Search for dark sector promptly decaying neutral particles in 140 fb^{-1} proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector

Tingyu Zhang

Department of Physics
The University of Tokyo

This dissertation is submitted for the degree of
Doctor of Philosophy

August 2024

Acknowledgements

I would like to express my deepest gratitude to my supervisor, Prof. Junichi Tanaka, for his unwavering support and insightful guidance throughout my doctoral studies. His profound expertise, thoughtful advice, and rigorous approach to research have been invaluable to the successful completion of this thesis.

I am also deeply grateful to the members of the thesis defense committee, Prof. Yutaka Ushiroda, Prof. Masashi Yokoyama, Prof. Masahiro Ibe, Prof. Taku Gunji, and Prof. Kentaro Yako, for their constructive feedback and diverse perspectives, which greatly enhanced the quality of this work.

Special thanks go to the faculties, colleagues, and friends at ICEPP, the University of Tokyo, for their collaboration and discussions that enriched my research experience. I am particularly thankful to Prof. Yuji Enari for his guidance on my work with the LAr calorimeter and physics analysis, as well as for introducing me to CERN and the delightful local restaurants. I am also grateful to Prof. Shion Chen, Prof. Yasuyuki Okumura, Prof. Masaya Ishino, Prof. Tatsuya Masubuchi, Prof. Takuya Nobe, and all the staff members for their constructive suggestions during meetings that helped me overcome many challenges in my work. My heartfelt thanks also go to Dr. Yi-Lin Yang, Dr. Reiyo Oishi, and Dr. Gen Tateno from Tanaka-lab for their assistance with my qualification work in LAr. Additionally, I am grateful to my friends and colleagues in Building 188 for their support in work and life, and for the good times we shared. In particular, I would like to thank Jiaqi Zang and Marin Furukawa for their companionship and the memorable moments we shared at CERN.

I am thankful to my analysis contacts in ATLAS, Dr. Cristiano Sebastiani, Dr. Iacopo Longarini, and Prof. Joseph Izen, for their expertise, coordination, and support in my physics analysis work. I also appreciate Dr. Elena Pompa Pacchi for showing me around Rome during conferences and my team members, Dr. Luis Pascual Dominguez and Bernardo Ricci, for their excellent collaboration in preparing our analysis for publication.

I extend my heartfelt thanks to all the friends I made during my time at CERN. The group of Dr. Zhelun Li, Dr. Miaoran Lu, Dr. Tianyu Yang, and Dr. Yousen Zhang made my two years at CERN among the most memorable experiences of my life. I cherish our discussions on work, life, love, and other abstract topics. Additionally, I would like to thank

Mengyang Li, Wenhao Ma, and Yanxi Gu for kindly driving us around Geneva, and Luxin Zhang, Pingxin Zhang, and Danning Liu for the wonderful gatherings. A special appreciation goes to Zihan Zhao for the enjoyable badminton matches and intellectual challenges and for showing me around the USTC campus.

Finally, I am profoundly grateful to my parents for their love, patience, and understanding.

Abstract

Several new physics models predict the existence of dark neutral particles, or dark photons, which decay promptly to collimated pairs of leptons, referred to as lepton jets. In the dark sector model, both couplings between Higgs and dark Higgs via the Higgs portal and couplings between photon and dark photons via the vector portal are favored to be probed in the LHC-ATLAS experiment.

In this thesis, we present a search for low-mass dark neutral particles, γ_d, in the range between 17 MeV and 0.24 GeV using 140 fb^{-1} of data collected from 2015 to 2018 during the LHC Run-2 pp collisions at $\sqrt{s} = 13$ TeV. By exploring the electron final states using channels of $H \to (\gamma_\mathrm{d} \to e^+e^-)(\gamma_\mathrm{d} \to e^+e^-) + X$ in the Falkowski-Ruderman-Volansky-Zupan (FRVZ) model and $H \to (\gamma_\mathrm{d} \to e^+e^-)(\gamma_\mathrm{d} \to e^+e^-)$ in the Hidden Abelian Higgs Model (HAHM), we extend the search limits for γ_d mass from the previous 0.25 GeV to 17 MeV.

No excess is observed, and we set an exclusion limit of 0.7% BR($H \to 2\gamma_\mathrm{d} + X$) for the FRVZ model and 0.07% BR($H \to 2\gamma_\mathrm{d}$) for the HAHM model at $m_{\gamma_d} = 17$ MeV. Compared to other experiments, this analysis extends the search limits for both the Higgs and vector portals.

Table of contents

Chapter 1

Introduction

The Standard Model (SM) forms the basis of modern particle physics, having achieved remarkable success in experimental verifications and predictions. However, the limitations within the SM necessitate the exploration of theories beyond the Standard Model (BSM). One promising path for BSM research is the study of dark matter.

Traditional approaches to dark matter research have focused on searches for dark matter candidates in models such as Weakly Interacting Massive Particles (WIMPs). Alternatively, dark matter can be explored model-independently through dark sector portals, representing groups of theoretical interactions and mixings with SM particles, leading to rich phenomenology.

Among these portals, the vector portal has been extensively explored by numerous experiments [1]. On the other hand, the ATLAS experiment at the Large Hadron Collider (LHC) offers the potential to provide unique insights into the Higgs portal since only LHC can produce Higgs bosons. This is possible thanks to the invisible decays of Higgs bosons, with an upper limit on the branching ratio of 11% [2]. One of the striking signatures is the scenario where dark photons (γ_d)s promptly decay into collimated pairs of leptons ($\ell = e,\ \mu$), often referred to as lepton jets (LJs).

Previous Higgs portal searches at the LHC have set exclusion limits for dark photons with masses greater than 250 MeV [3] [4] [5], focusing primarily on the muon final states. To expand to the lower undiscovered mass region, which is physics-motivated, we hereby present an analysis that explores the electron final states, referred to as electron lepton jets (eLJs). Through studies of the subdetector responses, we have introduced an innovative definition of eLJs: merged clusters with separately reconstructed tracks, thereby reducing the overwhelming background. In this thesis, we will search for eLJ signatures in the channel $H \to (\gamma_\mathrm{d} \to e^+e^-)(\gamma_\mathrm{d} \to e^+e^-)$ or $H \to (\gamma_\mathrm{d} \to e^+e^-)(\gamma_\mathrm{d} \to e^+e^-) + X$, where Xs are undetectable particles, using the 140 fb^{-1} full Run-2 pp collision data taken between 2015-2018 in the ATLAS experiment at the LHC.

The flow of the thesis is as follows. Chapter 2 overviews the SM and motivations for BSM theories, which drive our search for low-mass dark photons using eLJs. Chapter 3 reviews previous dark matter searches and the theoretical basis of dark sector models, focusing on the portals used in the benchmark models exploited in this thesis, and discusses the event signatures and topologies for the $H \to (\gamma_\mathrm{d} \to e^+e^-)(\gamma_\mathrm{d} \to e^+e^-)(+X)$ search channel. Chapter 4 describes the experimental setup for LHC and the ATLAS experiment, and discusses the standard reconstruction and identification of standard physics objects in ATLAS. Chapter 5 introduces the unique signatures and customizations to detect dark photons, including the simulation of events, kinematic distributions, eLJ object definitions, and trigger strategies. Chapter 6 presents the event selections for dark photons with such signatures. Chapter 7 details the data-driven background estimation method. Chapter 8 evaluates the systematic uncertainties arising in this analysis. Chapter 9 shows the results and discusses the interpretations. Then, Chapter 10 gives a conclusion.

Chapter 2

The Standard Model and Beyond

As the single theory that certainly represents one of the significant achievements of physical science, the Standard Model provides a comprehensive framework for understanding the fundamental interactions of elementary particles. Proposed by Glashow [6], Salam [7], and Weinberg [8] in the 1960s, the theory has made successful predictions and been extensively tested during the last 50 years. The experimental discovery of the scalar field Higgs Boson by ATLAS [9] and [10] CMS experiments at the Large Hadron Collider (LHC) in 2012 completed the last missing piece to the full empirical confirmation of the theory. This chapter introduces the theoretical basis of the Standard Model and the motivations for studies of Beyond Standard Model (BSM) theories.

2.1 Particles in the Standard Model

The SM describes three of the four fundamental forces: the electromagnetic interaction, the weak interaction, and the strong interaction, except for the gravitational interaction. Formulated in the relativistic Quantum Field Theory (QFT) framework, each particle is represented by a field, and the Lagrangian defines its dynamic. The fundamental nature of particles and interactions comes from symmetry principles, where the gauge invariant governs the SM [11] under the $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge symmetry. The $SU(3)_C$ non-abelian group is related with strong interaction, and the $SU(2)_L \times SU(1)_Y$ is associated with the electroweak interaction.

Based on their spin nature, particles can be categorized into two basic types: fermions with an integer spin, which follow Bose-Einstein statistics, and bosons with a half-integer spin, which follow Fermi-Dirac statistics.

Fermions, including leptons and quarks, have a spin value of 1/2 and are the building blocks of matter. Up until now, 12 flavors of elementary fermions have been found. They are sub-categorized into three generations:

- ***first generation***: electron (e^-), electron neutrino (ν_e), up (u), and down (d) quarks,
- ***second generation***: muon (μ^-), muon neutrino (ν_μ), charm (c) and strange (s) quarks, and
- ***third generation***: tau (τ^-), tau neutrino (ν_τ), top (t), and bottom (b) quarks.

Each generation is a heavier copy of the previous one, with similar charge and spin properties. Fermions that do not undergo strong interactions are known as leptons, which are e^-, ν_e, μ^-, ν_μ, τ^- and ν_τ. On the other hand, fermions that undergo both strong and electroweak interactions are known as quarks, which are u, d, c, s, t, and b. With the strong interaction color charge, quarks can only be detected in the color-free composite form called hadrons.

Gauge bosons, with spin value 1, including γ, W, Z bosons, and gluons, are the force-carrying particles that mediate interactions between fermions. Gluons mediate the strong interaction, and there are eight in total, corresponding to the number of generators in the $SU(3)_C$ gauge symmetry. The electroweak bosons W^+, W^- and Z^0 mediate the weak interaction. W^+ and W^- bosons couple only the left-handed (doublets under $SU(2)_L$) parts of fermions via the weak isospin, while Z^0 boson couples to both left-handed and right-handed (singlet under $SU(2)_L$) part of fermions via both isospin and hypercharge. The photon γ is massless and mediates the electromagnetic field.

The Higgs boson is a scalar particle (spin value 0) associated with the Higgs field, which gives mass to all particles. Through the process of spontaneous symmetry breaking [12] of the $SU(2)_L \times U(1)_Y$ to $U(1)_{EM}$ gauge symmetry, W and Z bosons acquire mass while γ remain massless, leading to weak interaction mediated to be short-range, and electromagnetic interaction to be long-range.

A summary of elementary particles in the SM is shown in Fig. 2.1.

2.2 Quantum Electrodynamics

Quantum electrodynamics (QED) [14] is the relativistic quantum field theory that describes the phenomena of charged particles interacting in exchange of photons. We can derive the general Lagrangian form of QED by the gauge invariance: we start from the Lagrangian of

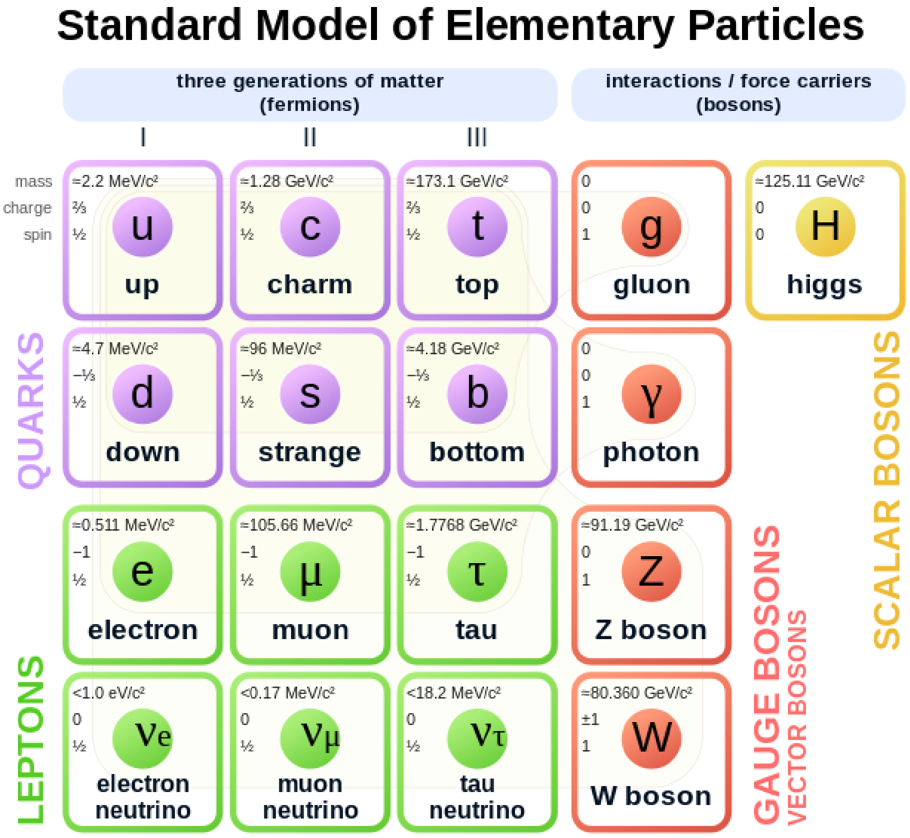


Fig. 2.1 A summary of elementary particles in the SM [13].

free Dirac fermion field $\psi(x)$:

$$\mathscr{L}_0 = i\bar{\psi}(x)\gamma^\mu\partial_\mu\psi(x) - m\bar{\psi}(x)\psi(x). \tag{2.1}$$

Under the local gauge transformation

$$\psi(x) \xrightarrow{U(1)} \psi'(x) \equiv \exp\{iQ\theta(x)\}\,\psi(x) \tag{2.2}$$

$$\partial_\mu\psi(x) \xrightarrow{U(1)} \exp\{iQ\theta(x)\}(\partial_\mu + iQ\partial_\mu\theta(x))\psi(x). \tag{2.3}$$

Here, Q is an arbitrary real charge constant, and $\theta(x)$ is the gauge phase transformation that depends on space-time coordinates. For gauge invariance to hold to cancel the extra term in

Eq. 2.1, we introduce a new gauge field $A_\mu(x)$ and the covariant derivative

$$A_\mu(x) \xrightarrow{U(1)} A'_\mu(x) \equiv A_\mu(x) - \frac{1}{e}\partial_\mu \theta(x), \tag{2.4}$$

$$D_\mu \psi(x) \equiv \left[\partial_\mu + ieQA_\mu(x)\right] \psi(x), \tag{2.5}$$

which satisfies the local gauge invariance we need. The Lagrangian then becomes:

$$\mathscr{L} = i\bar{\psi}(x)\gamma^\mu D_\mu \psi(x) - m\bar{\psi}(x)\psi(x) = \mathscr{L}_0 - eQA_\mu(x)\bar{\psi}(x)\gamma^\mu \psi(x). \tag{2.6}$$

The vertex of QED is the interaction between Dirac fermion ψ and gauge field $A_\mu(x)$. Consider $A_\mu(x)$ to be propagating, and we add the kinetic terms of the electromagnetic field:

$$\mathscr{L}_{Kin} = -\frac{1}{4}F_{\mu\nu}(x)F^{\mu\nu}(x), \tag{2.7}$$

where $F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu$ is the electromagnetic field strength tensor, which is invariant under gauge transformation of $A_\mu(x)$ in Eq. 2.4. It is important to note that a photon is massless so that a mass term like $\mathscr{L}_m = \frac{1}{2}m^2 A^\mu A_\mu$, which violates the gauge invariant requirement is not necessary. The complete Lagrangian form of propagating the QED field is as follows:

$$\mathscr{L}_{QED} = \bar{\psi}(x)(i\gamma^\mu\partial_\mu - m)\psi(x) - \frac{1}{4}F_{\mu\nu}(x)F^{\mu\nu}(x) - eQA_\mu(x)\bar{\psi}(x)\gamma^\mu \psi(x), \tag{2.8}$$

from which the Maxwell equations can be derived:

$$\partial_\mu F^{\mu\nu} = eQ\bar{\psi}\gamma^\nu \psi. \tag{2.9}$$

2.3 Quantum Chromodynamics

Quantum Chromodynamics (QCD) is the theory that describes the strong interaction of quarks, which is based on $SU(3)_C$ color symmetry.

In QCD, the color states of quark fields ψ_C are $SU(3)_C$ triplets, where C corresponds to red, green, and blue spinor elements. The free and classical QCD Lagrangian is constructed as follows:

$$\mathscr{L}_0 = i\bar{\psi}(x)\gamma^\mu\partial_\mu \psi(x) - m\bar{\psi}(x)\psi(x), \tag{2.10}$$

which is gauge symmetric globally. If we extend the gauge symmetry to local symmetry again:

$$\psi(x) \xrightarrow{SU(3)} \psi'(x) \equiv \exp\{i\theta_a(x)t_a\}\,\psi(x), \tag{2.11}$$

$$D_\mu \psi(x) \equiv \left[\partial_\mu - ig_0 t_a A_\mu(x)\right] \psi(x). \tag{2.12}$$

Here the quark field $\psi(x)$ couples to the gluon field $A_\mu(x)$ with a coupling strength g_0. t_a is the infinitesimal $SU(3)$ generator given by $T_a = \lambda_a/2$ where λ_a are the Gell-Mann matrices where $[t_a, t_b] = if_{abc}t_c$. We can define the QCD field tensor:

$$G^a_{\mu\nu} \equiv \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g_0 f_{abc} A^b_\mu A^c_\nu. \tag{2.13}$$

Unlike the QED tensor, the last term in $G^\mu\nu_a$ implies that gluons are color-charged and interact with each other. Including also the kinetic energy term, we reach the complete QCD Lagrangian:

$$\mathscr{L}_{QCD} = \bar{\psi}(x)(i\gamma^\mu\partial_\mu - m)\psi(x) - \frac{1}{4}G_{\mu\nu}(x)G^{\mu\nu}(x) + g_0 t_a A_\mu(x)\bar{\psi}(x)\gamma^\mu\psi(x). \tag{2.14}$$

One of the phenomena of QCD theory is color confinement, where detecting an isolated quark from a composite state is impossible. This is due to the coupling strength g_0 becoming very strong at high distances (low energy scales), and at some point, it is more energetically favorable to create a quark-antiquark pair. As a result, when individual quarks are produced in particle accelerators, we always observe clustered "jets" of many color-neutral particles.

Conversely, at high energy scales, the coupling strength g_0 becomes asymptotically weaker, leading to a phenomenon known as asymptotic freedom. Such free particle states allow perturbative calculations in the QCD theory.

2.4 Quantum Electroweak Theory

The unification of QED and weak interaction was developed in the 1960s by Sheldon Glashow [6], Abdus Salam [7], and Steven Weinberg [8] in the electroweak theory. We choose the $SU(2)_L \times U(1)_Y$ gauge group for such unification. The gauge fields corresponding to each generator are

$$\begin{aligned} &SU(2)_L \rightarrow W^1_\mu,\ W^2_\mu,\ W^3_\mu, \\ &U(1)_Y \rightarrow B_\mu. \end{aligned}$$

We can define the strength tensors for the abelian $U(1)_Y$ and non-abelian $SU(2)_L$ gauge fields as

$$\begin{aligned} W^i_{\mu\nu} &= \partial_\mu W^i_\nu - \partial_\nu W^i_\mu + g\varepsilon^{ijk} W^j_\mu W^k_\nu, \\ B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu. \end{aligned}$$

Let us also group the leptons to left-handed weak isospin doublet L ($T = 1/2$) and right-handed weak isospin singlet R ($T = 0$) as follows:

$$\begin{aligned} L &= \begin{pmatrix} \nu_L \\ l_L \end{pmatrix}, \\ R &= l_R, \end{aligned}$$

where $T_3 = +1/2$ and $T_3 = -1/2$ are the left-handed part of neutrino and charged lepton of any flavor ($l = e,\ \mu,\ \tau$). Since there is no observed right-handed neutrino component, the right-handed part is a weak isospin singlet.

We start from the free Lagrangian

$$\mathscr{L}_0 = \bar{L} i\gamma^\mu \partial_\mu L + \bar{R} i\gamma^\mu \partial_\mu R, \tag{2.15}$$

and introduce the covariant derivatives:

$$L : D_\mu \equiv \partial_\mu + i\frac{g}{2}\tau^i W^i_\mu + i\frac{g'}{2} Y B_\mu, \tag{2.16}$$

$$R : D_\mu \equiv \partial_\mu + i\frac{g'}{2} Y B_\mu, \tag{2.17}$$

where g and g' are the coupling constants associated with $SU(2)_L$ and $U(1)_Y$, respectively, Y is the hypercharge associated with $U(1)_Y$, and τ^i are Pauli matrices.

Therefore, the Lagrangian becomes

$$\mathscr{L}_{QEW} = \mathscr{L}_0 + \mathscr{L}_{CC} + \mathscr{L}_{NC}, \tag{2.18}$$

where the charged current part $\mathscr{L}_{CC}$ and the neutral current part $\mathscr{L}_{NC}$ are

$$\mathscr{L}_{CC} = -g\bar{L}\gamma^\mu \left(\frac{\tau^1}{2} W^1_\mu + \frac{\tau^2}{2} W^2_\mu\right) L, \tag{2.19}$$

$$\mathscr{L}_{NC} = -g\bar{L}\gamma^\mu \frac{\tau^3}{2} W^3_\mu L - \frac{g'}{2}(\bar{L}\gamma^\mu Y L + \bar{R}\gamma^\mu Y R) B_\mu. \tag{2.20}$$

First, $\mathscr{L}_{CC}$ can be rewritten as

$$\mathscr{L}_{CC} = -\frac{g}{2\sqrt{2}}[\bar{\nu}\gamma^{\mu}(1-\gamma^5)lW_{\mu}^{+} + \bar{l}\gamma^{\mu}(1-\gamma^5)\nu W_{\mu}^{-}], \tag{2.21}$$

where $W_{\mu}^{\pm}$ are the charged gauge bosons defined as

$$W_{\mu}^{\pm} = \frac{1}{\sqrt{2}}(W_{\mu}^{1} \pm W_{\mu}^{2}). \tag{2.22}$$

This $\mathscr{L}_{CC}$ shows the weakly charged current's $(V-A)$ structure [15].

Next, we look at $\mathscr{L}_{NC}$ and observe the structure of the Gell-Mann-Nishijima relation [16] [17]. To couple it with the electromagnetic current, we introduce rotation using the Weinberg angle [15]:

$$\begin{pmatrix} A_{\mu} \\ Z_{\mu} \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B_{\mu} \\ W_{\mu}^{3} \end{pmatrix} \tag{2.23}$$

where

$$\sin\theta_W = \frac{g'}{\sqrt{g^2+g'^2}}, \cos\theta_W = \frac{g}{\sqrt{g^2+g'^2}}, \tag{2.24}$$

and the Lagrangian becomes

$$\mathscr{L}_{NC} = -g\sin\theta_W(\bar{l}\gamma^{\mu}l)A_{\mu} - \frac{g}{2\cos\theta_W}\sum_{\psi=\nu,l}\overline{\psi}\gamma^{\mu}(g_V^i - g_A^i\gamma^5)\psi Z_{\mu}, \tag{2.25}$$

where g_V and g_A are vector and axial-vector coupling constants. We can find the electromagnetic current coupled to photon field A_{μ} and the electromagnetic charge

$$e = g\sin\theta_W = g'\cos\theta_W. \tag{2.26}$$

So far, we have introduced four massless gauge fields $W_{\mu}^{\pm}$, Z_{μ}, A_{μ}, and two massless fermions ν and l. How the gauge bosons acquire mass will be explained in Sec. 2.5.

2.5 Higgs Mechanism

In this section, we will add scalar fields to spontaneously break the symmetry [18] [19] and explain how mass of $W_{\mu}^{\pm}$ and Z_{μ} is given through the Higgs mechanism. Firstly, we introduce the scalar doublet

$$\Phi \equiv \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}. \tag{2.27}$$

The scalar Lagrangian is given by coupling as

$$\mathscr{L}_{\text{Scalar}} = \partial_\mu \Phi^\dagger \partial^\mu \Phi - V(\Phi^\dagger \Phi), \tag{2.28}$$

where the potential is given by

$$V(\Phi^\dagger \Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2. \tag{2.29}$$

Then we choose a constant vacuum expectation value (VEV) at the bottom of Higgs field potential as

$$\langle \Phi \rangle_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \tag{2.30}$$

where $v = \sqrt{-\mu^2/\lambda}$. In this way, when $SU(2)_L \times U(1)_Y \to U(1)_{EM}$ symmetry is spontaneously broken, VEV is conserved.

We parameterize the $SU(2)_L$ transformation by fluctuations of the Higgs boson H

$$\langle \Phi \rangle_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix}. \tag{2.31}$$

By introducing again the $SU(2)_L \times U(1)_Y$ covariant derivative as in Eqs. (2.16) and (2.17), the scalar Lagrangian in Eq. (2.28) can be written as

$$\mathscr{L}_{\text{Scalar}} = \left| \left(\partial_\mu + i\frac{g}{2}\tau^i W_\mu^i + i\frac{g'}{2} Y B_\mu \right) \begin{pmatrix} 0 \\ \frac{v+H}{\sqrt{2}} \end{pmatrix} \right|^2 - \mu^2 \frac{(v+H)^2}{2} - \lambda \frac{(v+H)^4}{4}. \tag{2.32}$$

Since only the mass contribution at the tree level is considered here, we ignore the H-interaction terms. With Eqs. (2.22) and (2.23) the Lagrangian becomes

$$\begin{aligned} \mathscr{L}_{\text{Scalar}} &= \frac{1}{2}\partial_\mu H \partial^\mu H + \frac{v^2 g^2}{4} \left((W_\mu^1)^2 + (W_\mu^2)^2 \right) + \frac{v^2}{8} \left(g W_\mu^3 - g' B_\mu \right)^2 + \text{other terms}, \\ &= \frac{1}{2}\partial_\mu H \partial^\mu H + \frac{g^2 v^2}{4} W_\mu^+ W^{-\mu} + \frac{1}{2} \frac{g^2 v^2}{4\cos^2\theta_W} Z_\mu Z^\mu + 0 \cdot A_\mu A^\mu + \text{other terms}. \end{aligned}$$

We can identify the mass of $W^{\pm}$ and Z

$$m_W = \frac{gv}{2},\ m_Z = \frac{gv}{2\cos\theta_W}. \tag{2.33}$$

It is also evident that the photon remains massless.

Using the approximation VEV of Higgs field of 246 GeV [20], the SM predicts the W mass to be $\sim$ 80 GeV and Z mass to be $\sim$ 90 GeV, which is consistent with precise experimental measurements [21].

2.6 Physics Beyond the Standard Model

The Standard Model is a remarkably successful theoretical framework in particle physics. It offers extensive predictive power and explains a wide range of phenomena. However, it also has limitations. Growing evidence suggests the need for new physics Beyond the Standard Model (BSM). Here we list a few phenomena that are not explained well by the SM:

- ***Hierarchy Problem***: as one of the most argued problems with the SM, the hierarchy problem is an essential question of scale. We saw how the Higgs field gives rise to the W and Z boson's mass. To reproduce the observed values of mass, one demands $\mu \sim O(100\text{ GeV})$, which has a large difference to the Planck scale (about 10^{19} GeV) corresponding to gravity. Unlike chiral symmetry to fermions that prohibits fermions from having a direct mass term and gauge symmetry to gauge bosons protecting the mass, the μ value for Higgs should be much larger like the Planck scale due to no symmetry to protect it. Even on the first-order correction

  $$\delta\mu^2 \sim \frac{\Lambda^2}{16\pi^2} \tag{2.34}$$

  is divergent in Λ. One has to impose a cancellation between the $O(M_{Planck})$ and $O(100\text{ GeV})$, which is clearly a fine-tuning and comes unnatural.

- ***Neutrino oscillations***: neutrino oscillation is the fact that a neutrino changes lepton flavor when propagating through space. For neutrino oscillation, the mass eigenstate must differ from a flavor eigenstate. While in the SM, no right-handed neutrinos can couple to the Higgs field to obtain mass. From the results of the Super-Kamiokande experiment in 1998 [22], it is evident that neutrino mass is not zero, thus requiring a modification of the SM.

- ***Dark matter and dark energy***: in modern cosmology, numerous observations suggest the presence of an undetected mass and an unknown form of energy. For instance, gravitational lensing [23], where gravity bends the path of light, reveals the effects of substantial dark matter halos. Additionally, the study of galactic rotation curves and the dynamics of galaxy clusters points to a significant quantity of invisible mass. It is necessary to align the observed cosmic microwave background spectrum with the total matter content and the universe's critical density, including dark energy. The universe is estimated to comprise 72.8% dark energy, 22.7% dark matter, and 4.5% ordinary matter [24].

- ***Gravity***: as two fundamental theories in physics, SM is widely considered incompatible with gravity, as described by General Relativity (GR) [25]. One difficulty in formulating a quantum theory of gravity is the emergence of infinite probabilities that cannot be neutralized through renormalization. Unlike the SM based on local gauge symmetries, GR is founded on symmetry under arbitrary smooth coordinate transformations. Additionally, any theory that seeks to unify GR and the SM must be valid from the electroweak scale (around 100 GeV) to the Planck scale (approximately 10^{19} GeV). Several efforts have been made to unify GR and the SM, often called a Theory of Everything (TOE). However, as of now, there is no consensus on any single candidate.

Although no experimental result has been accepted to contradict the Standard Model at the 5 σ level definitively, several experiments have yielded results suggesting the presence of new physics beyond the SM, which would motivate further detailed investigation. Below, we provide some notable results:

- ***Anomalous magnetic dipole moment***: the anomalous magnetic moment, which rises from the deviation between tree-level Feynman loops to the Dirac prediction of the magnetic moment of the particle, is an intrinsic property that provides a good test of the SM framework:
$$\alpha = \frac{g-2}{2}. \tag{2.35}$$
The most recent combined results from the BNL and FNAL muon g-2 experiment at Fermilab show that $\alpha_\mu = 116592059(22) \times 10^{-11}$ [26]. This value shows a discrepancy when compared to the SM prediction $\alpha_\mu = 116591810(43) \times 10^{-11}$ [27]. This suggests potential evidence for new physics. Theoretical work is ongoing to refine the SM predictions with calculations based on lattice QCD to further understand this discrepancy.

- ***Mass of** W **boson***: as one of the most critical parameters of the SM, the CDF collaboration reported a precise measurement of the W boson mass in 2022, finding $m_W = 80433.5 \pm 9.4$ MeV$/c^2$ [28]. This exceeds the SM prediction of $m_W = 80357 \pm 6$ MeV$/c^2$ [29] with a significance of 7 σ. However, the most recent result from the ATLAS experiment shows $m_W = 80366.5 \pm 15.9$ MeV$/c^2$ [30], which aligns with the Standard Model.

- ***X17 anomaly***: the X17 anomaly refers to an unexpected peak observed in data collected by the ATOMKI team [31] when protons were injected into thin lithium-7 targets, producing e^+ and e^- pairs. As shown in Fig. 2.2, the excess was observed at an opening angle of 140 degrees. This excess can be explained by a new vector gauge boson with a mass around 17 MeV, inspiring the name "X17" [32]. However, the X17 anomaly has not yet been confirmed by independent experiments and requires further investigation.

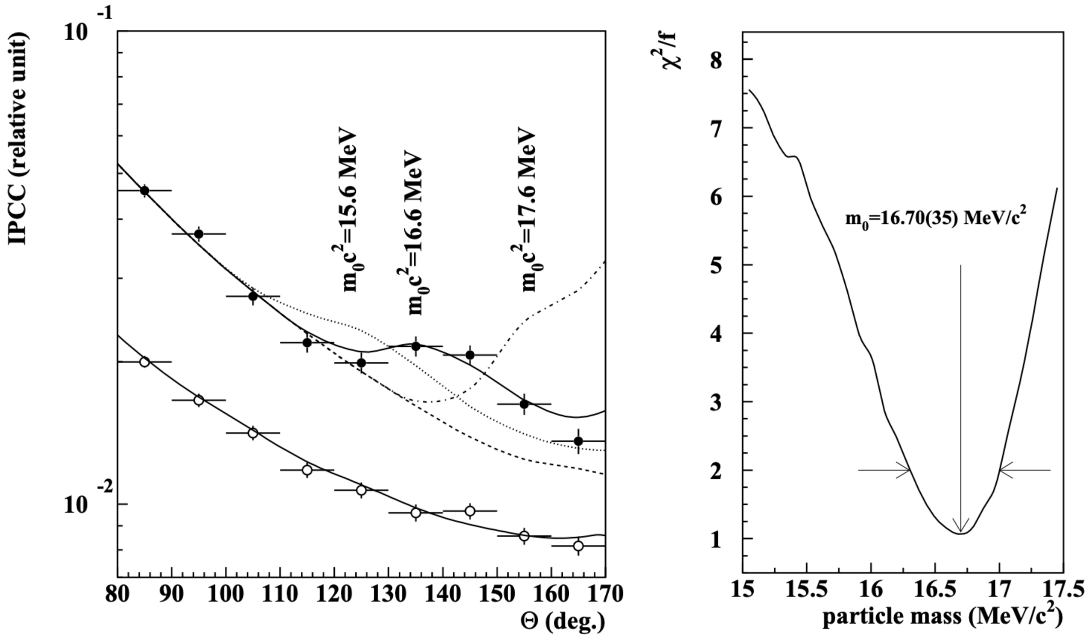
- ***R_D and R_{D^*} anomalies***: a number of experiments involving B meson decays have shown deviations in the ratio of the branching fraction R_D and R_{D^*} defined as:

  $$R_{D^{(*)}} = \frac{\mathrm{BR}(\bar{B} \to D^{(*)}\tau\bar{\nu}_\tau)}{\mathrm{BR}(\bar{B} \to D^{(*)}l\bar{\nu}_l)}, \tag{2.36}$$

  where $l = e$ or μ. An estimate of the world average by Heavy Flavour Averaging Group (HFAG) [33] shows that $R_D = 0.356 \pm 0.029$ and $R_{D^*} = 0.284 \pm 0.013$, which deviate from the SM prediction that $R_D = 0.300 \pm 0.008$ [34] and $R_{D^*} = 0.252 \pm 0.003$ [35].

So far, we have discussed various experimental evidence for BSM theories. One of the possible explanations is the existence of dark photons associated with Higgs bosons. Motivated by the X17 anomaly, and with the current dark photon searches at the LHC limited to $m_{\gamma_d} > 250$ MeV, the work in this thesis aims to extend the dark photons search range to 17 MeV using the $H \to (\gamma_\mathrm{d} \to e^+e^-)(\gamma_\mathrm{d} \to e^+e^-)(+X)$ channel in the ATLAS experiment.

The next chapter will introduce the theoretical framework for dark photons and their event topology in the dark sector models.



(a) (b)

Fig. 2.2 (a) Internal pair conversion correlation (IPCC) measured in the $^7\mathrm{Li}(p,e^+e^-)$ reaction at $E_p = 1.10$ MeV with $-0.5 < y < 0.5$ (closed circles) and $|y| \geq 0.5$ (open circles). Simulation results assuming boson masses of 15.6 MeV (dotted line), 16.6 MeV (full curve), 17.6 MeV (dashed-dotted line), and without boson (dashed line) are overlayed. (b) Determination of boson mass using χ^2/f method [31].

Chapter 3

Dark Matter Search at the LHC

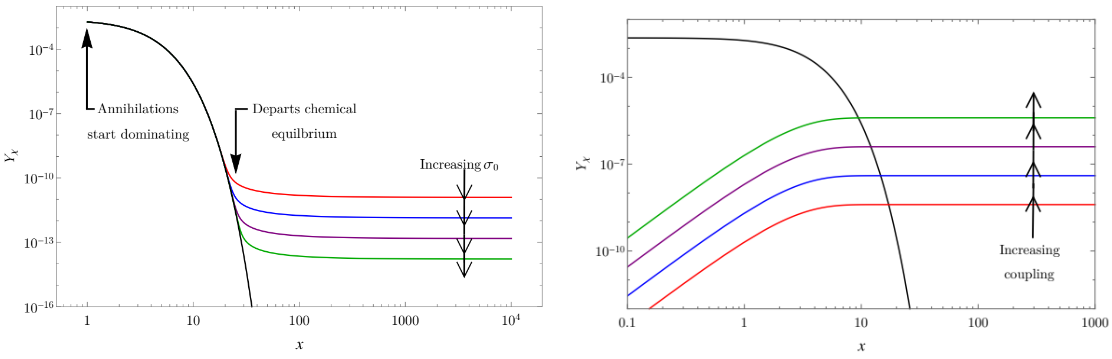
With rapid advancements in theoretical frameworks and improvements in experimental accuracy, we are pushed to the boundary of our current understanding of the universe. Irrefutable experimental evidence from astrophysics points to the existence of unknown dark matter and energy as explained in Section 2.6. Along with the pressing need for alternative BSM theories, it drives us to seek new solutions. One promising pathway is the study of dark matter to uncover new physics and find connections between particle physics and cosmology.

In this chapter, we explore the mechanisms of dark matter and introduce a unique method for investigating it using the ATLAS experiment at the LHC. This approach allows us to probe dark matter in previously unexplored regimes, potentially unlocking insights into new physics.

3.1 Introduction

Dark matter, as its name suggests, does not interact with ordinary matter and is detectable only through its gravitational effects. The first question we must address is how dark matter evolved to reach its current abundance of 22.7% of the universe's mass-energy content. Various scenarios explain its formation mechanism.

- **Freeze-out**: under the assumption that dark matter (χ) was in thermal equilibrium with particle bath (f) in the early universe, $\chi\chi \Leftrightarrow ff$, the production rate and annihilation rate were equal. The annihilation rate decreased as the universe cooled down. This continued until a critical point when the annihilation rate fell below the universe's expansion rate. At this point, the number of particles froze out, and their abundance became fixed. Assuming that annihilations are s-wave dominant, the process can be



(a) (b)

Fig. 3.1 Dark matter abundance in the (a) freeze-out, and (b) freeze-in processes [36].

described using the approximated Boltzmann transport equation:

$$\frac{dY_\chi}{dx} = -\frac{A}{x^2}(Y_\chi^2 - Y_{\chi,eq}^2), \tag{3.1}$$

$$A = \sqrt{\frac{\pi}{45}} g_* m_\chi M_{Pl} \sigma_0, \tag{3.2}$$

where Y_χ is the dark matter yield evolution, and $Y_{\chi,eq}$ is the yield at equilibrium state. $x \equiv \frac{m_\chi}{T}$ where T is the temperature of universe. g_* is the effective relativistic degrees of freedom contributing to the energy density, $M_{Pl} = 1.22 \times 10^{19}$ GeV is the Planck mass and σ_0 is the cross-section. A plot of yield Y as a function of x to visualize the freeze-out process is shown in Fig. 3.1 (a).

- **Freeze-in**: the alternative mechanism assumes the existence of a weak interaction between dark matter particles and SM bath particles. Through rare processes, the yield of dark matter accumulated. As the universe expanded, the rate reached a halt (freeze-in). The plot to visualize the freeze-in process is shown in Fig. 3.1 (b).

The second question is the properties and candidates for dark matter particles. Based on the information collected on dark matter, from the observational evidence of abundance in the universe, the dark matter candidate is expected to be stable with a lifetime comparable with the universe. Additionally, in the freeze-out scenario, the relic density at the initial thermal equilibrium is given by

$$\Omega_\chi \propto \frac{1}{<\sigma v>} \sim \frac{m_\chi^2}{g_\chi^4}, \tag{3.3}$$

where $< \sigma v >$ is the time-averaged annihilation cross section. It is observed that, by coincidence, the electroweak scale (around 100 GeV) and coupling give rise to the canonical value of the cross-section $< \sigma v >= 3.0 \times 10^{-26}$ cm^{3}s^{-1}. The fact that electroweak scale particles could make good dark matter candidates is the well-known *Weakly Interacting Massive Particle (WIMP) miracle*. Thus far, extensive efforts to search for WIMPs have been conducted through direct detection, indirect detection, and particle colliders. Fig. 3.2 schematically shows the different detection approaches.

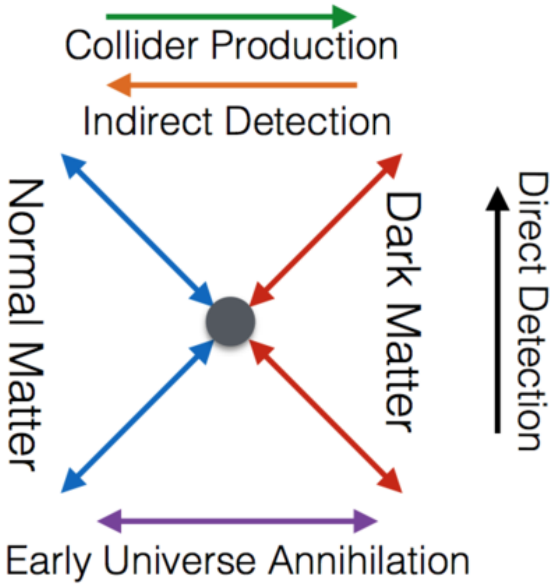


Fig. 3.2 Schematic overview of different dark matter detection approaches [37].

Direct detection methods aim to observe the effects of dark matter directly interacting with ordinary matter. Typically, detectors are shielded deep underground for minimal interference from cosmic rays and other background radiation. When dark matter wind travels from sources like a halo in the Milky Way and collides with a nucleus in the detector, it might impart energy to cause detectable recoils: vibrations detected as rising temperature, excitation, or ionization of recoiled atoms. Fig. 3.3 summarizes the spin-independent WIMP scattering limits and the signals of various models. We also see a neutrino floor where neutrinos produce an irreducible background, limiting the discovery potential of detectors.

