrow \ell^{\pm}\ell^{\pm}\nu^{n}\mathrm{q}\overline{\mathrm{q}}'$	24.44	3348
$\mathrm{pp} \rightarrow \mathrm{WWW} \rightarrow \ell^{\pm}\ell^{\pm}\ell^{\mp}\nu^{n}$	8.30	1137
$\mathrm{pp} \rightarrow \mathrm{WWZ} \rightarrow \ell^{\pm}\ell^{\mp}\nu^{n}\,\ell^{\pm}\ell^{\mp}$	1.63	223
$\mathrm{pp} \rightarrow \mathrm{WZZ} \rightarrow \ell^{\pm}\nu^{n}\,2(\ell^{\pm}\ell^{\mp})$	0.12	16.3
$\mathrm{pp} \rightarrow \mathrm{ZZZ} \rightarrow 3(\ell^{\pm}\ell^{\mp})\nu^{n}$	0.014	1.9
sum of above $\mathrm{VVV} \rightarrow 2/3/4/5/6\ell + (\nu)(\mathrm{q})$	34.50	4726

2 to 6, targeting decays from $\mathrm{W}^{\pm}\mathrm{W}^{\pm}\mathrm{W}^{\mp} \rightarrow \ell^{\pm}\ell^{\pm}\nu^{n}\mathrm{q}\overline{\mathrm{q}}'$ to $\mathrm{ZZZ} \rightarrow 3(\ell^{\pm}\ell^{\mp})\nu^{n}$. The selections in each category are then optimized to select one particular VVV decay and to suppress the major backgrounds in that category.

Events in the two lepton category are selected targeting the $\mathrm{W}^{\pm}\mathrm{W}^{\pm}\mathrm{W}^{\mp} \rightarrow \ell^{\pm}\ell^{\pm}\nu^{n}\mathrm{q}\overline{\mathrm{q}}'$ process. As a result, two leptons with the same charge are required in this category, along with additional final state jets. Since the WH(WW) signal is relatively softer and usually has less jets, events with only one jet are also allowed. Therefore, events are then categorized based on the number of jets: one jet or two jets. For two-jet events, the invariant mass of the two jets (m_{jj}) in signal events is usually close to the W mass, so events are further categorized based on whether the m_{jj} is inside or outside a window of ± 15 GeV centered on the W mass. There are two major backgrounds in the same-sign two leptons final state. The first one is the fully leptonic decay of the WZ process in which one of the three leptons is not reconstructed due to the limited acceptance of the detector and the limited lepton reconstruction efficiency. The second background contains a nonprompt lepton which originates from a heavy flavor hadron decay such as a top quark decay, from light flavor hadrons such as $\pi^{\pm}$, or from photon conversions. The probability of an

electron to be nonprompt is larger than that of a muon. As a result, events in the two-lepton category are further divided into three categories based on the flavor of the two leptons, to enhance the signal-to-background ratio: $e^{\pm}e^{\pm}$, $e^{\pm}\mu^{\pm}$, or $\mu^{\pm}\mu^{\pm}$.

Events in the three-lepton category are selected to target the $W^{\pm}W^{\pm}W^{\mp} \rightarrow \ell^{\pm}\ell^{\pm}\ell^{\mp}\nu^n$ process. In this final state, the WZ process is the dominant background. To suppress this background, events are categorized based on the number of same flavor and opposite sign (SFOS) lepton pairs in the event: 0 SFOS ($e^{\pm}e^{\pm}\mu^{\mp}$, $\mu^{\pm}\mu^{\pm}e^{\mp}$), 1 SFOS ($e^{\pm}e^{\mp}\mu^{\pm}$, $\mu^{\pm}\mu^{\mp}e^{\pm}$), and 2 SFOS ($e^{\pm}e^{\pm}e^{\mp}$, $\mu^{\pm}\mu^{\pm}\mu^{\mp}$). There are significantly less background events from WZ in the 0 SFOS category compared to 1 SFOS and 2 SFOS, so the 0 SFOS category has much better signal-to-background discrimination.

Events in the four-lepton category are selected to target the $W^{\pm}W^{\mp}Z \rightarrow \ell^{\pm}\ell^{\mp}\nu^n \ell^{\pm}\ell^{\mp}$ processes. Such events are required to have two SFOS leptons that have an invariant mass inside a Z mass window (±10 GeV around the Z mass). The remaining two leptons (called W-leptons) are required to have opposite signs of charge. In this category, the ZZ process is the dominant background. To suppress this background, the events are split into two categories based on the flavors of the W-leptons: $e\mu$, $ee/\mu\mu$. The $e\mu$ events have much less background from ZZ, with a small portion from $Z(ee/\mu\mu)Z(\tau\tau)$ decay. To suppress such events, the invariant mass of the $e\mu$ system ($m_{e\mu}$) is used to bin the events, as the $m_{e\mu}$ from $Z(\tau\tau)$ peaks around 50 GeV while the $m_{e\mu}$ from WWZ events is distributed over a wider range. The $ee/\mu\mu$ events have enormous backgrounds from the ZZ process. As a result, the invariant mass of the $ee/\mu\mu$ system ($m_{\ell\ell}$) is required to be outside a Z mass window (±10 GeV), and the p_T^{miss} and $p_T^{4\ell}$ of the event are used to bin the events in order to enhance the signal-to-background ratio.

Events in the five- and six-lepton categories are selected to target the $W^{\pm}ZZ \rightarrow \ell^{\pm}\nu^n 2(\ell^{\pm}\ell^{\mp})$ and $ZZZ \rightarrow 3(\ell^{\pm}\ell^{\mp})\nu^n$ processes, respectively. Such final states have very low statistics and are very clean in terms of backgrounds. As a result, for the five-lepton category we simply select 2 pairs of leptons from the Z and require an additional lepton, and for the six-lepton category we simply require six leptons in the final state.

For events in the two, three, and four leptons categories, both a sequential cut-based event selection method and a Boosted Decision Tree (BDT) based event selection method are used to enhance signal-to-background ratio (see Section 7.4); for events in five- and six-lepton categories, only a cut based method is used. A graphical

overview of the analysis strategy discussed above can be seen in Figure 7.3.

As we can see from Table 7.1, without considering the reconstruction efficiency, there are a total of 4726 signal events expected in all of the considered categories, which amount to approximately 3.5% of the total number of VVV events produced. The majority of the selected events expected are in the same-sign two-lepton and three-lepton categories, while only a few events are expected in the six-lepton category.
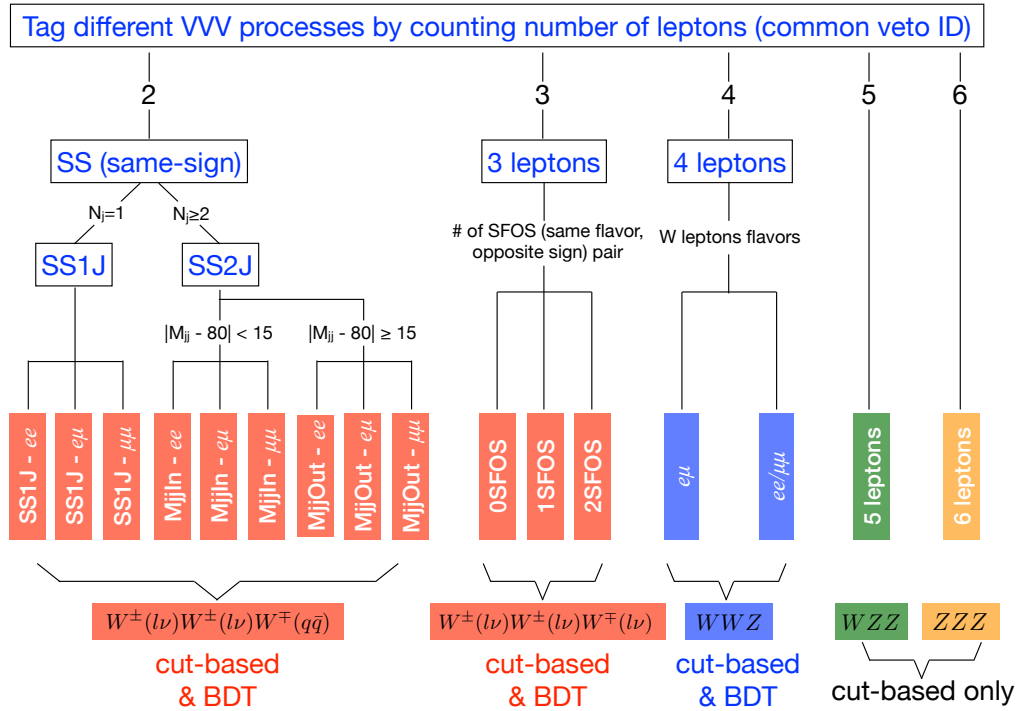
Figure 7.3: Overview of the triboson search strategy: events are categorized based on the number of leptons in the final states, and each set of final states is further divided into different categories to enhance the signal over background.

7.2 Event samples

This search uses the pp collision data collected by the CMS detector during the data taking years 2016, 2017, and 2018, corresponding to a total integrated luminosity of $137\,\mathrm{fb}^{-1}$.

The triggers used in this search are the dilepton triggers, which require the presence of two leptons passing some p_T thresholds. Depending on the flavors of the leptons, there are three different triggers being used: ee, eμ, and μμ triggers. The ee trigger requires the presence of at least two electrons (or positrons) of $p_\mathrm{T} > 23$ and 12 GeV,

and each of them has to pass some electron identification cuts, including cuts on the shower shapes, isolation, pixel matching, etc. The eμ trigger requires at least one electron of $p_{\mathrm{T}} > 23(8)$ GeV and at least one muon of $p_{\mathrm{T}} > 12(23)$ GeV, with both leptons passing the corresponding lepton identification cuts. The μμ trigger requires at least two isolated muons of $p_{\mathrm{T}} > 17$ and 8 GeV. An event that passes any of the above triggers is selected for this search.

In addition to the trigger requirement, the electrons and muons are also required to pass much tighter offline identification cuts, which are discussed in Section 7.3. To study the trigger efficiency with respect to the offline selection, and to correct for any mis-modeling of the trigger efficiency in MC, events triggered by orthogonal triggers ($p_{\mathrm{T}}^{\mathrm{miss}}$ triggers) are used. The trigger efficiency is defined as the number of events that pass the offline lepton selection and the trigger requirement divided by the number of events that pass the offline lepton selection. Simulated $t\bar{t}$ events are used to measure the trigger efficiency in Monte Carlo (MC).

The trigger efficiency of the ee, eμ, and μμ triggers, binned in the p_{T} of the two leptons, measured in both data and in $t\bar{t}$ MC for all three data taking years, are shown in Figures 7.4, 7.5, and 7.6. The ratio of the trigger efficiency in data to that in MC is used as a scale factor to correct for the MC yields in this search.

On average, the electron triggers have slightly higher efficiency than the muon triggers, mainly due to the inefficiency of the Level 1 (L1) muon triggers, as reported in [112] and discussed in Section 4.5. For all three triggers, the efficiency in MC is consistent with the data, with the ratio of efficiency to data consistent with 1.0 within an uncertainty of about 5%. This small inconsistency is corrected in all MC events, and the uncertainty is taken into account as one of the systematic uncertainties when we perform the fit for the result of this search.

The modeling of the standard model (SM) backgrounds and the VVV signal process rely on Monte Carlo (MC) simulations. The MADGRAPH5_AMC@NLO v2.2.2 (v2.4.2) generator [113] was used in 2016 (2017+2018) to generate the VVV and VH(VV) signal events and the single-top and diboson (VV, VVγ) background events. The next-to-leading order (NLO) POWHEG 2.0 generator [114–117] was used to generate VH(VV) signal events and single-top and diboson (VV) background events. The MADGRAPH5_AMC@NLO generator with the MLM scheme for matrix element and parton shower jet matching [118]wass used to generate SM $t\bar{t}$, $t\bar{t}$+X (X = W, Z, H), W+jets, Z+jets, Wγ, and $W^{\pm}W^{\pm}$ events. The NNPDF3.0 (3.1) [119, 120] parton distribution functions (PDFs) were used in the
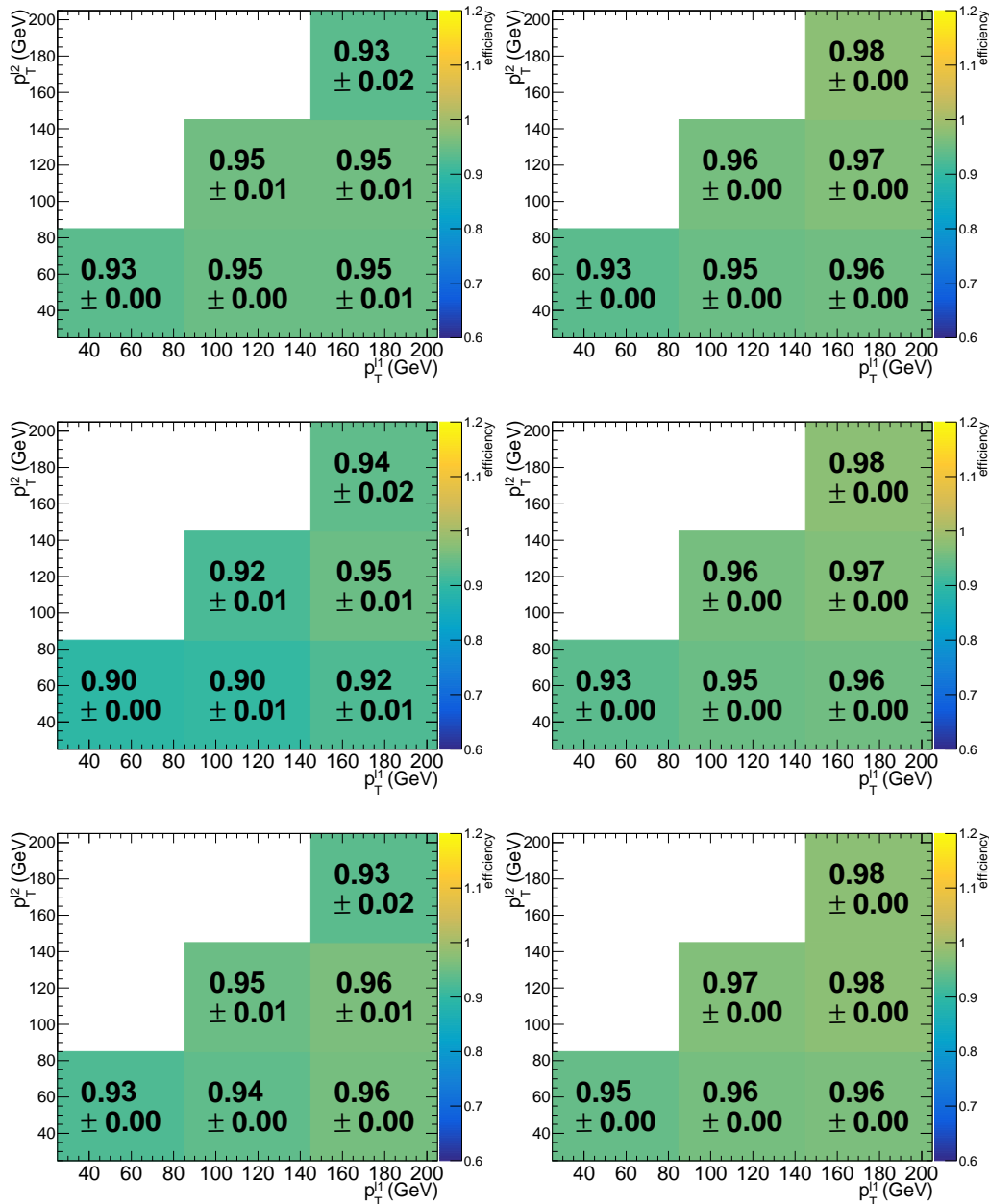
Figure 7.4: Trigger efficiency of the ee trigger for data (left column) and $t\bar{t}$ MC (right column) in 2016 (top), 2017 (middle), and 2018 (bottom), in bins of the two electrons' p_T. The x-axis is the p_T of the leading electron, $p_T^{\ell_1}$.

2016 (2017+2018) simulated samples. All MC simulations used PYTHIA 8.205 (8.230) [121] in 2016 (2017+2018) to model the parton showering, hadronization, and the underlying event, with parameters set by the CUETP8M1 tune [122] (for 2016) or the CP5 tune [123] (for 2017+2018). The interactions of all the particles
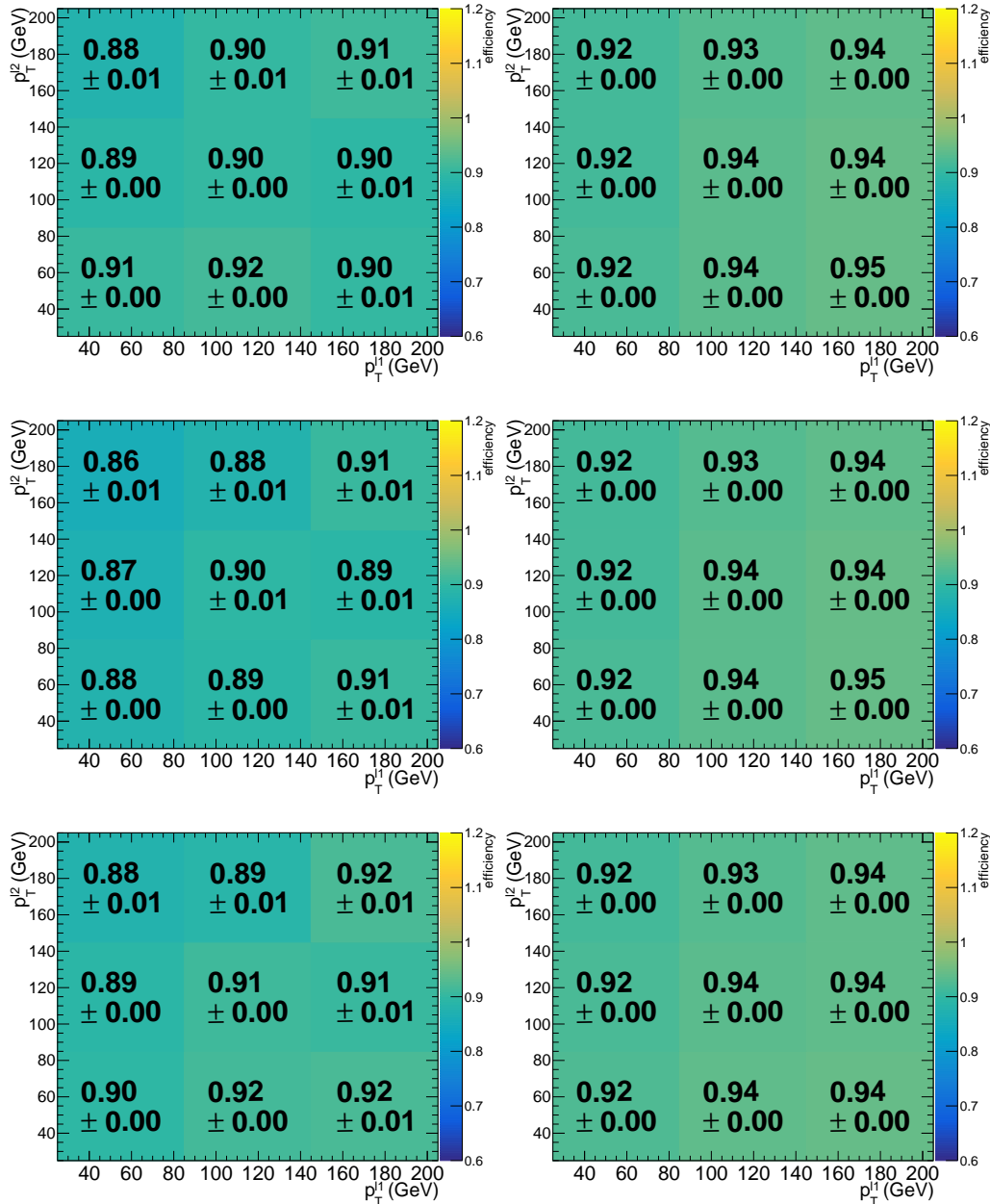
Figure 7.5: Trigger efficiency of the eμ trigger for data (left column) and $t\bar{t}$ MC (right column) in 2016 (top), 2017 (middle), and 2018 (bottom), in bins of the electron's and muon's p_T. The x-axis is the p_T of the leading electron, $p_T^{\ell_1}$.

in the generated events with the CMS detector are simulated with the GEANT4 [85] package, and the same reconstruction software used as is for data. Detector readout information from additional pp collisions within one single bunch crossing, referred to as pileup, are also simulated and included in each MC event.
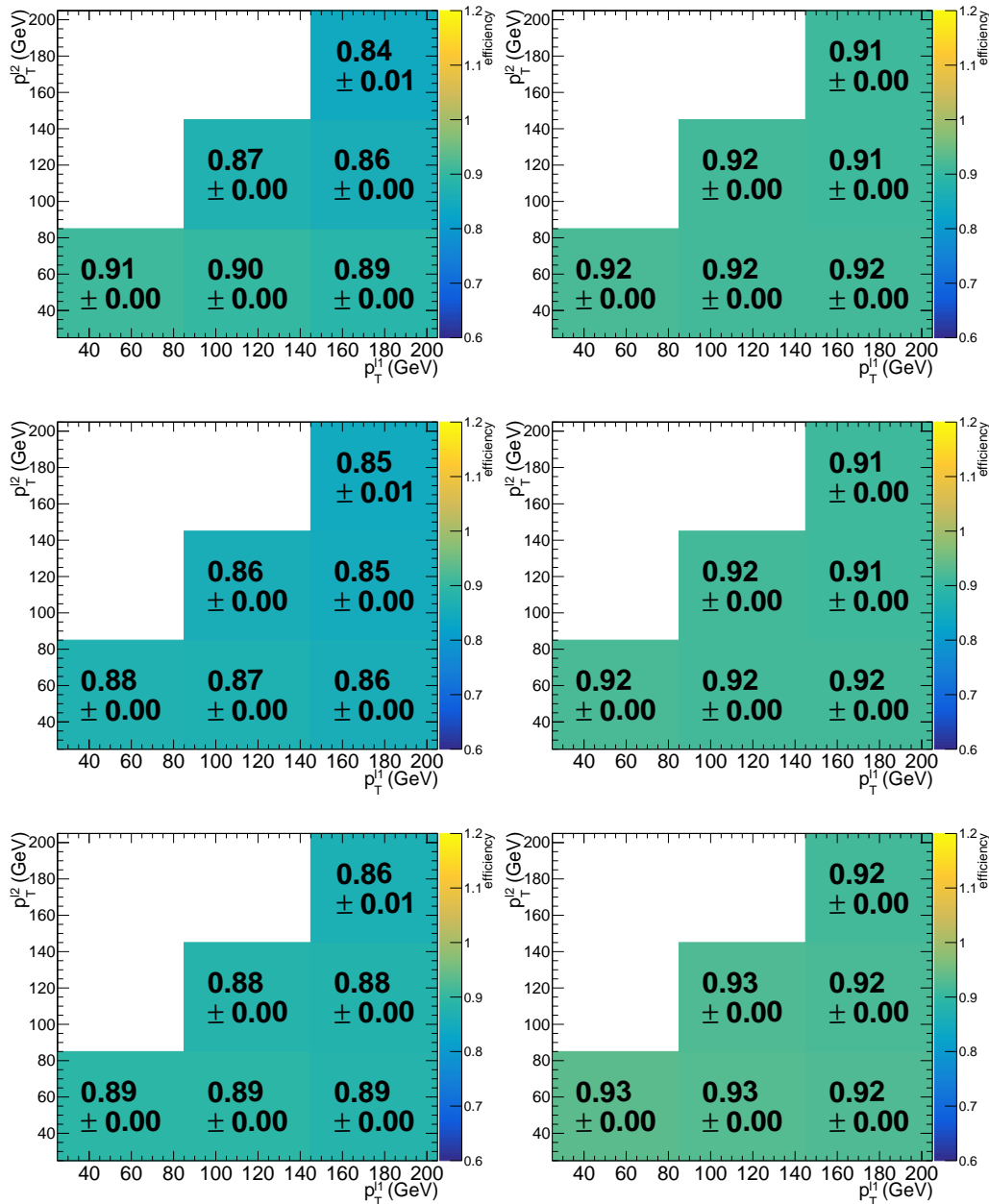
Figure 7.6: Trigger efficiency of the μμ trigger for data (left column) and $t\bar{t}$ MC (right column) in 2016 (top), 2017 (middle), and 2018 (bottom), in bins of the two muons' p_T. The x-axis is the p_T of the leading muon, $p_T^{\ell_1}$.

7.3 Objects reconstruction and identification

Physics objects are reconstructed based on the particle-flow (PF) algorithm [71], with details discussed in Section 4.6.

As electron objects are reconstructed by combining the momentum measurement

from the tracker and the energy measurement from the ECAL, and the tracker system only covers the pseudorapidity range up to $|\eta| < 2.5$, electrons in this search are only accepted up to $|\eta| < 2.5$, despite further coverage in η by ECAL (up to $|\eta| < 3.0$). To select electrons from other objects such as photon conversions and jets, a set of requirements is applied to the properties of the electron. This selection includes a multivariate (MVA) electron identification procedure which is trained by using the following variables as the main inputs: SC shower shapes ($\sigma_{i\eta i\eta}$, $\sigma_{i\phi i\phi}$, R_9), the fraction of the energy radiated as bremsstrahlung, $1/E - 1/p$, the ratio of the energy measured in the hadron calorimeter to the energy in the ECAL (H/E), any missing inner track hits, the conversion probability for the photons, $\Delta\eta$ and $\Delta\phi$ between the center of the SC (or the seed cluster) and the track. Apart from the selection based on the MVA output, additional selections on the isolation and the vertex of the electron are also applied.

The identification (ID) of muons relies on the parameters obtained by the muon reconstruction (such as the number of hits and the track fit χ^2) and on the matching between tracker tracks and standalone muon tracks. Two sets of muon IDs are used in this search, referred to as *Loose* and *Medium* working points, with the former aiming to select muons from the primary vertex and from decays of both light and heavy flavors, and the latter aiming to identify muons from the primary vertex and only heavy flavor decays. Additional requirements on the isolation and vertex of the muon are also implemented in the muon identification.

Two main types of lepton (electron and muon) ID requirements are defined in this search. The first one is referred to as *Veto* IDs, which are used to count the number of leptons and categorize events. The second type is referred to as *Nominal* IDs, which are different in different final states in order to better suppress various backgrounds and maintain high signal efficiencies. Leptons in one specific category which pass the *Veto* ID but later fail the *Nominal* ID are discarded (or in some cases, only used in control regions).

Table 7.2 summarizes the *Veto* ID requirements for both the electrons and muons applied in this search. $|d_z|$ ($|d_{xy}|$) is the longitudinal (transverse) impact parameter, defined as the distance from the center of the beam spot to the electron or muon in the longitudinal (transverse) plane. $I_{\mathrm{rel,R=0.3,Lep}}$ is the relative isolation, calculated by the ratio of the scalar sum of the p_{T}'s of all other PF particles (hadrons, photons and leptons) in a $\Delta R = 0.3$ cone around the lepton to the p_{T} of the lepton itself. The electron and muon are both required to have $p_{\mathrm{T}} > 10$ GeV to be considered as a

Veto lepton. The selection efficiency of this set of *Veto* ID criteria is about 90-98% for electrons and 93-99% for muons, depending on the p_T of the lepton.

Table 7.2: Electron and Muon *Veto* ID criteria used in the search and the corresponding average selection efficiencies for leptons. This common *Veto* ID guarantees orthogonality in the event selection among the different lepton bins used in the analyses.

	Electron	Muon
ID	MVA Loose	Loose
$|\eta|$	< 2.5	< 2.4
$|d_z|$	< 0.1 cm	< 0.1 cm
$|d_{xy}|$	< 0.05 cm	< 0.05 cm
$I_{\mathrm{rel,R=0.3,Lep}}$	< 0.4	< 0.4
efficiency (low p_T)	90%	93%
efficiency (high p_T)	98%	99%

For the *Nominal* lepton IDs of same-sign events, tighter selections are applied in addition to the requirements applied in the *Veto* lepton IDs, to better reject fake lepton backgrounds. Specifically, the electrons are required to pass a tighter cut on the MVA score (called the 80% working point) and the muons are required to pass a tighter cut (called the *Medium* working point). In addition, we tighten the relative isolation cuts to < 0.05 (0.04) for electrons (muons). We also require the 3D impact parameter of the electron (muon) to be smaller than 0.01 (0.015) cm relative to the primary vertex of the event, and we require the ratio of the 3D impact parameter to its estimated uncertainty for muons to be smaller than 4. The p_T of the electrons and muons in same-sign final states are both required to be greater than 25 GeV. The selection efficiency of the same-sign events' *Nominal* lepton ID, which is calculated after applying the *Veto* ID, is about 40-70% (20-30%) for an electron in the barrel (endcap) depending on the p_T of the electron and is about 55-90% for a muon, depending on the p_T of the muon.

For three-lepton events, the *Nominal* lepton IDs are slightly looser compared to the *Nominal* lepton ID of same-sign events. Specifically, the electrons are required to pass a 90% working point and muons are required to pass the *Medium* working point. The relative isolation cut is < 0.10 (0.15) for electrons (muons). The 3D impact parameter is required to be smaller than 0.015 cm for both electrons and muons. The ratio of the 3D impact parameter to its estimated uncertainty for muons is required to be smaller than 4. The p_T of the electrons and muons are required to be greater than 25/20 GeV depending on the event category (see Section 7.4).

The selection efficiency of the *Nominal* lepton ID for three-lepton events, which is calculated after applying the *Veto* ID, is about 60-80% (30-40%) for electrons in the barrel (endcap) depending on the p_T of the electron, and is about 75-97% for muons depending on the p_T of the muon.

For the *Nominal* lepton ID in four-lepton events, different requirements are applied based on whether the lepton is tagged as a lepton from W or Z decay, to optimize the signal-to-background ratio in these final states. For leptons from W decay, the electrons are required to pass an MVA 90% working point ID, and muons are required to pass the *Medium* working point ID. The ratio of the 3D impact parameter to its estimated uncertainty for both electrons and muons is required to be smaller than 4. The relative isolation is required to be smaller than 0.2 (0.15) for electrons (muons). For Z leptons, the electrons are required to pass an MVA *Loose* working point ID, and muons are required to pass the *Medium* working point ID. The ratio of the 3D impact parameter to its estimated uncertainty for both electrons and muons is required to be smaller than 0.20 (0.25) for electrons (muons). The ratio of the 3D impact parameter to its estimated uncertainty for both electrons and muons is required to be smaller than 4. The relative isolation is required to be smaller than 0.20 (0.15) for electrons (muons). The p_T of the two leptons from Z decay are required to be greater than 25 (15) GeV for the leading (subleading) lepton; and the p_T of the two charged leptons from W decay are required to be greater than 25 (15) GeV for the leading (subleading) one as well.

The selection efficiency of the lepton ID for leptons from W decay in four-lepton events, which is calculated after applying the *Veto* ID, is about 70-80% for electrons and about 70-97% for muons depending on the p_T of the electron or muon. The selection efficiency of the lepton ID for leptons from Z decay in four-lepton events, which is calculated after applying the *Veto* ID, is about 80-90% for electrons and about 80-97% for muons depending on the p_T of the electron or muon.

For five-lepton events, four of the leptons are required to pair with two Z bosons, and these leptons are required to pass the Z lepton ID as discussed above for four-lepton events. The last lepton which is tagged as being from the decay of a W boson, is required to pass the W lepton ID as discussed above for four-lepton events. The p_T of the leading (subleading) leptons associated with the two Z boson candidates is required to be greater than 25 (10) GeV. Furthermore, the p_T of the W candidate lepton is required to be greater than 10 GeV.

For six-lepton events, only the *Veto* lepton ID is applied to the six leptons. The p_T of the two leading leptons are required to be greater than 25 GeV, and the p_T of the rest of the four leptons are required to be greater than 10 GeV each.

Tables 7.3 and 7.4 list all the electron and muon cuts applied in the *Nominal* lepton ID for same-sign events, three-lepton events and four-lepton events (W lepton ID and Z lepton ID).

Table 7.3: Electron *Nominal* ID applied for different final states and the corresponding average selection efficiencies for prompt electrons. The efficiency is measured after applying the *Veto* ID.

	SS-ID	3ℓ-ID	W candidate (WID)	Z candidate (ZID)
Veto ID		Common *Veto* ID		
Working points	MVA 80%	MVA 90%	MVA 90%	MVA Loose
p_T	> 25 GeV	> 25/20 GeV	> 25(15) GeV	> 25(15) GeV
IP$_{3D}$	< 0.01 cm	< 0.015 cm	< $4\sigma_{\text{IP}_{3D}}$	< $4\sigma_{\text{IP}_{3D}}$
$I_{\text{rel,R}=0.3,\text{Lep}}$	< 0.05	< 0.10	< 0.20	< 0.20
efficiency (low p_T)	20-40%	30-60%	70%	80%
efficiency (high p_T)	30-70%	40-80%	80%	90%

Table 7.4: Muon *Nominal* ID applied for different final states and the corresponding average selection efficiencies for prompt muons. The efficiency is measured after applying the *Veto* ID.

	SS-ID	3ℓ-ID	W candidate (WID)	Z candidate (ZID)
Veto ID		Common *Veto* ID		
Working points		Medium working point		
p_T	> 25 GeV	> 20 GeV	> 25(15) GeV	> 25(15) GeV
IP$_{3D}$	< 0.015 cm	< 0.015 cm	< $4\sigma_{\text{IP}_{3D}}$	< $4\sigma_{\text{IP}_{3D}}$
$I_{\text{rel,R}=0.3,\text{Lep}}$	< 0.04	< 0.15	< 0.25	< 0.15
efficiency (low p_T)	55%	75%	70%	80%
efficiency (high p_T)	90%	97%	97%	97%

For all the *Veto* and *Nominal* lepton IDs discussed above, dedicated efficiencies are measured in bins of the lepton p_T and η for both data and MC. The differences of the efficiencies between data and MC are used to correct the MC event yields, and the corresponding uncertainties are also applied.

Jets are reconstructed from PF candidates, clustered with the anti-kt algorithm and with a cone size of 0.4 [74–76]. Only jets with transverse momentum p_T > 30 GeV

and $|\eta| < 2.5$ (within the tracker acceptance) are considered in this search. To reject noise and mis-measured jets, the selected jets have to fulfill several identification criteria, summarized in Table 7.5. The efficiency of this set of identification requirements is above 99% for the p_T and η range we consider in this search [124]. A series of corrections is applied to the measured four-momentum of the jet in order to correct for different effects in the jet measurement (such as pileup subtraction and jet flavor corrections). The corrections for jets reconstructed in data and MC are different. The details of the corrections are discussed further in [125].

Table 7.5: Jet identification selections for data taken from different years.

	2016	2017 and 2018
neutral hadronic energy fraction	< 0.99	< 0.90
neutral electromagnetic energy fraction	< 0.99	< 0.90
number of constituents	> 1	> 1
charged hadron fraction	> 0	> 0
charged multiplicity	> 0	> 0
charged EM fraction	< 0.99	-

To avoid double counting due to jets matched geometrically with a lepton, a jet that is matched to a lepton passing the *Veto* ID within $\Delta R < 0.4$ is not considered as a jet in the event.

To tag jets originating from b quarks, the loose working point of the deep combined secondary-vertex discriminator (DeepCSV) [126] is used. The DeepCSV is a multi-classification discriminator developed using a deep neural network with four hidden layers with 100 nodes each. The input variables for the training include the jet variables (p_T, η, number of selected tracks, etc), and the variables of the selected tracks (up to six selected tracks) as well as variables of the associated secondary vertex. Each node in the last layer of the neural network is a probability for a certain jet flavor category, for example $P(b)$, $P(c)$, $P(udsg)$. The b tagging used in this search requires the output discriminator $P(b)$ of the jet to be greater than 0.2217 in 2016, 0.1522 in 2017, and 0.1241 in 2018, which corresponds to about 84% efficiency for b jets and about 10% misidentification probability efficiency (for light-flavor jets). For tagging jets as b-tagged jets, a cut of $p_T > 20\,\mathrm{GeV}$ is used (instead of the > 30 GeV cut which is used to select other jets). To tag b hadrons with even lower p_T more efficiently, a soft b tag object [127] is used, which only uses a secondary vertex and does not require the presence of a jet.

The missing transverse momentum p_T^{miss} is defined as the negative sum of the transverse momentum of all PF particles reconstructed in an event. The jet energy corrections discussed above are propagated for the correction of the p_T^{miss} calculation. To avoid anomalous high p_T^{miss} events from detector noise, a series of filters designed to identify noise in the ECAL, HCAL, and Muon systems is applied to all events. Details of the p_T^{miss} reconstruction and performance in the CMS detector can be found in [128].

An "isolated track object" has been designed to reject events containing hadronic tau lepton decays. A majority of hadronic taus are single-prong decays and therefore the isolated track criterion is efficient in rejecting events containing hadronic taus. An isolated track can be either a PF lepton or a charged PF hadron, with the detailed requirements summarized in Table 7.6.

Table 7.6: Definition of an isolated track.

	If a lepton (e or μ)	If a charged PF hadron
$|d_z|$	< 0.1 cm	< 0.1 cm
p_T	> 5 GeV	> 10 GeV
ΔR with *Veto* leptons	> 0.01	> 0.01
p_T sum of charged PF candidate within $\Delta R < 0.3$	< 8 GeV or $0.2 \times p_T^{lepton}$	< 8 GeV or $0.1 \times p_T^{hadron}$

7.4 Event selection

This section discusses the full set of event selections for all the final states used in the triboson analysis. The selections presented below are optimized in terms of the signal-to-background ratio in each final state. For the same-sign (SS) dilepton, three-lepton and four-lepton final states, both a sequential cut approach and a multivariate approach using a Boosted Decision Tree (BDT) are used, and the results for both methods are reported in the results section. For the five and six-lepton final states, only a cut-based method is developed, as those two channels are already very clean, and very few events are available for further optimization.

7.4.1 Same-sign dilepton and three-lepton selections (cut-based)

As discussed in Section 7.1, events in the SS dilepton final state are first categorized based on the number of jets, into 1-jet and 2-jet categories. In the 2-jet category, events are then divided based on the dijet mass: the m_{jj}-in category for events with m_{jj} inside a W mass window, and the m_{jj}-out category for events with m_{jj} outside of the W mass window. Additionally, all categories are further divided based on the dilepton flavor ($e^{\pm}e^{\pm}$, $e^{\pm}\mu^{\pm}$, $\mu^{\pm}\mu^{\pm}$).

Different background compositions are present in each of the above categories. The main backgrounds in the SS categories and the corresponding suppression methods are:

- WZ backgrounds ("lost-lepton" backgrounds), in which one of the leptons from the Z decay is not reconstructed (or in events of $Z \rightarrow \tau\tau$ decay with one of the τ decaying to hadrons). This background is suppressed by requiring that there is no additional isolated track in the event (to suppress hadronic τ objects), and a maximum transverse mass between the lepton and the $\vec{p}_{T}^{miss}$ larger than 90 GeV. A cut on the minimum ΔR between any lepton and the jet pair is also applied: $\Delta R_{\ell j}^{min} < 1.5$.

- Nonprompt lepton background, which could originate from $t\bar{t}$ decay where the lepton originates from a b-hadron decay. The $t\bar{t}$ background events are largely reduced by rejecting events with b jets. We also require that the event have a large p_{T}^{miss}, as the background events in this category tend to have less energy carried away by neutrinos, compared to the signal.

- Irreducible backgrounds, such as those from $t\bar{t}W^{\pm}$ and $W^{\pm}W^{\pm}$ +jj processes, where one or both leptons originate from W decay. The $t\bar{t}W^{\pm}$ events are

reduced by rejecting events with a b-jet, and $W^\pm W^\pm$ +jj events are reduced by cuts on the mass and $\Delta\eta$ of the two leading jets: $m_{JJ} < 500$ GeV and $\Delta\eta_{JJ} < 2.5$.

- Other backgrounds, such as events due to lepton charge misreconstruction (a sign flip) in lepton pairs from a Z decay, which results in a fake same-sign event. These are suppressed by requiring that the dilepton mass be outside the Z mass window (± 20 GeV), and also be >20 GeV.

Table 7.7 summarizes all of the selections for the SS dilepton final state. Figures 7.7 and 7.8 show the distributions of some example variables that provide good signal to background discrimination for the SS 2-jet and 1-jet events, respectively.

Table 7.7: Event selection for the SS channel: for each category, we define three signal regions depending on the lepton flavor: $e^\pm e^\pm$, $e^\pm \mu^\pm$, and $\mu^\pm \mu^\pm$. This results in $3 \times 3 = 9$ signal regions for the SS channel.

Variable	m_{jj}-in and m_{jj}-out	1j	target background
Triggers	Pass the dilepton triggers		
Signal leptons	Exactly 2 SS leptons with $p_T > 25$ GeV		-
Jets	≥ 2 jets	1 jet	
Additional leptons	No additional *Veto* lepton		
Isolated tracks	No additional isolated track		
m_T^{max}	>90 GeV if not $\mu^\pm \mu^\pm$	>90 GeV	lost-lepton
$\Delta R_{\ell j}^{min}$	—	<1.5	
m_{jj} (closest ΔR)	$65 < m_{jj} < 95$ GeV or $\|m_{jj} - 80\,\text{GeV}\| \geq 15$ GeV	—	
m_{JJ} (leading jets)	<500 GeV	—	$W^\pm W^\pm$ +jj
$\Delta\eta_{JJ}$ (leading jets)	<2.5	—	
$m_{\ell\ell}$	$\|m_{\ell\ell} - m_Z\| > 20$ GeV if $e^\pm e^\pm$		charge flip
$m_{\ell\ell}$	>20 GeV		
b-tagging	no b-tagged jets and soft b-tag objects	$t\bar{t}$, $t\bar{t}W^\pm$	
p_T^{miss}	>45 GeV	$t\bar{t}$	

Within the three-lepton category, events are further categorized based on the number of SFOS lepton pairs (0, 1, or 2), since the dominant backgrounds (WZ and ZZ backgrounds) have one or two SFOS lepton pairs and the VVV signal has 0 SFOS pairs.

Different event selection requirements are then applied to each of these sub-categories, in order to maximize the background discrimination. The main backgrounds in the three-lepton sub-categories and the corresponding selections to suppress each background are:

- WZ backgrounds ("three-lepton" backgrounds) or ZZ backgrounds ("lost-lepton" backgrounds) that are mainly in the 1 SFOS and 2 SFOS categories. To suppress such backgrounds, the mass of the SFOS lepton pair is required to be outside the Z mass window (± 20 GeV) and the transverse mass of the 3rd lepton (for 1 SFOS events) or the maximum transverse mass of all leptons (for 2 SFOS events) is required to be larger than 90 GeV.

- Nonprompt lepton backgrounds, such as those from $t\bar{t}$ decay where a lepton originates from a b hadron decay. The $t\bar{t}$ events are largely reduced by rejecting events with a b-jet. In addition, for 1 SFOS and 2 SFOS events, we require that the $\Delta\phi$ between the p_{T} of three-lepton system and the $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ be greater than 2.5, as the signal peaks at π and the nonprompt background is flat between 0 and π.

- Irreducible backgrounds, such as those from $t\bar{t}W^{\pm}$ processes, with leptons originating from W decay. This background is reduced by rejecting events with a b-jet; in addition, we reject events with a b-jet in the 1 SFOS and 2 SFOS channels.

- Other backgrounds, such as $Z\gamma$ events with the photon misreconstructed as an electron; or misreconstruction of a lepton leading to a charge-sign flip. To suppress photon fakes, we cut on the p_{T} and mass of the three-lepton system, as this kind of background has low $p_{\mathrm{T}}^{3\ell}$, and $m_{3\ell}$ is inside the Z mass window. To suppress charge-sign flip backgrounds (mainly in 0 SFOS events where two same-sign electrons get reconstructed from a Z), we require the di-electron mass to be outside the Z mass window.

Table 7.8 summarizes all the selections for the SS dilepton final state. Figures 7.9 and 7.10 show the distributions of some example variables which provide good signal-to-background discrimination for the three-lepton events, before any selections are applied to those distributions.

Table 7.8: Event selection for the three-lepton channels.

Variable	0 SFOS	1 and 2 SFOS	target background
Triggers	Pass the dilepton triggers		
Signal leptons	Exactly 3 leptons with charge sum = $\pm 1e$		-
Lepton p_T	$p_T > 25/25/25$ GeV	$p_T > 25/20/20$ GeV	
Jets	≤ 1 jet	0 jet	
Additional leptons	No additional *Veto* lepton		
m_T^{3rd} (1 SFOS)	—	>90 GeV	
m_T^{max} (2 SFOS)	—	>90 GeV	three-lepton
m_{SFOS}	$m_{SFOS} > 20$ GeV and $\|m_{SFOS} - m_Z\| > 20$ GeV		
m_{SF}	$m_{SF} > 20$ GeV	—	
$p_T^{3\ell}$	—	>50 GeV	γ fakes
$m_{3\ell}$	$\|m_{3\ell} - m_Z\| > 10$ GeV		
b-tagging	no b-tagged jets and soft b-tag objects		$t\bar{t}$, $t\bar{t}W^{\pm}$
$\Delta\phi\left(\vec{p}_T^{\,3\ell}, \vec{p}_T^{\,miss}\right)$	—	>2.5	$t\bar{t}$
Dielectron mass	$\|m_{ee} - m_Z\| > 20$ GeV	—	charge flips
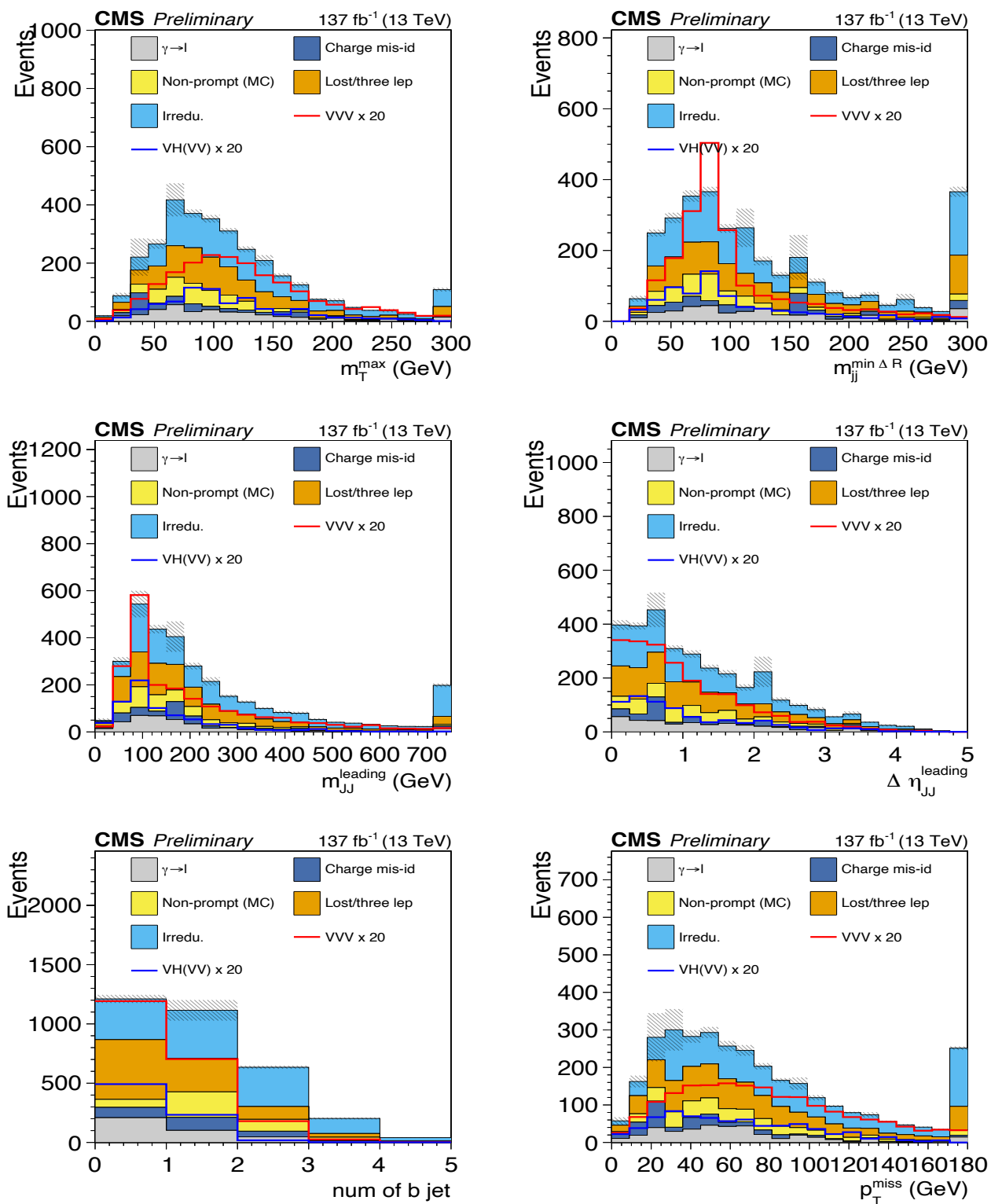
Figure 7.7: Some discriminant variables for the VVV and VH(VV) signal and the backgrounds in MC samples for the SS + 2 jets events. The selection applied in the plots requires two *Nominal* SS leptons with $p_T > 25\,\text{GeV}$, and two jets.
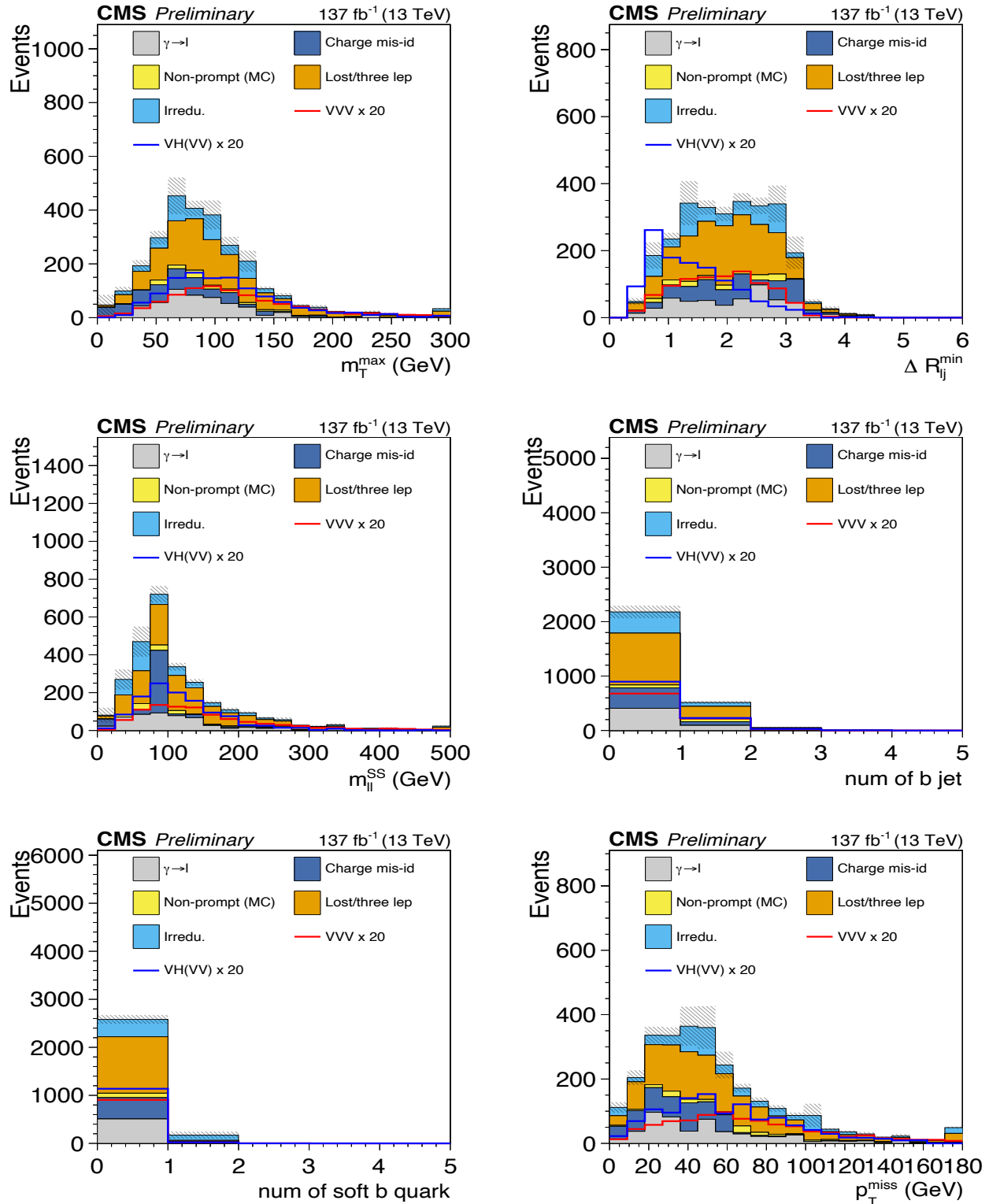
Figure 7.8: Some discriminant variables for the VVV and VH(VV) signal and the backgrounds in MC samples for the SS + 1 jet events. The selection applied in the plots requires two *Nominal* SS leptons with $p_T > 25$ GeV, and one jet.
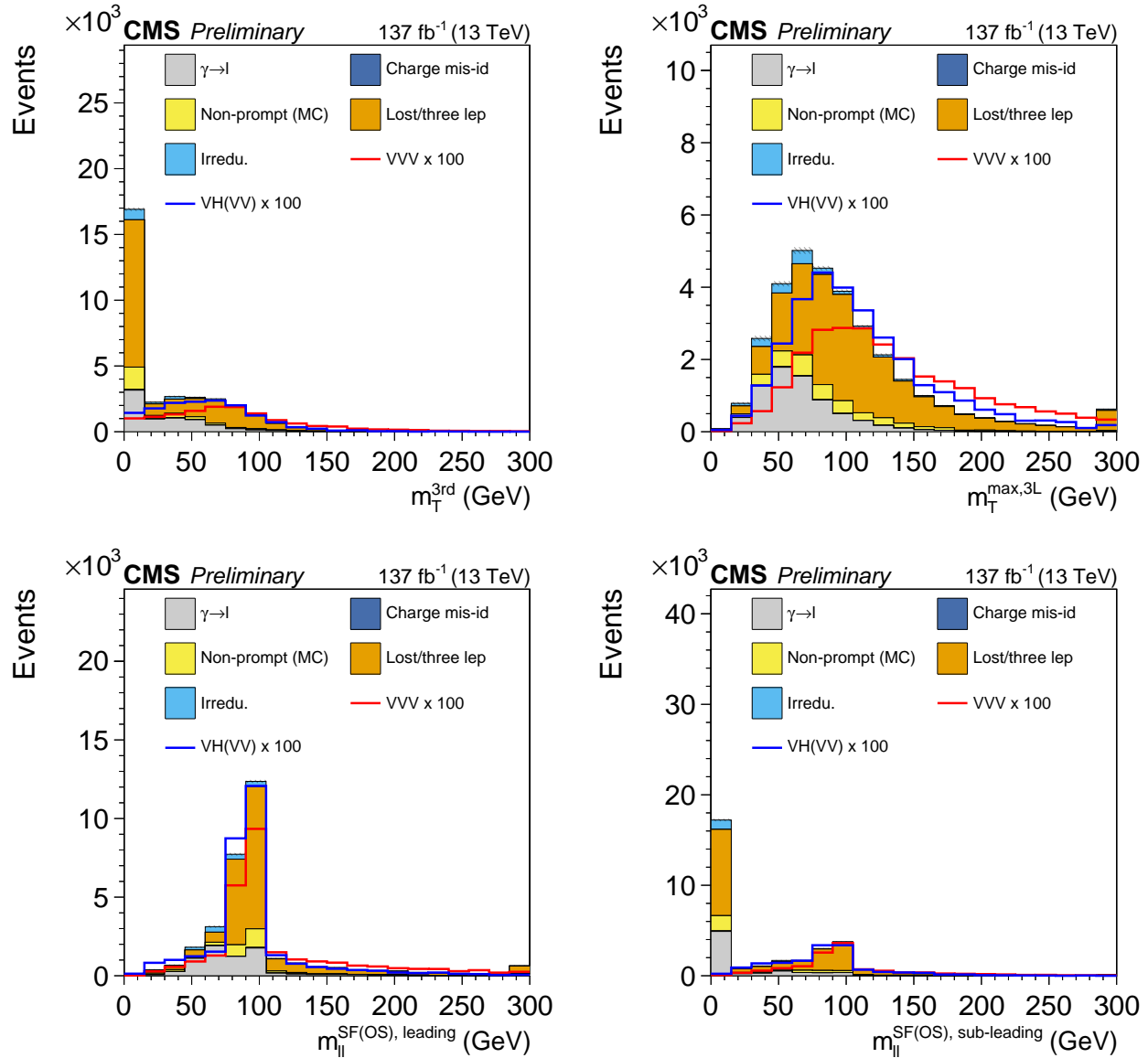
Figure 7.9: Some discriminant variables for the VVV and VH(VV) signal and the backgrounds in MC samples for the three-lepton events. The selection applied in the plots requires three *Nominal* leptons with $p_T > 25/20/20$ GeV.
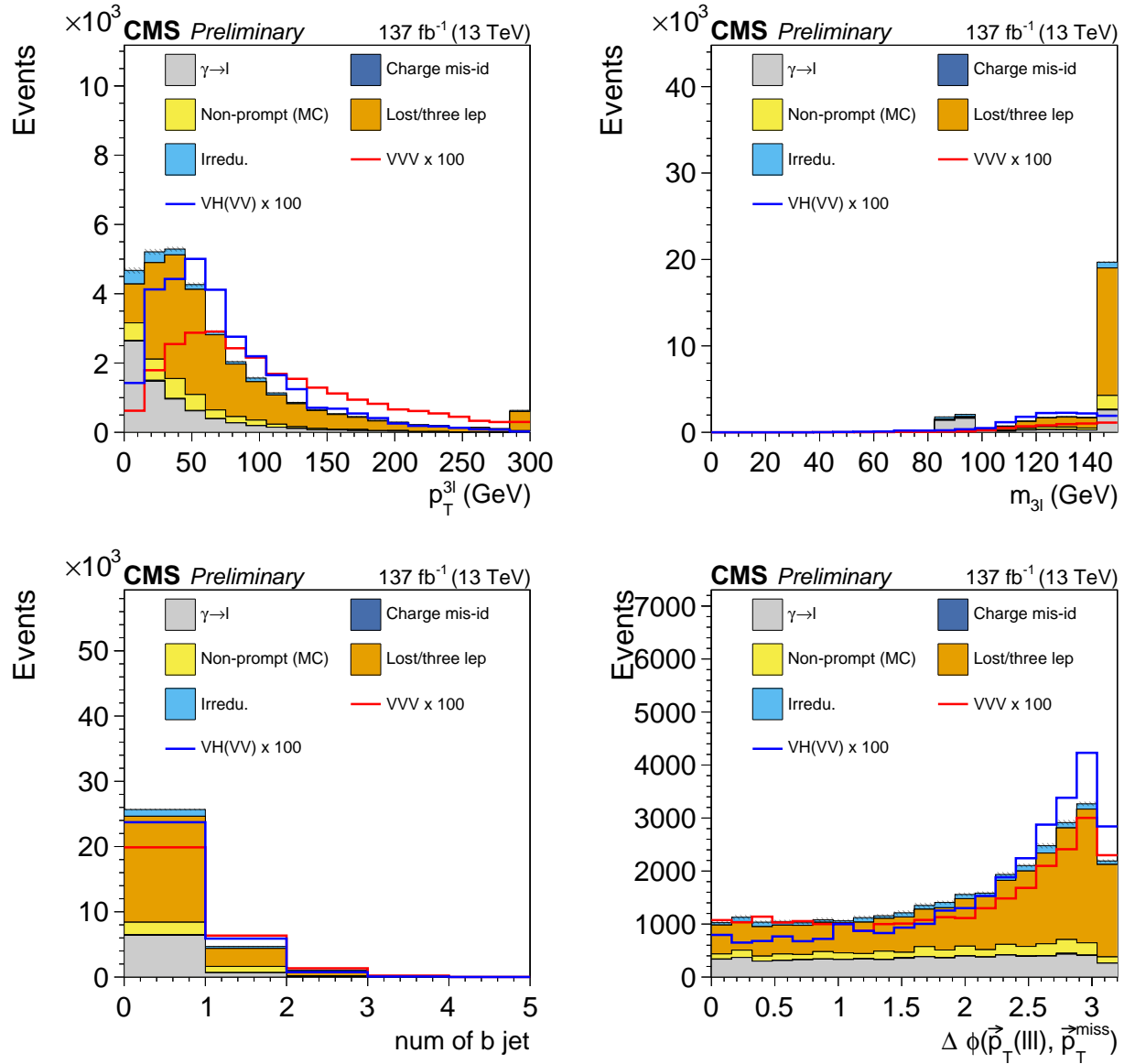
Figure 7.10: More discriminant variables for the VVV and VH(VV) signal and the backgrounds in MC samples for the three-lepton events. The selection applied in the plots requires three *Nominal* leptons with $p_T > 25/20/20$ GeV.