Indirect detection focuses on the byproducts of WIMP annihilation or decay rather than the WIMPs themselves. Experiments like the IceCube [39] and the SuperKamiokande [40] collaborations look for excess neutrino fluxes. Fig. 3.4 shows the WIMP spin-dependent cross-section upper limit from different experiments.

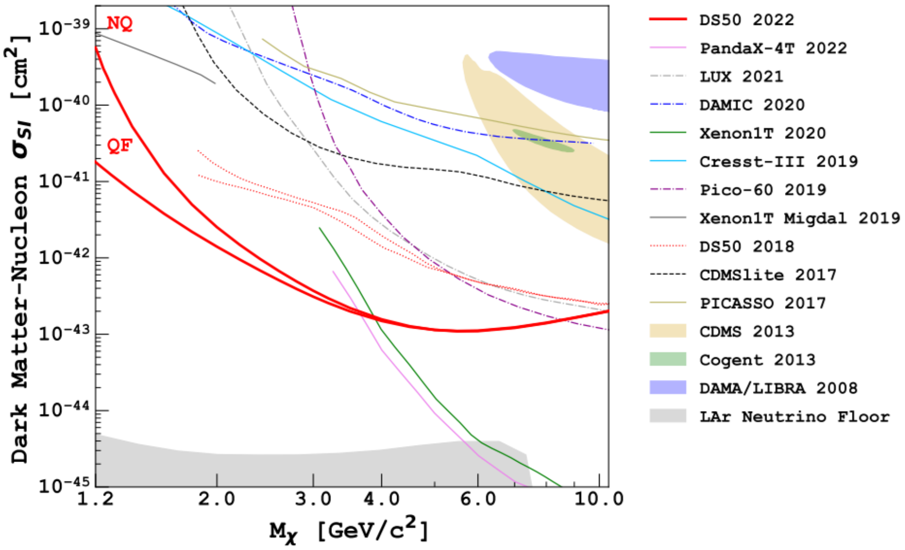


Fig. 3.3 WIMP nucleon spin-independent cross-section experimental bounds [38].

A final strategy to search for dark matter is in the products of high-energy particle colliders like the LHC, providing a completely different approach with much suppressed cosmic backgrounds. The dark matter particles would be undetectable in detectors unless they are produced in association with SM particles. Traditionally, ATLAS [42] and CMS [43] experiments in the LHC exploit the so-called mono-X signature where only one jet or vector boson is detected accompanied by sizeable missing energy in the detector. Fig. 3.5 shows the constraints set by the latest ATLAS mono-X search compared to direct detection experiments.

Extensive efforts were made to look for WIMP. Unfortunately, no evidence of new physics has been found yet. The increasingly stringent limits and trickier models motivate us to look for dark matter models outside the WIMP paradigm. LHC has been pushing the limit of high energy physics on the energy frontier, making TeV scale new physics more and more accessible. On the luminosity frontier, increasingly more data has also been collected. However, we are still left with another possibility: searching for dark matter at the currently reachable scale but from a different approach. One of the most promising methods is using dark sector models. Instead of model-specific dark matter particles, we are shifted to search for signatures of hidden portals linking dark matter and SM.

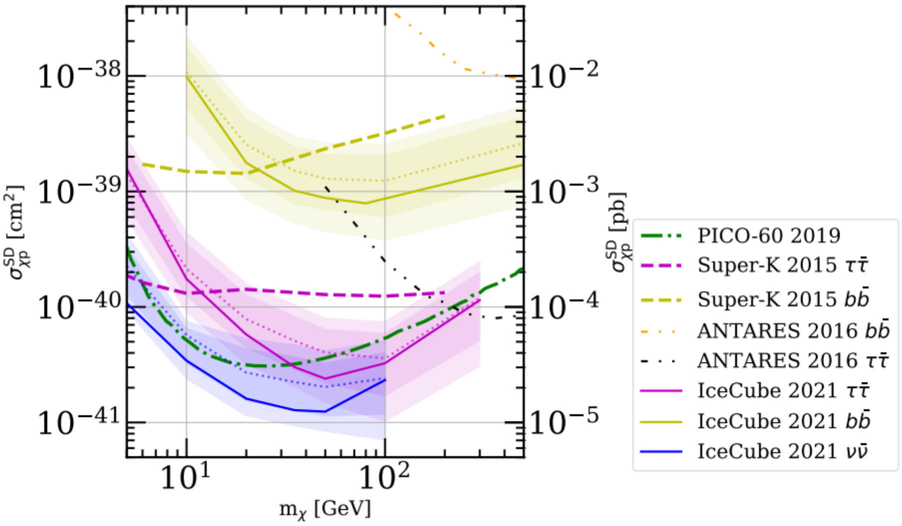


Fig. 3.4 90% upper limits (solid lines) and expected sensitivity (dotted lines) in the spin-dependent cross section as a function of WIMP mass. Figure from [41].

3.2 Dark Sector Models

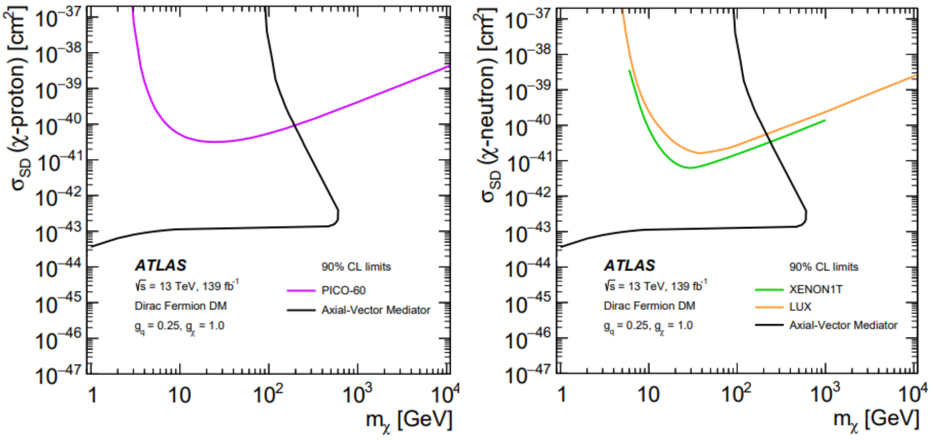
Unlike the WIMP paradigm, dark sector models assume interactions between the dark and visible sectors through the so-called portals, making dark matter particles experimentally detectable. Several well-motivated portals with specific couplings exist, including the vector (spin-1) portal, the neutrino (spin-1/2) portal, the Higgs (scalar) portal, and the axion (pseudo-scalar) portal. Dark sector models shift the focus from the specifics of dark matter composition to the mediator itself, leading to a new model-independent manner and providing rich phenomenology as channels to probe new physics with improved sensitivity.

In the simplest form of such models, a dark vector boson, which we can denote as γ_d and call a dark photon, is required to mediate by describing with the $U(1)_d$ gauge group. In addition, the latest results [2] show that the SM Higgs branching fraction to unknown final states is constrained to be less than 11%, offering unique insights into the Higgs portal through the exotic decay mode of Higgs bosons. As a result, the ATLAS experiment at LHC is motivated and favored to probe both the vector portal and the Higgs portal.

3.2.1 Vector Portal

In the presence of Abelian gauge symmetry $U(1)_d$, the mixing with SM hypercharge $U(1)_Y$ is described by [44]

$$\mathscr{L} = -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}A'_{\mu\nu}A'^{\mu\nu} + \frac{1}{2}\frac{\varepsilon}{\cos\theta_w}A'_{\mu\nu}B^{\mu\nu} + \frac{1}{2}m^2_{A',0}A'_\mu A'^\mu, \tag{3.4}$$



(a) (b)

Fig. 3.5 A comparison of the limit inferred from the ATLAS mono-X result (black line) with the constraints from direct-detection experiments on the spin-dependent (a) WIMP–proton scattering and (b) WIMP–neutron scattering cross sections as a function of the WIMP mass [42].

where $B_{\mu\nu}$ and $A'_{\mu\nu}$ are the field strength tensors for SM photon γ and dark photon γ_{d}, respectively. θ_w is the Weinberg mixing angle. The mixing term associated with the kinetic mixing parameter ε allows coupling between γ_{d} and γ, which can be arbitrarily small. To avoid interactions at the infinite range, we assume γ_{d} to acquire mass through spontaneous electroweak symmetry breaking (see Sec. 2.5), which originates a lifetime for γ_{d}. In such case, the decay width of γ_{d} decaying into pair of charged leptons is given by [45]

$$\Gamma_{\gamma_{\mathrm{d}}\to\bar{l}l} = \frac{4}{3}\varepsilon^2\alpha m_{\gamma_{\mathrm{d}}}\left(1-\frac{4m_l^2}{m_{\gamma_{\mathrm{d}}}^2}\right)^{1/2}\left(1+\frac{2m_l^2}{m_{\gamma_{\mathrm{d}}}^2}\right). \tag{3.5}$$

When $m_{\gamma_{\mathrm{d}}} > 2m_\pi$, hadronic decay is also possible, with a resulting decay width of

$$\Gamma_{\gamma_{\mathrm{d}}\to\mathrm{hadrons}} = \frac{4}{3}\varepsilon^2\alpha m_{\gamma_{\mathrm{d}}}\sqrt{1-\frac{4m_\mu^2}{m_{\gamma_{\mathrm{d}}}^2}}\left(1+\frac{2m_\mu^2}{m_{\gamma_{\mathrm{d}}}^2}\right)R(s=m_{\gamma_{\mathrm{d}}}^2), \tag{3.6}$$

where

$$R = \frac{\sigma_{e^+e^-\to\mathrm{hadrons}}}{\sigma_{e^+e^-\to\mu^+\mu^-}}. \tag{3.7}$$

Here R is evaluated at the center-of-mass energy $\sqrt{s} = m_{\gamma_{\mathrm{d}}}$. The branching ratios of γ_{d} are summarized in Fig. 3.6 for different masses. The resulting mean proper decay length of γ_{d}

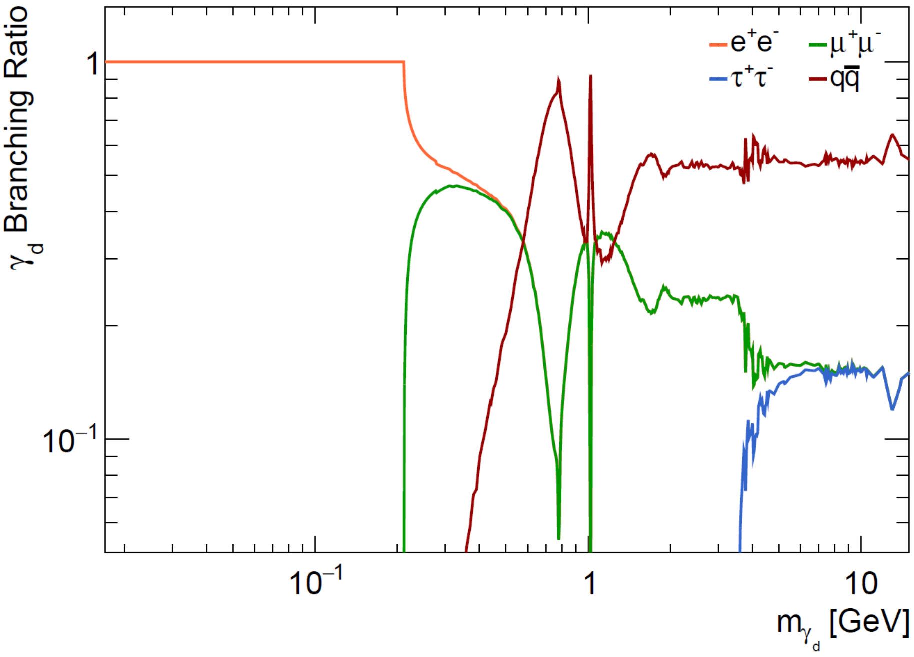


Fig. 3.6 Branching ratio of $\gamma_{\rm d}$ at different mass. For $m_{\gamma_{\rm d}} < 2m_\mu \sim 210$ MeV, $\gamma_{\rm d}$ decays to electrons only. For $m_{\gamma_{\rm d}} > 2m_\mu \sim 210$ MeV, as a consequence of lepton universality, electrons and muons have the same branching ratio. For $m_{\gamma_{\rm d}} > 2m_\pi \sim 279$ MeV, hadronic decays start to enter. The mass resonances correspond to $m_\rho \sim 776$ MeV, $m_\omega \sim 782$ MeV and $m_\phi \sim 1019$ MeV [46].

can be expressed as functions of ε and $m_{\gamma_{\rm d}}$ by the approximation relation [45]

$$c\tau_{\gamma_{\rm d}}\ [\text{mm}] \approx 0.1(\frac{10^{-4}}{\varepsilon})^2(\frac{100}{m_{\gamma_{\rm d}}\ [\text{MeV}]}). \tag{3.8}$$

A summary of latest bounds on the ε and $m_{\gamma_{\rm d}}$ is shown in Fig. 3.7. We can see an undiscovered region around $\varepsilon \lesssim 10^{-3}$ and $m_{\gamma_{\rm d}}$ between the $10 \sim 100$ MeV range where the electron decay is dominated. The target of the search in this thesis is to explore this uncovered region.

3.2.2 Higgs Portal

In analogy with the mechanism of the SM Higgs, $\gamma_{\rm d}$ could acquire mass from Higgs-like boson H_d with non-zero vacuum expectation value. Consider a dark Higgs field S and SM Higgs field ϕ, the coupling of such a dark Higgs boson to the visible sector can be described by

$$\mathscr{L}_{H,H_d}(\phi,S) = -\mu^2|\phi|^2 + \lambda|\phi|^2 - \mu_s|S|^2 + \lambda_s|S|^4 + \kappa|\phi|^2|S|^2, \tag{3.9}$$

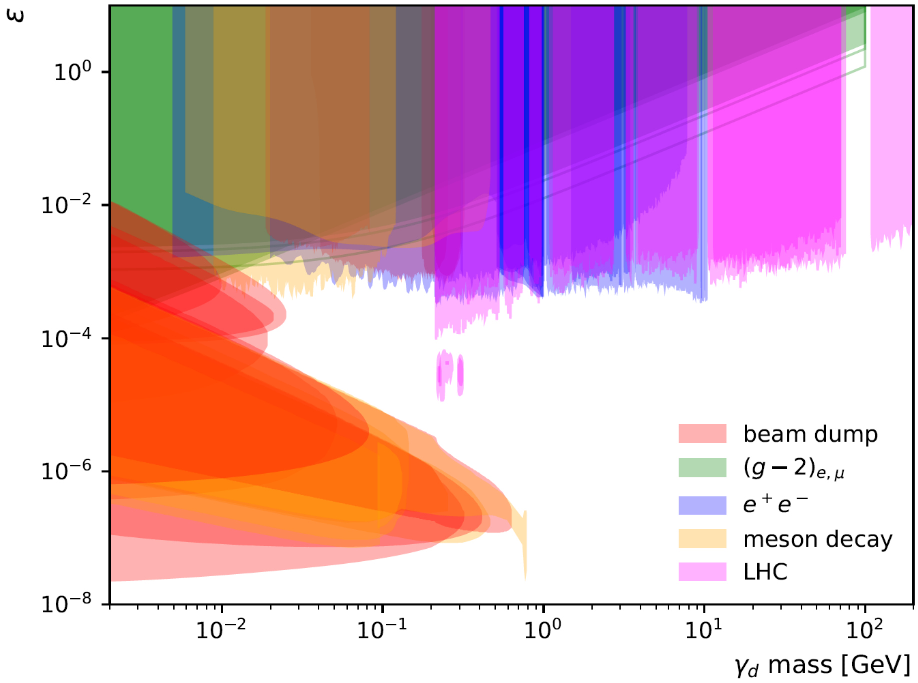


Fig. 3.7 Excluded regions in the ε, m_{γ_d} plane by bump hunt experiments (red), muon magnetic momentum measurements (green), e^+e^- colliders (blue), meson decays (yellow) and LHC (magenta) [47].

where κ is the mixing parameter between the Higgs Boson of the Standard Model and the Higgs Boson in the dark sector. The mixing between ϕ and S allows the decay of SM Higgs bosons to a pair of γ_ds via the so-called Higgs portal. The decay width of Higgs boson decaying into a pair of γ_d is given by [45]:

$$\Gamma_{H\to\gamma_\mathrm{d}\gamma_\mathrm{d}} = \theta_H^2 \times \frac{\alpha' m_H^3}{8m_{\gamma_\mathrm{d}}^2}\sqrt{1-\frac{4m_{\gamma_\mathrm{d}}^2}{m_H^2}}\left(1-\frac{4m_{\gamma_\mathrm{d}}^2}{m_H^2}+\frac{12m_{\gamma_\mathrm{d}}^4}{m_H^4}\right). \qquad (3.10)$$

The mixing angle $\theta_H \approx \frac{\kappa v w}{m_H^2 - m_{H_\mathrm{d}}^2}$, where v and w are the VEVs of SM Higgs and dark Higgs, and α' is the $U(1)_d$ structure constant. Through the κ term, the Higgs can also decay into a dark Higgs pair if $m_H > 2m_{H_\mathrm{d}}$, while the opposite is possible if $m_{H_\mathrm{d}} > 2m_H$. The decay width is given by [48]:

$$\Gamma_{H\to H_\mathrm{d}H_\mathrm{d}} = \frac{(m_H^2+2m_{H_\mathrm{d}}^2)^2\sin^2 2\theta_H}{128\pi m_H}\sqrt{1-\frac{4m_{H_\mathrm{d}}^2}{m_H^2}}\left(\frac{1}{w}\cos\theta_H+\frac{1}{v}\sin\theta_H\right)^2. \qquad (3.11)$$

The decay of Higgs into a dark Higgs and a dark photon is also possible.

In the dark photons production mode through Eq. 3.10, γ_{d} mixes with γ and decays into fermions via the vector portal, as mentioned in Sec. 3.2.1. This way, the Higgs portal dark sectors can be searched at the LHC together with the vector portal. The Higgs portal is particularly interesting for experiments at the LHC since Higgs bosons can be produced. In addition, since an upper limit of BR for the invisible decays of Higgs boson is set to $BR = 11\%$ [2], there is enough possibility to produce dark photons from Higgs bosons. In this analysis, we take the branching fraction as a free parameter to probe the Higgs portal instead of the mixing parameter κ.

3.2.3 Benchmark Models

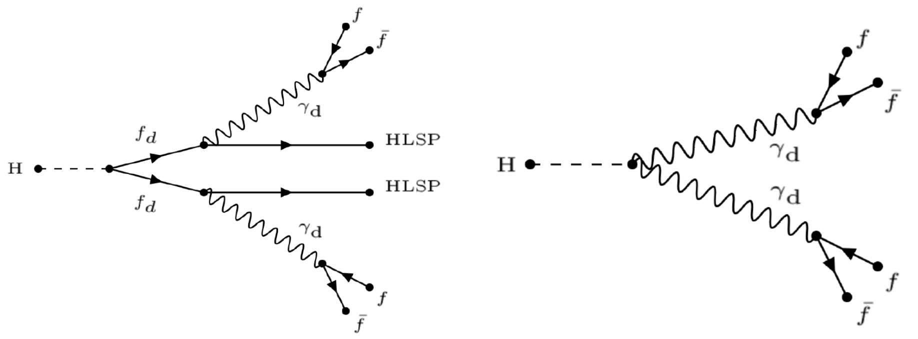
The search strategy for dark sectors focuses on utilizing the γ_{d} signature, which represents a wide range of scenarios, rather than relying on any specific model. We use simplified and widely recognized dark sector models as benchmarks to investigate vector and Higgs portals, aiming to optimize object selection and sensitivity throughout the analysis. In this thesis, we select the Hidden Abelian Higgs Model (HAHM) [49] and Falkowsky-Ruderman-Volansky-Zupan (FRVZ) model [50][51] as benchmarks.

FRVZ benchmark Model

The FRVZ benchmark model is a simplified dark sector model with a richer phenomenology due to the production of multiple BSM particles. As depicted in Fig. 3.8 (a), Higgs bosons decay to a pair of dark fermions f_{d}. Each f_{d} further decays until the production of the Hidden Lightest Stable Particle (HLSP) and a dark photon γ_{d}. Here, fermions f_d correspond to the Lightest Supersymmetric Particle (LSP) in the Minimal Supersymmetric Standard Model (MSSM) sector, $\tilde{N}_1$, and the HLSP corresponds to the “true LSP” in the original paper [50]. Dark photons are mixed with SM photons via the vector portal and finally decay into pairs of charged fermions. Various topologies are possible due to the BSM intermediate and hidden particles in the final state. In addition, depending on the strength of $U(1)_d$ coupling, more than one γ_{d} can be radiated, resulting in the multi γ_{d} production like a shower. In this study, we assume $\alpha_d \lesssim 0.01$ to avoid additional radiations of γ_{d} in the final state.

HAHM benchmark Model

The HAHM model is another simplified dark sector model. As depicted in Fig. 3.8 (b), in this model, the Higgs boson mixes with the dark Higgs to produce two γ_{d} via the Higgs portal. γ_{d} then decays into two charged fermions via the vector portal. The smaller number of particles involved makes simpler topology, allowing, e.g., the reconstruction of Higgs boson mass.



(a) (b)

Fig. 3.8 Feynman diagram of (a) FRVZ benchmark model and (b) HAHM benchmark model. Here *HLSP* stands for the hidden lightest stable particle, and f_{d} for dark fermion [52].

Search Signature

In addition to providing unique insights into the Higgs portal, the ATLAS experiment at the LHC is also motivated to explore the uncovered vector portal exclusion area, as shown in Fig. 3.7, particularly around $\varepsilon \lesssim 10^{-4}$ and $m_{\gamma_{\mathrm{d}}}$ in the $10 \sim 100$ MeV range. To maximize the ATLAS detector's potential to access this phase space, a dedicated γ_{d} signature is exploited by choosing the free parameters $m_{f_{\mathrm{d}}}$ and m_{HLSP} to be a few GeV, ensuring they are small relative to the Higgs boson, with $m_{HLSP} + m_{\gamma_{\mathrm{d}}} < m_{f_{\mathrm{d}}}$ (more details will be provided in Sec. 5.1). Similar assumptions are made in other experiments [4] [53]. As a result, the two γ_{d}s are boosted and well separated, and the charged fermions from γ_{d} are highly collimated, often falling into the same region of interest. In this thesis, since we focus on the search signature of low-mass dark photons, only pairs of electrons and muons are considered in the final states. Such a collimated lepton pair decaying from γ_{d} forms a **lepton jet**, or LJ. Fig. 3.9 and Fig. 3.10 show the event topology where two γ_{d}s from Higgs boson are produced back-to-back, and lepton pairs decaying from each γ_{d} form a lepton jet in the FRVZ and HAHM benchmark model, respectively.

Previous results at the LHC to explore Higgs and vector portals using the FRVZ model are shown in Fig. 3.11. This thesis is devoted to extending the search to cover the gap by exploiting the lepton jet signature in the ATLAS experiment. As mentioned, we focus on two dominant decay modes of γ_{d}, $\gamma_{\mathrm{d}} \to e^+e^-$ and $\gamma_{\mathrm{d}} \to \mu^+\mu^-$ for the low mass region (by 2 × charged pion mass). Since there are two γ_{d}s in an event, there are three possible channels: eLJ-eLJ, μLJ-μLJ and eLJ-μLJ. The author has worked on eLJ-eLJ, but the results from other channels will be shown in Chapter 9. We will explain the object building of both muLJ and eLJ until Sec. 5.4.1, but after that, we will focus on eLJ-eLJ.

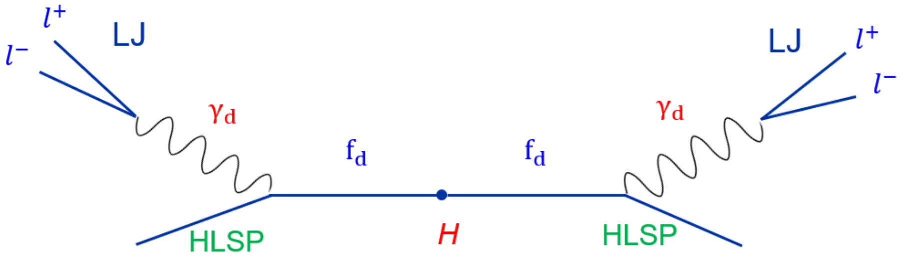


Fig. 3.9 Event topology of lepton jet signature of our interest in the FRVZ benchmark model. Two γ_d from Higgs boson are produced back-to-back, and the lepton pair decaying from γ_d are collimated.

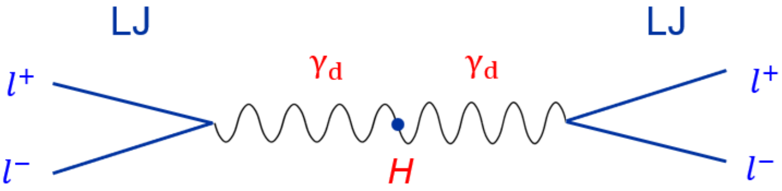


Fig. 3.10 Event topology of lepton jet signature of our interest in the HAHM benchmark model. Two γ_d from Higgs boson are produced back-to-back, and the lepton pair decaying from γ_d are collimated.

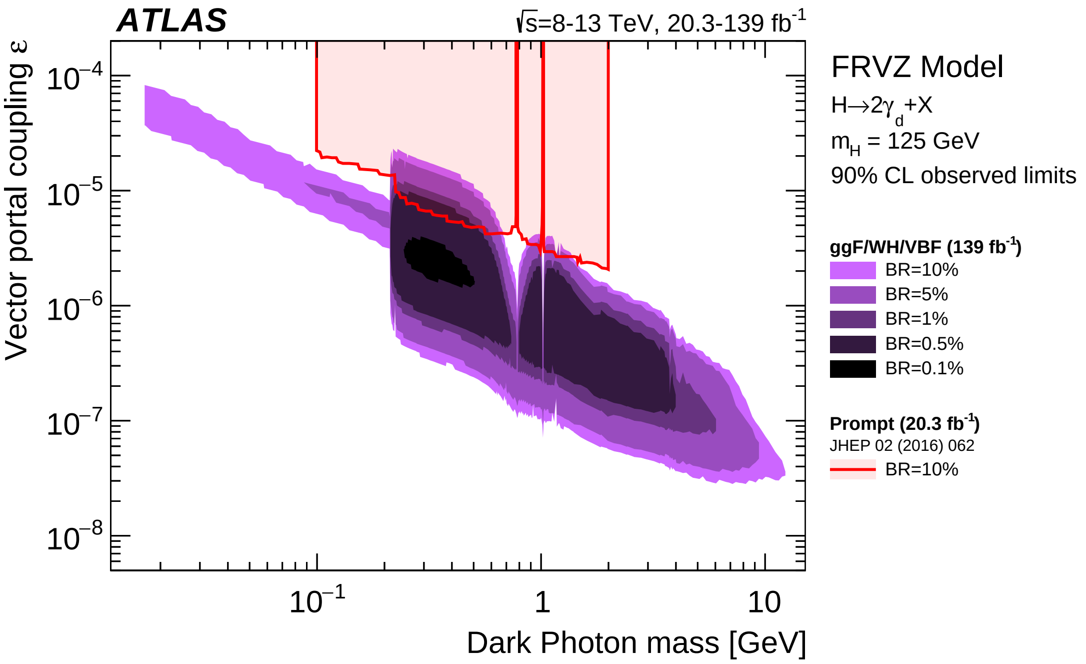


Fig. 3.11 Excluded areas in the ε, m_{γ_d} plane the FRVZ model with different $BR(H \to 2\gamma_d + X)$ assumptions [54].

Chapter 4

The ATLAS experiment at LHC

As one of the world's largest experiments in high-energy physics, the ATLAS experiment provides a unique chance to experimentally validate our current theoretical framework and search for new physics. This chapter briefly introduces the ATLAS detector at the LHC and the definition of standard physics objects.

4.1 Accelerator and the LHC

CERN, *the European Organization for Nuclear Research*, is one of the leading particle physics research organizations established in 1954. Located on the French-Swiss border near Geneva, CERN's accelerator complex consists of a series of machines designed to boost particle energies. During LHC Run-2 operations, the acceleration process begins at the Linear Accelerator 2 (LINAC2) [55], where protons from hydrogen gas are accelerated to 50 MeV. The proton bunches then pass through a series of pre-acceleration stages: the Proton Synchrotron Booster (PSB) [56], the Proton Synchrotron (PS) [57], and the Super Proton Synchrotron (SPS) [58], increasing their energies to 1.4 GeV, 25 GeV, and 450 GeV, respectively, before being injected into the final stage, the LHC ring, achieving the world's highest recorded collision energy of 13.6 TeV [59]. In the LHC ring, there are four locations where proton bunches collide:

- A Toroidal LHC Apparatus (***ATLAS***) [60] and Compact Muon Solenoid (***CMS***) [61] are general-purposed detectors designed to observe the phenomenon of Higgs boson and search for new physics beyond the SM.
- A Large Ion Collider Experiment (***ALICE***) [62] is designed to study the heavy ion collisions. By creating a state of quark-gluon plasma, properties of strong interaction can be explored.

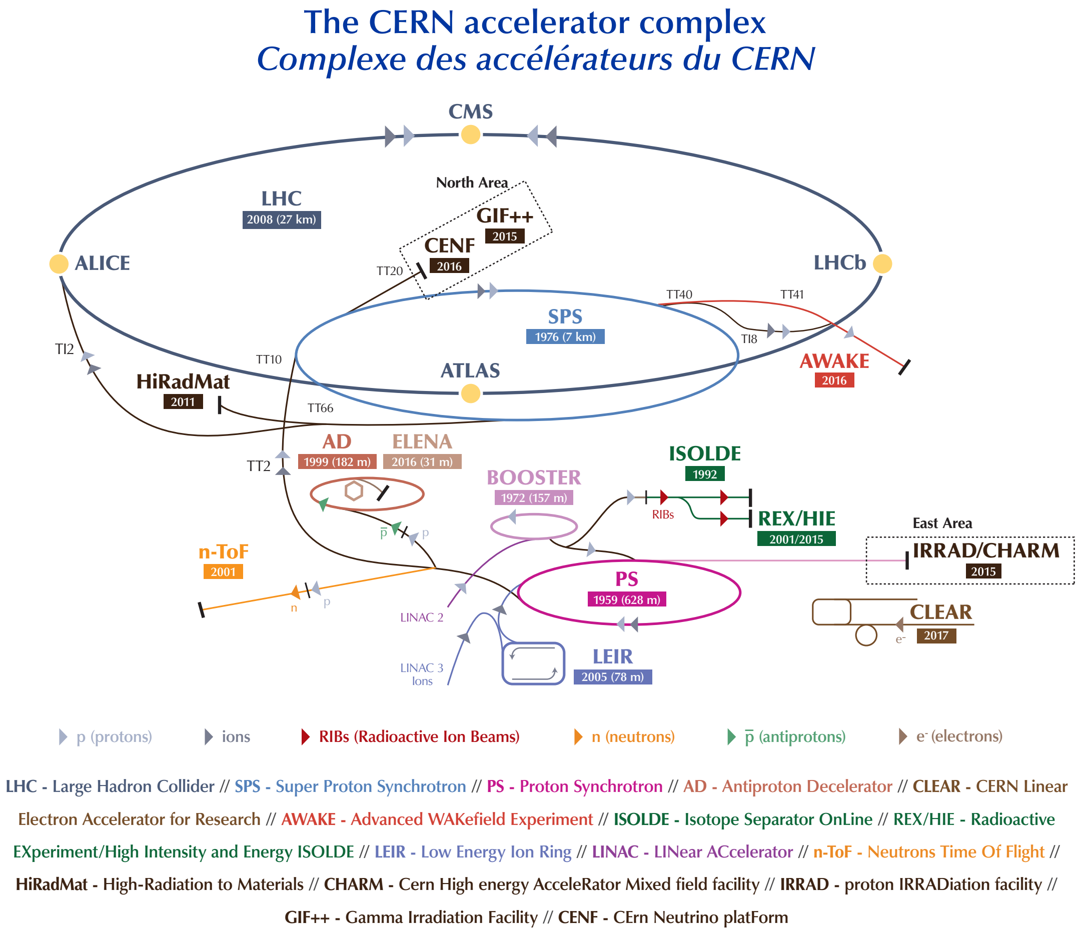


Fig. 4.1 Schematic overview of CERN accelerator complex [64].

- Large Hadron Collider beauty (***LHCb***) [63] is dedicated to the study of beauty quarks, as well as exotic hadrons, charm physics, electroweak physics, QCD physics, and other new physics searches.

4.1.1 pp collisions

In designing the LHC, proton-proton (pp) collisions were chosen to explore the high-energy frontier in considerations related to synchrotron radiation given by the relativistic Larmor formula

$$P_\gamma = \frac{q^2 c}{6\pi\varepsilon_0 c^3}\frac{a^2}{(m_0 c^2)^2} = \frac{q^2 c}{6\pi\varepsilon_0 c^3}\frac{1}{(m_0 c^2)^4}\frac{E^4}{\rho^2}, \tag{4.1}$$

which describes the energy loss by accelerated charged particles. Here, a is the acceleration, and ρ is the radius of curvature in circular motion. The mass term shows that electrons in an accelerator lose energy approximately 10^{13} times that of protons, making proton collisions good candidates for machines operating at very high energy. In the typical pp collisions, two partons interact primarily in *hard scattering* process where large momentum is transferred. Fig. 4.2 illustrates a pp collision. Here, the incoming protons are represented by the large green ellipse from the horizontal direction, and the hard scattering process is in the red circle. In addition, gluons and quarks can be emitted by the QCD radiation processes both before the collision as initial state radiation (ISR) and after the collision as final state radiation (FSR). ISR and FSR are later important in reconstructing and analyzing the collision events. The products further emit radiation, causing *parton showers* until hadronization shown as the green circles, resulting in bundles of hadrons detected as jets.

4.1.2 Data-taking

The operation of the LHC can be divided into data-taking periods of Run-1 (2011-2012), Run-2 (2015-2018), and the ongoing Run-3 (2023-), where stable beams are collided and recorded, and long-shutdown (LS) periods for repairs and upgrades as summarized in Fig. 4.3. The collider is operated at $\sqrt{s} = 7$ and 8 TeV during Run-1, 13 TeV during Run-2, and 13.6 TeV during the ongoing Run-3.

4.1.3 Luminosity

To measure the number of collisions produced in the detector, we define instantaneous luminosity L as the number of events per time per cross-section:

$$L = \frac{1}{\sigma}\frac{dN}{dt}. \tag{4.2}$$

A simple estimation is given by

$$L = \frac{kfN_1N_2}{4\pi\sigma_x\sigma_y}, \tag{4.3}$$

where N_1 and N_1 are the number of protons in the two bunches. k and f are the number and frequency of bunches circulating the LHC ring, and $\sigma_x(\sigma_y)$ are the $x(y)$ components of the cross-section. The instantaneous luminosity value is 1×10^{34} cm^{-2}s^{-1} with the design condition of LHC [66]. The integrated luminosity at a certain time is given by

$$L_{\text{int}} = \int L dt. \tag{4.4}$$

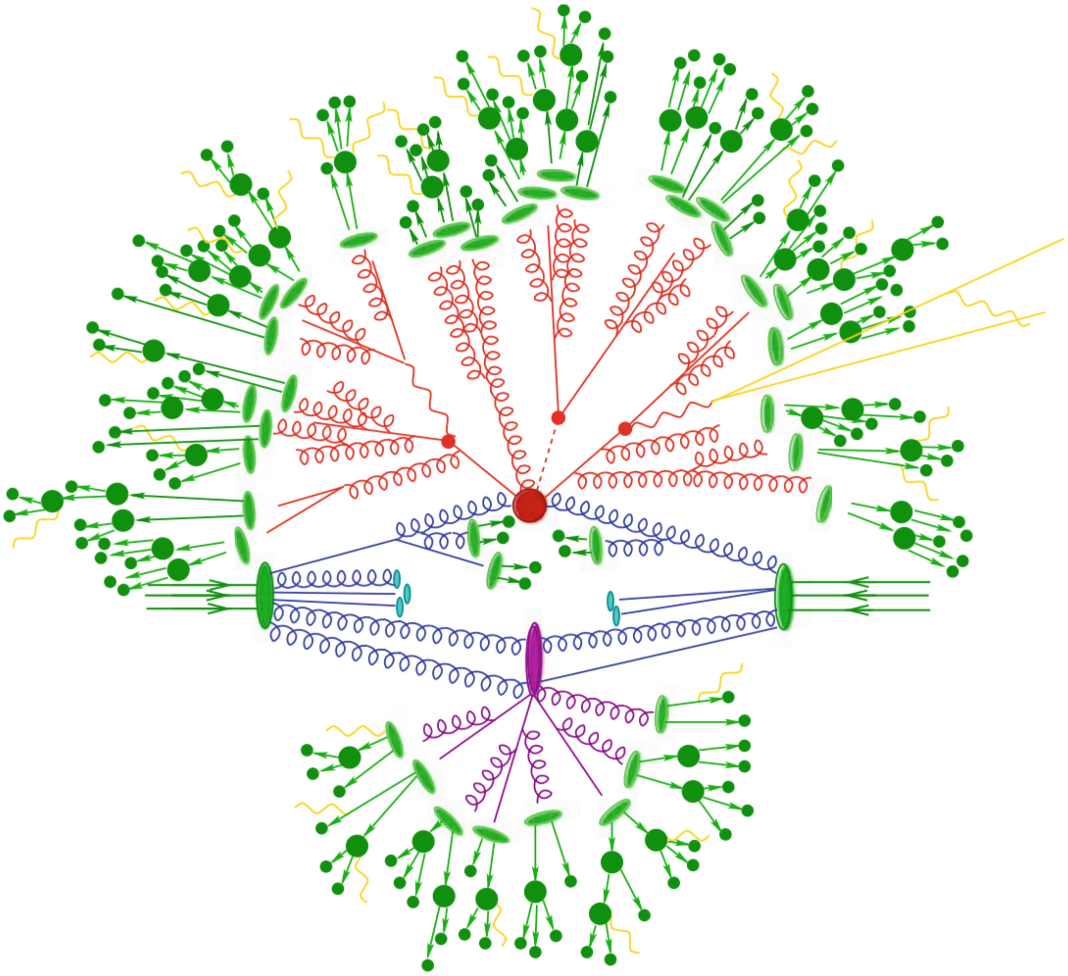


Fig. 4.2 Illustration of *pp* collision at LHC [65]. Two partons (large green ellipses) interact primarily in hard scattering (large red circle) and are accompanied by a secondary interaction (large purple ellipse). Green circles represent the hadron showering and hadronization. Electromagnetic radiation by charged particles is in yellow.

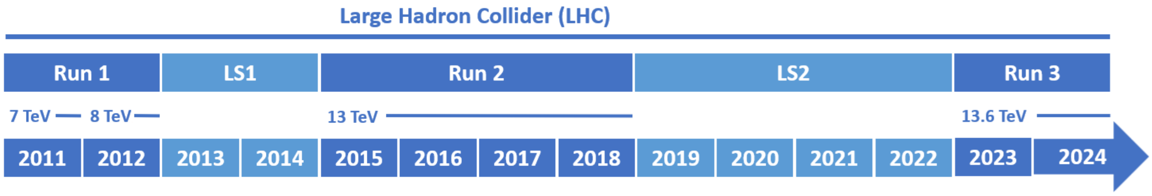
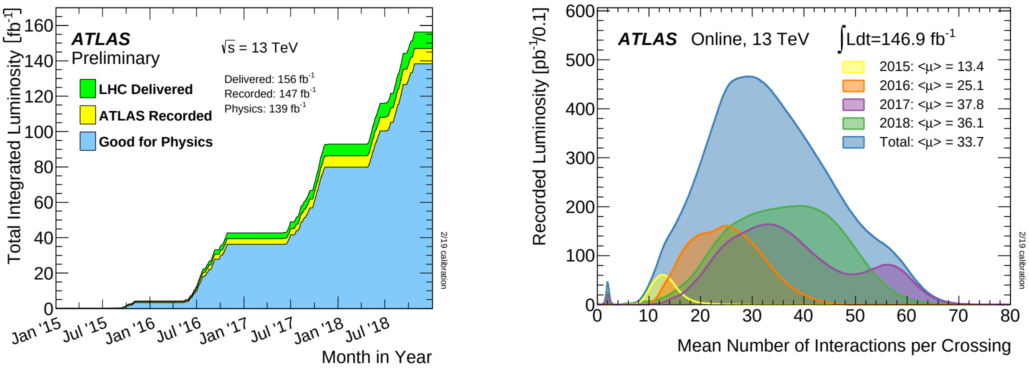


Fig. 4.3 Timeline of LHC operations.



(a) (b)

Fig. 4.4 (a) Total Integrated Luminosity for *pp* collisions, and (b) mean number of interactions per crossing $\langle\mu\rangle$ for the 2015-2018 LHC Run-2 *pp* collision data at 13 TeV center-of-mass energy [69].

Fig. 4.4 (a) shows the integrated luminosity during LHC Run-2. By combining the luminosity blocks of data collected where all ATLAS sub-detectors operated in good condition (good run list, or GRL), a total of $L_{\text{int}} = 139$ fb^{-1} is primarily measured with an uncertainty of 1.7% [67]. The latest study on luminosity determination recovers 1 fb^{-1} to have $L_{\text{int}} = 140$ fb^{-1} with an uncertainty of 0.83% [68]. In this thesis, we will use the latest total of $L_{\text{int}} = 140$ fb^{-1} *pp* collision data taken from 2015 to 2018 for physics analysis.

4.1.4 Bunch-crossing and Pile-up

At the LHC, protons are injected in consecutive bunches with approximately 7.5 m between each other. As a result, each revolution in the 27 km ring has 3564 bunch-crossings. For each bunch-crossing (BC), multiple protons could collide with each other at the interaction point, including both hard-scatter events we are interested in, as well as inelastic *pp* interactions. To quantitatively describe such a phenomenon, pile-up is used to describe the mean number of interactions per bunch-crossing:

$$\langle\mu\rangle = \frac{L_{\text{int}}\sigma}{R_{BC}}, \tag{4.5}$$

where R_{BC} is the bunch-crossing rate. Fig. 4.4 (b) shows the mean number of interactions per crossing during the 2015-2018 LHC Run-2 data-taking periods. Table. 4.1 summarizes the run parameters.

Table 4.1 Parameter values from 2015 to 2018.

Parameter	**2015**	**2016**	**2017**	**2018**
$\sqrt{s}$ [TeV]	13	13	13	13
Bunch spacing [n]	25	25	25	25
Peak luminosity [10^{33} cm^{-2}s^{-1}]	5	13	16	19
Integrated luminosity [fb^{-1}]	4.0	38.5	50.2	63.4
Pile-up ($\langle\mu\rangle$)	~ 13	~ 25	~ 38	~ 36

4.2 The ATLAS detector

4.2.1 Overview

The ATLAS detector is located at the Point-1 interaction point on the LHC ring. Cylindrical in shape, the detector is measured 46 m long, 25 m in diameter, 7000 t in weight, and sits in a cavern 100 m below ground. As shown in Fig. 4.5, it consists of a series of sub-detector layers:

- ***Inner Detector***: closest to the beam pipe, it consists of pixel detectors, a silicon tracker, and a transition radiation tracker to provide precise measurement of charged particle trajectories.
- ***Calorimeter***: the electromagnetic calorimeter captures electrons and photons in EM showers, while the hadronic calorimeter absorbs hadrons to measure jet energy depositions. Missing transverse energy (missing E_T) is also reconstructed in the calorimeter.
- ***Muon Spectrometer***: identification and measurement of muon tracks.
- ***Magnet system***: a central solenoid for the inner detector and superconducting toroids magnet for the muon spectrometer.

The details of each subsystem are explained in the following sections.

4.2.2 Coordinate system

The interaction point is the origin of the default coordinate system used in the ATLAS experiment. The z-axis runs along the beam direction, while the transverse x-y plane is defined such that the positive x-axis points towards the center of the LHC ring, and the y-axis points upwards to the surface. The part of the detector with positive z values is called the A-side, while the part with negative z values is called the C-side. The ϕ and θ are the azimuth and polar angles as defined in Fig. 4.6.

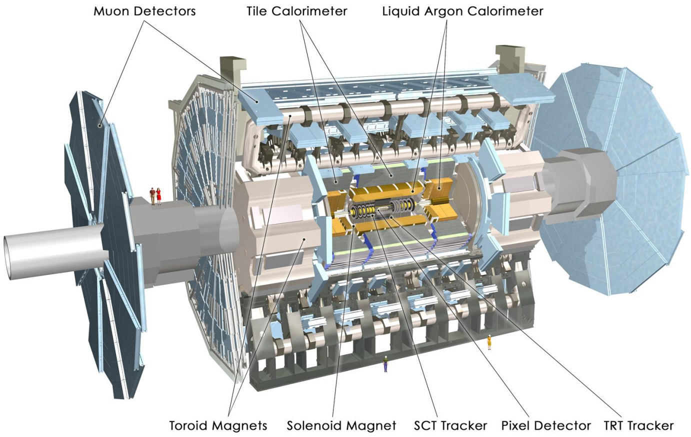


Fig. 4.5 Schematic view of the ATLAS detector [70].

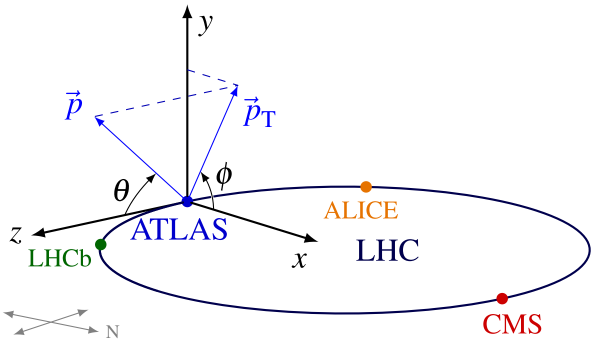


Fig. 4.6 Illustration of the coordinate system of the ATLAS detector [71].

Additionally, we introduce the spatial coordinate of *pseudorapidity* η to describe the angle of a particle relative to the beam axis, defined as

$$\eta \equiv -\ln\left[\tan\left(\frac{\theta}{2}\right)\right]. \tag{4.6}$$

The angular distance between objects in the detector can be described with $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

4.2.3 The magnet system

The magnet system is essential to the measurements of charged particle tracks. An intense magnetic field means more precise measurements of high-p_T object. The ATLAS experiment features a powerful magnet system with four large superconducting magnets. Fig. 4.7 shows the geometry of the magnet system.

- The ***central solenoid*** is 5.8 m long, 2.56 m in diameter, located inside the calorimeter in parallel with beamline, and provides a magnetic field of 2 T in z direction for the inner detector. The solenoid is structured with just 4.5 cm thickness to minimize the radiation length before the calorimeter.
- The ***barrel toroid*** consists of 8 separate coils to create a magnetic field in the ϕ direction up to 3.5 T for the track measurements in the barrel region of the muon spectrometer.
- Two ***end-cap toroids*** provide a magnetic field in the end-cap region of the muon spectrometer. The eight end-cap coils on each side are positioned interleaving with the barrel toroids.

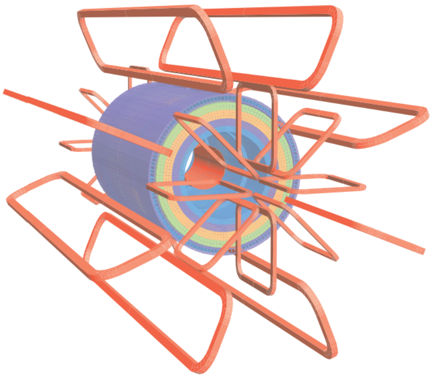
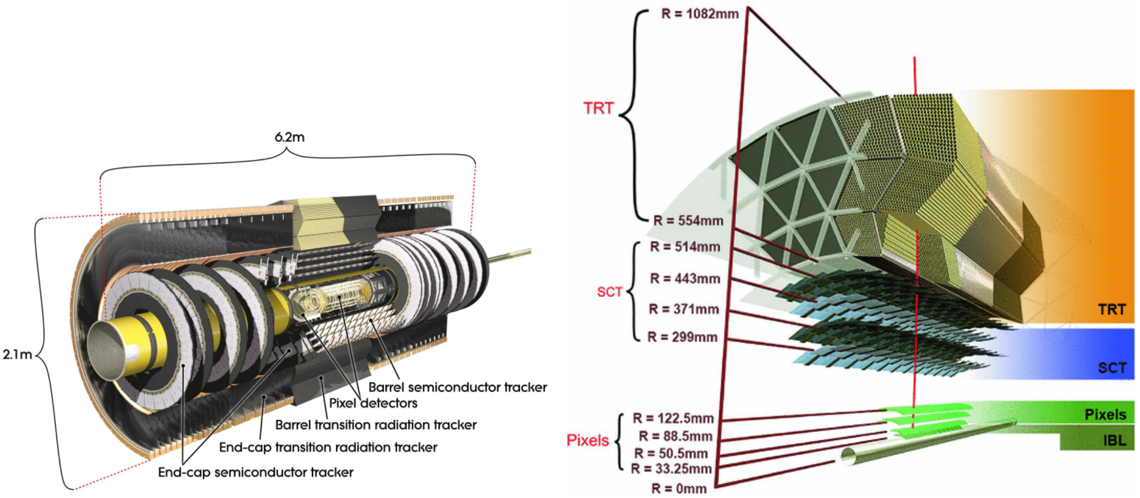


Fig. 4.7 Schematic view of magnet system in ATLAS [72]. The solenoid winds inside the calorimeter are purple. At the barrel region, eight toroid coils interleave with end-cap coils.

4.2.4 Inner Detector

The ATLAS inner detector (ID) [73] is designed to provide high-precision measurements of track transverse momentum and primary and secondary vertexes. The coverage of ID is $|\eta| < 2.5$ and full range of ϕ with a required p_T resolution of $\sigma_{p_\mathrm{T}}/p_\mathrm{T} = 0.05\% p_\mathrm{T}[\mathrm{GeV}] \oplus 1\%$ [72]. An overview of ID is shown in Fig. 4.8 (a), and increasing in radius, different layers are shown in Fig. 4.8 (b).

- ***Pixel Detector***: Using silicon microstrip and pixel technology, a total of 1744 pixel sensors are arranged in three barrel layers and two end-caps to cover the innermost three layers at 50.5 mm, 88.5 mm, and 122.5 mm in radius. Pixel detectors provide a 12 μm resolution in the $r-\phi$ plane and 66 μm in the z direction. After the Long Shutdown after Run-1 during 2013-2015, an additional layer, the insertable B-layer (IBL), was installed at a radius of 33.3 mm, significantly improving the ID performance.



(a) (b)

Fig. 4.8 Schematic view of different layers of Inner Detector [72] [74].

- ***Semiconductor Tracker (SCT)***: In the intermediate layers, SCT works with a Pixel detector to further provide a measurement of momentum, impact parameters, and vertex position. SCT consists of 4088 modules, each with 768 readout strips arranged in 4 barrels at 299 mm, 371 mm, 443 mm, and 514 mm, and two endcaps of 9 disks each. Slightly worse than pixel detector, SCT provides a resolution of 16 μm in $r-\phi$ plane and 580 μm in z.

- ***Transition Radiation Tracker (TRT)***: The outermost layer consists of straw detectors filled with Xenon gas. As the Xenon gas absorbs transition radiated photons, it produces distinct readout signals. These signals provide significant discrimination power between electrons and charged pions within the energy range of 1 GeV to 100 GeV. There are 52544 straw tubes arranged in 1.5 m length in the barrel region to cover $|\eta| < 1$, and 122880 straw tubes in the endcaps to cover 0.8 m $< |z| <$ 2.7 m and $1 < |\eta| < 2$, each with a diameter of 4 mm. By adding around 36 hits per track, TRT improves momentum measurement, and overall, TRT has a resolution of 130 μm.

4.2.5 Calorimeter

The calorimeter [75] is the critical component of the detector to provide precision measurement of energy deposition, where the fine granularity of the EM calorimeter is designed for electrons and photons. In contrast, the hadronic calorimeter absorbs the hadronic showers, preventing leakage to the muon spectrometer for jets. As shown in Fig. 4.9, the calorimeter system includes EM calorimeter, hadronic end-cap (HEC), forward calorimeter (FCal) built from liquid Argon (LAr) cryostats, and scintillator tile calorimeter:

- ***LAr EM Calorimeter***: the EM calorimeter can be divided into two parts: EM barrel (EMB) with $|\eta| < 1.475$ and two EM end-caps (EMEC) with $1.375 < |\eta| < 3.2$. The EM part is a lead-LAr detector with accordion shaped electrodes and lead absorbers, as shown in Fig. 4.10. The accordion geometry naturally allows full coverage of ϕ without cracking. The wave amplitude and folding angle vary with radius to provide linearity in ϕ. The EM calorimeter consists of a presampler and several active layers: three layers in the precise measurement region, $|\eta| < 1.35$ for EMB and $1.5 < |\eta| < 2.5$ for EMEC; two layers in the overlap $1.35 < |\eta| < 1.5$ region and high $|\eta| > 2.5$ region. Usually, the middle layer collects the largest fraction of energy from electrons and photons. The finest granularity is 0.025×0.025 in $\Delta\eta \times \Delta\phi$, in the middle layer of precision measurement regions.

- ***Scintillator Tile Calorimeter***: located right outside the EM calorimeter to measure hadronic showers in the barrel region, the tile has a barrel part in $|\eta| < 1.0$ and two extended barrels in $0.8 < |\eta| < 1.7$. It uses scintillating tiles as active materials and steel as an absorber. There are three layers, and the granularity is 0.1×0.1 in the first two layers and 0.2×0.2 in the last layer in $\Delta\eta \times \Delta\phi$.

- ***LAr Hadronic End-cap***: the HEC is located right behind EMEC to cover $1.5 < |\eta| < 3.2$. Parallel copper plates are interleaved with the LAr gaps to provide an active medium for the sampling. HEC consists of 4 layers with a coarse granularity of 0.1×0.1 and 0.2×0.2 in $\Delta\eta \times \Delta\phi$ in the low and high η regions. There are 5632 readout channels in total.

- ***LAr forward Calorimeter***: forward calorimeter covers $3.1 < |\eta| < 4.9$ as the extension of end-cap calorimeters, and to reduce background entering the muon spectrometer. FCal consists of 3 layers with 3524 readout channels on both sides.

4.2.6 Muon Spectrometer

The Muon Spectrometer (MS) [77] in ATLAS detects muon tracks in the presence of large superconducting toroid magnets and consists of separate high-precision tracking chambers and triggering chambers. The main components are shown in Fig. 4.11. At the barrel region, MS is structured with three cylindrical layers around the beam axis with the magnetic field provided by the large barrel toroid at $|\eta| < 1.4$. At the endcap region for $1.6 < |\eta| < 2.7$,

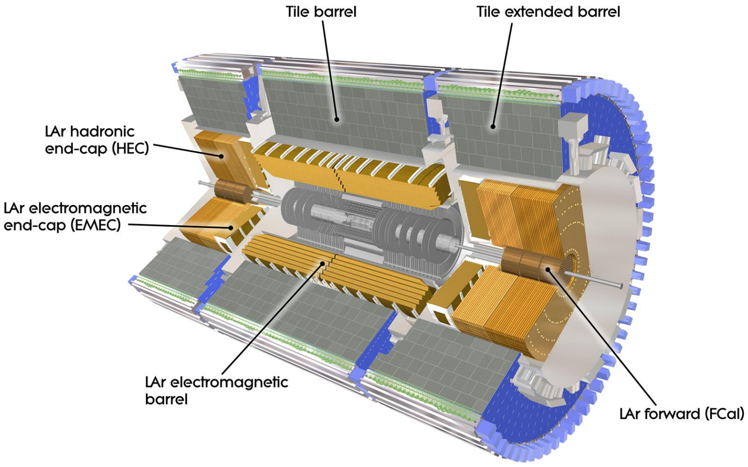


Fig. 4.9 Illustration of the calorimeter in ATLAS, which consists of LAr EM Calorimeter, Scintillator Tile Calorimeter, LAr Hadronic End-cap, and LAr forward Calorimeter [76].

the chambers are installed in three layers perpendicular to the beam with a magnetic field provided by the endcap toroids. Finally, in the transition area ($1.4 < |\eta| < 1.6$), barrel and endcap toroids contribute to the magnetic field. The precision tracking chambers include Monitored drift tubes (MDT) and Cathode strip chambers (CSC) while triggering chambers include Resistive plate chambers (RPC) and thin gap chambers (TGC). Precision tracking chambers function to precisely track the muon track coordinates from the magnetic field bending direction. In contrast, triggering chambers provide bunch-crossing identification (BCID) for synchronizing muon detection, fast trigger information for the later high-level triggers, and secondary coordinate information to complement MDT precision measurements.

- ***MDT*** covers the range of $|\eta| < 2.7$ except for the innermost layer where CSC replaces them for $2 < |\eta| < 2.7$. MDT is built from an aluminum drift tube with a diameter of 29.970 mm and 0.4 mm thickness, operating with a mixture of Argon and CO_2 gas at 3 bar to track muons passing the volume.

- ***CSC*** is designed to ensure safe operation at high particle fluxes, with higher counting rates of 1000 Hz/cm^2 compared with MDT's 150 Hz/cm^2, CSC covers $2.0 < |\eta| < 2.7$, and works in complementary with MDT to provide high spatial, time and double track resolution. CSC consists of a large and small chamber in ϕ is a multi-wire proportional chamber built with wires in the radial direction segmenting cathodes strips readouts both in parallel and perpendicular to the wire, allowing measurement of both coordinates of r and ϕ. CSC has a good timing resolution of 7 ns, a coordinate resolution of 60 μm in r, and 5 mm in ϕ.

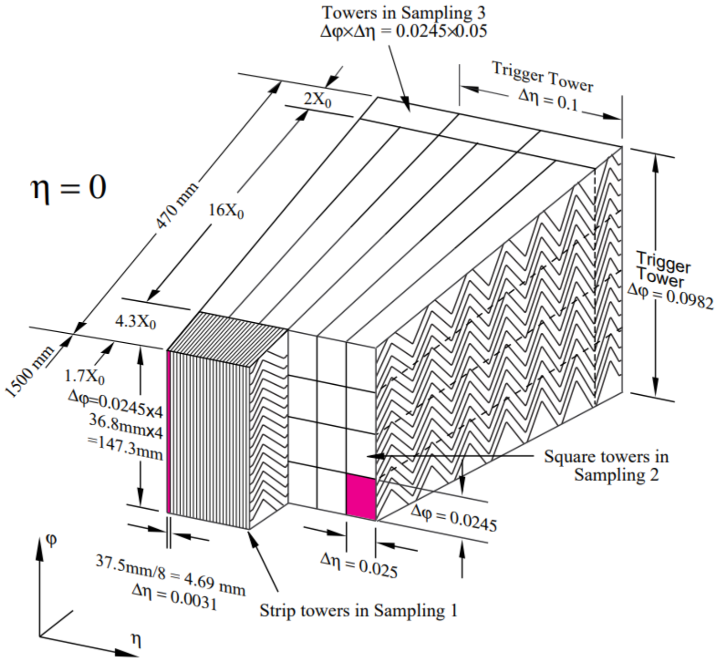


Fig. 4.10 Sketch of LAr EM barrel module with accordion shape. The granularity in η and ϕ is also shown [75].

- ***RPC*** is a gaseous detector consisting of resistive Bakelite plates in parallel kept at a distance of 2 mm, filled with a gas mixture of $C_2H_2F_4$, C_4H_10 and SF_6 on the barrel region $|\eta| < 1.05$. Three trigger stations are layered around the beam axis, where the inner two stations select low p_T tracks with 6 GeV $< p_T <$ 9 GeV and the outer station selects high p_T tracks with 9 GeV $< p_T <$ 35 GeV.

- ***TGC*** is a multi-wire proportional chamber with a narrow 2.8 mm gap between wires and cathode plates that covers the triggering in the endcap region $1.05 < |\eta| < 2.4$. The three layers of TGC, one in front of the second MDT layer (TGC1) and two behind (TGC2 and TGC3) work together on the fast coincidence of muon tracks to provide triggering information.

The overall MS system provides a resolution of $\sigma_{p_T}/p_T = 10\%$ at 1 TeV.

4.2.7 Trigger and Data Acquisition

During the Run-2 operation of LHC, proton bunches collide at a frequency of 40 MHz, while most events are soft inelastic pp collisions that are not of physics interest. Recording such a vast amount of data flow is inefficient and impractical. ATLAS's trigger and data acquisition (TDAQ) system are crucial components. It is designed to filter the vast amount

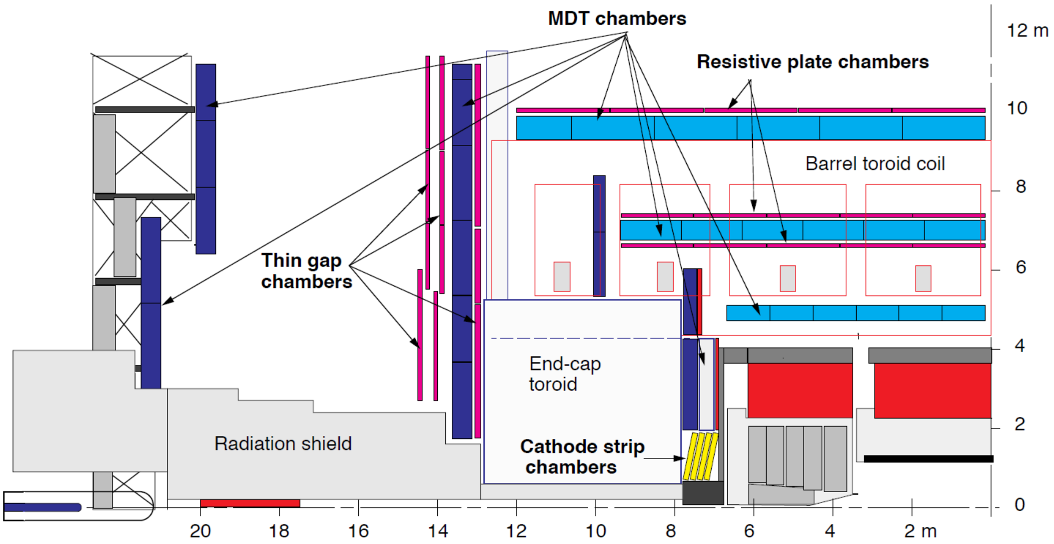


Fig. 4.11 Longitudinal projection of the Muon Spectrometer in ATLAS, which consists of Monitored drift tubes (MDT) and Cathode strip chambers (CSC) for precision tracking, and Resistive plate chambers (RPC) and Thin gap chambers (TGC) for triggering [77].

of data produced in *pp* collisions and select only the potentially interesting physics events. The system consists of a Level-1 (L1) Trigger, which processes based on fast sub-detector information, and a High-Level Trigger (HLT), which employs software to perform a finer selection. The data flow is reduced from 40 MHz to 100 kHz with the L1 trigger, and further down to 1000 Hz with the HLT trigger.

L1 Trigger

The L1 triggers [78] to perform initial selections through Field Programming Gate Arrays (FPGAs) using limited data from the detector to search for physics objects. It consists of two subsystems: the calorimeter system (L1Calo) and the muon system (L1Muon).

- ***L1Calo*** identifies electrons, photons, jets and τ leptons with high E_{T}, as well as missing E_{T} with a reduced granularity of 0.1×0.1 in $\Delta\eta \times \Delta\phi$ in the calorimeter, where muon activities found in the muon trigger system are not taken into account in the calculation of L1 missing E_{T}. The signal goes through the path in a sequence of pre-processors that digitizes the analog input and assigns bunch crossing: jet/energy processor to identify jets and cluster processor (CP) to identify electrons/photons and τ leptons above E_{T} threshold in parallel; L1 central trigger processor (CTP) to make the final decision.

- ***L1Muon*** identifies high E_T muons independently within each dedicated sub-detector segment: RPC for barrel region and TGC for end-cap region triggering. The trigger decision is based on coincidence across three trigger stations that track the path of a muon through the detector.

The L1 trigger also defines a region of interest (RoI). This RoI includes the geometric location and sub-detector information for the physically interesting objects that activate the trigger. This information is then passed to the High-Level Trigger (HLT) system for further analysis.

HLT Trigger

HLT trigger applies more complicated event selections in the processing computer farm to further reduce trigger rates. Seeded with detector information from L1 Trigger, HLT uses full granularity and precision of detector information within the RoI to perform, for example, a scan of the calorimeter for missing E_T building. Data flow with a reduced trigger rate of 1000 Hz after HLT selections is stored for physics analysis.

In the ATLAS experiment, a trigger menu [79] stores the chain that fires the trigger to record an event, specifying thresholds and selection criteria combining information from both L1 and HLT. For example, suppose two electrons fire the event with E_T larger than 17 GeV passing multivariate-based electron ID at "very loose" working point without including impact parameter d_0 information in the HLT algorithm, seeded by two isolated L1 EM clusters with E_T larger than 15 GeV. In that case, the trigger menu is named "HLT_2e17_lhvloose_nod0_L12EM15VHI". A prescaling is sometimes applied during the actual data taking, where a factor intentionally reduces the firing rate to manage the system frequency. In this thesis, we use un-prescaled triggers with objects of interest at the lowest E_T threshold to maximize the signal acceptance for the analysis as will be detailed in Sec. 5.5.

4.2.8 Event Simulations

Simulation of Monte Carlo (MC) events plays a crucial role in various aspects of physics analysis at the ATLAS experiment, from feasibility studies of the expected number of signals and backgrounds to evaluating systematical uncertainties. MC samples used in this thesis are explained in Section 5.1. Here we explain the MC simulation production steps in ATLAS:

Matrix element calculation

The simulation of proton-proton collision events starts from the Parton Distribution Functions (PDFs), which are the momentum distribution functions of partons within a proton. The asymptotic freedom [80] allows perturbative computation via matrix elements, and the outcome is a list of partons with high momentum. For example, in this thesis, the MadGraph event generator [81] has been used, which matches the matrix element to Next-to-Leading-Order (NLO) accuracy for the signal FRVZ and HAHM samples hard scattering simulation. POWHEG (Positive Weight Hardest Emission Generator) [82] is used for the generation of $t\bar{t}$ and single top processes.

Parton showering/hadronization

In this step, QCD processes at a lower energy scale, including the perturbative parton showering and non-perturbative hadronization processes, are simulated, leading to the generation of final state particles stable enough to enter the ATLAS detector, including ISR, FSR, etc. Interfaced with MadGraph, the PYTHIA8 generator [83] is used to simulate final decays in the signal samples. HERWIG (Hadron Emission Reactions With Interfering Gluons) (++) [84] is used for the EW-produced processes of W and Z bosons with additional jets. In addition, standalone generators like Sherpa (High-Energy Reactions of Particles) [85] are used for other SM processes in this analysis, providing both matrix element calculation and parton showering/hadronization.

Detector simulation

So far, the simulation has generated a list of truth particles. Factors such as interaction with detector material, sub-detector acceptance, and resolution must be considered to compare with actual experimental measurements. The detector simulation uses GEANT4 (Geometry And Tracking) [86], which integrates well with the ATLAS detector. GEANT4 simulates the process from making detector hits, considering any misalignment and distortions of the detector setup and background noise, such as beam gas and cavern background, to the HLT and reconstruction in the same treatment with experimental data.

Event weight

Each simulated event is assigned different sources of event weight:

- ***MC weight*** w_{mc} is calcurated to represent the probability of the simulated process in the event based on theoretical prediction and PDF. It can also include factors from computational order in the perturbation theory. For example, events generated beyond LO are assigned with weight to correct the cross-section distribution.

- ***Pile-up weight*** w_{pu} is the normalization associated with the reweighting that matches the simulated average number of interactions per bunch crossing with an observed number at each actual data-taking period.

The expected number of observed events from the simulation is given by

$$N = \sigma \sum_{j}^{2015-2016,2017,2018} L_j \sum_{i}^{n_j} \frac{w_{mc}^{i,j} \times w_{pu}^{i,j}}{w_{tot}}, \tag{4.7}$$

where L_j corresponds to the luminosity, n_j corresponds to the number of observed events after event selections in each data-taking year, and w_{tot} is the total simulated sum of $w_{mc}^{i,j} \times w_{pu}^{i,j}$.

4.3 Reconstruction of Physics Objects

The customization of a dark photon object with LJ signatures in ATLAS is based on standard physics objects of tracks, electrons, and muons. To suppress SM background processes, such as $t\bar{t}$, jets are also used. This section will discuss the standard reconstructing criteria for physics objects in the ATLAS experiment, focusing on these components.

4.3.1 Tracks and Primary Vertex

In the presence of a magnetic field from the solenoid in the ID, the trajectory of a track can be described by a perigee representation with five parameters with respect to a reference point, which is the average position of pp collisions (beam spot) as shown in Fig. 4.12. d_0 and z_0 are the impact parameters corresponding to the transverse and longitudinal distance to the beam spot. ϕ is the azimuthal angle and θ is the polar angle. The last parameter is the charge-to-momentum ratio q/p.

The track reconstruction algorithm [88] can be summarized as below:

- ***Track seeding and finding***: raw detector hits in the pixel and SCT detectors are first clustered and then grouped into sets of three to form various "space points" used for track seeds. A fourth space point from a different layer, with compatible curvature, confirms the track seed. A combinational Kalman filter is then applied to the extended path of each seed within the pixel and SCT detectors. This filter searches for adjacent clusters while smoothing the trajectory. Additionally, calorimeter information within the track seed's Regions of Interest (ROIs) is checked, and feedback is provided to track findings. This feedback helps recover efficiency lost due to the Bremsstrahlung process.

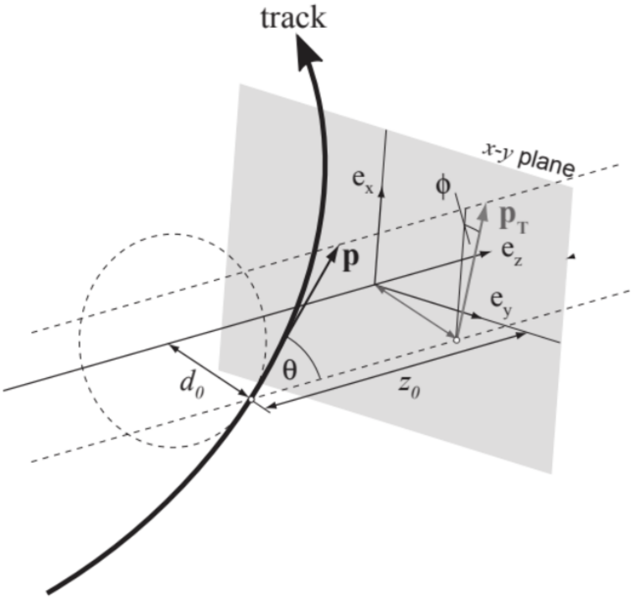


Fig. 4.12 Schematic view of perigee representation of track parameters [87].

- ***Ambiguity solving***: overlap between track hits is solved by assigning scores based on factors like quality for fit χ^2, number of hits missing from an ID layer (holes), transverse momentum p_{T}.
- ***TRT extension***: track finding algorithms are extended to TRT, and additional TRT hits are added to track if the refit has good quality. The additional measurements improve the momentum resolution and particle identification.

Basic working points are defined in Table. 4.2. In the analysis detailed in Chapter 5, tracks used to build LJ are required to satisfy the `Loose` working point.

As mentioned in Section 3.1, in each collision of pp bunches at LHC, we are mostly interested in the hard-scattering event, which is associated with one hard-scattering primary vertex (PV). To find PV, an adaptive multi-vertex fitter is used to find the vertex seeds iteratively, where the same track candidates can be used in multiple vertex candidates with different weights. PV is determined by the vertex with the largest sum of track p_{T} square [90].

4.3.2 Electrons

Electrons play an important role in building electronic lepton jets (eLJs), detailed in Section 5.4. This section will explain standard electron reconstruction, identification, and isolation.

Table 4.2 The ATLAS standard `Loose` and `Tight Primary` track working points for physics analysis [89].

`Loose Working Point`	`Tight Primary Working Point`
$p_\mathrm{T} > 0.5\,\mathrm{GeV}$	All the Loose requirements
$\lvert\eta\rvert < 2.5$	≥ 9 hits in Pixel/SCT for $\lvert\eta\rvert \leq 1.65$
≥ 7 clusters in Pixel/SCT	≥ 11 hits in Pixel/SCT for $\lvert\eta\rvert > 1.65$
≤ 2 holes in the Pixel/SCT	1 hit in IBL or in the innermost Pixel layer
≤ 1 hole in Pixel	No Pixel holes
$\lvert d_0\rvert < 2\,\mathrm{mm}$	-
$\lvert z_0 \sin\theta\rvert < 3\,\mathrm{mm}$	-

Reconstruction

The electron object is reconstructed as a cluster in EM calorimeter with associated charged track(s) from ID. It is performed in the following steps [91]:

- ***Topo-cluster reconstruction***: at the first step, topologically connected EM and hadronic clusters are connected using a '4-2-0' threshold of significance $S^{\mathrm{EM}}_{\mathrm{cell}} = \frac{E^{\mathrm{EM}}_{\mathrm{cell}}}{\sigma^{\mathrm{EM}}_{\mathrm{noise,\ cell}}}$. The clustering initializes from cells of the EM calorimeter with $S^{\mathrm{EM}}_{\mathrm{cell}} \geq 4$ and builds a proto-cluster. The presampler and first layer are excluded to avoid forming noise clusters. Each proto-cluster gets iterated to include neighboring cells with $S^{\mathrm{EM}}_{\mathrm{cell}} \geq 2$ and their nearest neighbors without significance requirement. Local maxima in proto-clusters are split into separate clusters. The EM fraction of total energy deposits $f_{\mathrm{EM}} > 0.5$ and EM energy $E_{\mathrm{EM}} > 400$ MeV is required.

- ***Track refitting***: reconstructed tracks, as described in the previous section, are considered for electron reconstruction. They are refitted using the Gaussian Sum Filter (GSF) algorithm [92] that describes the radiative energy losses of electrons with non-linear functions, allowing better estimations of the track parameters in the presence of bremsstrahlung.

- ***Track-cluster association***: tracks refitted by GSF are extrapolated to the middle layer of the EM calorimeter and matched to topo-clusters. The matching criteria include an angular separation with $\lvert\Delta\eta\rvert < 0.05$ and $-0.10 < q \times (\phi_{\mathrm{track}} - \phi_{\mathrm{topo\text{-}clus}}) < 0.05$, where q is the charge of the particle. The requirement on ϕ is asymmetric to account for potential energy loss due to photon radiation, which leads to a smaller value of measured track ϕ. When multiple tracks are matched to the same cluster, they are prioritized as follows: first, tracks with pixel hits; second, tracks with SCT hits; and third, by the smallest ΔR between the track and the middle layer of the EM calorimeter. The properties of electrons are then derived from the highest-ranked track.

- ***Supercluster reconstruction***: the supercluster is performed independently for electrons and photons and consists of two parts. Electron seed cluster is required to have $E_\mathrm{T} > 1$ GeV and matched ID track with at least four silicon detector hits, while photon seed cluster is required to have $E_\mathrm{T} > 1.5$ GeV with no requirement on tracks. Satellite cluster is then searched within a window of $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ around the seed cluster. For electrons, an additional wider window $\Delta\eta \times \Delta\phi = 0.125 \times 0.300$ also qualifies the satellite cluster if the best-matched track of the satellite cluster is the same as the seed cluster. Satellite clusters are added to the seed clusters to form the final superclusters.

- ***Object building***: in the last step, energy and position calibrations are applied, along with matching of tracks and conversion vertex (massless from two opposite-signed charged tracks) to superclusters, the final object building for physics analysis is as followed: for supercluster with associated tracks and no conversion vertex, it is built as an *electron*. For a supercluster with conversion vertex and no associated tracks, it is built as a *photon*. In other cases where a supercluster produces both an electron and a photon, the object is known as *ambiguous*.

Identification

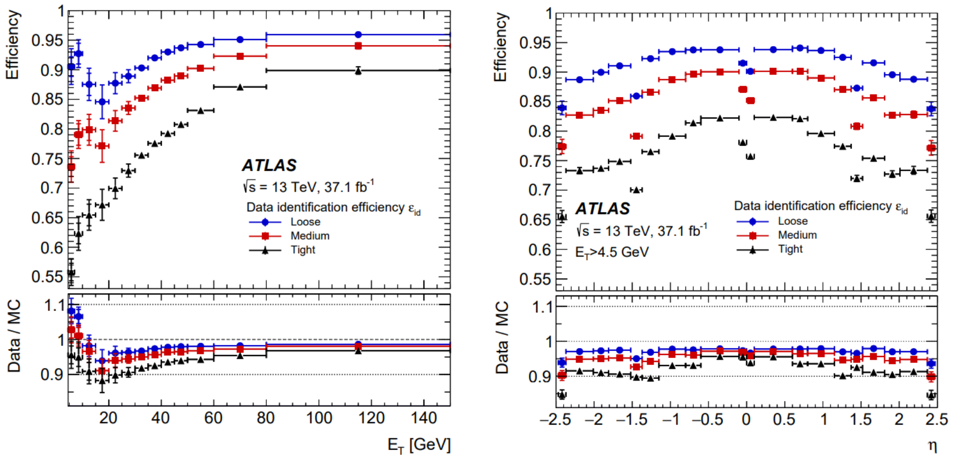
To further improve the purity of electron selections and suppress fake electrons from hadronic jets, photon conversions, and heavy flavor quarks, a likelihood (LH) based electron identification is introduced [91]. The likelihood function is built as

$$L_{S(B)}(\mathbf{x}) = \prod_{i=1}^{n} P_{S(B),i}(x_i), \tag{4.8}$$

where $\mathbf{x} = x_i$ is the vector of inputs. n is the number of variables, including showering shapes in the EM calorimeter, ID track qualities, and track-cluster matching information. P_i corresponds to a pdf in quantity i. The discriminant d_L is given by

$$d_L = -\tau^{-1} \ln\left(\frac{L_S + L_B}{L_S} - 1\right), \tag{4.9}$$

where τ is fixed to 15 [93]. In this way, the efficiency is smoothed, and three standard working points `loose`, `medium` and `tight` can be defined, which correspond to 93%, 88%, and 80%, respectively (as shown in Fig. 4.13). All the working points require tracks to have



(a) (b)

Fig. 4.13 Electron identification efficiency as a function of (a) transverse energy E_{T}, and (b) η in $Z \to ee$ events for different working points [91].

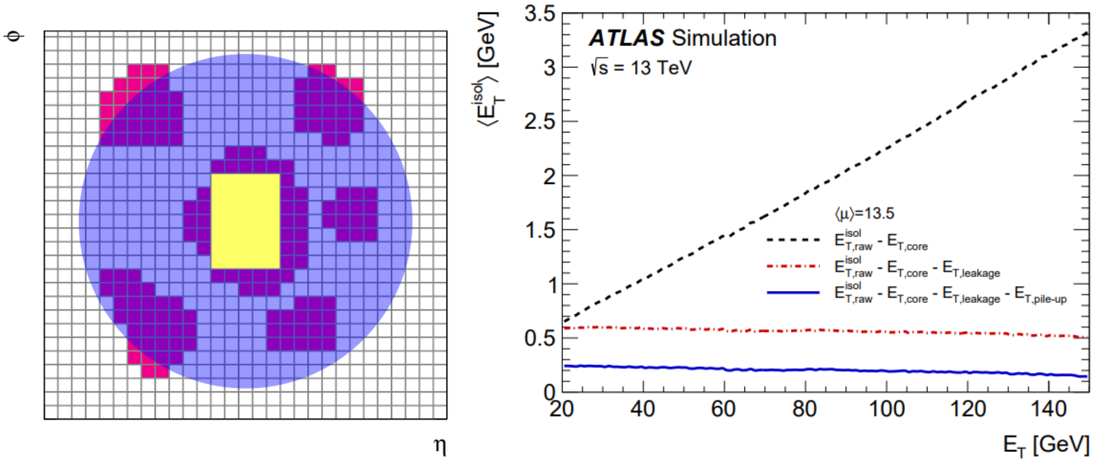
≥ 2 pixel hits and ≥ 7 pixel/silicon hits, where `medium` and `tight` require pixel hits in the innermost layer. In this thesis, we will apply the `medium` identification working point for electrons used to build LJs.

Isolation

Prompt production of electrons from the decay of W, Z, and Higgs bosons typically exhibit minimal detector activity in their vicinity. To quantify this, the term *isolation* is introduced. The isolation is defined as the sum of the transverse energy (E_{T}) or transverse momentum (p_{T}) within a narrow cone $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ around the electron. Generally, electrons that are promptly produced have low isolation values. In the ATLAS experiment, two types of isolation are considered:

- ***Calorimeter isolation***: the raw isolation energy $E_{\mathrm{T,raw}}^{\mathrm{iso}}$ is computed as the energy deposit in a fixed-size isolation cone (e.g., $\Delta R = 0.2$). The energy from prompt electron $E_{\mathrm{T,core}}^{\mathrm{iso}}$ is subtracted in a central $\Delta\eta \times \Delta\phi = 0.125 \times 0.175$ as shown in Fig. 4.14 (a). The leakage of prompt electron energy deposits outside the subtraction core $E_{\mathrm{T,leak}}^{\mathrm{iso}}$ is also evaluated by fitting to Crystal Ball function as a function of E_{T}. Lastly, the pile contribution to isolation is estimated by $E_{\mathrm{T,pile\text{-}up}}^{\mathrm{iso}}(\eta) = \rho_{\mathrm{median}}(\eta) \times (\pi\Delta R^2 - A_{\mathrm{core}})$, where $\rho_{\mathrm{median}}(\eta)$ is the median jet energy density in a unit area. The overall calorimeter isolation is given by

$$E_{\mathrm{T}}^{iso} = E_{\mathrm{T,\,raw}}^{iso} - E_{\mathrm{T,core}}^{iso} - E_{\mathrm{T,leak}}^{iso} - E_{\mathrm{T,pile\text{-}up}}^{iso}. \tag{4.10}$$



(a) (b)

Fig. 4.14 (a) Criteria of central cone subtraction ($\Delta\eta \times \Delta\phi = 0.125 \times 0.175$ in yellow) from isolation cone (in purple). (b) Calorimeter isolation E_{T}^{iso} in $\Delta R = 0.2$ cone as function of E_{T} for each component [91].

Contribution from each component of E_{T}^{iso} is shown in Fig. 4.14 (b).

- ***Track isolation***: track isolation $p_{\mathrm{T}}^{\mathrm{iso}}$ is computed as the sum of tracks p_{T} in isolation cone after removing tracks coming from the electron candidate. The finer granularity in ID allows a variable definition of isolation cone as

$$\Delta R = \min(\frac{10\ \mathrm{GeV}}{p_{\mathrm{T}}\ [\mathrm{GeV}]}, R_{\max}). \tag{4.11}$$

  Here, $R_{\max}$ is a fixed maximum isolation cone (e.g., $\Delta R = 0.2$), and tracks in the central $\Delta\eta \times \Delta\phi = 0.05 \times 0.1$ are considered as a part of the electron object, thus not included in the computation. The variable definition of track isolation in Eq. 4.11, which depends on electron p_{T} allows narrower cone sizes in the boosted decay signatures.