7.4.2 Same-sign dilepton and three-lepton selections (BDT-based)

To better suppress the backgrounds while maintaining high VVV signal efficiency, we trained two sets of Boosted Decision Trees (BDTs), one targeting backgrounds from events with prompt leptons (such as WZ, ZZ, $W^{\pm}W^{\pm}$), and another targeting backgrounds due to nonprompt or fake leptons (such as $t\bar{t}$ or photons misidentified as leptons). The two sets of BDTs are called prompt BDT and fake BDT, respectively. The event categories are kept the same as those used in the cut-based selection, and the kinematic selections applied in the cut-based method are replaced by cuts on the prompt BDT and fake BDT in each category.

Due to different available kinematic variables in the different categories, the BDTs are trained separately in the different categories: one for SS-2j events, one for SS-1j events, and one for 3ℓ events. The list of input variables used for training the different BDT's are summarized in Table 7.9.

Table 7.9: List of input variables for SS/3ℓ BDT training.

variables	prompt BDT			fake BDT		
	SS2J	SS1J	3ℓ	SS2J	SS1J	3ℓ
p_T^{miss}	✓	✓	✓	✓	✓	✓
leptons-p_T	✓	✓	✓	✓	✓	✓
leptons flavor	✓	✓	✓	✓	✓	✓
jets-p_T	✓	✓	✓	✓	✓	✓
$m_{\ell\ell}$ (SS (SF) pair for SS (3ℓ))	✓	✓	✓	✓	✓	✓
m_T^{max}	✓	✓	✓	✓	✓	✓
m_T^{min}	✓	✓		✓	✓	
$p_T(\ell\ell)$ (SS pair)	✓	✓		✓	✓	
m_{jj} (closest in ΔR)	✓			✓		
m_{JJ} (leading jets)	✓			✓		
m_{jj} (ΔR closest)	✓			✓		
$\Delta\eta_{JJ}$ (leading jets)	✓			✓		
ΔR_{JJ} (leading jets)	✓			✓		
ΔR_{jj} (closest in ΔR)	✓			✓		
number of jets	✓		✓	✓		✓
m_T^{3rd}			✓			✓
$m_{3\ell}$			✓			✓
$p_T^{3\ell}$			✓			✓
$\Delta\phi\left(\vec{p}_T^{\,3\ell},\vec{p}_T^{\,miss}\right)$			✓			✓
n_{SFOS}			✓			✓

The XGBoost open-source library [129] is used for the BDT training. To have

enough training statistics from simulation, events passing very loose preselection criteria are used for training. In SS2J, we require two SS leptons and at least two jets, in SS1J, we require two SS leptons and one jet, and for 3ℓ, we require 3 leptons and $|m_{\mathrm{SFOS}} - m_Z| > 20\,\mathrm{GeV}$. Table 7.10 summarizes the selections applied for the trained events and the approximate MC statistics that were used in the BDT training for signal and backgrounds (40% of the events are used for training in each case). The decision trees have a maximum depth of 3, and in total 400 trees (300 for SS1J training, where the statistics are low) are used and averaged for the BDT output. Some other parameters of the XGBClassifier are: booster = gbtree, learning_rate = 0.1, min_split_loss = 0, reg_lambda = 1, max_leaves = 0 (most of these are the default parameter choices in the XGBoost library).

The output of the BDT is the probability of the event to be a signal event (taken from the predict_proba function of the XGBClassifier, which is a logarithmic transformation of the output of the trees S: $\frac{1}{1+\exp(-S)}$), which has been constructed to provide a value ranging from 0 to 1.

Table 7.10: Selection applied for training events and approximate simulation statistics used for the SS1j, SS2j, and the 3ℓ BDT training.

selection	SS2J	SS1J	3ℓ
	two SS leptons ($p_{\mathrm{T}} > 25/25\,\mathrm{GeV}$)		3 leptons ($p_{\mathrm{T}} > 25/20/20\,\mathrm{GeV}$)
	$m_{\ell\ell} > 20\,\mathrm{GeV}$		$|m_{\mathrm{SFOS}} - m_Z| > 20\,\mathrm{GeV}$
	$N_{\mathrm{j}} \geq 2$	$N_{\mathrm{j}} = 1$	
signal MC	42×10^3	32×10^3	48×10^3
prompt background MC	667×10^3	185×10^3	357×10^3
fake background MC	11×10^3	3.7×10^3	28×10^3

Figures 7.11 and 7.12 show the ROC curve (a curve of signal vs. background efficiency obtained with different BDT cuts) and the BDT response of the SS1j, SS2j, and 3ℓ BDTs, respectively. Figure 7.14 shows the 2D distribution of the prompt and fake BDTs for different processes as well as different categories. From this, we can see that signal events are populated in regions with both large prompt BDT and large fake BDT score, while background events have either small prompt BDT or small fake BDT score. A rectangular cut on the two BDT scores can therefore be used to remove most of the backgrounds while maintaining high signal efficiency.

The ranking of input variables is given in Figure 7.13.

As mentioned above, we keep the overall classification of the twelve signal regions as in the cut-based approach, and we replace all selections (including selections on

$m_{\mathrm{T}}^{\mathrm{3rd}}/m_{\mathrm{T}}^{\mathrm{max}}$, $p_{\mathrm{T}}^{\mathrm{miss}}$, $\Delta\phi\left(\vec{p}_{\mathrm{T}}^{3\ell}, \vec{p}_{\mathrm{T}}^{\mathrm{miss}}\right)$, $\Delta\eta_{\mathrm{JJ}}$, m_{JJ} (leading jets), $\Delta R_{\ell\mathrm{j}}^{\mathrm{min}}$, and $p_{\mathrm{T}}^{3\ell}$) with selections on the prompt and fake BDT discriminants. Tables 7.11 and Table 7.12 summarize all the kinematics selections as well as the BDT cuts for the SS and 3ℓ channels in the BDT-based analysis. The selection cuts on the BDTs are optimized in each signal region (9 regions in SS, 3 in 3ℓ) individually to optimize the expected significance in each signal region. The final cuts on the prompt and fake BDTs are chosen by scanning both BDT discriminants simultaneously and choosing the cut on each BDT discriminant for which we get the best combined expected significance.

Table 7.11: Event selections for the SS channels using the BDT method: for each category, we define three signal regions depending on the lepton flavor: $e^{\pm}e^{\pm}$, $e^{\pm}\mu^{\pm}$, and $\mu^{\pm}\mu^{\pm}$. This results in $3 \times 3 = 9$ signal regions for the SS channels.

Variable	m_{jj}-in			m_{jj}-out			1j		
	$e^{\pm}e^{\pm}$	$e^{\pm}\mu^{\pm}$	$\mu^{\pm}\mu^{\pm}$	$e^{\pm}e^{\pm}$	$e^{\pm}\mu^{\pm}$	$\mu^{\pm}\mu^{\pm}$	$e^{\pm}e^{\pm}$	$e^{\pm}\mu^{\pm}$	$\mu^{\pm}\mu^{\pm}$
Trigger	Pass the dilepton triggers								
Signal leptons	Exactly 2 tight SS leptons with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$								
Additional leptons	No additional very loose lepton								
Isolated tracks	No additional isolated tracks								
$m_{\ell\ell}$	$>20\,\mathrm{GeV}$								
$m_{\ell\ell}$	$\|m_{\ell\ell} - m_{\mathrm{Z}}\| > 20\,\mathrm{GeV}$ if $e^{\pm}e^{\pm}$								
b-tagging	no b-tagged jets and soft b-tag objects								
Jets	≥ 2 jets			≥ 2 jets			1 jet		
m_{jj} (closest ΔR)	$65 < m_{\mathrm{jj}} < 95\,\mathrm{GeV}$			$\|m_{\mathrm{jj}} - 80\,\mathrm{GeV}\| \geq 15\,\mathrm{GeV}$			—		
Prompt BDT (>)	0.22	0.32	0.14	0.07	0.19	0.09	0.32	0.26	0.39
fake BDT (>)	0.00	0.89	0.77	0.75	0.82	0.73	0.68	0.89	0.64

Table 7.12: Event selections for the three-lepton channels of BDT method.

Variable	0 SFOS	1 SFOS	2 SFOS
Trigger	Pass dilepton triggers		
Signal leptons	3 *Nominal* leptons with charge sum = ±1e $p_{\mathrm{T}} > 25/25/25\,\mathrm{GeV}$ $\quad$ $p_{\mathrm{T}} > 25/20/20\,\mathrm{GeV}$		
Additional leptons	No additional very loose lepton		
m_{SFOS}	$m_{\mathrm{SFOS}} > 20\,\mathrm{GeV}$ and $\|m_{\mathrm{SFOS}} - m_{\mathrm{Z}}\| > 20\,\mathrm{GeV}$		
$m_{3\ell}$	$\|m_{3\ell} - m_{\mathrm{Z}}\| > 10\,\mathrm{GeV}$		
SF lepton mass	$>20\,\mathrm{GeV}$	—	—
Dielectron mass	$\|m_{\mathrm{ee}} - m_{\mathrm{Z}}\| > 20\,\mathrm{GeV}$	—	—
Jets	≤ 1 jet	0 jets	0 jets
b-tagging	No b-tagged jets and soft b-tag objects		
Prompt BDT (>)	0.28	0.60	0.38
fake BDT (>)	0.93	0.80	0.78
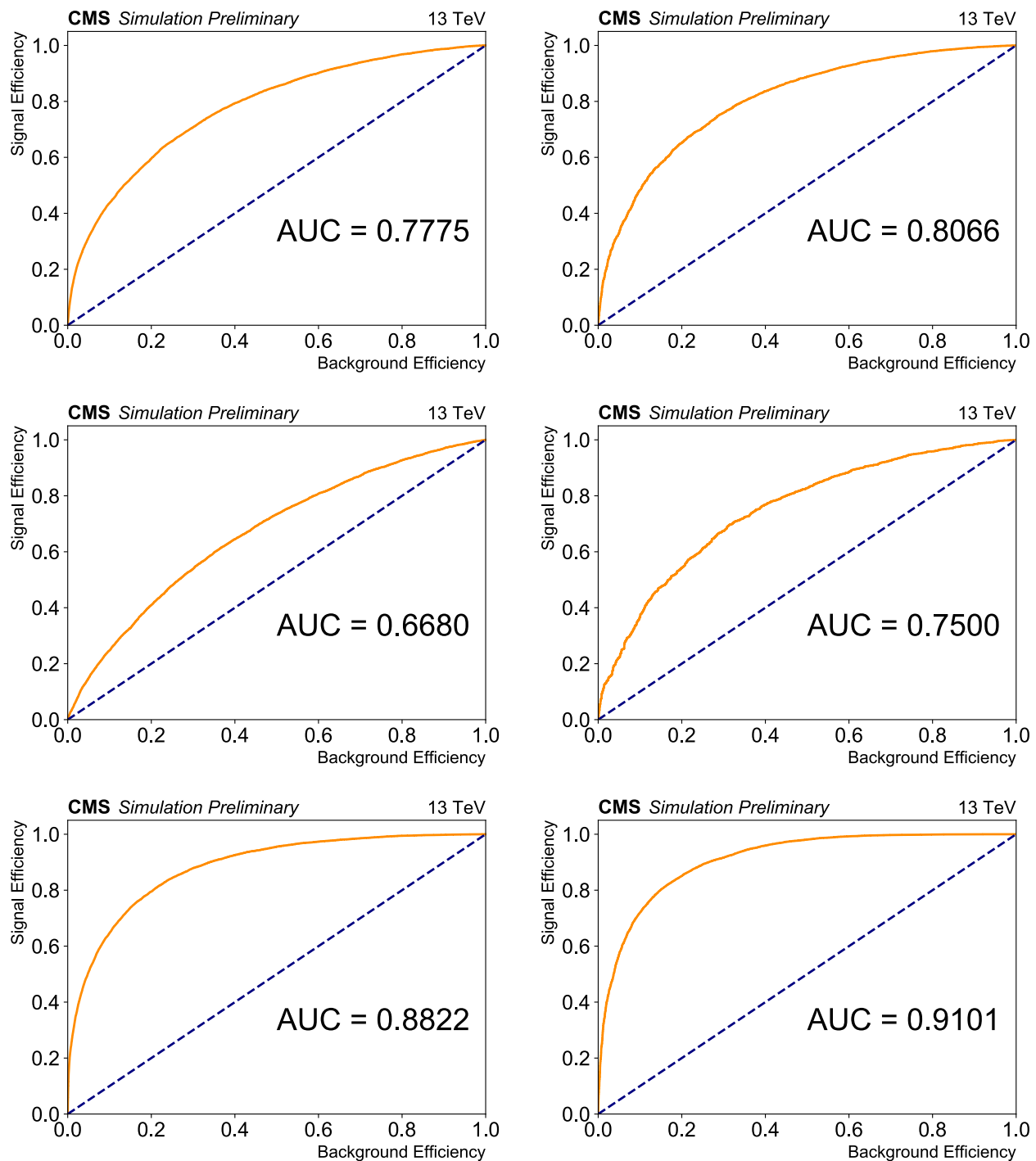
Figure 7.11: ROC curve corresponding to the SS/3ℓ BDT cuts. Left column: prompt BDT; right column: fake BDT. Top: SS2J; middle: SS1J; bottom: 3ℓ.
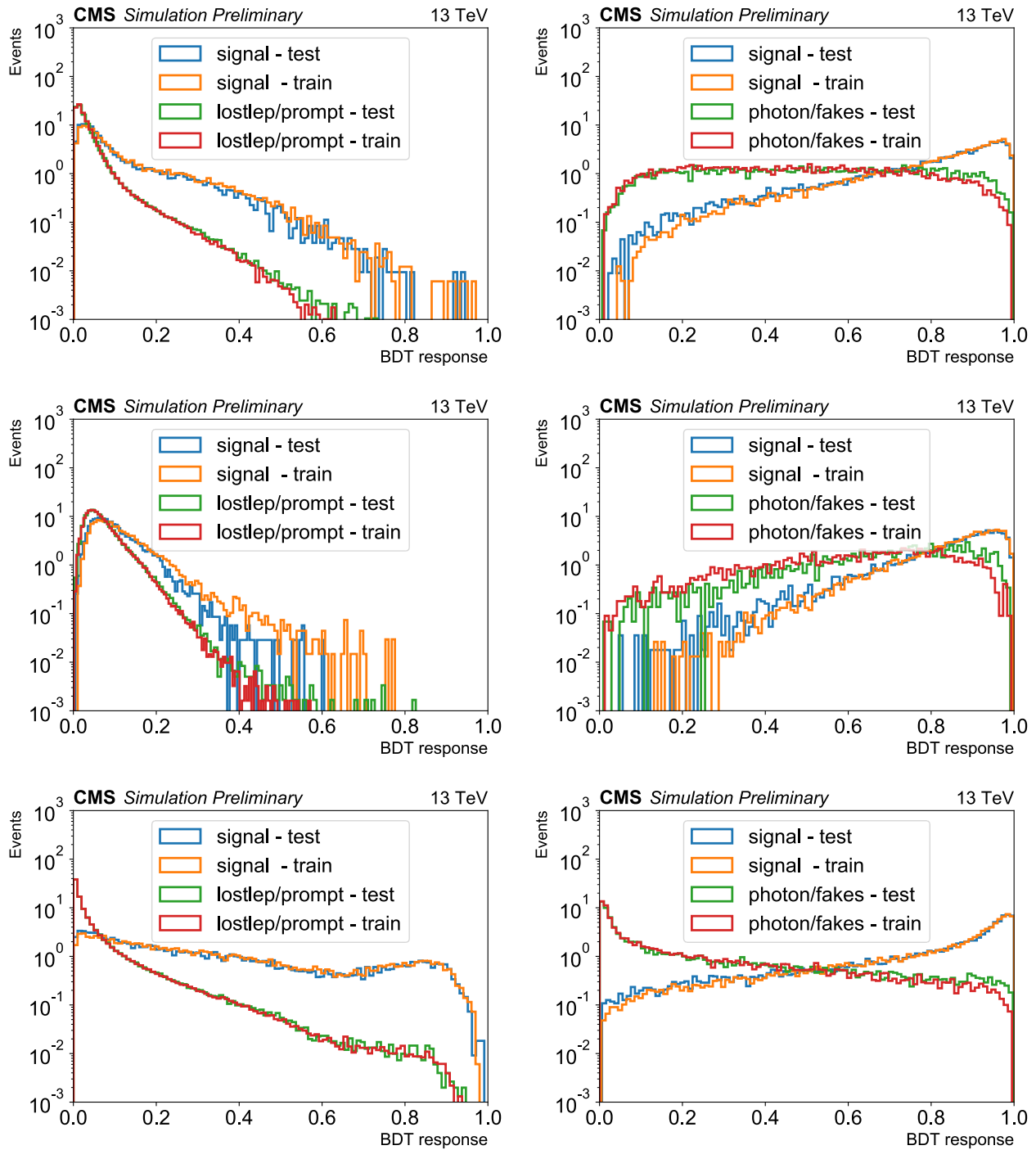
Figure 7.12: SS/3ℓ BDT response for signal and backgrounds in training and testing samples. Left column: prompt BDT; right column: fake BDT. Top: SS2J; middle: SS1J; bottom: 3ℓ.
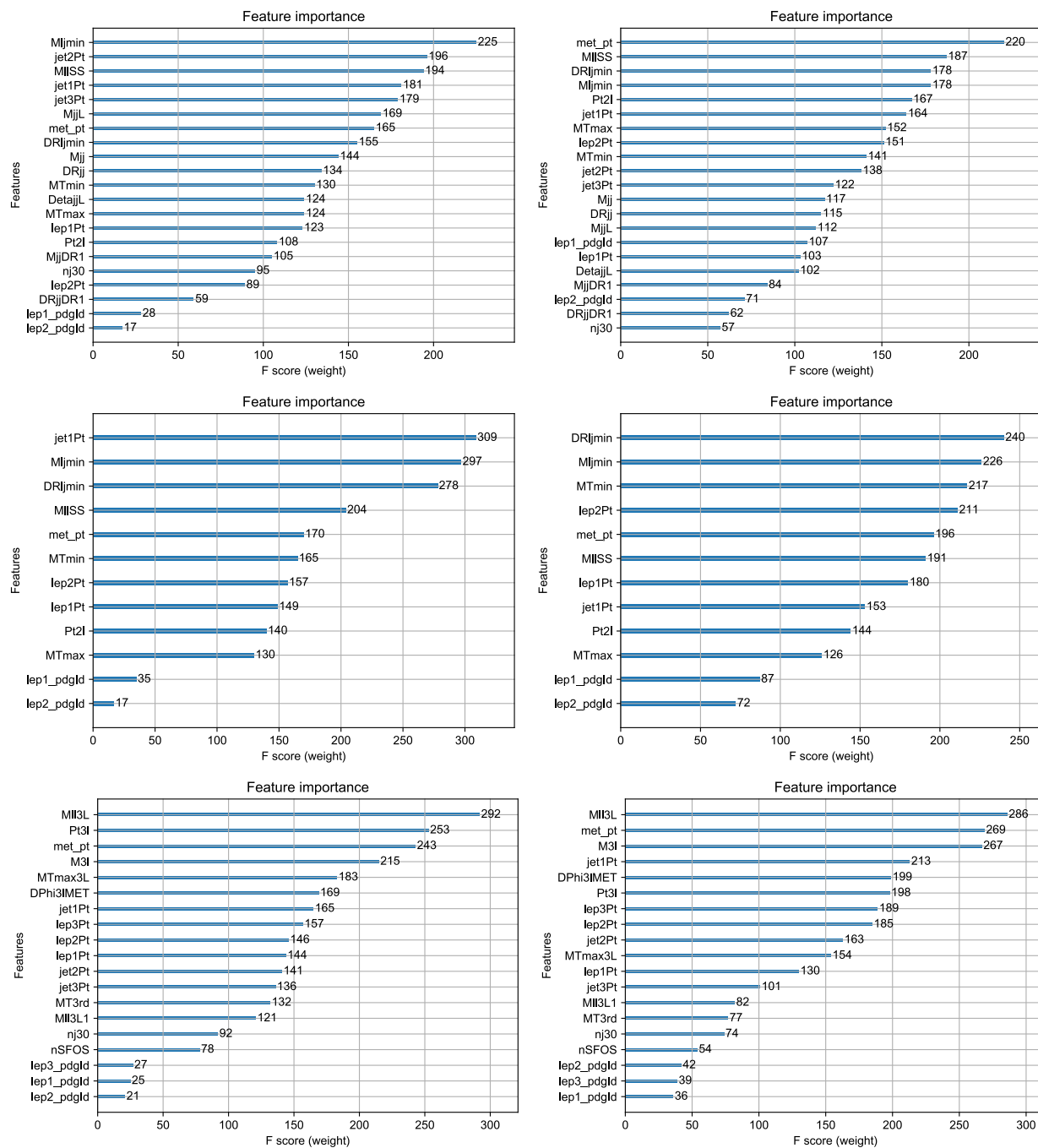
Figure 7.13: Input variable importance scores (calculated by the number of times the feature appears in a tree) for the SS1j, SS2j, and 3ℓ BDT training. Left column: prompt BDT; right column: fake BDT. Top: SS2J; middle: SS1J; bottom: 3ℓ.
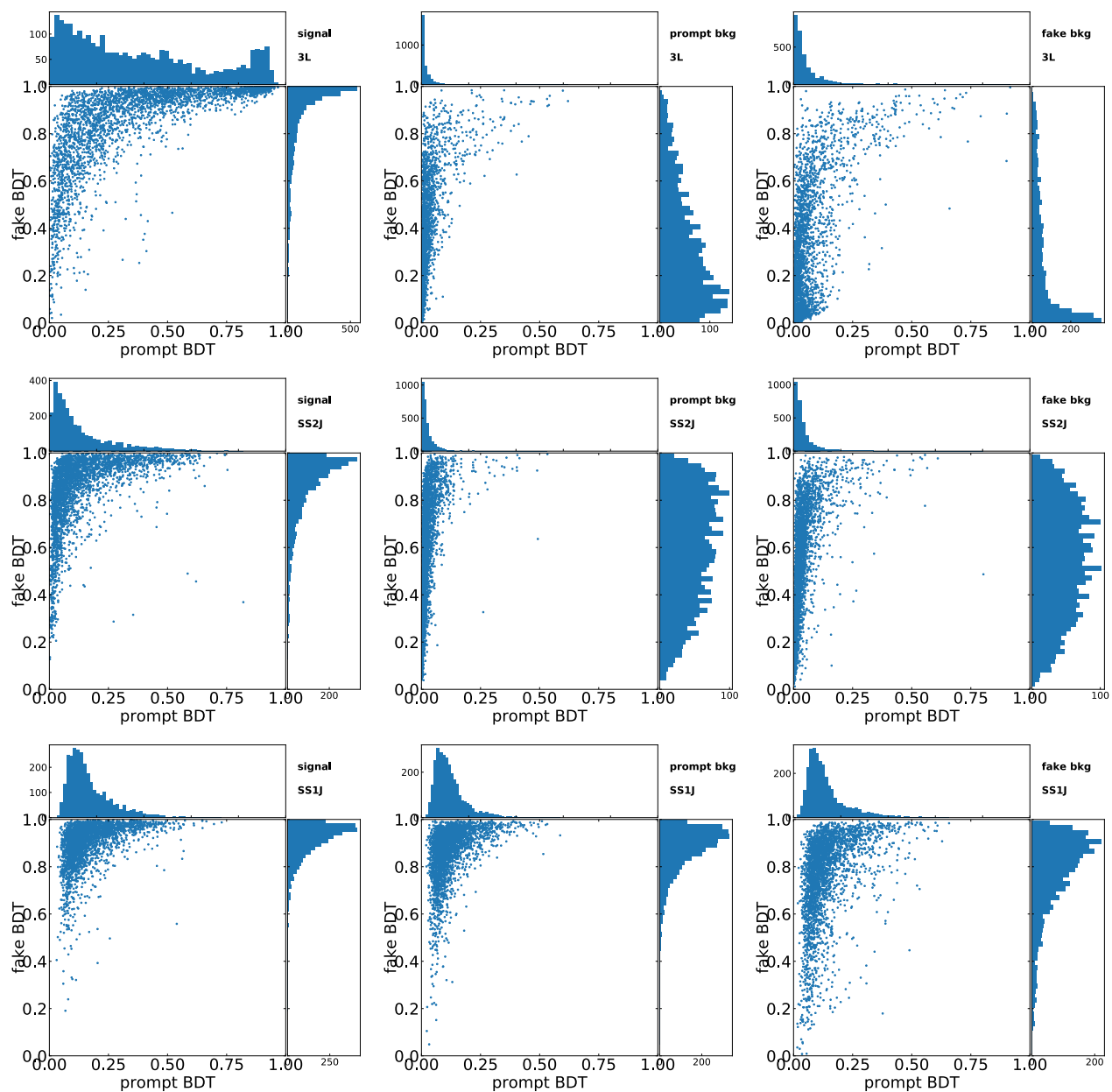
Figure 7.14: 2D distribution of the prompt BDT (X-axis) and fake BDT (Y-axis). Left: signal events; middle: prompt background events; right: fake background events. Top: three-lepton category (selection: three *Nominal* leptons with $p_T > 25/20/20$ GeV); middle: SS2J category (selection: two *Nominal* SS leptons with $p_T > 25$ GeV, and two jets); bottom: SS1J category (selection: two *Nominal* SS leptons with $p_T > 25$ GeV, and one jet).

7.4.3 Four-lepton selections (cut-based)

As discussed in Section 7.1, events in the four-lepton final state are further categorized based on the flavor of the two W candidate leptons, into the eμ category and the ee/μμ category. In both categories, ZZ events are the dominant backgrounds. In the eμ category, the main backgrounds are from $Z(\ell\ell)Z(\tau\tau)$ decay (where $\ell = e/\mu$) and in the ee/μμ category, the main backgrounds are from $Z(\ell\ell)Z(\ell\ell)$ decay (where $\ell = e/\mu$).

Several kinematic selections are applied to suppress the ZZ backgrounds. In the eμ category, where the ZZ decays as $Z(\ell\ell)Z(\tau\tau)$, the invariant mass of the eμ pair ($m_{e\mu}$) for ZZ events typically peaks around $50\,\text{GeV}$, and there are very few events above $100\,\text{GeV}$. This is in contrast to the invariant mass distribution of the eμ pair ($m_{e\mu}$) in WWZ events, which is much wider and which has more events in the high $m_{e\mu}$ region. Therefore, we use $m_{e\mu}$ as the main discriminating variable to categorize events in the eμ category. In addition, the m_{T2} variable [130] is utilized to suppress the ZZ background process. For our case (where the invisible particles are neutrinos), m_{T2} is defined as:

$$m_{T2} = \min_{\vec{p}_T^{\nu(1)} + \vec{p}_T^{\nu(2)} = \vec{p}_T^{\text{miss}}} \left[\max \left(m_T^{(1)}(\vec{p}_T^{\nu(1)}, \vec{p}_T^{e}), m_T^{(2)}(\vec{p}_T^{\nu(2)}, \vec{p}_T^{\mu}) \right) \right], \qquad (7.1)$$

In events where two particles are pair-produced and each of the particles decays semi-invisibly, one cannot simply build m_T. The idea behind the definition above is that for correct transverse masses $m_T^{(1,2)}$, the larger of the two is still smaller than or equal to the parent mass. The minimization over all possible splittings of p_T^{miss} to the transverse momenta of the unseen particles, $\vec{p}_T^{\nu(1,2)}$, ensures that each transverse mass does not exceed the parent mass. The m_{T2} value tends to be small for the ZZ background process, since the tau lepton decay products are generally collinear with each other. Figure 7.15 shows the $m_{e\mu}$ and m_{T2} distributions for both signal and background events in the eμ category, from which we can see that a combination of binning using those two variables can be constructed to give a good signal to background ratio.

In the ee/μμ category, the ZZ background decays as $Z(\ell\ell)Z(\ell\ell)$. Compared to the WWZ signal, such ZZ events will have a much smaller p_T^{miss}, since more of the energy in the signal events is carried away by the neutrinos from W-decay. Therefore, p_T^{miss} is the main discriminating variable in the ee/μμ category. In addition, the transverse momentum of the four-lepton system, $p_T^{4\ell}$, is also used as another discriminating variable. This variable is similar to p_T^{miss}, but as $p_T^{4\ell}$ is

measured purely from leptons, it is therefore likely to be more precise and well-modelled, and it gives complementary information to p_T^{miss}. Figure 7.16 shows the $p_T^{4\ell}$ and p_T^{miss} distributions for both signal and background events in the ee/μμ category, and the binning of ee/μμ events is then based on those two variables.

The subdominant background in the four-lepton final states are events from $t\bar{t}Z$ decay, where two leptons come from Z and two leptons come from the two t-quark decays. To suppress such backgrounds, the events are required to have zero b-tagged jets.

To remove any events stemming from low mass resonances coming from QCD processes with large cross sections, we require the invariant mass of any opposite charged lepton pair to be greater than 12 GeV.

Table 7.13 summarizes the event selection and binning for both the eμ and ee/μμ categories in the four-lepton final state. In total, there are four bins in the eμ category and three bins in the ee/μμ category.

Table 7.13: Event selections for the four-lepton channels (eμ and ee/μμ).

	eμ category	ee/μμ category
Trigger	Pass the dilepton triggers	
Z leptons	ZID SFOS pair, $p_T > 25/15$ GeV	
W leptons	WID OS pair, $p_T > 25/15$ GeV	
W leptons flavor	eμ	ee/μμ
b-tagging	No b-tagged jets	
$m_{\ell\ell}^{Z-tag}$ (Z pair)	$\|m_{\ell\ell}^{Z-tag} - m_Z\| < 10$ GeV	
$m_{\ell\ell}$ (WW pair)	—	$\|m_{\ell\ell} - m_Z\| > 10$ GeV
$m_{\ell\ell}^{OS}$	$m_{\ell\ell}^{OS} > 12$ GeV for any OS dilepton pair	
Bin 1	$m_{e\mu} < 40$ GeV	—
Bin 2	$40 < m_{e\mu} < 60$ GeV	—
Bin 3	$60 < m_{e\mu} < 100$ GeV	—
Bin 4	$m_{e\mu} > 100$ GeV, $m_{T2} > 25$ GeV	—
Bin A	—	$p_T^{miss} > 120$ GeV
Bin B	—	$70 < p_T^{miss} < 120$ GeV, $p_T^{4\ell} > 70$ GeV
Bin C	—	$70 < p_T^{miss} < 120$ GeV, $40 < p_T^{4\ell} < 70$ GeV
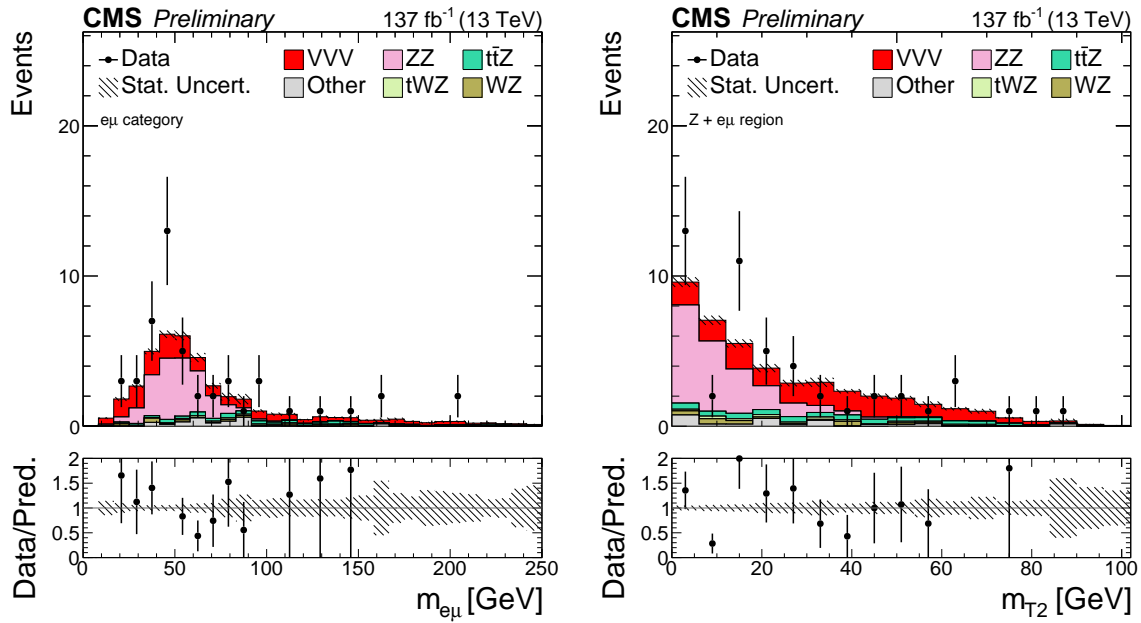
Figure 7.15: Distribution of the invariant mass of the two lepton candidates from two W decays $m_{e\mu}$ (left), and the distribution of m_{T2} (right) in the four-lepton final state $e\mu$ category.
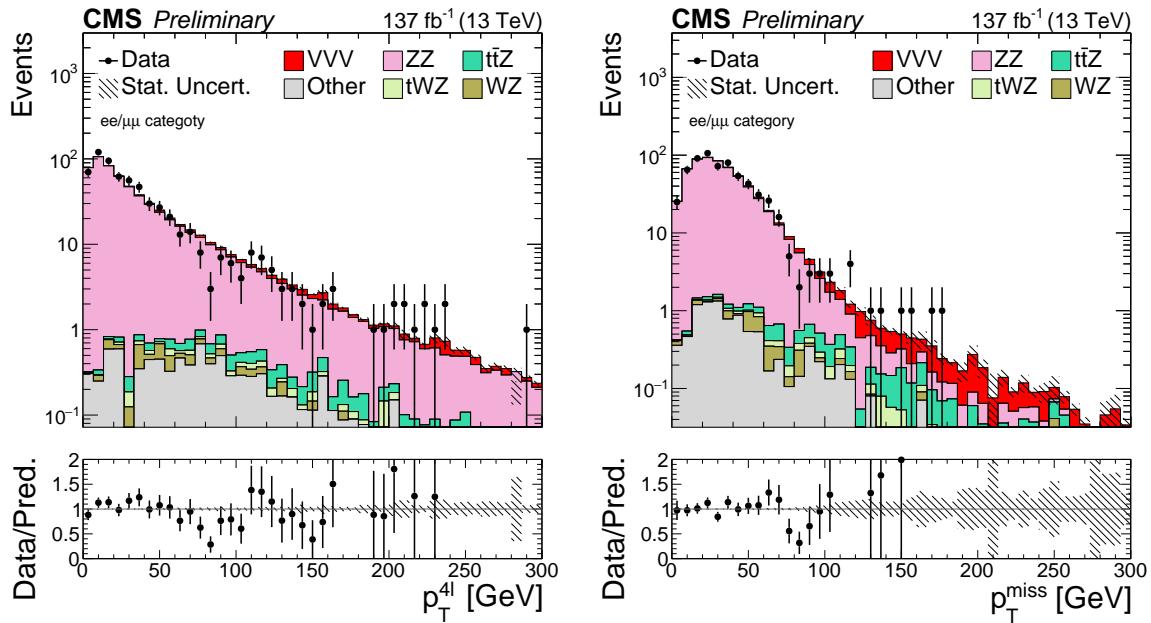
Figure 7.16: Distribution of the p_T of the four-lepton system $p_T^{4\ell}$ (left), and the distribution of p_T^{miss} (right) in the four-lepton final state $ee/\mu\mu$ category.

7.4.4 Four-lepton selections (BDT-based)

Similarly to the selection used for the same-sign dilepton and three-lepton final states, a BDT-based selection is also employed for the four-lepton final states. Two sets of BDTs are trained for the four-lepton events, targeting two major backgrounds, ZZ and $t\bar{t}Z$. Both the ZZ and $t\bar{t}Z$ BDTs are trained separately for the $e\mu$ and $ee/\mu\mu$ events. Table 7.14 lists all the input variables for the ZZ and $t\bar{t}Z$ BDT trainings. For $t\bar{t}Z$ events, the main discriminating power comes from the jet related variables, as there are two additional b-quark jets in the $t\bar{t}Z$ decay products. The lepton variables are used for both the ZZ BDT and $t\bar{t}Z$ BDT training. For the ZZ BDT, p_T^{miss}-related variables are also included in the training input.

It has been shown [131] that for $Z \rightarrow \tau\tau$ events, the angle between the neutrinos and the e/μ from τ lepton decay is typically small, which results in a large projection of $\vec{p}_T^{\text{miss}}$ on the e/μ direction; while this is usually not true for the WWZ signal events. Therefore, two variables are defined for the input to the ZZ BDT in this case:

$$P_\zeta = (\vec{p}_T^{\text{miss}} + \vec{p}_T^{\ell 1} + \vec{p}_T^{\ell 2}) \cdot \hat{\zeta} \qquad \text{and} \qquad P_\zeta^{\text{vis}} = (\vec{p}_T^{\ell 1} + \vec{p}_T^{\ell 2}) \cdot \hat{\zeta}, \qquad (7.2)$$

where the two leptons $\ell 1$ and $\ell 2$ denote the two W lepton candidates, and $\hat{\zeta}$ is a unit vector in the direction of the bisector of $\ell 1$ and $\ell 2$, as illustrated in Figure 7.17.
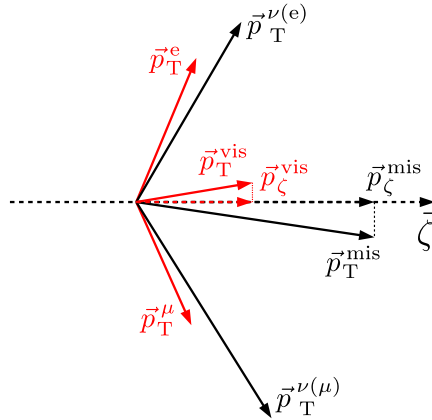
Figure 7.17: Projections of $\vec{p}_T^{\text{miss}}$ and leptons $\vec{p}_T$ on the $\hat{\zeta}$ direction.

The XGBoost open-source library [129] is used in the BDT training. ZZ, $t\bar{t}Z$, and WWZ MC events passing the preselection in Table 7.13 (the selections performed before the event binning part) are used in the training. For $t\bar{t}Z$ MC events, the b-veto selection is not applied to the training events in order to increase the training

Table 7.14: List of input variables for the four-lepton BDT training.

variables	$t\bar{t}Z$ BDT	ZZ BDT
$m_{\ell\ell}$ (WW pair)	✓	✓
p_T^Z	✓	✓
m_{T2}	✓	✓
$p_T^{\ell1}$ (W lepton)	✓	✓
$p_T^{\ell2}$ (W lepton)	✓	✓
$p_T^{4\ell}$	✓	✓
$\sum_i^4 p_T^i$	✓	✓
P_ζ		✓
P_ζ^{vis}		✓
p_T^{miss}		✓
$m_T^{\ell1}$ (W lepton)		✓
$m_T^{\ell2}$ (W lepton)		✓
$m_{4\ell}$		✓
$\min \Delta R(j, \ell3)$	✓	
$\min \Delta R(j, \ell4)$	✓	
leading jet p_T	✓	

statistics. In total, about 300,000 WWZ, 18,000 ZZ, and 25,000 $t\bar{t}Z$ MC events are used in the training for the eμ category, and about 28,000 WWZ and 810,000 ZZ MC events are used in the training for the ee/μμ category. The ZZ BDT and $t\bar{t}Z$ BDTs in the eμ category have a maximum depth of 3, and 400 trees are used and averaged to get the BDT output. For the ZZ BDT in the ee/μμ category, the maximum depth is 5 with 400 trees. Unlike the same-sign and three-lepton analysis, the four-lepton BDT output distribution is not put through the logarithmic transformation mentioned above. Figure 7.18 shows the $t\bar{t}Z$ BDT in the eμ category and the ZZ BDT in the eμ and ee/μμ categories, for both the signal and background samples, split into training and testing sets.

With the two BDTs, the $m_{\text{eμ}}$-m_{T2} (p_T^{miss}-$p_T^{4\ell}$)-based event binning in the cut-based method is replaced by ZZ BDT and $t\bar{t}Z$ BDT-based binning for eμ (ee/μμ) events, as the signals are enriched in the high ZZ BDT and high $t\bar{t}Z$ BDT regions. To achieve optimal signal-to-background discrimination, a series of 2D boundaries of the ZZ BDT and the $t\bar{t}Z$ BDT have been defined for the eμ category. The boundaries of the 2D binning have been optimized to maximize sensitivity (s/b) in a way similar to creating a decision tree, and the scanning of the boundaries stops when there are not enough MC events for optimization or the s/b does not increase significantly.
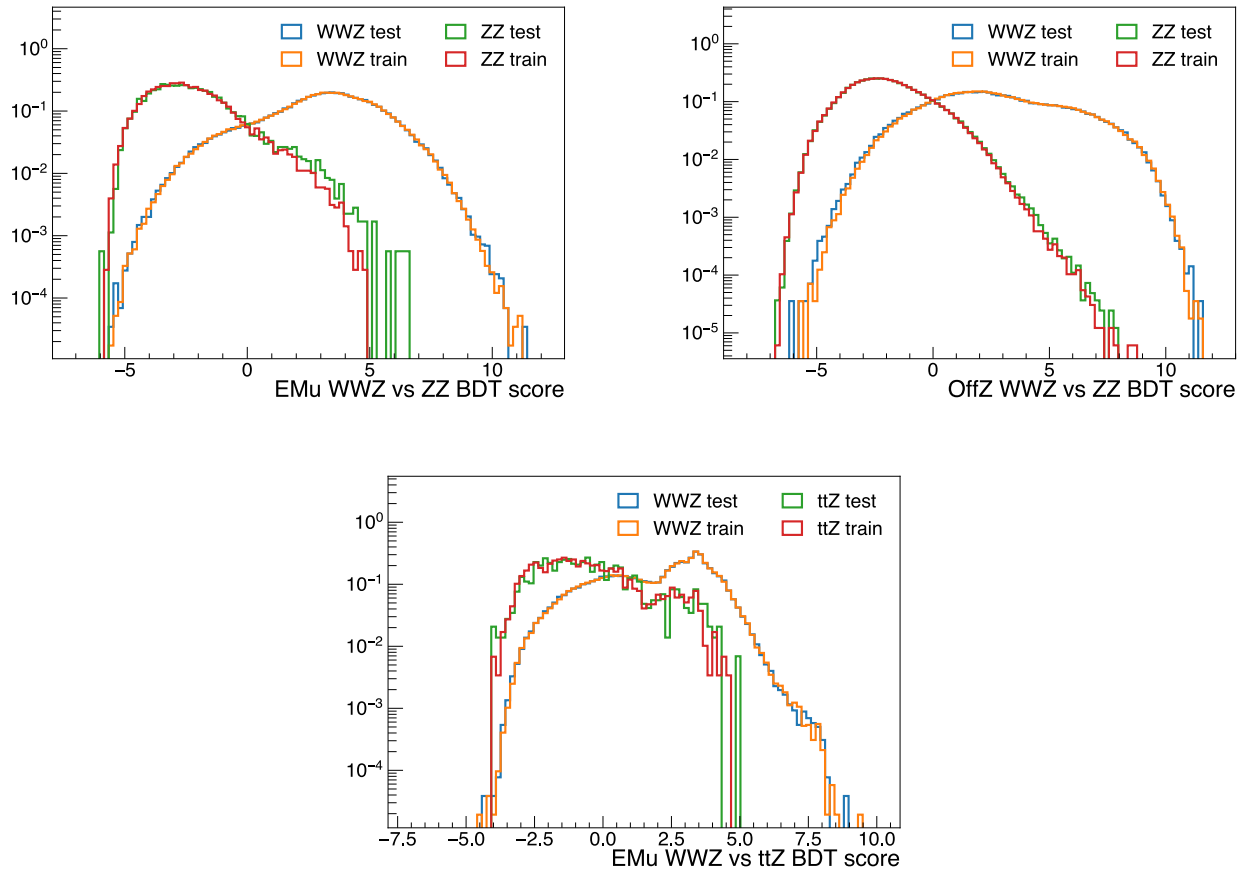
Figure 7.18: Top: $Z\overline{Z}$ BDT response in the four-lepton eμ (left) and ee/μμ (right) categories; bottom: $t\overline{t}Z$ BDT response in the four-lepton eμ category. The distributions for both the WWZ signal and background MC events are shown and are split into training and testing sets.

Figure 7.19 shows the 2D distributions of the $Z\overline{Z}$ and $t\overline{t}Z$ BDTs for signal and background events in the eμ category, with the optimized binning in the 2D plane. For the ee/μμ category, only the $Z\overline{Z}$ BDT is used, and therefore the events are binned based on the $Z\overline{Z}$ BDT score.

Table 7.15 summarizes the event binning based on the BDTs, together with all other selections for the BDT method.
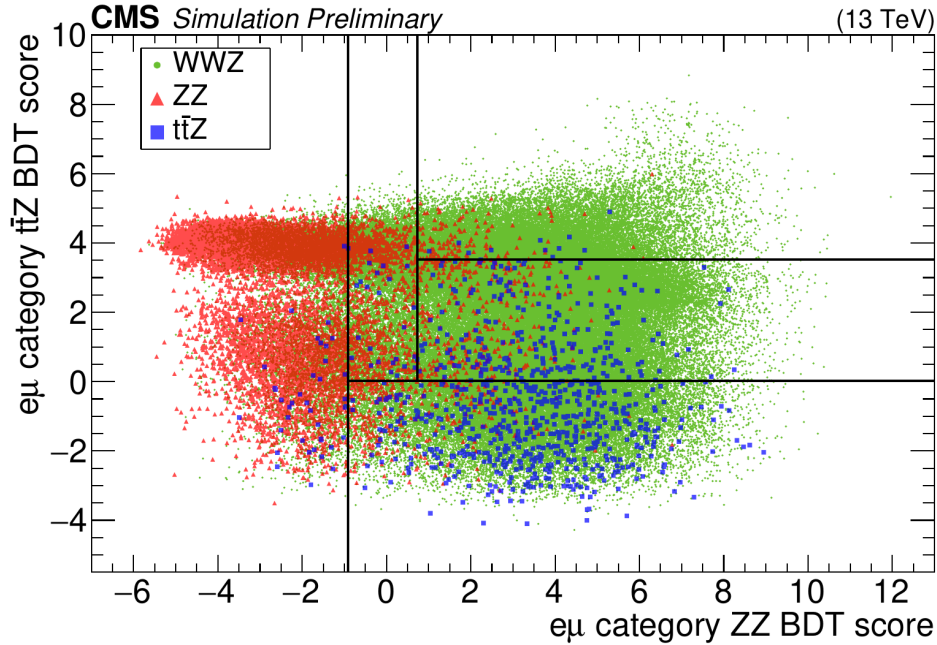
Figure 7.19: The ZZ BDT and $t\bar{t}Z$ BDT 2D distributions of the signal and background events in the four-lepton eμ category, together with the visualized event binning in the 2D plane.

Table 7.15: Event selections for the four-lepton channels (eμ and ee/$\mu\mu$) of the BDT method.

	eμ category	ee/$\mu\mu$ category
Trigger	Pass dilepton triggers	
Z leptons	ZID SFOS pair, $p_\mathrm{T} > 25/15\,\mathrm{GeV}$	
W leptons	WID OS pair, $p_\mathrm{T} > 25/15\,\mathrm{GeV}$	
W leptons flavor	eμ	ee/$\mu\mu$
b-tagging	No b-tagged jets	
$m_{\ell\ell}^{\mathrm{Z-tag}}$ (Z pair)	$|m_{\ell\ell}^{\mathrm{Z-tag}} - m_\mathrm{Z}| < 10\,\mathrm{GeV}$	
$m_{\ell\ell}$ (WW pair)	—	$|m_{\ell\ell} - m_\mathrm{Z}| > 10\,\mathrm{GeV}$
$m_{\ell\ell}^{\mathrm{OS}}$	$m_{\ell\ell}^{\mathrm{OS}} > 12\,\mathrm{GeV}$ for any OS dilepton pair	
Bin 1	ZZ BDT < -0.908	—
Bin 2	ZZ BDT > -0.908, $t\bar{t}Z$ BDT < 0.015	—
Bin 3	$-0.908 <$ ZZ BDT < 0.733, $t\bar{t}Z$ BDT > 0.015	—
Bin 4	ZZ BDT > -0.733, $0.015 <$ $t\bar{t}Z$ BDT < 3.523	—
Bin 5	ZZ BDT > -0.733, $t\bar{t}Z$ BDT > 3.523	—
Bin A	—	$0 <$ ZZ BDT < 3
Bin B	—	ZZ BDT > 3

7.4.5 Five- and six-lepton selections

Events in the five- and six-lepton final state categories are relatively clean and have much lower statistics, so a minimal set of selections has been applied, and only a cut-based selection has been explored for these categories.

For five-lepton events, we require five leptons with four passing the Z and one passing the W lepton ID criteria, as discussed in Section 7.3, and we require the p_T of the leptons from Z decay to be greater than $25/10$ GeV and the p_T of each charged lepton from W decay to be greater than 10 GeV. In addition, events with b-tagged jets are vetoed in order to reject backgrounds from t quarks (such as $t\bar{t}$, $t\bar{t}V$). To remove nonprompt lepton backgrounds from diboson or $t\bar{t}$ processes, we also require the transverse mass of the lepton from W decay to be greater than 50 GeV if the lepton is an electron.

For six-lepton events, we require six *Veto*-ID leptons as discussed in Section 7.3, with the leading two leptons having p_T greater than 25 GeV, and the rest of the four leptons having p_T greater than 10 GeV. In addition, to suppress backgrounds such as $t\bar{t}H$ events, we require the scalar sum of the p_T of all leptons to be greater than 250 GeV.

Figure 7.20 shows the distributions of the two discriminating variables used in the five and six lepton event selection: m_T of the W candidate charged lepton and scalar sum of lepton p_T.
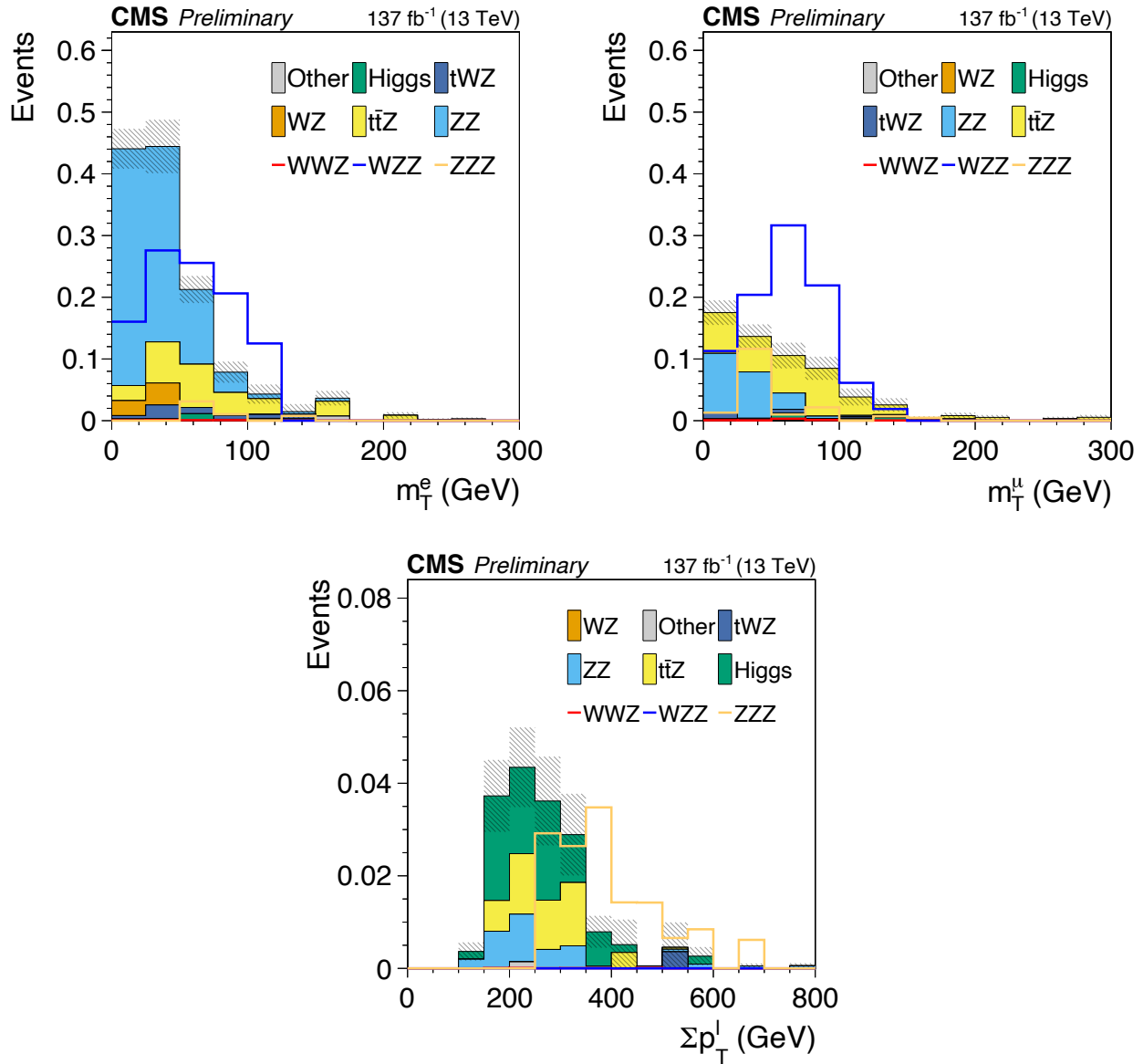
Figure 7.20: Top-left (top-right): distribution of the transverse mass of the W candidate electron (muon) for five-lepton events; bottom: distribution of scalar sum of the p_{T} of all leptons for six-lepton events.

7.5 Background estimation

7.5.1 Backgrounds in same-sign and three-lepton final states

For same-sign and three-lepton events, we expect the following types of backgrounds:

- **Lost/Three-lepton background:** This background is the most important in the SS channel, and stems mostly from WZ production where one lepton from the Z boson decay is not reconstructed ("lost") due to acceptance, ID, or isolation criteria. For the 3ℓ channel, this background comes through an off-shell Z boson.

- **Nonprompt lepton background:** This background contains one (two) prompt leptons and one nonprompt lepton in the SS (3ℓ) channel. The nonprompt lepton most often originates from a semileptonic heavy flavor (HF) hadron decay, but can also originate from light flavor (LF) hadron decays, light flavor hadron misidentification ($\pi^{\pm}$), Dalitz decay of neutral mesons, and conversions from non-isolated photons[1].

- **Irreducible background:** This background is very signal-like. In the SS channel, $W^{\pm}W^{\pm}$ can be produced via vector boson scattering (VBS), double parton scattering (DPS), or in $t\bar{t}+W$ associated production. In the 3ℓ channel, WWW can be produced via $t\bar{t}+W$.

- **Charge misidentification background:** This background mostly stems from events with a Z boson decay, where one lepton's charge is misidentified. This type of background is completely negligible for muons and very tiny for electrons.

- **Other backgrounds:** Other backgrounds, such as events with misidentified prompt photons, where the photons are misidentified as electrons or convert into an electron-positron pair, are expected to be small.

For the two major backgrounds, the nonprompt lepton background and the lost-lepton/three-lepton background, we have developed data-driven estimation methods to predict their contributions to the signal regions. For all other backgrounds, we still validate and normalize the backgrounds in dedicated control regions in data,

---

[1]Note that prompt isolated photons, e.g. isolated photons identified as an electron or converting to a pair of electrons, are identified using generator information and are part of the "Other backgrounds."

although the splitting of their yields into the signal regions is then performed with simulation.

The most sensitive regions are the same-sign $e^\pm\mu^\pm$, $\mu^\pm\mu^\pm$ channels, and the three-lepton 0 SFOS and 1 SFOS channels. In these regions, the background uncertainties are dominated by the nonprompt lepton background or by the lost-lepton/three-lepton background (mostly WZ). Another important background is the irreducible background. All other backgrounds are small. In this section, all background estimation techniques are discussed. The final result, including all background predictions is given in Section 7.7.