In ATLAS, standard working points for electrons include both calorimeter isolation and track isolation selections. They are summarized in Table. 4.3, where "cone20" and "varcone30" correspond to fixed cone size of $\Delta R = 0.2$ for calorimeter isolation, and $R_{\max} = 0.3$ for the variable cone size track isolation. In the analysis of this thesis, we select the `Loose_VarRad` isolation working point.

Table 4.3 Summary of `Loose_VarRad` and `Tight_VarRad` working point requirements for electrons. p_{T}^{e} is the p_{T} of electrons. "cone20" and "varcone30" corresponds to fixed cone size of $\Delta R = 0.2$ for calorimeter isolation, and $R_{\max} = 0.3$ for the variable cone size track isolation.

Isolation WP	Calorimeter	Track
`Loose_VarRad`	$\frac{E_{\mathrm{T}}^{\mathrm{iso,cone20}}}{p_{\mathrm{T}}^{e}} < 0.2$	$\frac{p_{\mathrm{T}}^{\mathrm{iso,varcone30}}}{p_{\mathrm{T}}^{e}} < 0.15$
`Tight_VarRad`	$\frac{E_{\mathrm{T}}^{\mathrm{iso,cone20}}}{p_{\mathrm{T}}^{e}} < 0.06$	$\frac{p_{\mathrm{T}}^{\mathrm{iso,varcone30}}}{p_{\mathrm{T}}^{e}} < 0.06$

4.3.3 Muons

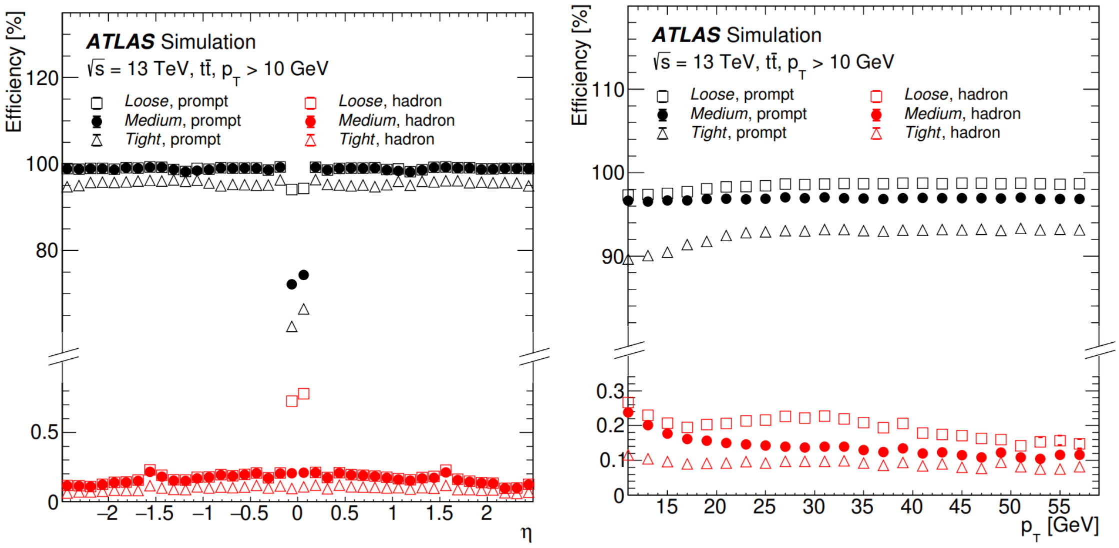
Muons are used to build muonic lepton jets (μLJs), which are not the focus of this thesis. To ensure orthogonality between channels when combining results, muons are discussed here for the purposes of muon vetoing and overlap removal in the reconstruction of eLJs, as will be detailed in Sec. 5.4.

Reconstruction

In the ATLAS experiment, the reconstruction of muons involves identifying muons from tracks in the ID and MS detectors, with additional information from the calorimeter. There are four types of reconstructed muons [94]:

- ***Combined (CB) muons***: reconstructed by performing a combined fit to tracks from both MS and ID, taking into account energy loss in the calorimeter. CB muons are reconstructed up in $|\eta| < 2.5$.
- ***Segment-tagged (ST) muons***: reconstructed based on tracks from ID that match at least one MS segment during extrapolation. Muon parameters are assigned directly from ID. ST muons are reconstructed up in $|\eta| < 2.7$.
- ***Calorimeter-tagged (CT) muons***: reconstructed based on tracks from ID that match energy deposits in the calorimeter matching a minimum ionizing particle. CT muons are optimized to recover acceptance in the gap region $|\eta| \leq 0.1$.
- ***MS Extrapolated (ME) muons***: reconstructed based on tracks from MS only, in case tracks from MS cannot be matched to ID. ME muons are reconstructed up in $|\eta| < 2.7$.

In the analysis in this thesis, CB muons are used.



(a) (b)

Fig. 4.15 Muon identification efficiency as a function of (a) transverse momentum p_{T}, and (b) η in $t\bar{t}$ events for different working points [95].

Identification

Like electrons, a muon identification algorithm is applied to select muons with high purity against hadrons. The q/p significance is a powerful variable that checks compatibility between ID and MS.

$$q/p\text{significance} = \frac{|q/p_{\mathrm{ID}} - q/p_{\mathrm{MS}}|}{\sqrt{\sigma^2(q/p_{\mathrm{ID}}) + \sigma^2(q/p_{\mathrm{MS}})}}, \tag{4.12}$$

where q and p are the charge and momentum of muons measured in the corresponding sub-detector, and σ is their uncertainties. Three selection working points are defined: `loose`, `medium` and `tight`. The `loose` WP accepts CB and ME muons in $|\eta| < 2.5$ with at least two precision stations and in the gap region $|\eta| < 0.1$ with one precision station. In addition, ST and CT muons in $|\eta| < 0.1$ are accepted. The `medium` WP selects CB and ME muons, including gap with one precision station, while the *tight* WP only accepts CB and ME muons with at least two precision stations. The efficiency of different WPs is shown in Fig. 4.15. In this thesis, `loose` identification WP is selected.

Isolation

Like the approach described for prompt electrons in Section 3.2, muon isolation involves measuring the surrounding activity to distinguish prompt muons from non-prompt ones

Table 4.4 Summary of different working point requirements for muons [96].

Isolation WP	Calorimeter and Track	Track p_{T} requirement
`PFlow_Loose`	$(p_{\mathrm{T}}^{\mathrm{varcone30}} + 0.4E_{\mathrm{T}}^{\mathrm{neflow20}}) < 0.16p_{\mathrm{T}}^{\mu}$	$p_{\mathrm{T}} > 500$ MeV
`PFlow_Tight`	$(p_{\mathrm{T}}^{\mathrm{varcone30}} + 0.4E_{\mathrm{T}}^{\mathrm{neflow20}}) < 0.045p_{\mathrm{T}}^{\mu}$	$p_{\mathrm{T}} > 500$ MeV
`Loose_VarRad`	$p_{\mathrm{T}}^{\mathrm{varcone30}} < 0.15p_{\mathrm{T}}^{\mu}$, $E_{\mathrm{T}}^{\mathrm{topocone20}}) < 0.3p_{\mathrm{T}}^{\mu}$	$p_{\mathrm{T}} > 1$ GeV
`Tight_VarRad`	$p_{\mathrm{T}}^{\mathrm{varcone30}} < 0.04p_{\mathrm{T}}^{\mu}$, $E_{\mathrm{T}}^{\mathrm{topocone20}}) < 0.15p_{\mathrm{T}}^{\mu}$	$p_{\mathrm{T}} > 1$ GeV

originating from heavy-flavor hadronic decays. In addition to the track isolation $p_{\mathrm{T}}^{\mathrm{varcone30}}$ and calorimeter isolation $E_{\mathrm{T}}^{\mathrm{topoetcone20}}$ used in electrons, we introduce an additional calorimeter-based isolation variable, $E_{\mathrm{T}}^{\mathrm{neflowiso20}}$. This variable is defined by excluding all energy deposits in the calorimeter that match ID tracks in the $\Delta R = 0.2$ muon isolation cone, thereby reducing the correlation with track-based isolation measures. Table 4.4 summarizes the different working points for muon isolation. In this thesis, the `PFlow_Loose` isolation working point is used.

4.3.4 Jets

In this analysis, jets are used to suppress SM background processes like $t\bar{t}$ and overlap removal as will be described in Sec. 4.3.5. We will discuss the reconstruction of jets, a jet vertex tagger to suppress pile-ups, and jet energy scale and resolution.

Reconstruction

Jets reconstruction [97] is performed on topo-clusters (see Sec. 4.3.2) using an anti-k_{T} clustering algorithm [98]. By defining the following metrics:

$$d_{i,j} = \min(k_{\mathrm{T},i}^{2p}, k_{\mathrm{T},j}^{2p}) \frac{\Delta^2 R_{i,j}}{R^2}, \tag{4.13}$$

$$d_{i,B} = k_{\mathrm{T},i}^{2p}. \tag{4.14}$$

Here k_{T} represents the transverse momentum, $d_{i,B}$ is the distance between topo-cluster i and beam, and $d_{i,j}$ is the distance between topo-cluster i and j. $\Delta^2 R_{i,j}$ corresponds to solid angle between topo-cluster i and j and R is fixed clustering size. The parameter $p = -1$ refers to anti-k_{T} algorithm, $p = 1$ refers to the sequential recombination k_{T} [99] and $p = 0$ refers to the Cambridge-Achen [100] algorithm. The algorithm iteratively tries to find the closest distance $d_{i,j}$ and combines the two clusters if $d_{i,j} < d_{i,B}$, namely the cluster-cluster distance is smaller than their cluster-beam distance. Otherwise, a jet is formed, and clusters in the jet are removed from the pool. We use a clustering size of $R = 0.4$ to reconstruct jets in ATLAS.

Jet Vertex Tagger

A jet vertex tagger (JVT) based on k-nearest neighbor (KNN) is applied to suppress jets reconstructed from pile-up. JVT is constructed by 2-D likelihood using the following two variables:

- the corrected jet vertex fraction (corrJVF):

  $$\mathrm{corrJVF} = \frac{\sum_l p_{\mathrm{T},l}^{\text{trk-jet}}(\mathrm{PV}_0)}{\sum_l p_{\mathrm{T},l}^{\text{trk-jet}}(\mathrm{PV}_0) + \frac{\sum_n \sum_l p_{\mathrm{T},l}^{\text{trk-jet}}(\mathrm{PV_n})}{(kn^{\mathrm{pu-trk}})}}, \tag{4.15}$$

  where $\sum_k p_{\mathrm{T},k}^{\text{trk-jet}}(\mathrm{PV}_0)$ is the sum of p_T for tracks associated with jet originated from hard-scattering PV_0, $\mathrm{PV_n}$ is the pile-up interaction and k scales the contribution of pile-up.

- R_{p_T} defined as the ratio of scalar sum of track p_T associated with jet from hard-scattering PV_0 to jet p_T:

  $$R_{p_\mathrm{T}} = \frac{\sum_l p_{\mathrm{T},l}^{\text{trk-jet}}(\mathrm{PV}_0)}{p_\mathrm{T}^{jet}}. \tag{4.16}$$

The contribution from the jets originating from the pile-up interaction is not negligible for low p_T jets. In this analysis, we require a JVT score > 0.11 for 20 GeV $< p_\mathrm{T} <$ 60 GeV to reduce the pile-up jets.

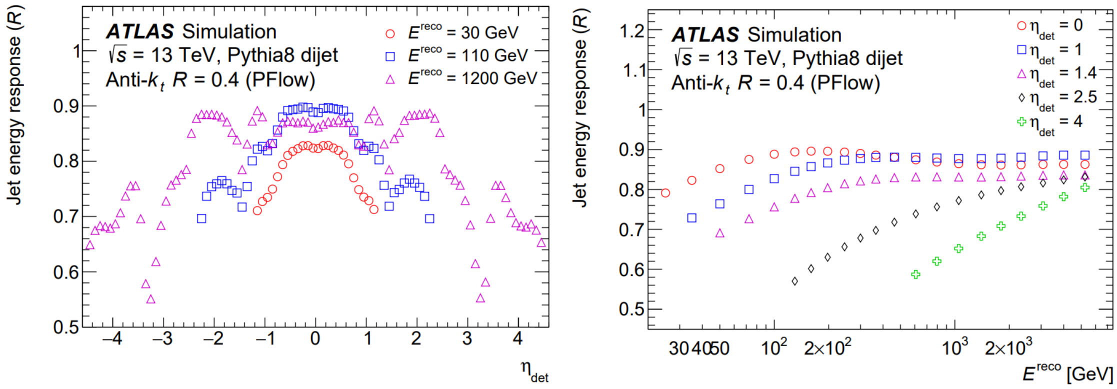
Jet Energy Scale and Resolution

To account for effects like detector response and energy loss in passive material, corrections are applied to the jet energy scale (JES) in the following steps [101]:

- ***pile-up contributions***: jet p_T due to pile-up contributions is corrected by

  $$p_\mathrm{T}^{\mathrm{corr}} = p_\mathrm{T} - \rho A_{\mathrm{jet}} - \alpha(N_{\mathrm{PV}} - 1) - \beta\mu, \tag{4.17}$$

  where the size of jet A_{jet} is determined by the relative number of ghost particles associated with a jet after clustering and ρ is the median p_T density in the y-ϕ plane. α



(a) (b)

Fig. 4.16 Average truth energy response $(E_{\text{reco}}/E_{\text{true}})$ as function of (a) reconstructed η and (b) reconstructed transverse energy E_{T} [102].

and β are the residual correction factors derived in bins of true p_{T} and η for N_{PV}, the number of reconstructed vertices, and μ is the average number of interactions per BC.

- ***absolute MC based calibration***: anti-k_{T} jets reconstructed from simulated dijet events are matched to truth jets after the pile-up corrections. The jet energy truth response $(E_{\text{reco}}/E_{\text{true}})$ is then fitted by a Gaussian function in fine bins of η and E_{T} (see Fig. 4.16).

- ***global sequential calibration***: a series of corrections based on calorimeter responses are applied. The corrections are derived from 5 different variables distinguishing quark-initiated and gluon-initiated jets: fraction of jet energy deposits in the first layer of tile calorimeter ($|\eta| < 1.7$), fraction of jet energy deposits in the third layer of LAr calorimeter ($|\eta| < 3.5$), number of ghost-associated tracks with $p_{\text{T}} > 1$ GeV ($|\eta| < 2.5$), average p_{T}-weighted transverse distance between the jet axis and ghost-associated tracks with $p_{\text{T}} > 1$ GeV ($|\eta| < 2.5$), and the number of ghost-associated muon track segments ($|\eta| < 2.7$).

- ***in situ jet calibrations***: jet response is corrected by the ratio of average jet p_{T} between data and MC simulation in the reference processes like Z/γ+jets by balancing the hadronic recoil in an event against the p_{T} of a calibrated Z boson or photon.

The jet energy resolution (JER) is parameterized in the functional form:

$$\frac{\sigma(p_T)}{p_T} = \frac{N}{p_T} \oplus \frac{S}{\sqrt{p_T}} \oplus C. \tag{4.18}$$

The noise (N) contribution corresponds to electronic noise and pile-ups. The stochastic (S) contribution corresponds to the statistical fluctuations in the energy deposits. The constant (C) corresponds to a constant correction from jet energy deposits in passive materials like cryostats and solenoid coils.

Systematic uncertainty arising from JES and JER are evaluated in Chapter 8.

4.3.5 Overlap Removal

Physics objects are reconstructed independently among sub-detectors, which could lead to the reuse of the same information. For instance, energy deposits detected in the calorimeter may be simultaneously reconstructed as electrons, muons, and hadronic jets. To prevent such ambiguity, ATLAS utilizes a technique called overlap removal (OR) to prioritize the assignment of physics objects when they are detected close to each other. Regarding electrons, muons, and hadronic jets, which are used in the search for prompt LJs, the standard OR criteria are applied in the following sequence:

- ***electron-electron OR***: when two electrons share the same track, an electron with lower p_{T} is rejected.
- ***muon-electron OR***: when muon and electron share the same track, the muon is rejected.
- ***electron-jet OR***: when electron and jet are close-by with angular separation $\Delta R < 0.2$, jet is rejected. Otherwise, the electron is rejected if $\Delta R < 0.4$.
- ***muon-jet OR***: when muon and jet are close by, jet with less than three associated tracks and with muon within $\Delta R < 0.2$ or with ghost associated track [103] is rejected. Otherwise, muon is rejected if $\Delta R < 0.4$.

In this thesis, we optimize OR criteria by removing *electron-electron OR* to adapt to the collimated signature of LJs decaying from γ_{d}.

Chapter 5

Prompt Dark Photon Signatures

This chapter explores the unique signatures of prompt dark photons that decay into LJs. These LJs are highly specialized objects derived from standard reconstructed physics objects described in Section 4.3.

Section 5.1 provides an overview of the data and simulated MC events used in this analysis. The kinematic properties of the LJs are presented in Section 5.2. Section 5.3 discusses the customization of physics objects, and Section 5.4 details the reconstruction of prompt LJs for both electron and muon channels. Later in the thesis, we will focus on the electron channel only. Section 5.5 describes the trigger strategy.

5.1 Data and Simulated Events

We have simulated two signal benchmark model MC samples and SM background MC samples to fully exploit the search limits of the complete Run-2 data in ATLAS. They are used to optimize selection cuts, help understand the contributions from different background processes, and interpret the search results.

Full Run-2 data

The full dataset from Run-2 pp collisions at a center-of-mass energy of $\sqrt{s} = 13$ TeV listed in ATLAS centrally distributed GRLs (see Section. 4.1.3) is used. The data corresponds to 140 fb^{-1} integrated luminosity with a 0.8% uncertainty [104], which is summarized in Table 5.1.

Table 5.1 Full dataset in run-2 pp collisions at center-of-mass energy of $\sqrt{s} = 13$ TeV for physics analysis. The data corresponds to 140 fb^{-1} integrated luminosity.

Year	**Data Period**	**Luminosity** (fb^{-1})
2018	period B-Q	58.79
2017	period B-K	44.31
2016	periods A-L	33.40
2015	periods D-J	32.44

Signal MC samples

Benchmark models FRVZ and HAHM (see Sec. 3.2.3) involving both the Higgs portal and vector portal are simulated. Higgs bosons generated through gluon-gluon fusion production mode [105] with cross-section of $\sigma = 48.52$ pb [106] for $m_H = 125.09$ GeV. Higgs boson production through vector boson fusion processes with a cross-section of $\sigma = 3.78$ pb is not included in this analysis, and we suppress it by applying jet-veto. The MG5_aMC@NLO v2.2.3 generator [81], interfaced with the PYTHIA8 generator [83], is used for simulation of the Higgs boson to subsequent dark-sector particles, along with parton showering/hadronization processes to final products. The detector response is simulated through GEANT4 [86] and then reconstructed and processed similarly with data.

Signal MC samples are generated with dark photon γ_d masses ranging from 17 MeV to 40 GeV, aiming to extend the boundaries in previous searches [4]. The decay length is set to $c\tau < 0.01$ mm to optimize for the search for promptly decaying γ_ds. For the FRVZ signal samples, the masses of f_d and HLSP are selected to be small compared with the Higgs boson but well below the threshold where γ_d is kinematically allowed ($m_{\mathrm{HLSP}} + m_{\gamma_\mathrm{d}} = m_{f_\mathrm{d}}$). This allows the production of dark photons that are well-boosted and separated in the transverse plane, creating a unique signature for our target, as will be discussed in Section 5.2. Dark photons decay to fermions according to the branching ratios in Fig. 3.6. Table 5.2 and Table 5.3 summarize the parameters used in generating the FRVZ and HAHM samples, respectively.

SM background MC samples

The main SM background processes of W+jets, Z+jets, $t\bar{t}$, single top, multijets, γ+jets and WW, WZ, ZZ dibosons have been generated to help analyze the composition of backgrounds entering the signal region and to refine the strategy for selection cuts. Simulation of the W+jets, Z+jets, WW, WZ, and ZZ events have been generated using SHERPA 2.2.1 [85] with the NNPDF 3.0 NNLO [107] PDF set. Single top and $t\bar{t}$ MC samples are generated using

Table 5.2 Parameters used for the MC simulation of $H \to 2\gamma_{\mathrm{d}} + X$, according to the FRVZ benchmark model. The cross-section corresponds to BR($H \to 2\gamma_{\mathrm{d}} + X$)= 100%. Each dataset consists of 390000 events.

m_H [GeV]	$m_{\gamma_{\mathrm{d}}}$ [GeV]	m_{HLSP} [GeV]	m_{fd} [GeV]	σ [pb]
125	0.017	2	5	48.52
125	0.03	2	5	48.52
125	0.06	2	5	48.52
125	0.1	2	5	48.52
125	0.24	2	5	48.52
125	0.4	2	5	48.52
125	0.9	2	5	48.52
125	2	2	10	48.52
125	6	4	25	48.52
125	10	6	35	48.52
125	15	10	45	48.52
125	25	10	45	48.52
125	40	7	55	48.52

Table 5.3 Parameters used for the MC simulation of $H \to 2\gamma_{\mathrm{d}}$, according to the HAHM benchmark model. The cross-section corresponds to BR($H \to 2\gamma_{\mathrm{d}}$)= 100%. Each dataset consists of 390000 events.

m_H [GeV]	$m_{\gamma_{\mathrm{d}}}[GeV]$	σ [pb]
125	0.017	48.52
125	0.1	48.52
125	0.4	48.52
125	2	48.52
125	10	48.52
125	15	48.52
125	25	48.52
125	40	48.52

Table 5.4 Summary of generator information for SM background processes. The generator cross-sections are used to generate samples.

Physics process	Generator	PDF set	σ Normalization	Parton shower	Tune
$W(\to \ell\nu)+$ jets	SHERPA 2.2.11	NNPDF3.0NNLO	NNLO	SHERPA	SHERPA
$Z(\to \ell\ell)+$ jets	SHERPA 2.2.11	NNPDF3.0NNLO	NNLO	SHERPA	SHERPA
$t\bar{t}$	POWHEG-BOX V2	CT10/CTEQ6L1	NNLO	PYTHIA8	PERUGIA2012
single-top	POWHEG-BOX V2	CT10/CTEQ6L1	NNLO	PYTHIA8	PERUGIA2012
di-boson	SHERPA 2.2.1	NNPDF3.0NNLO	NNLO	SHERPA	SHERPA
multijet	PYTHIA8	NNPDF2.3LO	LO	PYTHIA8	A14
$\gamma+$ jets	PYTHIA8	NNPDF2.3LO	LO	PYTHIA8	A14

POWHEG-BOX V2 [82] and PYTHIA8 [108] with the PERUGIA2012 [109] tune for parton showering and hadronization, and CT10/CTEQ6L1 [110] [111] PDF sets. Multijet and γ+jets samples are generated with PYTHIA8 in different p_T slices of leading parton and photons, respectively. Like signal MC simulation, a full detector simulation is performed using the GEANT4 [86] toolkit. The generator information is summarized in Table 5.4.

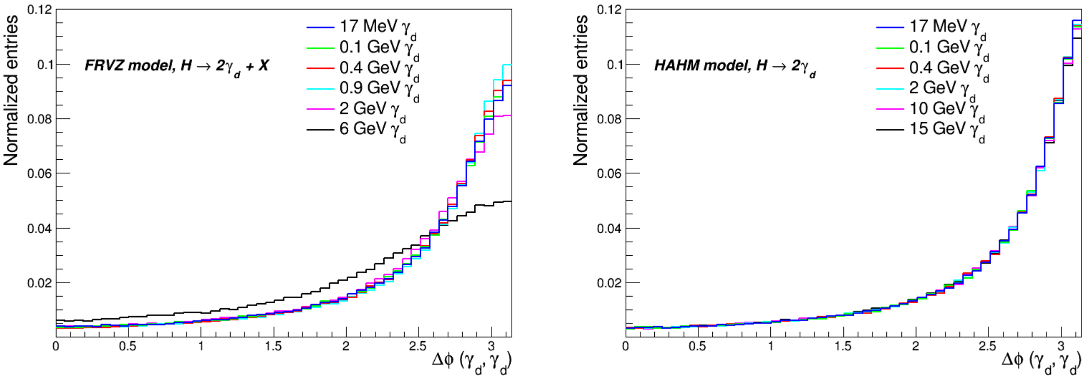
5.2 Dark Photon Kinematics

From the event topologies shown in Fig. 3.8 (b) and Fig. 3.8 (a) for the HAHM and FRVZ models, we expect dark photons to be produced back-to-back and decay to collimated pairs of leptons, namely, the LJ signature. In this section, signal MC samples are used to study the kinematic properties of dark photons and their decay to electrons without any cuts.

Angular separation $\Delta\phi$ between two γ_ds in both FRVZ and HAHM benchmark MC samples is shown in Fig. 5.1. It is observed that two γ_d are well separated in the transverse plane. As γ_d mass gets larger, around 6 GeV for the FRVZ model, $\Delta\phi$ starts to get flattened. Due to the absence of hidden particle final states in the HAHM model, two γ_d are more separated in the transverse plane.

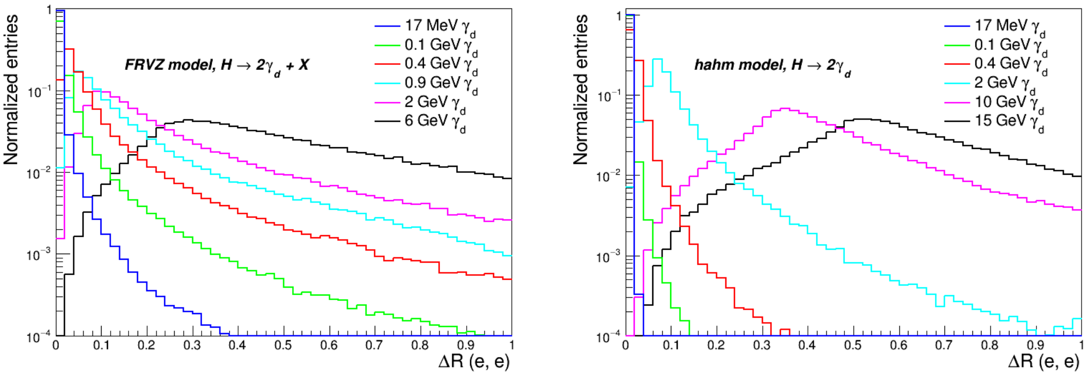
Angular separation ΔR between electron pairs decaying from γ_d is shown in Fig. 5.2. It is observed that electron pairs are highly collimated, especially at low γ_d mass region, resulting in the unique signature of LJs we target. Electrons separate as the γ_d mass increases, and we are less sensitive (more details will be given in Sec. 5.4).

Transverse momentum p_T of γ_d, and the decay product electrons are shown in Fig. 5.3 and Fig. 5.4, respectively. It is observed that both γ_d and electron p_T are similar among different γ_d mass points, while for HAHM γ_d and electrons are more boosted due to the absence of HLSP, resulting in a better acceptance and efficiency.



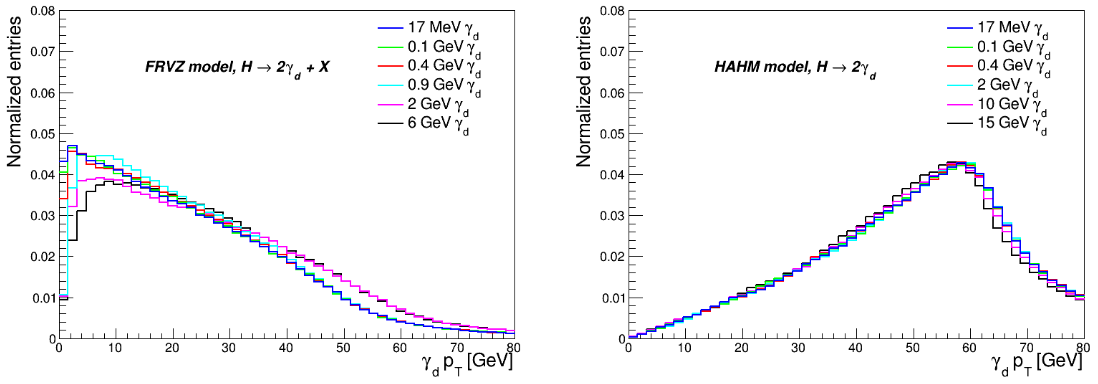
(a) (b)

Fig. 5.1 Normalized distribution of angular separation $\Delta\phi$ between two γ_{d} in the benchmark (a) FRVZ and (b) HAHM signal MC simulations.



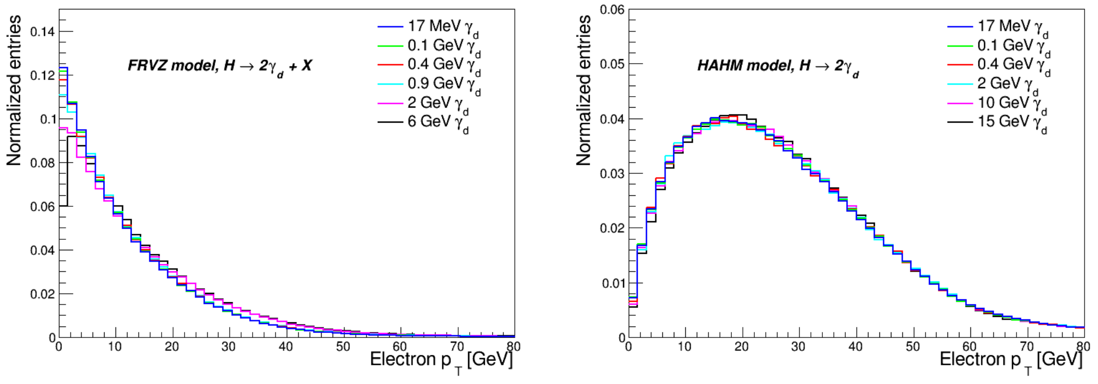
(a) (b)

Fig. 5.2 Normalized distribution of angular separation ΔR between two electrons decaying from γ_{d} in the benchmark (a) FRVZ and (b) HAHM signal MC simulations.



(a) (b)

Fig. 5.3 Normalized distribution of transverse momentum p_{T} of γ_{d} in the benchmark (a) FRVZ and (b) HAHM signal MC simulations.



(a) (b)

Fig. 5.4 Normalized distribution of transverse momentum p_{T} of electrons decaying from γ_{d} in the benchmark (a) FRVZ and (b) HAHM signal MC simulations.

5.3 Physics Objects in Prompt Lepton Jets

The construction of LJs begins with physics objects in the standard ATLAS reconstruction. Table 5.5 summarizes the requirements for electrons and muons to build LJs. The p_T requirements for electrons and muons are set to the lowest thresholds in the phase space where calibrations are valid to maximize signal acceptance. Due to the hadronic fakes, events involving electrons present a less clean signature than those involving muons. Therefore, we use the `Medium` identification WP for electrons and the `CB Loose` identification WP for muons. The collimated signatures of LJs motivate the study of modification to the isolation requirements. As will be discussed in Sec. 5.3.1, we apply the standard `Loose_VarRad` isolation WP for electrons, while for muons, we use a modified `PFlowLoose_VarRad` isolation WP.

Table 5.6 summarizes the requirements for tracks associated with electrons and muons. We require the `Loose` quality (as detailed in Table 4.2), with additional criteria on track quality d_0 and z_0 to further suppress non-prompt lepton fakes.

Table 5.5 Summary of requirements for muons and electrons. Electrons reconstructed in the transition region between the barrel and the endcap of the calorimeter system at $1.37 < |\eta| < 1.52$ are excluded due to poor performance. `PFlowLoose_VarRad` isolation WP is modified, as will be explained in Sec. 5.3.1.

Object	p_T [GeV]	$\|\eta\|$	Identification	Isolation
Electron	> 4.5	$\|\eta\| < 1.37$ or $1.52 < \|\eta\| < 2.5$	`Medium`	`Loose_VarRad`
Muon	> 3	$\|\eta\| < 2.5$	`CB Loose`	modified `PFlowLoose_VarRad`

Table 5.6 Summary of requirements for muon and electron associated tracks.

Associated Track	p_T [GeV]	$\|\eta\|$	Quality	$\|d_0\|/\sigma_{d_0}$	$\|z_0\| \sin\theta$ [mm]
Electron-track	> 0.5	< 2.5	`Loose`	< 5	< 0.5
Muon-track	> 0.5	< 2.5	`Loose`	< 3	< 0.5

5.3.1 Close-by Isolation Corrections

As mentioned in Sec. 4.3, the isolation is defined as the surrounding energy of a muon or electron used to suppress non-prompt particles, typically accompanied by nearby additional particles. While the standard isolation identifies prompt objects well in many other analyses, the nature of LJs to build from highly collimated leptons presents a challenge. This section introduces a correction method for isolation when dealing with prompt LJs.

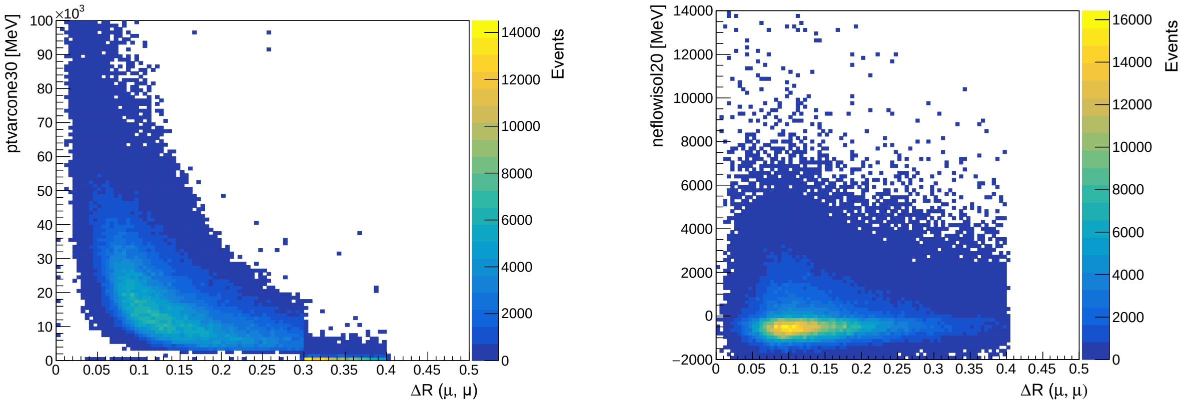


Fig. 5.5 The (a) ptvarcone30 and the (b) neflowisol20 distributions for muons as a function of their truth opening angle ΔR. Muons are required to be matched to truth muons from γ_d, which has a mass of $m_{\gamma_d} = 2$ GeV.

For muons, the particle-flow based (PFlow) isolation WP, `PFlowLoose_VarRad`, is composed of two elements: a charged component, `ptvarcone30` using raw track data, and a neutral component, `neflowisol20`, which relies on energy deposits in the calorimeter. Fig. 5.5 shows the isolation distributions as a function of the truth ΔR between muons. It is noted that while `neflowisol20` has a standard isolation energy scale, `ptvarcone30` has much higher values in the tens of GeV range, which leads to failures in the isolation WP criteria. Such large isolation is primarily due to the other muon from γ_d being counted as `ptvarcone30`. On the other hand, in the case of calorimeter isolation, there are no EM objects for muons. To correct the contribution from the other muon from γ_d, PFlow objects associated with any other muon located near the original muon candidate from the isolation calculation are excluded. Fig. 5.6 illustrates the `ptvarcone30` distributions for both the standard and corrected variables. As expected, the contributions in the standard variable are consistent with the p_T of the nearby muon. Fig. 5.8 (a) shows the WP efficiency before and after applying corrections.

A similar study was also performed for the case of electrons. Fig. 5.7 shows isolation as a function of truth ΔR between electrons decaying from γ_d at 0.9 GeV. It is noted that in regions sensitive for the electron channel ($m_{\gamma_\mathrm{d}} < 1$ GeV), most electrons decaying from γ_d fall within the central subtraction region ($\Delta\eta \times \Delta\phi = 0.125 \times 0.175$ for calorimeter based isolation and $\Delta\eta \times \Delta\phi = 0.05 \times 0.1$ for track based isolation, (see Fig. 4.14) where they will not contribute to the isolation computation. This allows the electron to pass standalone isolation WP selections without correction. In addition, as we will see in Sec. 5.3.2, collimated electrons tend to merge to only one cluster, in which case no additional tracks can be subtracted by correction. Fig. 5.8 (b) shows the isolation WP efficiency before and after

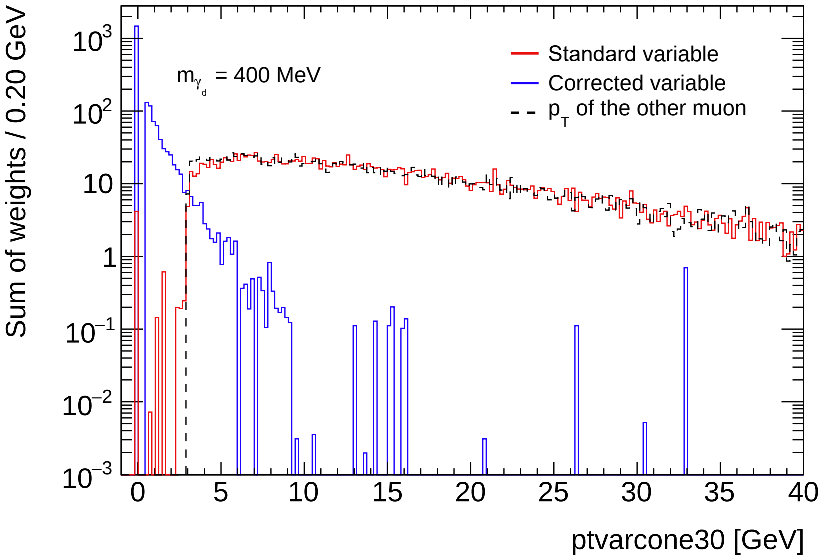


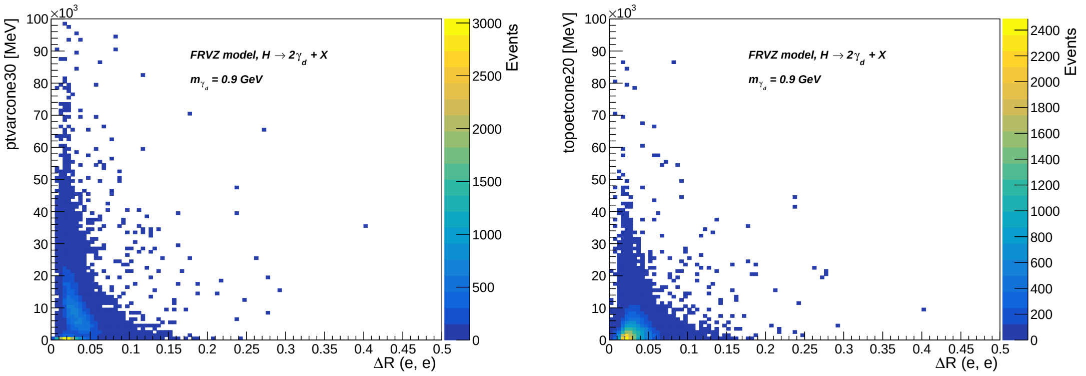
Fig. 5.6 Comparison of the standard and corrected $p_\mathrm{T}^\mathrm{iso,varcone30}$ isolation variable. p_T of the other muon in the dashed line aligns well with a nominal variable in red.

applying corrections. Since we focus on $m_{\gamma_\mathrm{d}} < 240$ MeV where there is no improvement after corrections. As a result, we apply corrected `PFlowLoose_VarRad` WP to muon objects and the standalone `Loose_VarRad` to electron objects to build LJs.

5.3.2 Merged EM clusters

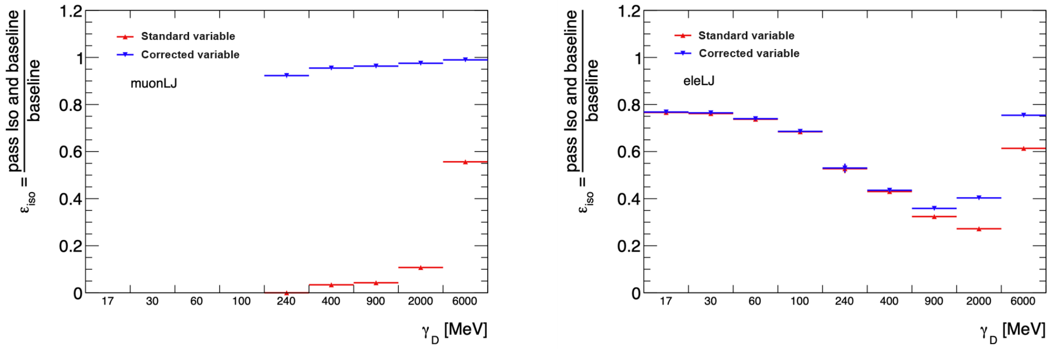
The collimated nature of LJs presents challenges to the standard method of defining objects in the ATLAS detector. This is especially problematic when an electron enters the reconstruction region of another electron from the same decay. This section will discuss the responses of ID and EM calorimeters and the unique signature of objects used to build LJs.

In the EM calorimeter, superclusters are formed from topo-clusters to build electrons (see Sec. 4.3.2). Suppose the separation angle between electron pairs from γ_d decay is large enough (i.e., $\Delta R \gtrsim 0.1$). In that case, they are recognized as distinct EM clusters close to each other (expected for a high-mass γ_d). However, if the electrons are very collimated, their energy deposits merge into one single EM cluster (low-mass γ_d). From Fig. 5.9 (a), which shows the number of reconstructed clusters from electron pairs inside LJ, the cluster merging occurs in $\sim 90\%$ of $\gamma_\mathrm{d} \to ee$ decays when γ_d mass is lower than 6 GeV. Fig. 5.10 (a) shows the number of EM clusters as a function of truth ΔR between electrons for FRVZ signal MC samples for γ_d mass at 0.1 GeV for the merged cluster scenario, while Fig. 5.10 (b) for FRVZ signal MC samples for γ_d mass at 6 GeV is an example of separate close-by clusters scenario.



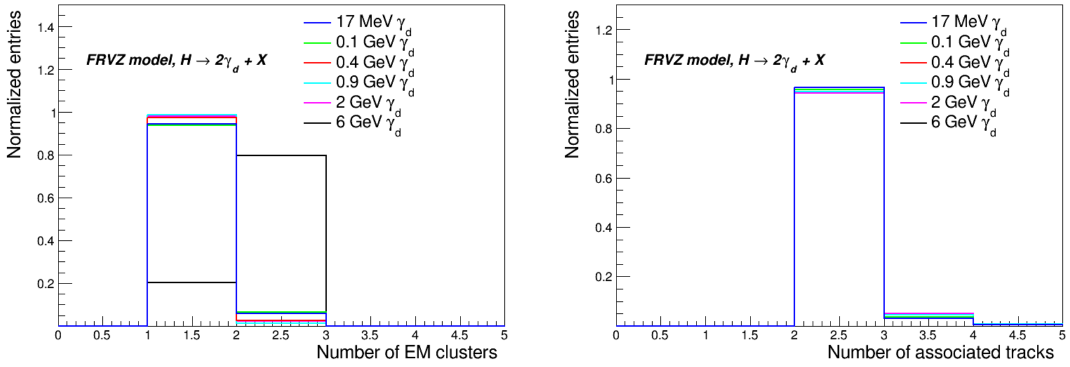
(a) (b)

Fig. 5.7 The (a) ptvarcone30 and the (b) topoetcone20 distributions for electrons as a function of their truth opening angle ΔR. Electrons are required to be matched to truth electrons from γ_d, which has a mass of $m_{\gamma_\mathrm{d}} = 0.9$ GeV.



(a) (b)

Fig. 5.8 Isolation WP efficiency for (a) muon and (b) electron before and after corrections. The corrected isolation WP is used for muons, while the standard isolation WP is used for electrons in this thesis.



(a) (b)

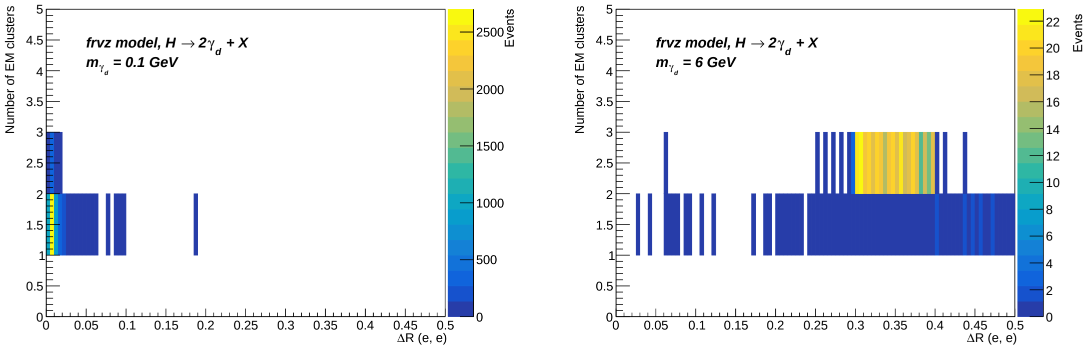
Fig. 5.9 Number of (a) EM clusters and (b) associated tracks inside LJ at different γ_{d} mass in the FRVZ benchmark model.

Tracks, on the other hand, have a better resolution due to the high granularity of ID, and even in the merged cluster scenario, collimated electron tracks from γ_{d} can be reconstructed separately. In addition, GSF-refitted tracks are associated with clusters (see Sec. 4.3.1). As shown in Fig. 5.9 (b), electrons from LJs tend to have two associated tracks. The signature of LJ with merged cluster and two associated tracks provides a window of sensitivity to explore low mass γ_{d}s.

5.4 Prompt Lepton Jets Reconstruction

This section will explore the reconstruction method for γ_{d}s that promptly decay into collimated pairs of leptons, referred to as prompt LJs. The LJs are constructed based on the merged EM cluster signature discussed in Section 5.3.2, using standard physics objects such as electrons, muons, and tracks in the following procedures.

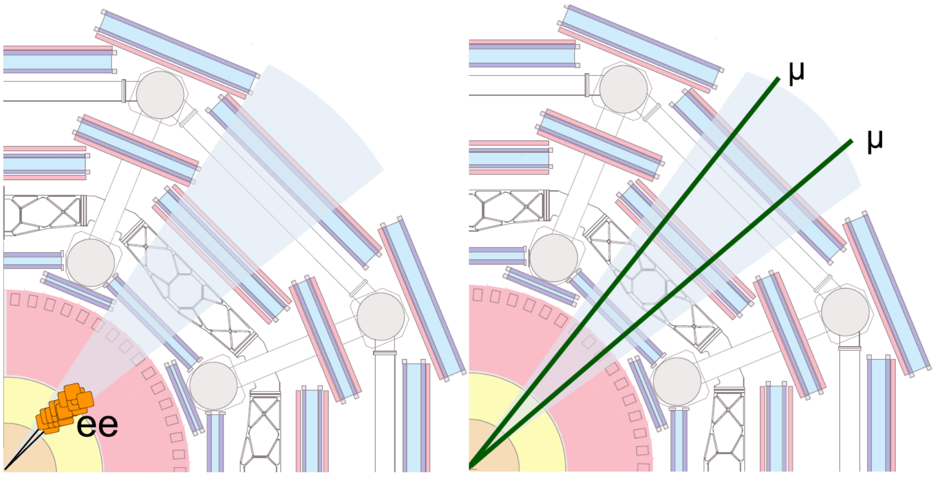
Electrons and muons are first clustered using the Cambridge-Aachen clustering algorithm [112] by selecting the highest p_{T} seed from the pool of reconstructed electrons and muons that meet the requirements outlined in Table 5.5. It then searches for additional leptons within a $\Delta R = 0.4$ radius around this seed. Each lepton contributes to updating the four-momentum vector of the LJ candidate and is subsequently removed from the pool while the updated LJ stays in. This iterative process continues until no leptons remain in the pool. There is no upper limit on the number of leptons in LJ, allowing the object to be sensitive to scenarios with more γ_{d}s radiated. Additionally, for each LJ candidate, tracks that meet the requirements summarized in Table 5.6 are also added. LJs are classified into two exclusive categories, as illustrated in Fig. 5.11, the following:



(a) (b)

Fig. 5.10 Number of EM clusters as a function of truth ΔR between electrons decaying from FRVZ γ_d with the mass assumption of (a) 0.1 GeV and (b) 6 GeV. The peak at 1 cluster for 0.1 GeV corresponds to the merged cluster scenario, while the peak at 2 clusters for 6 GeV corresponds to the separated close-by clusters scenario.

- ***Electronic Lepton-Jet*** (eLJ): If at least one electron, at least two associated tracks and no muons are found inside the cone.
- ***Muonic Lepton-Jet*** (**μLJ**): If at least two muons and no electrons are found inside the cone.



(a) electronic LJ (b) muonic LJ

Fig. 5.11 Illustration of reconstructed (a) eLJ, and (b) μLJ.

From the event topology of the FRVZ and HAHM benchmark models, two LJs are produced back-to-back in the transverse plane. In this analysis, we select two objects of interest: ***leading eLJ*** for the eLJ with largest p_T, and ***furthest eLJ*** for the eLJ with largest angular separation $\Delta\phi$ with leading eLJ in the transverse plane. Based on this configuration, we define three distinct search channels:

- eLJ-eLJ ***Electron Channel***: If both the leading and the furthest LJs are both eLJs.
- μLJ-μLJ ***Muon Channel***: If both the leading and the furthest LJs are both μLJs.
- eLJ-μLJ ***Mixed Channel***: If the leading and the furthest LJs are eLJs and μLJs each.

In this thesis, we will focus on the search for dark photons decaying into electron final states using the electron channel. Results from the muon and mixed channel are shown in Chapter 9.

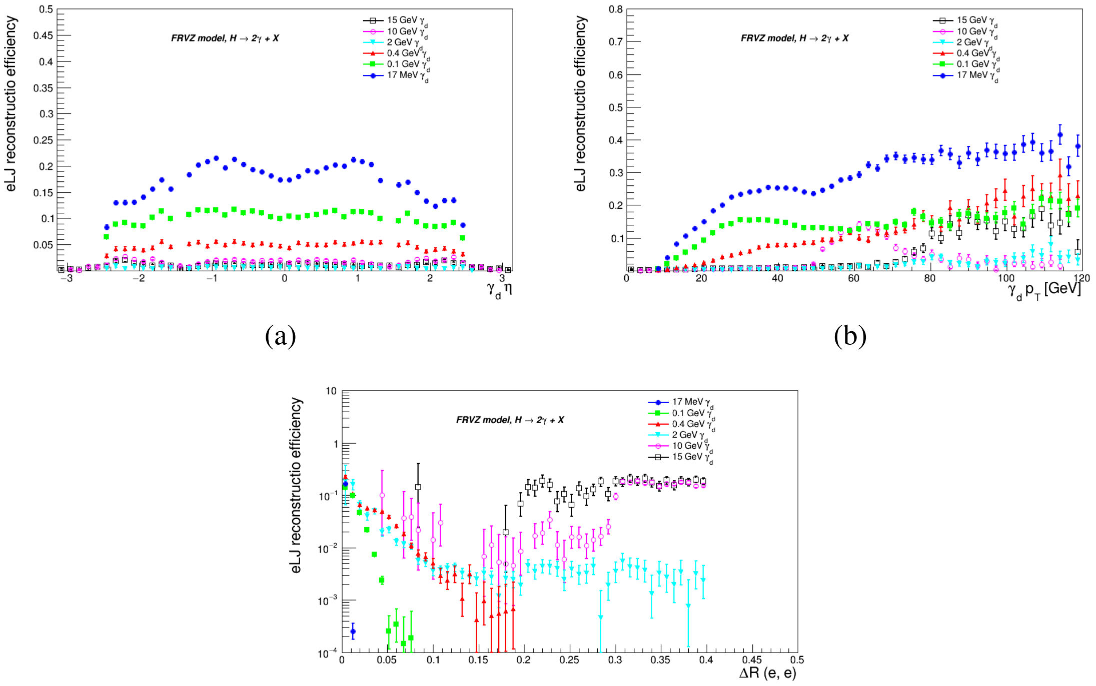
5.4.1 eLJ Reconstruction Efficiency

Reconstruction efficiency of eLJ is also shown as function of truth η, p_{T}, and ΔR between electrons decaying from γ_{d} for the FRVZ and HAHM benchmark model in Fig. 5.12 (b) and Fig. 5.13 (b). The drop in efficiency at $|\eta| = 1.5$ in Fig. 5.12 (a) and Fig. 5.13 (a) corresponds to transition regions from calorimeter barrel to end-cap. The eLJ reconstruction efficiencies as functions of ΔR between electrons from γ_{d} decays is shown in Fig. 5.12 (c) and Fig. 5.13 (c). In general, eLJ reconstruction efficiency increases with larger p_{T} and drops with larger ΔR between electrons for $m_{\gamma_{\mathrm{d}}} < 2$ GeV. This results in higher efficiency for the HAHM model compared with the FRVZ model.

The eLJ compositions introduce an additional complication. We investigate the reconstruction efficiency of eLJs composed of either a single reconstructed electron (forming a merged cluster) or two separate reconstructed electrons (forming distinct clusters). As shown in Fig. 5.14, the efficiency distribution is relatively flat in η for eLJs composed of a single electron. In contrast, eLJ consisting of two electrons has higher efficiency for $|\eta| < 1.5$. This is due to the better calorimeter resolution in the barrel region, which allows close-by clusters to be reconstructed separately. Fig. 5.15 and Fig. 5.16 show the reconstruction efficiency as functions of γ_{d} p_{T} and ΔR between electrons from γ_{d} decays at different mass points.

5.5 Triggers

A combination of single and multi-electron triggers are used to trigger eLJs. This section will discuss the details of trigger strategies and efficiencies.



(a)

(b)

(c)

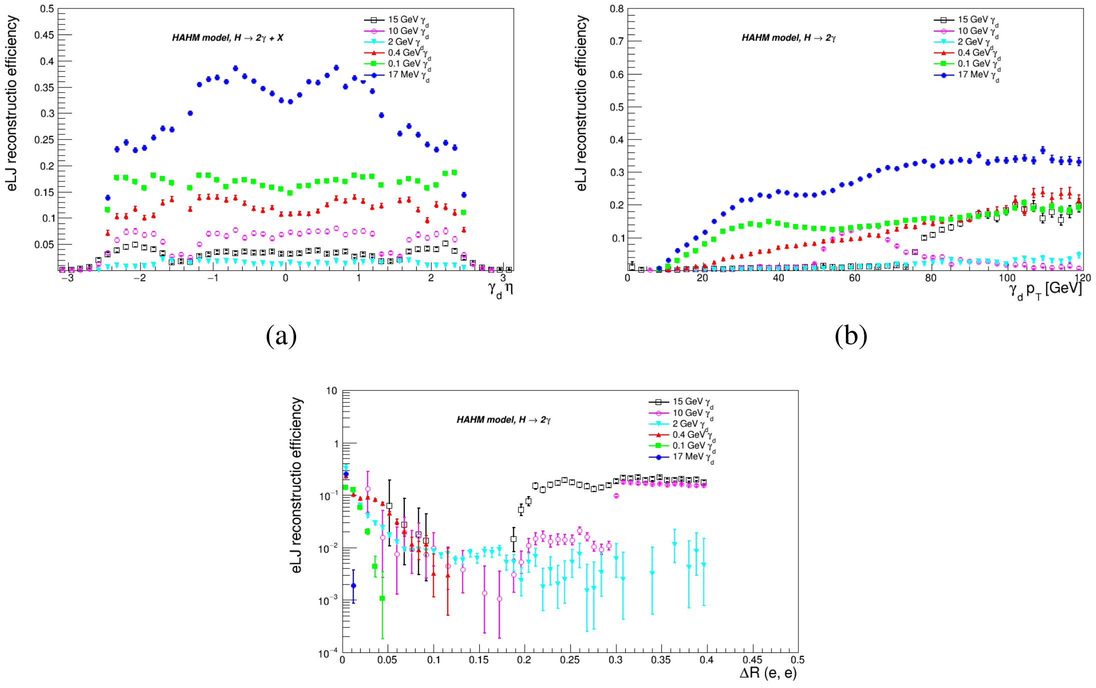
Fig. 5.12 *e*LJ reconstruction efficiency as function of (a) truth η of γ_d, (b) truth p_T of γ_d, and (c) truth ΔR between electrons decaying from γ_d at different γ_d mass in the FRVZ benchmark model.

5.5.1 Trigger Choice

To maximize signal efficiency, we select a combination of the lowest un-prescaled single-electron and di-electron triggers for different data-taking periods, as listed in Table 5.7. The trigger names follow the conventions explained in Sec. 4.2.7.

5.5.2 Trigger Efficiency

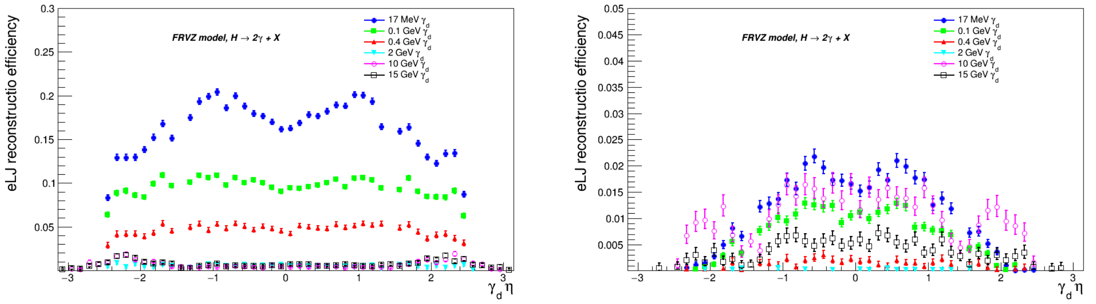
Fig. 5.17 shows the trigger efficiency of single-electron, multi-electron, and combination of any triggers as functions of p_T of 17 MeV γ_d in the FRVZ benchmark model. It is noticed that even with the tighter isolation requirement and higher p_T threshold, single electron triggers exhibit better trigger efficiency. With the complementary di-electron triggers, overall efficiency is maximized. Trigger efficiencies of single-electron and di-electron triggers for different mass points are shown in Fig. 5.18 and Fig. 5.19, respectively. Here, events are pre-selected, requiring both γ_d to decay into a pair of electrons.



(a) (b)

(c)

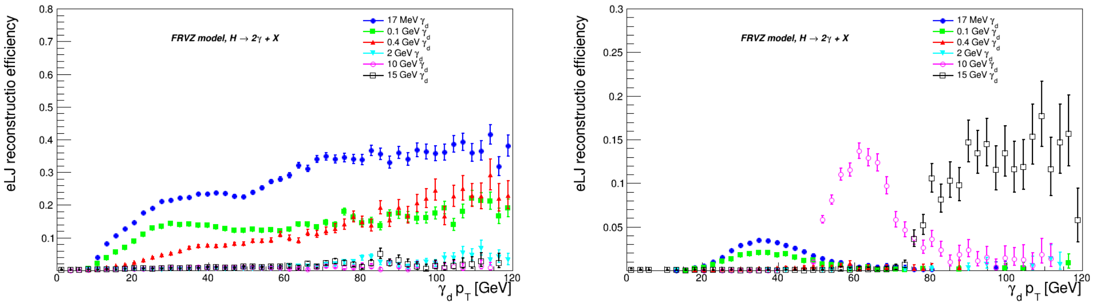
Fig. 5.13 eLJ reconstruction efficiency as function of (a) truth η of $\gamma_{\rm d}$, (b) truth $p_{\rm T}$ of $\gamma_{\rm d}$, and (c) truth ΔR between electrons decaying from $\gamma_{\rm d}$ at different $\gamma_{\rm d}$ mass in the HAHM benchmark model.



(a) (b)

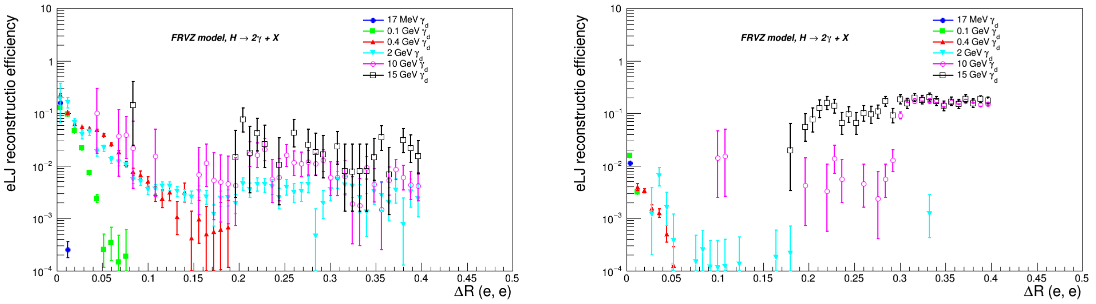
Fig. 5.14 Reconstruction efficiency in the category of (a) 1 electron and (b) 2 electrons inside eLJ as a function of truth η of $\gamma_{\rm d}$ at different $\gamma_{\rm d}$ mass in the FRVZ benchmark model.

The overall trigger efficiency for the FRVZ benchmark model is summarized in Table 5.10. The drop in electron trigger efficiency around 2 GeV is due to the transition from merged clusters to separate clusters scenario, where the standard object reconstruction becomes ineffective, as indicated in Fig. 5.9 (a).



(a) (b)

Fig. 5.15 Reconstruction efficiency in the category of (a) 1 electron and (b) 2 electrons inside *e*LJ as a function of truth p_{T} of γ_{d} at different γ_{d} mass in the FRVZ benchmark model.



(a) (b)

Fig. 5.16 Reconstruction efficiency in the category of (a) 1 electron and (b) 2 electrons inside *e*LJ as a function of truth ΔR between electrons decaying from γ_{d} at different γ_{d} mass in the FRVZ benchmark model.

Similarly, the trigger efficiency for the HAHM benchmark model is summarized in Table 5.11. As expected, the HAHM model generally shows higher efficiency due to the more significant boost of γ_{d} from the event topology. Note that the drop in electron trigger efficiency at the transition mass point of $m_{\gamma_d} = 2$ GeV also appears in the HAHM benchmark model.

The trigger efficiency in each data-taking year is shown in Table 5.8 and Table 5.9. These tables show the impact of varying p_{T} thresholds across different years. For each event in the MC simulations, we apply a scale factor, calibrated from data using $Z \to ee$ events, based on the combination of triggers outlined in Table. 5.7.

Table 5.7 List electron triggers used in the *e*LJ-*e*LJ channel for the corresponding data-taking periods.

Type	Data-taking periods	Trigger
Single-electron	2015	HLT_e24_lhmedium_L1EM20VH HLT_e60_lhmedium HLT_e120_lhloose
	2016 - 2018	HLT_e26_lhtight_nod0_ivarloose HLT_e60_lhmedium_nod0 HLT_e140_lhloose_nod0
Di-electron	2015	HLT_2e12_lhvloose_L12EM10VH
	2016	HLT_2e17_lhvloose_nod0
	2017 (only B5-B8)	HLT_2e24_lhvloose_nod0
	2017 (except B5-B8)	HLT_2e17_lhvloose_nod0_L12EM15VHI
	2018	HLT_2e17_lhvloose_nod0_L12EM15VHI

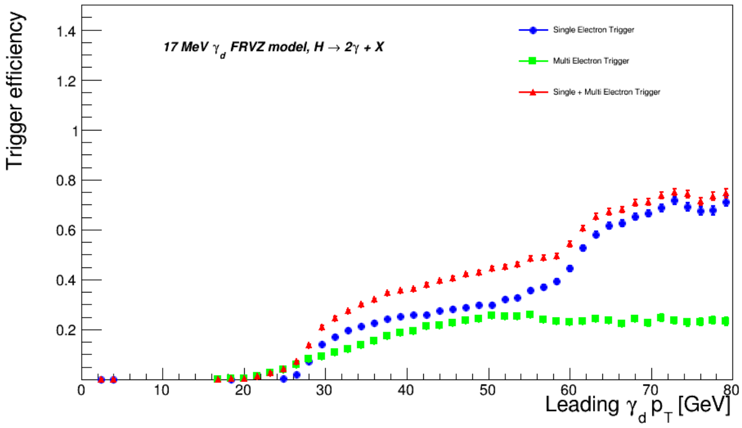
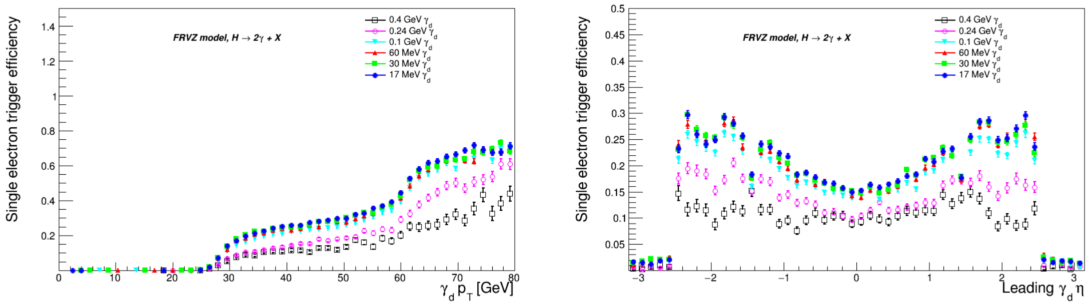


Fig. 5.17 Trigger efficiencies for γ_d at 17 MeV in the FRVZ signal model, shown for single-electron, multi-electron, and logical OR of any triggers, shown as a function of truth γ_d p_T .

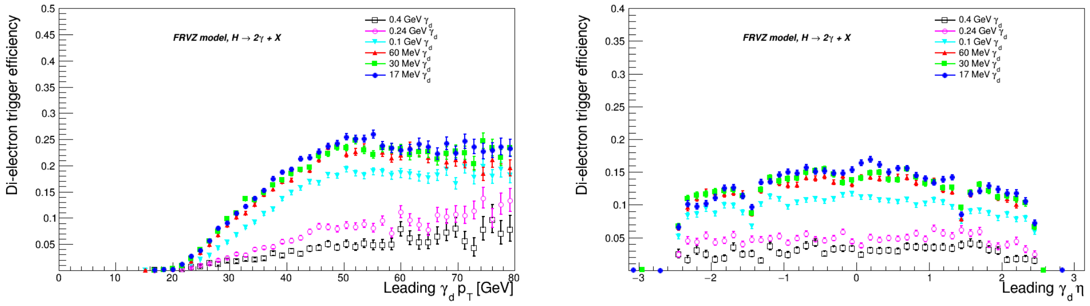
Table 5.8 Table of electron trigger efficiency for FRVZ benchmark model in different data-taking years.

FRVZ γ_d mass	17 MeV	30 MeV	60 MeV	0.1 GeV	0.24 GeV	0.4 GeV	0.9 GeV	2.0 GeV	6.0 GeV
2015	41%	41%	39%	36%	24%	18%	17%	31%	53%
2016	26%	25%	25%	22%	14%	10%	8.4%	13%	33%
2017	23%	22%	22%	19%	14%	10%	7.4%	8.1%	26%
2018	24%	23%	23%	20%	14%	10%	7.3%	8.4%	26%



(a) (b)

Fig. 5.18 Single-electron trigger efficiencies for the FRVZ signal model, shown for different mass points and as a function of the $\gamma_{\rm d}$ (a) $p_{\rm T}$ and (b) η.



(a) (b)

Fig. 5.19 Multi-electron trigger efficiencies for the FRVZ signal model, shown for different mass points and as a function of the $\gamma_{\rm d}$ (a) $p_{\rm T}$ and (b) η.

Table 5.9 Table of electron trigger efficiency for HAHM benchmark model in different data-taking years.

HAHM $\gamma_{\rm d}$ mass	17 MeV	0.1 GeV	0.4 GeV	2 GeV	10 GeV	15 GeV	25 GeV	40 GeV
2015	80%	75%	46%	36%	84%	85%	85%	91%
2016	72%	68%	37%	20%	78%	80%	81%	84%
2017	69%	66%	41%	13%	72%	76%	78%	81%
2018	69%	66%	40%	13%	73%	77%	80%	83%

Table 5.10 Table of electron trigger efficiency for FRVZ benchmark model at different $\gamma_{\rm d}$ mass points.

FRVZ $\gamma_{\rm d}$ mass	17 MeV	30 MeV	60 MeV	0.1 GeV	0.24 GeV	0.4 GeV	0.9 GeV	2.0 GeV	6.0 GeV
Single-electron	18%	18%	16%	16%	12%	9.3%	6.4%	7.7%	22%
Di-electron	12%	11%	11%	8.4%	4.1%	2.5%	1.7%	3.0%	11%
Combined	25%	24%	24%	21%	15%	11%	7.5%	9.5%	26%

Table 5.11 Table of electron trigger efficiency for HAHM benchmark model at different γ_d mass points.

HAHM γ_d mass	17 MeV	0.1 GeV	0.4 GeV	2 GeV	10 GeV	15 GeV	25 GeV	40 GeV
Single-electron	58%	56%	34%	12%	67%	72%	73%	75%
Di-electron	52%	46%	18%	5%	56%	63%	66%	69%
Combined	70%	67%	39%	15%	74%	78%	80%	83%

5.5.3 Trigger-matching

To ensure that the electrons firing the triggers are the same ones building the eLJs, we implement additional trigger-matching criteria as follows:

1. Offline reconstructed electron has p_T above HLT threshold of ≥ 1 GeV.
2. $\Delta R < 0.2$ between offline reconstructed electron and HLT trigger objects.

In addition, HLT objects are required to be matched to offline electrons within *e*LJ. In the event fired by di-electron triggers, two HLT objects are needed to match electrons from different *e*LJs, while single-electron triggers require an HLT object to match electrons from any *e*LJ. Fig. 5.20 shows the trigger matching efficiency defined as $\varepsilon = \frac{N_\mathrm{trigger\ matched}}{N_\mathrm{triggered}}$ for the FRVZ and HAHM models, where we see $> 85\%$ efficiency at different γ_d mass points.

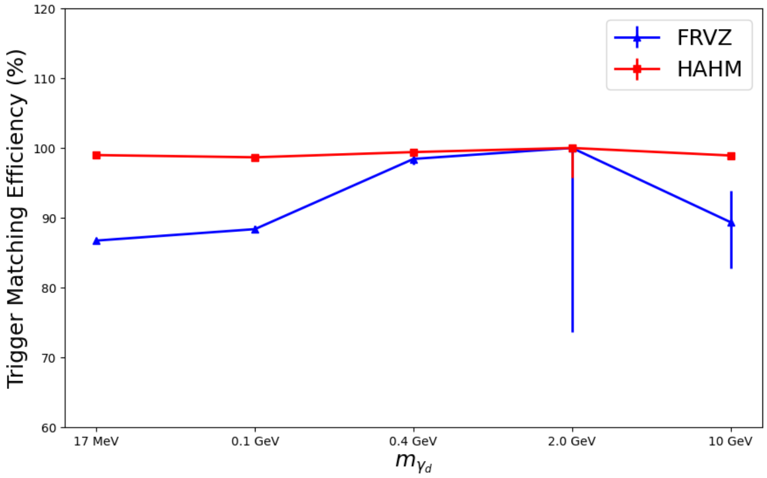


Fig. 5.20 Trigger matching efficiencies for the FRVZ signal model, shown for different mass points.

Chapter 6

Event Selection

This chapter will discuss the event selections in the search for promptly decaying dark photons conducted through the *e*LJ-*e*LJ channel. MC simulations of the signal and SM background are used to optimize the selections and understand the background compositions.

Sec. 6.1 defines the event selections for the signal region. Sec. 6.2 describes the background processes and presents the yields for both signal and background events. Sec. 6.3 defines the control region in data and compares it with MC simulation for the variables used in event selections.

6.1 Signal Region Definition

Based on the eLJ signatures of back-to-back γ_ds, and collimated pairs of electrons from γ_d decays as described in Sec. 3.2.3, we apply the following ***pre-selections*** to each processed event:

- the event must be listed in the *good run list* where all sub-detectors were in good data-taking conditions (see Sec. 4.1.3).
- the presence of a good primary vertex is required (see Sec. 4.3.1).
- at least two *e*LJs must be reconstructed (see Sec. 5.4).
- the event must be triggered by logical OR of single-electron and di-electron triggers and matched (see Sec. 5.5).

In each event, as mentioned in Sec. 5.4, the analysis selects two objects of interest: ***leading eLJ*** for the *e*LJ with the largest p_T, and ***furthest eLJ*** for the *e*LJ with most significant angular separation $\Delta\phi$ in the transverse plane with respect to leading *e*LJ.

To suppress the dominant SM background events due to the random overlap of electrons with additional tracks and reduce subdominant SM background events to a negligible level, the following sets of requirements are applied:

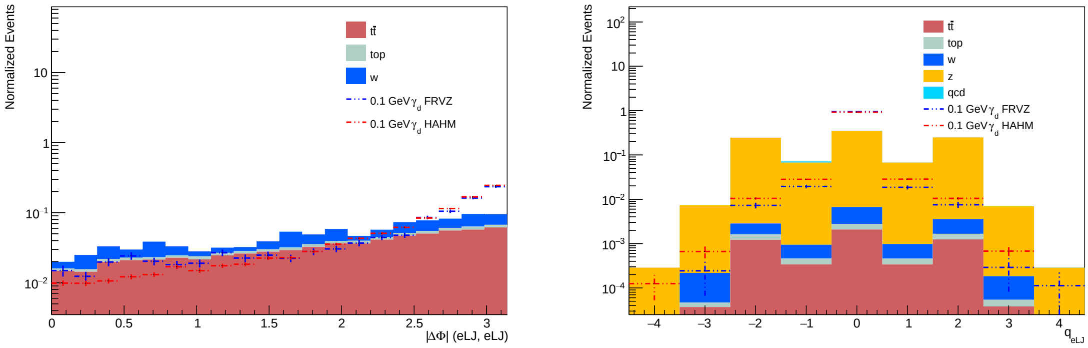
- ***eLJ selections*** applied for both *e*LJs:
  - reconstructed with exactly one EM cluster;
  - the leading track has $p_{\mathrm{T}} > 5$ GeV;
  - $|\eta| < 1.37$;
  - sum of charge for leading two tracks $q_{e\mathrm{LJ}} = 0$.
- ***event selections*** applied for each event:
  - $|\phi^{\mathrm{lead}}_{e\mathrm{LJ}} - \phi^{\mathrm{furthest}}_{e\mathrm{LJ}}| > 2.5$;
  - combined invariant mass $m_{e\mathrm{LJ}\text{-}e\mathrm{LJ}}$ must satisfy the Z-mass veto, hence $m_{e\mathrm{LJ}\text{-}e\mathrm{LJ}} \notin [80, 100]$ GeV;
  - mass imbalance $m_{\mathrm{imb}} = \frac{|m^{\mathrm{lead}}_{e\mathrm{LJ}} - m^{\mathrm{furthest}}_{e\mathrm{LJ}}|}{m^{\mathrm{lead}}_{e\mathrm{LJ}} + m^{\mathrm{furthest}}_{e\mathrm{LJ}}} < 0.8$, where $m_{e\mathrm{LJ}}$ is built from leading two tracks in *e*LJ with electron mass hypothesis.
- ***other selection***:
  - no jet with $p_{\mathrm{T}} > 40$ GeV.

The reason why these cuts are required is explained below. The definition of *e*LJ includes both scenarios where merged and separate clusters are found within the cone. Since the *e*LJ-*e*LJ channel is particularly relevant for low-mass γ_{d} below $2m_{\mu} \sim 0.21$ GeV, which lies within the merged cluster regime as shown in Figs. 5.2 and 5.9, we require exactly one electron to help simplify the event compositions in the SM background while retaining more than 90% of the signal yield (see Sec. 5.3.2).

The two *e*LJs are expected to be back-to-back from the signal event topology, while this is not the case for generic SM background processes. As we can see from Fig. 6.1 (a), signal *e*LJs are more separated in $\Delta\phi$.

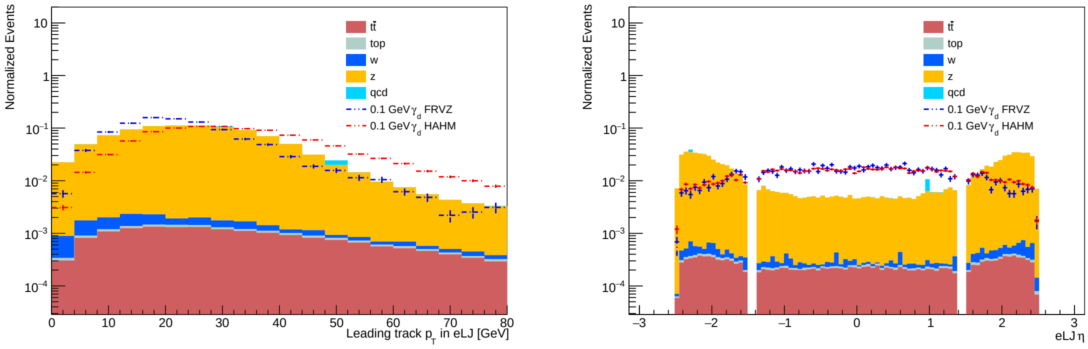
Two electron tracks from neutral particle decays are expected to possess opposite signed charges, while random tracks overlapping electrons can have random signed charges. Fig. 6.1 (b) shows the $q_{e\mathrm{LJ}}$ for signal and background processes.

Fig 6.2 (a) shows the leading track p_{T} distributions for signal and background. $p_{\mathrm{T}} > 5$GeV is selected to optimize the signal-to-background ratio.



(a) (b)

Fig. 6.1 Distributions at pre-selections level of (a) $|\Delta\phi|$ between two eLJs, and (b) $q_{e\mathrm{LJ}}$ for SM processes, FRVZ and HAHM signal with $\gamma_\mathrm{d} = 0.1$ GeV.



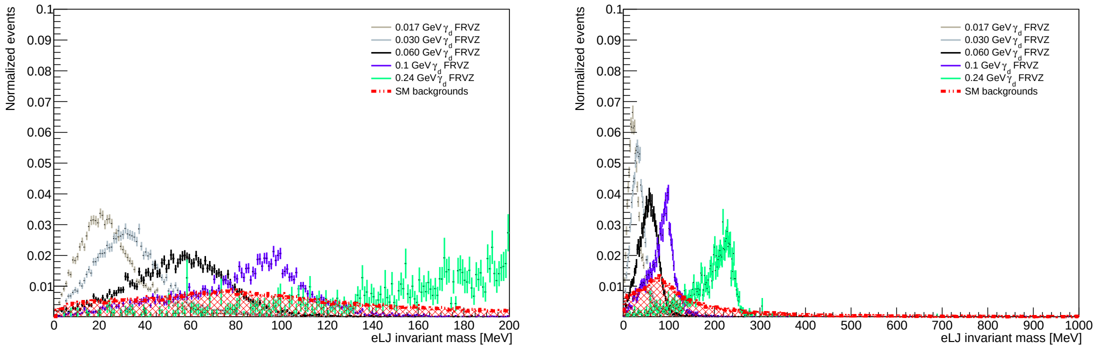
(a) (b)

Fig. 6.2 Distributions at pre-selections level of (a) leading track p_T, and (b) η for SM processes, FRVZ and HAHM signal with $\gamma_\mathrm{d} = 0.1$ GeV.

From Fig 6.2 (b), it is observed that SM processes are dominant in the high-η region, whereas signal events are more frequently found in the low-η region. This is due to the higher level of jet backgrounds in the end-cap region ($|\eta| > 1.5$), which increases the chance of mis-associating a random jet with an electron, thereby faking an eLJ.

The invariant mass of eLJ can be reconstructed from the leading two associated tracks with electron mass hypothesis. Fig. 6.3 shows the reconstructed mass for eLJ in both signal and SM background processes. Signal eLJ peaks at the corresponding γ_d mass, while background eLJ exhibits a flatter kinematic distribution. Consequently, invariant mass imbalance provides discrimination over SM processes, as shown in Fig. 6.4 (a).

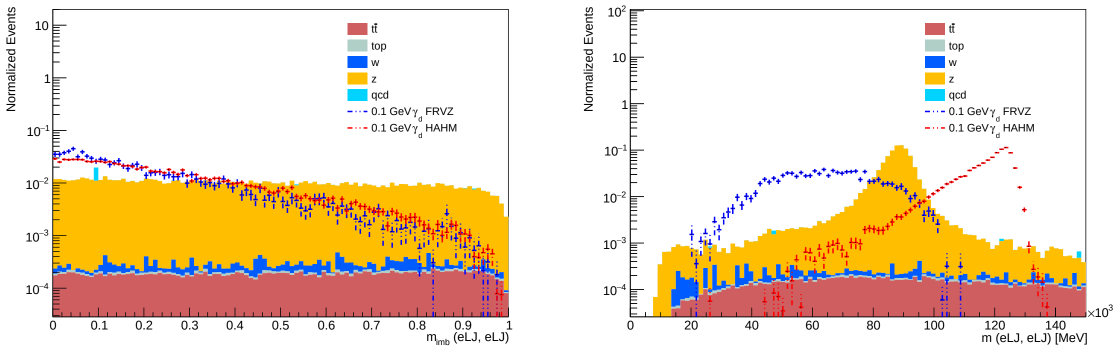
In addition, it is observed that prompt electrons from Z boson decays can sometimes be misidentified as two distinct eLJs in the presence of additional tracks. This process can



(a)

(b)

Fig. 6.3 Distributions at pre-selections level of *e*LJ invariant mass in (a) [0, 200 MeV], and (b) [0, 1000 MeV] range for SM processes, FRVZ and HAHM signal with different γ_d mass points.



(a)

(b)

Fig. 6.4 Distributions at pre-selections level of (a) *e*LJ invariant mass imbalance and (b) *e*LJ-*e*LJ invariant mass for SM processes, FRVZ and HAHM signal with $\gamma_d = 0.1$ GeV.

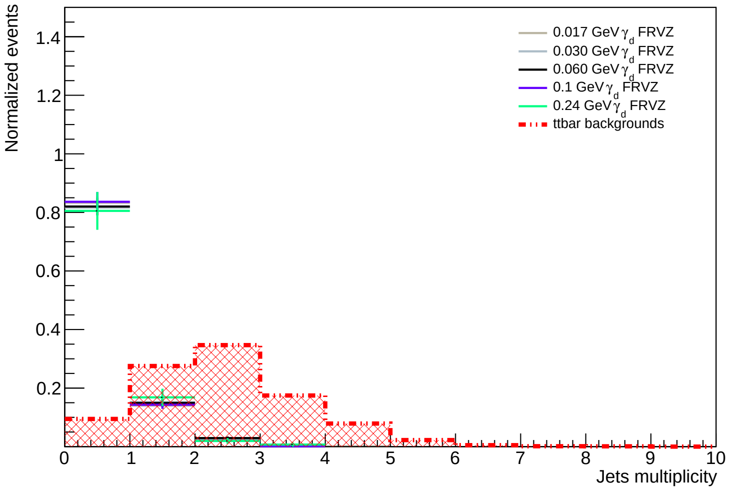


Fig. 6.5 Distributions at pre-selections level of the number of jets with $p_{\mathrm{T}} > 40$ GeV for $t\bar{t}$ processes and FRVZ signal with different γ_{d} mass points.

be suppressed by rejecting events where the invariant mass of the two *e*LJs is around the *Z* boson mass, as illustrated in Fig. 6.4 (b).