To predict the lost-lepton background in the SS channels, we define the WZ control region (CR) as follows: In addition to the two signal leptons, we require the presence of a third lepton. The third lepton needs to pass only the *Nominal-3ℓ* selection criteria. The third lepton has to form a SFOS lepton pair with one of the other two leptons and the pair has to be compatible with being from a Z boson, namely $|m_{\mathrm{SFOS}} - m_Z| < 20\,\mathrm{GeV}$. For the WZ CR of the cut-based method, we apply all the other selections on the SS channel as in Table 7.7, except that we drop the selection $65 < m_{\mathrm{jj}} < 95\,\mathrm{GeV}$ or apply its inverse ($|m_{\mathrm{jj}} - 80\,\mathrm{GeV}| > 15\,\mathrm{GeV}$) for the ≥ 2 jet categories, in order to have sufficient statistics in the control region. The m_{jj} variable will, to first order, have the same distribution in both the signal and control regions, as both jets should stem from intial state radiation (ISR) for the dominant WZ process, or from the same matrix element process (as in the subdominant $t\bar{t}Z$ background). We evaluate the experimental uncertainties on the m_{jj} modeling by varying the jet energy corrections and validating it in data. For the WZ CR of the BDT-based method, we do not merge the m_{jj}-in and m_{jj}-out categories as the BDT cuts on those two categories are different.

To predict the three-lepton background (from WZ or ZZ) in the three lepton channels, we define the WZ control region (CR) as follows. we keep all the selections for the 3ℓ channel as in Table 7.8 (Table 7.12) for the cut-based (BDT-based) method. However, we invert the requirement on the m_{SFOS} variable, i.e. instead of vetoing events where the two leptons are compatible with originating from a Z boson decay, we select such events. By doing so, we have no direct data-driven background estimation for the 0 SFOS channel. However, since a statistically meaningful control region is hard to find (if available at all), we rely on the predictions from the 1 and 2 SFOS channels to extrapolate the 3-lepton background to the 0 SFOS channel.

The control regions mentioned above are mutually exclusive to each other (due to

the requirement on the number of jets), and also orthogonal to the signal regions. However, two SS signal regions with ≥ 2 jets (for the two m_{jj} bins) have an identical control region for the cut-based method. Thus, all nuisances between the two regions for this background are 100% correlated.

Using those control regions, the estimation of the lost-lepton background (in the same-sign channels) or the three-lepton background (in the three-lepton channels) is straightforward:

$$N_{\text{pred}}(\text{lost-lepton/3-lepton}) = \{N_{\text{CR}}(\text{data}) - N_{\text{CR}}^{\text{sim.}}(\text{other bkg})\} \times \text{TF}_{\text{CR}\to\text{SR}}, \quad (7.3)$$

where the transfer factor is obtained from simulation (using WZ MC), but is validated in the control and signal region sidebands.

Figure 7.21 shows the data and MC event yields in the WZ control regions for both the cut-based and BDT-based methods, and good data/MC agreement is observed. The statistical uncertainty of the events in the CR, as well as the statistical uncertainty and systematic uncertainty when measuring the transfer factor, are considered and propagated to the lost-lepton/3ℓ background prediction.
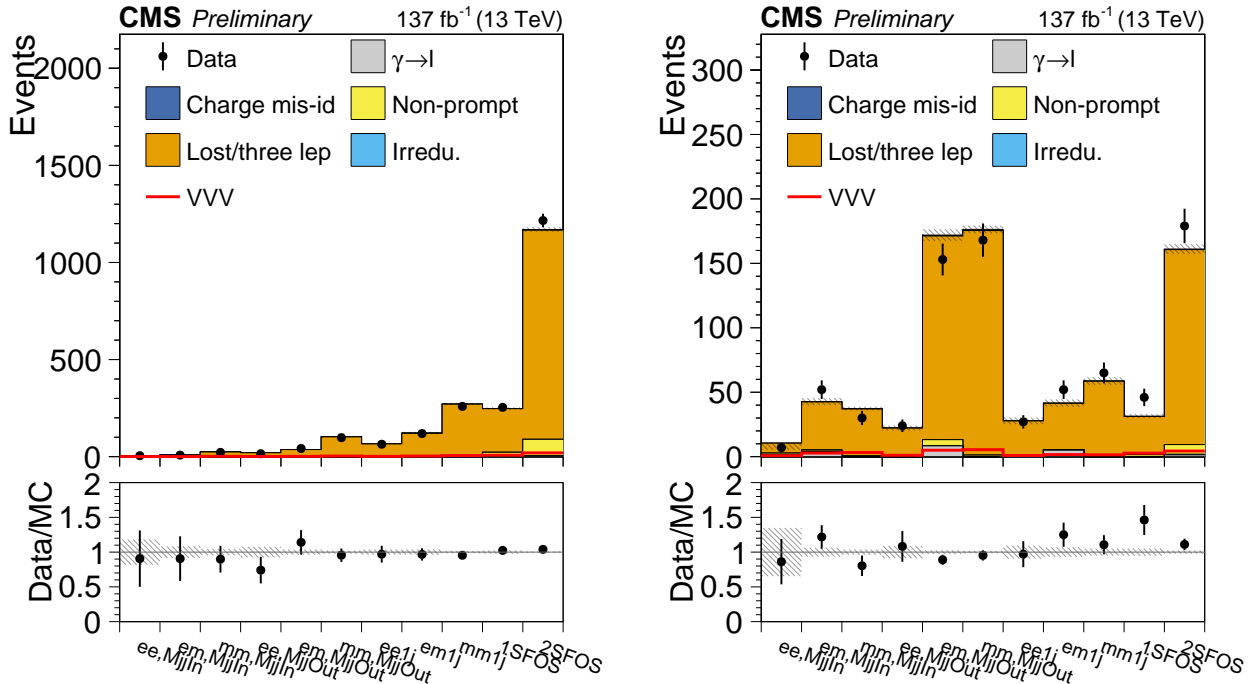
Figure 7.21: Data and simulation yields in the lost-lepton/3ℓ (WZ) control regions for the cut-based (left) and BDT-based (right) methods.

Backgrounds from misidentified leptons are estimated using the tight-to-loose ratio method, also known as the fake rate method. The method estimates the number of misidentified leptons in the signal region by extrapolating across the lepton isolation variable. In addition to the *Nominal* lepton IDs used to define the signal regions (see Tables 7.3 and 7.4), loose lepton IDs are defined by loosening the isolation criteria of the tight IDs. For both electrons and muons, the relative isolation cuts in the *Loose* ID is set to be < 0.4.

The ratio between the number of misidentified leptons passing a *Nominal* ID (or to be called as tight ID) to a *Loose* ID is called the fake rate. The fake rate is measured by using separate control data samples collected with a set of appropriately prescaled single lepton triggers. The region where the fake rate is measured is called the measurement region. Taking the same-sign dilepton final state channels as an example, the fake rate is then used to predict the number of misidentified lepton events in two tight lepton events (i.e. the signal regions) by extrapolating from a data sample of lepton events with one tight lepton plus one loose-but-not-tight lepton. The latter region where the fake rate is applied to estimate the number of misidentified lepton events in the signal region is called the application region. For the 3ℓ channels, the application region events contain two tight leptons and one loose-but-not-tight lepton.

Several effects have to be considered when doing this estimation: when measuring the fake rate, we need to make sure that the kinematics of the mother parton is similar in the measurement region and the application region. Residual effects (e.g. due to different flavor composition for the mother parton) are assessed in a closure test.

All other backgrounds, including events with mismeasured lepton charge or prompt photons misidentified as electrons, and irreducible backgrounds such as $W^\pm W^\pm$ and $t\bar{t}W^\pm$ events, are directly determined from simulation, and validated and corrected by events in the corresponding validation regions.

The $t\bar{t}W^\pm$ background estimation is validated from a control sample with at least one b-tagged jet. We find that the best selection for a reasonably pure sample is to select events with two SS leptons, at least four jets, and at least one b-tagged jet. As events with ≥ 2 b-tagged jets have a stronger ability to constrain this background, we use the distribution of the b-tagged jet multiplicity, N_b, to normalize this background. To achieve the best accuracy for the $t\bar{t}W^\pm$ estimation, we decided to fit the distribution of $t\bar{t}$+X and other backgrounds to the data distribution of N_b. Note that since there

is a non-negligible contribution of $t\bar{t}H \, (\to WW)$, we fit the $t\bar{t}W^{\pm} + t\bar{t}H$ processes together. In the fits, we test if the fit results are stable among the three years, and the test reveals nothing suspicious. Figure 7.22 shows the N_b distribution for both data and MC in the validation for the $t\bar{t}W^{\pm}$ background. A fit to this distribution (floating $t\bar{t}$+X vs. other backgrounds) reveals that the $t\bar{t}W^{\pm}$ background estimated from simulation needs to be scaled up by a factor of 1.67 ± 0.18.
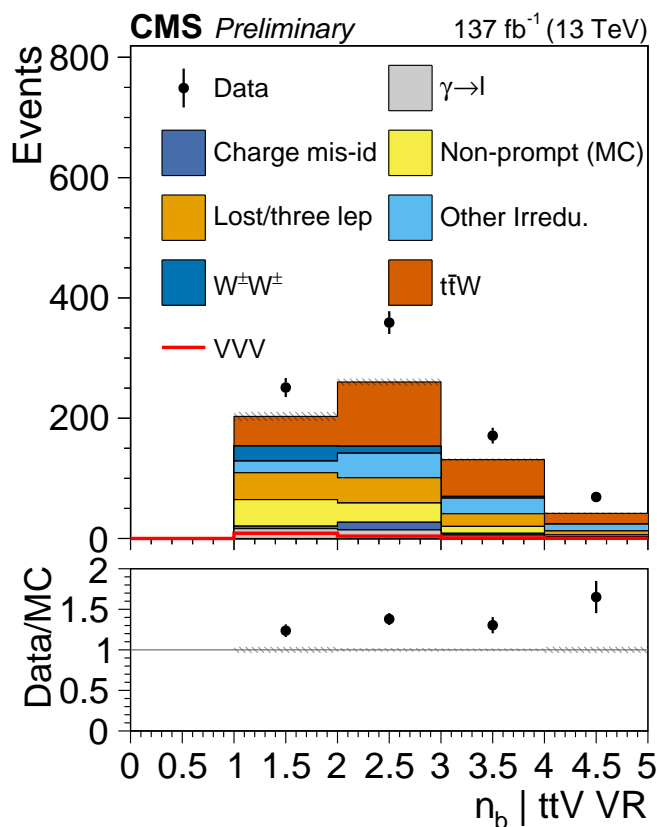
Figure 7.22: Distribution of N_b for both data and MC in the $t\bar{t}W^{\pm}$ validation region.

The fake photon background has multiple sources, however a common feature is a low lepton momentum and/or close proximity to a real lepton. This is why the selection criteria such as the constraints on p_T^{miss}, m_T, or $m_{\ell\ell}$ reduce this background dramatically. For the 3ℓ signal regions, one frequent background source is $Z \to \ell\ell$ with a radiated photon. Therefore, a large fraction of the background is rejected by $|m_{3\ell} - m_Z| > 10 \, \text{GeV}$ for events with 3ℓ. To validate the background due to prompt photons misidentified as electrons for the cut-based method, we define the validation region as follows: select 3ℓ events, apply the signal region selection, but invert the selection criteria on $p_T^{3\ell}$ and $m_T^{\text{max}}/m_T^{\text{3rd}}$ and drop the $m_{3\ell}$ criteria, and for

the BDT-based method, invert the fake and prompt BDT cuts (requiring the events to have prompt BDT < 0.28 and fake BDT < 0.78, motivated by the prompt and fake BDT distributions). After such selections, we find that the photon fakes are very pure in the region near $m_{3\ell} \approx m_Z$, as shown in Figure 7.23 for the cut-based method and Figure 7.24 for the BDT-based method.

A fit to the $m_{3\ell}$ distribution with photon fake backgrounds against all other backgrounds floating is then performed, and the fit reveals that for the cut-based method, the photon fake backgrounds need to be scaled from MC yields by a factor of 0.991 ± 0.072 for 2016, 0.508 ± 0.025 for 2017 and 0.551 ± 0.022 for 2018; and for the BDT-based method, the photon fake backgrounds need to be scaled from MC yields by a factor of 0.550 ± 0.082, 0.259 ± 0.030, and 0.281 ± 0.026 for 2016, 2017, and 2018 respectively.

The validation region to validate the $W^{\pm}W^{\pm}$ background modelling in MC is constructed by inverting either the m_{JJ} or the $\Delta\eta_{JJ}$ criteria (or both). In Figure 7.25, we show the m_{JJ} variable for $\Delta\eta_{JJ} > 2.5$ (left) and $\Delta\eta_{JJ}$ for $m_{JJ} > 500$ GeV. The model agrees very well with the data in this region, therefore we do not apply any correction to the WW+jj background. The statistical precision of these two validation regions is about 10%, and we add this as a systematic uncertainty for this process.

To validate the charge mismeasurement background modeling in MC, we look at the dielectron invariant mass near the Z mass peak for events having two SS electrons (and no other selection criteria). As can be appreciated from Figure 7.26, the Z mass peak from charge mismeasurements is well reproduced in simulation, within a precision of about 20%. We therefore take the backgrounds due to charge mismeasurements from simulation without any correction with a 20% systematic uncertainty.

7.5.2 Backgrounds in four-lepton final states

There are six main background components to the four-lepton final state. They are ZZ, $t\bar{t}Z$, tWZ, WZ, Higgs, and other rare SM processes.

The largest background contribution to the four-lepton final state comes from the ZZ background process, which has a cross section that is three orders of magnitude larger than that of the signal process. The contribution from this background process in the signal region is estimated from the control region (CR) events that identify two Z boson candidates.

The second largest background contribution comes from the $t\bar{t}Z$ background pro-
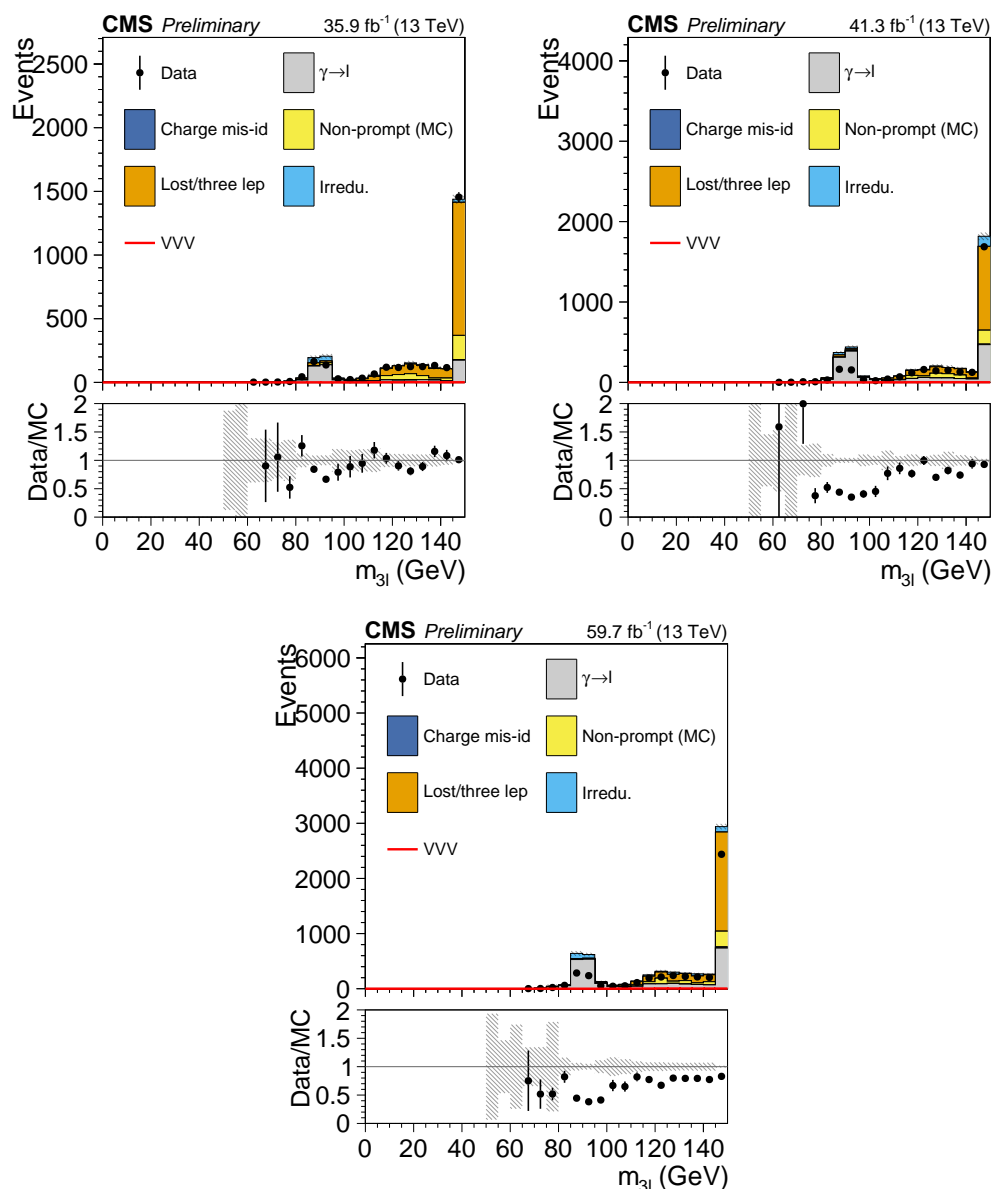
Figure 7.23: Distribution of $m_{3\ell}$ after inverting both the $m_T^{\max}/m_T^{3rd}$ and $p_T^{3\ell}$ criteria for 2016 (top-left), 2017 (top-right), and 2018 (bottom) for the cut-based method.

cess, which has a cross section that is approximately 50 times larger than that of the signal process. However, the $t\bar{t}Z$ background process events are reduced by requiring that there are no b-tagged jets in each event. The contribution from the $t\bar{t}Z$ background process in the signal region is estimated from the CR events that require the additional presence of b-tagged jets.

The main usage of the CRs is to determine the normalization scale factors to be used
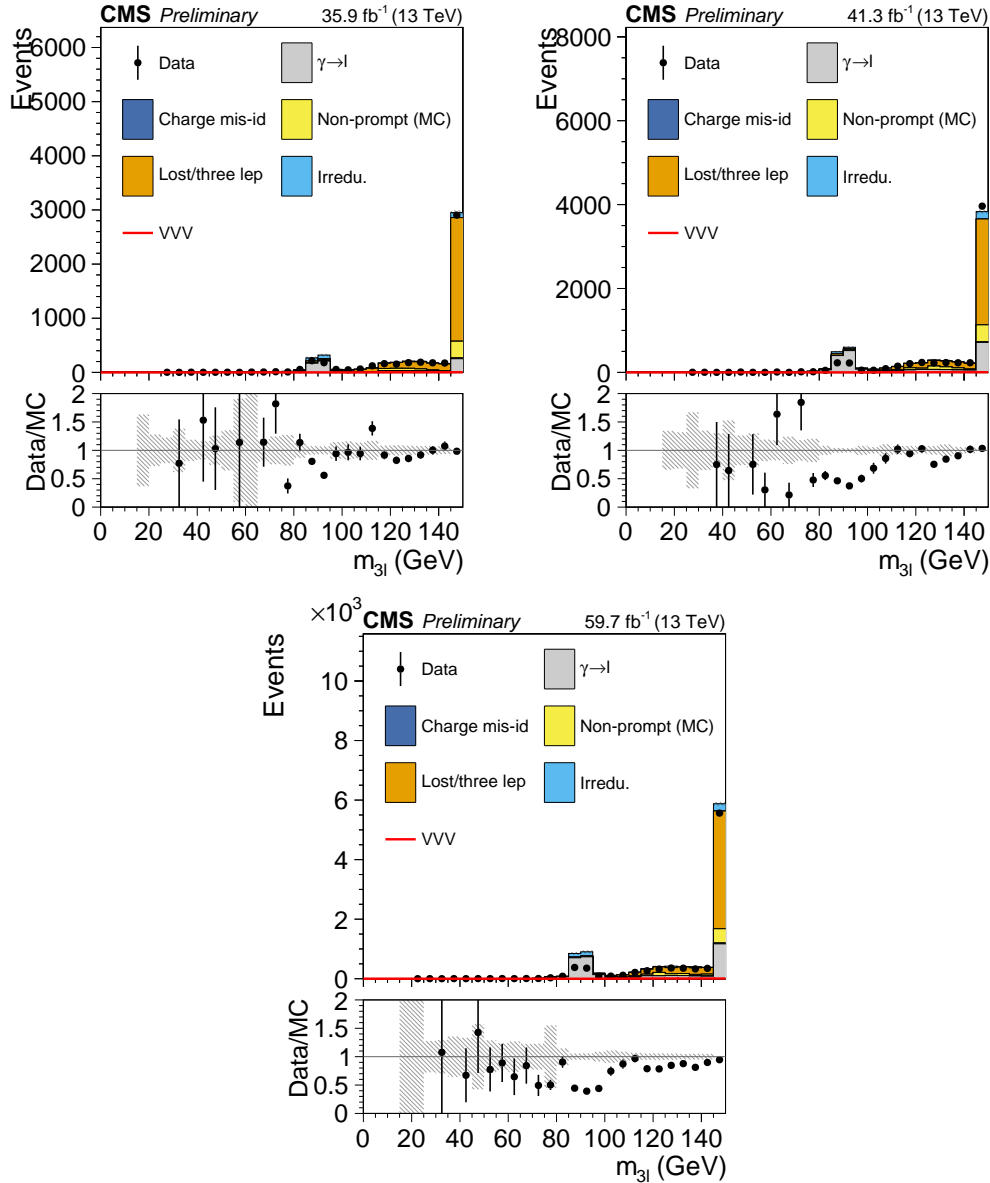
Figure 7.24: Distribution of $m_{3\ell}$ after inverting both the m_T^{max}/m_T^{3rd} and $p_T^{3\ell}$ criteria for 2016 (top-left), 2017 (top-right), and 2018 (bottom) for the BDT-based method.

in estimating each background's contribution to one of the signal regions . For each background process and its corresponding control region CR_i, the normalization scale factor NSF_{B_i,CR_i} is defined as the ratio of the observed to the expected yields in the control region CR_i, for the background process B_i being estimated. The observed yields in each CR_i, N_{B_i,CR_i}^{data} for each of the background processes under consideration, are corrected by subtracting the events in this CR_i due to all the background processes $B_j, j \neq i$, other than the one under consideration, where some
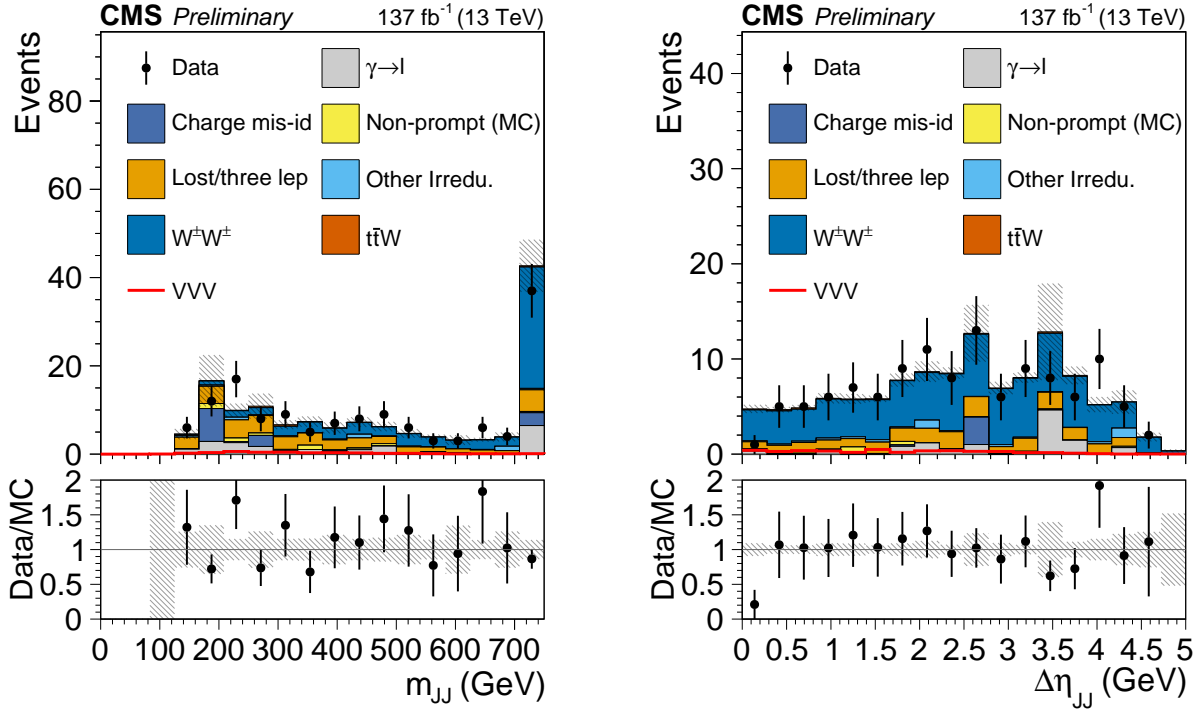
Figure 7.25: m_{JJ} for $\Delta\eta_{JJ} > 2.5$ (left) and $\Delta\eta_{JJ}$ for $m_{JJ} > 500\,\text{GeV}$ (right) as validation of the $W^\pm W^\pm$ VBS background.

of the events selected for the other control regions $CR_j, j \neq i$ can, according to the simulation, end up in the current control region CR_i. Once these subtractions are done, the estimate of the contribution of the background B_i to the signal region can then be written as

$$N_{\text{SR},B_i}^{\text{pred}} = N_{\text{SR},B_i}^{\text{MC}} \cdot \underbrace{N_{B_i,CR_i}^{\text{data}}/N_{B_i,CR_i}^{\text{MC}}}_{\text{Normalization scale factor}(\text{NSF}_{B_i})} = N_{B_i,CR_i}^{\text{data}} \cdot \underbrace{N_{\text{SR},B_i}^{\text{MC}}/N_{B_i,CR_i}^{\text{MC}}}_{\text{Transfer factor}(\text{TF}_{B_i,CR_i \to \text{SR},B_i})}$$

$$(7.4)$$

where $N_{B_i,CR_i}^{\text{data}}$ and N_{B_i,CR_i}^{MC} are the observed and simulation-estimated yields in the control region CR_i due to background process B_i and $N_{\text{SR},B_i}^{\text{MC}}$ is the yield in the signal region due background process B_i estimated from the simulation. Once the normalization scale factor for background process B_i (NSF_{B_i}) is determined, the simulated events of the background process being estimated are scaled according to the value of NSF_{B_i}.

Effectively, what is being done is to extrapolate from the observed number of
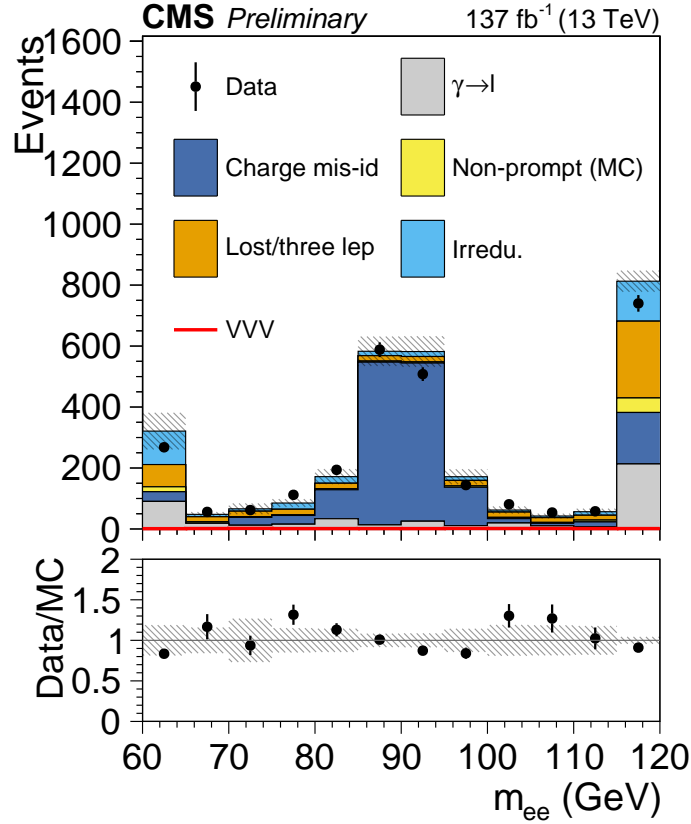
Figure 7.26: Distribution of the dielectron invariant mass m_{ee} for events with an SS dielectron pair.

events in the data CR, $N^{\text{data}}_{\text{B}_i,\text{CR}_i}$, by the transfer factor ($\text{TF}_{\text{B}_i,\text{CR}_i \to \text{SR},\text{B}_i}$) obtained from the simulation. The various sources of systematic uncertainty associated with the $\text{TF}_{\text{B}_i,\text{CR}_i \to \text{SR},\text{B}_i}$ often cancel out if the effect of the systematic uncertainty is correlated between the signal and the control region. This can lead to a smaller total uncertainty on important background processes in the signal regions.

In practice, what is being done for the analysis is obtaining the central background estimate yields by taking the $N^{\text{MC}}_{\text{SR},\text{B}_i}$ and multiplying it by NSF_{B_i}. For the uncertainty on those estimated yields, the statistical uncertainties from the $N^{\text{data}}_{\text{B}_i,\text{CR}_i}$'s are taken in addition to the uncertainties from the TF's.

The ZZ control region selections are nearly identical to the ones used for the signal region in the $ee/\mu\mu$ category, except they do not require the $p_{\text{T}}^{\text{miss}}$ cut and have an inverted requirement on the $m_{\ell\ell}$ (i.e. $|m_{\ell\ell} - m_{\text{Z}}| < 10$ GeV). This region will be referred to as the ZZ control region (ZZ CR). The purity of events from the

ZZ background process in the ZZ CR is ≈100%. The normalization scale factor NSF_{ZZ} obtained from the entire Run 2 dataset is 0.99 ± 0.02, which shows a good consistency between the simulation and the data. There are slight variations from year to year, but the NSF_{ZZ}'s in each year's dataset are consistent with unity. The NSF_{ZZ} factors are applied separately for each year's simulated dataset.

To ensure proper background estimation, it is important to check that the extrapolation transfer factors $\text{TF}_{ZZ\ CR \to e\mu\ SR}$ and $\text{TF}_{ZZ\ CR \to ee/\mu\mu\ SR}$ obtained from the simulation are modeled well. There are several variables to be checked for any potential mismodeling problem. From the ZZ CR to the signal regions in the eµ category, there are several key quantities that are being extrapolated, such as the lepton flavors, $m_{\ell\ell}$, and m_{T2} distributions. Since the lepton efficiencies are generally well modeled, within a few percent per lepton, and they are corrected using the scale factors, the mismodeling in the extrapolation across lepton flavors is expected to be negligible. The modeling of a few key kinematic distributions has been checked, as illustrated in Figure 7.27.

For the BDT-based method, a similar approach is taken to estimate the ZZ backgrounds. In the BDT eµ channel, the ZZ CR is defined in the same way as the ZZ CR used in the cut-based method, and a corresponding BDT distribution in the ZZ CR is used to assess the individual transfer factor per BDT bin, where each bin in the control region acts as a control region for each bin in the signal region. In the BDT ee/µµ channel, the ZZ control region is defined from the ee/µµ events with a BDT score of less than 0 (which were discarded and not used in the ee/µµ signal region).

To estimate the $t\bar{t}Z$ background contribution in the signal regions, events with more than one b-tagged jet and the W lepton candidates with eµ flavor are selected. This region will be referred to as the b-tag CR. In this region, the purity of the $t\bar{t}Z$ background process is ≈80%. The expected yields of the $t\bar{t}Z$ background process from simulation generally agree well with the observed data. The normalization scale factor for the $t\bar{t}Z$ background process $\text{NSF}_{t\bar{t}Z}$ has been found to be 1.08 ± 0.17 for the Run 2 dataset. For the final estimate, the $\text{NSF}_{t\bar{t}Z}$ factors computed for each year are applied to their respective data sets. The modeling of key discriminating variables is tested in the b-tag CR. Figure 7.28 shows the m_{T2}, $m_{\ell\ell}$, p_T^{miss}, and $p_T^{4\ell}$ distributions.

Finally, there are the WZ, tWZ, and Higgs background processes. These background processes are subdominant in the signal region (they contribute less than
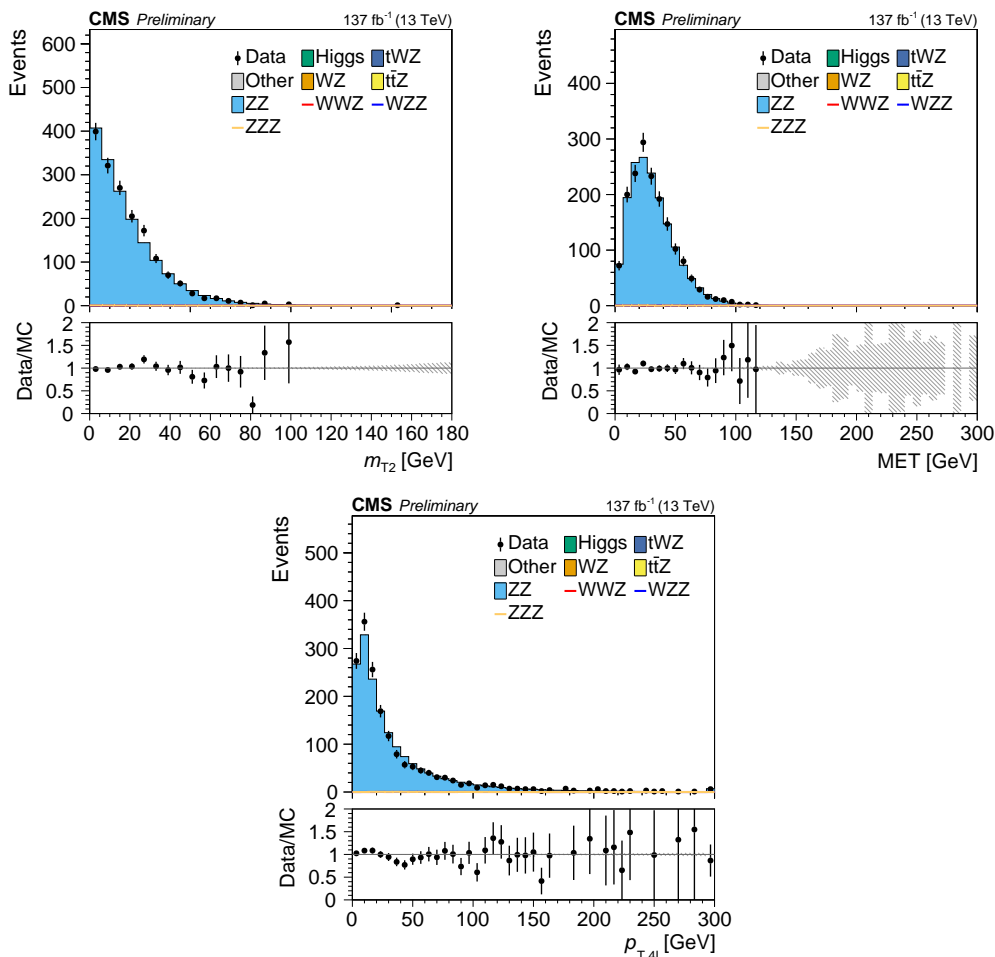
Figure 7.27: Distributions of m_{T2} (top-left), p_T^{miss} (top-right), and $p_T^{4\ell}$ (bottom) for the data and MC in the four-lepton ZZ CR.

10% to the total background). There are also other rare SM background processes that have minor contributions. These background processes are extrapolated directly from the simulation, using appropriate validation regions to check the modeling. For the WZ backgrounds, this process enters the four lepton signal region with a misidentified lepton. To check such modelling in the MC, we checked the lepton isolation variable distribution of events with three *Nominal* ID leptons and one lepton that passes the *Veto* ID except for the lepton isolation criteria. No clear deviation between data and MC is observed, so the distributions generally exhibit good modeling. We take the statistical uncertainty of those events (about 28%) as the uncertainty in the WZ background estimation.

To validate the tWZ backgrounds, we define a validation region similar to the $t\bar{t}Z$
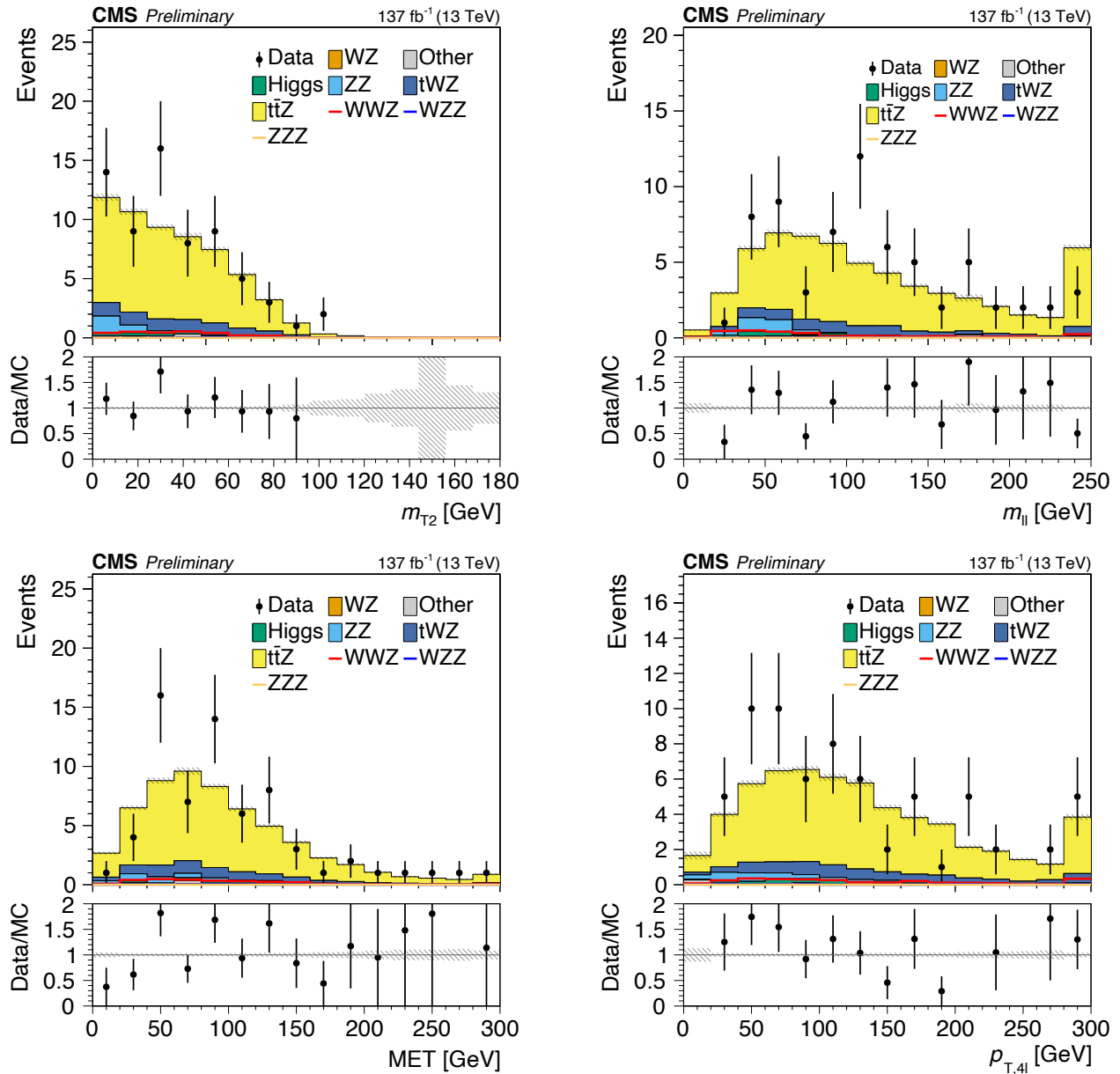
Figure 7.28: Distributions of m_{T2} (top left), $m_{\ell\ell}$ (top right), p_T^{miss} (bottom left), $p_T^{4\ell}$ (bottom right) for the data and MC in the four-lepton b-tag CR.

CR, except that only one b-tagged jet is required in this case. By doing so, the tWZ and $t\bar{t}Z$ events in this region are similar to each other, but the overall tWZ contribution among the backgrounds is small. When we perform a fit to the number of jets (N_j) distribution in the tWZ CR, the fit reveals that the tWZ component needs to be scaled up by a factor of 1.47, with a very large uncertainty on this number (1.47 ± 2.13) due to the smallness of the tWZ contribution to this control

region. As a result, we decide to take the MC yield of tWZ directly as the tWZ background estimation, and assign a 47% uncertainty to this estimate.

7.5.3 Backgrounds in five- and six-lepton final states

For the five lepton final states, the largest source of background is the ZZ process where the additional lepton is often a misidentified jet or is indicative of the presence of a B-hadron that decayed semileptonically through a cascade to a soft lepton. To ascertain the origin of this lepton, the generator objects that were matched to this lepton within a cone of $\Delta R < 0.3$ were studied. The lepton was found to be associated with either:

- a heavy flavor B or D hadron decay, which was the dominant source of the fake lepton, or

- π^0s decaying to photons with one or more photons converting to lepton pairs; or Dalitz decay of $\pi^0 \rightarrow e^+ e^- \gamma$. This effect is seen to be subdominant.

To further understand this background, we constructed a control region to study the agreement of the data with simulations where one lepton is known to be fake. In this context, a fakeable lepton is one that fails the isolation requirements placed on the leptons in the signal region. Due to the lack of events in a five-lepton selection where one of the leptons is fake (achieved by requiring a non-isolated additional lepton), the analysis was performed using a four-lepton selection. The central idea behind the construction of this control region is that the exploration of the WZ process (where the W and Z bosons decay leptonically) in a four lepton selection is topologically similar to the ZZ process (where both Z bosons decay leptonically) in a five-lepton selection. In order to minimize contamination from Drell-Yan (DY) and ZZ events, we use only the three muon and one electron selection to isolate the WZ signal. The other lepton flavor combinations were explored. Since these event selections do not produce pure WZ events, they were dropped from consideration. The additional selection requirements applied are the following:

- a b-tagged jet veto

- requiring a Z-boson candidate, where the Z-boson candidate is constructed by combining any two pairs of muons

- the p_T of the electron >15 GeV, to reduce the DY background

The distributions of p_T^{miss} are shown in Figure 7.29. The data and the total simulated background distributions agree within the statistical uncertainties. As a result, we use the statistical uncertainty in this control region (about 30%) for the ZZ background uncertainty in the five-lepton final states.
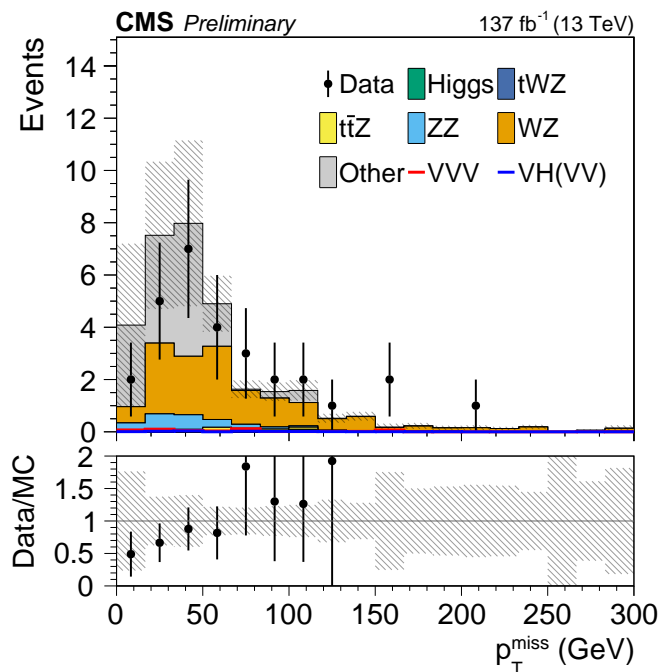
Figure 7.29: The distribution of p_T^{miss} in the WZ control region, where one of the leptons is required to be non-isolated.

In the six-lepton channel, the background processes are determined entirely from simulations. The total background contribution in this channel is less than one event, and this does not require additional control regions to be constructed in order to study these minuscule contributions.

7.6 Systematic uncertainties

The dominant systematic uncertainties come from the uncertainties related to the background estimates, discussed in Section 7.5. The remaining uncertainties are related to instrumental effects and theoretical uncertainties, such as the electron/muon ID efficiency scale factor ("eleSF"/"muoSF" in the tables below), the trigger efficiency scale factor (triggerSF), the jet energy scale (JES), the luminosity uncertainty, the b tagging scale factor, the statistics of simulated events (Sim. statistics), the normalization and factorization scale Q^2, the PDF uncertainties, and the α_s variations.

Tables 7.16 and 7.17 summarize the typical sizes of the systematics on each back-

ground process in the same-sign and three-lepton final states from various sources, for the cut-based and BDT-based approach. In general, the contribution from irreducible, charge flips, and $\gamma \rightarrow$ nonprompt ℓ backgrounds are very small, so the relatively large uncertainties in those backgrounds do not matter in the final result.

The typical total uncertainty on the lost/3ℓ (nonprompt ℓ) backgrounds are about 20% (30%) in the most sensitive categories, but they are still small enough such that the result is still statistics dominated. Tables 7.18 and 7.19 summarize the typical sizes of the systematics on each background process in the four-lepton final states from various sources for the cut-based approach. Tables 7.20 and 7.21 summarize the typical sizes of the systematics on each background process in the four-lepton final states from various sources for the BDT approach. For the dominant background processes, ZZ ($t\bar{t}Z$) backgrounds, the total systematic uncertainty is about 5-10% (10-20%). Tables 7.22 and 7.23 summarize the systematics on the background process for the five and six lepton final states. The dominant systematics in five- and six-lepton final states are the systematics on the ZZ backgrounds where the additional lepton is a misidentified lepton, for which the uncertainty is estimated to be 30% in the five-lepton category and 60% in the six-lepton category.

In most final states, the dominant uncertainty on the signal prediction comes from the statistics in MC due to the limited number of simulated events that pass the event selection. This uncertainty is about 5-30% for the same-sign/3ℓ, 3-5% for the four-lepton, and 0.4-11% for the five- and six-lepton final states.

For the signal predictions, we apply normalization and factorization scale, PDF, and α_s variations to obtain the uncertainties on the signal acceptance. Various experimental uncertainties (similar to the sources for background uncertainties) are also evaluated. Table 7.24 summarizes the systematics on the signal process for the cut-based analysis, and Table 7.25 summarizes the systematics on the signal process for the BDT analysis. Table 7.26 summarizes the systematics on the signal process for the cut-based approach. Table 7.27 summarizes the systematics on the signal process for the BDT approach. Table 7.28 summarizes the systematics on the signal process for the five- and six-lepton final states.

All the sources of systematic uncertainties mentioned above are treated carefully for correlation among data taken from different years and among different final states. For example, all systematics are fully correlated among different years except luminosity uncertainty, as most of those systematics are related to measurements that are done the same way in different years. All systematics are considered fully

correlated among different final states, except the simulation statistics uncertainty and background estimation method related uncertainties.

Table 7.16: Summary of the typical background systematic uncertainties (in percent) for the SS and 3ℓ final states, for the cut-based analysis.

Uncertainty	Lost/three ℓ	Nonprompt ℓ	Irreducible	Charge flips	$\gamma \rightarrow$ nonprompt ℓ
Sim. statistics	3–15	—	6–26	22–90	30–100
Bkg. contam.	0–2	3–20	—	—	—
eleSF	0–2.4	—	—	—	—
muoSF	0–1.2	—	—	—	—
triggerSF	3	—	—	—	—
pileup	0–1	—	—	—	—
Q^2	0–2	—	—	—	—
PDF	0–1	—	—	—	—
JEC	0–3	—	—	—	—
WZ/ttZ fraction	<1	—	—	—	—
data validation	1–10	—	10–14	20	8–15
Fake rate e	—	21–44 (SS), 2–3 (3ℓ)	—	—	—
Fake rate μ	—	3–15	—	—	—
Closure test e	—	19–29 (SS), 2–4 (3ℓ)	—	—	—
Closure test μ	—	7–30	—	—	—

Table 7.17: Summary of the typical background systematic uncertainties (in percent) for the SS and 3ℓ final states, for the BDT based analysis.

Uncertainty	Lost/three ℓ	Nonprompt ℓ	Irreducible	Charge flips	$\gamma \rightarrow$ nonprompt ℓ
Sim. statistics	5–25	—	5–15	27–100	35–340
Bkg. contam.	1–35	10–70	—	—	—
pileup	0–2	—	—	—	—
Q^2	1–13	—	—	—	—
PDF	0–2	—	—	—	—
JES	0–60	—	—	—	—
WZ/ttZ fraction	<1	—	—	—	—
data validation	1–10	—	10–21	20	7–19
Fake rate e	—	16–65 (SS), 2–10 (3ℓ)	—	—	—
Fake rate μ	—	1–20	—	—	—
Closure test e	—	20–30 (SS), 2–7 (3ℓ)	—	—	—
Closure test μ	—	3–30	—	—	—

Table 7.18: Summary of the typical systematic uncertainties (in percent) on the background estimates in the eμ SR (denoted by − if not applicable or if the size is smaller than 0.1%).

Systematics	$t\bar{t}Z$	ZZ	tWZ	Higgs	WZ	Other
b-tag CR data statistics	12	-	-	-	-	-
ZZ CR data statistics	-	2.4	-	-	-	-
b tagging scale factor (HF)	-	-	4.9	3.4	-	0.5
b tagging scale factor (LF)	-	-	0.9	0.3	-	0.2
JES	-	-	2.3	12	-	0.1
Pileup reweighting	-	-	1.4	4.3	-	0.5
α_s	-	-	0.0	0.1	-	2.0
PDF	-	-	0.4	0.6	-	1.1
Q^2	-	-	-	15	-	2.9
Luminosity	2.5	2.5	2.5	2.5	2.5	2.5
Systematics on $\text{TF}_{\text{b-tag CR}\rightarrow\text{e}\mu}$	10	-	-	-	-	-
Systematics on $\text{TF}_{\text{ZZ CR}\rightarrow\text{e}\mu}$	-	4.5	-	-	-	-
Systematics on WZ validation	-	-	-	-	30	-
triggerSF	3.2	3.2	3.2	3.2	3.2	3.2
eleSF	-	-	3.3	0.1	-	2.1
muoSF	-	-	3.2	4.3	-	3.8
Sim. statistics	-	-	6.2	82	28	43

Table 7.19: Summary of the typical systematic uncertainties (in percent) on the background estimates in the $ee/\mu\mu$ SR (denoted by $-$ if not applicable or if the size is smaller than 0.1%).

Systematics	$t\bar{t}Z$	ZZ	tWZ	Higgs	WZ	Other
b-tag CR data statistics	12	-	-	-	-	-
ZZ CR data statistics	-	2.4	-	-	-	-
b tagging scale factor (HF)	-	-	4.1	1.5	-	0.0
b tagging scale factor (LF)	-	-	0.8	0.9	-	1.8
JES	-	-	3.9	6.9	-	1.9
Pileup reweighting	-	-	1.3	2.6	-	41
α_s	-	-	0.1	0.5	-	0.3
PDF	-	-	-	0.9	-	0.1
Q^2	-	-	0.2	2.7	-	8.8
Luminosity	2.5	2.5	2.5	2.5	2.5	2.5
Systematics on $TF_{\text{b-tag CR}\rightarrow e\mu}$	10	-	-	-	-	-
Systematics on $TF_{\text{ZZ CR}\rightarrow e\mu}$	-	8.9	-	-	-	-
Systematics on WZ validation	-	-	-	-	30	-
triggerSF	3.2	3.2	3.2	3.2	3.2	3.2
eleSF	-	-	3.2	4.3	-	3.5
muoSF	-	-	3.3	3.0	-	0.7
Sim. statistics	-	-	7.6	26	50	81

Table 7.20: Summary of the typical systematic uncertainties (in percent) on the background estimates in the eµ SR in the BDT approach (denoted with − if not applicable or if the size is smaller than 0.1%).

Systematics	$t\bar{t}Z$	ZZ	tWZ	Higgs	WZ	Other
b-tag CR data statistics	12	-	-	-	-	-
ZZ CR data statistics	-	2.4	-	-	-	-
b tagging scale factor (HF)	-	-	4.9	1.5	-	0.5
b tagging scale factor (LF)	-	-	0.8	0.7	-	0.2
JES	-	-	2.5	0.7	-	0.3
Pileup reweighting	-	-	1.6	4.5	-	1.1
α_s	-	-	-	0.2	-	1.7
PDF	-	-	0.2	0.9	-	0.9
Q^2	-	-	-	15	-	1.9
Luminosity	2.5	2.5	2.5	2.5	2.5	2.5
Systematics on $\text{TF}_{\text{b-tag CR}\to e\mu}$	29	-	-	-	-	-
Systematics on $\text{TF}_{\text{ZZ CR}\to e\mu}$	-	2.2	-	-	-	-
Systematics on WZ validation	-	-	-	-	30	-
triggerSF	3.2	3.2	3.2	3.2	3.2	3.2
eleSF	-	-	3.2	2.9	-	2.8
muoSF	-	-	3.2	3.9	-	3.3
Sim. statistics	-	-	5.7	30	22	36

Table 7.21: Summary of the typical systematic uncertainties (in percent) on the background estimates in the ee/μμ SR for the BDT approach (denoted by − if not applicable or if the size is smaller than 0.1%).

Systematics	$t\bar{t}Z$	ZZ	tWZ	Higgs	WZ	Other
b-tag CR data statistics	12	-	-	-	-	-
ZZ CR data statistics	-	2.4	-	-	-	-
b tagging scale factor (HF)	-	-	4.2	0.9	-	0.4
b tagging scale factor (LF)	-	-	0.8	1.1	-	1.8
JES	-	-	2.4	1.3	-	0.6
Pileup reweighting	-	-	1.5	3.6	-	41
α_s	-	-	-	0.5	-	0.4
PDF	-	-	0.3	0.3	-	0.1
Q^2	-	-	0.0	0.1	-	8.3
Luminosity	2.5	2.5	2.5	2.5	2.5	2.5
Systematics on $\mathrm{TF}_{\text{b-tag CR}\to e\mu}$	11	-	-	-	-	-
Systematics on $\mathrm{TF}_{\text{ZZ CR}\to e\mu}$	-	3.0	-	-	-	-
Systematics on WZ validation	-	-	-	-	30	-
triggerSF	3.2	3.2	3.2	3.2	3.2	3.2
eleSF	-	-	3.0	4.6	-	3.7
muoSF	-	-	3.4	3.3	-	0.8
Sim. statistics	-	-	6.4	16	19	73

Table 7.22: Summary of the typical systematic uncertainties (in percent) on the background estimates in the five-lepton signal region.

Systematics	$t\bar{t}Z$	ZZ	tWZ	Other
b tagging scale factor (HF)	4.7	0.00	0.00	0.00
b tagging scale factor (LF)	0.01	0.00	0.00	0.00
WZ CR uncertainty	-	30	-	-
Luminosity	2.5	2.5	2.5	2.5
JES	0.01	0.02	0.00	26.0
Pileup reweighting	0.01	0.1	0.00	0.02
PDF	1.0	1.0	0.05	0.06
Q^2	0.02	0.01	0.01	0.01
α_s	0.02	0.01	0.01	0.01
eleSF and muoSF	3.6	4.1	5.3	7.4
triggerSF	2.0	2.0	2.0	2.0
Sim. statistics	22	5.0	34	60

Table 7.23: Summary of the typical systematic uncertainties (in percent) on the background estimates in the six-lepton signal region.

Systematics	$t\bar{t}Z$	ZZ	$t\bar{t}H$
b tagging scale factor (HF)	3.13	0.02	0.06
b tagging scale factor (LF)	4.16	0.05	0.00
WZ CR uncertainty	-	60	-
Luminosity	2.5	2.5	2.5
Pileup reweighting	17	0.01	6.2
PDF	7.6	0.05	5.3
Q^2	0.02	0.01	15
α_s	60	0.01	0.01
eleSF and muoSF	18	7.5	6.4
triggerSF	2.0	2.0	2.0
Sim. statistics	100	28	26

Table 7.24: Summary of the typical signal systematic uncertainties (in percent) for the SS and 3ℓ final states, for the cut-based analysis.

Uncertainty	Typical range
Sim. statistics	4–19
eleSF	0–5
muoSF	0–2
triggerSF	1–3
pileup	0–2
Q^2	2–14
PDF	2–7
α_s	1
JEC	1–5
b-tagging HF	0–2
b-tagging LF	1–2
Luminosity	2.5

Table 7.25: Summary of the typical WWW signal systematic uncertainties (in percent) for the SS and 3ℓ final states, for the BDT based analysis.

Uncertainty	Typical range
Sim. statistics	4–32
eleSF	2–6
muoSF	0-1
triggerSF	3
pileup	0–5
Q^2	4–8
PDF	1–6
α_s	0-3
JES	0–20
b-tagging HF	0–1
b-tagging LF	0–1
Luminosity	2.5

Table 7.26: Summary of the typical systematic uncertainties (in percent) on the number of signal events in the four-lepton signal regions (denoted by − if not applicable or if the size is smaller than 0.1%).

Systematics	eμ SR	ee/μμ SR
b tagging scale factor (HF)	0.2	0.2
b tagging scale factor (LF)	0.8	0.9
JES	0.8	0.4
Pileup reweighting	0.1	1.9
α_s	0.1	0.1
PDF	0.1	0.8
Q^2	0.2	0.9
MET smearing	0.1	0.5
Luminosity	2.5	2.5
triggerSF	3.2	3.2
eleSF	3.3	3.0
muoSF	3.2	3.5
Sim. statistics	2.8	4.1

Table 7.27: Summary of the typical systematic uncertainties (in percent) on the signal events in the four-lepton signal regions in the BDT approach (denoted by − if not applicable or if the size is smaller than 0.1%).

Systematics	eμ SR	ee/μμ SR
b tagging scale factor (HF)	0.2	0.1
b tagging scale factor (LF)	0.8	0.8
JES	0.6	0.8
Pileup reweighting	-	2.7
α_s	0.2	0.3
PDF	0.1	0.2
Q^2	0.6	0.0
MET smearing	-	-
Luminosity	2.5	2.5
triggerSF	3.2	3.2
eleSF	3.3	3.0
muoSF	3.2	3.3
Sim. statistics	3.7	5.0

Table 7.28: Summary of the typical systematic uncertainties (in percent) on signal events in the five and six lepton signal regions (denoted by - if not applicable or if the size is smaller than 0.1%).