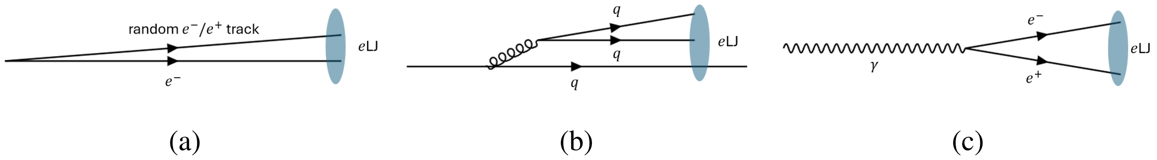
Jet multiplicity is used as a discriminating variable to reduce the impact of $t\bar{t}$ processes. According to the number distribution of jets with $p_{\mathrm{T}} > 40$ GeV shown in Fig. 6.5, the jet multiplicity requirement effectively suppresses $t\bar{t}$ events by a factor of 10 while the signal events are accepted with a rate more than 80%.

6.2 Expected Signal yields and Background processes

Table. 6.1 and Table. 6.2 present the number of events remaining after each selection step for the FRVZ and HAHM benchmark signal models, respectively, as described in the previous section. The MC events are scaled to 140 fb^{-1}, with the signal MC events assuming $\sigma = \sigma_{\mathrm{ggF}}$ and $BR(H \to 2\gamma_{\mathrm{d}} + X = 0.5\%)$. The event selections are optimized for collimated *e*LJ signatures, resulting in higher efficiency at low γ_{d} masses. Based on the signal yields obtained after event selections, we focus on γ_{d} mass less than 0.24 GeV for the FRVZ model and 0.4 GeV for the HAHM model.

To understand the misidentification of SM processes, as illustrated in Fig. 6.6, we first categorize fake *e*LJs into the following types:

- ***electron-fake*** *e*LJ: when a real electron, promptly decaying from a hard-scattering event (e.g., from *W* or *Z* bosons), overlaps with additional tracks that are occasionally



(a) (b) (c)

Fig. 6.6 Illustration of (a) electron-fake eLJ, (b) jet-fake eLJ, and (c) photon conversions eLJ.

mis-associated with the electron, a fake eLJ can be formed from the real electron. Here, the additional tracks could come from EM showering or random jets.

- ***jet-fake*** eLJ: clusters and multiple tracks induced from jets get reconstructed as fake eLJ.
- ***photon conversions*** eLJ: photons get converted internally or within materials and decay to collimated pairs of electrons to form fake eLJ.

SM processes generate fake eLJ-eLJ events using two such fake eLJs. The specifics are as follows: $t\bar{t}$, single-top, and W+jets events from one electron-fake eLJ and one jet-fake eLJ; Z+jets events from two electron-fake eLJs; multijet events from two jet-fake eLJs; γ+jets events from one photon conversion eLJ and one jet-fake eLJ.

Table 6.3 presents the number of events from SM process simulations for $t\bar{t}$, single-top, Z+jets, W+jets. Table 6.4 shows the number of multijet events, where some MC events are scaled to 10^6 due to large cross-sections at low p_{T} slices. To ensure the effective suppression of multijet processes, we perform additional checks on the event selections. This involves relaxing some selection criteria to increase the statistical sample size. Table 6.5 shows the number of events for low p_{T} slices of multijet processes without applying identification or isolation selections on the electrons used to construct eLJs. This shows an additional suppression factor of 10^5 beyond the identification and isolation requirements. Therefore, multijet processes are effectively suppressed by our event selections. Other dibosons, γ+jets, and $Z\gamma$ processes are suppressed to be less than 10% of remaining backgrounds based on MC simulations, mostly due to `medium` electron identification and the jet veto requirement. The remaining background is dominated by Z+jets, where both eLJs are electron-fake eLJs originating from Z boson decays. The actual number of background events entering the SR is estimated using data only as detailed in Chapter 7. The contribution from other processes is accounted for by systematic uncertainties obtained from validation regions in the background estimation method.

Table 6.1 Signal events remaining after each cut applied in the *e*LJ-*e*LJ channel. Events are generated according to the FRVZ model and are normalized assuming a branching ratio $B(H \to 2\gamma_d) = 0.5\%$, with $L = 140$ fb^{-1} and $\sigma = \sigma_{\mathrm{ggF}}$.

m_{γ_d} [GeV]	0.017	0.03	0.06	0.1	0.24	0.4	0.9	2	6
No cut	33500 ± 70	33500 ± 70	33500 ± 70	33500 ± 70	12500 ± 40	6600 ± 30	5400 ± 30	5400 ± 30	3200 ± 20
2 *e*LJs	962 ± 11	774 ± 9.9	567 ± 8.4	416 ± 7.2	107 ± 3.6	27 ± 1.8	4.3 ± 0.71	0.60 ± 0.25	6.7 ± 0.86
Trigger Matched	833 ± 10	654 ± 9.0	485 ± 7.8	366 ± 6.7	101 ± 3.6	27 ± 1.8	4.3 ± 0.71	0.60 ± 0.25	6.5 ± 0.85
1 electron in *e*LJ	753 ± 9.6	583 ± 8.5	432 ± 7.3	324 ± 6.3	96 ± 3.4	26 ± 1.8	4.2 ± 0.70	0.60 ± 0.25	0.23 ± 0.17
Leading track $p_{\mathrm{T}} > 5$ GeV	753 ± 9.6	583 ± 8.5	432 ± 7.3	324 ± 6.3	96 ± 3.4	26 ± 1.8	4.2 ± 0.70	0.60 ± 0.25	0.23 ± 0.17
*e*LJ $\lvert\eta\rvert < 1.37$	481 ± 7.7	357 ± 6.7	265 ± 5.8	182 ± 4.7	59 ± 2.7	17 ± 1.5	2.7 ± 0.57	0.19 ± 0.14	0.11 ± 0.11
$q_{e\mathrm{LJ}} = 0$	480 ± 7.7	357 ± 6.7	265 ± 5.8	182 ± 4.7	59 ± 2.7	17 ± 1.5	2.7 ± 0.57	0.19 ± 0.14	0.11 ± 0.11
$\Delta\Phi(e\mathrm{LJ}, e\mathrm{LJ}) > 2.5$	250 ± 5.5	185 ± 4.8	140 ± 4.2	106 ± 3.6	28 ± 1.8	7.2 ± 0.94	1.4 ± 0.41	0 ± 0	0 ± 0
Z mass veto	208 ± 5.0	156 ± 4.4	116 ± 3.8	91 ± 3.3	23 ± 1.7	4.8 ± 0.76	1.1 ± 0.36	0 ± 0	0 ± 0
$m_{e\mathrm{LJ}}$ imbalance < 0.8	203 ± 5.0	154 ± 4.4	115 ± 3.8	90 ± 3.3	23 ± 1.7	4.6 ± 0.75	0.88 ± 0.32	0 ± 0	0 ± 0
No jet with $pt > 40$ GeV	176 ± 4.6	132 ± 4.1	100 ± 3.6	79 ± 3.1	20 ± 1.5	4.4 ± 0.73	0.73 ± 0.28	0 ± 0	0 ± 0

Table 6.2 Signal events yield after each selection applied in the *e*LJ-*e*LJ channel. Events are generated according to the HAHM model and are normalized assuming a branching ratio $B(H \to 2\gamma_d) = 0.5\%$, with $L = 140$ fb^{-1} and $\sigma = \sigma_{\mathrm{ggF}}$.

m_{γ_d} [GeV]	0.017	0.1	0.4	2	10	15	25
No cut	33500 ± 70	33500 ± 70	6600 ± 30	5500 ± 30	3500 ± 20	3400 ± 20	3300 ± 20
2 *e*LJs	4200 ± 23	1600 ± 14	240 ± 5.5	4.3 ± 0.70	115 ± 3.7	24 ± 1.7	17 ± 1.4
Trigger Matched	4200 ± 23	1600 ± 14	230 ± 5.5	4.3 ± 0.70	114 ± 3.7	23 ± 1.7	16 ± 1.4
1 electron in *e*LJ	3900 ± 22	1500 ± 14	230 ± 5.4	4.2 ± 0.70	2.3 ± 0.50	8.3 ± 1.0	10 ± 1.1
Leading track $p_{\mathrm{T}} > 5$ GeV	3900 ± 22	1500 ± 14	230 ± 5.4	4.2 ± 0.70	2.3 ± 0.50	8.3 ± 1.0	10 ± 1.1
*e*LJ $\lvert\eta\rvert < 1.37$	2600 ± 18	770 ± 9.9	130 ± 4.2	2.9 ± 0.57	0.25 ± 0.18	1.3 ± 0.38	1.4 ± 0.43
$q_{e\mathrm{LJ}} = 0$	2600 ± 18	770 ± 9.9	130 ± 4.2	2.9 ± 0.57	0.12 ± 0.12	0.88 ± 0.32	1.3 ± 0.41
$\Delta\Phi(e\mathrm{LJ}, e\mathrm{LJ}) > 2.5$	1600 ± 15	460 ± 7.6	65 ± 2.9	0.56 ± 0.26	0 ± 0	0.18 ± 0.14	0.42 ± 0.21
Z mass veto	1600 ± 14	420 ± 7.3	62 ± 2.8	0.56 ± 0.26	0 ± 0	0.062 ± 0.062	0.19 ± 0.13
$m_{e\mathrm{LJ}}$ imbalance < 0.8	1500 ± 14	420 ± 7.3	62 ± 2.8	0.56 ± 0.26	0 ± 0	0.062 ± 0.062	0.19 ± 0.13
No jet with $pt > 40$ GeV	1300 ± 13	360 ± 6.8	52 ± 2.5	0.56 ± 0.26	0 ± 0	0.062 ± 0.062	0.17 ± 0.13

Table 6.3 Events yield after each selection applied in the *e*LJ-*e*LJ channel, for different MC background samples and full Run-2 data. Events are normalized to $L = 140$ fb^{-1}.

	$t\bar{t}$	single-top	Z+jets	W+jets	Data
No cut	$1.3 \times 10^8 \pm 2400$	$4.7 \times 10^5 \pm 71$	$1.3 \times 10^9 \pm 27000$	$9.1 \times 10^9 \pm 130000$	/
2 *e*LJs	1550 ± 16	241 ± 5.9	130360 ± 760	1164 ± 100	123043
Trigger Matched	1390 ± 15	239 ± 5.8	129470 ± 750	1110 ± 93	121746
1 electron in *e*LJ	1380 ± 15	238 ± 5.8	129140 ± 750	1111 ± 93	121298
Leading track $p_{\mathrm{T}} > 5$ GeV	1380 ± 15	238 ± 5.8	129060 ± 750	1110 ± 93	121239
*e*LJ $\lvert\eta\rvert < 1.37$	369 ± 7.6	52 ± 2.6	12650 ± 240	147 ± 36	17960
$q_{e\mathrm{LJ}} = 0$	256 ± 6.3	36 ± 2.2	8370 ± 180	119 ± 35	12866
$\Delta\Phi(e\mathrm{LJ}, e\mathrm{LJ}) > 2.5$	90 ± 3.7	13 ± 1.4	6600 ± 170	29 ± 7.2	9446
Z mass veto	81 ± 3.5	12 ± 1.3	1222 ± 73	29 ± 7.0	2452
$m_{e\mathrm{LJ}}$ imbalance < 0.8	64 ± 3.1	11 ± 1.2	1038 ± 67	26 ± 6.9	1978
No jet with $pt > 40$ GeV	6.2 ± 0.96	2.1 ± 0.53	936 ± 66	19 ± 6.6	1704

6.3 Checks of Selection Variables

To verify and compare the shapes of event selection variables, we define a control region to compare MC simulations of SM background with data. The dominant background process

Table 6.4 Multijet events remaining after each cut are applied in the *e*LJ-*e*LJ channel in different p_{T} slices. Events are normalized to $L = 140$ fb^{-1}. Large events scales are due to large cross sections at low jet p_{T} slices.

multijet slices jet $p_{\mathrm{T}}[GeV]$	Slice 0 [0, 20]	Slice 1 [20, 60]	Slice 2 [60, 160]	Slice 3 [160, 400]	Slice 4 [400, 800]	Slice 5 [800, 1300]	Slice 6 [1300, 1800]
No cut	$10^{16} \pm 5 \times 10^{11}$	$3 \times 10^{14} \pm 10^{10}$	$3 \times 10^{12} \pm 3 \times 10^{8}$	$4 \times 10^{10} \pm 10^{6}$	$4 \times 10^{8} \pm 10^{5}$	$9 \times 10^{6} \pm 400$	$3 \times 10^{5} \pm 40$
2 *e*LJs	0 ± 10^{6}	0 ± 10^{6}	11000 ± 7200	560 ± 210	110 ± 12	17 ± 1.1	2.2 ± 0.17
Trigger Matched	0 ± 10^{6}	0 ± 10^{6}	1100 ± 1100	300 ± 150	85 ± 11	12 ± 0.89	1.5 ± 0.15
1 electron in *e*LJ	0 ± 10^{6}	0 ± 10^{6}	1100 ± 1100	300 ± 150	84 ± 11	12 ± 0.89	1.5 ± 0.15
Leading track $p_{\mathrm{T}} > 5$ GeV	0 ± 10^{6}	0 ± 10^{6}	1100 ± 1100	300 ± 150	84 ± 11	12 ± 0.89	1.5 ± 0.15
*e*LJ $\lvert\eta\rvert < 1.37$	0 ± 10^{6}	0 ± 10^{6}	0 ± 1100	90 ± 63	41 ± 7.3	8.0 ± 0.75	1.4 ± 0.15
$q_{e\mathrm{LJ}} = 0$	0 ± 10^{6}	0 ± 10^{6}	0 ± 1100	90 ± 63	35 ± 6.8	5.4 ± 0.61	0.80 ± 0.11
$\Delta\Phi(e\mathrm{LJ}, e\mathrm{LJ}) > 2.5$	0 ± 10^{6}	0 ± 10^{6}	0 ± 1100	0 ± 45	17 ± 4.5	4.0 ± 0.50	0.67 ± 0.11
Z mass veto	0 ± 10^{6}	0 ± 10^{6}	0 ± 1100	0 ± 45	17 ± 4.5	4.0 ± 0.50	0.67 ± 0.11
$m_{e\mathrm{LJ}}$ imbalance < 0.8	0 ± 10^{6}	0 ± 10^{6}	0 ± 1100	0 ± 45	14 ± 4.3	3.4 ± 0.46	0.58 ± 0.10
No jet with $pt > 40$ GeV	0 ± 10^{6}	0 ± 10^{6}	0 ± 1100	0 ± 45	3.8 ± 2.0	0.84 ± 0.21	0.082 ± 0.025

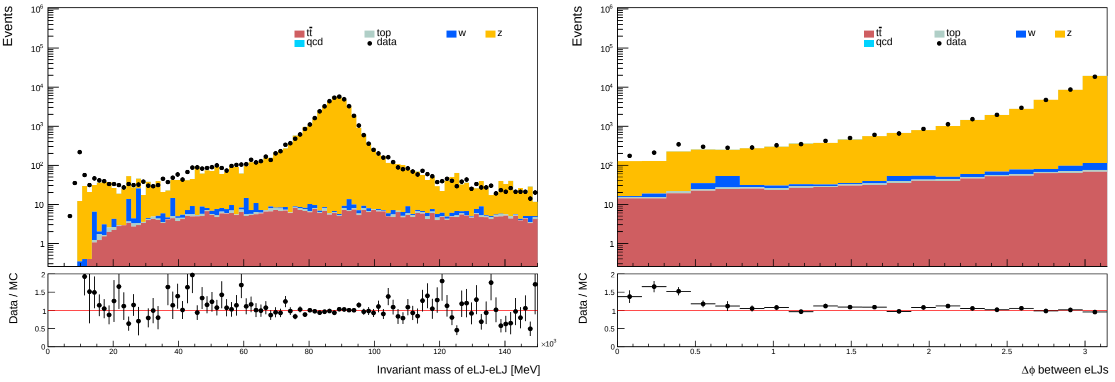
Table 6.5 Multijet events remaining after each cut are applied in the *e*LJ-*e*LJ channel without applying identification or isolation selections on the electrons used to construct *e*LJs. The suppression factor of these event selections is estimated to be 10^5. Events are normalized to $L = 140$ fb^{-1}.

multijet slices jet $p_{\mathrm{T}}[GeV]$	Slice 0 [0, 20]	Slice 1 [20, 60]	Slice 2 [60, 160]
No cut	$10^{16} \pm 5 \times 10^{11}$	$3 \times 10^{14} \pm 10^{10}$	$3 \times 10^{12} \pm 3 \times 10^{8}$
2 *e*LJs (`medium` ID and `loose` isolation)	0 ± 10^{6}	0 ± 10^{6}	11000 ± 7200
2 *e*LJs (no ID or isolation)	$1.97 \times 10^{13} \pm 2.22 \times 10^{10}$	$3.63 \times 10^{11} \pm 4.32 \times 10^{8}$	$3.86 \times 10^{11} \pm 1.50 \times 10^{8}$
Trigger Matched	$2.58 \times 10^{9} \pm 2.39 \times 10^{8}$	$5.12 \times 10^{7} \pm 5.01 \times 10^{6}$	$4.73 \times 10^{7} \pm 1.45 \times 10^{6}$
1 electron in *e*LJ	$1.67 \times 10^{9} \pm 1.95 \times 10^{8}$	$3.09 \times 10^{7} \pm 4.05 \times 10^{6}$	$2.42 \times 10^{7} \pm 1.12 \times 10^{6}$
Leading track $p_{\mathrm{T}} > 5$ GeV	$1.67 \times 10^{9} \pm 1.95 \times 10^{8}$	$3.09 \times 10^{7} \pm 4.05 \times 10^{6}$	$2.41 \times 10^{7} \pm 1.12 \times 10^{6}$
*e*LJ $\lvert\eta\rvert < 1.37$	$2.32 \times 10^{8} \pm 6.90 \times 10^{7}$	$1.11 \times 10^{7} \pm 2.49 \times 10^{6}$	$5.97 \times 10^{6} \pm 4.38 \times 10^{5}$
$q_{e\mathrm{LJ}} = 0$	$1.28 \times 10^{8} \pm 4.78 \times 10^{7}$	$8.45 \times 10^{6} \pm 2.30 \times 10^{6}$	$4.46 \times 10^{6} \pm 3.92 \times 10^{5}$
$\Delta\Phi(e\mathrm{LJ}, e\mathrm{LJ}) > 2.5$	$1.28 \times 10^{8} \pm 4.78 \times 10^{7}$	$8.45 \times 10^{6} \pm 2.30 \times 10^{6}$	$3.63 \times 10^{6} \pm 3.80 \times 10^{5}$
Z mass veto	$7.10 \times 10^{7} \pm 3.82 \times 10^{7}$	$4.60 \times 10^{6} \pm 1.46 \times 10^{6}$	$3.04 \times 10^{6} \pm 3.58 \times 10^{5}$
$m_{e\mathrm{LJ}}$ imbalance < 0.8	$7.10 \times 10^{7} \pm 3.82 \times 10^{7}$	$3.01 \times 10^{6} \pm 1.16 \times 10^{6}$	$2.57 \times 10^{6} \pm 3.44 \times 10^{5}$
No jet with $pt > 40$ GeV	$7.10 \times 10^{7} \pm 3.82 \times 10^{7}$	$3.01 \times 10^{6} \pm 1.16 \times 10^{6}$	$2.38 \times 10^{6} \pm 3.35 \times 10^{5}$

involves additional tracks overlapping with prompt electrons decaying from Z bosons. Since the charge of these extra tracks is random relative to the electron, we can define a control region by requiring $|q_{e\mathrm{LJ}}| > 0$. This CR should, in principle, contain the same background processes as the SR but with negligible signal events (see Fig. 6.1 (b)).

Additionally, a scale factor is applied to the background MC simulations in the CR since our focus is on comparing the shapes of the variables rather than their normalization. The scale factor, determined to be 0.76, is derived from the ratio of data to MC simulations by applying the pre-selections defined in Sec. 6.1 and requiring $|q_{e\mathrm{LJ}}| > 0$.

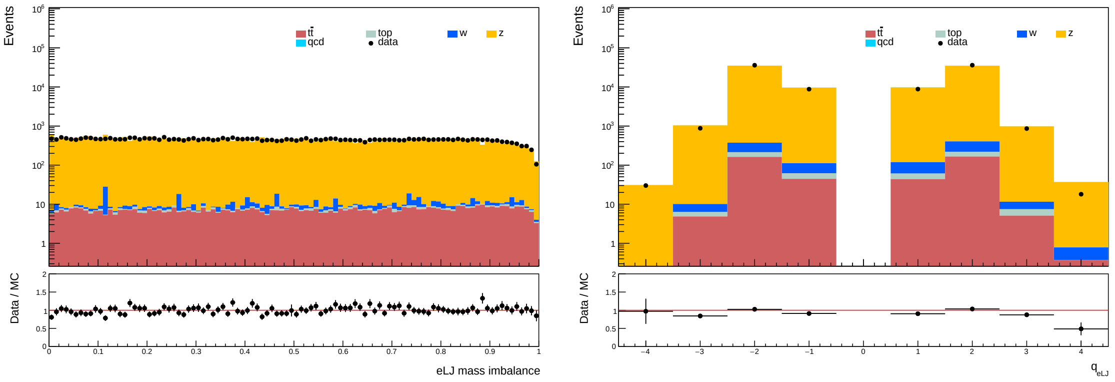
Fig. 6.7 and Fig. 6.8 present the comparisons between data and MC simulations for the event topology variables of the leading and furthest *e*LJs. The reasonable agreement between data and MC simulations validates our understanding of background compositions. Fig. 6.9



(a) (b)

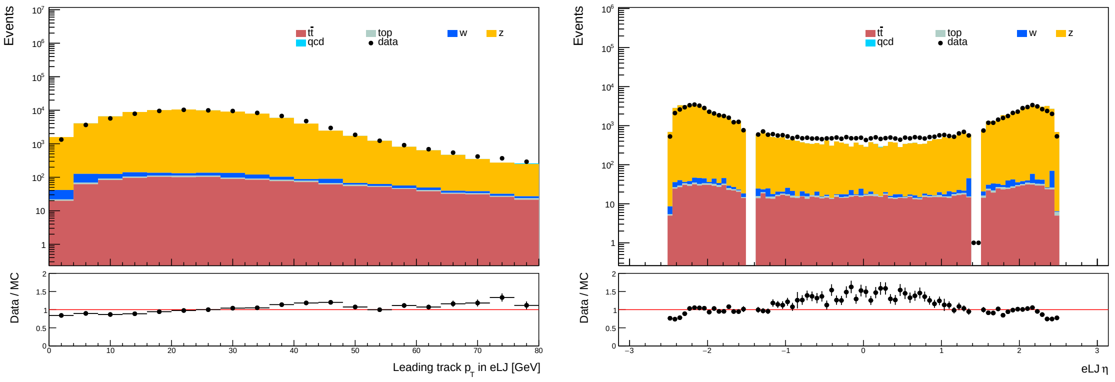
Fig. 6.7 MC simulation to data comparisons for (a) eLJ-eLJ invariant mass, and (b) $\Delta\phi$ between eLJs. Here, MC simulations are scaled to the observed number of events in the CR requiring pre-selections and $|q_{e\mathrm{LJ}}| > 0$.

shows the kinematic variables of leading track p_{T} and γ_{d} η. In this case, the MC simulation shows some deviations from the data. These mismodelings motivate us to use data rather than MC simulations to estimate the number of events entering the SR. In Chapter 7, we will introduce the fully data-driven background estimation approach used in this analysis.



(a) (b)

Fig. 6.8 MC simulation to data comparisons for (a) eLJ-eLJ invariant mass, and (b) $\Delta\phi$ between eLJs. Here, MC simulations are scaled to the observed number of events in the CR requiring pre-selections and $|q_{e\mathrm{LJ}}| > 0$.



(a) (b)

Fig. 6.9 MC simulation to data comparisons for (a) leading track p_{T}, and (b) η of eLJs. Here, MC simulations are scaled to the observed number of events in the CR requiring pre-selections and $|q_{e\mathrm{LJ}}| > 0$.

Chapter 7

Background estimation

Based on what we observed in Sec. 6.3, MC simulation-based methods are unreliable in the search involving customized prompt dark photons in this thesis. Therefore, we employ entirely data-driven approaches for background estimation. Sec. 7.1 introduces the data-driven ABCD method [113]. Sec. 7.2 discusses the selection of variables and definitions of the signal region and control regions in the ABCD plane. Sec. 7.3 validates the background estimation method by defining three additional validation regions and confirming the ABCD assumptions.

7.1 ABCD Method

A data-driven ABCD method is selected for background estimation. In this method, data is divided into four regions (A, B, C, and D) based on two uncorrelated variables, x and y. Region A is the SR enriched by signal events, while regions B, C, and D are the three CRs dominated by background events. A schematic view of ABCD methods is shown in Fig. 7.1.

These variables are chosen to discriminate the signal and background events effectively. Under the assumption that the signal is well-contained in region A, we can estimate the number of backgrounds in region A using the number of observed events in region B, C, and D as:

$$N_A^{\text{bkg}} = \frac{N_B \cdot N_C}{N_D}.$$

Here, N_A^{bkg} represents the estimated number of background events in the SR A, and N_B, N_C, and N_D are the observed numbers of events in the CR B, C, and D. Note that the method remains robust in the presence of multiple background sources in the plane, provided that the events from these different backgrounds result in a uniform distribution across the plane.

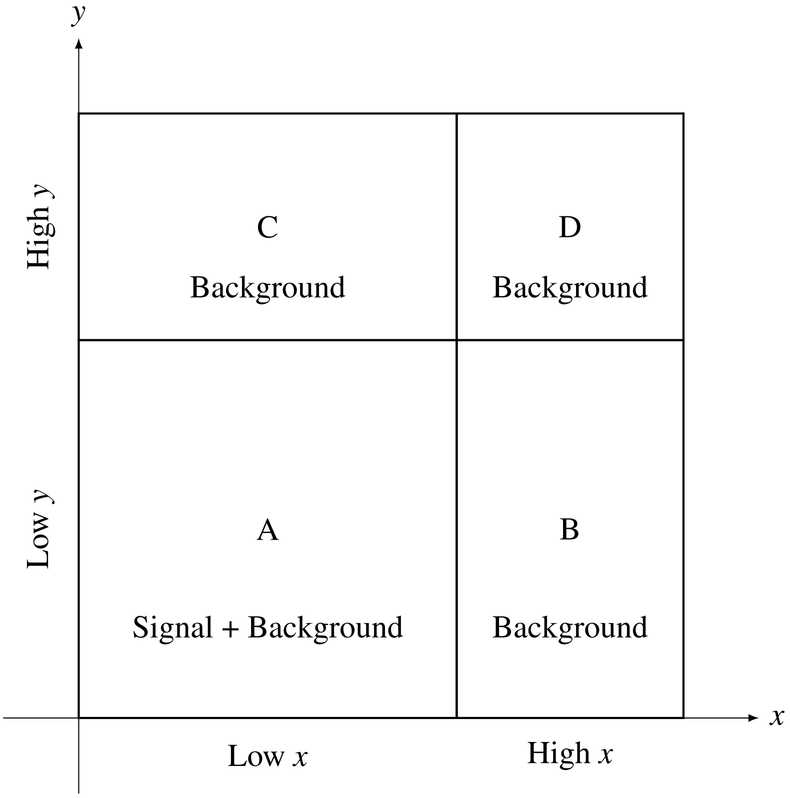


Fig. 7.1 Schematic view of ABCD background estimation method. A is the SR, while B, C, and D are the CRs.

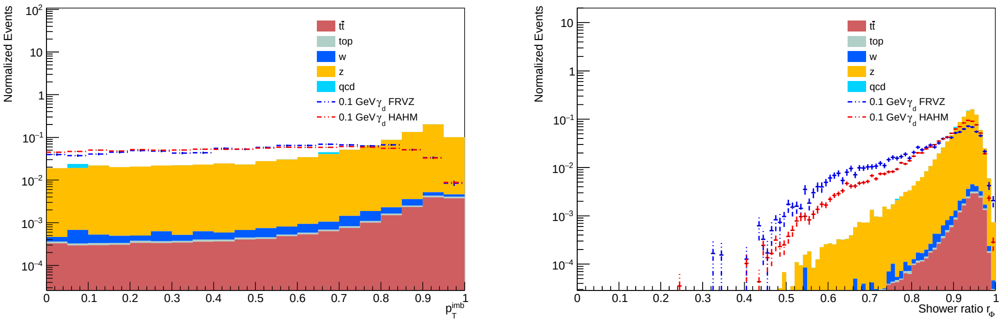
7.2 Selection of Variables

The choice of variables x and y is critical for the success of the ABCD method. These variables are required to be:

1. sensitive in kinematic features of signal events, to ensure that signal events are well contained in region A;
2. background events propagate uniformly in all regions.

As mentioned in Sec. 6.2, the primary source of residual background events comes from the prompt production of electron pairs with random crossing of tracks like Z+jets. For this analysis, we define the variables as follows:

- $\boldsymbol{x}$: leading eLJ $p_{\mathrm{T}}^{\mathrm{imb}}$ represents the p_{T} imbalance of the leading two tracks in the leading eLJ, where $p_{\mathrm{T}}^{\mathrm{imb}} = \frac{|p_{\mathrm{T}}^{\mathrm{lead}} - p_{\mathrm{T}}^{\mathrm{sublead}}|}{p_{\mathrm{T}}^{\mathrm{lead}} + p_{\mathrm{T}}^{\mathrm{sublead}}}$;
- $\boldsymbol{y}$: furthest eLJ shower ratio R_{ϕ} represents the ratio of 3×3 to 3×7 ($\eta \times \phi$) cell energy depositions in the middle layer of the EM calorimeter around the furthest eLJ.



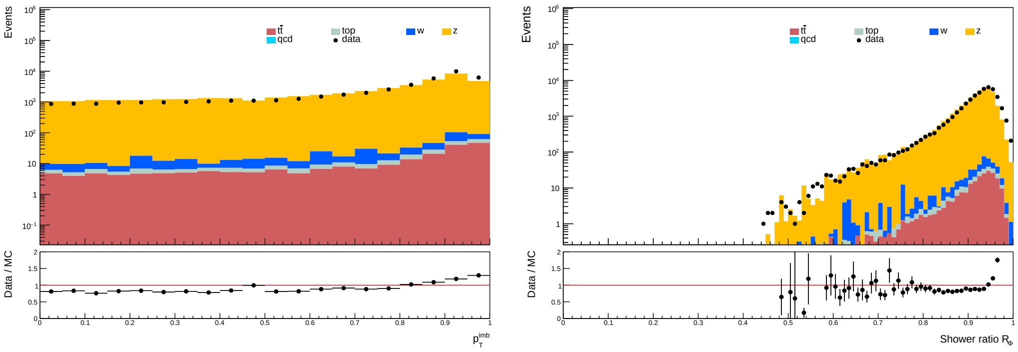
(a) (b)

Fig. 7.2 (a) Track p_{T} imbalance $p_{\mathrm{T}}^{\mathrm{imb}}$ of leading two tracks of leading eLJ, and (b) electron shower ratio R_ϕ of furthest eLJ for simulated SM processes and signals with $m_{\gamma_{\mathrm{d}}} = 0.1$ GeV. Distributions here are shown by requiring pre-selections.

The choice of ABCD variables, $p_{\mathrm{T}}^{\mathrm{imb}}$ and R_ϕ from different eLJs, ensures their lack of correlation. For eLJ candidates from signal events, both the leading and subleading tracks originate from an electron pair decaying from a dark photon, which is expected to be balanced in p_{T}. In contrast, for background events, the leading track originating from a prompt electron usually has a p_{T} much larger than the subleading track from other sources due to random overlapping. Additionally, eLJs from dark photons exhibit a broader spread of energy deposits in the EM calorimeter due to merging two electrons into one cluster. In contrast, the energy is much more centered in the case of prompt electrons. Fig. 7.2 shows both the leading eLJ $p_{\mathrm{T}}^{\mathrm{imb}}$ and the furthest eLJ R_ϕ. These variables are further validated by comparing MC simulations with full Run-2 data in the background dominant CR by requiring that $|q_{e\mathrm{LJ}}| > 0$, as shown in Fig. 7.3. The shape of data is in reasonable agreement with that of Z+jets, which indicates the dominant background to be estimated using data is due to events from two electron-fake eLJs.

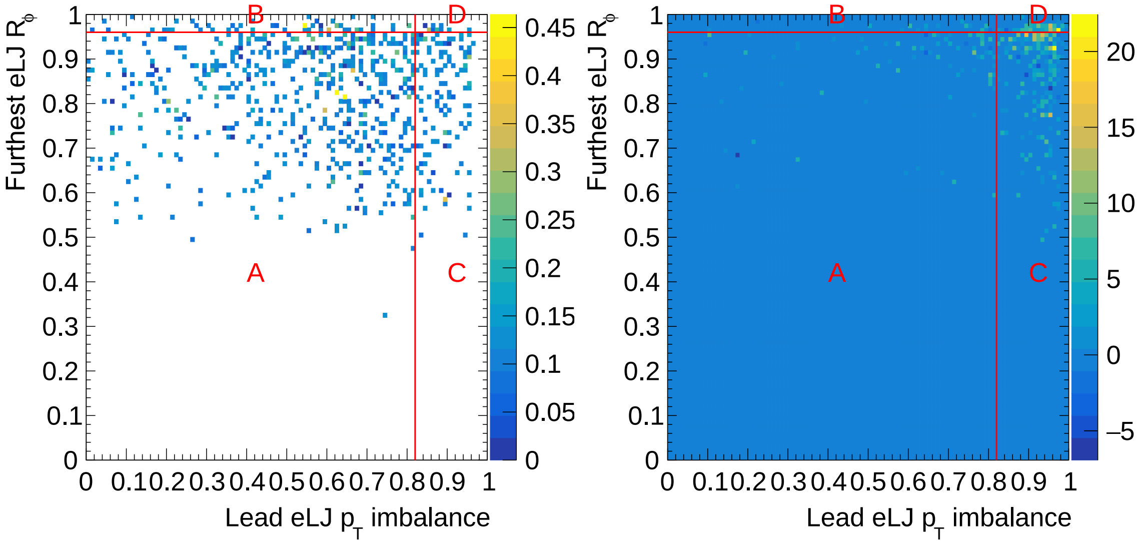
Fig. 7.4 shows 2D plots of the ABCD planes for the reference signal sample (FRVZ with γ_d mass set to 17 MeV) and background MC simulations. In the ABCD planes, 80% of the signal events populate in region A, while 80% of the background events populate regions B, C, and D. Note that background MC simulations are used solely to verify the ABCD method. Their event counts are not included in the final estimation.

The choices of threshold values of leading eLJ $p_{\mathrm{T}}^{\mathrm{imb}}$ and furthest eLJ R_ϕ are determined with the criteria to maximize signal to background sensitivity while keeping signal leakage defined as the number of signal events over background in regions B/C/D to below 10%. Fig. 7.5 shows the signal leakage of 17 MeV γ_{d} events in the scan over values of ABCD variables, from which we see the leakage is mainly affecting region B. Fig. 7.6 shows



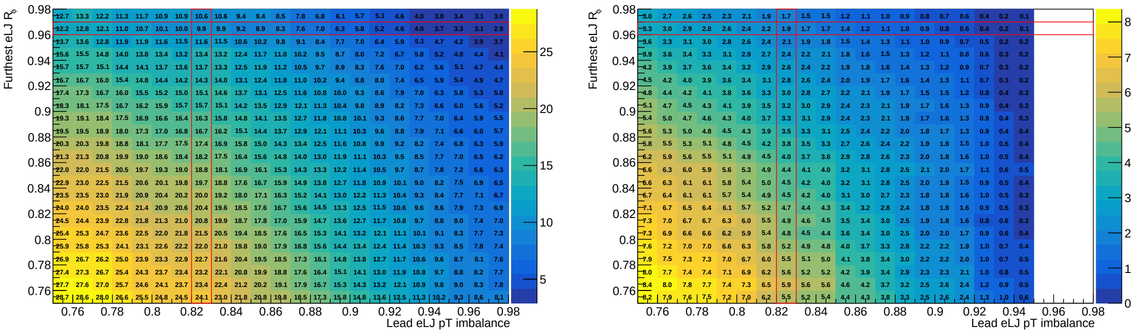
(a) (b)

Fig. 7.3 MC simulation to data comparisons for (a) track p_{T} imbalance $p_{\mathrm{T}}^{\mathrm{imb}}$ of leading two tracks of leading eLJ, and (b) electron shower ratio R_ϕ of furthest eLJ in the CR by requiring pre-selections and $|q_{e\mathrm{LJ}}| > 0$.



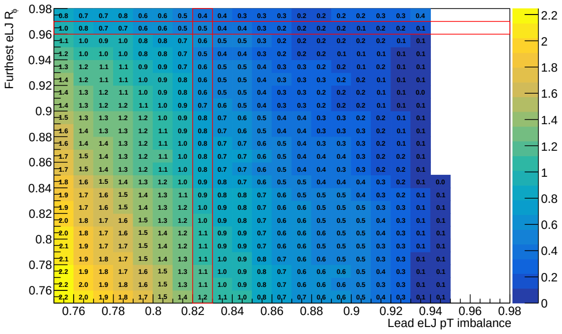
(a) (b)

Fig. 7.4 ABCD plane of leading eLJ $p_{\mathrm{T}}^{\mathrm{imb}}$ and furthest eLJ R_ϕ for (a) a FRVZ benchmark signal sample with $m_{\gamma_d} = 0.1$ GeV, (b) background MC samples. Negative entries corresponds to events with negative weights in MC.



(a) Percentage leakage in CR B

(b) Percentage leakage in CR C



(c) Percentage leakage in CR D

Fig. 7.5 Percentage ratio of the number of 17 MeV FRVZ γ_d normalized to BR($H \to 2\gamma_\mathrm{d} + X$) of 0.5% over data in CR B/C/D.

$S/\sqrt{B_{exp}}$ in final SR A, where we observe peak at $R_\phi = 0.96$ due to shape in Fig. 7.2 (b). Considering all factors, we choose the ABCD definition given in Table 7.1.

Yields in the ABCD are reported in Table 7.2 and Table 7.3 for the FRVZ and the HAHM benchmark models and background MC simulations. Here, the expected N_A is computed from $N_B \times N_C / N_D$.

Table 7.1 Definition of the signal and control regions used in the background estimation.

Requirement / Region	A (SR)	B	C	D
leading eLJ $p_\mathrm{T}^\mathrm{imb}$	< 0.82	< 0.82	> 0.82	> 0.82
furthest eLJ R_ϕ	< 0.96	> 0.96	< 0.96	> 0.96

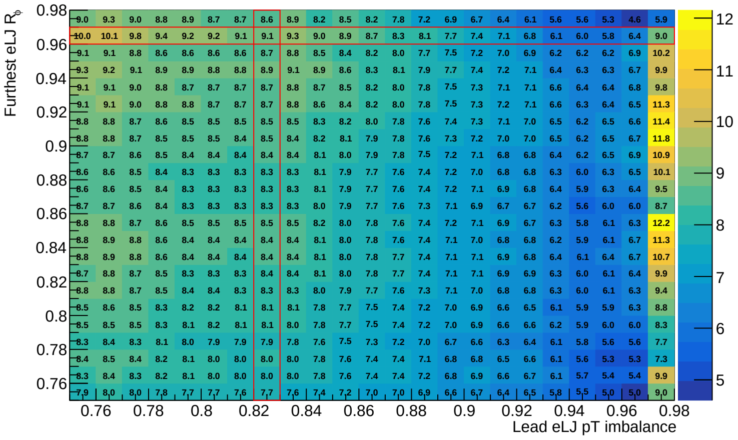


Fig. 7.6 Significance ($S/\sqrt{N_{exp}}$) of the number of events 17 MeV FRVZ $\gamma_{\rm d}$ normalized to BR($H \to 2\gamma_{\rm d} + X$) of 0.5% in SR A over the squared root of the expected number of background in data ($N_{exp} = N_B \times N_C / N_D$).

Table 7.2 ABCD yields in FRVZ signal and background simulations. Signal yields are normalized to a BR($H \to 2\gamma_{\rm d} + X$) of 0.5%. Uncertainties are statistical only, while the uncertainty on the number of expected events in region A is obtained from the propagation of the statistical uncertainty on regions B, C, and D.

Region	FRVZ Signal samples					Background
	$\gamma_{\rm d}$ mass [GeV]	0.017	0.03	0.06	0.1	0.24	
N_A (expected)							190
N_A (SR)		138.65 ± 4.11	101.65 ± 3.56	76.21 ± 3.12	58.60 ± 2.65	15.29 ± 1.35	139 ± 50
N_B		14.06 ± 1.31	11.17 ± 1.23	7.84 ± 0.97	6.03 ± 0.92	2.47 ± 0.55	39 ± 12
N_C		20.22 ± 1.58	16.73 ± 1.46	13.81 ± 1.30	12.72 ± 1.23	1.60 ± 0.43	564 ± 50
N_D		2.58 ± 0.55	1.79 ± 0.54	1.66 ± 0.52	1.20 ± 0.35	0 ± 0	158 ± 25

Table 7.3 ABCD yields in HAHM signal and background simulations The signal assumes a BR($H \to 2\gamma_{\rm d}$) of 0.5%. Uncertainties are statistical only, while the uncertainty on the number of expected events in region A is obtained from the propagation of the statistical uncertainty in regions B, C, and D.

Region	HAHM Signal samples				Background
	$\gamma_{\rm d}$ mass [GeV]	0.017	0.1	0.4	
N_A (expected)					190
N_A (SR)		1150.23 ± 12.15	278.19 ± 5.94	36.94 ± 2.12	139 ± 50
N_B		71.83 ± 3.04	24.87 ± 1.76	5.05 ± 0.76	39 ± 12
N_C		90.38 ± 3.33	48.35 ± 2.59	8.72 ± 1.08	564 ± 50
N_D		5.29 ± 0.79	4.22 ± 0.70	1.76 ± 0.48	158 ± 25

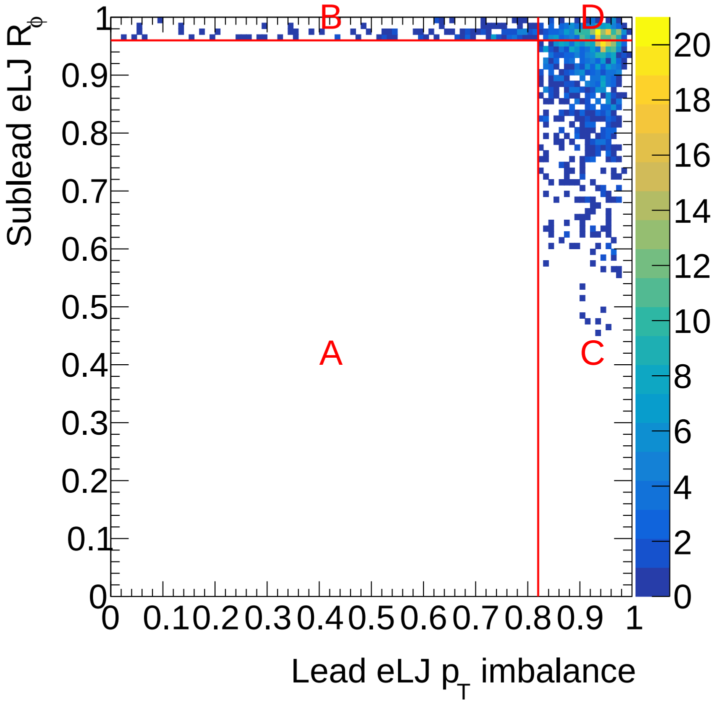


Fig. 7.7 ABCD plane of leading *e*LJ p_T^{imb} and furthest *e*LJ R_ϕ for full Run-2 data. The SR A is blinded in the data at this stage of the background estimation, with the unblinded data to be presented in Sec. 9.1.

7.3 Validation of the Background Estimation Method

We validate the ABCD background estimation method using data in sub-regions of the ABCD plane as shown in Fig. 7.7. Two such regions are constructed from the combination of the B, C, and D regions, namely the BD and the CD sub-planes, which are further divided into four sub-regions: BD1, BD2, BD3, and BD4 for the BD sub-plane and CD1, CD2, CD3, and CD4 for the CD sub-plane, as illustrated in Fig. 7.8. The prediction of the ABCD method is tested in the BD1 and CD1 regions, where the vertical (horizontal) boundary is varied in the test of BD (CD) regions. We fix the other ABCD variable in each case at a nominal value. Specifically, the furthest *e*LJ R_ϕ is set to 0.98 in the BD sub-plane, and the leading *e*LJ p_T^{imb} is fixed at 0.90 in the CD sub-plane. In each combination of leading *e*LJ p_T^{imb} and the furthest *e*LJ R_ϕ cut, expected number of BD1 (CD1) is compared with observed number of BD2×BD3/BD4 (CD2×CD3/CD4).

Fig. 7.9 (a) shows the observed versus expected number of events in BD1 as a function of the p_T^{imb} cut. Similarly, Fig. 7.9 (b) shows the observed versus expected number in CD1,

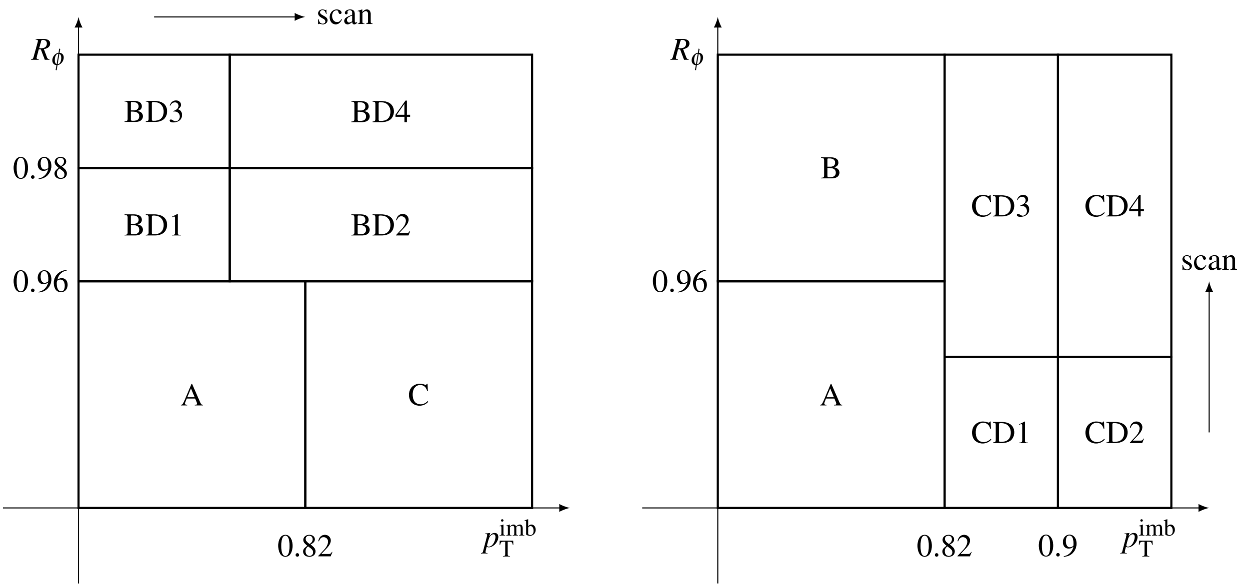
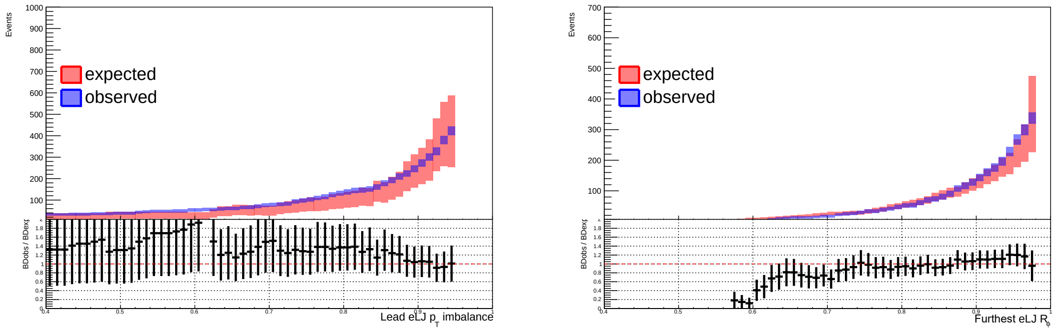


Fig. 7.8 Illustration of sub-region definitions in BD and CD control region.

scanned by the R_ϕ cut. For values of $p_{\mathrm{T}}^{\mathrm{imb}}$ and R_ϕ around the final SR threshold, the number of expected and observed events agrees within 1 σ.

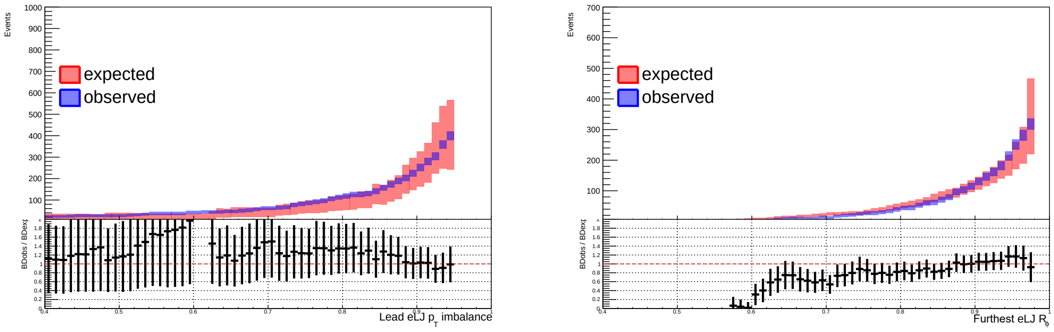
The impact of the signal contamination in regions B, C, and D is evaluated by checking the observed vs expected number of events in BD and CD sub-planes after subtracting the simulated signal contributions. Fig. 7.10 shows the robustness of the ABCD method under the assumption of 17 MeV FRVZ γ_{d} with BR($H \to 2\gamma_{\mathrm{d}} + X$) of 0.5% signal contamination. The agreement is better when the signal contamination is subtracted. Therefore, we check the impact on the expected limit due to the signal contamination using the fitting framework as will be described in Sec. 9.3.1.



(a) BD validation

(b) CD validation

Fig. 7.9 Observed and expected values of the number of events in data in the *e*LJ-*e*LJ channel BD and CD sub-planes as a function of the cut on $p_{\mathrm{T}}^{\mathrm{imb}}$ and R_ϕ, respectively.



(a) BD validation (b) CD validation

Fig. 7.10 Observed and expected values of the number of events in data in the eLJ-eLJ channel BD and CD sub-planes as a function of the cut on $p_{\mathrm{T}}^{\mathrm{imb}}$ and R_{ϕ}, respectively, after subtracting the simulated signal contributions of 17 MeV γ_{d} with BR$(H \to 2\gamma_{\mathrm{d}} + X)$ of 0.5%.

In addition to the BD/CD sub-planes closure checks, to further validate the background estimation method, we define three additional validation regions orthogonal to the ABCD plane used in Fig. 7.7 to confirm the validity of $p_{\mathrm{T}}^{\mathrm{imb}}$ and R_{ϕ} in all phase spaces including region A. Each validation region is independently defined as follows:

- ***VR1***: 80 GeV $< m_{e\mathrm{LJ}-e\mathrm{LJ}} <$ 100 GeV;
- ***VR2***: $|q_{e\mathrm{LJ}}| > 0$;
- ***VR3***: $|\eta| > 1.5$.

Other event selections on SR are applied to these VRs.

As shown in Table 7.4, which summarizes the background compositions of the SR and three VRs from MC simulations, both the SR and VRs are predominantly composed of Z+jets events.

Table 7.4 Composition of background processes from MC simulation in SR and VRs.

Region	$t\bar{t}$	single-top	W+jets	Z+jets	Multijets
SR	0.30%	0.22%	1.98%	97.00%	0.50%
VR1	0.01%	0.00%	0.03%	99.95%	0%
VR2	0.23%	0.20%	0.83%	98.15%	0.58%
VR3	0.10%	0.11%	2.10%	84.85%	12.84%

Table 7.5, Table 7.6, and Table 7.7 present the yields in the ABCD plane for the three validation regions: VR1, VR2, and VR3, respectively. The deviations between expected N_A

Table 7.5 Signal and data yields in VR1 defined by requiring the mass of two *e*LJ to be within the Z-mass window normally vetoed in the definition of the SR. Signal events weighted with a BR=0.5% for FRVZ and 0.5% for HAHM.

Region	FRVZ					HAHM			Run-2 data
γ_d mass [GeV]	0.017	0.03	0.06	0.1	0.24	0.017	0.1	0.4	
N_B	2.43 ± 0.55	1.90 ± 0.46	2.04 ± 0.49	0.64 ± 0.25	0.83 ± 0.32	0 ± 0	0.12 ± 0.12	0 ± 0	318
N_C	2.84 ± 0.63	2.98 ± 0.61	3.39 ± 0.68	1.86 ± 0.47	0.57 ± 0.26	6.59 ± 0.95	4.69 ± 0.74	0.48 ± 0.25	2486
N_D	1.17 ± 0.37	0.38 ± 0.22	0.71 ± 0.29	0.37 ± 0.21	0 ± 0	0 ± 0	0 ± 0	0 ± 0	2249
N_A	29.88 ± 1.92	17.59 ± 1.43	14.49 ± 1.42	9.46 ± 1.20	3.79 ± 0.67	50.62 ± 2.55	26.61 ± 1.76	1.98 ± 0.48	427
N_A (expected)									352 ± 22

Table 7.6 Signal and data yields in VR2 defined by requiring the charge of the *e*LJ to be different from zero. Signal events weighted with a BR=0.5% for FRVZ and 0.5% for HAHM.

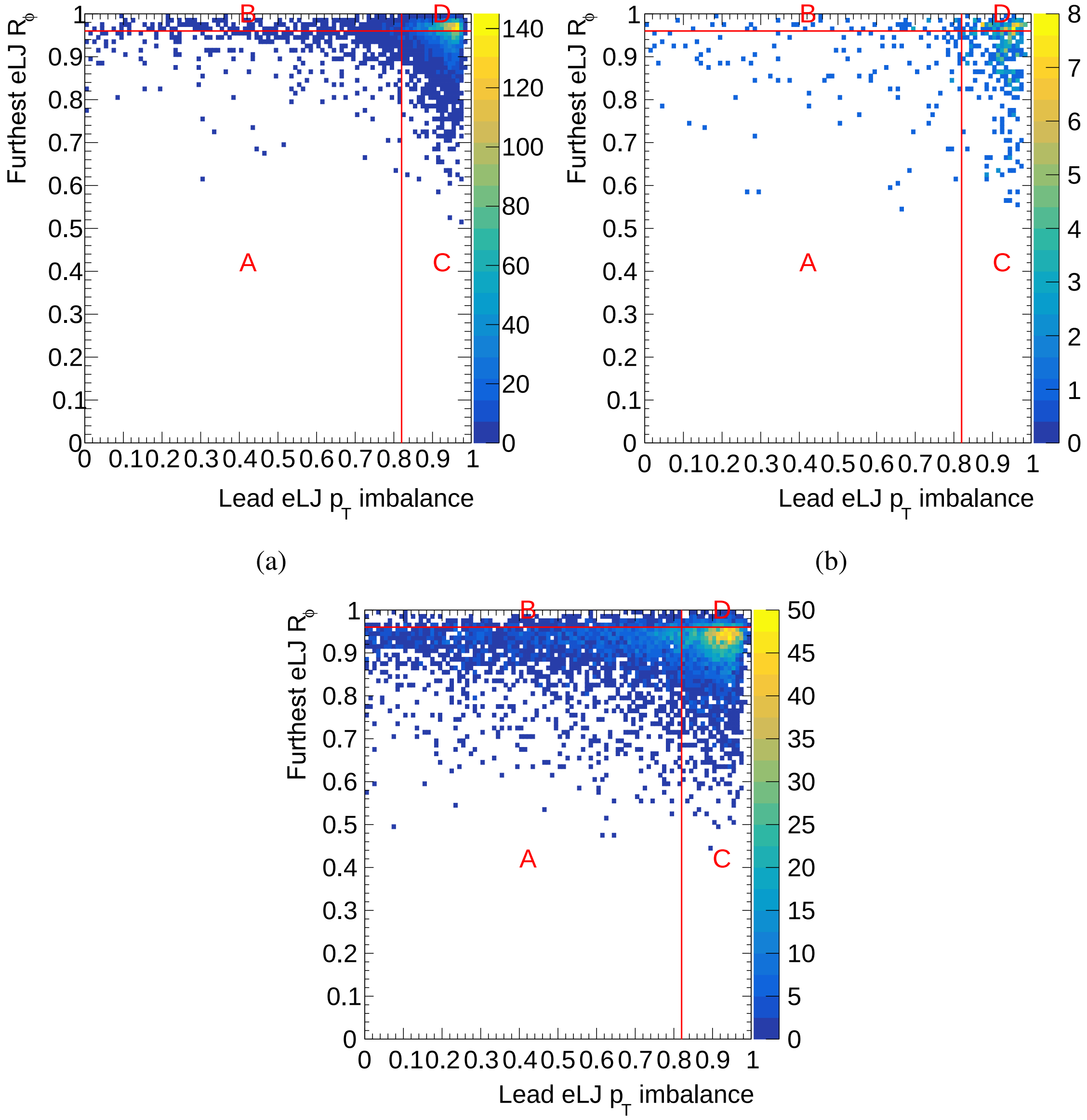
Region	FRVZ					HAHM			Run-2 data
γ_d mass [GeV]	0.017	0.03	0.06	0.1	0.24	0.017	0.1	0.4	
N_B	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	44
N_C	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.23 ± 0.16	0 ± 0	0 ± 0	246
N_D	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	106
N_A	0 ± 0	0.09 ± 0.09	0 ± 0	0 ± 0	0 ± 0	0.64 ± 0.26	0.88 ± 0.33	0 ± 0	101
N_A (expected)									102 ± 19

and observed N_A in the VRs are considered sources of non-closure systematic uncertainties, which will be explained in Sec. 8.3. Fig. 7.11 illustrate the population of events from the VRs entering the ABCD regions, as defined in Table 7.1.

Moreover, to validate the prediction of the ABCD method, we also compare the expected number of events with the observed number of events using the scan method for all 3 VRs. For each VR, we perform a scan of both $p_{\mathrm{T}}^{\mathrm{imb}}$ and R_{ϕ}, fixing the other variable to the threshold value defined for the SR in Table 7.1. Fig. 7.12, Fig. 7.13, and Fig. 7.14 show the observed versus expected number of events in VR1, VR2, and VR3, respectively. The expected values in region A are consistent with the observed numbers.

Table 7.7 Signal and data yields in VR3 defined by reversing the selection on the η of the *e*LJ. Signal events weighted with a BR=0.5% for FRVZ and 0.5% for HAHM.

Region	FRVZ					HAHM			Run-2 data
γ_d mass [GeV]	0.017	0.03	0.06	0.1	0.24	0.017	0.1	0.4	
N_B	2.78 ± 0.56	2.50 ± 0.53	1.56 ± 0.42	1.83 ± 0.46	0.08 ± 0.08	16.26 ± 1.40	11.82 ± 1.35	1.66 ± 0.44	353
N_C	13.41 ± 1.33	14.38 ± 1.36	9.96 ± 1.08	11.05 ± 1.12	2.24 ± 0.51	58.56 ± 2.75	48.44 ± 2.47	6.34 ± 0.99	3599
N_D	0 ± 0	0.31 ± 0.18	0.14 ± 0.14	0.09 ± 0.09	0.06 ± 0.06	2.12 ± 0.64	1.63 ± 0.44	0 ± 0	390
N_A	83.25 ± 3.14	69.68 ± 2.90	51.11 ± 2.45	45.57 ± 2.37	9.35 ± 1.03	623.15 ± 8.80	301.41 ± 6.13	31.48 ± 1.94	2948
N_A (expected)									3257 ± 245

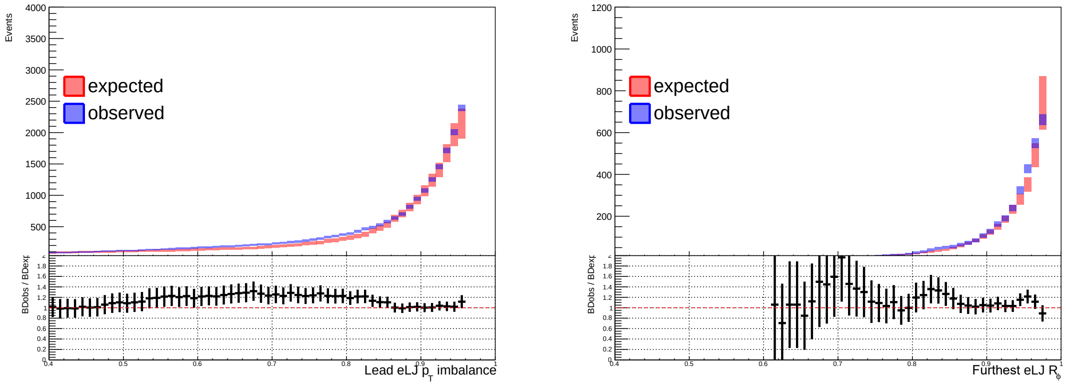


(a)

(b)

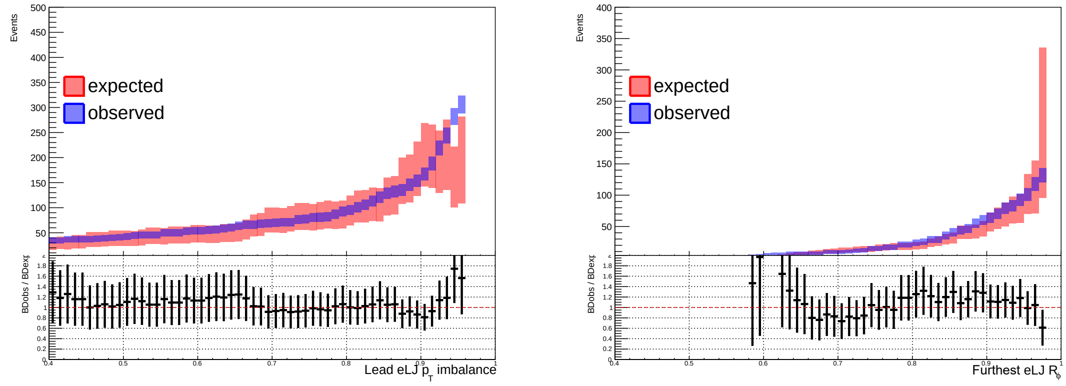
(c)

Fig. 7.11 ABCD plane of leading eLJ $p_{\mathrm{T}}^{\mathrm{imb}}$ and furthest eLJ R_ϕ for (a) VR1 on-shell Z boson mass window, (b) VR2 $|q_{e\mathrm{LJ}}| > 0$, and (c) VR3 $|\eta| > 1.5$ end-cap.



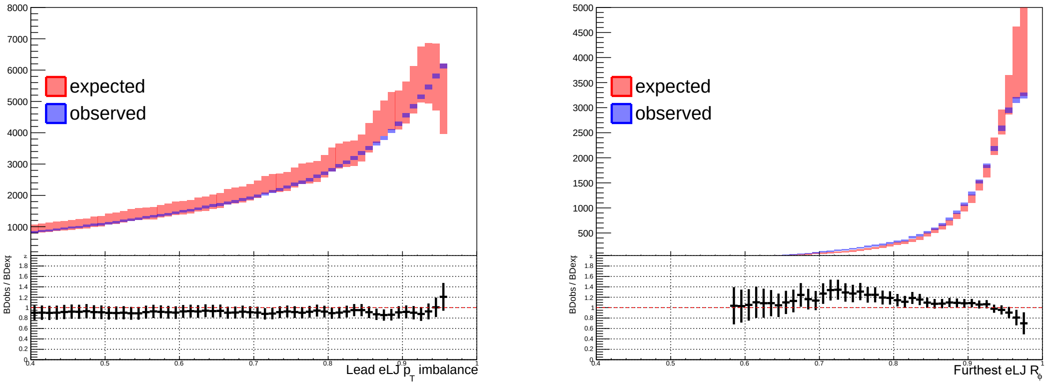
(a) $p_{\mathrm{T}}^{\mathrm{imb}}$ scan (b) R_ϕ scan

Fig. 7.12 Observed vs. expected values of the number of events in data in the *e*LJ-*e*LJ channel on-shell Z boson mass window validation region as a function of the cut on $p_{\mathrm{T}}^{\mathrm{imb}}$ and R_ϕ, respectively.



(a) $p_{\mathrm{T}}^{\mathrm{imb}}$ scan (b) R_ϕ scan

Fig. 7.13 Observed vs. expected values of the number of events in data in the *e*LJ-*e*LJ channel $|q_{e\mathrm{LJ}}| > 0$ validation region as a function of the cut on $p_{\mathrm{T}}^{\mathrm{imb}}$ and R_ϕ, respectively.



(a) $p_{\mathrm{T}}^{\mathrm{imb}}$ scan

(b) R_ϕ scan

Fig. 7.14 Observed vs. expected values of the number of events in data in the *e*LJ-*e*LJ channel $|\eta| > 1.5$ end-cap validation region as a function of the cut on $p_{\mathrm{T}}^{\mathrm{imb}}$ and R_ϕ, respectively.

Chapter 8

Systematic Uncertainties

In this chapter, we evaluate the systematic uncertainties. We perform an event counting experiment on the final yields in the signal region to derive the analysis results. The number of signal events is obtained from MC simulations, while the number of background events is estimated using the ABCD data-driven method explained in Chapter 7. Systematic effects from experimental and theoretical sources, which may affect the expected number of signal events, are considered. Additionally, we assign a non-closure systematic uncertainty to the background estimation method. The likelihood function used to fit the data includes all systematic uncertainties as nuisance parameters. These parameters are then profiled in the final fit, as will be explained in Sec. 9.2.

8.1 Experimental Uncertainties

The following experimental sources of uncertainties are considered and evaluated as impacts on signal yields in the MC simulation for γ_{d} mass below 0.24 GeV in the FRVZ model and 0.4 GeV for the HAHM model.

8.1.1 Standard Objects

Per-object scale factors are applied to correct the modeling of electrons in terms of reconstruction efficiency, isolation, and identification by comparing p_{T} and η distribution of signal and e from Z events [114]. The uncertainties on these scale factors are treated as systematic uncertainties to the final signal selection efficiencies. The most significant impact is $\pm 1.6\%$ for $m_{\gamma_{\mathrm{d}}} = 0.06$ GeV hypothesis of the FRVZ model.

8.1.2 Triggers

The efficiency of single and di-electron triggers listed in Table 5.7 are calibrated by the $Z \to ee$ events [114]. The systematic impact is evaluated to be negligible for the HAHM model and ±0.2% for the FRVZ model.

8.1.3 Luminosity

The systematic uncertainty on the integrated luminosity of Run-2 data from 2015 to 2018 is 0.8% [104]. The uncertainty is calculated based on the luminosity scale calibrations using x-y beam separation scans of the LUCID2 detector [115].

8.1.4 Pile-up Reweighting

To account for the multiple interactions per bunch crossing mentioned in Sec. 4.1.4, the pile-up distribution in each data-taking period is modeled in the MC simulations and corrected by data in a pile-up reweighting procedure. The systematic uncertainties associated with pile-up reweighting are evaluated by shifting the default data scale factor 1/1.03 to up-variation 1/0.99 and down variation 1/1.07. The uncertainties propagate through the event selections and result in, at most, a ±5.5% effect for $m_{\gamma_\mathrm{d}} = 0.24$ GeV hypothesis of the FRVZ model.

8.1.5 $e\gamma$ energy scale and resolution

The systematic uncertainties concerning the electron kinematic uncertainties arising from the $e\gamma$ energy scale and resolution are also evaluated. The most significant impact is ±0.95% for $m_{\gamma_\mathrm{d}} = 0.1$ GeV hypothesis of the FRVZ model.

8.1.6 Jet Energy Scale and Jet Energy Resolution

Systematic uncertainties from JES and JER (see Sec. 4.3.4) are measured by varying a reduced set of 3 nuisance parameters for JES and 6 nuisance parameters for JER associated with the calibration procedures [116]. The most significant impact is ±3.1% for $m_{\gamma_\mathrm{d}} = 0.1$ GeV hypothesis of the FRVZ model.

8.1.7 MC Statistics

The uncertainties in the final signal yields due to MC statistics are also considered. These uncertainties become dominant compared to other systematic effects in the high mass region, where acceptance is lower. The most significant uncertainty observed is ±9.3% for $m_{\gamma_\mathrm{d}} = 0.24$ GeV hypothesis of the FRVZ model.

8.2 Theoretical Uncertainties

The theoretical uncertainties for the FRVZ and HAHM signal processes are considered. A general n-th order perturbation theory computation of inclusive cross section in pp collisions can be described by:

$$\sigma^{(n)} = PDF(x_1, \mu_F, Q) \otimes PDF(x_2, \mu_F, Q) \otimes \hat{\sigma}^{(n)}(x_1, x_2, \mu_R, Q), \tag{8.1}$$

where

$$\hat{\sigma} = \alpha_s \hat{\sigma}^{(0)} + \alpha_s^2 \hat{\sigma}^{(1)} + ... + \alpha_s^n \hat{\sigma}^{(n)} + \mathscr{O}(\alpha_s^{n+1}). \tag{8.2}$$

Here μ_F and μ_R are the factorization and renormalization scales. The parton distribution functions at energy scale Q can be estimated by the DGLAP equations [117] [118] [119]:

$$dPDF(x, Q^2)/d\log Q^2 = P(\alpha_s)(x, z) PDF(x/z, Q^2), \tag{8.3}$$

where the splitting function $P(\alpha_s)$ can be perturbative expanded by

$$P(\alpha_s) = \alpha_s p^{(LO)} + \alpha_s^2 p^{(NLO)} + ... \tag{8.4}$$

We consider two sources of theoretical uncertainties from the perturbative QCD cross-section computation.

The remaining uncertainties from higher orders $\mathscr{O}(\alpha_s^{n+1})$ are estimated by performing scale variations on renormalization and factorization. A 7-point scale variation of $\{\mu_F, \mu_R\} \times \{0.5, 0.5\}, \{0.5, 1\}, \{1, 0.5\}, \{1, 1\}, \{1, 2\}, \{2, 1\}, \{2, 2\}$ is considered, and the maximum among these 7 variations is used to evaluate scale uncertainties.

PDFs are extracted from experimental data using global fitting methods. The uncertainties arising from experimental uncertainties and the choice of functional forms in the PDF fitting procedures are incorporated by PDF error ensembles. Each ensemble features slight variations in parameterization. For event generation in both FRVZ and HAHM signal MC simulations, the NNPDF2.3 [120] set is used, where an ensemble of 101 internal PDF variations is included. The sample standard deviation from the central value is assigned as the PDF uncertainty to the final signal yields.

Table 8.1 Summary table of the systematic uncertainties on the final yield of FRVZ signal MC events.

m_{γ_d} (GeV)	ID (%)	Iso (%)	Reco (%)	Trigger (%)	Lumi (%)	PRW (%)	$e\gamma$ reso (%)	$e\gamma$ scale (%)	JER (%)	JES (%)	MC stat (%)	PDF (%)	Scale (%)	Total (%)
0.017	1.5	0.16	0.84	0.20	0.8	0.87	0.12	0.21	1.2	2.5	3.1	3.4	5.5	7.9
0.03	1.6	0.19	0.88	0.23	0.8	2.6	0.33	0.76	0.41	2.2	3.6	3.4	5.7	8.5
0.06	1.6	0.18	0.85	0.21	0.8	2.9	0.78	0.68	1.5	2.2	4.2	3.4	5.6	9.0
0.10	1.5	0.15	0.79	0.20	0.8	2.4	0.95	0.67	0.54	3.1	4.7	3.4	5.6	9.2
0.24	1.3	0.13	0.72	0.13	0.8	5.5	0.90	0.15	1.5	2.6	9.3	3.4	5.5	13

Table 8.2 Summary table of the systematic uncertainties on the final yield of HAHM signal MC events.

m_{γ_d} (GeV)	ID (%)	Iso (%)	Reco (%)	Trigger (%)	Lumi (%)	PRW (%)	$e\gamma$ reso (%)	$e\gamma$ scale (%)	JER (%)	JES (%)	MC stat (%)	PDF (%)	Scale (%)	Total (%)
0.017	0.43	0.031	0.85	0.019	0.8	1.2	0.12	0.056	0.33	2.2	1.1	3.4	5.7	7.3
0.1	0.48	0.042	0.80	0.026	0.8	0.51	0.091	0.22	0.69	2.3	2.2	3.4	5.9	7.7
0.4	0.42	0.029	0.85	0.019	0.8	2.5	0.37	0.0012	1.4	2.5	5.8	3.4	5.6	9.6

Both scale and PDF uncertainties are implemented by a reweighting method, which approximates the probability of a sample produced with another PDF:

$$w = \frac{PDF_{new}(x_1, \mu_F, Q) \otimes PDF_{new}(x_2, \mu_F, Q)}{PDF_{old}(x_1, \mu_F, Q) \otimes PDF_{old}(x_2, \mu_F, Q)}. \tag{8.5}$$

Fig. 8.1 and Table 8.1 summarize the experimental and theoretical systematic uncertainties for different γ_d mass hypotheses in the FRVZ benchmark model. Similarly, Fig. 8.2 and Table 8.2 summarize the uncertainties for the HAHM benchmark models.

8.3 Non-closure Uncertainties

In this analysis, we have adapted the entirely data-driven ABCD method for background estimation as explained in Chapter 7. Consequently, the expected background event yield uncertainty in SR A is affected by statistical uncertainties propagated from CRs B, C, and D, rather than the theoretical and experimental effects discussed for the signal events. We have included non-closure uncertainties to account for the difference between the predicted and observed events in the VRs as an additional systematic uncertainty due to the adaption of the ABCD method.

To evaluate the background estimation method, three VRs with background compositions similar to the SR have been constructed: VR1, which is within the Z mass window; VR2, where the eLJ charge requirement is inverted; and VR3, in the end-cap region as explained in Sec. 7.3. For each VR, the non-closure uncertainty is determined by the maximum observed

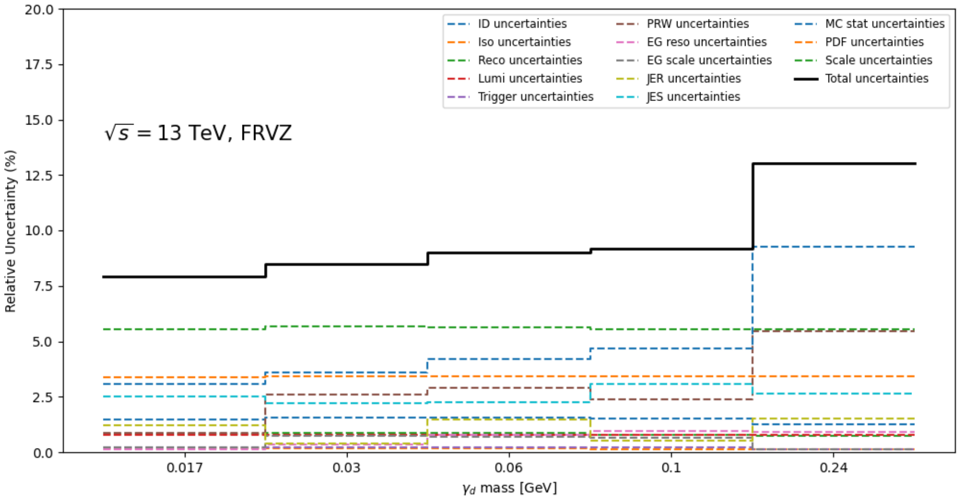


Fig. 8.1 Contributions from the different systematic sources of uncertainties to the signal yields for the FRVZ benchmark model.

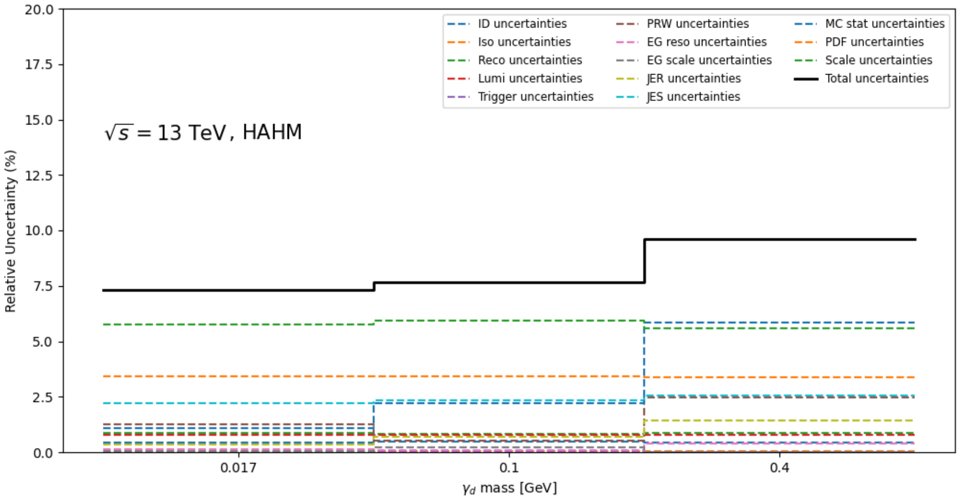
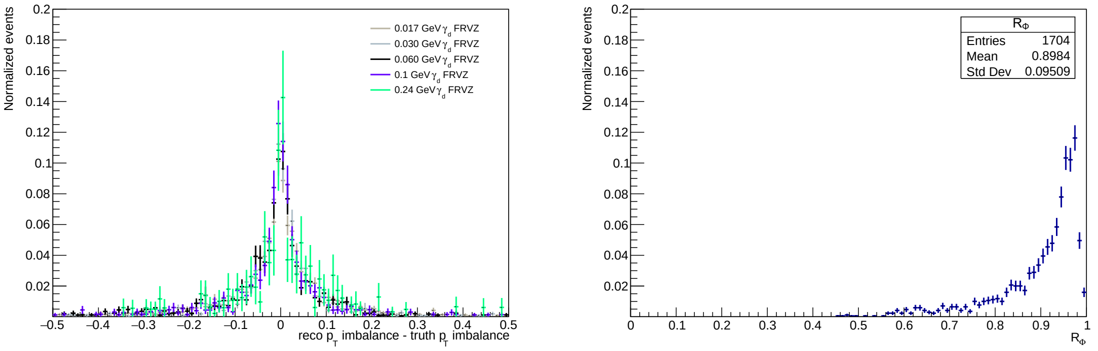


Fig. 8.2 Contributions from the different systematic sources of uncertainties to the signal yields for the HAHM benchmark model.

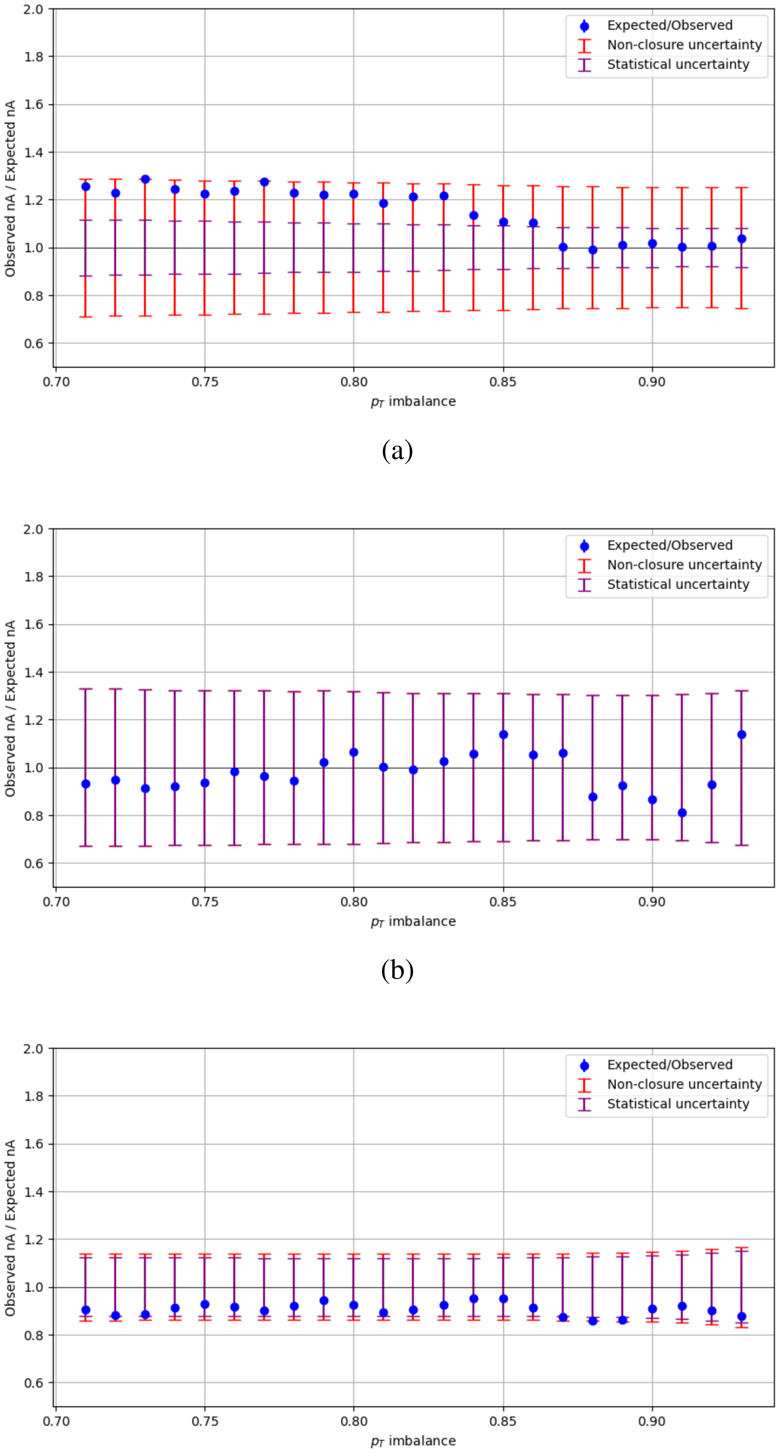


(a) (b)

Fig. 8.3 (a) Distribution for measured $p_{\mathrm{T}}^{\mathrm{imb}}$ - truth $p_{\mathrm{T}}^{\mathrm{imb}}$ for events entering the ABCD plane for the FRVZ benchmark model. The mean σ is 0.11. (b) Distribution for R_{ϕ} entering the ABCD plane for data. The standard deviation is 0.1.

deviation from the expected event yield, calculated as $N_A = N_B \times N_C/N_D$, considering the statistical uncertainties. To account for the resolution of the measured $p_{\mathrm{T}}^{\mathrm{imb}}$ shown in Fig. 8.3 (a), which has an average $\sigma = 0.11$, we evaluate within the range $0.71 < p_{\mathrm{T}}^{\mathrm{imb}} < 0.93$, while fixing the showering ratio R_{ϕ} at the nominal threshold of 0.96. Similarly, we evaluate the variation of R_{ϕ} within the range $0.86 < R_{\phi} < 0.99$ while fixing the $p_{\mathrm{T}}^{\mathrm{imb}}$ at the nominal threshold of 0.82. The range is selected based on the standard deviation of 0.1 in the actual data distribution as shown in Fig. 8.3 (b). Fig. 8.4 and Fig. 8.5 show the non-closure systematic uncertainties evaluated through scans of $p_{\mathrm{T}}^{\mathrm{imb}}$ in VR1, VR2, and VR3. In addition, we perform the closure checks on BD and CD sub-planes by scanning over $0.71 < p_{\mathrm{T}}^{\mathrm{imb}} < 0.93$ and $0.86 < R_{\phi} < 0.99$, respectively, as shown in Fig. 8.6.