Systematics	five lepton SR	six lepton SR
b tagging scale factor (HF)	-	-
b tagging scale factor (LF)	-	-
JES	0.05	-
Pileup reweighting	1.8	0.6
α_s	0.2	0.2
PDF	0.1	0.2
Q^2	0.2	0.6
Lepton SF	3-4	3.6
Luminosity	2.5	2.5
triggerSF	2.0	2.0
Sim. statistics	11	0.4

7.7 Results

The observed data, together with the signal and background predictions (pre-fit), are shown in Figure 7.30, for both the cut-based and BDT-based methods. The corresponding tabulated results are found in Tables 7.29, 7.31, 7.30, and 7.32. In general, the observed number of events (N_{obs}) agrees well with the sum of the signal and background predictions (N_{pred}) in all channels. The pull, defined as $\frac{N_{obs} - N_{pred}}{\sqrt{\sigma_{obs}^2 + \sigma_{pred}^2}}$, is less than 1σ in most channels.

Several different fitting strategies are employed to interpret the observed data relative to the expected signal and backgrounds. First of all, we have eight different VVV signals in total: WWW-onshell, WH(WW), WWZ-onshell, ZH(WW), WZZ-onshell, WH(ZZ), ZZZ-onshell, and ZH(ZZ), and each of them may have a different signal strength (μ, defined as the ratio of the measured production cross section times branching fraction to the corresponding SM prediction). Different treatments of these eight signal strengths are applied during the fit, depending on the interpretation goals, in several scenarios:

1. A common signal strength is used for all eight signals, μ_{VVV}. This scenario is used mainly for the establishment of the first observation of VVV production regardless of how the three bosons are produced and what combinations of W and Z they are.

2. A common signal strength is used for all VVV-onshell signals, $\mu_{VVV-onshell}$, and all the VH(VV) processes are treated as backgrounds (i.e. their signal strengths are fixed to 1.0 and they are only allowed to fluctuate within the systematic uncertainties, by $\pm 1\sigma$). This is to measure the VVV-onshell cross section, with the assumption of the VH(VV) process being produced as predicted by the SM.

3. In total, four different signal strengths are used: μ_{WWW} (for WWW-onshell and WH(WW) signals), μ_{WWZ} (for WWZ-onshell and ZH(WW) signals), μ_{WZZ} (for WZZ-onshell and WH(ZZ) signals), μ_{ZZZ} (for ZZZ-onshell and ZH(ZZ) signals). In this case, we are measuring the production cross sections for each VVV process at the same time, where we do not distinguish the VVV-onshell and VH(VV) processes.

4. Same as 3, but treating all the VH(VV) processes as backgrounds. In this case, we are measuring the production cross section for each of the VVV-onshell

processes at the same time.

The most straightforward fitting scenario with all eight signal strengths floating simultaneously is not among the above scenarios, and this is because the result from current selection does not have enough distinguishing power between VVV-onshell and VH(VV) signals. However, in the later part of this section, we will show a result (not optimal) from a fit with all the eight signal strengths floating and present the 2D likelihood of the VVV-onshell and VH(VV) signals.

Figure 7.31 and 7.32 show the postfit signal and background yields in each signal region for both the cut-based and the BDT-based methods, using the 3rd fitting strategy above. The corresponding tabulated results can be found in Tables 7.33, 7.35, 7.34, and 7.36. The result of the fits for the four signal strengths μ_{WWW}, μ_{WWZ}, μ_{WZZ} and μ_{ZZZ} are shown in Figure 7.31. As shown, the best fit values for μ_{WWW} are μ_{WWZ} are close to 1.0, within the uncertainties. While the central value of μ_{WZZ} is larger than 1 and the value μ_{ZZZ} is at 0.0, due to an excess (deficit) of observed events in the five (six)-lepton final states; the statistics in each of these channels is very small and the results are consistent with the SM predictions within the uncertainties.

Using the four different fitting strategies mentioned above, we measure the cross sections for the individual VVV signal channels and for the combined VVV signal, for the two cases where the VH(VV) process is taken as part of the signal, and as part of the background. Figure 7.33 shows the measured cross sections in these two cases, for both the cut-based and BDT-based methods, and in the corresponding tabulated form which is shown in Table 7.37.

We can see that all the measured cross sections agree with the SM predictions, and the uncertainties are mostly statistically dominated. The center values of the cross sections from the BDT-based method agree with the results from the cut-based method within the uncertainties, and the total uncertainty on the VVV cross section is 25-26% (23-25%) for the cut-based (BDT-based) method.

The statistical significances L for the different triboson processes are also computed using the asymptotic approximation [132]:

$$
L = \begin{cases} -2 \ln \lambda(0) & \hat{\mu} \geq 0 \,, \\ 0 & \hat{\mu} < 0 \,, \end{cases} \tag{7.5}
$$

where $\lambda(0)$ is the profile likelihood ratio for $\mu = 0$, and $\hat{\mu}$ is the best fit value of μ.

The observed (expected) significance for the combined VVV production is 5.9 (5.7)σ, and the significances for the individual triboson production processes are found in Table 7.38.

It is worth noting that in the above fitting strategies, the VH(VV) signal strength is either fixed to 1.0 (when VH(VV) is treated as background) or the same as the VVV-onshell signal strength (when VH(VV) is treated as part of the signal). However, the assumptions about VH(VV) in both cases are not necessarily valid, as the VH(VV) process has not been experimentally established yet. In similar analyses where VH(VV) is the main target process [133, 134], the VVV-onshell process is treated as the background and its signal strength is thus fixed to 1.0. A 2D fit of the two signal strengths is needed to distinguish the VVV-onshell and VH(VV) processes. Figure 7.34 shows such a fit for the WWW, WWZ, and combined VVV signals, from which one can see that in most cases the observed VVV-onshell signal is smaller than the SM prediction and the VH(VV) signal is larger than SM, but the deviations are mostly within twice the uncertainty in each case. The likelihood maps in Figure 7.34 are also diagonal. Therefore, it is very hard to distinguish the two signal processes, since the event selections in this search are designed such that the VVV-onshell and VH(VV) signals are similarly distributed in different signal regions. A dedicated event selection is needed to identify VVV-onshell enriched and VH(VV) enriched signal regions in order to perform a better 2D fit of the two signals, and this is discussed in Section 7.9.
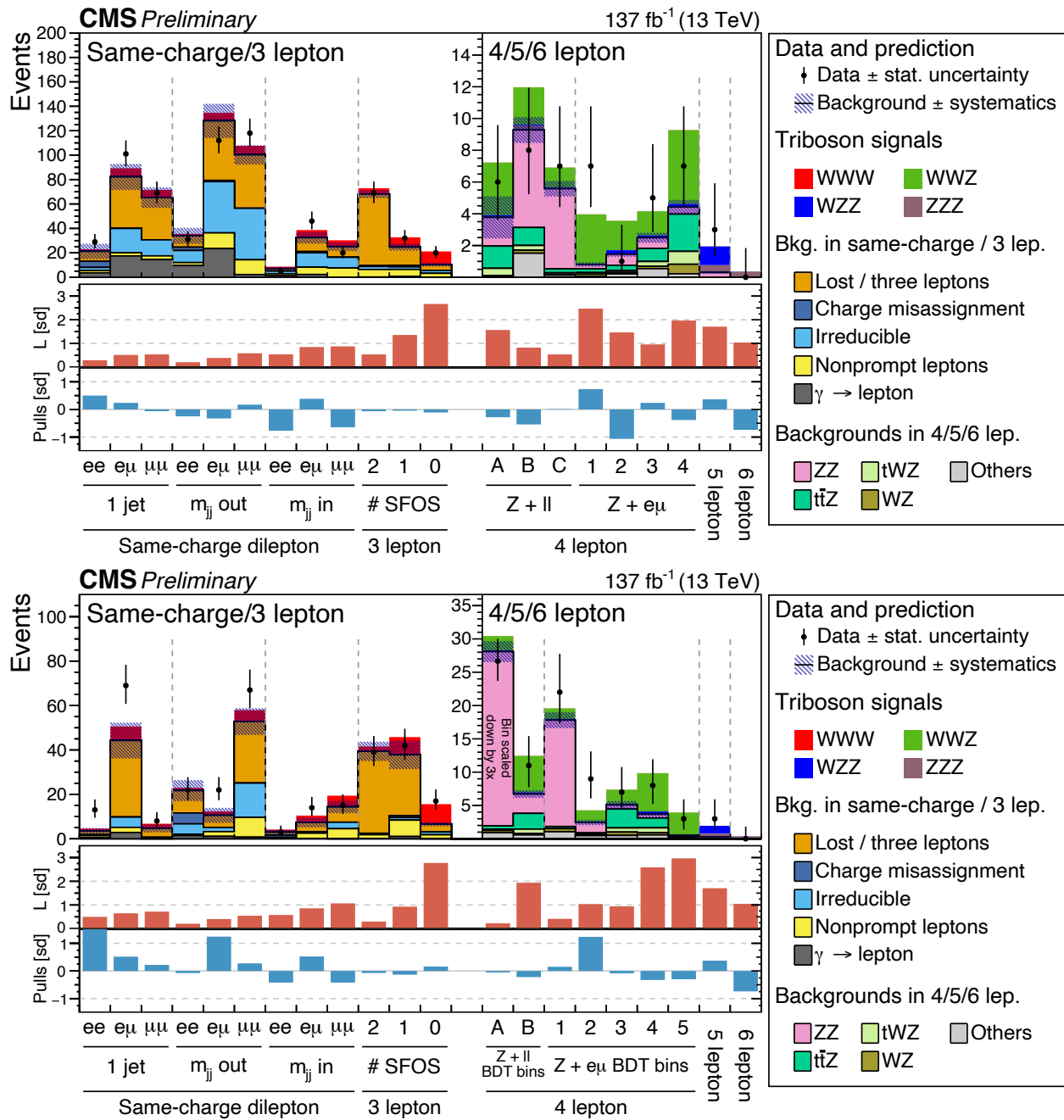
Figure 7.30: The observed events, background, and signal predictions (pre-fit) in all signal regions. The VVV signal is stacked on top of the total background and is based on the SM theoretical cross section. The yields of the top (bottom) plot are based on the cut-based (BDT-based) method. The middle panel shows the expected signal significance of each signal region, which is calculated with both the VVV and VH(VV) processes treated as signals from a single bin data. The bottom panel shows the pulls in each signal region $(\frac{N_{obs}-N_{pred}}{\sqrt{\sigma_{obs}^2+\sigma_{pred}^2}})$.

Table 7.29: The number of expected signal and background events, estimated by using the background estimation methods (pre-fit) discussed in Section 7.5, for the SS/3ℓ final states corresponding to 137 fb^{-1} for the cut-based analysis. The last two rows show the pull and expected significance in each bin (each significance is obtained from a single bin data).

Signal region	SS m_{jj}-in			SS m_{jj}-out			SS 1j		3ℓ		
	$e^\pm e^\pm$	$e^\pm \mu^\pm$	$\mu^\pm \mu^\pm$	$e^\pm e^\pm$	$e^\pm \mu^\pm$	$\mu^\pm \mu^\pm$	$e^\pm \mu^\pm$	$\mu^\pm \mu^\pm$	0SFOS	1SFOS	2SFOS
lost/3ℓ	1.9±0.5	11.4±2.3	8.7±1.0	9.3±2.4	48.7±7.3	43.8±4.4	42.0±4.6	34.6±2.7	4.6±0.9	15.1±1.6	58.6±2.7
Irreducible	2.2±0.5	11.8±3.1	8.7±1.7	9.9±1.4	42.0±5.4	42.1±5.1	19.9±7.7	13.1±2.2	2.4±0.8	2.1±1.2	2.5±0.8
Nonprompt	1.4±1.1	6.0±2.6	7.5±2.3	2.5±1.8	12.9±7.7	12.1±4.0	2.6±1.5	3.1±1.0	3.2±0.8	5.8±1.6	6.2±1.6
Charge flips	< 0.1	1.1±1.0	< 0.1	2.7±3.4	1.0±0.4	< 0.1	0.2±0.1	< 0.1	< 0.1	1.1±0.9	0.7±0.2
γ → nonprompt ℓ	1.5±0.6	2.2±0.8	< 0.1	9.5±4.5	23.4±7.4	2.2±0.8	17.5±5.4	14.4±7.6	< 0.1	0.6±0.8	0.2±0.2
Background sum	7.0±1.4	32.4±4.8	24.9±3.0	33.9±6.5	128±13	100±7	82.1±10.6	65.2±8.4	10.2±1.5	24.7±2.8	68.1±3.2
WWW onshell	1.2±0.3	3.9±0.7	4.4±0.6	1.1±0.3	4.8±0.7	5.0±0.6	2.2±0.4	2.1±0.4	7.1±0.9	4.7±0.7	3.0±0.5
WH→WWW	0.3±0.3	1.9±1.4	0.7±0.3	0.4±0.4	1.6±0.6	1.9±0.7	4.9±2.0	3.7±1.3	3.3±1.5	3.1±1.0	1.4±0.6
WWW total	1.5±0.4	5.8±1.5	5.1±0.7	1.6±0.4	6.5±0.9	6.8±0.9	7.1±2.1	5.8±1.4	10.4±1.8	7.8±1.3	4.5±0.8
WWZ onshell	< 0.1	0.3±0.1	0.2±0.1	< 0.1	< 0.1	0.1±0.0	< 0.1	< 0.1	0.3±0.2	0.2±0.2	0.2±0.2
ZH→WWZ	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	0.4±0.3	< 0.1	0.5±0.4	0.2±0.1	< 0.1	< 0.1
WWZ total	< 0.1	0.3±0.1	0.2±0.1	< 0.1	< 0.1	0.5±0.3	< 0.1	0.5±0.4	0.5±0.2	0.2±0.2	0.2±0.2
WZZ onshell	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
WH→WZZ	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
WZZ total	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
ZZZ onshell	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
ZH→ZZZ	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
ZZZ total	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
VVV onshell	1.2±0.3	4.2±0.7	4.6±0.6	1.1±0.3	4.8±0.7	5.1±0.6	2.2±0.4	2.1±0.4	7.4±0.9	4.9±0.8	3.2±0.5
VH→VVV	0.3±0.3	1.9±1.4	0.7±0.3	0.4±0.4	1.6±0.6	2.2±0.8	4.9±2.0	4.2±1.4	3.4±1.5	3.1±1.0	1.4±0.6
VVV total	1.5±0.4	6.1±1.5	5.3±0.7	1.6±0.4	6.5±0.9	7.3±1.0	7.1±2.1	6.3±1.4	10.9±1.8	8.0±1.3	4.7±0.8
Total	8.5±1.5	38.5±5.0	30.2±3.1	35.5±6.5	134±14	107±7	89.2±10.8	71.5±8.5	21.1±2.3	32.7±3.1	72.8±3.3
Observed	5	46	20	31	112	118	101	69	20	32	69
pull	-0.77	0.38	-0.64	-0.06	-0.25	-0.32	0.23	-0.06	-0.06	-0.04	-0.10
L (expected)	0.54	0.84	0.87	0.20	0.38	0.58	0.51	0.53	2.66	1.35	0.53

Table 7.30: The number of expected signal and background events, estimated by using the background estimation methods (pre-fit) discussed in Section 7.5, for the SS/3ℓ final states corresponding to 137 fb^{-1} for the BDT-based analysis. The last two rows show the pull and expected significance in each bin (each significance is obtained from a single bin data).

Signal region	SS m_{jj}-in			SS m_{jj}-out				SS 1j		3ℓ		
	$e^\pm e^\pm$	$e^\pm \mu^\pm$	$\mu^\pm \mu^\pm$	$e^\pm e^\pm$	$e^\pm \mu^\pm$	$\mu^\pm \mu^\pm$	$e^\pm e^\pm$	$e^\pm \mu^\pm$	$\mu^\pm \mu^\pm$	0 SFOS	1 SFOS	2 SFOS
Lost/3ℓ	1.3±1.0	4.2±1.2	6.9±1.9	10.1±2.8	5.4±3.0	27.7±3.7	1.7±0.7	34.4±6.3	4.1±1.8	3.5±0.7	27.4±5.8	36.9±4.2
Irreducible	1.0±0.1	0.6±0.1	2.9±0.2	4.7±0.4	1.9±0.2	15.4±1.1	0.4±0.0	4.6±0.2	0.5±0.1	1.3±0.1	1.2±0.1	0.3±0.0
Nonprompt	0.5±0.6	2.4±1.8	4.5±1.9	0.6±0.7	2.2±1.3	8.9±4.6	0.9±1.0	2.3±1.0	0.2±0.1	1.8±0.6	7.4±2.4	1.7±1.1
Charge flips	< 0.1	< 0.1	< 0.1	4.9±3.7	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	0.9±0.8	0.3±0.1
$\gamma \rightarrow$ nonprompt ℓ	0.1±0.2	< 0.1	< 0.1	1.4±0.5	1.1±0.4	0.7±0.4	0.5±0.5	2.8±4.9	< 0.1	< 0.1	1.0±0.4	< 0.1
Background sum	2.9±1.1	7.2±2.2	14.3±2.7	21.7±4.7	10.5±3.3	52.8±6.0	3.5±1.4	44.2±8.0	4.7±1.8	6.5±0.9	37.8±6.4	39.3±4.3
WWW onshell	0.8±0.1	2.0±0.3	4.0±0.4	0.8±0.1	0.8±0.2	2.8±0.3	0.3±0.1	1.0±0.1	0.3±0.1	5.8±0.6	3.8±0.5	1.6±0.2
WH→WWW	0.3±0.2	1.0±0.4	1.1±0.3	0.5±0.2	1.0±0.5	2.2±0.6	0.8±0.3	5.3±1.5	1.7±0.6	2.8±0.8	4.1±0.9	0.5±0.3
WWW total	1.1±0.2	3.0±0.5	5.1±0.5	1.2±0.2	1.8±0.5	5.0±0.6	1.1±0.3	6.3±1.6	2.1±0.6	8.6±1.0	7.9±1.0	2.1±0.4
WWZ onshell	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	0.2±0.1	< 0.1	< 0.1
ZH→WWZ	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	0.2±0.1	0.1±0.1	< 0.1
WWZ total	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	0.4±0.1	0.1±0.1	< 0.1
WZZ onshell	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
WH→WZZ	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	0.1±0.2	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
WZZ total	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	0.1±0.2	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
ZZZ onshell	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
ZH→ZZZ	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
ZZZ total	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
VVV onshell	0.8±0.1	2.0±0.3	4.0±0.4	0.8±0.1	0.8±0.2	2.8±0.3	0.3±0.1	1.0±0.1	0.3±0.1	6.0±0.6	3.8±0.5	1.6±0.2
VH→VVV	0.3±0.2	1.0±0.4	1.1±0.3	0.5±0.2	1.0±0.5	2.3±0.6	0.8±0.3	5.3±1.5	1.7±0.6	3.0±0.8	4.2±0.9	0.5±0.3
VVV total	1.1±0.2	3.0±0.5	5.1±0.5	1.2±0.2	1.8±0.5	5.1±0.7	1.1±0.3	6.3±1.6	2.1±0.6	9.0±1.0	8.0±1.0	2.1±0.4
Total	4.0±1.2	10.3±2.2	19.4±2.8	22.9±4.7	12.3±3.4	57.9±6.0	4.6±1.4	50.5±8.2	6.8±1.9	15.5±1.3	45.8±6.5	41.3±4.3
Observed	3	14	15	22	22	67	13	69	8	17	42	39
pull	-0.41	0.52	-0.41	-0.07	1.24	0.27	2.14	0.51	0.21	-0.07	-0.13	0.15
L (expected)	0.57	0.85	1.06	0.20	0.40	0.54	0.49	0.64	0.71	2.77	0.92	0.29

Table 7.31: The number of expected signal and background events, estimated by using the background estimation methods (pre-fit) discussed in Section 7.5, for the 4/5/6-lepton final states corresponding to $137\,\mathrm{fb}^{-1}$ for the cut-based analysis. The last two rows show the pull and expected significance in each bin (each significance is obtained from a single bin data).

Signal	$4\ell\,\mathrm{e\mu}$					$4\ell\,\mathrm{ee/\mu\mu}$		5ℓ	6ℓ
region	bin 1	bin 2	bin 3	bin 4	bin A	bin B	bin C		
ZZ	0.3±0.0	0.7±0.0	0.7±0.0	0.4±0.0	1.8±0.2	6.1±0.6	5.1±0.5	0.3±0.1	< 0.1
$t\bar{t}Z$	0.2±0.0	0.3±0.1	0.8±0.1	2.3±0.4	1.4±0.2	1.1±0.2	0.2±0.0	0.1±0.0	< 0.1
tWZ	0.1±0.0	0.1±0.0	0.3±0.0	0.8±0.1	0.5±0.1	0.3±0.1	< 0.1	< 0.1	< 0.1
WZ	0.2±0.1	< 0.1	0.1±0.1	0.6±0.2	< 0.1	0.2±0.1	< 0.1	< 0.1	< 0.1
Higgs	< 0.1	0.2±0.1	< 0.1	< 0.1	< 0.1	1.0±0.3	0.1±0.1	< 0.1	< 0.1
Other	< 0.1	< 0.1	0.5±0.2	0.2±0.1	< 0.1	0.5±0.5	< 0.1	< 0.1	< 0.1
Background sum	0.8±0.1	1.3±0.1	2.5±0.3	4.4±0.4	3.7±0.3	9.3±0.8	5.4±0.5	0.4±0.1	0.0±0.0
WWW onshell	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
WH→WWW	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
WWW total	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
WWZ onshell	0.5±0.1	0.6±0.1	1.2±0.1	4.5±0.3	2.4±0.3	1.3±0.1	0.7±0.1	< 0.1	< 0.1
ZH→WWZ	2.6±0.2	1.2±0.1	0.3±0.0	0.1±0.0	0.9±0.1	1.0±0.1	0.6±0.0	< 0.1	< 0.1
WWZ total	3.1±0.2	1.8±0.1	1.5±0.1	4.6±0.4	3.3±0.3	2.4±0.1	1.2±0.1	< 0.1	< 0.1
WZZ onshell	< 0.1	< 0.1	< 0.1	0.2±0.1	< 0.1	< 0.1	< 0.1	1.2±0.2	< 0.1
WH→WZZ	< 0.1	0.2±0.2	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
WZZ total	< 0.1	0.2±0.2	< 0.1	0.2±0.1	< 0.1	< 0.1	< 0.1	1.2±0.2	< 0.1
ZZZ onshell	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	0.4±0.1	0.3±0.0
ZH→ZZZ	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	0.2±0.1	< 0.1	< 0.1	< 0.1
ZZZ total	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	0.2±0.1	< 0.1	0.4±0.1	0.3±0.0
VVV onshell	0.5±0.1	0.6±0.1	1.2±0.1	4.7±0.4	2.4±0.3	1.3±0.1	0.7±0.1	1.5±0.2	0.3±0.0
VH→VVV	2.6±0.2	1.5±0.2	0.3±0.0	0.1±0.0	0.9±0.1	1.2±0.1	0.6±0.0	< 0.1	< 0.1
VVV total	3.1±0.2	2.1±0.2	1.5±0.1	4.8±0.4	3.3±0.3	2.5±0.2	1.2±0.1	1.5±0.2	0.3±0.0
Total	3.9±0.2	3.4±0.3	4.0±0.3	9.2±0.6	7.0±0.4	11.8±0.8	6.7±0.5	2.0±0.2	0.3±0.0
Observed	7	1	5	7	6	8	7	3	0
pull	0.73	-1.06	0.23	-0.38	-0.28	-0.54	0.01	0.36	-0.74
L (expected)	2.47	1.46	0.95	1.96	1.56	0.82	0.53	1.70	1.04

Table 7.32: The number of expected signal and background events, estimated by using the background estimation methods (pre-fit) discussed in Section 7.5, for the 4/5/6-lepton final states corresponding to $137\,\text{fb}^{-1}$ for the BDT-based analysis. The last two rows show the pull and expected significance in each bin.

Signal			4ℓ eμ			4ℓ ee/$\mu\mu$		5ℓ	6ℓ
region	bin 1	bin 2	bin 3	bin 4	bin 5	bin A	bin B		
ZZ	16.0±1.1	1.6±0.1	0.6±0.1	0.6±0.1	0.2±0.0	78.5±4.7	2.9±0.3	0.3±0.1	< 0.1
t$\bar{\text{t}}$Z	0.3±0.1	0.1±0.1	2.8±0.5	1.5±0.3	< 0.1	1.6±0.2	2.4±0.4	0.1±0.0	< 0.1
tWZ	< 0.1	< 0.1	0.6±0.1	0.8±0.1	< 0.1	0.5±0.1	0.7±0.1	< 0.1	< 0.1
WZ	0.5±0.2	0.2±0.1	0.5±0.2	0.3±0.2	0.1±0.1	1.0±0.3	0.2±0.1	< 0.1	< 0.1
Higgs	1.0±0.4	< 0.1	0.3±0.1	0.1±0.1	< 0.1	2.2±0.4	0.5±0.2	< 0.1	< 0.1
Other	< 0.1	0.4±0.2	0.2±0.1	0.4±0.3	< 0.1	0.5±0.5	< 0.1	< 0.1	< 0.1
Background sum	17.7±1.2	2.3±0.3	5.1±0.6	3.7±0.5	0.3±0.1	84.3±4.7	6.6±0.6	0.4±0.1	0.0±0.0
WWW onshell	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
WH→WWW	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
WWW total	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
WWZ onshell	0.3±0.1	0.4±0.1	1.6±0.4	4.3±0.5	1.2±0.2	3.1±0.4	3.8±0.5	< 0.1	< 0.1
ZH→WWZ	1.2±0.1	1.3±0.1	0.5±0.1	1.5±0.1	2.2±0.1	3.4±0.2	1.7±0.1	< 0.1	< 0.1
WWZ total	1.6±0.1	1.7±0.1	2.2±0.4	5.8±0.5	3.4±0.2	6.5±0.5	5.5±0.5	< 0.1	< 0.1
WZZ onshell	< 0.1	< 0.1	< 0.1	0.2±0.1	< 0.1	0.2±0.1	< 0.1	1.2±0.2	< 0.1
WH→WZZ	< 0.1	< 0.1	< 0.1	0.2±0.2	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
WZZ total	< 0.1	< 0.1	< 0.1	0.4±0.2	< 0.1	0.2±0.1	< 0.1	1.2±0.2	< 0.1
ZZZ onshell	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	0.1±0.0	< 0.1	0.4±0.1	0.3±0.0
ZH→ZZZ	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	0.1±0.1	< 0.1	< 0.1
ZZZ total	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	0.1±0.0	0.1±0.1	0.4±0.1	0.3±0.0
VVV onshell	0.3±0.1	0.4±0.1	1.6±0.4	4.5±0.5	1.2±0.2	3.5±0.4	3.8±0.5	1.5±0.2	0.3±0.0
VH→VVV	1.2±0.1	1.3±0.1	0.5±0.1	1.7±0.2	2.2±0.1	3.4±0.2	1.8±0.2	< 0.1	< 0.1
VVV total	1.6±0.1	1.7±0.1	2.2±0.4	6.2±0.6	3.4±0.2	6.9±0.5	5.6±0.5	1.5±0.2	0.3±0.0
Total	19.3±1.2	4.0±0.3	7.3±0.7	9.8±0.7	3.7±0.3	91.2±4.7	12.2±0.8	2.0±0.2	0.3±0.0
Observed	22	9	7	8	3	80	11	3	0
pull	0.05	1.23	-0.08	-0.32	-0.30	-0.14	-0.22	0.36	-0.74
L (expected)	0.14	1.03	0.94	2.58	2.96	0.67	1.93	1.70	1.04

Figure 7.31: The observed events, background, and signal yields after the fit in all signal regions. The fit is performed with four different signal strengths floating at the same time, which corresponds to the WWW, WWZ, WZZ, and ZZZ processes, where the VVV-onshell and VH(VV) signals are not distinguished. The predicted VVV signal is stacked on top of the total background and is from the fit. The yields of the top (bottom) plot are based on the cut-based (BDT-based) method. The middle panel shows the expected signal significance of each signal region, which is calculated with the total VVV+VH taken as signal from a single bin fit. The bottom panel shows the pulls in each signal region ($\frac{N_{obs}-N_{pred}}{\sqrt{\sigma_{obs}^2+\sigma_{pred}^2}}$).

Figure 7.32: The observed events, background, and signal yields after the fit in all signal regions. The fit is performed with four different signal strengths floating at the same time, which corresponds to the WWW, WWZ, WZZ, and ZZZ processes, where the VVV-onshell and VH(VV) signals are not distinguished. The predicted VVV signal is stacked on top of the total background and is from the fit. The yields of the top (bottom) plot are based on the cut-based (BDT-based) method. The middle panel shows the observed signal significance of each signal region, which is calculated with the total VVV+VH taken as signal from a single bin fit. The bottom panel shows the pulls in each signal region ($\frac{N_{\mathrm{obs}} - N_{\mathrm{pred}}}{\sqrt{\sigma_{\mathrm{obs}}^2 + \sigma_{\mathrm{pred}}^2}}$).

Table 7.33: The number of expected signal and background events, estimated by using the background estimation methods discussed in Section 7.5, after simultaneously fitting the four different VVV signal strengths WWW, WWZ, WZZ, ZZZ, for the SS/3ℓ final states corresponding to 137 fb^{-1} for the cut-based analysis. The last two rows show the pull and observed significance in each bin (each significance is obtained from a single bin fit).

Signal region	SS m_{jj}-in			SS m_{jj}-out			SS 1j			3ℓ		
	$e^\pm e^\pm$	$e^\pm \mu^\pm$	$\mu^\pm \mu^\pm$	$e^\pm e^\pm$	$e^\pm \mu^\pm$	$\mu^\pm \mu^\pm$	$e^\pm e^\pm$	$e^\pm \mu^\pm$	$\mu^\pm \mu^\pm$	0SFOS	1SFOS	2SFOS
lost/3ℓ	1.8±0.4	10.9±2.0	8.7±1.0	8.8±1.7	46.0±6.2	44.8±4.4	8.4±1.3	43.5±4.4	34.5±2.7	4.6±0.8	15.1±1.5	58.3±2.4
Irreducible	2.1±0.4	13.0±3.6	8.4±1.4	9.8±1.4	41.1±4.5	42.8±4.7	2.6±0.6	22.8±8.6	13.2±1.9	2.5±0.9	2.2±1.2	2.5±0.8
Nonprompt	1.3±0.9	5.8±2.4	6.8±2.2	2.3±1.3	12.0±6.1	11.2±3.8	1.8±2.9	2.4±1.3	2.8±1.1	3.0±0.9	5.7±1.6	5.9±1.6
Charge flips	< 0.1	1.2±2.0	< 0.1	2.6±1.6	1.0±0.5	< 0.1	6.9±4.7	0.2±0.1	< 0.1	< 0.1	1.1±1.3	0.7±0.2
$\gamma \rightarrow$ nonprompt ℓ	1.4±0.4	2.3±0.9	< 0.1	8.6±3.1	19.2±5.1	2.3±0.9	3.8±1.1	19.7±6.0	13.8±7.0	< 0.1	0.6±0.7	0.2±0.3
Background sum	6.6±1.2	33.3±5.2	23.9±2.8	32.1±4.3	119±11	101±7	23.6±5.8	88.7±11.4	64.4±7.8	10.1±1.5	24.7±2.9	67.6±3.1
WWW onshell	1.0±0.5	3.3±1.5	3.5±1.6	0.9±0.5	3.9±1.8	4.1±1.9	0.5±0.3	1.8±0.8	1.7±0.9	5.9±2.6	3.8±1.7	2.5±1.2
WH$\rightarrow$WWW	0.2±0.3	1.9±1.5	0.6±0.4	0.4±0.4	1.3±0.8	1.7±1.0	0.8±0.5	4.5±2.7	3.3±2.0	3.0±1.7	2.7±1.5	1.3±0.8
WWW total	1.2±0.6	5.1±2.2	4.1±1.6	1.3±0.6	5.3±2.0	5.7±2.1	1.4±0.6	6.3±2.8	5.0±2.2	8.8±3.1	6.6±2.3	3.8±1.4
WWZ onshell	< 0.1	0.3±0.2	0.2±0.1	< 0.1	< 0.1	0.1±0.1	< 0.1	< 0.1	< 0.1	0.3±0.2	0.2±0.2	0.2±0.1
ZH$\rightarrow$WWZ	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	0.3±0.3	< 0.1	< 0.1	0.4±0.4	0.2±0.1	< 0.1	< 0.1
WWZ total	< 0.1	0.3±0.2	0.2±0.1	< 0.1	< 0.1	0.4±0.3	< 0.1	< 0.1	0.4±0.4	0.4±0.2	0.2±0.2	0.2±0.1
WZZ onshell	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
WH$\rightarrow$WZZ	< 0.1	0.2±0.6	< 0.1	< 0.1	0.2±8.5	0.2±0.5	0.2±1.0	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
WZZ total	< 0.1	0.2±0.6	< 0.1	< 0.1	0.2±8.5	0.2±0.5	0.2±1.0	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
ZZZ onshell	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
ZH$\rightarrow$ZZZ	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
ZZZ total	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
VVV onshell	1.0±0.5	3.5±1.5	3.7±1.6	0.9±0.5	3.9±1.8	4.2±1.9	0.5±0.3	1.8±0.8	1.7±0.9	6.1±2.6	4.0±1.8	2.7±1.2
VH$\rightarrow$VVV	0.2±0.3	2.0±1.6	0.6±0.4	0.4±0.4	1.5±8.5	2.2±1.1	1.0±1.1	4.5±2.7	3.7±2.0	3.1±1.7	2.7±1.5	1.3±0.8
VVV total	1.2±0.6	5.5±2.2	4.2±1.6	1.3±0.6	5.4±8.7	6.3±2.2	1.5±1.2	6.3±2.8	5.4±2.2	9.3±3.1	6.8±2.3	3.9±1.4
Total	7.8±1.3	38.8±5.7	28.2±3.3	33.5±4.4	124±14	107±7	25.1±5.9	95.0±11.8	69.8±8.1	19.3±3.4	31.4±3.7	71.5±3.4
Observed	5	46	20	31	112	118	29	101	69	20	32	69
pull	-0.66	0.36	-0.54	-0.14	-0.19	0.16	0.28	0.11	-0.02	0.05	0.03	-0.04
L (observed)	0.00	1.74	0.00	0.00	0.00	1.36	1.17	1.30	0.33	2.41	1.24	0.10

Table 7.34: The number of expected signal and background events, estimated by using the background estimation methods after fitting the four different VVV signals (in total four signal strengths: WWW, WWZ, WZZ, ZZZ) simultaneously, for the SS/3ℓ final states corresponding to 137 fb^{-1}, for the BDT-based analysis. The last two rows show the pull and observed significance in each bin (each significance is obtained from a single bin fit).
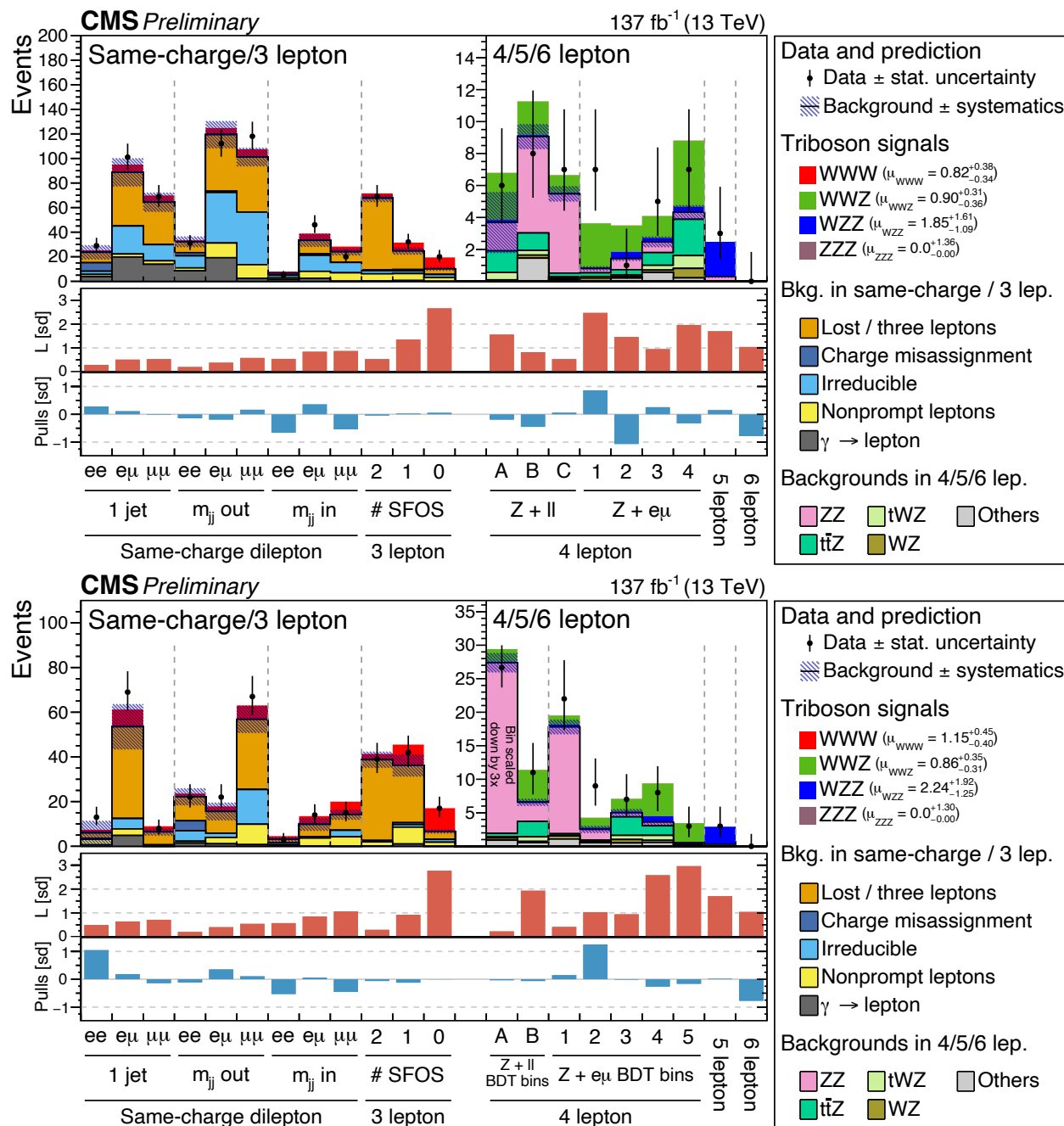
Signal region	SS m_{jj}-in			SS m_{jj}-out			SS 1j			3ℓ		
	$e^\pm e^\pm$	$e^\pm \mu^\pm$	$\mu^\pm \mu^\pm$	$e^\pm e^\pm$	$e^\pm \mu^\pm$	$\mu^\pm \mu^\pm$	$e^\pm e^\pm$	$e^\pm \mu^\pm$	$\mu^\pm \mu^\pm$	0SFOS	1SFOS	2SFOS
lost/3ℓ	1.4±0.9	5.5±1.6	7.0±1.7	10.7±2.6	9.7±3.6	31.4±3.8	2.5±1.1	41.0±6.1	5.8±1.6	3.5±0.7	25.6±4.2	36.1±3.1
Irreducible	1.0±0.1	0.6±0.1	2.9±0.2	4.7±0.4	1.9±0.2	15.5±1.2	0.4±0.0	4.6±0.2	0.5±0.1	1.3±0.1	1.2±0.1	0.3±0.0
Nonprompt	0.6±0.6	3.6±2.4	4.2±1.5	0.8±1.0	2.8±1.5	9.1±4.5	2.5±5.2	2.9±1.4	0.2±0.1	1.8±0.5	7.5±2.3	1.8±1.1
Charge flips	< 0.1	< 0.1	< 0.1	4.5±2.5	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	0.8±1.2	0.3±0.1
γ → nonprompt ℓ	0.1±0.2	< 0.1	< 0.1	1.4±0.5	1.1±0.4	0.7±0.4	0.6±1.2	4.8±8.0	< 0.1	< 0.1	1.0±0.4	< 0.1
Background sum	3.1±1.1	9.7±2.9	14.2±2.3	22.1±3.8	15.6±4.0	56.8±6.0	5.9±5.4	53.4±10.1	6.4±1.6	6.6±0.9	36.2±5.0	38.6±3.3
WWW onshell	0.9±0.4	2.3±0.9	4.6±1.7	0.9±0.4	1.0±0.6	3.3±1.3	0.3±0.2	1.2±0.4	0.4±0.2	6.7±2.4	4.3±1.6	1.8±0.7
WH→WWW	0.4±0.3	1.3±0.9	1.2±0.5	0.5±0.3	1.3±1.3	2.7±1.2	1.1±0.8	6.5±3.1	2.2±1.1	3.4±1.6	5.0±2.1	0.6±0.6
WWW total	1.3±0.5	3.7±1.3	5.8±1.7	1.5±0.5	2.3±1.4	6.0±1.7	1.4±0.8	7.7±3.1	2.5±1.1	10.1±2.9	9.3±2.6	2.4±0.9
WWZ onshell	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	0.2±0.1	< 0.1	< 0.1
ZH→WWZ	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	0.3±0.9	< 0.1	< 0.1	< 0.1	0.1±0.1	< 0.1	< 0.1
WWZ total	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	0.3±0.9	< 0.1	< 0.1	< 0.1	0.3±0.1	< 0.1	< 0.1
WZZ onshell	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
WH→WZZ	0.2±0.3	< 0.1	< 0.1	< 0.1	< 0.1	0.3±0.9	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	0.2±0.9
WZZ total	0.2±0.3	< 0.1	< 0.1	< 0.1	< 0.1	0.3±0.9	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	0.2±0.9
ZZZ onshell	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
ZH→ZZZ	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
ZZZ total	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
VVV onshell	0.9±0.4	2.3±0.9	4.6±1.7	0.9±0.4	1.0±0.6	3.3±1.3	0.3±0.2	1.2±0.4	0.4±0.2	6.9±2.4	4.3±1.6	1.8±0.7
VH→VVV	0.5±0.4	1.3±0.9	1.2±0.5	0.5±0.3	1.3±1.3	3.0±1.4	1.1±0.8	6.5±3.1	2.2±1.1	3.6±1.6	5.0±2.1	0.8±1.1
VVV total	1.5±0.6	3.7±1.3	5.8±1.7	1.5±0.5	2.3±1.4	6.3±1.9	1.4±0.8	7.7±3.1	2.5±1.1	10.4±2.9	9.3±2.6	2.6±1.3
Total	4.5±1.2	13.4±3.1	20.0±2.9	23.6±3.8	17.8±4.2	63.0±6.3	7.4±5.5	61.1±10.6	9.0±2.0	17.0±3.0	45.4±5.6	41.2±3.5
Observed	3	14	15	22	22	67	13	69	8	17	42	39
pull	-0.53	0.06	-0.46	-0.12	0.35	0.11	1.005	0.18	-0.14	0.00	-0.12	-0.06
L (observed)	0.04	1.63	0.15	0.05	1.98	1.45	2.96	2.16	1.07	3.14	0.49	0.00

Table 7.35: The number of expected signal and background events, estimated by using the background estimation methods after fitting simultaneously the different VVV signals (in total four signal strengths: WWW, WWZ, WZZ, ZZZ), for the 4/5/6-lepton final states corresponding to $137\,\mathrm{fb}^{-1}$, for the cut-based analysis. The last two rows show the pull and observed significance in each bin (each significance is obtained from a single bin fit).

Signal	4ℓ eμ				4ℓ ee/$\mu\mu$			5ℓ	6ℓ
region	bin 1	bin 2	bin 3	bin 4	bin A	bin B	bin C		
ZZ	0.3±0.0	0.7±0.0	0.7±0.0	0.4±0.0	1.8±0.2	6.0±0.6	5.0±0.5	0.3±0.1	< 0.1
t$\bar{\mathrm{t}}$Z	0.2±0.0	0.3±0.1	0.8±0.1	2.3±0.4	1.4±0.2	1.1±0.2	0.2±0.0	0.1±0.0	< 0.1
tWZ	< 0.1	0.1±0.0	0.3±0.0	0.8±0.1	0.5±0.1	0.3±0.1	< 0.1	< 0.1	< 0.1
WZ	0.2±0.1	< 0.1	0.1±0.2	0.6±0.2	< 0.1	0.2±0.1	< 0.1	< 0.1	< 0.1
Higgs	< 0.1	0.2±0.1	< 0.1	< 0.1	< 0.1	1.0±0.2	0.1±0.1	< 0.1	< 0.1
Other	< 0.1	< 0.1	0.6±0.3	0.2±0.1	< 0.1	0.5±0.4	< 0.1	< 0.1	< 0.1
Background sum	0.7±0.1	1.3±0.1	2.5±0.3	4.3±0.4	3.6±0.3	9.1±0.8	5.3±0.5	0.4±0.1	0.0±0.0
WWW onshell	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
WH→WWW	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
WWW total	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
WWZ onshell	0.5±0.2	0.5±0.2	1.1±0.4	4.0±1.6	2.1±0.9	1.2±0.4	0.6±0.2	< 0.1	< 0.1
ZH→WWZ	2.3±0.9	1.1±0.4	0.3±0.1	0.1±0.1	0.8±0.3	0.9±0.4	0.5±0.2	< 0.1	< 0.1
WWZ total	2.8±0.9	1.6±0.5	1.4±0.4	4.1±1.6	2.9±1.0	2.1±0.6	1.1±0.3	< 0.1	< 0.1
WZZ onshell	< 0.1	< 0.1	0.1±0.1	0.4±0.3	0.2±0.2	0.1±0.1	< 0.1	2.2±1.5	< 0.1
WH→WZZ	< 0.1	0.4±0.3	0.1±0.2	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
WZZ total	< 0.1	0.4±0.3	0.2±0.2	0.4±0.3	0.2±0.2	0.1±0.1	< 0.1	2.2±1.5	< 0.1
ZZZ onshell	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
ZH→ZZZ	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
ZZZ total	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
VVV onshell	0.5±0.2	0.5±0.2	1.2±0.4	4.4±1.6	2.3±0.9	1.3±0.5	0.6±0.2	2.2±1.5	< 0.1
VH→VVV	2.3±0.9	1.5±0.5	0.4±0.3	0.1±0.1	0.8±0.3	0.9±0.4	0.5±0.2	< 0.1	< 0.1
VVV total	2.8±0.9	2.0±0.6	1.6±0.5	4.5±1.6	3.1±1.0	2.2±0.6	1.1±0.3	2.2±1.5	< 0.1
Total	3.5±0.9	3.3±0.6	4.1±0.6	8.8±1.7	6.7±1.0	11.3±1.0	6.4±0.6	2.6±1.5	0.0±0.0
Observed	7	1	5	7	6	8	7	3	0
pull	0.86	-1.07	0.26	-0.32	-0.19	-0.45	0.06	0.15	-0.78
L (observed)	4.09	0.00	0.74	1.12	1.05	0.00	0.57	2.39	0.00

Table 7.36: The number of expected signal and background events, estimated by using the background estimation methods after fitting simultaneously the different VVV signals (in total four signal strengths: WWW, WWZ, WZZ, ZZZ), for the 4/5/6-lepton final states corresponding to 137 fb^{-1}, for the BDT-based analysis. The last two rows show the pull and observed significance in each bin (each significance is obtained from a single bin fit).

Signal	4ℓ eμ					4ℓ ee/$\mu\mu$		5ℓ	6ℓ
region	bin 1	bin 2	bin 3	bin 4	bin 5	bin A	bin B		
ZZ	15.9±1.0	1.6±0.1	0.6±0.1	0.6±0.1	0.2±0.0	76.4±4.3	2.9±0.3	0.3±0.1	< 0.1
t$\bar{\text{t}}$Z	0.2±0.1	0.1±0.1	2.8±0.5	1.4±0.2	< 0.1	1.5±0.3	2.3±0.3	0.1±0.0	< 0.1
tWZ	< 0.1	< 0.1	0.6±0.1	0.7±0.1	< 0.1	0.5±0.1	0.7±0.1	< 0.1	< 0.1
WZ	0.5±0.2	0.2±0.2	0.5±0.2	0.3±0.3	0.1±0.1	1.0±0.4	0.2±0.1	< 0.1	< 0.1
Higgs	1.0±0.4	0.1±0.4	0.3±0.1	0.1±0.1	< 0.1	2.2±0.4	0.5±0.2	< 0.1	< 0.1
Other	< 0.1	0.4±0.3	0.2±0.2	0.4±0.2	< 0.1	0.5±0.5	< 0.1	< 0.1	< 0.1
Background sum	17.7±1.1	2.4±0.5	5.0±0.6	3.6±0.4	0.3±0.1	82.2±4.3	6.5±0.5	0.4±0.1	0.0±0.0
WWW onshell	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
WH→WWW	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
WWW total	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
WWZ onshell	0.3±0.1	0.4±0.2	1.4±0.7	3.6±1.5	1.0±0.5	2.7±1.2	3.2±1.4	< 0.1	< 0.1
ZH→WWZ	1.1±0.5	1.1±0.5	0.5±0.2	1.3±0.5	1.8±0.8	2.9±1.2	1.5±0.6	< 0.1	< 0.1
WWZ total	1.3±0.5	1.5±0.5	1.9±0.8	4.9±1.6	2.9±0.9	5.6±1.7	4.7±1.5	< 0.1	< 0.1
WZZ onshell	0.2±0.2	< 0.1	0.2±0.2	0.4±0.4	< 0.1	0.5±0.4	0.2±0.2	2.6±1.8	< 0.1
WH→WZZ	0.2±0.3	0.2±0.3	< 0.1	0.5±0.5	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
WZZ total	0.4±0.3	0.2±0.3	0.2±0.2	0.9±0.7	< 0.1	0.5±0.4	0.2±0.2	2.6±1.8	< 0.1
ZZZ onshell	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
ZH→ZZZ	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
ZZZ total	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
VVV onshell	0.5±0.2	0.4±0.2	1.6±0.8	4.0±1.5	1.0±0.5	3.2±1.3	3.4±1.4	2.6±1.8	< 0.1
VH→VVV	1.2±0.5	1.3±0.6	0.5±0.2	1.7±0.8	1.8±0.8	2.9±1.2	1.5±0.6	< 0.1	< 0.1
VVV total	1.7±0.6	1.7±0.6	2.1±0.8	5.8±1.7	2.9±0.9	6.1±1.8	4.8±1.5	2.6±1.8	< 0.1
Total	19.4±1.2	4.1±0.8	7.1±1.0	9.4±1.8	3.2±0.9	88.2±4.7	11.3±1.6	3.0±1.8	0.0±0.0
Observed	22	9	7	8	3	80	11	3	0
pull	0.05	1.25	-0.02	-0.26	-0.16	-0.11	-0.06	0.02	-0.78
L (observed)	0.31	3.12	0.80	1.90	2.33	0.00	1.48	2.39	0.00
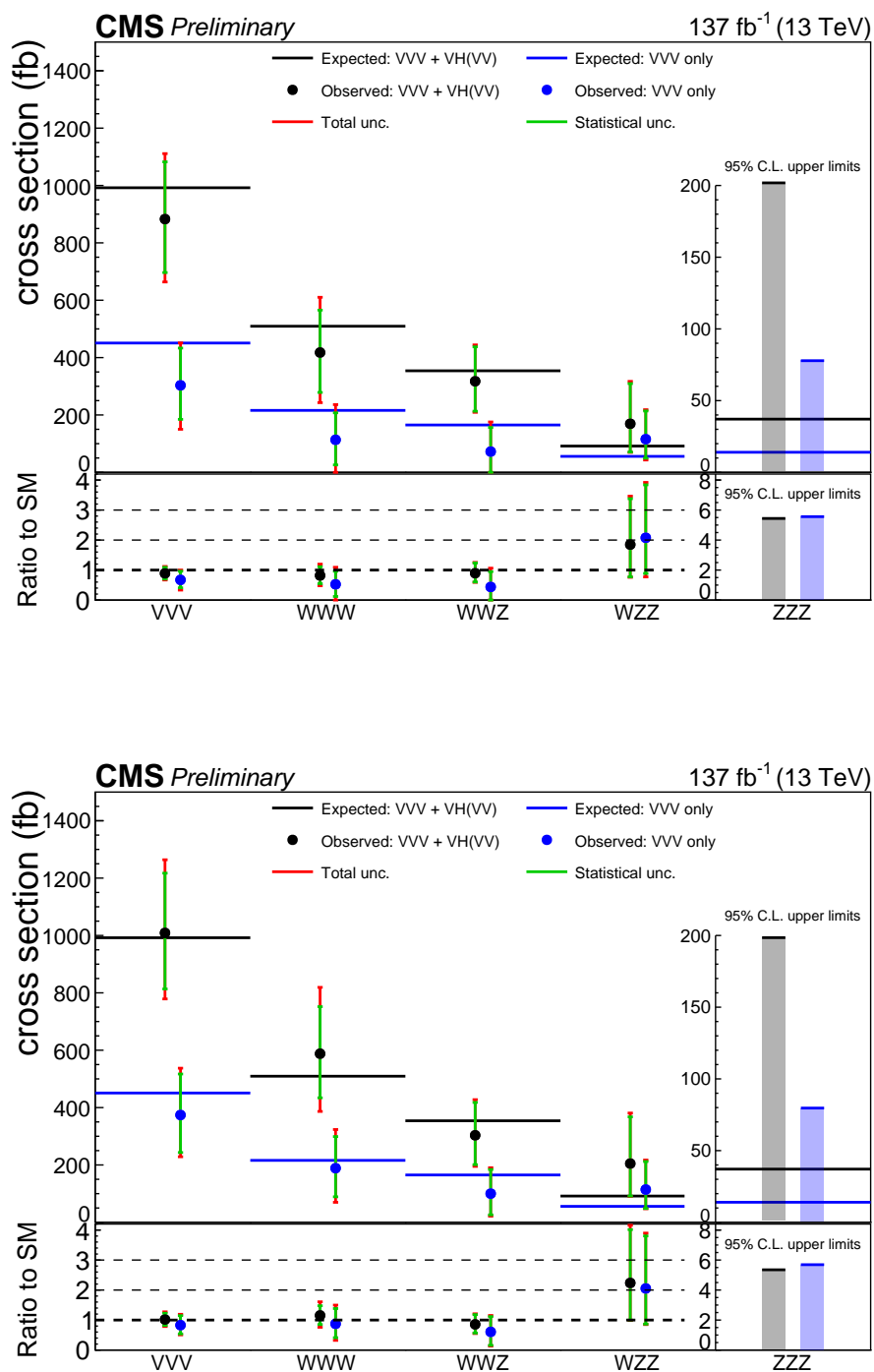
Figure 7.33: The measured cross sections for the different VVV processeses, together with the corresponding SM predictions, and the ratios between each of the measured cross sections and the corresponding SM predictions. The top plot shows the result from the cut-based method, and the bottom plot shows the result from the BDT-based method. For the ZZZ process, the 95% C.L. upper limits on the cross section is shown.

Table 7.37: The measured triboson cross sections σ and the corresponding SM predictions. The uncertainties listed in the brackets are statistical and systematic respectively. The combined VVV cross section is calculated from the fit with a single signal strength μ_{VVV}, and the individual cross sections are calculated from a simultaneous fit with four signal strengths, as discussed in the text. For the ZZZ process, 95% confidence level upper limits are reported.

Process	$\sigma_{\text{measured, BDT}}$ (fb)	$\sigma_{\text{measured, cut}}$ (fb)	σ_{SM} (fb)
	VH(VV) contributions as signal		
VVV	$1010^{+260}_{-230}\left(^{+210}_{-200}{}^{+150}_{-120}\right)$	$880^{+230}_{-220}\left(^{+200}_{-190}{}^{+110}_{-120}\right)$	992
WWW	$590^{+230}_{-200}\left(^{+160}_{-150}{}^{+160}_{-130}\right)$	$420^{+190}_{-170}\left(^{+150}_{-140}{}^{+120}_{-110}\right)$	509
WWZ	$300^{+120}_{-110}\left(^{+120}_{-100}{}^{+50}_{-40}\right)$	$320^{+130}_{-110}\left(^{+120}_{-100}{}^{+40}_{-30}\right)$	354
WZZ	$200^{+180}_{-110}\left(^{+160}_{-110}{}^{+70}_{-20}\right)$	$170^{+150}_{-100}\left(^{+140}_{-100}{}^{+50}_{-20}\right)$	92
ZZZ	<200	<200	37
	VH(VV) contributions as background		
VVV	$370^{+160}_{-150}\left(^{+140}_{-130}{}^{+80}_{-60}\right)$	$300^{+150}_{-150}\left(^{+130}_{-120}{}^{+70}_{-100}\right)$	451
WWW	$190^{+130}_{-120}\left(^{+110}_{-100}{}^{+80}_{-70}\right)$	$110^{+120}_{-120}\left(^{+90}_{-90}{}^{+80}_{-80}\right)$	216
WWZ	$100^{+90}_{-80}\left(^{+80}_{-70}{}^{+30}_{-30}\right)$	$70^{+100}_{-80}\left(^{+80}_{-70}{}^{+60}_{-30}\right)$	165
WZZ	$110^{+100}_{-70}\left(^{+100}_{-70}{}^{+30}_{-10}\right)$	$120^{+100}_{-70}\left(^{+100}_{-70}{}^{+30}_{-30}\right)$	56
ZZZ	<80	<80	14

Table 7.38: The observed (expected) significance (σ) for the different and combined triboson processes, for the two different (cut-based, BDT-based) analysis methods, and for the two different fitting strategies (treating VH(VV) as part of the signal or as a background).

Process	VH(VV) as signal		VH(VV) as background	
	cut-based	BDT-based	cut-based	BDT-based
WWW	2.54 (2.94)	3.33 (3.09)	0.96 (1.82)	1.63 (1.87)
WWZ	3.53 (3.62)	3.35 (4.09)	0.87 (2.17)	1.31 (2.15)
WZZ	1.55 (0.70)	1.71 (0.69)	1.68 (0.76)	1.71 (0.76)
ZZZ	0.00 (0.90)	0.00 (0.89)	0.00 (0.90)	0.00 (0.90)
VVV	5.02 (5.37)	5.67 (5.88)	2.34 (3.52)	2.90 (3.48)
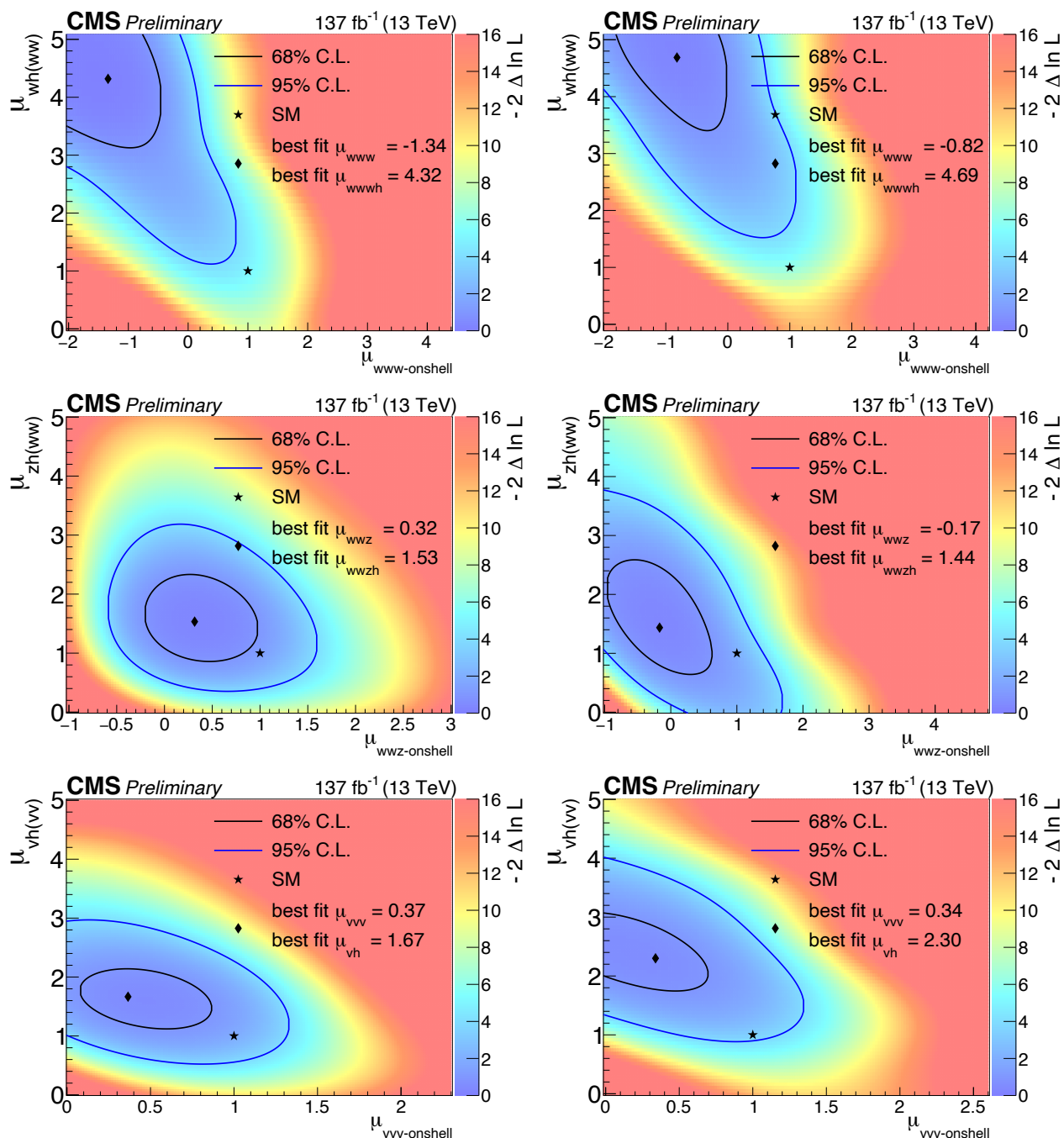
Figure 7.34: 2D likelihood contour plots ($-2\Delta \ln L$ where L is the likelihood) as a function of the VVV-onshell and VH(VV) signal strengths $\mu_{\text{VVV}-\text{onshell}}$ and $\mu_{\text{VH(VV)}}$. The plots in the left column (right column) are from the cut-based (BDT-based) method. The top plots are the likelihood contours versus the signal strength for the WWW-onshell and WH(WW) processes; the middle plots are for the WWZ-onshell and ZH(WW) processes; and the bottom plots are for the combined VVV-onshell and VH(VV) processes. The point with the best fit signal strengths $\mu_{\text{VVV}-\text{onshell}}$ and $\mu_{\text{VH(VV)}}$ is shown as a diamond in each of the plots together with a star marking the SM prediction at (1.0, 1.0). Each of the plots also shows the curves which correspond to the 68% and 95% C.L. contours.

7.8 Summary

In summary, we have observed the production of three massive vector bosons for the first time, with $137\,\mathrm{fb}^{-1}$ of proton-proton collision data collected by the CMS detector. The observed (expected) significance of the VVV production is 5.9 (5.7) standard deviation. We have also found evidence of WWW production and WWZ production individually for the first time, with the observed (expected) significance of 3.3 (3.4)σ and 3.1 (4.1)σ, respectively. The measured VVV cross section is $1.01^{+0.21}_{-0.20}(\mathrm{stat})^{+0.15}_{-0.12}(\mathrm{syst}) = 1.01^{+0.26}_{-0.23}$ pb, consistent with the SM prediction. The individually measured WWW, WWZ, WZZ production cross sections and an upper limit on the ZZZ production cross section are also reported.

7.9 Outlook

As the current results are still limited by statistics uncertainty, with more integrated luminosity to come in LHC Run 3 and at the HL-LHC, the precision of the triboson measurements reported in this thesis are expected to be further improved, including the possible observation of the individual VVV production channels. Making projections of the results obtainable with more data is not so straightforward, because several of the systematic uncertainties are not very easy to scale, as some of the background estimation-related uncertainties come from the statistical uncertainties in the data in the control regions. In addition, a major part of the uncertainties in the signals come from the simulation statistics which could be improved by using larger simulated samples.

Table 7.39 shows a conservative quick projection, in which all the background and signal systematic uncertainties are kept the same as their current values, and we simply scale the background and signal yields by the projected luminosity. We can see from this table that by the end of Run 3 (with a total of $300\,\mathrm{fb}^{-1}$), we expect to observe WWW and WWZ productions (with VH(VV) as signal), and we also expect to observe VVV-onshell production. By the end of HL-LHC ($3000\,\mathrm{fb}^{-1}$), we expect to see evidence for WZZ and ZZZ productions as well. With $3000\,\mathrm{fb}^{-1}$ data, the VVV cross section measurement precision is expected to reach the 10% level, where the uncertainty will be dominated by systematic errors (including an estimated 4% statistical uncertainty and a 9% systematic uncertainty). Figure 7.35 shows the likelihood ($-2\Delta \ln L$) scan over different signal strengths with the uncertainties broken down to the statistical and systematic uncertainties. As we can see in the figure, this search will be dominated by systematic uncertainty for WWW and WWZ production at the HL-LHC. As discussed above, with improved statistics in

the control region data, and more simulated events, the systematic uncertainty can be further improved relative to the results shown in Table 7.39, raising the prospect of establishing solid evidence (beyond 3σ) for WZZ and ZZZ by the end of the HL-LHC program.

Table 7.39: Expected sensitivity (significance $L(\sigma)$ and signal strength μ) of triboson measurements for different VVV productions with different values of integrated luminosity ($\mathcal{L}$).

$\mathcal{L}$(fb^{-1})	quantity	VVV	WWW	WWZ	WZZ	ZZZ
137	$L(\sigma)$: VH(VV) as signal	5.9	3.1	4.1	0.7	0.9
	μ: VH(VV) as signal	$1.00^{+0.25}_{-0.22}$	$1.00^{+0.40}_{-0.36}$	$1.00^{+0.36}_{-0.31}$	$1.00^{+1.67}_{-1.48}$	$1.00^{+2.88}_{-0.84}$
	$L(\sigma)$: VH(VV) as bkg	3.5	1.8	2.2	0.8	0.9
	μ: VH(VV) as bkg	$1.00^{+0.38}_{-0.34}$	$1.00^{+0.62}_{-0.56}$	$1.00^{+0.59}_{-0.51}$	$1.00^{+1.66}_{-1.34}$	$1.00^{+2.94}_{-0.84}$
300	$L(\sigma)$: VH(VV) as signal	7.8	4.0	5.6	1.0	1.3
	μ: VH(VV) as signal	$1.00^{+0.18}_{-0.17}$	$1.00^{+0.31}_{-0.28}$	$1.00^{+0.26}_{-0.23}$	$1.00^{+1.10}_{-0.99}$	$1.00^{+1.72}_{-0.87}$
	$L(\sigma)$: VH(VV) as bkg	4.8	2.4	3.0	1.1	1.3
	μ: VH(VV) as bkg	$1.00^{+0.28}_{-0.25}$	$1.00^{+0.47}_{-0.43}$	$1.00^{+0.42}_{-0.37}$	$1.00^{+1.08}_{-0.90}$	$1.00^{+1.74}_{-0.86}$
3000	$L(\sigma)$: VH(VV) as signal	11.3	6.1	7.9	2.5	3.0
	μ: VH(VV) as signal	$1.00^{+0.10}_{-0.10}$	$1.00^{+0.19}_{-0.18}$	$1.00^{+0.13}_{-0.11}$	$1.00^{+0.41}_{-0.36}$	$1.00^{+0.45}_{-0.36}$
	$L(\sigma)$: VH(VV) as bkg	8.2	3.7	5.9	2.8	3.0
	μ: VH(VV) as bkg	$1.00^{+0.15}_{-0.14}$	$1.00^{+0.28}_{-0.27}$	$1.00^{+0.21}_{-0.19}$	$1.00^{+0.40}_{-0.34}$	$1.00^{+0.45}_{-0.36}$

With the first establishment of VVV production, many interesting previously unexplored areas of physics, both within and beyond the SM, can be studied, using VVV production as a tool.

The first area that can be studied concerns the individual gauge-gauge and Higgs-gauge couplings. As stated earlier, these two couplings are still not determined separately, because the current event selection is not designed to isolate the VVV-onshell and VH(VV) processes. In order to discriminate between the two processes, several kinematic quantities can be utilized, among which the most significant features of VH(VV) are that the two vector bosons from Higgs boson decay tend to be softer than the 3$^{\text{rd}}$ vector boson, and the opening angle of the softest pair of vector bosons also tends to be smaller, compared to the VVV-onshell case.

As a result, variables such as the jet p_{T} and minimum jet-lepton opening angle in the same-sign final state, the $p_{\text{T}}^{\text{miss}}$ in the 3ℓ final state and the invariant mass of the two W lepton candidates in the four-lepton eμ final state, could be used to bin the events into different signal regions. Figure 7.36 shows the distributions of these variables as examples. Even without dedicated binning in such variables, once the accumulated

luminosity is doubled as projected after LHC Run 3 by 2024, the significances of the individual VH(VV) and VVV-onshell signals from a 2D simultaneous fit are each expected to be around 5σ, and this analysis is expected to yield a simultaneous observation of the two processes.

In addition to the measurement of the SM-predicted gauge-gauge and Higgs-gauge couplings, VVV production can also be used as a tool to search for many BSM physics, both directly and indirectly. The four-boson contact interaction in Figure 7.2 could be a low-energy approximation of a more precise model with a mediator, and new physics beyond the SM could thus change this coupling constant. A search for anomalous quartic gauge-boson couplings (aQGC) using WWW production (same-sign two-lepton and three-lepton final states) performed by CMS with 2016 data [101] has put promising limits on several dimension-8 anomalous coupling term coefficients [135]. Similar searches could be performed with the full Run 2 data set and with more hadronic or semi-hadronic final states included, as such searches require a sufficient statistical sample of highly boosted vector bosons. In addition to such indirect BSM searches for aQGC, direct searches for heavy mediators (such as the radion [136]) modifying the four-boson vertex can also be performed.
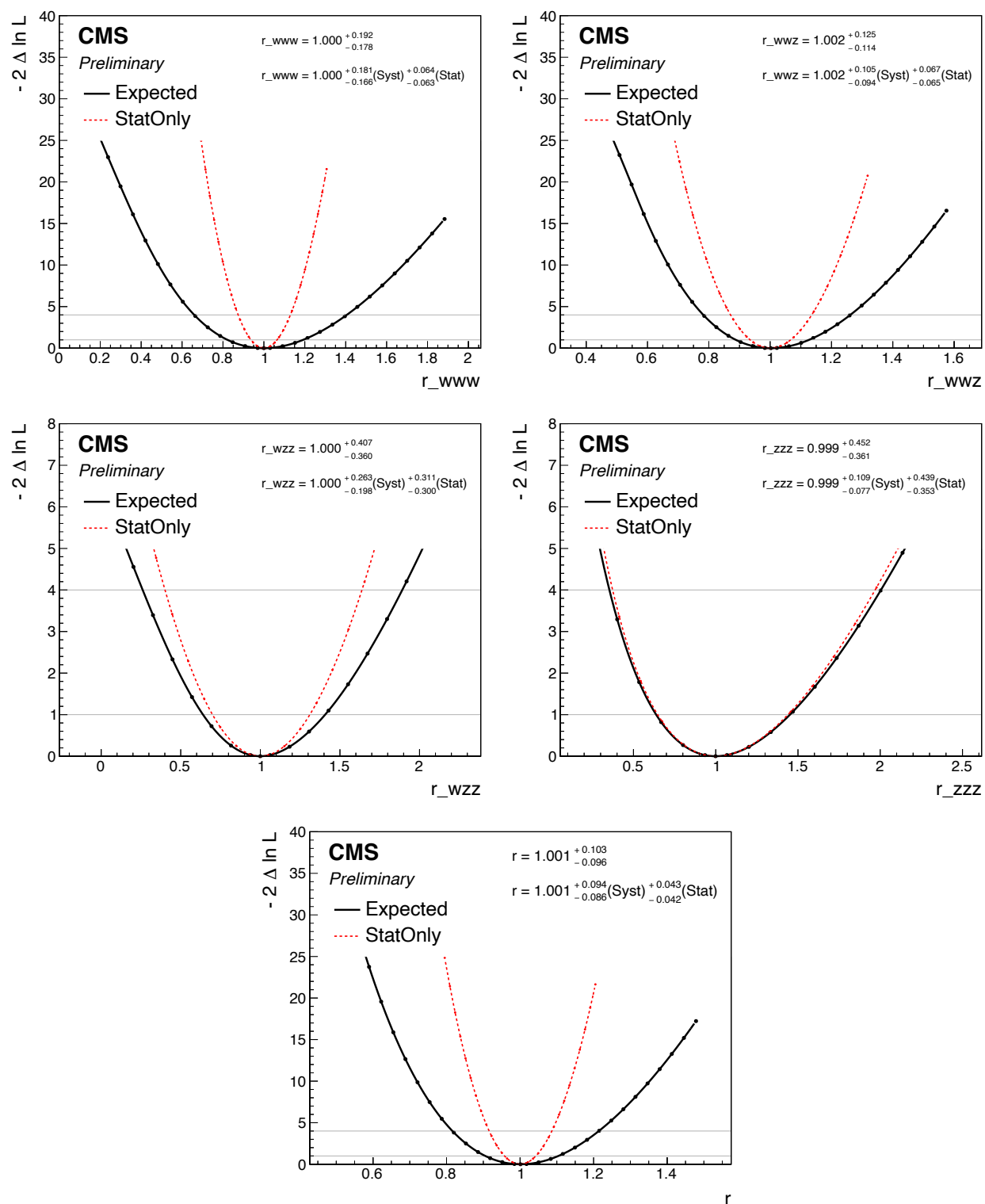
Figure 7.35: $-2\Delta \ln L$ as a function of the signal strength for different signals: WWW (top-left), WWZ (top-right), WZZ (center-left), ZZZ (center-right), and the combined VVV signal (bottom). The plots are projections of the results expected with $3000\,\text{fb}^{-1}$ of data at the HL-LHC. VH(VV) is included as part of the signal in all plots.
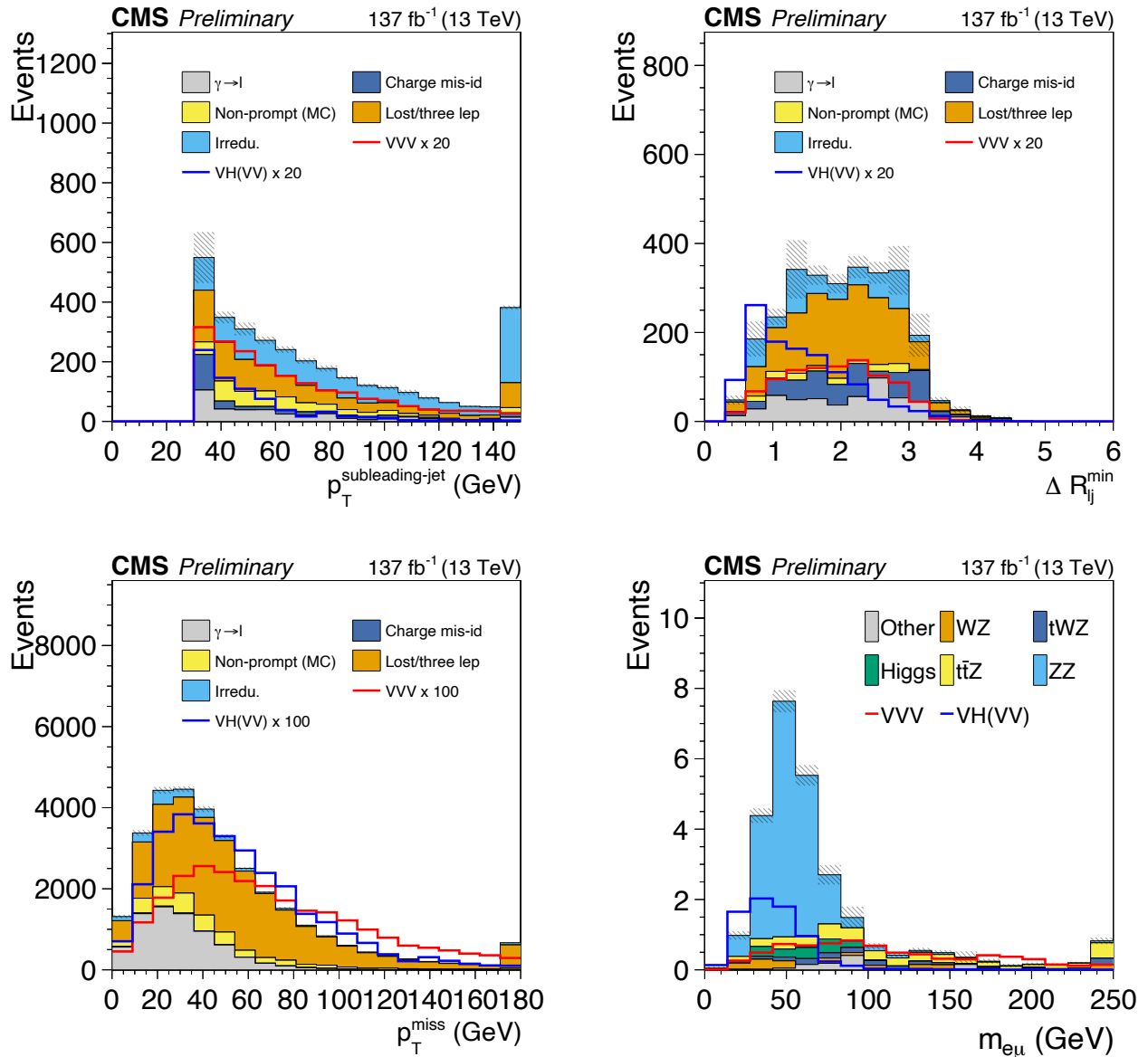
Figure 7.36: Distributions of some useful variables to discriminate the VVV-onshell signal from the VH(VV) signal - top left: subleading jet p_T in the SS + 2 jets preselection; top right: $\Delta R_{\ell j}^{min}$ in SS + 1 jet preselection; bottom left: p_T^{miss} in the 3ℓ preselection region; bottom right: invariant mass of the two W lepton candidates in the four-lepton eμ final state.

Part IV

Search for long-lived particles using delayed photons in CMS at the LHC

Chapter 8

SEARCH FOR LONG-LIVED PARTICLES USING DELAYED PHOTONS WITH 2016 DATA IN CMS

There are many SM unstable particles which are long-lived and have proper decay lengths between a few millimeters and a few meters. The decay products of such long-lived particles can be detected by CMS, but exhibit different signatures from prompt particle decays. Such long-lived particles (LLP) are also predicted in many models of physics beyond the SM, such as supersymmetry (SUSY) with gauge-mediated SUSY breaking (GMSB) [16–25] or hidden valley models [137].

In this chapter, we present a search for LLPs that decay to a photon and a weakly interacting particle, using a benchmark scenario of GMSB, commonly referred to as the "Snowmass Points and Slopes 8" (SPS8) [138]. In this scenario, the LLPs we are interested in are the next-to-lightest SUSY particle (NLSP), the neutralino ($\widetilde{\chi}_1^0$), which is produced via cascade decays of pair-produced squarks and gluinos and decays to the lightest SUSY particle (LSP), the gravitino ($\widetilde{G}$), as shown in Figure 8.1. The mass of the neutralino is linearly related to the SUSY scale, Λ, and is also affected by other parameters such as the messenger mass. In the SPS8 model, these parameters are fixed and Λ is a free parameter, and the value of Λ also determines the primary production mode and decay rate of the NLSP. The dominant decay mode of the NLSP in SPS8 is to a photon and gravitino, resulting in a final state with one or two photons and missing transverse momentum (p_T^{miss}).

Previously, a search was performed by CMS using p_T^{miss} and the photon arrival time at the ECAL in single photon plus p_T^{miss} events with $4.9\,\text{fb}^{-1}$ of data at $\sqrt{s} = 7\,\text{TeV}$ [139], which excluded such a GMSB scenario for neutralino masses below 220 GeV and proper decay length $c\tau$ up to 6 m. ATLAS also reported a similar search with the same GMSB model by using the photon arrival time at their ECAL and the longitudinal impact parameter of the projected photon trajectory Δz_γ relative to the center of the interaction region, in diphoton events with $4.8\,\text{fb}^{-1}$ data at $\sqrt{s} = 7\,\text{TeV}$ [140]. The search excluded such GMSB models for neutralino masses up to 240 GeV and $c\tau$ up to 15 m. Followup searches with 8 TeV data have been performed by both CMS [141] and ATLAS [142], which excluded such GMSB models for a neutralino mass up to 300 GeV and 430 GeV, and $c\tau$ up to 25 m and 30
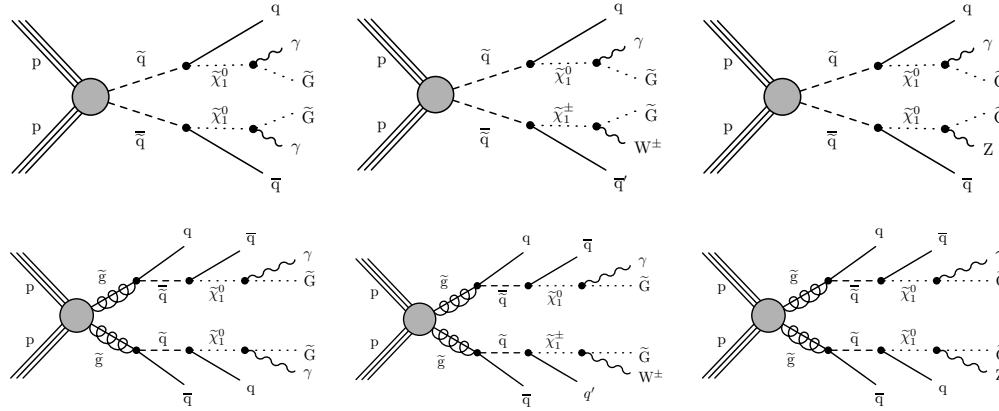
Figure 8.1: Example Feynman diagrams for SUSY processes that result in diphoton (left) and single photon (middle and right) final states via squark (upper) and gluino (lower) pair-production at the LHC.

m, respectively. Figure 8.2 summarizes the exclusion regions from all the previous searches by CMS and ATLAS. As we see from the previous results obtained with Run 1 data shown in the figure, the sensitivity for small $c\tau$ signal models is greatly improved by requiring two photons in the events, as this greatly reduces the prompt photon SM backgrounds.

In this search, we use CMS data taken in 2016 with a diphoton trigger, and select events with at least two photons. A dedicated identification is applied to the leading photon to ensure high efficiency for time delayed (or vertex displaced) photons, and to suppress prompt photons by cutting on the electromagnetic shower shapes. The arrival time of the leading photon in the ECAL is then used as a shape variable in the fit to extract the signal yield, together with another variable, the missing transverse momentum (p_T^{miss}). The results are then combined with a similar search that uses CMS data taken in 2017 with a single photon trigger, that has better sensitivity to signals with large $c\tau$.

The remaining sections of this chapter are organized as follows: Section 8.1 discusses the data sample and the triggers used in this analysis. The object reconstruction and identification, especially the identification details for out-of-time photons, are given in Section 8.2. Section 8.3 is dedicated to the measurement of the CMS ECAL timing resolution, and the necessary corrections to match MC events with data. The event selections are given in Section 8.4. The background estimation as well as the systematic uncertainties are presented in Sections 8.5 and 8.6, and the final results are given in Section 8.7.
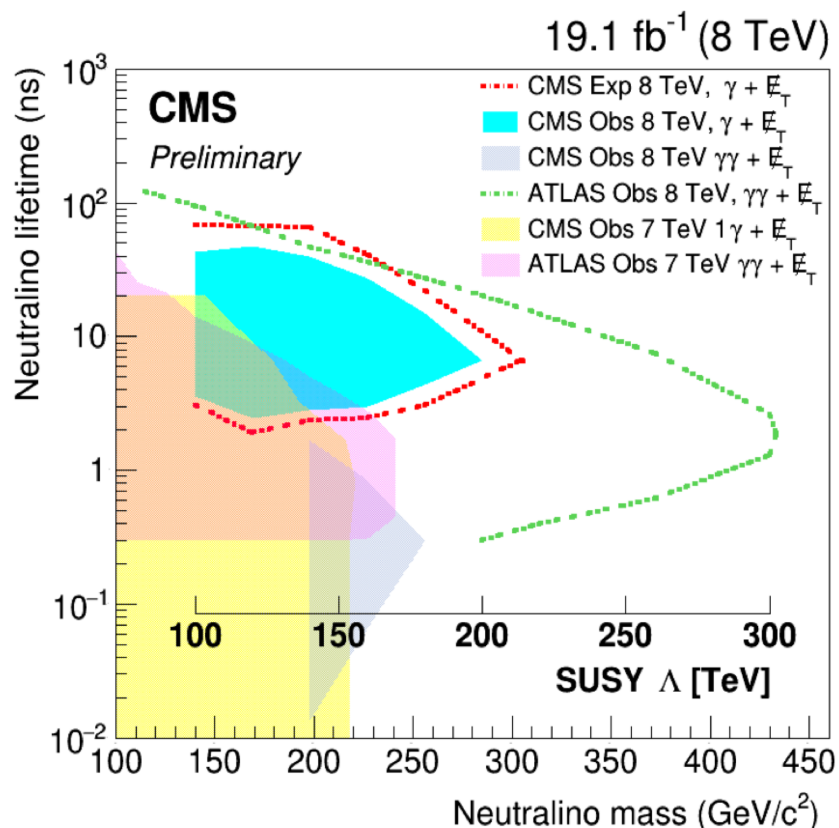
Figure 8.2: Summary of exclusion results in the plane of the proper lifetime of a neutralino NLSP in a GMSB model versus its mass, from searches with LHC Run 1 data by both CMS and ATLAS [141].

8.1 Event samples

The data used in this search corresponds to a total integrated luminosity of $35.9\,\text{fb}^{-1}$ of proton-proton collisions collected by the CMS detector in 2016. The result of a combined search with 2016 and 2017 data will be presented in the next chapter.

The GMSB SPS8 signal model samples are produced by first generating a tabulated form of the particle spectra in a set of SUSY Les Houches Accord (SLHA) files [143]. The purpose of the SLHA project is to define a set of conventions for supersymmetric extensions of the Standard Model together with an accord specifying generic file structures for 1) supersymmetric model specifications and input parameters, 2) EW scale supersymmetric mass and coupling spectra, and 3) decay tables, to provide a universal interface between spectrum calculation programs, decay packages, and high energy physics event generators using the ISASUGRA toolkit [144] in the ISAJET v7.87 generator package [145]. ISASUGRA is a subprogram of ISAJET that calculates the mass-spectrum and decay-channels for a number of SUSY models, including mSUGRA. A grid of samples with different SUSY scale Λ and neutralino lifetime $c\tau$

values is used to construct a set of SLHA tables, where the parameters used to generate the tables follow the SPS8 recommendations. The resulting SLHA files are then passed to PYTHIA v8.212 [121] to generate the benchmark signal models for each of the designated Λ and $c\tau$ values. Table 8.1 lists all of the GMSB signal samples used in this search, along with the masses of the gluino, neutralino, and gravitino, and the cross sections for different Λ and $c\tau$ values.

Table 8.1: Table of the generated GMSB SPS8 signal models, and the corresponding mass points and cross sections. For each Λ point, a grid of signal models with $c\tau = 0.1, 0.5, 1, 2, 4, 6, 8, 10, 12,$ and 100 m is generated. The mass of the gravitino $M_{\widetilde{G}}$ is proportional to the square root of $c\tau$ (as can be seen in Eq. 2.10 and 2.11, we have $M_{\widetilde{G}} \propto \langle F \rangle$, and $c\tau \propto \langle F \rangle^2$); and the mass of gluino and neutralino are determined by Λ.

Λ	$M_{\widetilde{g}}$	$M_{\widetilde{\chi}_1^0}$	$M_{\widetilde{G}}$ (10^{-7} GeV) for different $c\tau$	cross section
(TeV)	(GeV)	(GeV)	0.1, 0.5, 1, 2, 4, 6, 8, 10, 12, 100 m	(fb)
100	838	139	1.5, 3.4, 4.9, 6.9, 9.7, 12, 14, 15, 17, 49	2160 ± 20
150	1207	212	4.7, 10, 15, 21, 30, 36, 42, 47, 51, 148	228 ± 2
200	1565	285	10, 22, 32, 45, 64, 78, 90, 100, 110, 318	44.5 ± 0.3
250	1915	358	18, 40, 57, 81, 114, 140, 162, 180, 198, 571	12.6 ± 0.1
300	2260	430	29, 65, 92, 130, 183, 225, 259, 290, 318, 917	4.45 ± 0.03
350	2599	503	43, 96, 136, 193, 273, 334, 386, 430, 473, 1364	1.78 ± 0.01
400	2935	576	61, 136, 192, 271, 384, 470, 543, 610, 665, 1920	0.778 ± 0.005

The MADGRAPH5_AMC@NLO v2.2.2 generator [113] is used to generate the standard model background events, including top quark and $t\bar{t}$ production at next-to-leading order (NLO) in quantum chromodynamics (QCD), and QCD multijet, γ+jets, W+jets, and Z+jets productions at leading order (LO). SHERPA v2.2.4 [146, 147] is used to generate diphoton events, which include Born processes with up to three additional jets, as well as box-diagram processes at LO precision. The NNPDF3.0 [119, 120] parton distribution functions (PDFs) are used in the simulated samples. All the generated MC samples are processed with GEANT4 [85] to simulate the interactions of particles with the CMS detector. Pileup interactions are also simulated and included in each MC event.

This search uses an unprescaled diphoton trigger, which requires two isolated photon objects with transverse energy greater than 42 and 25 GeV for the leading and subleading photons. Additional requirements, such as cuts on photon shower shape variables (R_9, $\sigma_{i\eta i\eta}$) and ratio of energies in HCAL and ECAL (H/E), are also applied on both photon objects in the trigger. The details of these cuts are discussed in Section 8.2, together with the offline identification criteria for photons.

8.2 Object reconstruction and identification

Jet, p_T^{miss}, and photon objects are used in this search. A particle-flow (PF) algorithm [71] is used to reconstruct all the physics objects in CMS, with details discussed in Section 4.6.

Jets are reconstructed from PF candidates, clustered with the anti-k_T algorithm with a cone size of 0.4 [74–76]. Jets are required to pass the identification criteria summarized in Table 7.5 and to have transverse momentum $p_T > 30$ GeV and $|\eta| < 3.0$. To avoid double counting, jets within a $\Delta R = 0.3$ cone centered on either of the two leading photon objects are not used in the selection.

The missing transverse momentum p_T^{miss} is defined as the negative sum of the transverse momentum of all PF particles reconstructed in an event. To avoid anomalous high-p_T^{miss} events from detector noise, all events have a series of filters applied, which are designed to identify noise from the ECAL, HCAL, and Muon systems. It is also noteworthy that we realized that there is a particular issue when applying the noise filters to our GMSB signal models, in the case where one of the signal photons is produced outside of the ECAL acceptance but within the acceptance of the other CMS detectors such as the HCAL or Muon system, since the photon can shower and deposit energy in those systems. Such events yield an isolated energy object in a single detector, such as only hits in the HCAL but no hit in the ECAL. In such cases, the designed HCAL (or Muon) noise filter removes those events because they have the same topological signature as a real HCAL noise event. This is a small effect for small-lifetime signal models, but for large-lifetime signals up to 10% of the signal events are removed by such filters.

Photons are reconstructed from the energy deposits in the ECAL, which covers a fiducial region of $|\eta| < 1.4442$ in the barrel (EB) and $1.566 < |\eta| < 2.5$ in the endcaps (EE). The excluded $1.4442 < |\eta| < 1.566$ region includes two parts, with $1.4442 < |\eta| < 1.479$ covering the last two crystals at each end of the barrel, and $1.479 < |\eta| < 1.566$ covering the outer 5 rings of crystals (the first outer ring of trigger towers) in the endcap. As the photon reconstruction is not optimal in this transition region between the barrel and endcap, photons in this region are not considered.

The details of photon energy reconstruction have been discussed in Section 4.3. To distinguish and select good photons from other objects such as jets and electrons, a set of requirements are imposed on variables related to the photon showers, isolation, and signals in other detectors such as the HCAL and tracker. The list of variables

considered in the photon identification procedure are:

- **R_9**: the energy sum in the 3×3 set of crystals centered on the most energetic crystal in a supercluster (SC), divided by the total energy of the SC. The R_9 of good and unconverted photons is very close to 1.0; and the R_9 of objects such as jets is much smaller. A cut on $R_9 > 0.9$ is applied to all photons in this search.

- **$\sigma_{i\eta i\eta}$**: the width of the seed basic cluster of the photon in the η direction, computed with weighted crystal η indices, where the weights are related to the energy deposited in each crystal, as defined in Eq. 8.1. The distribution of $\sigma_{i\eta i\eta}$ for photons peaks below 0.01 and the $\sigma_{i\eta i\eta}$ for jets is usually larger than 0.01.

- **H/E**: the energy deposited in the HCAL in a $\Delta R = 0.15$ cone around the photon direction, divided by the SC energy measured in the ECAL. For photons, there is generally no significant energy deposited in the cone in the HCAL, and so H/E is in general very small.

- **ChargedHadPFIso**: the scalar p_T sum of all charged hadron PF candidates in a $\Delta R = 0.30$ cone around the photon direction.

- **TrackIso**: the scalar p_T sum of all tracks in a $\Delta R = 0.30$ cone around the photon direction.

- **NeutralHadPFIso**: the scalar p_T sum of all neutral hadron PF candidates in a $\Delta R = 0.30$ cone around the photon direction.

- **HCALIso**: the scalar sum of transverse energy in HCAL deposits in a $\Delta R = 0.30$ cone around the photon direction.

- **PhotonPFIso**: the scalar p_T sum of all photon PF candidates in a $\Delta R = 0.30$ cone around the photon direction.

- **ECALIso**: the scalar sum of transverse energy in ECAL deposits in a $\Delta R = 0.30$ cone around the photon direction.

- **S_{major}**: the weighted energy along the semi-major axis of the elliptical photon shower, as defined by the eigenvalues in Eq. 8.1, where $S_{\phi\phi}$, $S_{\eta\eta}$, and $S_{\eta\phi}$ are the second central moments of the spatial distribution of the energy deposits

in the ECAL in η-ϕ coordinates, and the logarithmic weight w_i in the equation is the same as the weight used to calculate the cluster position [148].

- S_{minor}: the weighted energy along the semi-minor axis of the elliptical photon shower, as defined by the eigenvalues in Eq. 8.1.

$$S_{\frac{\text{major}}{\text{minor}}} = \frac{S_{\phi\phi} + S_{\eta\eta} \pm \sqrt{(S_{\phi\phi} - S_{\eta\eta})^2 + 4S_{\eta\phi}^2}}{2} \tag{8.1}$$

$$\text{where } S_{\eta\phi} = \sum_{i=1}^{N} w_i(\eta_i - <\eta>)(\phi_i - <\phi>)$$

$$S_{\eta\eta} = \sum_{i=1}^{N} w_i(\eta_i - <\eta>)(\eta_i - <\eta>)$$

$$S_{\phi\phi} = \sum_{i=1}^{N} w_i(\phi_i - <\phi>)(\phi_i - <\phi>)$$

$$<\eta> = \frac{\sum_i w_i \eta_i}{\sum_i w_i}, \quad <\phi> = \frac{\sum_i w_i \phi_i}{\sum_i w_i}$$

$$w_i = \max\left[4.2 + \ln(\frac{E_i}{E_{\text{cluster}}}); 0\right].$$

The distributions of the above identification variables for signal photons and fake photons (from jets) are shown in Figures 8.3 and 8.4. For signal photons, two types of photon collections are shown: the GED photon and OOT photon. The GED photon refers to the "general event description" photons, which only collects in-time photons with a time delay (with respect to a prompt photon from the primary vertex) smaller than 3 ns. To cover our delayed photon signals, a second photon collection, called out-of-time (OOT) photons is designed for this search, which collects photons that are not saved in the GED photon collection.

As shown in these figures, the shapes of the photon ID variable distributions are very different for GED and for OOT photons. This is especially true for the shower shape variables, where the OOT (displaced) photons have a non-zero incident angle with respect to the long axis of the crystal, and thus the EM shower is non-circular and has a elliptical shape. In addition, when running the reconstruction for OOT photons, the full particle flow reconstruction is not run on those photons and therefore the HCAL-related information for those photons is not available, including the H/E and HCALIso values. Due to the differences between GED and OOT photons,

different identification criteria are applied to the GED and OOT photons, and these requirements are optimized separately in order to maximize the signal efficiencies for in-time and out-of-time signal photons.

Table 8.2 summarizes the full set of photon identification cuts for both GED and OOT photons used in this search. To compensate for the effects of the energy deposited due to pileup, all the isolation variables are corrected event by event, by using the effective pileup transverse momentum density ρ (median of the energy density per unit area in the rapidity-azimuth plane) times an effective area (A_{eff}, the area of the isolation region), i.e. the isolation variables are subtracted by the pileup contribution by ρA_{eff} [149]. These corrections are measured separately for GED and OOT photons and have different values for different pseudorapidities of the photons. Table 8.3 summarizes the corresponding effective areas used in this set of photon identification criteria, which are obtained by fitting the isolation dependency on the ρ in with γ+jet events. The signal efficiency of the GED (OOT) photon ID has been measured to be about 65% (80%), while the efficiency for picking up fake GED (OOT) photons from QCD multijet events has been measured to be about 5% (10%). Figures 8.5, 8.6, 8.7, and 8.8 show the corresponding efficiencies as a function of the photon p_{T}, η, and arrival time, from which we can see that the efficiencies are relatively flat for different values of those variables in all cases.

Table 8.2: Photon identification cuts used in the delayed photon search. All the isolation variables are ρ corrected as described in the text.

Variable	GED photon ID	OOT photon ID
H/E	< 0.0269	—
$\sigma_{i\eta i\eta}$	< 0.017	< 0.017
ChargedHadPFIso	<0.202	
TrackIso	—	$< 2.0 + 0.005996 p_{\text{T}}^{\gamma}$
NeutralHadPFIso	$< 0.264 + 0.0148 p_{\text{T}}^{\gamma} + 1.7 \times 10^{-5} p_{\text{T}}^{\gamma 2}$	
PhotonPFIso	$< 2.362 + 0.0047 p_{\text{T}}^{\gamma}$	
ECALIso	—	$< 3.0 + 0.007132 p_{\text{T}}^{\gamma}$
S_{major}	—	< 1.3

As the GED and OOT photons are reconstructed independently with the same reconstruction algorithm, but just starting from a different seed crystal (GED photon reconstruction only starts with a seed which has a time smaller than 3 ns, while OOT photon reconstruction starts with a seed which has a time larger than 3 ns), in most cases the GED and OOT reconstruction will not select the same photon. But in

Table 8.3: Effective areas used to correct the isolation variables in the GED and OOT photon ID procedures.

η range	GED photon ID			OOT photon ID		
	ChargedHadPFIso	NeutralHadPFIso	PhotonPFIso	TrackIso	NeutralHadPFIso	ECALIso
$\lvert\eta\rvert < 1.0$	0.0360	0.0597	0.1210	—	—	—
$1.0 < \lvert\eta\rvert < 1.479$	0.0377	0.0807	0.1107	—	—	—
$\lvert\eta\rvert < 0.8$	—	—	—	0.2304	0.0488	0.1555
$0.8 < \lvert\eta\rvert < 1.479$	—	—	—	0.1969	0.0523	0.1362

rare cases (less than 5%), the reconstructed GED and OOT photon clusters might partially overlap with each other, i.e. there may be some shared crystals. In such cases, either one or both of the two clusters is incomplete and is only part of the full photon cluster. To avoid double counting, a cluster-overlap removal procedure is applied to the two photon collections. To do this, we put all the GED and OOT photons together, and compare every pair of photons with one from the GED collection and one from the OOT collection. If the ΔR between the two photons is smaller than 0.3, then we consider that the photons are overlapped and we only keep the photon with the larger p_T, and remove the photon with smaller p_T from the list of photon objects, before continuing the cluster-overlap procedure.

As mentioned above, the OOT photons are not included in the PF candidate list because the full particle flow reconstruction is not run on OOT photons. As a result, if we decide to keep an OOT photon in the above cluster-overlap cleaning procedure, the $\vec{p}_\mathrm{T}$ of the OOT photon is subtracted from the $\vec{p}_\mathrm{T}^{\,\mathrm{miss}}$. For GED photons, if a GED photon is removed during the cleaning procedure, the $\vec{p}_\mathrm{T}$ of the GED photon is added back into $\vec{p}_\mathrm{T}^{\,\mathrm{miss}}$ since it was included in the $\vec{p}_\mathrm{T}^{\,\mathrm{miss}}$ calculation by default. When we decide to remove an OOT photon or keep a GED photon in the overlap cleaning, nothing is done to the $\vec{p}_\mathrm{T}^{\,\mathrm{miss}}$. After this correction for $p_\mathrm{T}^{\mathrm{miss}}$, the resolution of the $p_\mathrm{T}^{\mathrm{miss}}$ (measured by the standard deviation of the distribution of the difference between the reconstructed and the generated $p_\mathrm{T}^{\mathrm{miss}}$ in MC) is improved by about 20%.
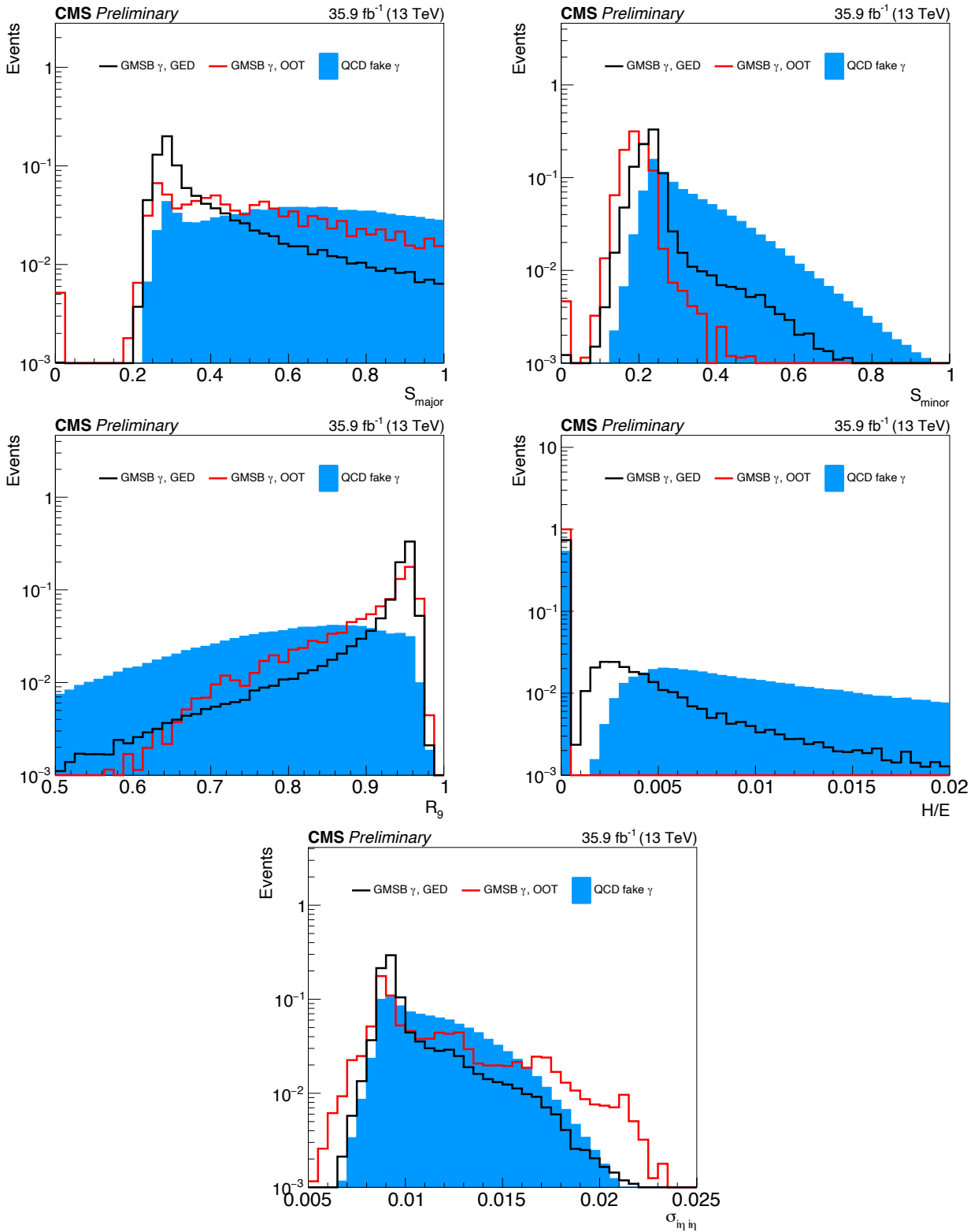
Figure 8.3: Distributions of some photon identification variables for GMSB signal photons (GED and OOT photons) from a representative point (Λ = 250 TeV and $c\tau$ = 2 m) and fake photons in QCD multijet MC events. Top row from left to right: S_{major}, S_{minor}; center row from left to right: R_9, H/E; bottom plot is $\sigma_{i\eta i\eta}$.
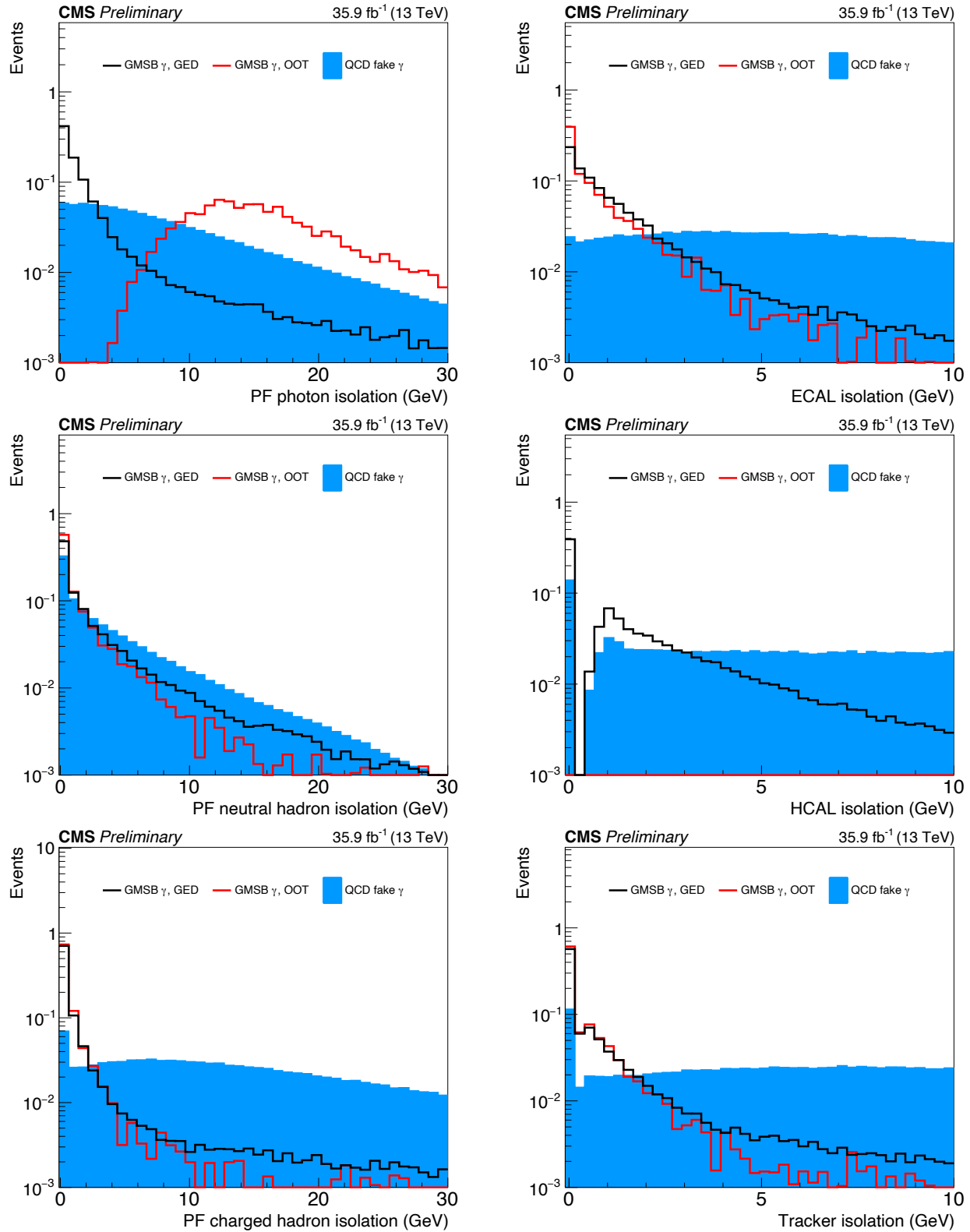
Figure 8.4: Distributions of some photon identification variables for GMSB signal photons (GED and OOT photons) from a representative point ($\Lambda = 250$ TeV and $c\tau = 2$ m) and fake photons in QCD multijet MC events. Left column from top to bottom: PhotonPFIso, NeutralHadPFIso, ChargedHadPFIso; right column from top to bottom: ECALIso, HCALIso, TrackIso.
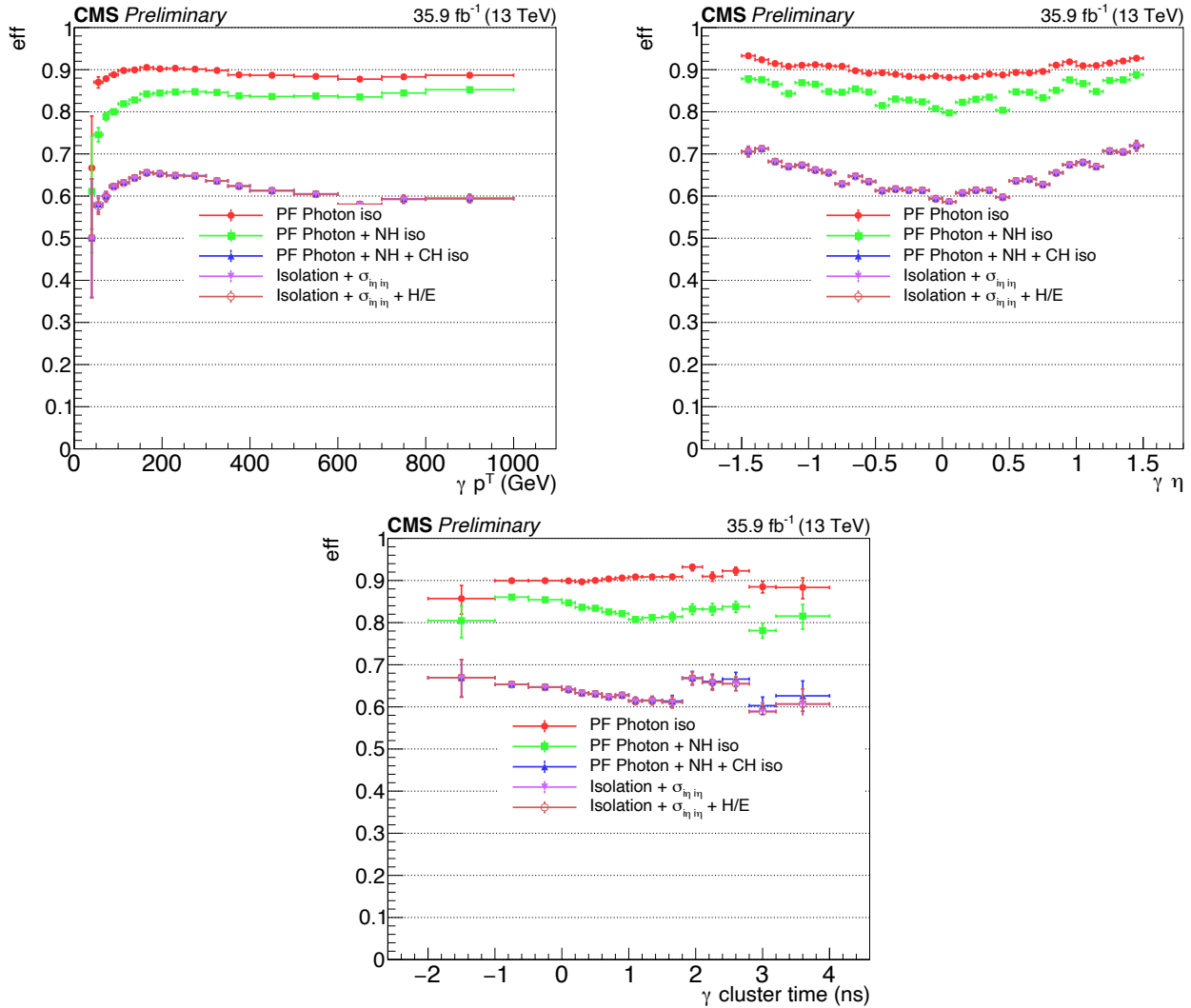
Figure 8.5: The GED photon ID signal efficiency as a function of the photon p_T (top-left), η (top-right), and arrival time (bottom). The efficiencies are measured using the GMSB $\Lambda =$ 200 TeV signal samples. The denominator is the number of reconstructed (RECO) photons which are matched to generated (GEN) photons, where a match requires that the RECO and GEN photon clusters have $\Delta R < 0.3$, and $\Delta p_\mathrm{T}/p_\mathrm{T}^\mathrm{GEN} < 0.3$. The numerator is the number of such matched photons that pass the corresponding GED photon ID selection criteria. The efficiency of the cut-flow in the GED photon ID cut sets are shown in the plots, where the last one (Isolation + $\sigma_{i\eta i\eta}$ + H/E) is the whole set of GED photon ID cuts.
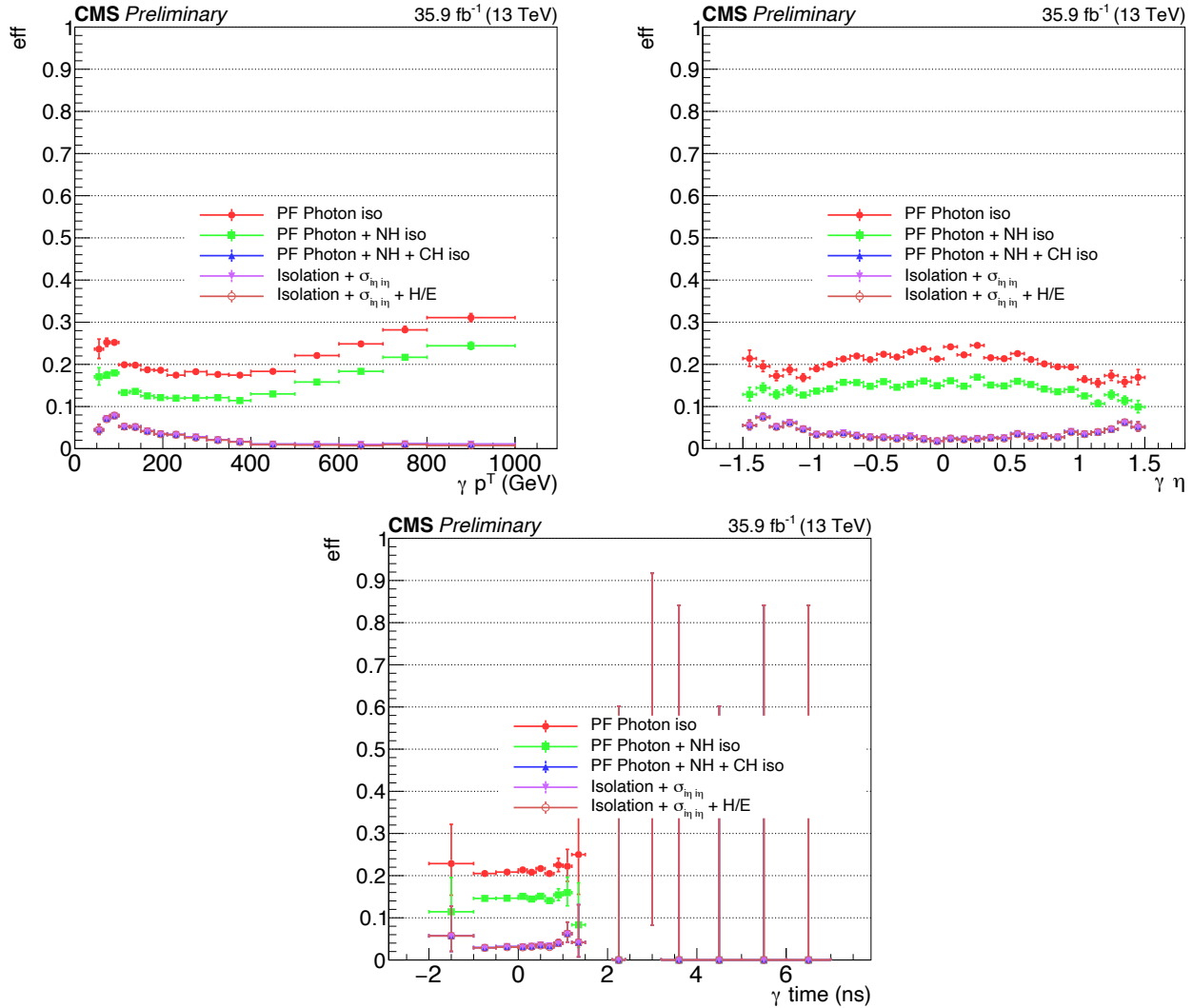
Figure 8.6: GED photon ID background efficiency as a function of the photon p_T (top-left), η (top-right), and arrival time (bottom). The efficiencies are measured using QCD MC signal samples. The denominator is the number of reconstructed (RECO) photons which have $p_T > 70$ GeV. The numerator is the number of such photons that pass the corresponding GED photon ID selection. The efficiency of the cut-flow in the GED photon ID cut sets are shown in the plots, where the last one (Isolation + $\sigma_{i\eta i\eta}$ + H/E) is the whole set of GED photon ID cuts.
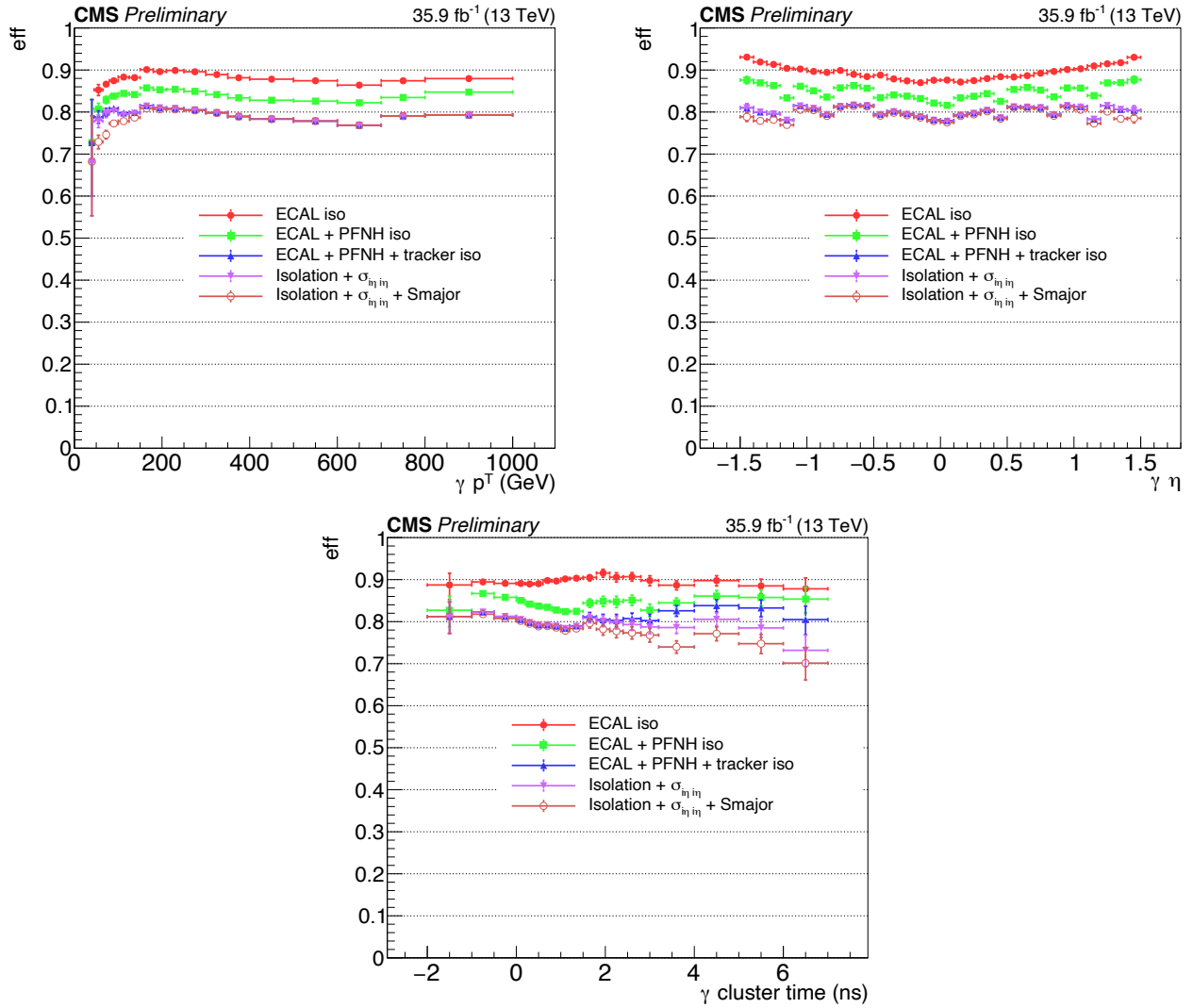
Figure 8.7: OOT photon ID signal efficiency as a function of the photon p_T (top-left), η (top-right), and arrival time (bottom). The efficiencies are measured using the GMSB Λ = 200 TeV signal samples. The denominator is the number of reconstructed (RECO) photons which are matched to generated (GEN) photons, where we require that the RECO and GEN photon clusters are matched within $\Delta R < 0.3$, and $\Delta p_T / p_T^{GEN} < 0.3$. The numerator is the number of such matched photons that pass the corresponding OOT photon ID selection. The efficiency of the cut-flow in the OOT photon ID cut sets are shown in the plots, where the last one (Isolation + $\sigma_{i\eta i\eta}$ + S_{major}) is the whole set of OOT photon ID cuts.
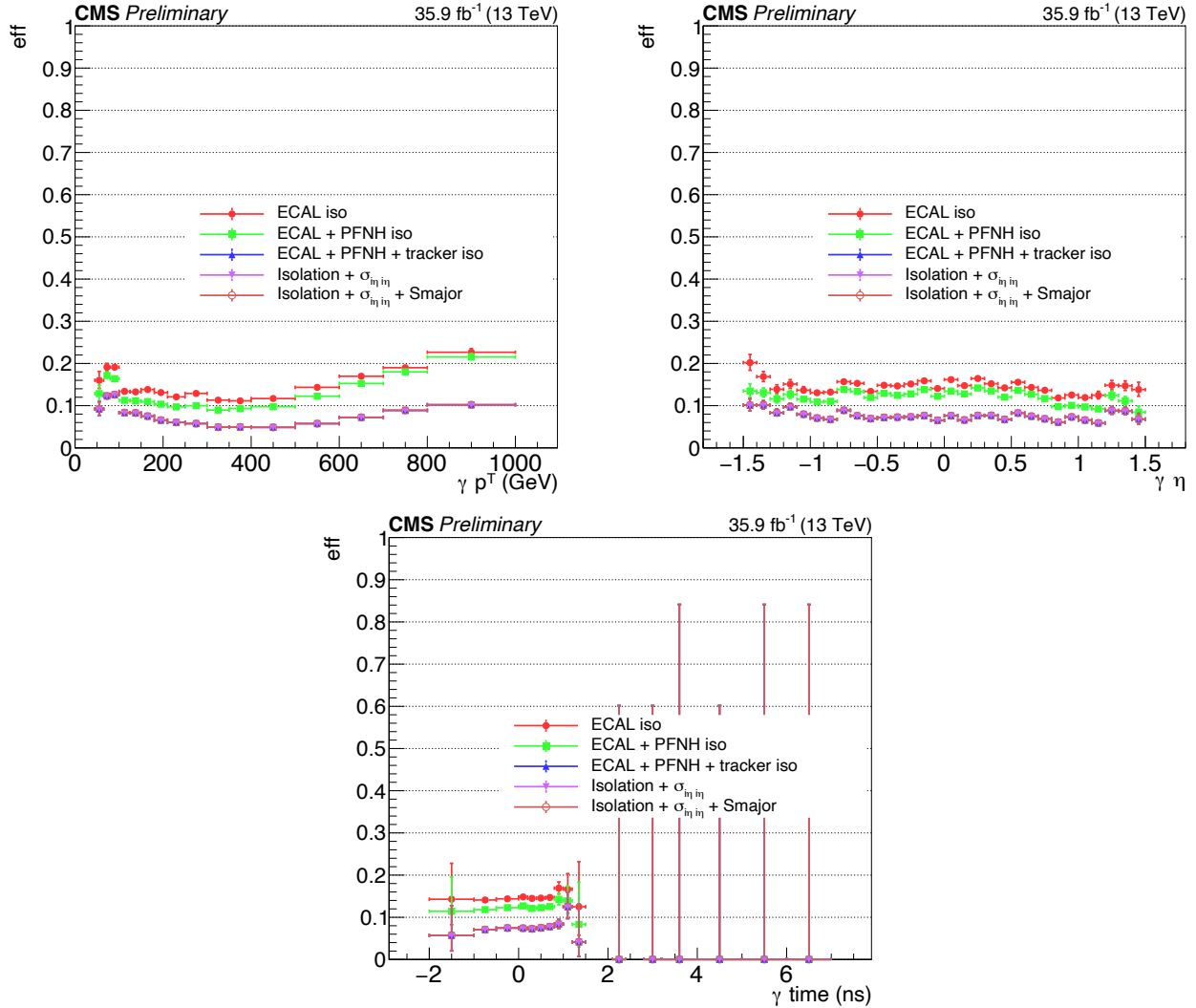
Figure 8.8: OOT photon ID background efficiency as a function of the photon p_T (top-left), η (top-right), and arrival time (bottom). The efficiencies are measured using QCD MC signal samples. The denominator is the number of reconstructed (RECO) photons which have $p_T > 70$ GeV. The numerator is the number of such photons that pass the corresponding OOT photon ID selection. The efficiency of the cut-flow in the OOT photon ID cut sets are shown in the plots, where the last one (Isolation + $\sigma_{i\eta i\eta}$ + S_{major}) is the whole set of OOT photon ID cuts.

8.3 ECAL timing measurement

This search uses the arrival time of the photon at the ECAL as one of the key discriminants to extract delayed photon signals from the SM backgrounds.

As mentioned in Section 4.3, each ECAL crystal's signal output pulse is digitized into 10 consecutive samples spaced 25 ns apart. Our task is to estimate the time at which the pulse reaches its maximum amplitude, t_{max}. Since the amplitude of each of the 10 samples depends on the distance from that sample's time to t_{max} as well as the pulse maximum amplitude, by looking at the ratio of that sample's amplitude and the amplitude of its neighboring samples, we can get an estimate of t_{max}, which we denote as $t_{max,k}$ (where the index k denotes the sample bin index in the pulse, which ranges from 0 to 9). And an uncertainty-weighted average of those 10 estimations, is defined as the final estimate of the timestamp of that crystal's signal pulse, which we denote as t_{ECAL}^i (index "i" is the crystal index). A detailed description of the ECAL timestamp reconstruction, as well as how the $t_{max,k}$ is estimated, can be found in [150].

The timestamp uncertainty for crystal i is parameterized as:

$$\sigma_i^2 = \left(\frac{N}{A_i/\sigma_{N_i}} \right)^2 + C^2, \tag{8.2}$$

where A_i is the amplitude of the signal detected by crystal i, σ_{N_i} is the pedestal noise for crystal i, and N and C are constants which are determined from a dedicated fit to the data. As shown in Equation 8.2, there are two main sources of uncertainty in the ECAL timestamp measurement, the noise term and the constant term, which are added in quadrature. The noise term is proportional to the noise amplitude and inversely proportional to the signal amplitude of the channel. The constant term includes the systematic uncertainty in the above time estimation method and other effects such as the variation of the depth of the EM shower starting point in the crystal.

To measure the values of the constants N and C in the above formula, we compare the time difference (Δt) of two neighboring crystals that belong to the same EM shower and have similar energy deposits, specifically by requiring that the ratio of the two energies be between 0.8 and 1.2. The EM shower is selected from photon objects with very basic identification requirements to avoid showers from jets. In addition, to avoid noisy channels and gain switch effects in the readout, we require the energy deposited in each crystal to be between 1 and 120 GeV. We then fit the

Δt distribution with a Gaussian function, and take the fitted $\sigma(\Delta t)$ as the resolution. Different fits are performed in bins of the effective amplitude/noise ratio in the pair of crystals (crystal 1 and 2), defined as:

$$A_{\text{eff}}/\sigma_{\text{N}} = \frac{(A_1/\sigma_{\text{N}_1})(A_2/\sigma_{\text{N}_2})}{\sqrt{(A_1/\sigma_{\text{N}_1})^2 + (A_2/\sigma_{\text{N}_2})^2}}. \tag{8.3}$$

The $\sigma(\Delta t)$ is then the sum (in quadrature) of uncertainties of the two timestamps from the two crystals parameterized in Eq. 8.2, and therefore the $\sigma(\Delta t)$ can be parameterized as the following:

$$\sigma^2(\Delta t) = \left(\frac{N}{A_{\text{eff}}/\sigma_{\text{N}}}\right)^2 + 2C^2, \tag{8.4}$$

where the uncertainties of the two terms are taken directly from the fit of the $\sigma(\Delta t)$ vs. $A_{\text{eff}}/\sigma_{\text{N}}$ curve.

A detailed study revealed that $\sigma(\Delta t)$ depends on whether the two neighboring crystals belong to the same readout unit (each readout unit consists of a grid of 5×5 crystals) or not, as there is clock jitter among the readout units. We refer to the $\sigma(\Delta t)$ obtained for pairs of crystals in the same readout unit as the intrinsic time resolution of the crystals, as the inter-unit clock jitter does not enter.

Figure 8.9 shows the result of the single crystal's timing resolution measurement obtained with the 2016 data. Two additional curves and fitting results are also shown, which are the measurements of $\sigma(\Delta t)$ for two neighboring crystals that belong to different ECAL readout units, and the combined result that includes both cases: of two neighboring crystals that belong to the same or different readout units. We can see that once the clock jitter is included, the combined resolution gets slightly worse. The constants N and C in Eqn. 8.2 are measured to be $N = 31.6 \pm 1.2$ ns, $C = 0.077 \pm 0.001$ ns for a pair of crystals in the same readout unit, and $N = 32.9 \pm 1.2$ ns, $C = 0.085 \pm 0.001$ ns for the combined result. As the clock jitter is not simulated in the MC simulation, we will take this effect into account when we make corrections to the MC timestamp in a later step.

From the time resolution of each ECAL crystal (parameterized in Eq. 8.2, with the values of the constants N and C shown in Figure 8.9), the time resolution of any ECAL crystal hit can then be estimated given the energy deposited and the pedestal noise for that hit. And based on this estimation, the arrival time of the whole photon
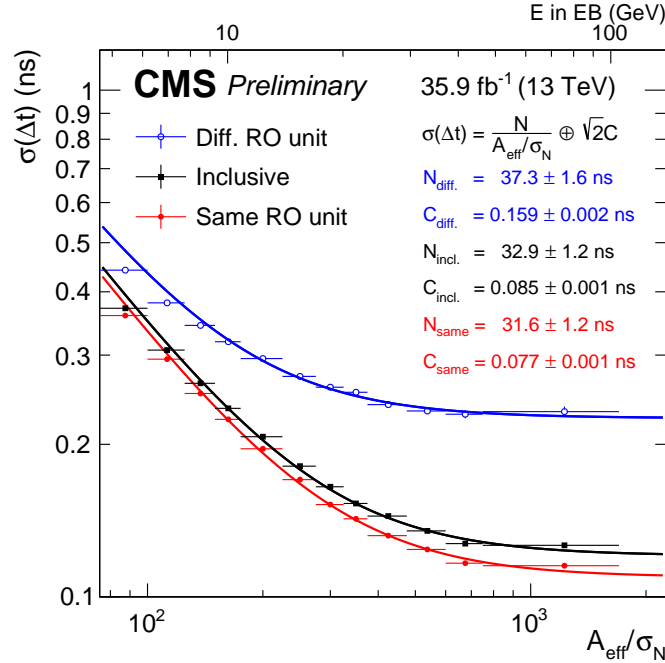
Figure 8.9: The resolution of the time difference ($\sigma(\Delta t)$) of two neighboring ECAL crystals as a function of the effective amplitude/noise ratio of the two crystals (A_{eff}/σ_N). The measurement is performed in three cases based on whether the two crystals belong to the same or different ECAL readout (RO) units or a mix of the two cases. The dependence of $\sigma(\Delta t)$ on A_{eff}/σ_N is fitted by the function in Eqn. 8.4, and the results of the fits are shown in the plot. The top ticks on the x-axis show the approximate ECAL energy deposited (in GeV) for the corresponding A_{eff}/σ_N, given the average pedestal noise (about 62 MeV) in 2016.

cluster can then be estimated as the resolution-weighted average of the times of all the crystal hits in the cluster:

$$ t_{\text{ECAL}} = \frac{\sum_i \frac{t^i_{\text{ECAL}}}{\sigma_i^2}}{\sum_i \frac{1}{\sigma_i^2}}, \tag{8.5} $$

where t^i_{ECAL} is the timestamp of the signal pulse in the crystal and σ_i is the resolution of that timestamp estimated with the method presented above.

For photons from prompt particle decays at the primary vertex, the distribution of their arrival times peaks at zero, as they are defined with respect to the collision time (which is given by the LHC clock) minus L/c, where L is the distance from the detector center to the photon cluster center, and c is the speed of light in vacuum. The width of this distribution is about 400 ps, and it has several contributions (in order of importance):

- The collision time with respect to the time given by the LHC clock. As this time is not measured directly by CMS with the current detector, this time has a spread of about 200 ps (given by the beam spot time spread).

- The clock jitter of the ECAL readout unit, which contributes about 150 ps.

- The intrinsic time resolution of each ECAL channel, which corresponds to Eqn. 8.2 and is typically 100 to 150 ps for high energy hits.

- The fact that the primary vertex position differs from the center of the detector. This is corrected by replacing L with the distance from the reconstructed primary vertex position to the photon cluster position in the above definition (a time-of-flight correction).

- The intrinsic jitter of the LHC clock, which is very small (sub-10 ps) [151].

As some factors (such as the clock jitter) which affect the ECAL timing (both the mean value and the resolution) are not fully simulated in MC, the timestamps reconstructed with the above method are slightly different for the EM shower cluster in the data and MC. To measure this difference, the mean and standard deviation of the electron cluster timestamp are measured as a function of the electron energy. Figure 8.10 shows the result of this measurement for both data and MC electron clusters. From the plot, we see that the MC timestamp needs to be shifted and smeared to match the data. The shift and smearing corrections for the MC timestamps are applied based on the energy of the EM cluster, using the data and MC differences shown in Figure 8.10.

8.4 Event selection

Events with at least two photons with p_T larger than 70 and 40 GeV are selected. Both GED and OOT photons are considered in the selection. The leading photon is required to be in the barrel region of the ECAL ($|\eta| < 1.4442$), and it has to pass the ID cuts as listed in Table 8.2. In addition, this photon has to pass $R_9 > 0.9$ and $S_{minor} < 0.4$ cuts to further suppress fake photons from jets. Furthermore, a conversion-safe electron veto cut [149] is also applied on this photon to suppress electrons that are misidentified as photons. The sub-leading photon can be either in the barrel or endcap region, and only very loose ID cuts (a trigger-like selection) are applied. The requirement of the presence of two photons rejects the majority of background events from pp collisions that have a prompt photon, such as QCD multijet events or γ + jets events.
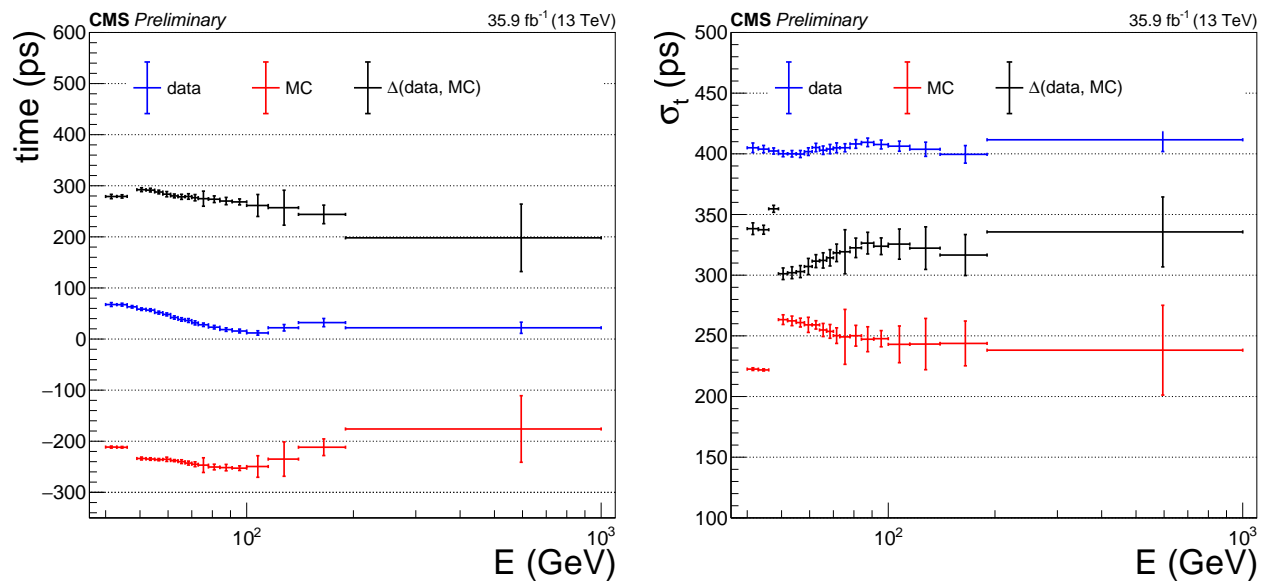
Figure 8.10: The mean (left) and standard deviation (right) of the electron cluster times in the data and MC of 2016, and the difference between the data and MC, in bins of the electron cluster energy.

In order to suppress non-collisional background events such as hits from beam halo muon-induced photons in the ECAL, events are also required to have at least three jets, with each jet having $p_T > 30$ GeV, $|\eta| < 3.0$ and to pass the tight jet ID criteria, which is the same jet ID used in the VVV search as described in Section 7.3 and Table 7.5. The non-collisional background events are likely to produce photons with non-zero arrival times at the ECAL, which is similar to what our delayed photons signals would look like.

Table 8.4 summarizes the full set of selections used in this search. The selection efficiency times acceptance for different GMSB SPS8 signal samples are shown in Figure 8.11. As expected, the efficiency times acceptance is high for small $c\tau$ and large Λ samples. Photons in those samples have larger probability to be produced inside the ECAL and the photons are also more boosted due to the large Λ. The efficiency times acceptance is about 20% (0.15%) for small (large) $c\tau$ samples. A breakdown of this efficiency into steps of cuts mentioned above can be found in Tables 8.5, 8.6, 8.7, and 8.8 for different signal models.

Table 8.4: Summary of event selection used in the delayed photon search.

trigger	signal trigger as discussed in Section 8.1
leading photon	$p_\mathrm{T} > 70\,\mathrm{GeV}$ $\|\eta\| < 1.4442$ GED or OOT ID as in Table 8.2 $S_\mathrm{minor} < 0.4$ $R_9 > 0.9$ conversion-safe electron veto
sub-leading photon	$p_\mathrm{T} > 40\,\mathrm{GeV}$ $\|\eta\| < 1.4442$ or $1.566 < \|\eta\| < 2.5$ very loose ID (trigger like)
jets	nJets ≥ 3 $p_\mathrm{T} > 30\,\mathrm{GeV}$ $\|\eta\| < 3.0$ Jet ID as discussed in Section 8.2
$p_\mathrm{T}^\mathrm{miss}$ filters	as discussed in Section 8.2

Table 8.5: Event selection cut-flow efficiency for GMSB SPS8 signal samples of $c\tau = 10\,\mathrm{cm}$ and varying Λ (unit of efficiency: %; unit of Λ : TeV).

selection	$\Lambda = 100$	$\Lambda = 200$	$\Lambda = 300$	$\Lambda = 400$
-	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00
+ Signal Trigger	62.46 ± 0.33	61.78 ± 0.31	62.42 ± 0.32	65.13 ± 0.32
+ 1st Photon p_T and η cut	59.32 ± 0.31	61.13 ± 0.31	62.23 ± 0.32	65.05 ± 0.32
+ 1st Photon ID	21.98 ± 0.16	31.22 ± 0.20	38.93 ± 0.23	43.35 ± 0.25
+ 1st Photon Electron Veto	20.51 ± 0.16	29.44 ± 0.19	36.80 ± 0.22	40.85 ± 0.24
+ nJets≥ 3	17.75 ± 0.15	22.50 ± 0.16	27.29 ± 0.19	30.15 ± 0.19
+ $p_\mathrm{T}^\mathrm{miss}$ filters	17.61 ± 0.14	22.26 ± 0.16	27.02 ± 0.18	29.82 ± 0.19
+ 2nd Photon p_T and η cut	16.20 ± 0.14	21.04 ± 0.16	25.74 ± 0.18	28.61 ± 0.19
+ 2nd Photon ID	12.23 ± 0.12	17.33 ± 0.14	22.58 ± 0.16	25.29 ± 0.17
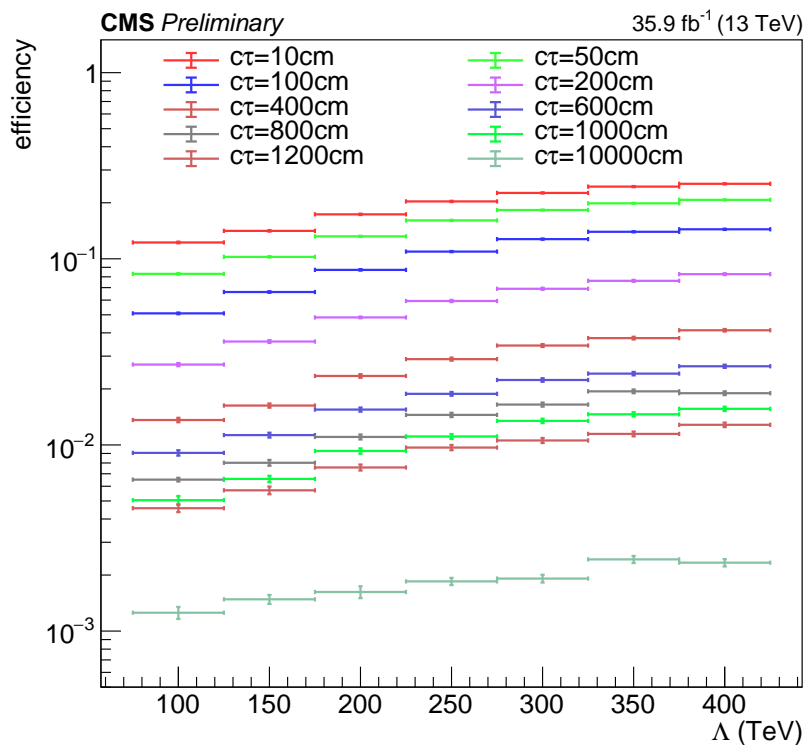
Figure 8.11: Selection efficiency times acceptance for GMSB SPS8 signal models of different $c\tau$ and Λ. The x-axis is the Λ of the signal sample, and different curves are for different values of $c\tau$ of the signal sample.

Table 8.6: Event selection cut-flow efficiency for GMSB SPS8 signal samples of $c\tau$ = 100 cm and varying Λ (unit of efficiency: %; unit of Λ : TeV).

selection	$\Lambda = 100$	$\Lambda = 200$	$\Lambda = 300$	$\Lambda = 400$
-	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00
+ Signal Trigger	35.62 ± 0.15	40.39 ± 0.19	45.26 ± 0.21	48.79 ± 0.23
+ 1st Photon p_T and η cut	33.15 ± 0.15	39.71 ± 0.19	44.96 ± 0.21	48.63 ± 0.23
+ 1st Photon ID	11.07 ± 0.08	18.28 ± 0.11	24.87 ± 0.14	28.08 ± 0.16
+ 1st Photon Electron Veto	9.63 ± 0.07	16.58 ± 0.11	22.86 ± 0.13	25.96 ± 0.15
+ nJets$\geq$ 3	8.10 ± 0.06	12.21 ± 0.09	16.48 ± 0.11	18.67 ± 0.13
+ p_T^{miss} filters	7.64 ± 0.06	11.58 ± 0.09	15.72 ± 0.11	17.68 ± 0.12
+ 2nd Photon p_T and η cut	6.83 ± 0.06	10.76 ± 0.08	14.84 ± 0.10	16.90 ± 0.12
+ 2nd Photon ID	5.09 ± 0.05	8.71 ± 0.07	12.75 ± 0.09	14.40 ± 0.11

Table 8.7: Event selection cut-flow efficiency for GMSB SPS8 signal samples of $c\tau$ = 1000 cm and varying Λ (unit of efficiency: %; unit of Λ : TeV).

selection	Λ = 100	Λ = 200	Λ = 300	Λ = 400
-	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00
+ Signal Trigger	10.05 ± 0.12	10.17 ± 0.10	11.20 ± 0.11	12.44 ± 0.12
+ 1st Photon p_{T} and η cut	9.01 ± 0.11	9.66 ± 0.10	10.79 ± 0.10	12.14 ± 0.12
+ 1st Photon ID	2.42 ± 0.05	3.57 ± 0.06	5.14 ± 0.07	6.01 ± 0.08
+ 1st Photon Electron Veto	1.118 ± 0.040	2.145 ± 0.046	3.26 ± 0.05	3.81 ± 0.06
+ nJets≥ 3	0.914 ± 0.036	1.474 ± 0.038	2.044 ± 0.045	2.34 ± 0.05
+ $p_{\mathrm{T}}^{\mathrm{miss}}$ filters	0.867 ± 0.035	1.390 ± 0.037	1.895 ± 0.043	2.158 ± 0.049
+ 2nd Photon p_{T} and η cut	0.732 ± 0.032	1.204 ± 0.034	1.656 ± 0.041	1.964 ± 0.047
+ 2nd Photon ID	0.504 ± 0.026	0.928 ± 0.030	1.346 ± 0.036	1.563 ± 0.042

Table 8.8: Event selection cut-flow efficiency for GMSB SPS8 signal samples of $c\tau$ = 10000 cm and varying Λ (unit of efficiency: %; unit of Λ : TeV).

selection	Λ = 100	Λ = 200	Λ = 300	Λ = 400
-	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00
+ Signal Trigger	6.32 ± 0.06	5.65 ± 0.07	5.345 ± 0.049	6.00 ± 0.05
+ 1st Photon p_{T} and η cut	5.59 ± 0.06	5.19 ± 0.06	4.991 ± 0.047	5.71 ± 0.05
+ 1st Photon ID	1.425 ± 0.031	1.668 ± 0.038	2.208 ± 0.031	2.849 ± 0.037
+ 1st Photon Electron Veto	0.265 ± 0.013	0.437 ± 0.019	0.573 ± 0.015	0.732 ± 0.018
+ nJets≥ 3	0.234 ± 0.012	0.287 ± 0.015	0.301 ± 0.011	0.370 ± 0.013
+ $p_{\mathrm{T}}^{\mathrm{miss}}$ filters	0.231 ± 0.012	0.285 ± 0.015	0.295 ± 0.011	0.364 ± 0.013
+ 2nd Photon p_{T} and η cut	0.180 ± 0.011	0.231 ± 0.014	0.239 ± 0.010	0.291 ± 0.011
+ 2nd Photon ID	0.125 ± 0.009	0.162 ± 0.011	0.191 ± 0.009	0.233 ± 0.010

8.5 Background estimation

The missing transverse momentum p_T^{miss} and photon arrival time t_γ are used as the final discriminating variables to distinguish signal from background using an ABCD method as described below. With the ABCD method, only observed data and simulated signal samples are needed to estimate the background yield, avoiding issues from the fact that some backgrounds which contribute to events with time delayed photons are not simulated, such as photons from satellite bunches which are spaced about 2.5 ns apart from the main bunches. Events are first divided into four bins in the 2D plane of p_T^{miss} and t_γ (rectangular binning), which we refer to as bin A, B, C and D, where bin A has low p_T^{miss} and low t_γ, bin B has high p_T^{miss} and low t_γ, bin C has high p_T^{miss} and high t_γ, and bin D has low p_T^{miss} and high t_γ, as illustrated in Figure 8.12. As we will demonstrate later, the p_T^{miss} and t_γ are uncorrelated for the backgrounds that pass our selection. As a result, the backgrounds in bin C can be estimated by $\text{Bkg}_B \times \text{Bkg}_D / \text{Bkg}_A$. The sum of signal and background events in the four bins (N_A, N_B, N_C, N_D) can be expressed as:

$$N_A = \text{Bkg}_A + \mu \times \text{Sig}_A \tag{8.6}$$
$$N_B = c_1 \times \text{Bkg}_A + \mu \times \text{Sig}_B$$
$$N_C = c_1 \times c_2 \times \text{Bkg}_A + \mu \times \text{Sig}_C$$
$$N_D = c_2 \times \text{Bkg}_A + \mu \times \text{Sig}_D,$$

where:

- Bkg_A is the background yield in bin A (unknown);

- c_1 is the ratio between the backgrounds in B and A, and c_2 is the ratio between the backgrounds in D and A (unknown);

- μ is the overall signal strength (unknown);

- Sig_A, Sig_B, Sig_C, and Sig_D are the predicted signal yields in bins A, B, C and D (taken directly from signal MC).

A fit with input from the observed data and expected signal yields in the four bins is then performed to extract the four unknown variables (c_1, c_2, Bkg_A, μ), by
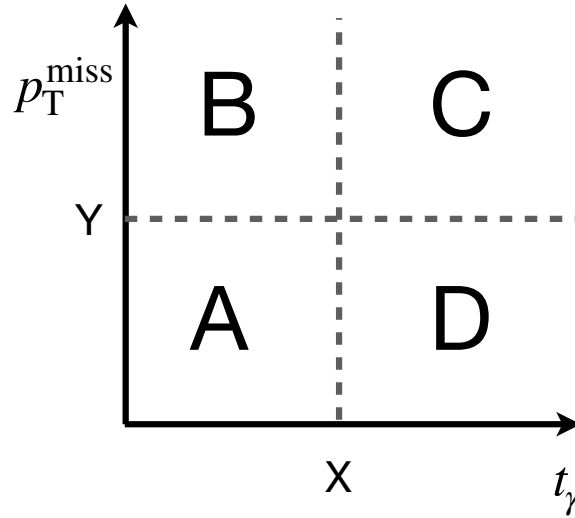
Figure 8.12: Illustration of four bins A, B, C and D dividing the t_γ and p_T^{miss} 2D plane. The boundaries X and Y are optimized for different signal models.

maximizing the following likelihood:

$$L = \prod_i^{\text{ABCD}} \text{Poisson}(\text{Obs}_i|N_i) \times \prod_j^{\text{nuisance}} \text{Constraints}(\sigma_j, \hat{\sigma}_j) \qquad (8.7)$$

where Obs_i is the observed events in each of the four bins, and σ_j are the nuisance parameters for systematic uncertainties with either Gaussian or log normal constraints.

An important precondition of using the ABCD method is to make sure that p_T^{miss} and t_γ are uncorrelated and independent for background events. For non-collisional backgrounds, such as from beam halo or cosmic ray muons, the events are usually enriched in the large p_T^{miss} and large t_γ region, which results in a highly correlated 2D distribution of p_T^{miss} and t_γ. However, those backgrounds are reduced to a negligible level by the number of jet and number of photons requirements. The dominant backgrounds come from pp collisions, for which the p_T^{miss} and t_γ are uncorrelated and the t_γ distributions are the same for low p_T^{miss} and high p_T^{miss} events.

To verify that the p_T^{miss} and t_γ are uncorrelated and independent, two control regions (CR) are defined, QCD CR and γ + jets CR, where each of them is defined in the same way as the signal region, but with one cut inverted. The γ + jets CR is defined by inverting the number of jets required (nJets < 3), and the QCD CR is defined by inverting the photon ID requirement on the leading photon. For the events in the

two control regions, we measure the Pearson correlation coefficient between p_T^{miss} and t_γ (the covariance of the two variables divided by the product of their standard deviations), and find that the correlation coefficients are below 1% in both cases, even for events in the large time regions. In addition, we also see that the t_γ shapes are the same for different p_T^{miss} values, as shown in Figure 8.13.
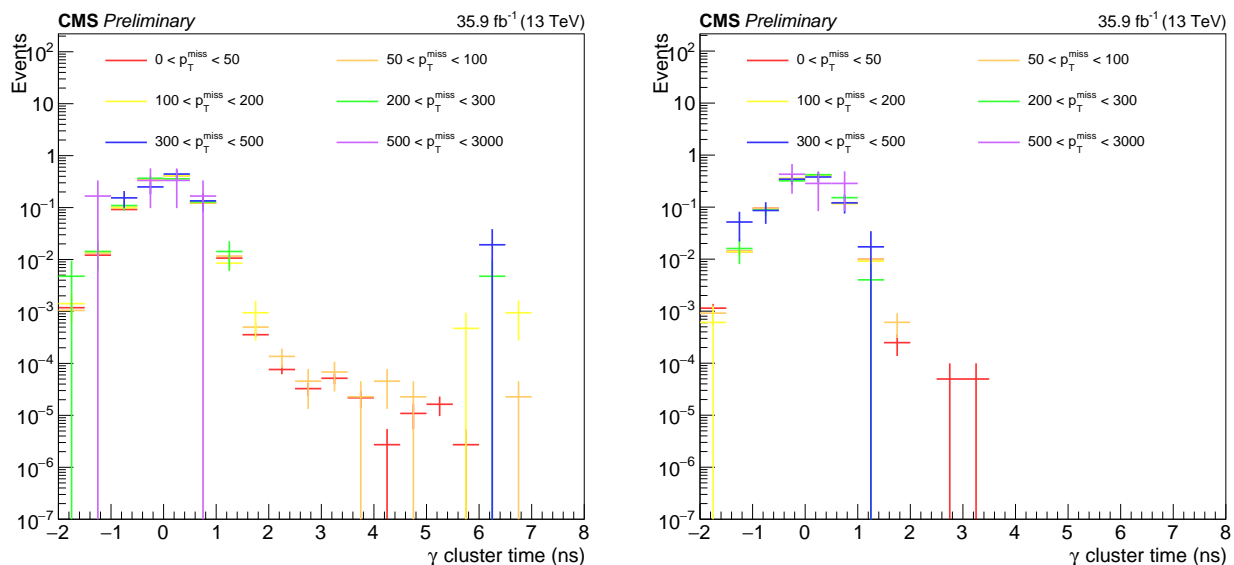
Figure 8.13: Distribution of t_γ for events with different p_T^{miss}, from the data in the control regions γ + jets CR (left) and QCD CR (right). All the distributions are normalized to have an integral of 1 in order to compare the shapes of the different curves.

With the four bins A, B, C and D defined above, the delayed photon signals will be enriched in bin C, i.e. the bin with large p_T^{miss} and large t_γ. Because different signal models with different $c\tau$ and Λ have different t_γ and p_T^{miss} distributions, the optimal binning to define the four bins (i.e. the time split and p_T^{miss} split) are different. The p_T^{miss} and t_γ boundaries that define the four bins A, B, C and D are chosen to yield the optimal expected sensitivity. For each signal model of a given Λ and $c\tau$, we scan over all the combinations of the following p_T^{miss} and t_γ splits:

- p_T^{miss} split candidates: 50, 100, 150, 200, 300, 500, 1000 GeV;

- t_γ split candidates: 0, 0.5, 1.0, 1.5, 2.0, 3.0, 5.0 ns.

For each p_T^{miss} and t_γ split combination candidate, we take the expected signal yields in bins A, B, C and D from signal MC, and the expected background yield in the four

bins from data, in a blinded way. First of all, as bin A is the least sensitive bin and the signal-to-background ratio in bin A is too small, we directly take the observed events in bin A, N_A, as the predicted background in bin A. The backgrounds in the other three bins (B, C, and D) are then estimated based on N_A and the weighted ratio $r_{B/A}$ between the high p_T^{miss} and low p_T^{miss} bins, or the ratio $r_{D/A}$ between the high t_γ and low t_γ bins, where the weights are obtained from p_T^{miss} and t_γ template shapes, respectively. The p_T^{miss} and t_γ template shapes are obtained from the data in a signal-depleted t_γ or p_T^{miss} region, where the regions are defined respectively as $|t_\gamma| < 1$ ns or $p_T^{miss} < 100$ GeV.

From those templates, we obtain the ratios $r_{B/A}$ ($r_{D/A}$) by dividing the number of events with p_T^{miss} (t_γ) larger than the p_T^{miss} (t_γ) split candidate by the number of events with p_T^{miss} (t_γ) smaller than the p_T^{miss} (t_γ) split candidate. Figure 8.14 shows the t_γ and p_T^{miss} templates and also the ratios $r_{B/A}$ and $r_{D/A}$ for some representative p_T^{miss} and t_γ split candidates. The background estimates in bins B, C, and D are then calculated as $N_A r_{B/A}$, $N_A r_{B/A} r_{D/A}$, and $N_A r_{D/A}$, respectively. The expected 95% upper limit on the signal strength is then calculated based on the signal and background estimates, and the optimal p_T^{miss} and t_γ split combination is chosen from the one which gives the best expected upper limit on the signal strength for that signal model. It turns out that the optimal ABCD binning boundaries for similar signals with similar $c\tau$ or Λ are the same or very close to each other. To simplify the analysis, we choose the same binning boundary for similar signal models, as summarized in Table 8.9. As we can see from the table, the t_γ split is small for small $c\tau$ signal models, as those signals have smaller t_γ, and the p_T^{miss} split is small for small Λ signal models, as the p_T of the neutrinos are smaller for those signals.

Table 8.9: Optimal ABCD binning boundaries for different signal models.

$c\tau$ (m)	Λ (TeV)	p_T^{miss} split (GeV)	t_γ split (ns)
≤ 0.1	any	250	0.0
> 0.1	≤ 300	100	1.5
> 0.1	> 300	150	1.5

8.6 Systematic uncertainties

The dominant uncertainty in this search comes from the ABCD fit. The optimized ABCD binnings shown in Table 8.9 yield zero or very few observed events in the most sensitive bin (bin C). Performing a fit with few expected signal events and near-zero background predictions is very challenging, and leads to a large statistical
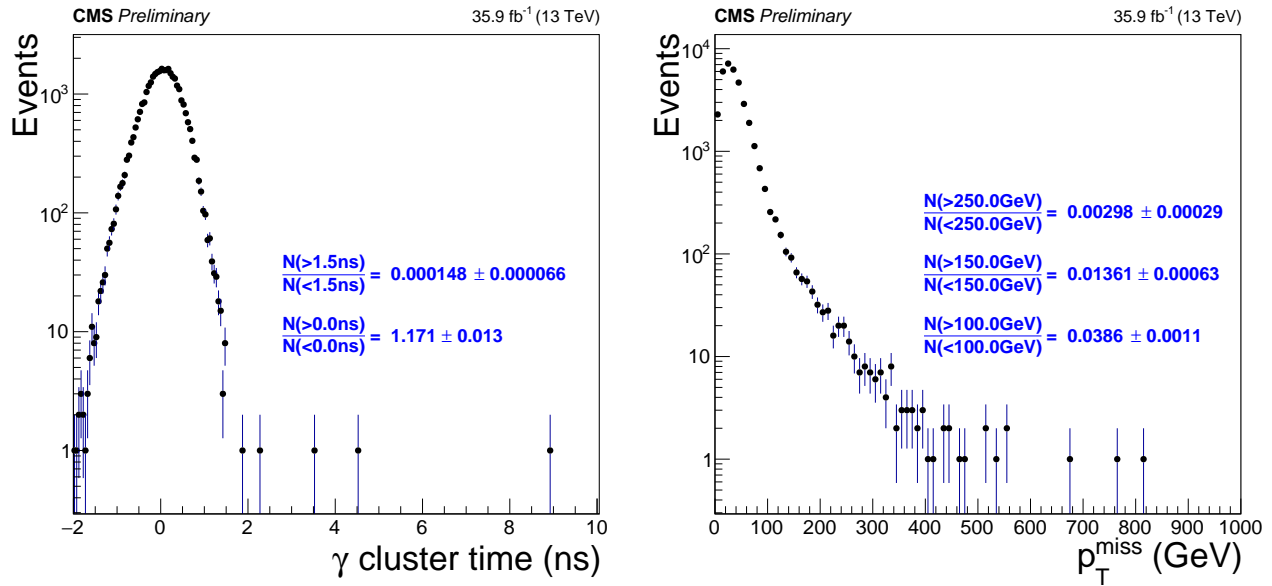
Figure 8.14: t_γ (left) and p_T^{miss} (right) templates derived from the data using the signal selection: for the t_γ template, events are selected with $p_T^{miss} < 100\,\text{GeV}$; for the p_T^{miss} templates, events are selected with $|t_\gamma| < 1\text{ns}$. Also shown in the plots are the estimates of the ratio of the number of events in bin B divided by the number in bin A ($r_{B/A}$) and the corresponding ratio for bins D and A ($r_{D/A}$), for different t_γ and p_T^{miss} splits. These ratios are used to predict the number of events in bins B, D, and C given the number of events in bin A.

uncertainty.

There are several subdominant effects that affect the signal predictions from simulation in each of the four bins A, B, C, and D. The effects include those which affect the overall signal normalization in each of the four bins such as the integrated luminosity uncertainty, and those which affect the event migration among the four bins. The latter category includes effects which might change the p_T^{miss} or t_γ shapes, such as the timing correction in MC, the energy scale and resolution of the photon and jet measurements, and the trigger and photon ID efficiencies along with the corresponding uncertainties. For all cases that affect the p_T^{miss} or t_γ shapes, dedicated measurements are performed to evaluate the data and MC differences; the resulting corrections are then applied to all simulated signal events, and the corresponding uncertainties are propagated to the signal yield predictions as uncertainties in the p_T^{miss} or t_γ shapes. For example, the photon and jet energy scale and resolution and the trigger and ID efficiency uncertainties are applied to the p_T^{miss} shape uncertainties, and the MC time correction uncertainties (in the shift correction and the

resolution correction) are applied to the t_γ shape uncertainty.

The background prediction method relies on the independence of p_T^{miss} and t_γ, which has been verified by checking the t_γ shape for different p_T^{miss} values, as shown in Figure 8.13. An additional systematic uncertainty on the background prediction has been evaluated to account for any additional dependence between the two variables. In order to do this, we look at the number of observed events in bins A, B, C, and D in the QCD and γ + jets control regions (CRs): N_A^{CR}, N_B^{CR}, N_C^{CR}, and N_D^{CR}. If p_T^{miss} and t_γ are independent, we know that the predicted observed events in bin C should be approximately $N_A^{\text{CR,pred}} = N_B^{\text{CR}} N_D^{\text{CR}} / N_A^{\text{CR}}$. We then compare N_C^{CR} with $N_A^{\text{CR,pred}}$ for all binning options in Table 8.9 in both the QCD and γ+jets CRs. It has been found that for $c\tau \leq 0.1$ m binning, the difference of N_C^{CR} and $N_A^{\text{CR,pred}}$ is about 2%. But for $c\tau > 0.1$ m binning, the statistics in bin C can be too small, leading to a difference of nearly 100% (i.e. there are cases where $0 \sim N_C^{\text{CR}} < N_A^{\text{CR,pred}}$). Because of this observation, we decided to assign a 2% (90%) uncertainty on the background prediction in bin C for $c\tau \leq 0.1$ m ($c\tau > 0.1$ m) models.

A summary of the systematic uncertainties for both the signal and background predictions and the typical values of the uncertainties are shown in Table 8.10.

Table 8.10: Summary of systematic uncertainties in the delayed photon search. Also included are notes on whether each source affects the signal yields (Sig) or the background (Bkg) estimates, and to which bins each uncertainty applies.

source	Sig/Bkg	Bins	uncertainty
Integrated luminosity	Sig	A,B,C,D	2.5%
Photon energy scale	Sig	A,B,C,D	1%
Photon energy resolution	Sig	A,B,C,D	1%
Jet energy scale	Sig	A,B,C,D	1.5%
Jet energy resolution	Sig	A,B,C,D	1.5%
Photon time bias	Sig	A,B,C,D	1.5%
Photon time resolution	Sig	A,B,C,D	0.5%
Trigger efficiency	Sig	A,B,C,D	2%
Photon identification	Sig	A,B,C,D	2%
Closure in bin C ($c\tau \leq 0.1$ m)	Bkg	C	2%
Closure in bin C ($c\tau > 0.1$ m)	Bkg	C	90%

8.7   Results

As discussed in Section 8.5, the signal and background yields are estimated from a fit to the observed data in bins of p_T^{miss} and t_γ. Figure 8.15 shows the p_T^{miss} and t_γ distributions after event selection for both the observed data and the signal MC. The

plots show that both the p_T^{miss} and t_γ shapes are very different for the signal versus the data (which has been proven to consist mainly of backgrounds, as described below), while the p_T^{miss} (t_γ) shapes for different t_γ (p_T^{miss}) regions are almost identical in the data. This further proves that it is valid to use the ABCD method discussed above to fit the signal and background yields with these two variables.
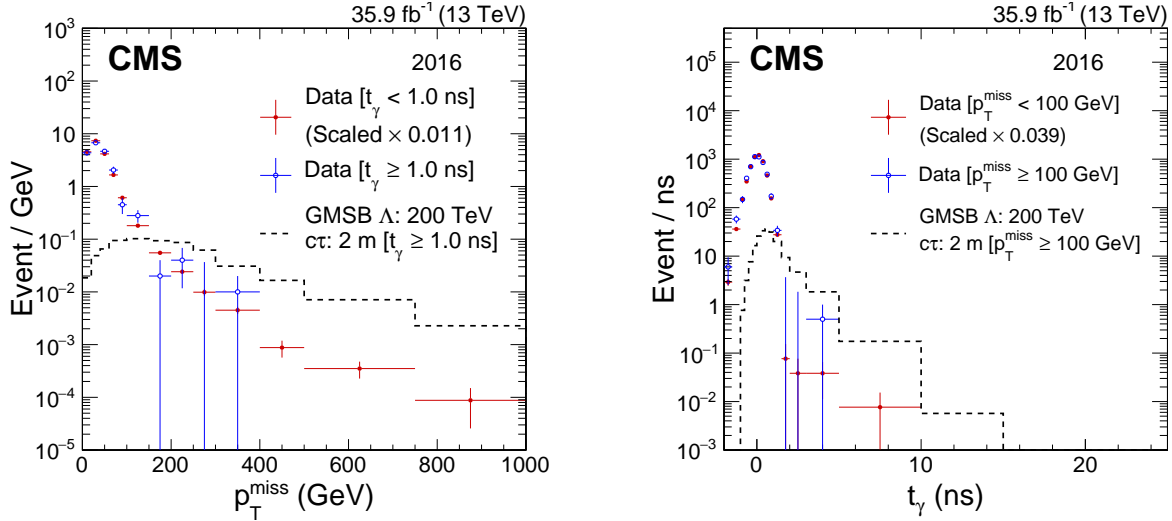

Figure 8.15: The p_T^{miss} (left) and t_γ (right) distributions after the event selection, shown for the data and for a representative signal benchmark (GMSB: $\Lambda = 200\,\text{TeV}$, $c\tau = 2m$). The p_T^{miss} distribution for data is separated into events with $t_\gamma \geq 1$ ns (blue) and $t_\gamma < 1$ ns (red), scaled to match the total number of events with $t_\gamma \geq 1$ ns. The t_γ distribution for data is separated into events with $p_T^{miss} \geq 100\,\text{GeV}$ (blue, darker) and $p_T^{miss} < 100\,\text{GeV}$ (red, lighter), scaled to match the total number of events with $p_T^{miss} \geq 100\,\text{GeV}$. The signal (black, dotted) is shown in the left plot only for events with $t_\gamma \geq 1$ ns, and in the right plot only for events with $p_T^{miss} \geq 100\,\text{GeV}$. The entries in each bin are normalized by the bin width. The horizontal bars on data indicate the bin boundaries. The last bin in each plot includes overflow events.

From these distributions, we count the number of events in bins A, B, C, and D, for each of the signal models, as summarized in Tables 8.11, 8.12, 8.13, 8.14, 8.15, and 8.17 which show the signal yields estimated from the MC simulation (pre-fit). As shown, the expected signal yields in the each of the four bins decrease as the $c\tau$ of the signal model increases (because of decreasing acceptance) or as the Λ of the signal model increases (because of a decreasing production cross section). Table 8.18 shows the observed data compared to the background-only fit yields in bins A, B, C and D. No statistically significant excess or deficit relative to the background expectation is observed in any of the bins.

The observed data is then used to perform the signal+background fit for each signal model to obtain the post-fit background and signal yields, and the fit results are interpreted as the upper limits on the cross sections of signal production for various Λ and $c\tau$ parameter points in the GMSB SPS8 scenario.

The modified frequentist criterion $\mathrm{CL_s}$ [152–154] with the profile likelihood ratio test statistic determined by toy experiments is used to evaluate the observed and expected upper limits at the 95% confidence level on the signal production cross sections. In the toy experiments, we evaluate the confidence level to exclude different injected signal strengths r, the ratio of the measured to the predicted signal production cross section. The exclusion limit on r is taken where the 95% confidence level is reached, which corresponds to $\mathrm{CL_s} = 0.05$.

As an example, Figure 8.16 shows the $\mathrm{CL_s}$ value as a function of the signal strength r obtained from toy experiments (10000 toys) for the expected and observed limits for a representative signal point ($\Lambda = 350$ TeV, $c\tau = 2$ m). The corresponding 95% confidence level upper limits are taken from such plots.

Summary plots of the observed and expected 95% confidence level upper limits on the signal production cross section as a function of the mass of the neutralino (which is linearly related to Λ) for signal models with different values of $c\tau$ are shown in Figures 8.17 and 8.18. Also shown in these plots is the theoretical cross section, which decreases exponentially as a function of the neutralino mass. Signal models with an observed upper limit cross section smaller than the theoretical cross section are excluded at the 95% confidence level.

Figure 8.19 shows the upper limits on the cross section in the 2D plane of Λ versus $c\tau$. Also shown in Figure 8.19 are the exclusion boundaries in the 2D plane of Λ versus $c\tau$ from this search, and the exclusion boundaries from previous ATLAS and CMS searches. It can be seen that with $35.9\,\mathrm{fb}^{-1}$ data taken from 2016, models with neutralino masses up to 300, 425, 250, and 200 GeV are excluded for neutralino $c\tau$ of 0.1, 1, 10, and 100 m, respectively. The result is much better than the CMS Run 1 result, and similar to the $20.3\,\mathrm{fb}^{-1}$ result from the ATLAS Run 1 search, with slightly more exclusion power than the ATLAS result in the large $c\tau$ region.

Table 8.11: Signal yield prediction (pre-fit) in the bins A, B, C, and D estimated from signal MC after event selection for the GMSB SPS8 $\Lambda = 100$ TeV models with varying $c\tau$. See Table 8.9 for the t_γ-p_T^{miss} splits for each signal point.

$c\tau$ (cm)	Yield in Bin A	Yield in Bin B	Yield in Bin C	Yield in Bin D
10	2200 ± 43	928 ± 27	1746 ± 38	4590 ± 63
50	1600 ± 27	4590 ± 47	141 ± 8	88 ± 6
100	1012 ± 21	2638 ± 35	180 ± 9	106 ± 7
200	469 ± 19	1416 ± 33	152 ± 11	72 ± 8
400	230 ± 13	703 ± 23	86 ± 8	39 ± 6
600	130 ± 10	485 ± 19	76 ± 8	20.3 ± 3.9
800	83.3 ± 4.6	361 ± 10	47.7 ± 3.5	15.4 ± 1.9
1000	64 ± 8	286 ± 17	31 ± 6	12.2 ± 3.6
1200	54 ± 7	248 ± 13	30.8 ± 4.9	18.9 ± 3.8
10000	14.8 ± 2.8	85 ± 7	1.1 ± 0.8	1.2 ± 0.8

Table 8.12: Signal yield prediction (pre-fit) in bins A, B, C, and D estimated from signal MC after event selection for GMSB SPS8 $\Lambda = 150$ TeV and varying $c\tau$. See Table 8.9 for the t_γ-p_T^{miss} splits for each signal point.

$c\tau$ (cm)	Yield in Bin A	Yield in Bin B	Yield in Bin C	Yield in Bin D
10	220.2 ± 4.3	139.8 ± 3.4	293.7 ± 4.9	506 ± 7
50	164.9 ± 2.9	618.4 ± 5.8	42.2 ± 1.4	14.9 ± 0.8
100	100.8 ± 2.2	373.2 ± 4.3	50.7 ± 1.5	19.3 ± 0.9
200	52.1 ± 2.0	189.9 ± 4.0	37.0 ± 1.7	15.6 ± 1.1
400	20.3 ± 1.3	86.7 ± 2.7	18.9 ± 1.2	7.3 ± 0.8
600	13.9 ± 1.0	57.2 ± 2.1	15.3 ± 1.1	5.0 ± 0.6
800	8.3 ± 0.8	43.3 ± 1.9	11.1 ± 0.9	2.87 ± 0.48
1000	5.3 ± 0.6	35.6 ± 1.7	9.7 ± 0.8	3.2 ± 0.5
1200	4.9 ± 0.7	31.2 ± 1.8	8.2 ± 0.9	2.29 ± 0.48
10000	1.33 ± 0.22	9.7 ± 0.6	1.19 ± 0.21	0.19 ± 0.08

Table 8.13: Signal yield prediction (pre-fit) in bins A, B, C, and D estimated from signal MC after event selection for GMSB SPS8 $\Lambda = 200\,\text{TeV}$ and varying $c\tau$. See Table 8.9 for the t_γ-p_T^miss splits for each signal point.

$c\tau$ (cm)	Yield in Bin A	Yield in Bin B	Yield in Bin C	Yield in Bin D
10	45.4 ± 0.8	37.0 ± 0.7	83.1 ± 1.1	111.3 ± 1.3
50	30.30 ± 0.47	161.0 ± 1.1	16.10 ± 0.34	3.43 ± 0.15
100	20.20 ± 0.42	96.3 ± 0.9	18.10 ± 0.40	4.54 ± 0.20
200	10.75 ± 0.41	48.7 ± 0.8	13.93 ± 0.47	3.68 ± 0.24
400	4.19 ± 0.27	23.7 ± 0.6	7.87 ± 0.38	1.75 ± 0.18
600	2.39 ± 0.19	15.7 ± 0.5	5.42 ± 0.29	1.21 ± 0.13
800	1.63 ± 0.16	11.20 ± 0.42	3.98 ± 0.25	0.87 ± 0.11
1000	1.32 ± 0.14	9.02 ± 0.37	3.45 ± 0.23	0.76 ± 0.10
1200	1.20 ± 0.14	7.44 ± 0.35	2.76 ± 0.21	0.61 ± 0.10
10000	0.23 ± 0.05	2.04 ± 0.17	0.31 ± 0.06	0.049 ± 0.026

Table 8.14: Signal yield prediction (pre-fit) in bins A, B, C, and D estimated from signal MC after event selection for GMSB SPS8 $\Lambda = 250\,\text{TeV}$ and varying $c\tau$. See Table 8.9 for the t_γ-p_T^miss splits for each signal point.

$c\tau$ (cm)	Yield in Bin A	Yield in Bin B	Yield in Bin C	Yield in Bin D
10	12.3 ± 0.23	13.93 ± 0.25	34.29 ± 0.4	31.58 ± 0.39
50	7.89 ± 0.13	56.50 ± 0.39	7.23 ± 0.13	1.19 ± 0.05
100	5.45 ± 0.10	34.10 ± 0.25	8.53 ± 0.12	1.39 ± 0.05
200	2.67 ± 0.11	16.91 ± 0.28	6.21 ± 0.16	1.09 ± 0.07
400	1.17 ± 0.07	7.77 ± 0.18	3.41 ± 0.12	0.71 ± 0.05
600	0.62 ± 0.05	5.08 ± 0.15	2.47 ± 0.10	0.335 ± 0.039
800	0.483 ± 0.046	3.92 ± 0.13	1.88 ± 0.09	0.269 ± 0.034
1000	0.338 ± 0.039	3.10 ± 0.11	1.34 ± 0.07	0.231 ± 0.032
1200	0.352 ± 0.040	2.53 ± 0.10	1.30 ± 0.07	0.180 ± 0.028
10000	0.062 ± 0.010	0.612 ± 0.031	0.149 ± 0.015	0.0112 ± 0.0043

Table 8.15: Signal yield prediction (pre-fit) in bins A, B, C, and D estimated from signal MC after event selection for GMSB SPS8 $\Lambda = 300$ TeV and varying $c\tau$. See Table 8.9 for the t_γ-p_T^{miss} splits for each signal point.

$c\tau$ (cm)	Yield in Bin A	Yield in Bin B	Yield in Bin C	Yield in Bin D
10	3.91 ± 0.08	6.27 ± 0.10	15.84 ± 0.17	10.08 ± 0.13
50	2.390 ± 0.040	23.20 ± 0.13	3.260 ± 0.047	0.370 ± 0.015
100	1.620 ± 0.037	14.20 ± 0.11	4.05 ± 0.05	0.440 ± 0.019
200	0.851 ± 0.038	6.86 ± 0.11	2.89 ± 0.07	0.366 ± 0.024
400	0.359 ± 0.024	3.27 ± 0.07	1.62 ± 0.05	0.202 ± 0.018
600	0.219 ± 0.019	2.09 ± 0.05	1.124 ± 0.043	0.124 ± 0.014
800	0.162 ± 0.016	1.54 ± 0.05	0.799 ± 0.036	0.129 ± 0.014
1000	0.146 ± 0.015	1.253 ± 0.044	0.667 ± 0.032	0.086 ± 0.011
1200	0.057 ± 0.009	0.994 ± 0.040	0.583 ± 0.030	0.060 ± 0.009
10000	0.0167 ± 0.0034	0.221 ± 0.012	0.065 ± 0.0060	0.0037 ± 0.0016

Table 8.16: Signal yield prediction (pre-fit) in bins A, B, C, and D estimated from signal MC after event selection for GMSB SPS8 $\Lambda = 350$ TeV and varying $c\tau$. See Table 8.9 for the t_γ-p_T^{miss} splits for each signal point.

$c\tau$ (cm)	Yield in Bin A	Yield in Bin B	Yield in Bin C	Yield in Bin D
10	1.311 ± 0.029	2.814 ± 0.043	7.78 ± 0.07	3.76 ± 0.05
50	1.700 ± 0.022	9.38 ± 0.05	1.410 ± 0.020	0.250 ± 0.008
100	1.140 ± 0.019	5.780 ± 0.044	1.660 ± 0.023	0.370 ± 0.010
200	0.568 ± 0.019	2.810 ± 0.043	1.248 ± 0.028	0.267 ± 0.013
400	0.246 ± 0.012	1.305 ± 0.029	0.706 ± 0.021	0.145 ± 0.009
600	0.144 ± 0.010	0.813 ± 0.023	0.497 ± 0.018	0.098 ± 0.008
800	0.113 ± 0.008	0.636 ± 0.020	0.408 ± 0.016	0.086 ± 0.007
1000	0.081 ± 0.007	0.488 ± 0.017	0.298 ± 0.013	0.070 ± 0.006
1200	0.060 ± 0.006	0.402 ± 0.016	0.238 ± 0.012	0.0335 ± 0.0046
10000	0.0143 ± 0.0020	0.106 ± 0.005	0.0311 ± 0.0030	0.0050 ± 0.0012

Table 8.17: Signal yield prediction (pre-fit) in bins A, B, C, and D estimated from signal MC after event selection for GMSB SPS8 $\Lambda = 400\,\mathrm{TeV}$ and varying $c\tau$. See Table 8.9 for the t_γ-p_T^{miss} splits for each signal point.

$c\tau$ (cm)	Yield in Bin A	Yield in Bin B	Yield in Bin C	Yield in Bin D
10	0.494 ± 0.011	1.392 ± 0.020	3.826 ± 0.035	1.359 ± 0.020
50	0.610 ± 0.009	4.370 ± 0.025	0.710 ± 0.009	0.0948 ± 0.0036
100	0.420 ± 0.009	2.650 ± 0.023	0.820 ± 0.012	0.120 ± 0.0049
200	0.223 ± 0.007	1.367 ± 0.020	0.617 ± 0.013	0.104 ± 0.005
400	0.091 ± 0.005	0.638 ± 0.013	0.363 ± 0.010	0.0619 ± 0.0041
600	0.0557 ± 0.0039	0.404 ± 0.010	0.241 ± 0.008	0.0375 ± 0.0032
800	0.0359 ± 0.0031	0.281 ± 0.008	0.183 ± 0.007	0.0278 ± 0.0027
1000	0.0311 ± 0.0031	0.232 ± 0.008	0.149 ± 0.006	0.0242 ± 0.0027
1200	0.0254 ± 0.0026	0.196 ± 0.007	0.123 ± 0.005	0.0131 ± 0.0019
10000	0.0048 ± 0.0008	0.0414 ± 0.0023	0.0174 ± 0.0015	0.0020 ± 0.0005

Table 8.18: Observed number of events (N_{obs}) and predicted background yields from the background-only fit ($N_{bkg}^{post-fit}$) in bins A, B, C, and D in data for the different t_γ and p_T^{miss} bin boundaries summarized in Table 8.9. Uncertainties in the $N_{bkg}^{post-fit}$ values are the postfit uncertainties. The propagation of the systematic uncertainties is handled during the fit, and therefore they are included in the postfit uncertainties.

Bin boundary [t_γ (ns), p_T^{miss} (GeV)]		A	B	C	D
(0, 250)	N_{obs}	16 139	41	62	18 826
	$N_{bkg}^{post-fit}$	16 130 ± 110	47.5 ± 4.8	55.6 ± 5.6	18 830 ± 130
(1.5, 100)	N_{obs}	33 760	1302	1	5
	$N_{bkg}^{post-fit}$	33 760 ± 160	1303 ± 37	0.29 ± 0.28	5.7 ± 2.2
(1.5, 150)	N_{obs}	34 595	467	0	6
	$N_{bkg}^{post-fit}$	34 600 ± 170	467 ± 22	0.08 ± 0.08	5.9 ± 2.3
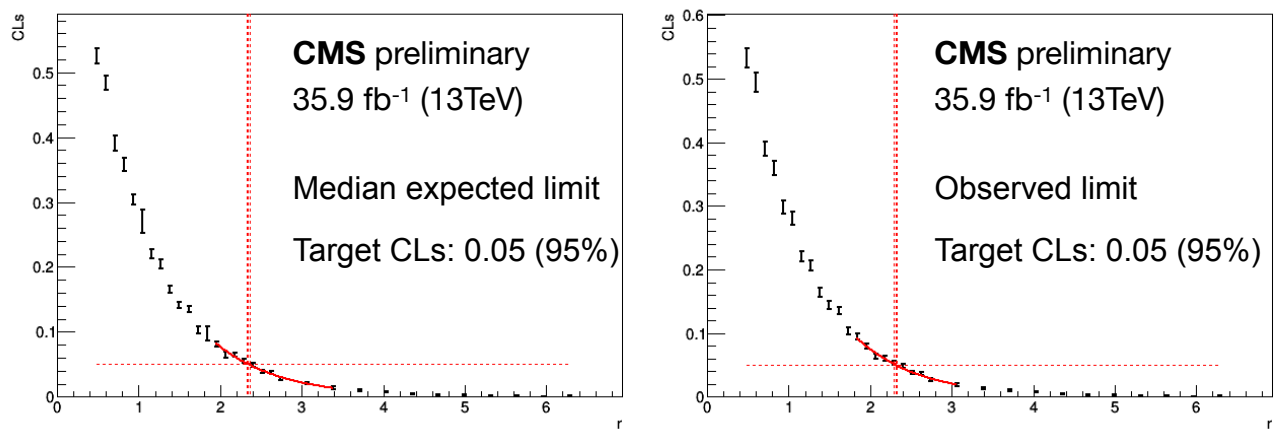
Figure 8.16: CL_s as a function of signal strength r obtained from toy experiments for the median expected limit (left plot) and the observed limit (right plot) for an example signal model ($\Lambda = 350$TeV and $c\tau = 200$cm). The 95% confidence level upper limit on the signal strength is taken as the x-axis value where the y-axis is equal to 0.05.
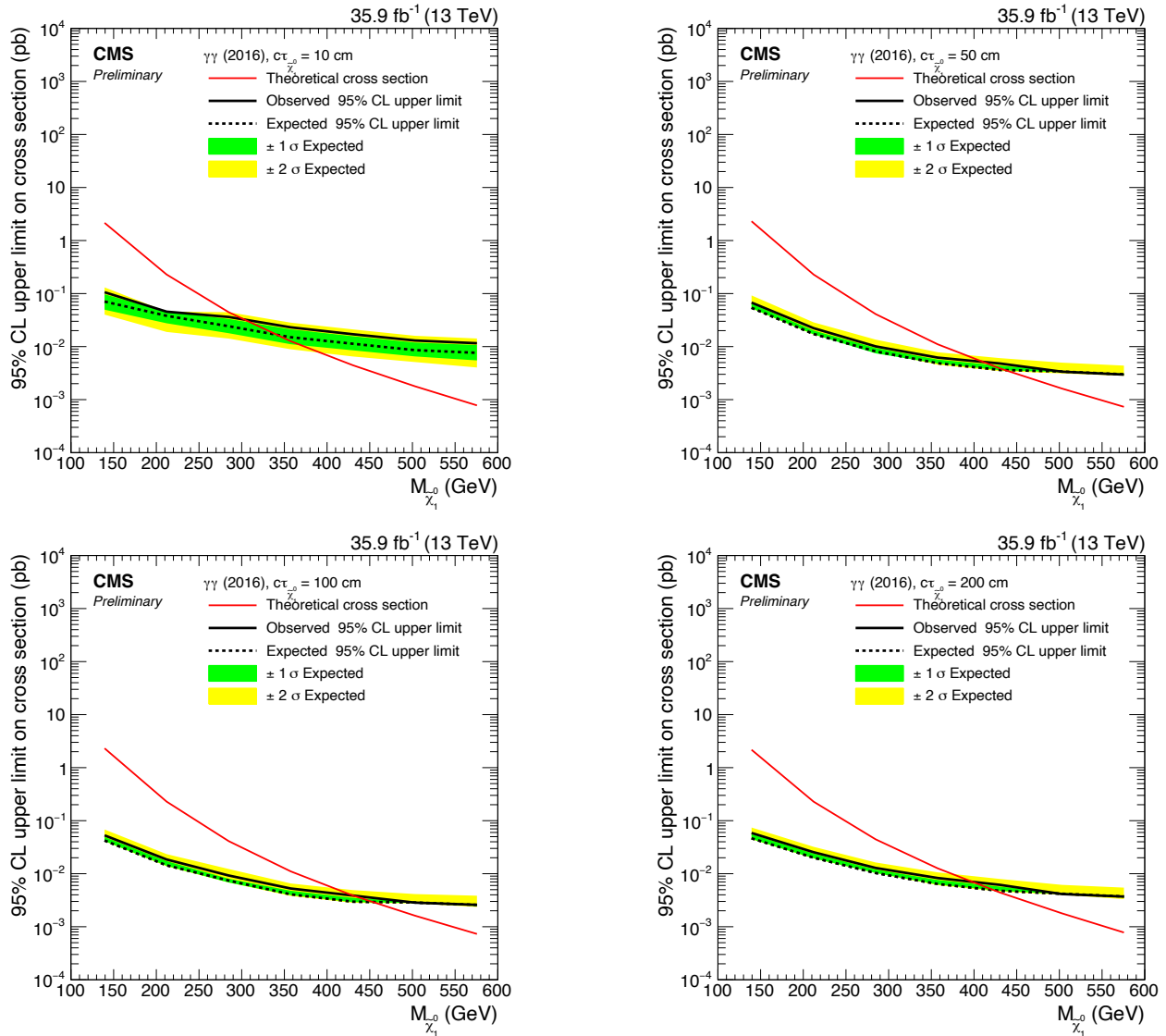
Figure 8.17: The observed and expected 95% confidence level upper limits on the GMSB SPS8 signal production cross section, together with the theoretical signal production cross section, as a function of the neutralino mass, for neutralino signals with different $c\tau$. The corresponding $c\tau$ of the signal for each plot are: 10 cm (top-left), 50 cm (top-right), 100 cm (bottom-left), and 200 cm (bottom-right).
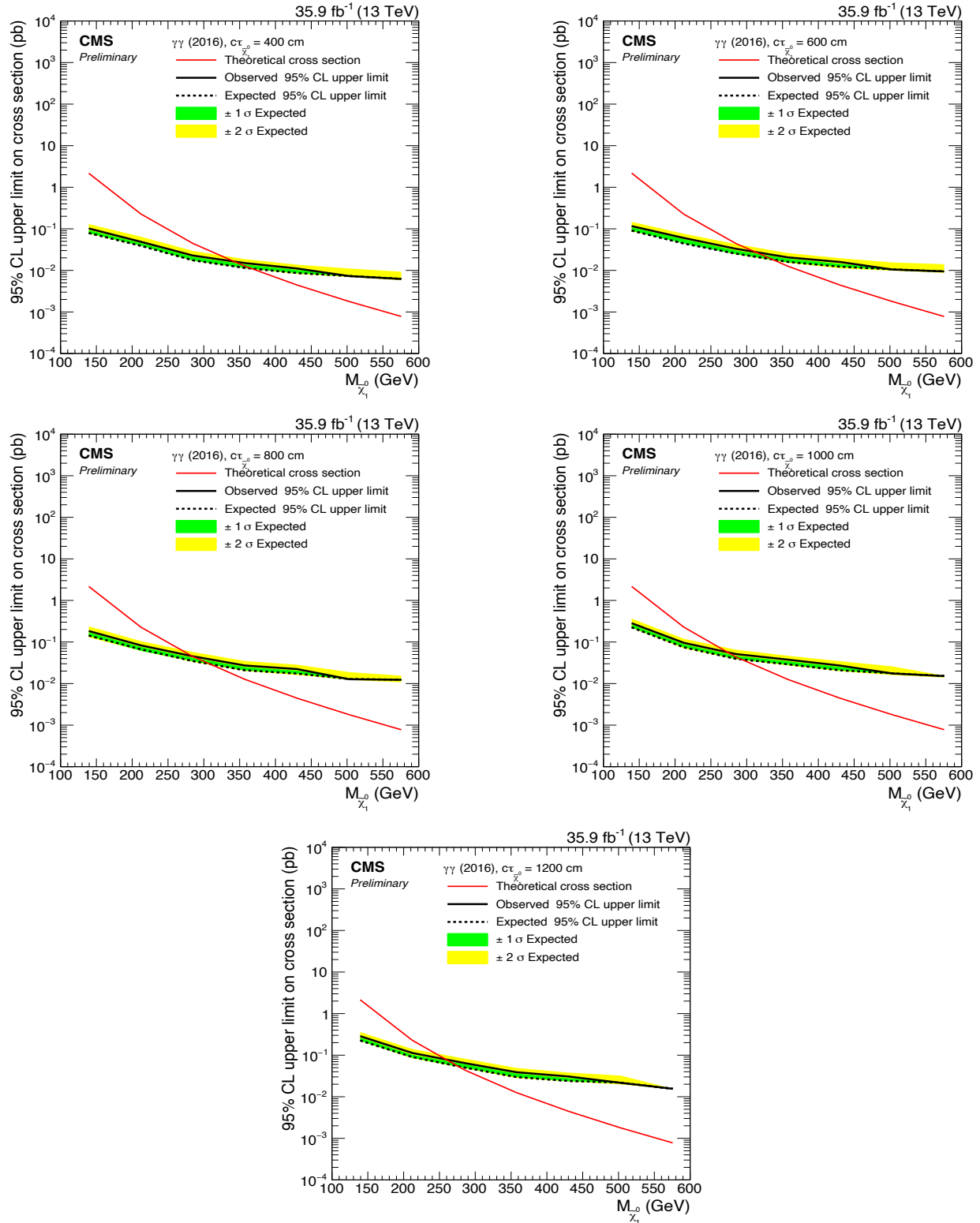
Figure 8.18: The observed and expected 95% confidence level upper limits on the GMSB SPS8 signal production cross section, together with the theoretical signal production cross section, as a function of neutralino mass, for neutralino signals with different $c\tau$. The corresponding $c\tau$ of the signal for each plot are: 400 cm (top-left), 600 cm (top-right), 800 cm (center-left), 1000 cm (center-right), and 1200 cm (bottom).
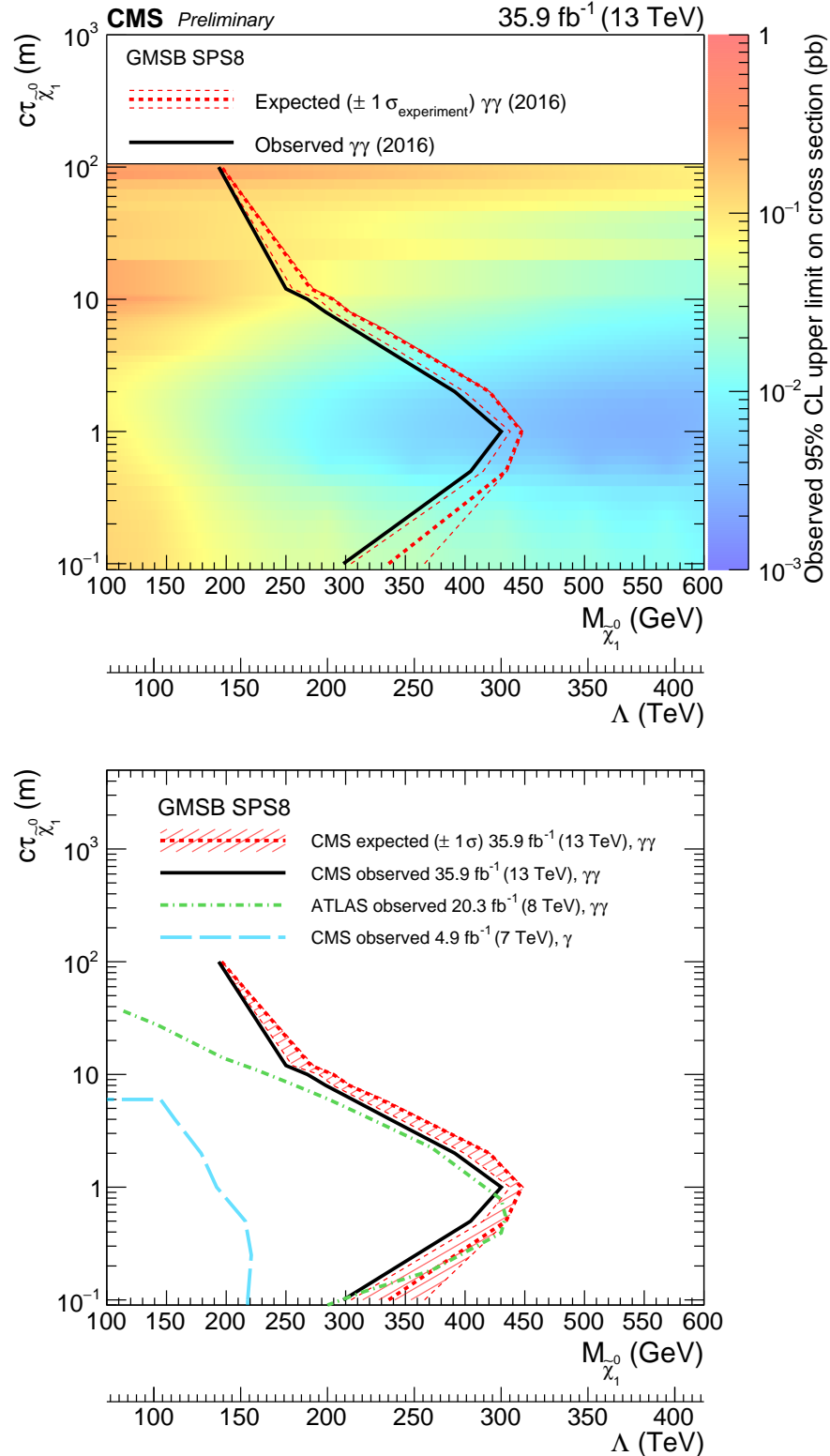
Figure 8.19: Upper plot: the color map shows the observed 95% confidence level upper limit on the signal cross section as a function of the Λ (or neutralino mass) and $c\tau$ of the signal models; the solid and dashed curves show the exclusion boundary in the $c\tau$ versus Λ 2D plane. Lower plot: the exclusion boundary from this search compared to the exclusion boundaries from the previous ATLAS and CMS searches in the $c\tau$ vs Λ 2D plane.

Chapter 9

RESULTS OF A COMBINED SEARCH WITH 2016 AND 2017 DATA

A similar search with 2017 data was performed using a very similar analysis strategy, event selection, and background estimation methods.

In the search with the 2017 data, a dedicated single photon trigger was developed and used. The dedicated trigger selects one photon with $p_T > 60$ GeV, and uses relatively tight cuts on the photon ID in the trigger, including a tighter cut (compared to the 2016 trigger) on the shower shape $\sigma_{i\eta i\eta}$, and additional cuts on the shower shape variables S_{major} and S_{minor}. In addition, the scalar sum of the p_T of all jets (HT) is required to be greater than 350 GeV in the trigger.

Due to the difference in the triggers, slightly different selections are applied to the 2016 and 2017 searches. In 2017, events are divided into 2 categories, the single photon category and diphoton category. The 2017 diphoton category has the same selection on photons as the selection used for the 2016 data: requiring the presence of at least two photons with p_T greater than 70/40 GeV. The 2017 single photon category includes events that do not enter the diphoton category, and the event has to have at least one photon with p_T greater than 70 GeV. In addition, the 2017 photon ID is slightly tighter than the 2016 photon ID, due to tighter cuts applied in the 2017 trigger. In addition, a cut of HT> 400 GeV is also applied to the 2017 events to match the cut in the trigger.

For comparison, Figures 9.1 and 9.2 show the upper limits on the cross section from the different searches and categories: 2016 (diphoton), 2017 single photon, and 2017 diphoton. The corresponding 2D exclusion region from these categories are shown in Figure 9.3. As we can see from those plots, the results from the 2016 search and the 2017 diphoton category are very similar. They are both much better than the 2017 single photon category in the small $c\tau$ region, due to the much better background suppression with the two photon selection. In the larger $c\tau$ region, the single photon category is better than the diphoton category due to the higher acceptance.

These results (2016, 2017 single photon category, 2017 diphoton category) are then combined, with most of the systematic uncertainties (except the uncorrelated trigger

efficiency uncertainty between 2016 and 2017) fully correlated due to the same method being used in the measurements. Figure 9.4 shows the upper limits on the cross section and the exclusion region in the $c\tau$ versus Λ 2D plane. The combination of the 2016 and 2017 searches gives a much better result compared to the previous ATLAS and CMS results, at all values of $c\tau$ and Λ.

In summary, a search for long-lived particles (neutralinos) decaying to delayed photons is performed using pp collision data taken during 2016, corresponding to an integrated luminosity of $35.9\,\mathrm{fb}^{-1}$. The search makes use of CMS ECAL timing measurements, which have a typical time precision of about 100-150 ps for a single high energy crystal hit, and a 300 to 400 ps time resolution for an electron or photon cluster as a whole. Backgrounds are estimated by use of a 2D fit of the photon arrival times at the ECAL t_γ and the missing transverse momentum $p_\mathrm{T}^\mathrm{miss}$, where the delayed photon signal events are enriched in the region with large t_γ and large $p_\mathrm{T}^\mathrm{miss}$. The results are interpreted with a benchmark GMSB SPS8 model, and upper limits on the signal production cross sections are presented. The results are combined with a similar search using 2017 data ($41.5\,\mathrm{fb}^{-1}$), which uses a different set of triggers but a very similar analysis strategy, event selection, and background estimation methods. Neutralinos with proper decay lengths of 0.1, 1, 10, and 100 m and masses up to 320, 525, 360, and 215 GeV are excluded at 95% confidence level with the combined result, respectively. We extend the previous best limits in the neutralino proper decay length by up to one order of magnitude, and in the neutralino mass by up to 100 GeV.
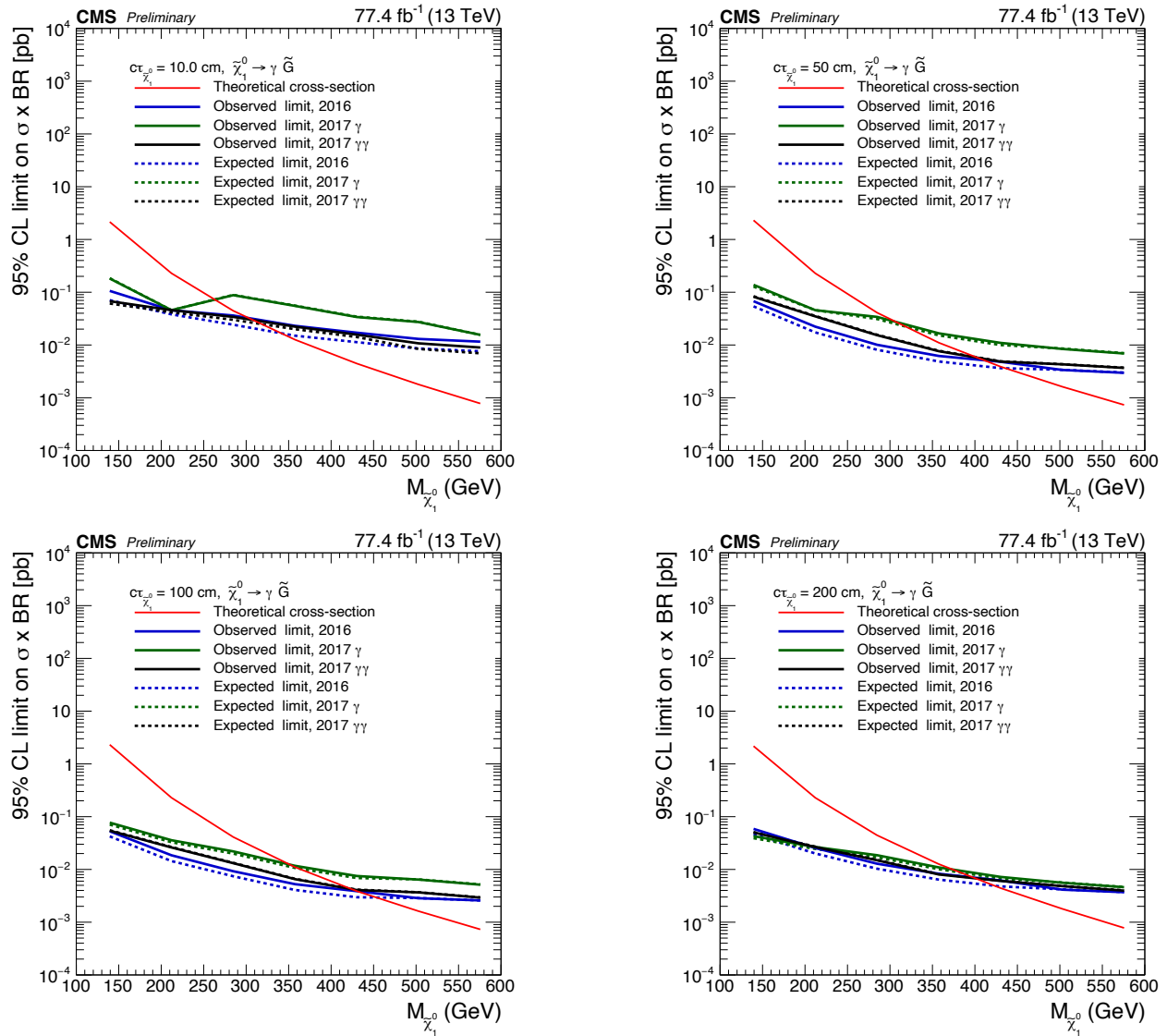
Figure 9.1: The observed and expected 95% CL limits on the GMSB SPS8 signal production cross section, together with the theoretical signal production cross section, as a function of neutralino mass, for neutralino signals with different $c\tau$. The corresponding $c\tau$ of the signal for each plot are: 10 cm (top-left), 50 cm (top-right), 100 cm (bottom-left), and 200 cm (bottom-right). Results from the 2016, 2017 single photon category (γ), and the 2017 diphoton category ($\gamma\gamma$) are shown separately in each of the plots for comparison.
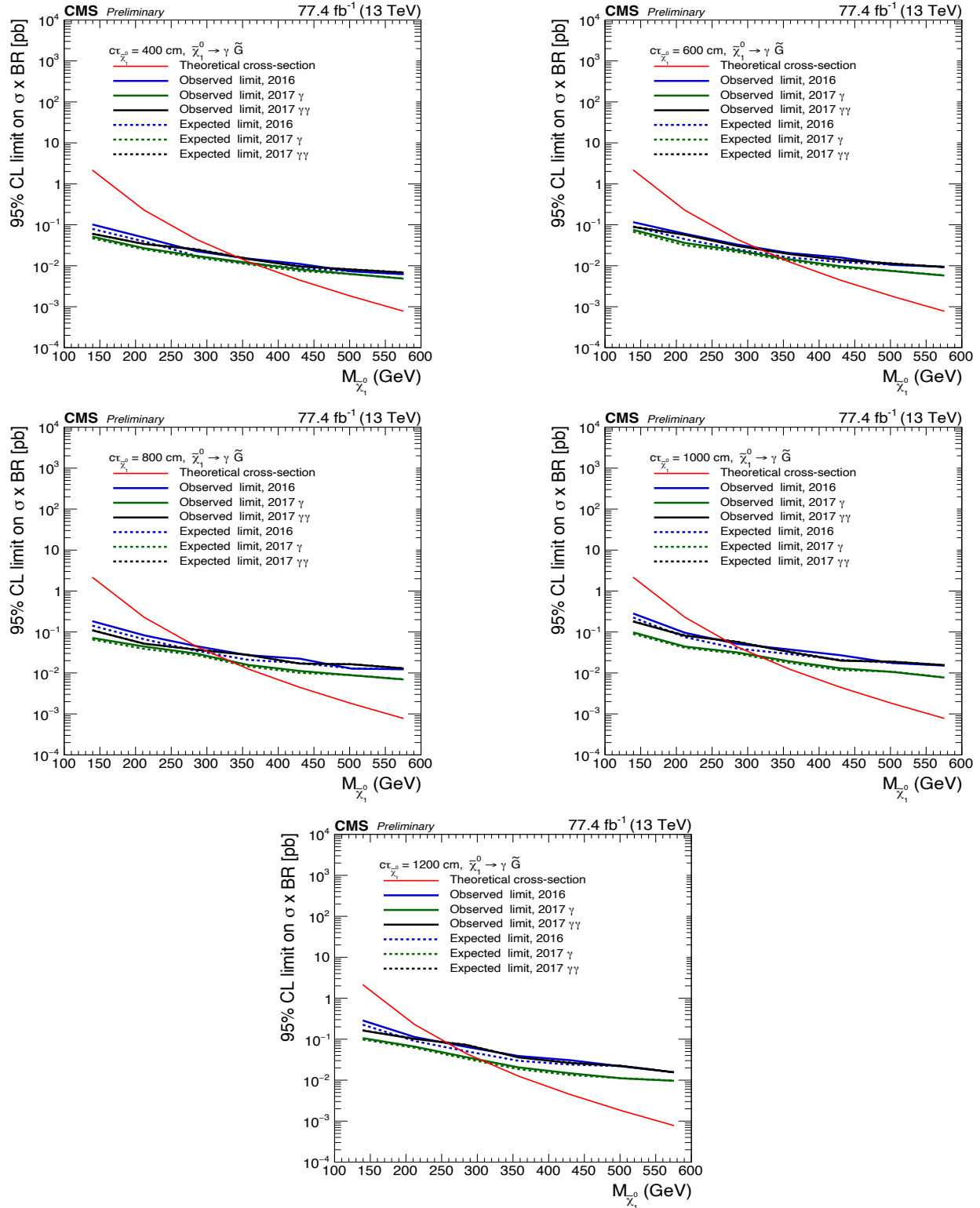
Figure 9.2: The observed and expected 95% CL limits on the GMSB SPS8 signal production cross section, together with the theoretical signal production cross section, as a function of neutralino mass, for neutralino signals with different $c\tau$. The corresponding $c\tau$ of the signal for each plot are: 400 cm (top-left), 600 cm (top-right), 800 cm (center-left), 1000 cm (center-right), and 1200 cm (bottom). Results from the 2016, 2017 single photon category (γ), and the 2017 diphoton category ($\gamma\gamma$) are shown separately in each of the plots for comparison.
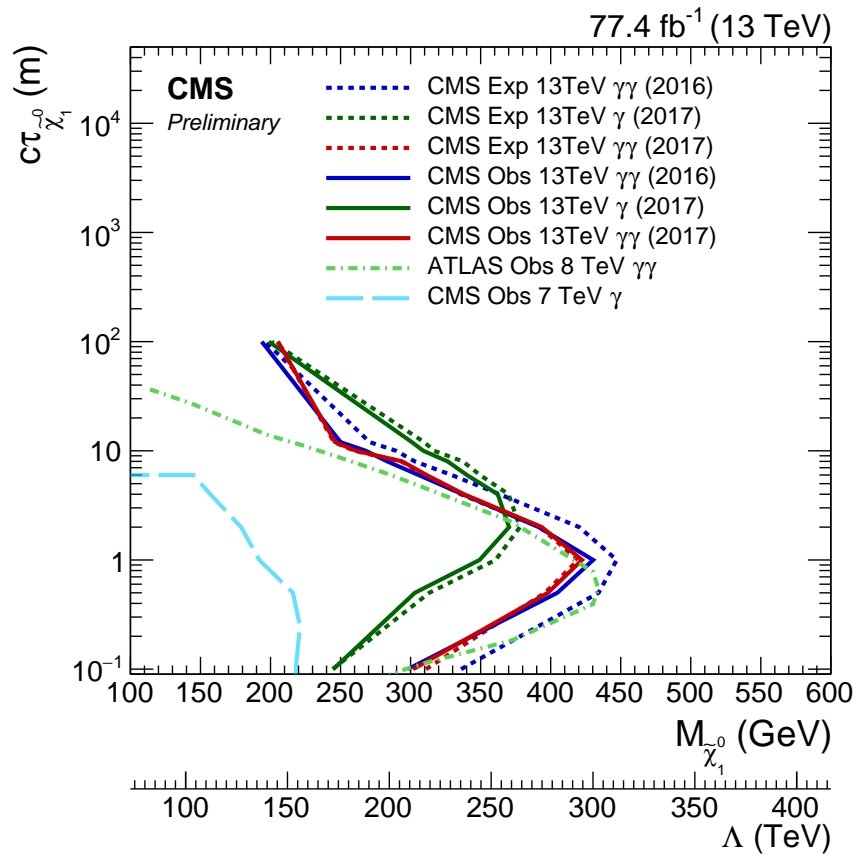
Figure 9.3: The exclusion boundary in the Λ and $c\tau$ 2D plane from various searches: 2016 (diphoton), 2017 single photon category, 2017 diphoton category, and the previous ATLAS and CMS searches.
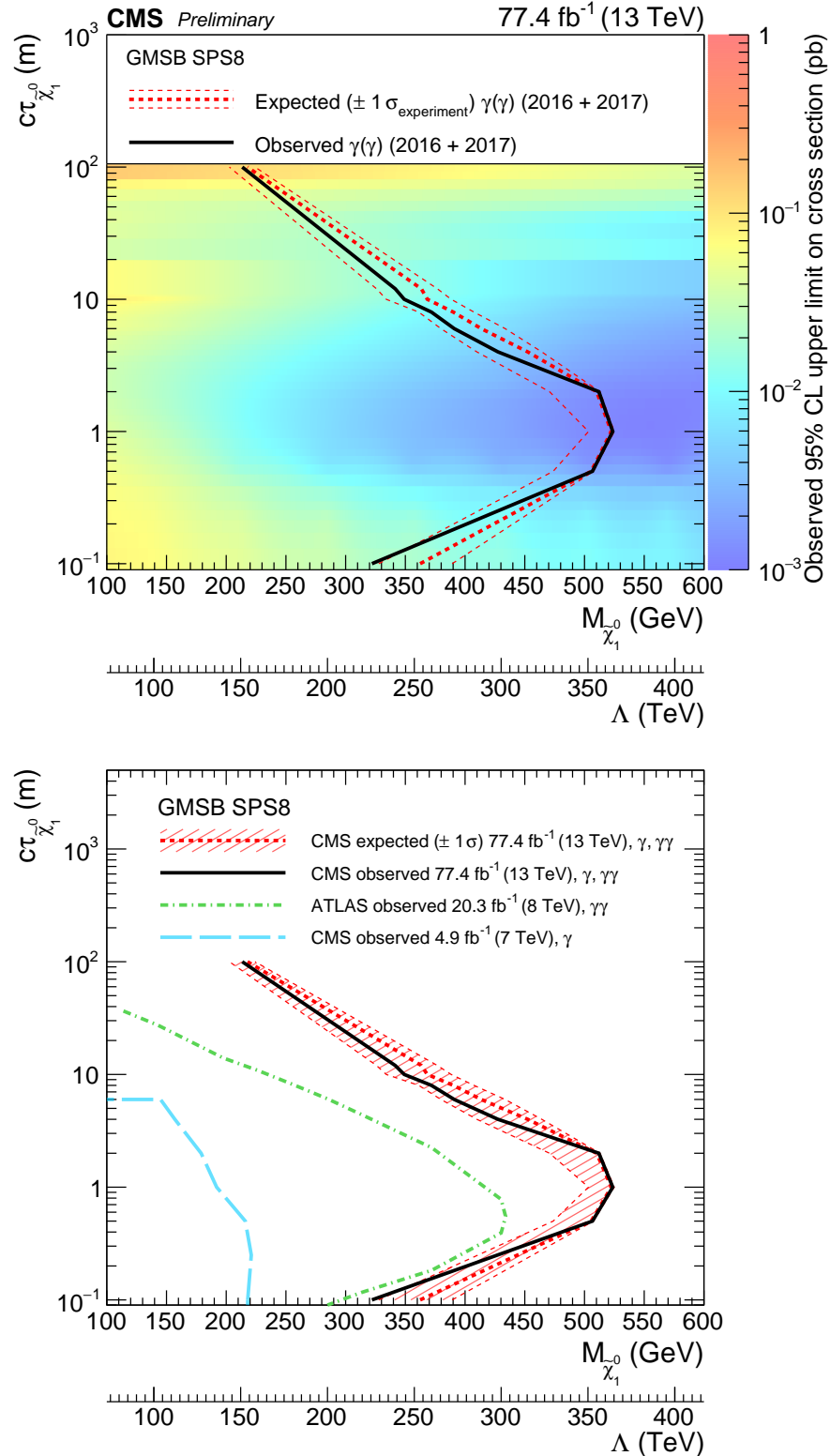
Figure 9.4: Upper plot: the color map shows the observed 95% confidence level upper limit on the signal cross section as a function of the Λ (or neutralino mass) and $c\tau$ of the signal models; the solid and dashed curves show the exclusion boundary in the the 2D plane of Λ and $c\tau$. The results are from combined 2016 and 2017 searches. Lower plot: the exclusion boundary from this search compared to the exclusion boundaries from previous ATLAS and CMS searches in the $c\tau$ versus Λ 2D plane.

C h a p t e r   10

TOWARDS FUTURE SEARCHES

An on-going analysis with the 2018 data that uses a similar strategy is being performed at the time of this thesis writing. As this analysis is completed and when more data is taken in CMS at LHC Run3, we can expect forthcoming improved limits, especially for models with large $c\tau$.

As a simple estimate of the expected improvement, if we assume the same selection efficiency and detector performance (i.e. timing resolution, etc.) in 2018 and during Run 3 as in 2017, and if we simply scale the event yields in different bins A, B, C, and D by the corresponding luminosity, we can get the estimated upper limits on the signal strength r for different luminosity scenarios, as shown in Table 10.1 for a few representative signal models.

We can see from the table that with a total of 300 fb^{-1} from Run 2 and Run 3 expected by 2024, for the most sensitive signals with $c\tau$ around 2 m, we can expect to exclude signal models with Λ beyond 400 TeV, which corresponds to a limit on the neutralino mass of 580 GeV. Similar (but a bit smaller) improvements are also expected for signal models with smaller $c\tau$. .

Table 10.1: Expected 95% confidence level upper limits on the signal strength r from the 2016 and 2017 analyses, and simple estimates of the expected upper limits if we add in the 2018 and Run 3 data, assuming the same event selection efficiency and timing performance as in 2017.

signal model		2016	2017	2016-17	2018	2016-18	2016-18-Run 3
Λ (TeV)	$c\tau$ (m)	35.9 fb^{-1}	41.5 fb^{-1}	77.4 fb^{-1}	59.7 fb^{-1}	137.1 fb^{-1}	300 fb^{-1}
350	2	1.7	1.2	0.8	0.9	0.5	0.3
400	2	3.5	2.3	1.5	1.7	0.9	0.5
350	0.1	6.2	3.3	3.0	2.7	2.0	1.2

As the photon arrival time at ECAL t_γ is the most important discriminant in this search, an improved ECAL time measurement is expected to improve the sensitivity significantly, especially for small $c\tau$ signal models where the t_γ for signal event is very small and is more difficult to be distinguished from prompt photon backgrounds.

As discussed in Section 8.3, the majority of the uncertainties in the CMS ECAL time measurement comes from the electronics (including readout unit and clock jitter).

The crystals themselves have great potential for improved timing measurements due to their fast rise time and large amplitude. During the Phase-II upgrade, designed for the High Luminosity LHC that is scheduled to start running in 2028, all the CMS ECAL readout electronics will be replaced by a system that samples the pulse at a much higher rate (from the current 40 MHz to 160 MHz) [93], which means that there will be more samples for each pulse to be used in the timing measurement. The ECAL clock will also be upgraded such that the jitter will be below the 10 ps level. With these upgrades, the intrinsic time resolution for high energy electrons/photons is expected to be around 30 ps; an order of magnitude better than the ECAL readout that was used in this analysis (and which will also continue to be used during the LHC Run 3 in 2022-24).

With the Phase-II ECAL upgrade, we expect to extend our reach to smaller $c\tau$ signal models, and exclude such models for much larger Λ values. Preliminary studies [155] have shown that the exclusion boundary of Λ is expected to reach around 500 TeV for $c\tau$ down to 0.1 cm with $1000\,\mathrm{fb}^{-1}$ of data with the upgraded ECAL, which is much larger than the current reach of approximately 200 TeV.

However, even with the upgraded ECAL, we will still not have a measurement of the beam spot timestamp, which has a time spread of about 200 ps as discussed in Section 8.3, and this could dominate the total electron/photon time resolution measurement. Fortunately, as part of the CMS Phase-II upgrade, the new precision MIP Timing Detector (MTD) which is now under development as discussed in Chapter 5, will be installed between the ECAL and tracker, and is expected to provide timestamps of charged particles with 30 ps resolution [155, 156].

With the timestamps of all (or most) charged particles measured, the primary vertex time can then be reconstructed with a projected resolution at the 30 ps level. In such cases, the overall time resolution of electrons/photons will also be at the 30 ps level, which will make this search sensitive to even smaller $c\tau$ models. Preliminary studies [155] have shown that the exclusion boundary of Λ is expected to be around 700 TeV for $c\tau$ down to 0.1 cm with $1000\,\mathrm{fb}^{-1}$ data with the MTD, which would not be possible with the upgraded ECAL alone. Figure 10.1, which is taken from Ref. [155], shows a summary of the expected exclusion regions with different detector upgrade configurations, as discussed above.

Besides the above discussed improvements for better timing measurements, the sensitivity of this search can also be improved with better analysis techniques. First of all, the current photon ID as discussed in Section 8.2 was optimized mainly

against fake photons from QCD, and this led to having many prompt photons passing the photon ID requirement. Several shower shape variables can be used to discriminate the delayed photon against prompt photons, such as $\sigma_{i\eta i\eta}$, S_{major}, and S_{minor}. Currently the photon ID only applies simple cuts on those variables, which is not very effective in removing the majority of the prompt photon backgrounds.

A preliminary study of an improved photon ID using machine learning techniques with those variables as input for training has shown that the final limit can be improved substantially (around a 25% improvement). In addition, further improvement can be gained with improved triggers. Currently for 2017, the diphoton category uses a single photon trigger with relatively tight photon ID cuts, requiring the same tight cuts to be applied offline for the two photon category. If instead the diphoton category were to use a diphoton trigger which has no or very loose photon ID cuts in the trigger to increase the signal yields in the diphoton category, preliminary studies have shown that about a 30% improvement in the limits on Λ, and therefore on the neutralino mass could be expected.
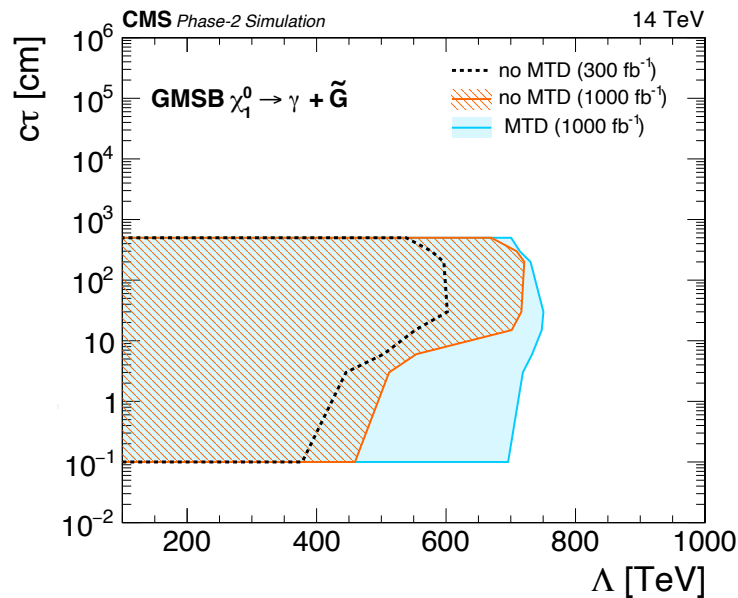
Figure 10.1: Exclusion region in the $c\tau$ versus Λ 2D plane with different detector upgrade and luminosity scenarios: the dashed black line shows the expected result with only the ECAL upgrade and $300\,\text{fb}^{-1}$ of data; the orange line shows the expected result with only the ECAL upgrade with $1000\,\text{fb}^{-1}$ of data; the blue line shows the expected result with the ECAL upgrade and the MTD detector installed, with $1000\,\text{fb}^{-1}$ of data. The plot is taken from Ref. [155].

Part V

Conclusion

Chapter 11

CONCLUSION

In this thesis, we first presented the power of the standard model of particle physics in describing the fundamental particles and their interactions (weak, strong, electromagnetic) and predicting the experimental observations. Among the predictions, the production of three massive vector bosons is a good test of the non-Abelian structure of the gauge field as well as the couplings between the Higgs boson and gauge bosons.

We also presented the motivation for new physics beyond the standard model, in particular a supersymmetric extension of the standard model which solves the hierarchy problem and also includes a good candidate for dark matter. This extension predicts new particles with masses between the GeV and TeV scale and motivates a wide range of searches at the LHC.

We then introduced the LHC machine which delivers proton-proton collisions at $\sqrt{s}$ = 13 TeV with high luminosity. We also reviewed the CMS detector which is designed as a compact high field general purpose detector, with excellent tracking for charged particles and energy measurements for neutral and charged particles. In particular, we presented details of the energy reconstruction and calibration of the electromagnetic calorimeter, with a focus on the inter-calibration of different channels with $\pi^0 \to \gamma\gamma$ events which is crucial to achieve an energy resolution of the calorimeter at about the 2% level.

We then presented in detail the first observation of a rare electroweak process: the production of three massive vector bosons (VVV, with $V = W, Z$) in proton-proton collisions at $\sqrt{s}$ = 13 TeV. The analysis presented in this thesis led to the first observation of a combined VVV production with an observed (expected) significance of 5.7 (5.9) standard deviations. This observation opens a new program of rich physics with VVV processes, including the search for anomalous triple-gauge or quartic-gauge couplings in the context of an effective field theory, and the prospect of constraining the gauge-gauge and Higgs-gauge couplings in the same final state, as well as direct searches for new two- or three-boson resonances with heavy mediators.

We also presented a search for long-lived particles that decay to photons inside the

CMS ECAL region. The timing measurement of the CMS ECAL is employed to discriminate the time delayed signal photons from prompt photon backgrounds. A dedicated identification and the timing measurement of such photons are presented in this thesis. The results are then interpreted with a gauge-mediated supersymmetry breaking model, where the neutralino is the long-lived particle that decays into a photon plus a gravitino. The limits on the neutralino's mass and proper decay length presented in this thesis extend the previous limits by an order of magnitude in the neutralino's proper decay length, and by about 100 GeV in its mass.

Both of these major physics investigations will continue to advance at CMS as the LHC physics program continues with Run 3 in 2022-24, and at the HL-LHC scheduled to begin in 2028.

As many long-lived particle searches need a precision timing detector for both charged and neutral particles, especially in the challenging high pileup environment of colliders such as the HL-LHC, and since physics object reconstruction will also benefit from additional timing information of the particle, we have reviewed a new detector for the CMS Phase-2 upgrade of the ECAL, the MIP Timing Detector (MTD), and in particular we have presented the test beam and GEANT4 simulation studies of the timing performance of the barrel layer (LYSO+SiPM) sensors with different sensor sizes, which provided key input for the final design choice for the timing layer.

In addition, in the appendix, we have presented studies of the timing performance of a Cadmium-Telluride sensor in electromagnetic showers, which has been shown to be a good candidate for the active material in a sampling calorimeter for experiments at a future collider. Overall, it has been demonstrated with test beam data that a time resolution at and below 30 ps is achievable with various sensor choices, which can play a key role in opening new chapters of physics investigations both at the HL-LHC and at other future colliders.

Part VI

Appendix

A p p e n d i x A

PRECISION TIMING DETECTORS WITH
CADMIUM-TELLURIDE SENSORS

Calorimeters with excellent timing resolution have been a focal point of recent research and development, among which scintillating crystals as well as silicon-based sampling calorimeters have been the design choices of greatest interest.

Silicon-based sensors have very high sensitivity to charged particles in an electromagnetic (EM) shower, but their sensitivity to low energy (keV range) photons which compose the majority of the shower is limited. This has motivated the consideration of alternative sensors with higher atomic numbers whose higher sensitivity to low energy photons could possibly provide better performance, both for energy and timing measurements.

In this appendix, test beam measurements of EM showers with a Cadmium-Telluride (CdTe) sensor are presented. As shown, a time resolution of 25 picoseconds (ps) and an energy resolution of 18% have been achieved with a single CdTe sensor placed near the peak of an EM shower initiated by high energy electrons.

The measurements were performed at two CERN test beam facilities: the East-Area T9 beam line which provided electrons of energy in the range of 2-7 GeV, and the North-Area H2 beam line which provided 50-200 GeV electrons. The beams are composed of a mixture of electrons and pions, with an electron fraction of about 75% in the H2 beam line and about 10% in the T9 line.

Similar test setups were used in the two beam lines. Figure A.1 shows the setup at the T9 beam line, together with a photo of the CdTe sensor being tested. The CdTe sensor under test was a $10 \times 10 \times 1 \, \text{mm}^3$ Schottky type diode purchased from Acrorad [157]. The CdTe sensor was placed inside a small copper box with 0.3 mm thick walls, and was operated at a bias voltage of 700 V. The dark current was measured to be 3 nA to 6nA depending on the environment of the experimental area. Wire chambers upstream of the test setup (not shown in the figure) were used to measure the horizontal and vertical beam position. A micro-channel plate photomultiplier (MCP-PMT) with a time resolution measured in situ of less that 10 ps was placed downstream of the wire chamber to serve as a time reference. The time reference detector used at the T9 beamline was a Hamamatsu R3809U

MCP-PMT, while in the H2 beamline a Photek 240 MCP-PMT was used. As the Photek 240 MCP-PMT contains a significant amount of material (about $1.8X_0$), it was placed downstream of the CdTe sensor at H2.

The EM shower was produced by a tungsten and lead absorber $6X_0$ thick in the H2 setup, and by a $2X_0$ thick absorber in the T9 setup. At T9, since the electron fraction in the beam is relatively low, an LYSO crystal together with an MCP-PMT placed downstream of the CdTe sensor was used to select electrons. The setup was triggered by a $1 \times 1\,\mathrm{cm}^2$ plastic scintillator coupled to a photomultiplier in the T9 beamline, and by a $4 \times 4\,\mathrm{cm}^2$ plastic scintillator at H2.
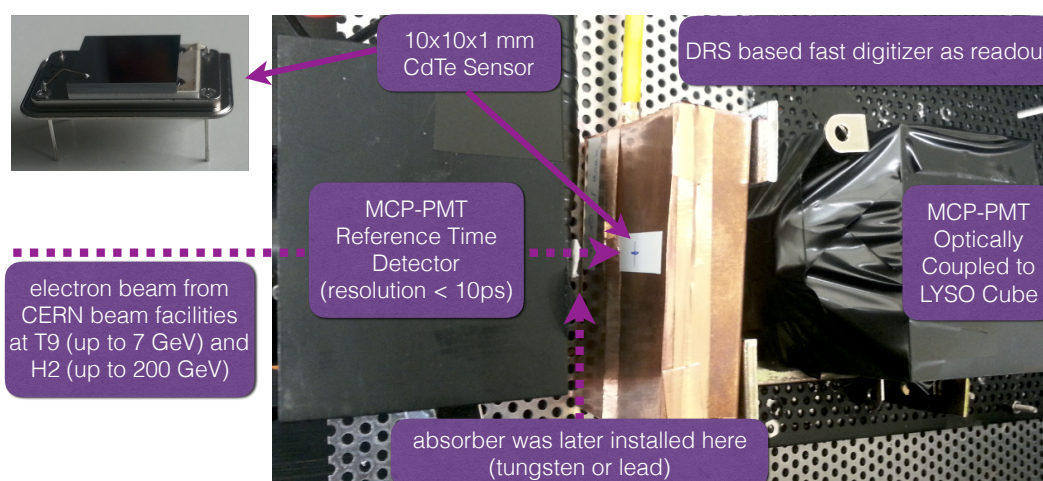
Figure A.1: Experimental test setup for the CdTe sensor at the CERN T9 beam line.

The signal from the CdTe sensor was amplified by a Hamamatsu C5594 amplifier with a bandwidth of 1.5 GHz and a gain of 36 dB. For the test at H2, where the beam energy was much higher, a 10 dB attenuator was connected to the output signal of the CdTe sensor to fit the signal within the dynamic range of the amplifier. All signals were digitized by a CAEN V1742 digitizer [84] at a sampling rate of 5 GHz. The baseline pedestal of each channel of the digitizer was measured by the time samples outside of the signal window, and was subtracted from the signal pulses. An example pulse from the CdTe sensor from a 100 GeV electron beam is shown in Figure A.2.

The total charge from the CdTe channel was obtained by integrating the pulse over the entire 200 ns time window recorded by the digitizer. Figure A.3 shows the distribution of the total charge for a 2 GeV electron beam with a $2X_0$ absorber (at T9) and a 100 GeV electron beam with a $6X_0$ absorber (at H2) in front of the CdTe
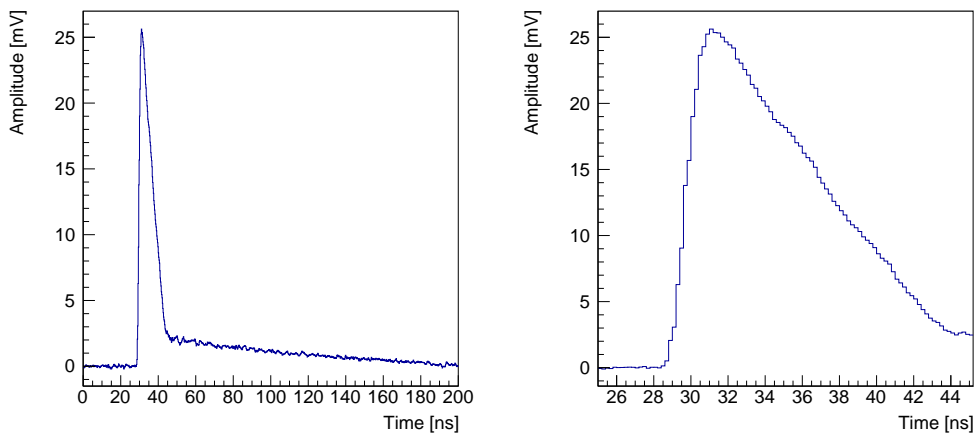
Figure A.2: An example signal pulse from the CdTe sensor for EM showers produced by a $6X_0$ tungsten absorber in a 100 GeV electron beam . The left plot is the pulse in the entire 200 ns window from the digitizer, and the right plot is the zoom in of the early part of the pulse.

sensor. The two distributions were fitted with a gaussian function. For the 100 GeV electron beam, the standard deviation from the fit is about 18% of the mean of the total charge. Figure A.4 shows the fitted mean charge as a function of the electron beam energy. We can see that the charge scales up as the beam energy increases. The result of the resolution of the energy measurement is also encouraging as the measurement is performed using only a single small sensor in one single layer. Further study of the longitudinal shower profile and having more sensors covering the transverse area of the shower is expected to give much better energy resolution.

The timestamp of the shower hitting the CdTe sensor was reconstructed as the time at which the pulse reached 30% of the pulse maximum amplitude. This timestamp is compared to the reference time measured by the MCP-PMT. Figure A.5 shows an example distribution of the time difference between the digitized CdTe output pulse and the MCP-PMT pulse. A resolution of 45 ps is achieved in this case without any corrections applied.

In order to improve the time resolution, the timestamp from the CdTe channel is corrected based on the amplitude of the pulse, as illustrated in Figure A.6. In addition, we also observed that the timestamp has a strong dependence on the position of the beam particle, as shown in Figure A.7. The signal from the CdTe channel arrives about 100 ps later when the beam position is far away from the anode wire bond position compared to the case when the beam position is close to
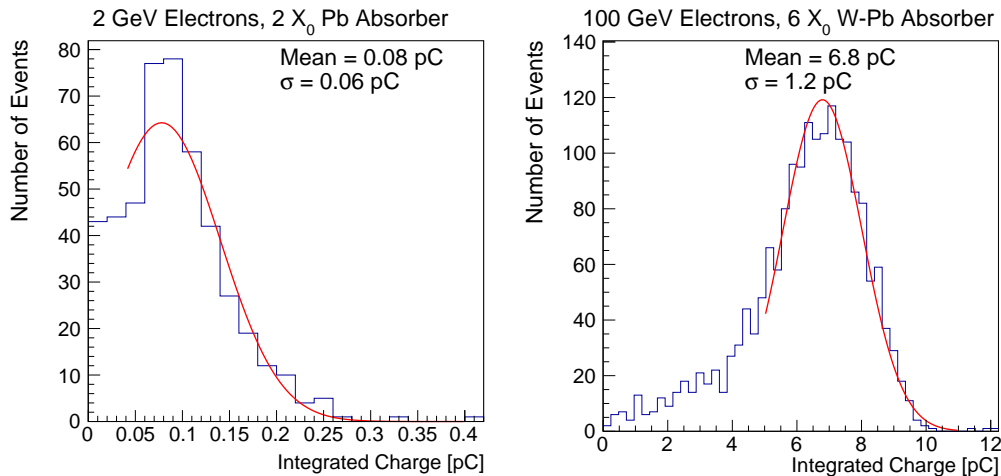
Figure A.3: Distribution of the total charges collected in the CdTe sensor, for a 2 GeV electron beam with a $2X_0$ lead absorber (left) and for a 100 GeV electron beam with a $6X_0$ tungsten and lead absorber (right) placed in front of the sensor.
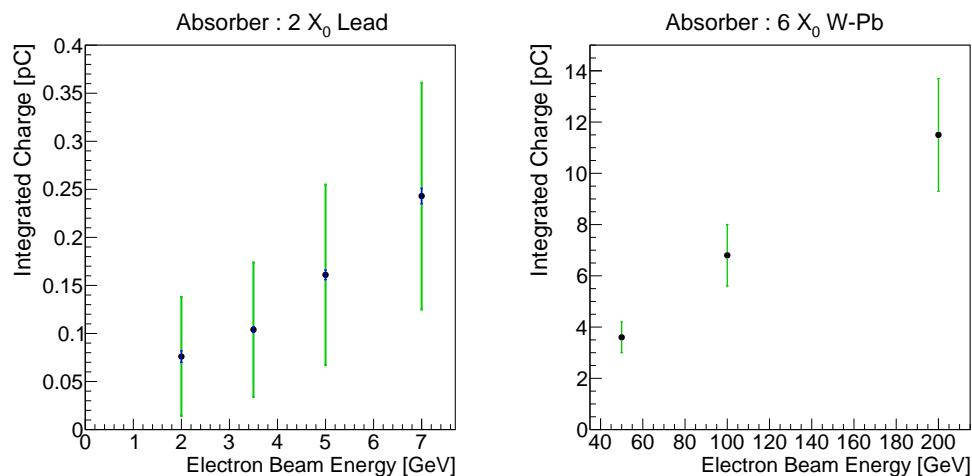
Figure A.4: The mean collected charge from the CdTe sensor as a function of the electron beam energy, for 2 - 7 GeV beams with a $2X_0$ absorber in front of the sensor on the left, and for 50 - 200 GeV beams with a $6X_0$ absorber on the right. The green bars show the resolution of the charge measurement at each beam energy.

the wire bond. Therefore, a position correction on the timestamp is also applied based on Figure A.7.

After the amplitude correction and beam position correction, the time resolution for a 100 GeV incident electron is improved from 45 ps to 25 ps, as shown in Figure A.8. Similarly to what has been observed with the LYSO+SiPM sensor, the time resolution also depends on the position of the beam. In this case it is related
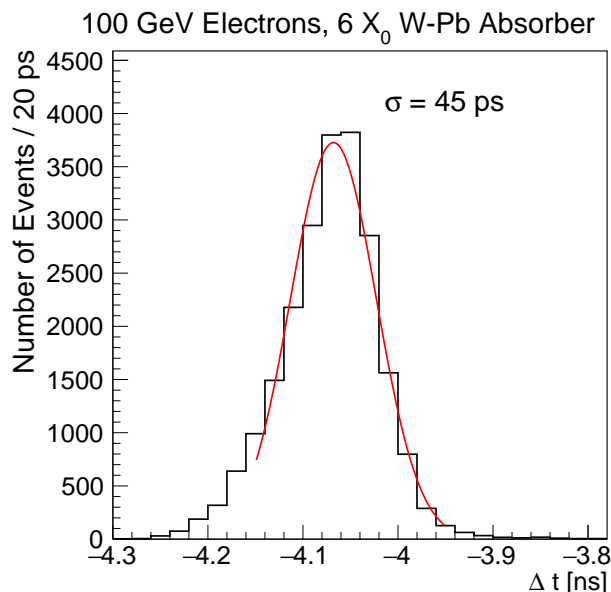
Figure A.5: Distribution of the time difference between the CdTe digitized pulse channel and the MCP-PMT pulse for a 100 GeV incident electron beam with a $6X_0$ absorber. The distribution is fitted with a gaussian function. The standard deviation of the fitted gaussian is 45 ps.

to the distance of the beam position to the anode wire bond in the CdTe sensor. Figure A.9 shows the resolution dependence on the distance. The best resolution is around 20 ps for the beam position near the anode wire bond in the sensor.

In summary, the energy and timestamp of an EM shower has been measured with a $10 \times 10 \times 1\,\text{mm}^3$ CdTe sensor. An energy resolution of about 18% was achieved with a single CdTe sensor behind a $6X_0$ absorber for an incident electron beam of 100 GeV. The time resolution has been measured to be 20-33 ps across the entire sensor, where the best case (20 ps) is obtained when the beam position is close to the anode wire bond of the sensor.

The excellent time and energy resolution obtained with a single sensor in an EM shower in this study indicate that CdTe could be a good choice for the active material of a sampling calorimeter for future collider experiments.
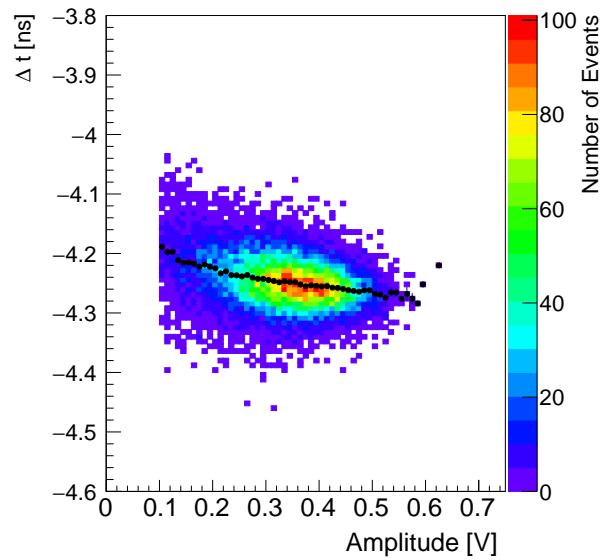
Figure A.6: The dependence of the timestamp on the amplitude of the pulse for a 100 GeV electron beam with a $6X_0$ absorber. A mild dependence is seen, therefore a correction on the timestamp based on the amplitude is applied to correct for this dependence.
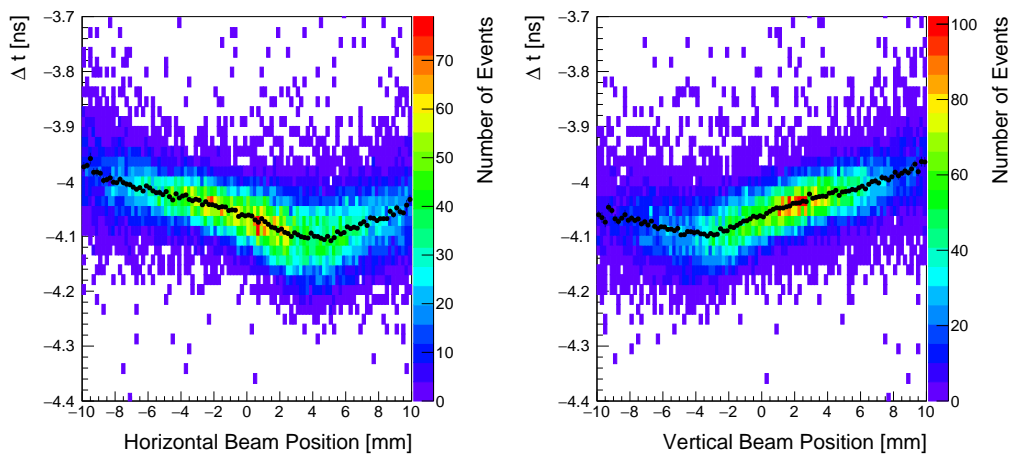
Figure A.7: The dependence of the timestamp on the beam particle position. The left plot shows the dependence on the horizontal position and the right plot shows the dependence on the vertical position of the beam.
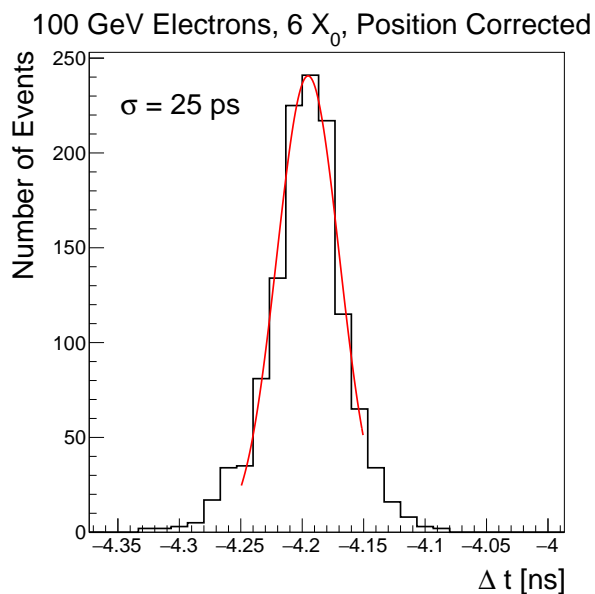
Figure A.8: Distribution of the time difference from CdTe channel and MCP-PMT channel for 100 GeV electron beam with a $6X_0$ absorber after the amplitude correction and beam position correction. A 25 ps resolution is achieved.
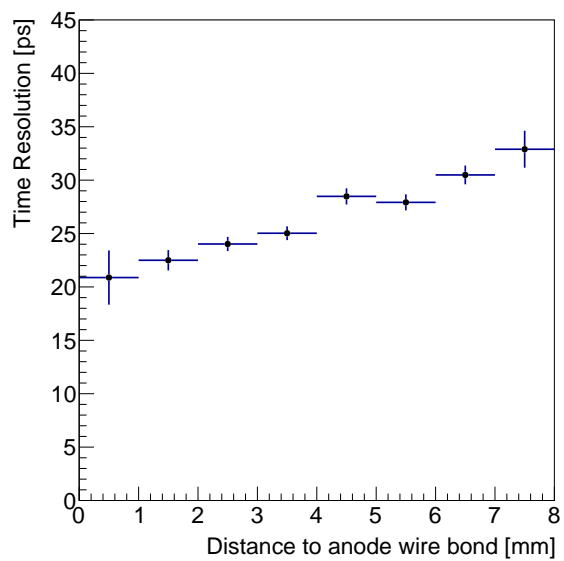
Figure A.9: The time resolution of the CdTe sensor as a function of the distance from the beam position to the wire bond location on the sensor.

BIBLIOGRAPHY

[1] CMS Collaboration. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC." In: *Physics Letters B* 716.1 (Sept. 2012), pp. 30–61. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex]. URL: http://dx.doi.org/10.1016/j.physletb.2012.08.021.

[2] ATLAS Collaboration. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC." In: *Physics Letters B* 716.1 (Sept. 2012), pp. 1–29. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex]. URL: http://dx.doi.org/10.1016/j.physletb.2012.08.020.

[3] CMS Collaboration. "A measurement of the Higgs boson mass in the diphoton decay channel." In: *Physics Letters B* 805 (June 2020), p. 135425. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2020.135425. arXiv: 2002.06398 [hep-ex]. URL: http://dx.doi.org/10.1016/j.physletb.2020.135425.

[4] *The Standard Model of Particle Physics*. URL: https://www.symmetrymagazine.org/standard-model/ (visited on 08/15/2020).

[5] Particle Data Group. "Review of particle physics." In: *Progress of Theoretical and Experimental Physics* 2020.8 (Aug. 2020). 083C01. ISSN: 2050-3911. DOI: 10.1093/ptep/ptaa104. eprint: https://academic.oup.com/ptep/article-pdf/2020/8/083C01/33653179/ptaa104.pdf. URL: https://doi.org/10.1093/ptep/ptaa104.

[6] D. Buttazzo et al. "Investigating the near-criticality of the Higgs boson." In: *Journal of High Energy Physics* 2013.12 (Dec. 2013). ISSN: 1029-8479. DOI: 10.1007/jhep12(2013)089. arXiv: 1307.3536 [hep-ph]. URL: http://dx.doi.org/10.1007/JHEP12(2013)089.

[7] S. P. Martin. "A Supersymmetry Primer." In: *Advanced Series on Directions in High Energy Physics* (July 1998), pp. 1–98. ISSN: 1793-1339. DOI: 10.1142/9789812839657_0001. arXiv: hep-ph/9709356 [hep-ph]. URL: http://dx.doi.org/10.1142/9789812839657_0001.

[8] S. Dimopoulos. "Soft supersymmetry breaking and the supersymmetric standard model." In: *Nuclear Physics B - Proceedings Supplements* 101.1-3 (Aug. 2001), pp. 183–194. ISSN: 0920-5632. DOI: 10.1016/s0920-5632(01)01504-3. arXiv: hep-th/0105034 [hep-th]. URL: http://dx.doi.org/10.1016/S0920-5632(01)01504-3.

[9] S. Chang et al. "Phenomenology of induced electroweak symmetry breaking." In: *Journal of High Energy Physics* 2015.3 (Mar. 2015). ISSN: 1029-

8479. DOI: 10.1007/jhep03(2015)017. arXiv: 1411.6023 [hep-ph]. URL: http://dx.doi.org/10.1007/JHEP03(2015)017.

[10] P. Fayet and J. Iliopoulos. "Spontaneously broken supergauge symmetries and goldstone spinors." In: *Physics Letters B* 51.5 (1974), pp. 461–464. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(74)90310-4. URL: http://www.sciencedirect.com/science/article/pii/0370269374903104.

[11] L. O'Raifeartaigh. "Spontaneous symmetry breaking for chirals scalar superfields." In: *Nuclear Physics B* 96.2 (1975), pp. 331–352. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(75)90585-4. URL: http://www.sciencedirect.com/science/article/pii/0550321375905854.

[12] L. Hall, J. Lykken, and S. Weinberg. "Supergravity as the messenger of supersymmetry breaking." In: *Phys. Rev. D* 27 (10 May 1983), pp. 2359–2378. DOI: 10.1103/PhysRevD.27.2359. URL: https://link.aps.org/doi/10.1103/PhysRevD.27.2359.

[13] A. H. Chamseddine, R. Arnowitt, and P. Nath. "Locally supersymmetric grand unification." In: *Phys. Rev. Lett.* 49 (14 Oct. 1982), pp. 970–974. DOI: 10.1103/PhysRevLett.49.970. URL: https://link.aps.org/doi/10.1103/PhysRevLett.49.970.

[14] L. Alvarez-Gaumé, J. Polchinski, and M. B. Wise. "Minimal low-energy supergravity." In: *Nuclear Physics B* 221.2 (1983), pp. 495–523. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(83)90591-6. URL: http://www.sciencedirect.com/science/article/pii/0550321383905916.

[15] R. Barbieri, S. Ferrara, and C.A. Savoy. "Gauge models with spontaneously broken local supersymmetry." In: *Physics Letters B* 119.4 (1982), pp. 343–347. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(82)90685-2. URL: http://www.sciencedirect.com/science/article/pii/0370269382906852.

[16] L. Alvarez-Gaumé, M. Claudson, and M. B. Wise. "Low-energy supersymmetry." In: *Nuclear Physics B* 207.1 (1982), pp. 96–110. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(82)90138-9. URL: http://www.sciencedirect.com/science/article/pii/0550321382901389.

[17] G. F. Giudice and R. Rattazzi. "Theories with gauge mediated supersymmetry breaking." In: *Physics Reports* 322.6 (Dec. 1999), pp. 419–499. ISSN: 0370-1573. DOI: 10.1016/s0370-1573(99)00042-3. arXiv: hep-ph/9801271 [hep-ph]. URL: http://dx.doi.org/10.1016/S0370-1573(99)00042-3.

[18] S. Dimopoulos et al. "Experimental signatures of low-energy gauge mediated supersymmetry breaking." In: *Physical Review Letters* 76.19 (May 1996), pp. 3494–3497. ISSN: 1079-7114. DOI: 10.1103/physrevlett.76.3494. arXiv: hep-ph/9601367 [hep-ph]. URL: http://dx.doi.org/10.1103/PhysRevLett.76.3494.

[19] P. Fayet. "Mixing between gravitational and weak interactions through the massive gravitino." In: *Physics Letters B* 70 (1977), p. 461. DOI: 10.1016/0370-2693(77)90414-2. URL: http://dx.doi.org/10.1016/0370-2693(77)90414-2.

[20] H. Baer et al. "Signals for the minimal gauge-mediated supersymmetry breaking model at the Fermilab Tevatron collider." In: *Physical Review D* 55.7 (Apr. 1997), pp. 4463–4474. ISSN: 1089-4918. DOI: 10.1103/physrevd.55.4463. arXiv: hep-ph/9610358 [hep-ph]. URL: http://dx.doi.org/10.1103/PhysRevD.55.4463.

[21] H. Baer et al. "Reach of Tevatron upgrades in gauge-mediated supersymmetry breaking models." In: *Physical Review D* 60.5 (July 1999). ISSN: 1089-4918. DOI: 10.1103/physrevd.60.055001. arXiv: hep-ph/9903333 [hep-ph]. URL: http://dx.doi.org/10.1103/PhysRevD.60.055001.

[22] S. Dimopoulos, S. Thomas, and J. D. Wells. "Sparticle spectroscopy and electroweak symmetry breaking with gauge-mediated supersymmetry breaking." In: *Nuclear Physics B* 488.1-2 (Mar. 1997), pp. 39–91. ISSN: 0550-3213. DOI: 10.1016/s0550-3213(97)00030-8. arXiv: hep-ph/9609434 [hep-ph]. URL: http://dx.doi.org/10.1016/S0550-3213(97)00030-8.

[23] J. R. Ellis, J. L. Lopez, and D. V. Nanopoulos. "Analysis of LEP constraints on supersymmetric models with a light gravitino." In: *Physics Letters B* 394.3-4 (Feb. 1997), pp. 354–358. ISSN: 0370-2693. DOI: 10.1016/s0370-2693(97)00019-1. arXiv: hep-ph/9610470 [hep-ph]. URL: http://dx.doi.org/10.1016/S0370-2693(97)00019-1.

[24] M. Dine et al. "New tools for low energy dynamical supersymmetry breaking." In: *Physical Review D* 53.5 (Mar. 1996), pp. 2658–2669. ISSN: 1089-4918. DOI: 10.1103/physrevd.53.2658. arXiv: hep-ph/9507378 [hep-ph]. URL: http://dx.doi.org/10.1103/PhysRevD.53.2658.

[25] G. F. Giudice and R. Rattazzi. "Gauge-mediated supersymmetry breaking." In: *Perspectives on supersymmetry*. World Scientific, Singapore, 1998, p. 355. DOI: 10.1142/9789812839657_0015. URL: http://dx.doi.org/10.1142/9789812839657_0015.

[26] LHC SUSY Cross Section Working Group. *SUSY Cross Sections for 13, 14, 33 and 100 TeV pp Collisions*. URL: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/SUSYCrossSections (visited on 08/20/2020).

[27] CMS Collaboration. *CMS Supersymmetry Physics Results*. URL: https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsSUS (visited on 08/20/2020).

[28] *SUSY July 2020 Summary Plot Update*. Tech. rep. ATL-PHYS-PUB-2020-020. Geneva: CERN, July 2020. URL: http://cds.cern.ch/record/2725258.

[29] L. Evans and P. Bryant. "LHC machine." In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08001–S08001. DOI: 10.1088/1748-0221/3/08/s08001. URL: https://doi.org/10.1088/1748-0221/3/08/S08001.

[30] J. P. Blewett. "200 GeV intersecting storage accelerators." In: *Proceedings of the 8th international conference on high-energy accelerators.* Vol. 28. 1971. URL: https://lss.fnal.gov/conf/C710920/p501.pdf.

[31] E. Mobs. "The CERN accelerator complex - 2019. Complexe des accélérateurs du CERN - 2019." In: (July 2019). General Photo. URL: https://cds.cern.ch/record/2684277.

[32] O. S. Brüning et al. *LHC Design Report.* CERN Yellow Reports: Monographs. Geneva: CERN, 2004. DOI: 10.5170/CERN-2004-003-V-1. URL: https://cds.cern.ch/record/782076.

[33] A Radi. "Studying total proton–proton cross section collision at large hadron collider using gene expression programming." In: 869 (June 2017), p. 012049. DOI: 10.1088/1742-6596/869/1/012049. URL: https://doi.org/10.1088/1742-6596/869/1/012049.

[34] TOTEM Collaboration. "First measurement of elastic, inelastic and total cross-section at $\sqrt{s}$ = 13 TeV by TOTEM and overview of cross-section data at LHC energies." In: *The European Physical Journal C* 79.103 (Feb. 2019). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-019-6567-0. arXiv: 1712.06153 [hep-ex]. URL: https://doi.org/10.1140/epjc/s10052-019-6567-0.

[35] M. Benedikt, D. Schulte, and F. Zimmermann. "Optimizing integrated luminosity of future hadron colliders." In: *Phys. Rev. ST Accel. Beams* 18 (10 Oct. 2015), p. 101002. DOI: 10.1103/PhysRevSTAB.18.101002. URL: https://link.aps.org/doi/10.1103/PhysRevSTAB.18.101002.

[36] CMS Collaboration. *CMS Luminosity - Public Results*. URL: https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults (visited on 08/09/2020).

[37] Apollinari G. et al. *High-Luminosity Large Hadron Collider (HL-LHC): Technical Design Report V. 0.1.* CERN Yellow Reports: Monographs. Geneva: CERN, 2017. DOI: 10.23731/CYRM-2017-004. URL: https://cds.cern.ch/record/2284929.

[38] T. Sakuma. "Cutaway diagrams of CMS detector." In: (May 2019). URL: http://cds.cern.ch/record/2665537.

[39] J. C. Lottin et al. "Conceptual design of the CMS 4 Tesla solenoid." In: *Advances in Cryogenic Engineering: Part A*. Ed. by Peter Kittel. Boston, MA: Springer US, 1996, pp. 819–826. ISBN: 978-1-4613-0373-2. DOI: 10.1007/978-1-4613-0373-2_106. URL: https://doi.org/10.1007/978-1-4613-0373-2_106.

[40] CMS Collaboration. *CMS physics: technical design report volume 1: detector performance and software*. Technical Design Report CMS. There is an error on cover due to a technical problem for some items. Geneva: CERN, 2006. URL: https://cds.cern.ch/record/922757.

[41] F. Kircher et al. "Final design of the CMS solenoid cold mass." In: *16th International Conference on Magnet Technology (MT-16)*. Oct. 1999.

[42] CMS Collaboration. "Precise Mapping of the Magnetic Field in the CMS Barrel Yoke using Cosmic Rays." In: *Journal of Instrumentation* 5.03 (Mar. 2010), T03021–T03021. ISSN: 1748-0221. DOI: 10.1088/1748-0221/5/03/t03021. arXiv: 0910.5530 [physics.ins-det]. URL: http://dx.doi.org/10.1088/1748-0221/5/03/T03021.

[43] A. Dominguez et al. *CMS technical design report for the Pixel detector upgrade*. Tech. rep. CERN-LHCC-2012-016. CMS-TDR-11. Additional contacts: Jeffrey Spalding, Fermilab, Jeffrey.Spalding@cern.ch Didier Contardo, Universite Claude Bernard-Lyon I, didier.claude.contardo@cern.ch. Sept. 2012. URL: https://cds.cern.ch/record/1481838.

[44] *Sketches of the CMS Tracker Detector*. URL: https://twiki.cern.ch/twiki/bin/view/CMSPublic/DPGResultsTRK (visited on 08/26/2020).

[45] S. Spannagel. *Status of the CMS Phase I Pixel detector upgrade*. Tech. rep. CMS-CR-2015-272. CMS-CR-2015-272. Geneva: CERN, Oct. 2015. DOI: 10.1016/j.nima.2016.03.028. URL: https://cds.cern.ch/record/2102893.

[46] V. Karimäki et al. *The CMS tracker system project: technical design report*. Technical Design Report CMS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/368412.

[47] *Tracker Material Budget plots*. URL: https://twiki.cern.ch/twiki/bin/view/CMSPublic/TrackerMaterialBudgetplots (visited on 08/26/2020).

[48] *Atomic and nuclear properties of lead tungstate*. URL: https://pdg.lbl.gov/2017/AtomicNuclearProperties/HTML/lead_tungstate.html (visited on 08/25/2020).

[49] CMS Collaboration. "Energy calibration and resolution of the CMS electromagnetic calorimeter in pp collisions at $\sqrt{s}$ = 7 TeV." In: *Journal of Instrumentation* 8.09 (Sept. 2013), P09009–P09009. ISSN: 1748-0221. DOI: 10.1088/1748-0221/8/09/p09009. arXiv: 1306.2016 [hep-ex]. URL: http://dx.doi.org/10.1088/1748-0221/8/09/P09009.

[50] CMS Collaboration. *Reconstruction of signal amplitudes in the CMS electromagnetic calorimeter in the presence of overlapping proton-proton interactions.* 2020. arXiv: 2006.14359 [physics.ins-det].

[51] CMS Collaboration. "Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s}$ = 8 TeV." In: *Journal of Instrumentation* 06 (June 2015), P06005–P06005. ISSN: 1748-0221. DOI: 10.1088/1748-0221/10/06/p06005. arXiv: 1502.02701 [physics.ins-det]. URL: http://dx.doi.org/10.1088/1748-0221/10/06/P06005.

[52] "ECAL 2016 refined calibration and Run 2 summary plots." In: (Apr. 2020). URL: https://cds.cern.ch/record/2717925.

[53] *The CMS electromagnetic calorimeter project: Technical Design Report.* Technical Design Report CMS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/349375.

[54] CMS Collaboration. "Performance of the CMS Hadron Calorimeter with cosmic ray muons and LHC beam data." In: *Journal of Instrumentation* 5.03 (Mar. 2010), T03012–T03012. ISSN: 1748-0221. DOI: 10.1088/1748-0221/5/03/t03012. arXiv: 0911.4991 [physics.ins-det]. URL: http://dx.doi.org/10.1088/1748-0221/5/03/T03012.

[55] CMS Collaboration. "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s}$ = 13 TeV." In: *Journal of Instrumentation* 13.06 (June 2018), P06015–P06015. ISSN: 1748-0221. DOI: 10.1088/1748-0221/13/06/p06015. arXiv: 1804.04528 [physics.ins-det]. URL: http://dx.doi.org/10.1088/1748-0221/13/06/P06015.

[56] CMS Collaboration. "Performance of the CMS Drift Tube chambers with cosmic rays." In: *Journal of Instrumentation* 5.03 (Mar. 2010), T03015–T03015. ISSN: 1748-0221. DOI: 10.1088/1748-0221/5/03/t03015. arXiv: 0911.4855 [physics.ins-det]. URL: http://dx.doi.org/10.1088/1748-0221/5/03/T03015.

[57] G. Charpak and F. Sauli. "High-accuracy, two-dimensional read-out in multiwire proportional chambers." In: *Nuclear Instruments and Methods* 113.3 (1973), pp. 381–385. ISSN: 0029-554X. DOI: https://doi.org/10.1016/0029-554X(73)90503-X. URL: http://www.sciencedirect.com/science/article/pii/0029554X7390503X.

[58] The CMS collaboration. "The performance of the CMS muon detector in proton-proton collisions at sqrt(s) = 7 TeV at the LHC." In: *Journal of Instrumentation* 8.11 (Nov. 2013), P11002–P11002. ISSN: 1748-0221. DOI: 10.1088/1748-0221/8/11/p11002. arXiv: 1306.6905 [physics.ins-det]. URL: http://dx.doi.org/10.1088/1748-0221/8/11/P11002.

[59] CMS Collaboration. "RPC detector performance results for 2016 and 2017." In: (Feb. 2018). URL: https://cds.cern.ch/record/2305542.

[60] CMS Collaboration. "The CMS experiment at the CERN LHC." In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08004–S08004. DOI: 10.1088/1748-0221/3/08/s08004. URL: https://doi.org/10.1088/1748-0221/3/08/s08004.

[61] CMS Collaboration. *CMS TriDAS project: technical design report, volume 1: the trigger systems.* Technical Design Report CMS. URL: https://cds.cern.ch/record/706847.

[62] S. Cittolin, A. Rácz, and P. Sphicas. *CMS The TriDAS Project: technical design report, volume 2: data acquisition and High-Level Trigger. CMS trigger and data-acquisition project.* Technical Design Report CMS. Geneva: CERN, 2002. URL: http://cds.cern.ch/record/578006.

[63] CMS Collaboration. "The CMS trigger system." In: *Journal of Instrumentation* 12.01 (Jan. 2017), P01020–P01020. ISSN: 1748-0221. DOI: 10.1088/1748-0221/12/01/p01020. arXiv: 1609.02366 [physics.ins-det]. URL: http://dx.doi.org/10.1088/1748-0221/12/01/P01020.

[64] CMS Collaboration. *Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s}$ = 13 TeV.* 2020. arXiv: 2006.10165 [hep-ex].

[65] *Virtex-7 FPGA.* URL: https://www.xilinx.com/products/silicon-devices/fpga/virtex-7.html (visited on 08/21/2020).

[66] R. Frazier et al. "A demonstration of a Time Multiplexed Trigger for the CMS experiment." In: *Journal of Instrumentation* 7.01 (Jan. 2012), pp. C01060–C01060. DOI: 10.1088/1748-0221/7/01/c01060. URL: https://doi.org/10.1088/1748-0221/7/01/c01060.

[67] CMS Collaboration. "Performance of the reconstruction and identification of high-momentum muons in proton-proton collisions at $\sqrt{s}$ = 13 TeV." In: *Journal of Instrumentation* 15.02 (Feb. 2020), P02027–P02027. ISSN: 1748-0221. DOI: 10.1088/1748-0221/15/02/p02027. arXiv: 1912.03516 [physics.ins-det]. URL: http://dx.doi.org/10.1088/1748-0221/15/02/P02027.

[68] "Level-1 E/Gamma trigger performance in Run II." In: (July 2019). URL: http://cds.cern.ch/record/2682202.

[69] CMS Collaboration. "Selecting VBF Higgs in the Level-1 trigger at CMS." In: (Feb. 2018). URL: https://cds.cern.ch/record/2305546.

[70] "Level-1 Muon Trigger Performance." In: (July 2018). URL: https://cds.cern.ch/record/2629852.

[71] CMS Collaboration. "Particle-flow reconstruction and global event description with the CMS detector." In: *Journal of Instrumentation* 12.10 (Oct. 2017), P10003–P10003. ISSN: 1748-0221. DOI: 10.1088/1748-0221/12/10/p10003. arXiv: 1706.04965 [physics.ins-det]. URL: http://dx.doi.org/10.1088/1748-0221/12/10/P10003.

[72] S. R. Davis. "Interactive slice of the CMS detector." In: (Aug. 2016). URL: https://cds.cern.ch/record/2205172.

[73] R. Frühwirth. "Application of Kalman filtering to track and vertex fitting." In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 262.2 (1987), pp. 444–450. ISSN: 0168-9002. DOI: https://doi.org/10.1016/0168-9002(87)90887-4. URL: http://www.sciencedirect.com/science/article/pii/0168900287908874.

[74] M. Cacciari and G. P. Salam. "Dispelling the N^3 myth for the k_T jet-finder." In: *Physics Letters B* 641.1 (Sept. 2006), pp. 57–61. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2006.08.037. arXiv: hep-ph/0512210 [hep-ph]. URL: http://dx.doi.org/10.1016/j.physletb.2006.08.037.

[75] M. Cacciari, G. P. Salam, and G. Soyez. "The anti-k_T jet clustering algorithm." In: *Journal of High Energy Physics* 2008.04 (Apr. 2008), pp. 063–063. ISSN: 1029-8479. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph]. URL: http://dx.doi.org/10.1088/1126-6708/2008/04/063.

[76] M. Cacciari, G. P. Salam, and G. Soyez. "FASTJET user manual." In: *The European Physical Journal C* 72.3 (Mar. 2012). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-012-1896-2. arXiv: 1111.6097 [hep-ph]. URL: http://dx.doi.org/10.1140/epjc/s10052-012-1896-2.

[77] CMS Collaboration. "Performance of tau-lepton reconstruction and identification in CMS." In: *Journal of Instrumentation* 7.01 (Jan. 2012), P01001–P01001. ISSN: 1748-0221. DOI: 10.1088/1748-0221/7/01/p01001. arXiv: 1109.6034 [physics.ins-det]. URL: http://dx.doi.org/10.1088/1748-0221/7/01/P01001.

[78] CMS Collaboration. "Reconstruction and identification of tau lepton decays to hadrons and tau neutrino at CMS." In: *Journal of Instrumentation* 11.01 (Jan. 2016), P01019–P01019. ISSN: 1748-0221. DOI: 10.1088/1748-0221/11/01/p01019. arXiv: 1510.07488 [physics.ins-det]. URL: http://dx.doi.org/10.1088/1748-0221/11/01/P01019.

[79] CMS Collaboration. "Performance of reconstruction and identification of τ leptons decaying to hadrons and ν_τ in pp collisions at $\sqrt{s}$ = 13 TeV." In: *Journal of Instrumentation* 13.10 (Oct. 2018), P10005–P10005. ISSN: 1748-0221. DOI: 10.1088/1748-0221/13/10/p10005. arXiv: 1809.02816 [hep-ex]. URL: http://dx.doi.org/10.1088/1748-0221/13/10/P10005.

[80] CMS Collaboration. *A MIP timing detector for the CMS Phase-2 upgrade.* Tech. rep. CERN-LHCC-2019-003. CMS-TDR-020. Geneva: CERN, Mar. 2019. URL: https://cds.cern.ch/record/2667167.

[81] CMS Collaboration. *Technical proposal for a MIP timing detector in the CMS experiment Phase 2 upgrade.* Tech. rep. CERN-LHCC-2017-027. LHCC-P-009. Geneva: CERN, Dec. 2017. URL: https://cds.cern.ch/record/2296612.

[82] S. Kwan et al. "The pixel tracking telescope at the Fermilab test beam facility." In: *Nucl. Instrum. Methods Phys. Res., A* 811.FERMILAB-PUB-14-212-CD (July 2014), 162–169. 8 p. DOI: 10.1016/j.nima.2015.12.003. URL: https://cds.cern.ch/record/2709555.

[83] *FTBF Silicon Telescope.* URL: https://ftbf.fnal.gov/si-pixel-telescope/ (visited on 08/28/2020).

[84] *CAEN V1742.* URL: https://www.caen.it/products/v1742/ (visited on 08/28/2020).

[85] S. Agostinelli et al. "GEANT4—a simulation toolkit." In: *Nucl. Instrum. Meth. A* 506 (2003), p. 250. DOI: 10.1016/S0168-9002(03)01368-8. URL: http://dx.doi.org/10.1016/S0168-9002(03)01368-8.

[86] *LYSO Material Data Sheet.* URL: https://www.crystals.saint-gobain.com/products/prelude-420-LYSO (visited on 08/28/2020).

[87] X. Qu, L. Zhang, and R. Zhu. "Radiation induced color centers and light monitoring for lead tungstate crystals." In: *IEEE Transactions on Nuclear Science* 47.6 (2000), pp. 1741–1747.

[88] CMS Collaboration. "CMS ECAL response to laser light." In: (Mar. 2019). URL: https://cds.cern.ch/record/2668200.

[89] CMS Collaboration. "CMS ECAL with 2017 data." In: (May 2018). URL: https://cds.cern.ch/record/2319285.

[90] "ECAL Detector Performance, 2011 Data." In: (May 2012). URL: http://cds.cern.ch/record/1457922.

[91] "ECAL 2018 refined calibration and Preshower Run 2 performance plots." In: (Nov. 2019). URL: https://cds.cern.ch/record/2701337.

[92] *The Phase-2 Upgrade of the CMS Endcap Calorimeter.* Tech. rep. CERN-LHCC-2017-023. CMS-TDR-019. Geneva: CERN, Nov. 2017. URL: https://cds.cern.ch/record/2293646.

[93] CMS Collaboration. *The Phase-2 upgrade of the CMS barrel calorimeters.* Tech. rep. CERN-LHCC-2017-011. CMS-TDR-015. Geneva: CERN, Sept. 2017. URL: https://cds.cern.ch/record/2283187.

[94] CMS Collaboration. "Measurements of production cross sections of WZ and same-sign WW boson pairs in association with two jets in proton-proton collisions at $\sqrt{s}$ = 13 TeV." In: (2020). arXiv: 2005.01173 [hep-ex].

[95] ATLAS Collaboration. "Observation of electroweak production of a same-sign W boson pair in association with two jets in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector." In: *Physical Review Letters* 123.16 (Oct. 2019). ISSN: 1079-7114. DOI: 10.1103/physrevlett.123.161801. arXiv: 1906.03203 [hep-ex]. URL: http://dx.doi.org/10.1103/PhysRevLett.123.161801.

[96] CMS Collaboration. "Measurements of production cross sections of WZ and same-sign WW boson pairs in association with two jets in proton-proton collisions at $\sqrt{s}$ = 13 TeV." In: (2020). arXiv: 2005.01173 [hep-ex].

[97] ATLAS Collaboration. "Observation of electroweak $W^{\pm}Z$ boson pair production in association with two jets in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector." In: *Physics Letters B* 793 (June 2019), pp. 469–492. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2019.05.012. arXiv: 1812.09740 [hep-ex]. URL: http://dx.doi.org/10.1016/j.physletb.2019.05.012.

[98] CMS Collaboration. *Evidence for vector boson scattering in events with four leptons and two jets in proton-proton collisions at $\sqrt{s}$ = 13 TeV.* Tech. rep. CMS-PAS-SMP-20-001. Geneva: CERN, 2020. URL: http://cds.cern.ch/record/2718812.

[99] ATLAS Collaboration. "Observation of electroweak production of two jets and a Z-boson pair with the ATLAS detector at the LHC." In: (2020). arXiv: 2004.10612 [hep-ex].

[100] ATLAS Collaboration. "Evidence for the production of three massive vector bosons with the ATLAS detector." In: *Physics Letters B* 798 (Nov. 2019), p. 134913. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2019.134913. arXiv: 1903.10415 [hep-ex]. URL: http://dx.doi.org/10.1016/j.physletb.2019.134913.

[101] CMS Collaboration. "Search for the production of $W^{\pm}W^{\pm}W^{\mp}$ events at $\sqrt{s}$ = 13 TeV." In: *Physical Review D* 100.1 (July 2019). ISSN: 2470-0029. DOI: 10.1103/physrevd.100.012004. arXiv: 1905.04246 [hep-ex]. URL: http://dx.doi.org/10.1103/PhysRevD.100.012004.

[102] CMS Collaboration. *Observation of the production of three massive gauge bosons at $\sqrt{s}$ = 13 TeV.* 2020. arXiv: 2006.11191 [hep-ex].

[103] A. Lazopoulos, K. Melnikov, and F. Petriello. "QCD corrections to triboson production." In: *Physical Review D* 76.1 (July 2007). ISSN: 1550-2368. DOI: 10.1103/physrevd.76.014001. arXiv: hep-ph/0703273 [hep-ph]. URL: http://dx.doi.org/10.1103/PhysRevD.76.014001.

[104] T. Binoth et al. "NLO QCD corrections to triboson production." In: *Journal of High Energy Physics* 2008.06 (June 2008), pp. 082–082. ISSN: 1029-8479. DOI: 10.1088/1126-6708/2008/06/082. arXiv: 0804.0350 [hep-ph]. URL: http://dx.doi.org/10.1088/1126-6708/2008/06/082.

[105] V. Hankele and D. Zeppenfeld. "QCD corrections to hadronic WWZ production with leptonic decays." In: *Physics Letters B* 661.2-3 (Mar. 2008), pp. 103–108. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2008.02.014. arXiv: 0712.3544 [hep-ph]. URL: http://dx.doi.org/10.1016/j.physletb.2008.02.014.

[106] F. Campanario et al. "QCD corrections to charged triple vector boson production with leptonic decay." In: *Physical Review D* 78.9 (Nov. 2008). ISSN: 1550-2368. DOI: 10.1103/physrevd.78.094012. arXiv: 0809.0790 [hep-ph]. URL: http://dx.doi.org/10.1103/PhysRevD.78.094012.

[107] S. Dittmaier, A. Huss, and G. Knippen. "Next-to-leading-order QCD and electroweak corrections to WWW production at proton-proton colliders." In: *Journal of High Energy Physics* 2017.9 (Sept. 2017). ISSN: 1029-8479. DOI: 10.1007/jhep09(2017)034. arXiv: 1705.03722 [hep-ph]. URL: http://dx.doi.org/10.1007/JHEP09(2017)034.

[108] T. Han and S. Willenbrock. "QCD correction to the pp → WH and ZH total cross-sections." In: *Phys. Lett.* B273 (1991), pp. 167–172. DOI: 10.1016/0370-2693(91)90572-8.

[109] O. Brein, A. Djouadi, and R. Harlander. "NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders." In: *Physics Letters B* 579.1-2 (Jan. 2004), pp. 149–156. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2003.10.112. arXiv: hep-ph/0307206 [hep-ph]. URL: http://dx.doi.org/10.1016/j.physletb.2003.10.112.

[110] M. L. Ciccolini, S. Dittmaier, and M. Kramer. "Electroweak radiative corrections to associated WH and ZH production at hadron colliders." In: *Physical Review D* 68.7 (Oct. 2003). ISSN: 1089-4918. DOI: 10.1103/physrevd.68.073003. arXiv: hep-ph/0306234 [hep-ph]. URL: http://dx.doi.org/10.1103/PhysRevD.68.073003.

[111] D. de Florian et al. *Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector.* CERN Report CERN-2017-002-M. 2016. DOI: 10.23731/CYRM-2017-002. arXiv: 1610.07922 [hep-ph].

[112] Donato, S. "CMS trigger performance." In: *EPJ Web Conf.* 182 (2018), p. 02037. DOI: 10.1051/epjconf/201818202037. URL: https://doi.org/10.1051/epjconf/201818202037.

[113] J. Alwall et al. "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations." In: *Journal of High Energy Physics* 2014.7 (July 2014). ISSN: 1029-8479. DOI: 10.1007/jhep07(2014)079. arXiv: 1405.0301 [hep-ph]. URL: http://dx.doi.org/10.1007/JHEP07(2014)079.

[114] P. Nason. "A new method for combining NLO QCD with shower Monte Carlo algorithms." In: *Journal of High Energy Physics* 2004.11 (Nov. 2004), pp. 040–040. ISSN: 1029-8479. DOI: 10.1088/1126-6708/2004/11/040. arXiv: hep-ph/0409146 [hep-ph]. URL: http://dx.doi.org/10.1088/1126-6708/2004/11/040.

[115] S. Frixione, P. Nason, and C. Oleari. "Matching NLO QCD computations with parton shower simulations: the POWHEG method." In: *Journal of High Energy Physics* 2007.11 (Nov. 2007), pp. 070–070. ISSN: 1029-8479. DOI: 10.1088/1126-6708/2007/11/070. arXiv: 0709.2092 [hep-ph]. URL: http://dx.doi.org/10.1088/1126-6708/2007/11/070.

[116] S. Alioli et al. "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX." In: *Journal of High Energy Physics* 2010.6 (June 2010). ISSN: 1029-8479. DOI: 10.1007/jhep06(2010)043. arXiv: 1002.2581 [hep-ph]. URL: http://dx.doi.org/10.1007/JHEP06(2010)043.

[117] E. Re. "Single-top Wt-channel production matched with parton showers using the POWHEG method." In: *The European Physical Journal C* 71.2 (Feb. 2011). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-011-1547-z. arXiv: 1009.2450 [hep-ph]. URL: http://dx.doi.org/10.1140/epjc/s10052-011-1547-z.

[118] J. Alwall et al. "Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions." In: *The European Physical Journal C* 53.3 (Dec. 2007), pp. 473–500. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-007-0490-5. arXiv: 0706.2569 [hep-ph]. URL: http://dx.doi.org/10.1140/epjc/s10052-007-0490-5.

[119] R. D. Ball et al. "Parton distributions for the LHC Run II." In: *Journal of High Energy Physics* 2015.4 (Apr. 2015). ISSN: 1029-8479. DOI: 10.1007/jhep04(2015)040. arXiv: 1410.8849 [hep-ph]. URL: http://dx.doi.org/10.1007/JHEP04(2015)040.

[120] R. D. Ball et al. "Parton distributions from high-precision collider data." In: *Eur. Phys. J. C* 77 (2017), p. 663. DOI: 10.1140/epjc/s10052-017-5199-5. arXiv: 1706.00428 [hep-ph].

[121] T. Sjöstrand et al. "An introduction to PYTHIA 8.2." In: *Computer Physics Communications* 191 (June 2015), pp. 159–177. ISSN: 0010-4655. DOI: 10.1016/j.cpc.2015.01.024. arXiv: 1410.3012 [hep-ph]. URL: http://dx.doi.org/10.1016/j.cpc.2015.01.024.

[122] CMS Collaboration. "Event generator tunes obtained from underlying event and multiparton scattering measurements." In: *The European Physical Journal C* 76.3 (Mar. 2016). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-016-3988-x. arXiv: 1512.00815 [hep-ex]. URL: http://dx.doi.org/10.1140/epjc/s10052-016-3988-x.

[123] CMS Collaboration. "Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements." In: *The European Physical Journal C* 80.1 (Jan. 2020). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-019-7499-4. arXiv: 1903.12179 [hep-ex]. URL: http://dx.doi.org/10.1140/epjc/s10052-019-7499-4.

[124] CMS Collaboration. *Jet algorithms performance in 13 TeV data.* Tech. rep. CMS-PAS-JME-16-003. Geneva: CERN, 2017. URL: http://cds.cern.ch/record/2256875.

[125] CMS Collaboration. "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV." In: *Journal of Instrumentation* 12.02 (Feb. 2016), P02014–P02014. ISSN: 1748-0221. DOI: 10.1088/1748-0221/12/02/p02014. arXiv: 1607.03663 [hep-ex]. URL: http://dx.doi.org/10.1088/1748-0221/12/02/P02014.

[126] CMS Collaboration. "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV." In: *Journal of Instrumentation* 13.05 (May 2017), P05011–P05011. ISSN: 1748-0221. DOI: 10.1088/1748-0221/13/05/p05011. arXiv: 1712.07158 [physics.ins-det]. URL: http://dx.doi.org/10.1088/1748-0221/13/05/P05011.

[127] CMS Collaboration. "Search for direct production of supersymmetric partners of the top quark in the all-jets final state in proton-proton collisions at $\sqrt{s}$ = 13 TeV." In: *Journal of High Energy Physics* 2017.10 (Oct. 2017). ISSN: 1029-8479. DOI: 10.1007/jhep10(2017)005. arXiv: 1707.03316 [hep-ex]. URL: http://dx.doi.org/10.1007/JHEP10(2017)005.

[128] CMS Collaboration. *Performance of missing energy reconstruction in 13 TeV pp collision data using the CMS detector.* Tech. rep. CMS-PAS-JME-16-004. Geneva: CERN, 2016. URL: http://cds.cern.ch/record/2205284.

[129] T. Chen and C. Guestrin. *XGBoost: a scalable tree boosting system.* Aug. 2016. DOI: 10.1145/2939672.2939785. arXiv: 1603.02754 [cs.LG]. URL: http://dx.doi.org/10.1145/2939672.2939785.

[130] C. G. Lester and D. J. Summers. "Measuring masses of semi-invisibly decaying particles pair produced at hadron colliders." In: *Physics Letters B* 463.1 (Sept. 1999), pp. 99–103. ISSN: 0370-2693. DOI: 10.1016/s0370-

2693(99)00945-4. arXiv: hep-ph/9906349 [hep-ph]. URL: http://dx.doi.org/10.1016/S0370-2693(99)00945-4.

[131] CMS Collaboration. "Measurement of the $Z\gamma^* \rightarrow \tau\tau$ cross section in pp collisions at $\sqrt{s} = 13$ TeV and validation of τ lepton analysis techniques." In: *The European Physical Journal C* 78.9 (Sept. 2018). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-018-6146-9. arXiv: 1801.03535 [hep-ex]. URL: http://dx.doi.org/10.1140/epjc/s10052-018-6146-9.

[132] G. Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics." In: *The European Physical Journal C* 71.2 (Feb. 2011). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-011-1554-0. arXiv: 1007.1727 [physics.data-an]. URL: http://dx.doi.org/10.1140/epjc/s10052-011-1554-0.

[133] CMS Collaboration. "Measurements of properties of the Higgs boson decaying to a W boson pair in pp collisions at $\sqrt{s} = 13$ TeV." In: *Physics Letters B* 791 (Apr. 2019), pp. 96–129. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2018.12.073. arXiv: 1806.05246 [hep-ex]. URL: http://dx.doi.org/10.1016/j.physletb.2018.12.073.

[134] *Measurements of properties of the Higgs boson in the four-lepton final state in proton-proton collisions at $\sqrt{s} = 13$ TeV.* Tech. rep. CMS-PAS-HIG-19-001. Geneva: CERN, 2019. URL: http://cds.cern.ch/record/2668684.

[135] C. Degrande. "A basis of dimension-eight operators for anomalous neutral triple gauge boson interactions." In: *Journal of High Energy Physics* 2014.2 (Feb. 2014). ISSN: 1029-8479. DOI: 10.1007/jhep02(2014)101. arXiv: 1308.6323 [hep-ph]. URL: http://dx.doi.org/10.1007/JHEP02(2014)101.

[136] K. Agashe et al. "Flavor universal resonances and warped gravity." In: *Journal of High Energy Physics* 2017.1 (Jan. 2017). ISSN: 1029-8479. DOI: 10.1007/jhep01(2017)016. arXiv: 1608.00526 [hep-ph]. URL: http://dx.doi.org/10.1007/JHEP01(2017)016.

[137] M. J. Strassler and K. M. Zurek. "Echoes of a hidden valley at hadron colliders." In: *Physics Letters B* 651.5-6 (Aug. 2007), pp. 374–379. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2007.06.055. arXiv: hep-ph/0604261 [hep-ph]. URL: http://dx.doi.org/10.1016/j.physletb.2007.06.055.

[138] B. C. Allanach et al. "The Snowmass Points and Slopes: Benchmarks for SUSY Searches." In: *The European Physical Journal C* 25.1 (Sept. 2002), pp. 113–123. ISSN: 1434-6052. DOI: 10.1007/s10052-002-0949-3. arXiv: hep-ph/0202233 [hep-ph]. URL: http://dx.doi.org/10.1007/s10052-002-0949-3.

[139] CMS Collaboration. "Search for long-lived particles in events with photons and missing energy in proton-proton collisions at $\sqrt{s}$ = 7 TeV." In: *Physics Letters B* 722.4-5 (May 2013), pp. 273–294. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2013.04.027. arXiv: 1212.1838 [hep-ex]. URL: http://dx.doi.org/10.1016/j.physletb.2013.04.027.

[140] ATLAS Collaboration. "Search for non-pointing photons in the diphoton and $E_{\mathrm{T}}^{\mathrm{miss}}$ final state in $\sqrt{s}$ = 7 TeV proton-proton collisions using the ATLAS detector." In: *Physical Review D* 88.1 (July 2013). ISSN: 1550-2368. DOI: 10.1103/physrevd.88.012001. arXiv: 1304.6310 [hep-ex]. URL: http://dx.doi.org/10.1103/PhysRevD.88.012001.

[141] *Search for long-lived neutral particles in the final state of delayed photons and missing energy in proton-proton collisions at $\sqrt{s}$ = 8 TeV*. Tech. rep. CMS-PAS-EXO-12-035. Geneva: CERN, 2015. URL: http://cds.cern.ch/record/2063495.

[142] ATLAS Collaboration. "Search for non-pointing and delayed photons in the diphoton and missing transverse momentum final state in 8 TeV pp collisions at the LHC using the ATLAS detector." In: *Physical Review D* 90.11 (Dec. 2014). ISSN: 1550-2368. DOI: 10.1103/physrevd.90.112005. arXiv: 1409.5542 [hep-ex]. URL: http://dx.doi.org/10.1103/PhysRevD.90.112005.

[143] P. Skands et al. *SUSY Les Houches Accord: interfacing SUSY spectrum calculators, decay packages, and event generators.* July 2004. DOI: 10.1088/1126-6708/2004/07/036. arXiv: hep-ph/0311123 [hep-ph]. URL: http://dx.doi.org/10.1088/1126-6708/2004/07/036.

[144] *Tools to scan the mSUGRA phasespace.* URL: https://wiki.nikhef.nl/atlas/Tools_to_scan_the_mSUGRA_phasespace (visited on 09/06/2020).

[145] H. Baer et al. *ISAJET 7.69: A Monte Carlo Event Generator for pp, $\bar{p}p$, and e^+e^- Reactions.* 2003. arXiv: hep-ph/0312045 [hep-ph].

[146] E. Bothmann et al. "Event Generation with Sherpa 2.2." In: *SciPost Physics* 7.3 (Sept. 2019). ISSN: 2542-4653. DOI: 10.21468/scipostphys.7.3.034. arXiv: 1905.09127 [hep-ph]. URL: http://dx.doi.org/10.21468/SciPostPhys.7.3.034.

[147] T. Gleisberg et al. "Event generation with SHERPA 1.1." In: *Journal of High Energy Physics* 2009.02 (Feb. 2009), pp. 007–007. ISSN: 1029-8479. DOI: 10.1088/1126-6708/2009/02/007. arXiv: 0811.4622 [hep-ph]. URL: http://dx.doi.org/10.1088/1126-6708/2009/02/007.

[148] T.C. Awes et al. "A simple method of shower localization and identification in laterally segmented calorimeters." In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 311.1 (1992), pp. 130–138. ISSN: 0168-

9002. DOI: https://doi.org/10.1016/0168-9002(92)90858-2. URL: http://www.sciencedirect.com/science/article/pii/0168900292908582.

[149] CMS Collaboration. "Performance of photon reconstruction and identification with the CMS detector in proton-proton collisions at $\sqrt{s}$ = 8 TeV." In: *Journal of Instrumentation* 10.08 (Aug. 2015), P08010–P08010. ISSN: 1748-0221. DOI: 10.1088/1748-0221/10/08/p08010. arXiv: 1502.02702 [physics.ins-det]. URL: http://dx.doi.org/10.1088/1748-0221/10/08/P08010.

[150] CMS Collaboration. "Time reconstruction and performance of the CMS Electromagnetic Calorimeter." In: *Journal of Instrumentation* 5.03 (Mar. 2010), T03011–T03011. ISSN: 1748-0221. DOI: 10.1088/1748-0221/5/03/t03011. arXiv: 0911.4044 [physics.ins-det]. URL: http://dx.doi.org/10.1088/1748-0221/5/03/T03011.

[151] S. Baron et al. "Jitter impact on clock distribution in LHC experiments." In: *Journal of Instrumentation* 7.12 (Dec. 2012), pp. C12023–C12023. DOI: 10.1088/1748-0221/7/12/c12023. URL: https://doi.org/10.1088/1748-0221/7/12/C12023.

[152] T. Junk. "Confidence level computation for combining searches with small statistics." In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 434.2-3 (Sept. 1999), pp. 435–443. ISSN: 0168-9002. DOI: 10.1016/s0168-9002(99)00498-2. arXiv: hep-ex/9902006 [hep-ex]. URL: http://dx.doi.org/10.1016/S0168-9002(99)00498-2.

[153] A. L. Read. "Presentation of search results: the CL_s technique." In: *Journal of Physics G: Nuclear and Particle Physics* 28.10 (Sept. 2002), pp. 2693–2704. DOI: 10.1088/0954-3899/28/10/313. URL: https://doi.org/10.1088/0954-3899/28/10/313.

[154] The ATLAS Collaboration, The CMS Collaboration, and The LHC Higgs Combination Group. *Procedure for the LHC Higgs boson search combination in Summer 2011.* Tech. rep. CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11. Geneva: CERN, Aug. 2011. URL: https://cds.cern.ch/record/1379837.

[155] CMS Collaboration. *A MIP Timing Detector for the CMS Phase-2 upgrade.* Tech. rep. CERN-LHCC-2019-003. CMS-TDR-020. Geneva: CERN, Mar. 2019. URL: https://cds.cern.ch/record/2667167.

[156] Z. Zhang. "Performance of the CMS precision electromagnetic calorimeter at LHC Run II and prospects for High-Luminosity LHC." In: *Journal of Instrumentation* 13.04 (Apr. 2018), pp. C04013–C04013. DOI: 10.1088/1748-0221/13/04/c04013. URL: https://doi.org/10.1088/1748-0221/13/04/C04013.

[157] *Acrorad Co., Ltd.* URL: https://www.acrorad.co.jp/ (visited on 08/31/2020).