The total systematic uncertainty is calculated as the quadratic sum of non-closure systematic uncertainties from maximum deviations observed in each region to 19%. After implementing the non-closure systematic uncertainties, the expected and observed events in N_A are consistent within a 1 σ.

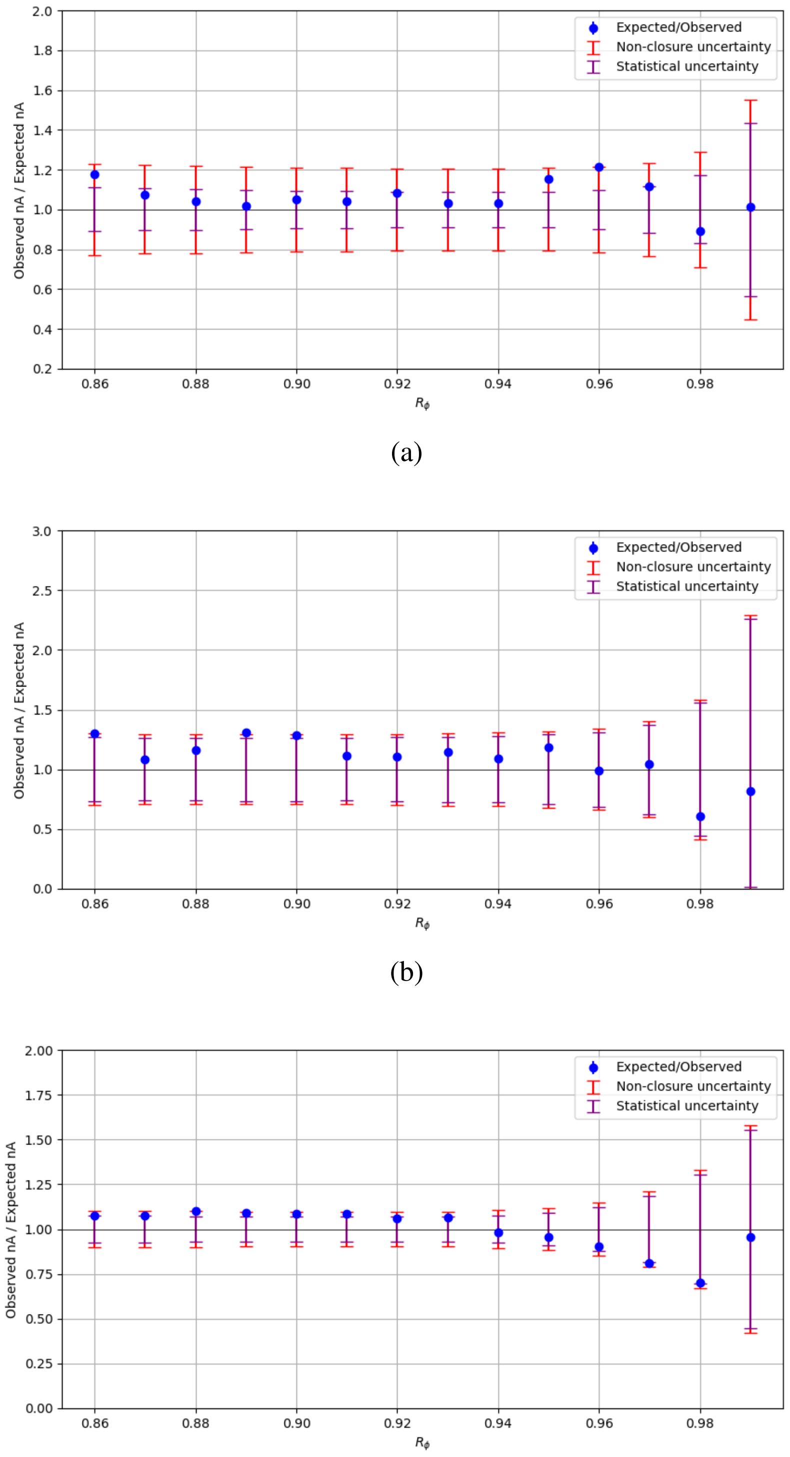


(a)

(b)

(c)

Fig. 8.4 Ratio of observed N_A / expected N_A at different $p_{\mathrm{T}}^{\mathrm{imb}}$ thresholds around nominal cut of 0.82 for (a) VR1 within the Z mass window, (b) VR2 where the eLJ charge requirement is inverted, and (c) VR3 in the end-cap region. Statistical uncertainties propagated from regions B, C, and D are shown in purple. Non-closure uncertainties are shown in red.

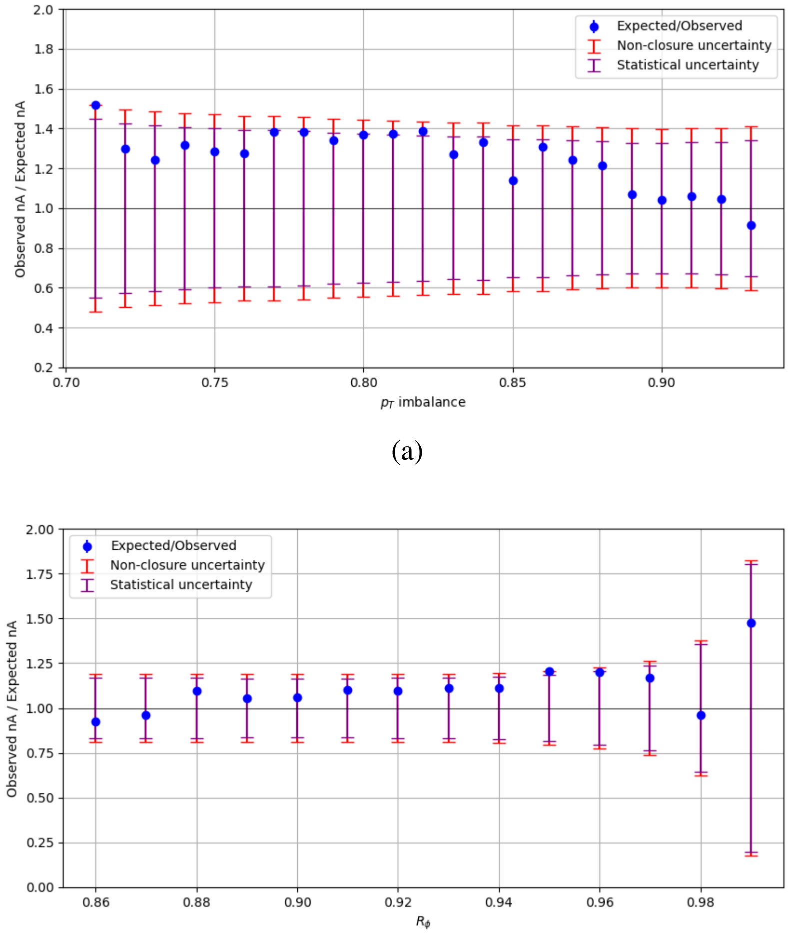


(a)

(b)

(c)

Fig. 8.5 Ratio of observed N_A / expected N_A at different R_ϕ thresholds around nominal cut of 0.96 for (a) VR1 within the Z mass window, (b) VR2 where the eLJ charge requirement is inverted, and (c) VR3 in the end-cap region. Statistical uncertainties propagated from regions B, C, and D are shown in purple. Non-closure uncertainties are shown in red.



(a)

(b)

Fig. 8.6 Ratio of observed N_A / expected N_A for (a) BD sub-plane at different $p_{\mathrm{T}}^{\mathrm{imb}}$ thresholds around the nominal cut of 0.82 (b) CD sub-plane at different R_ϕ thresholds around the nominal cut of 0.96. Statistical uncertainties propagated from regions B, C, and D are shown in purple. Non-closure uncertainties are shown in red.

Chapter 9

Results and Discussions

9.1 Results

Fig. 9.1 shows the final event distribution of the full Run-2 data in the ABCD plane. The observed number of events and the expected number of events in SR A is summarized in Table 9.1. Here, the expected number of events in region A is estimated by $N_A = N_B \times N_C / N_D$, assuming that the ABCD plane is populated by background events only. Overall, the observed 351 events in A are consistent with the expected 303 ± 33 (stat.) ± 58 (syst.) with no significant excess observed. Statistical uncertainty of the expected number in A is propagated from the standard error in regions B, C, and D and dominated by the number in B. Systematic uncertainty of 19% due to non-closure obtained in VRs is described in Sec. 8.3.

Table 9.1 Summary of expected and observed events entering the ABCD plane.

Region	B	C	D	Expected A	Observed A
Number of events	125	862	356	303 ± 33 (stat.) ± 58 (syst.)	351

As an additional cross-check, Fig. 9.2 shows the eLJ invariant mass distribution in the final SR, compared with the prediction of MC simulations SM background processes. No unexpected behavior is observed.

9.2 Likelihood fit

Since the unblinded data does not show any significant excess compared to the estimated background obtained using the background-only ABCD method, we will use these results to establish exclusion upper limits for the production cross-section times the Higgs branching fraction $\sigma \times B$ for γ_{d} production in both the FRVZ and HAHM processes. We apply a

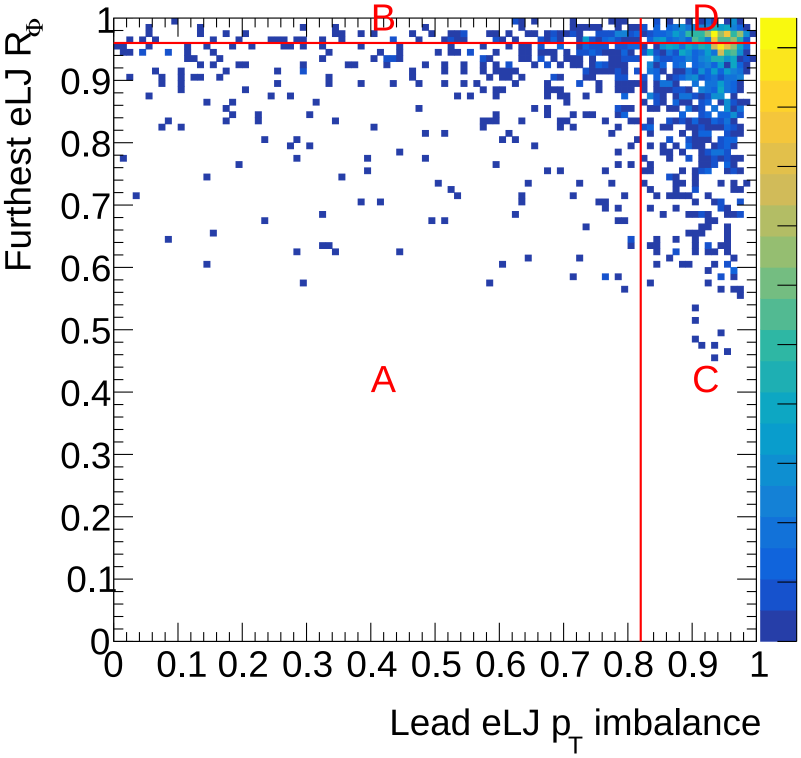
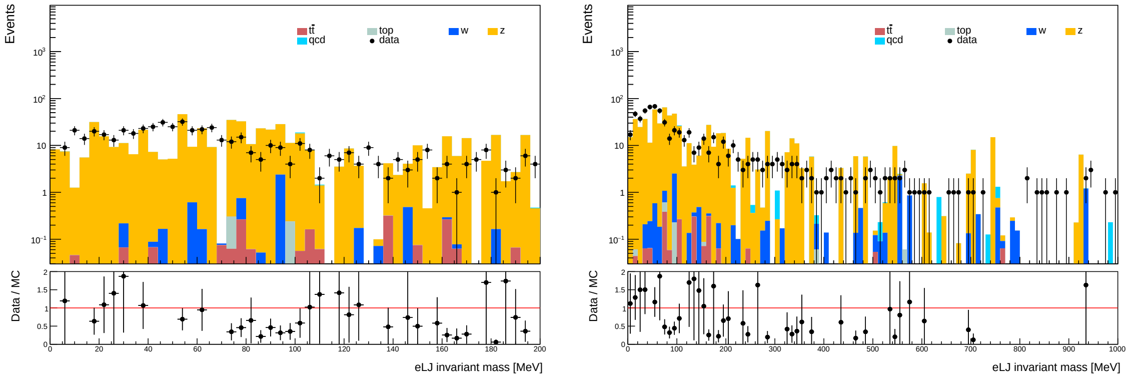


Fig. 9.1 Event distribution of the full Run-2 data in the ABCD plane.



(a)

(b)

Fig. 9.2 Full Run-2 data to MC simulation comparisons for *e*LJ invariant mass in the range (a) [0, 200 MeV], and (b) [0, 1000 MeV] in the final SR A. The MC simulations are scaled to the observed number of data for shape comparison. No unexpected behavior is observed.

likelihood-based modified ABCD method, which simultaneously fits signal and background yields in the four regions. The likelihood function that describes the expected number of signal and background events is defined by the product of Poisson functions, taking the form:

$$\mathscr{L}(n_A, n_B, n_C, n_D | s, b, \tau_B, \tau_C) = \prod_{i=A,B,C,D} \frac{e^{-N_i} N_i^{n_i}}{n_i!}, \tag{9.1}$$

where n_A, n_B, n_C, and n_D are the observed number of events in each region in data. N_i are the expected number in each region, which are linear combinations of signal and background events defined as follows:

$$\begin{aligned} N_A &= s + b \\ N_B &= s\varepsilon_B + b\tau_B \\ N_C &= s\varepsilon_C + b\tau_C \\ N_D &= s\varepsilon_D + b\tau_C\tau_B \end{aligned}$$

Here, s and b correspond to the number of signal and background events in region A. The signal contamination in regions B, C, and D are described by ε_i, obtained from signal MC simulations. τ_B and τ_C are the nuisance parameters that define the ABCD ansatz for background events.

All the parameters s, b, τ_B, τ_C are allowed to float to fit the observed data in four regions. The systematic uncertainties discussed in Chapter 8 are included in the fit as follows: experimental and theoretical effects on the signal yields are included in the nuisance parameter s by multiplying the fit likelihood with Gaussian probability density functions (PDFs), where the variance of these PDFs corresponds to 68% of the systematic uncertainties. Similarly, the non-closure effects on the observed number of A is similarly implemented to the nuisance parameter b by multiplying the fit likelihood with Gaussian PDFs.

The upper limit on signal strength μ defined as the ratio of fitted s to input s with BR($H \to 2\gamma_{\mathrm{d}} + X$)=0.5% is obtained with the CL$_\mathrm{S}$ method [121] by performing a global simultaneous fit based on the profile likelihood method [122] to normalize the observed data in all the regions to the signal expectation. The presence of any excess generated by the signal is then evaluated by comparing the estimated number to the observed number of events in SR A.

Possible signal contamination in the CRs is considered by performing the simultaneous ABCD fit to signal and background. The impact of signal contamination in this analysis will be investigated in Sec. 9.3.

9.3 Exclusion Limits

Using the fitting procedures described in Section 9.2, we obtain $\mathrm{CL_S}$ as functions of the signal strength for different γ_d masses in both the FRVZ and HAHM models. The results are shown in Figures 9.3 and 9.4, respectively.

Figure 9.5 presents the upper limit on BR($H \to 2\gamma_\mathrm{d} + X$) at the 95% CL for the FRVZ model, while Figure 9.6 presents the upper limit on BR($H \to 2\gamma_\mathrm{d}$) for the HAHM model. It is observed that the observed limit is within 1 σ of the expected limit, and the exclusion power is stronger for lower mass γ_d in both benchmark models. This is expected since this analysis is optimized for collimated eLJ signatures targeting the low mass regime.

9.3.1 Impact of signal contamination

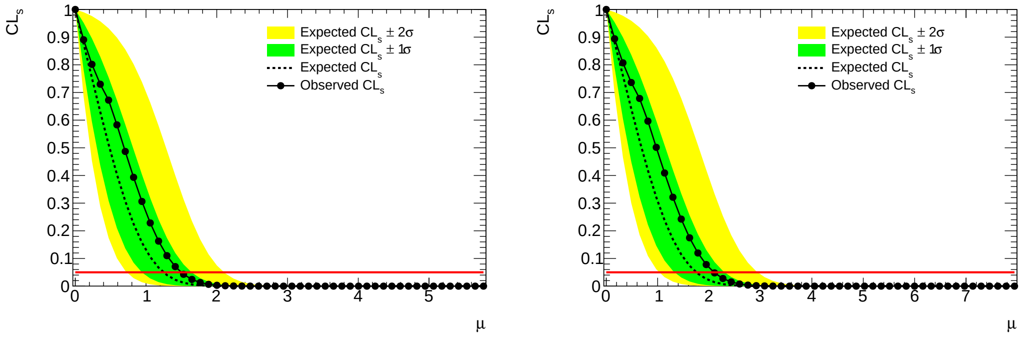
As we saw from Sec. 7.3, the ABCD agreement is better when subtracting the signal contamination. To study the impact of signal contamination in regions B, C, and D on the expected limits, simultaneous fitting to signal and background numbers in the ABCD plane is performed by neglecting the signal contribution in these regions. The result of the fit is shown in Fig. 9.7 for the FRVZ model and Fig. 9.8 for the HAHM model labeled as *no leakage*. In comparison, the actual limits shown in Fig. 9.5 and Fig. 9.6 are labeled as *with leakage*. The limit fittings converge for both scenarios, and it is observed that the exclusion limit, when subtracting the signal contamination, is approximately 1 σ more powerful than the actual case with leakage. For future developments, the leakage could be optimized to improve the search power.

9.3.2 Limits from Muon and Mixed Channel

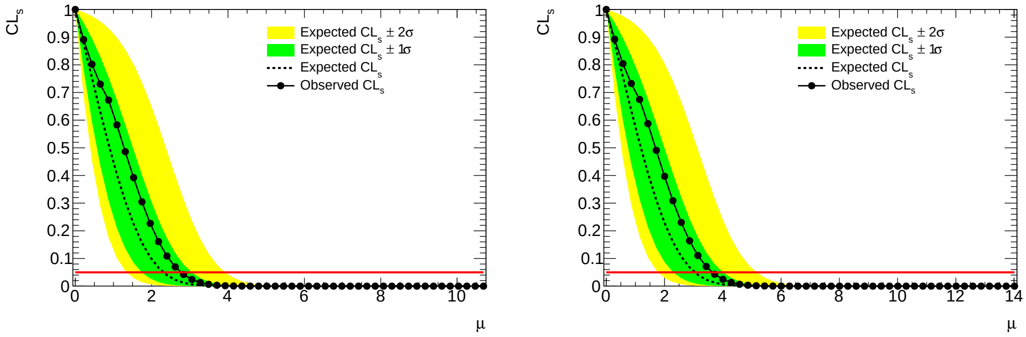
In addition to the electron channel discussed in this thesis, the muon channel (μLJ-μLJ) and mixed channel (eLJ-μLJ) are exploited to provide exclusion limits in the higher mass range of $[0.24, 10]$ GeV, as shown in Fig. 9.9 for both the FRVZ and HAHM models. Due to the excellent mass resolution of μLJ compared with eLJ, the muon channel and mixed channel perform a bump hunting on the invariant mass of μLJ, leading to a more competitive exclusion limit in their phase space.

9.3.3 Higgs portal interpretation

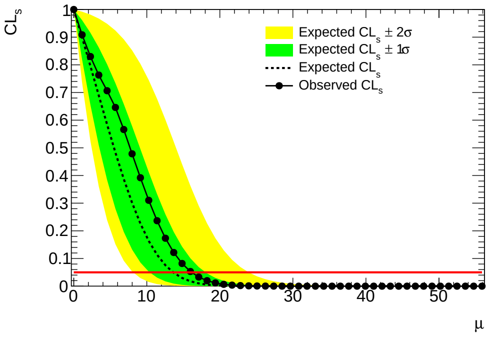
Until now, several search efforts for dark photons in the Higgs portal interpretations in the promptly decaying scenario ($c\tau < 0.01$ mm) have been conducted by the ATLAS [3] [4]



(a) FRVZ 17 MeV γ_d (b) FRVZ 30 MeV γ_d

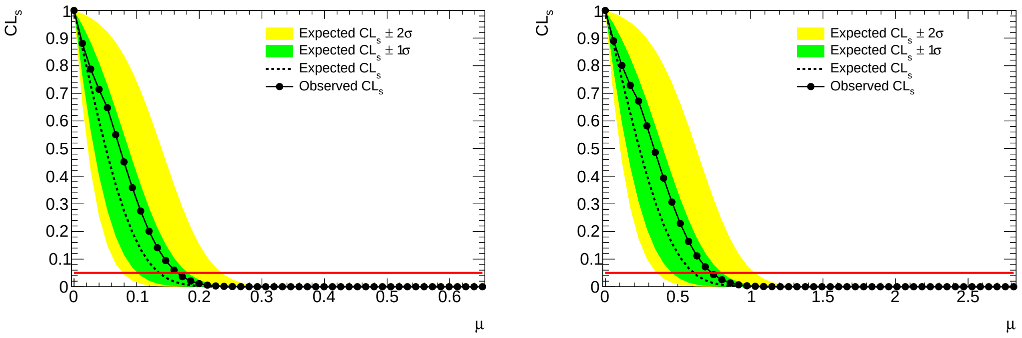


(c) FRVZ 60 MeV γ_d (d) FRVZ 100 MeV γ_d

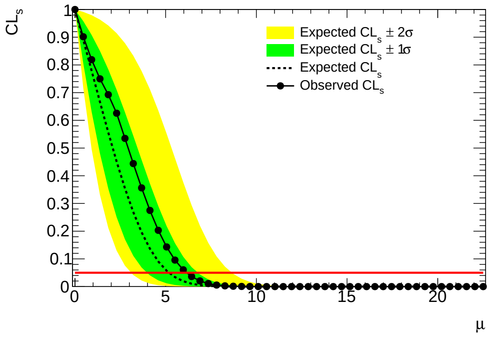


(e) FRVZ 240 MeV γ_d

Fig. 9.3 CL$_\mathrm{S}$ as functions of the signal strength μ for different γ_d masses in the FRVZ benchmark model. The signal strength $\mu = 1$ refers to BR($H \to 2\gamma_\mathrm{d} + X$)=0.5%. The upper limit at 95% is highlighted in red.



(a) HAHM 17 MeV γ_d (b) HAHM 100 MeV γ_d



(c) HAHM 400 MeV γ_d

Fig. 9.4 $\mathrm{CL_S}$ as functions of the signal strength μ for different γ_d masses in the HAHM benchmark model. The signal strength $\mu = 1$ refers to BR($H \to 2\gamma_\mathrm{d}$)=0.5%. The upper limit at 95% is highlighted in red.

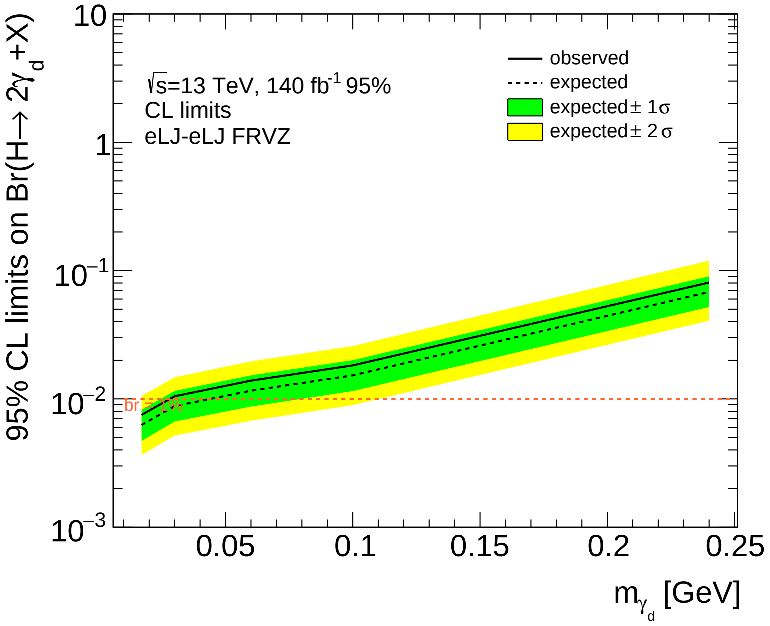


Fig. 9.5 Expected and observed limits at the 95% CL on BR($H \to 2\gamma_{\mathrm{d}} + X$) for the FRVZ model promptly decaying γ_{d} ($c\tau < 0.01$ mm). Systematic uncertainties are included. For $m_{\gamma_{\mathrm{d}}} > 2m_{\mu}$, the branching ratio of muon decays is also taken into account according to Fig. 3.6.

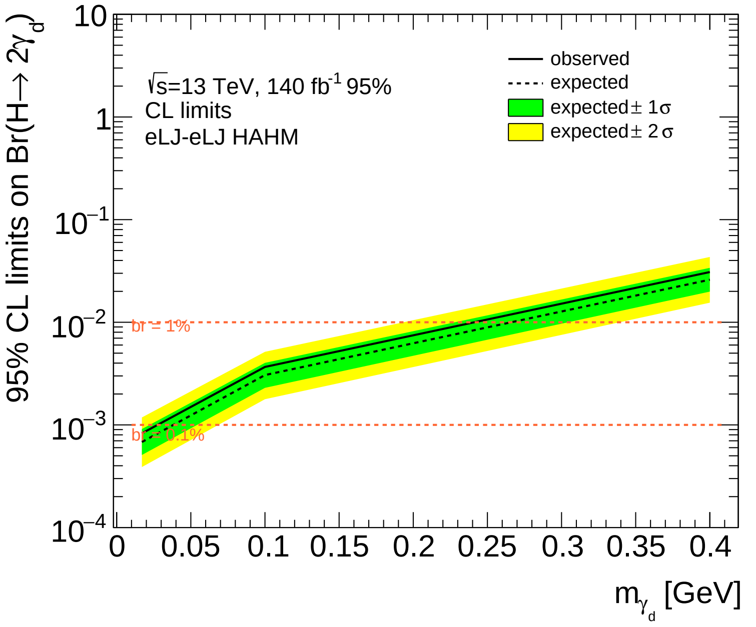


Fig. 9.6 Expected and observed limits at the 95% CL on BR($H \to 2\gamma_{\mathrm{d}}$) for the HAHM model promptly decaying γ_{d} ($c\tau < 0.01$ mm). Systematic uncertainties are included. For $m_{\gamma_{\mathrm{d}}} > 2m_{\mu}$, the branching ratio of muon decays is also taken into account according to Fig. 3.6.

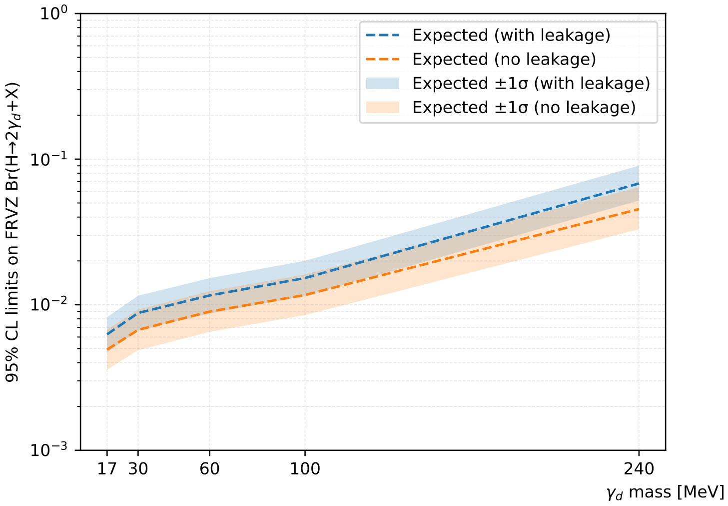


Fig. 9.7 Expected and observed limits at the 95% CL on BR($H \to 2\gamma_\mathrm{d} + X$) for the FRVZ model promptly decaying γ_d ($c\tau < 0.01$ mm). Systematic uncertainties are included. Signal contributions in regions B, C, and D are zero in the label *no leakage*. For $m_{\gamma_\mathrm{d}} > 2m_\mu$, the branching ratio of muon decays is also taken into account according to Fig. 3.6.

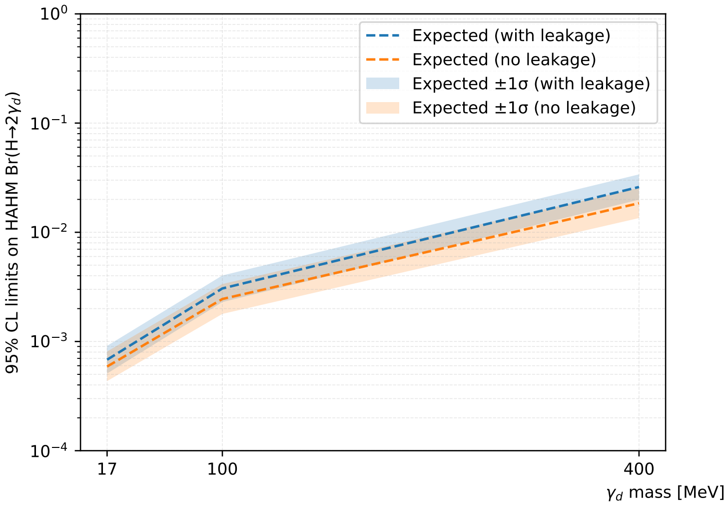
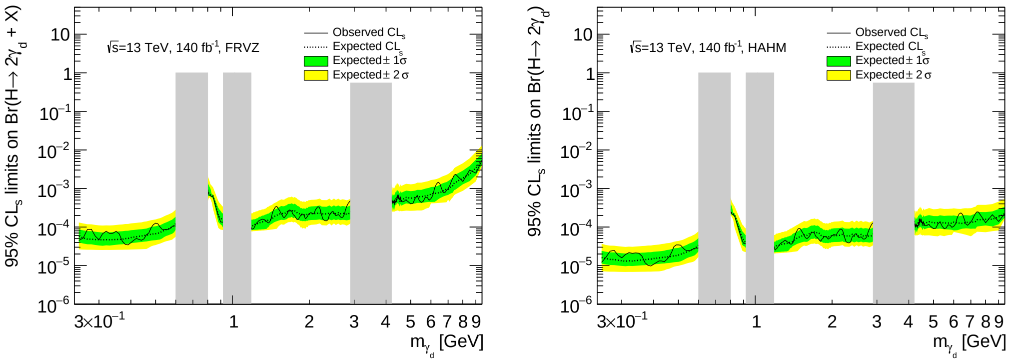


Fig. 9.8 Expected and observed limits at the 95% CL on BR($H \to 2\gamma_\mathrm{d}$) for the HAHM model promptly decaying γ_d ($c\tau < 0.01$ mm). Systematic uncertainties are included. Signal contributions in regions B, C, and D are zero in the label *no leakage*. For $m_{\gamma_\mathrm{d}} > 2m_\mu$, the branching ratio of muon decays is also taken into account according to Fig. 3.6.



(a) (b)

Fig. 9.9 95% CL exclusion limit (a) on $BR(H \to 2\gamma_d + X)$ for the FRVZ model, and (b) on $BR(H \to 2\gamma_d)$ for the HAHM model in the μLJ-μLJ and eLJ-μLJ combined channel of full run-2 pp collision data collected in the ATLAS experiment. Resonant states are blinded

.

and CMS [5] experiments at the LHC, and the excluded phase-spaces are limited to γ_{d} mass larger than 0.25 GeV. By exploring signatures of γ_{d}s that promptly decay into the more challenging electron final states, this analysis pushes the search limit to the previously unexplored $[0.017, 0.24]$ GeV lower mass range benchmarked by the FRVZ and HAHM models. It provides the first exclusion limit for the Higgs portal in the FRVZ and HAHM dark sector model, which is way below the BR($H \to inv$)$<$ 11% constraint. Furthermore, this analysis can potentially be sensitive in interpreting even lower mass range and eventually limited by the resolution of ID tracking performance, which is 130 μm in the TRT in the ATLAS experiment.

9.3.4 Vector portal interpretation

In this analysis, the search results for γ_{d}s have been optimized for the promptly decaying scenarios, specifically where $c\tau < 0.01$ mm. To compare these results with constraints from other experiments, the exclusion limits are interpreted at 90% CL in the vector portal using Eq. 3.5, assuming the production of γ_{d}s through the Higgs portal in the FRVZ and HAHM models. Fig. 9.10 shows the exclusion contours as a function of ε and $m_{\gamma_{\mathrm{d}}}$ for the FRVZ model, while Fig. 9.11 shows them for the HAHM model. It is shown that this analysis extends the exclusions to unsearched regions of $\varepsilon \sim 10^{-3}$ in the vector portal through the Higgs portal in both the FRVZ and HAHM models, compared with other vector-portal-only experiments and the displaced search in ATLAS which focuses on decaying of γ_{d} outside the inner detector. Note that $c\tau < 0.01$ mm is a conservative estimate, and this analysis could be sensitive to γ_{d}s with longer lifetimes (smaller ε). For future improvements, we can perform

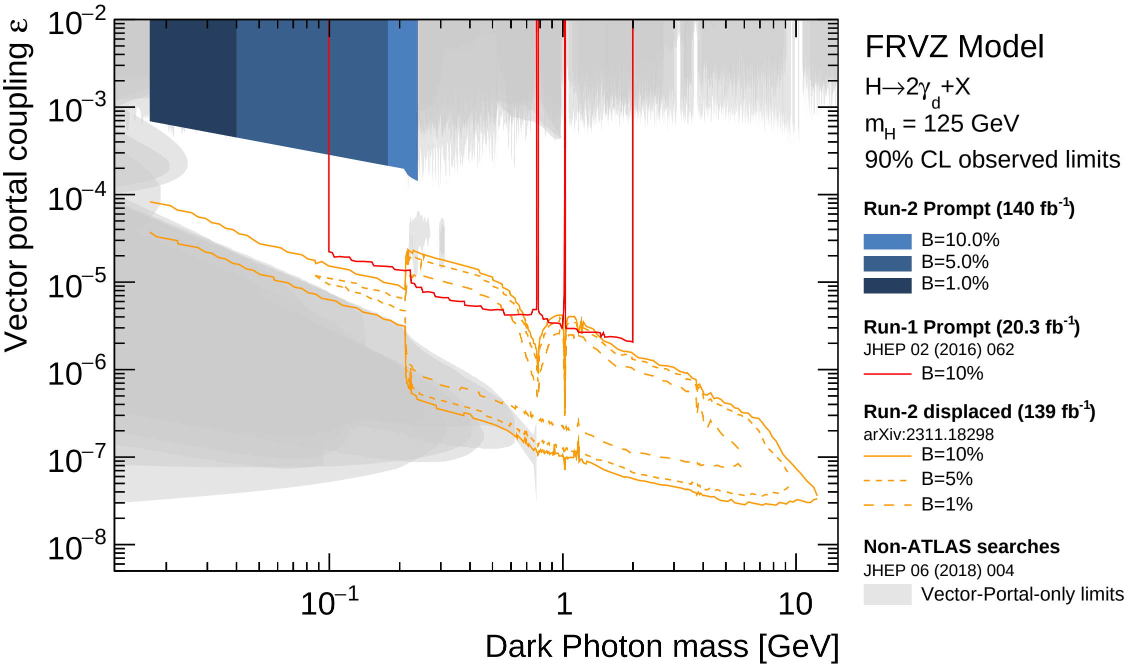


Fig. 9.10 90% exclusion contours as a function of ε and m_{γ_d} for the FRVZ model in the assumption of γ_d decay length $c\tau < 0.01$ mm
. The limits are obtained by assuming a decay branching ratio of Higgs boson into γ_ds ranging between 10% to 1%, where the darker colors correspond to stronger choices of branching ratio. The figure also shows previous ATLAS Run-1 prompt [4], ATLAS Run-2 dispaced [53], and other vector-portal-only limits.

lifetime reweighting to extrapolate signal efficiency using samples generated with different γ_d mean decay lifetimes, which will fully exploit the limits of this analysis in the vector portal.

9.4 Future developments

The work in this thesis lays the groundwork for future analyses at the LHC. By utilizing the scenario of merged clusters with separate tracks, which provide a rich phenomenology, we can access and probe collimated decays of dark photons and other new physics particles. A potential improvement comes from the intermediate scenario where electrons from dark photon decays are close but not merging in the calorimeter. As shown in this thesis, constructing such objects using standard definitions results in lower efficiency. In future upgrades like HL-LHC, dedicated triggers targeting such signatures could be developed based on the unique shower shapes and tracking information. Additionally, cluster reconstruction and

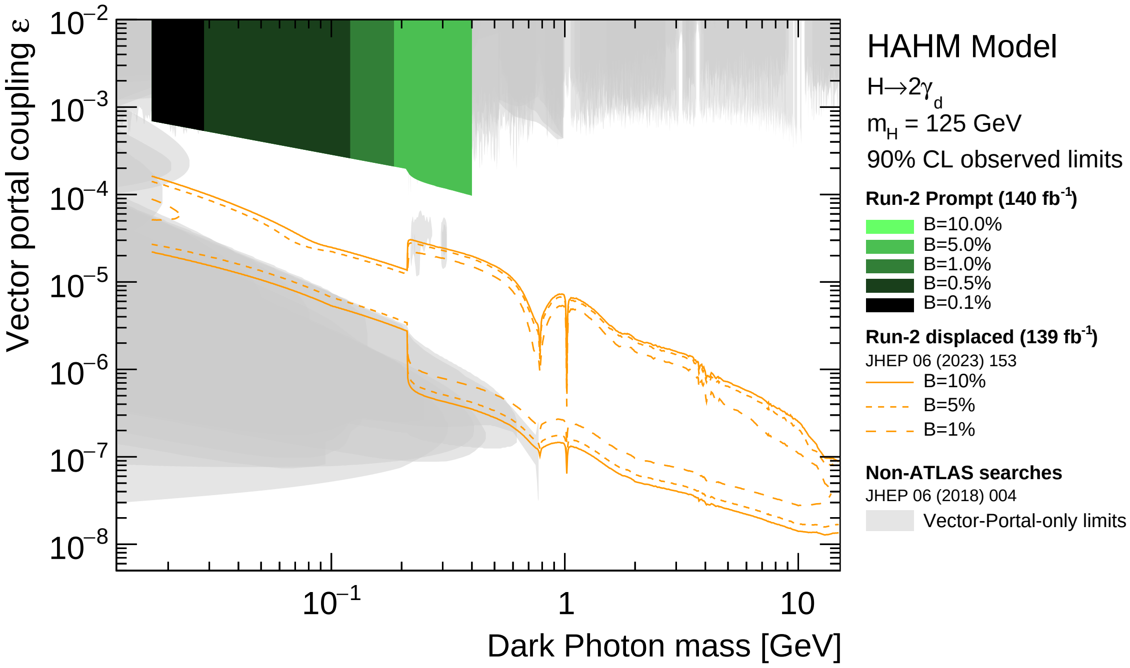


Fig. 9.11 90% exclusion contours as a function of ε and m_{γ_d} for the HAHM model in the assumption of γ_d decay length $c\tau < 0.01$ mm. The limits are obtained by assuming a decay branching ratio of Higgs boson into γ_ds ranging between 10% to 0.1%, where the darker colors correspond to stronger choices of branching ratio. The figure also shows previous ATLAS Run-2 dispaced [53], and other vector-portal-only limits.

isolation criteria can be further optimized. While this analysis focuses on promptly decaying scenarios ($\varepsilon \gtrsim 10^{-3}$), studying the effects of vertex fitting from ID tracks can help bridge the gap between prompt and displaced analyses ($\varepsilon \lesssim 10^{-4}$).

Chapter 10

Conclusion

In this thesis, we have presented a search for dark photons that promptly decay into electron final states with a decay length $c\tau < 0.01$ mm in 140 fb^{-1} of pp collision data collected at $\sqrt{s} = 13$ TeV with the ATLAS detector. The search investigates evidence of dark sector Higgs portal and vector portal production of γ_{d}s using the FRVZ and HAHM models. Specifically, in the FRVZ model, the sensitivity is explored under the assumption that the masses of HLSP and f_{d} are a few GeV, with the condition that $m_{\text{HLSP}} + m_{\gamma_{\text{d}}} < m_{f_{\text{d}}}$ (detailed mass parameters are summarized in Table 5.2). Similar assumptions are made in other experiments [4] [53]. The analysis targets the physics-motivated low-mass dark photons by exploiting unique signatures from collimated electron pairs in the ATLAS detector. Utilizing the resolution difference between the inner detector and the calorimeter, we have developed an innovative criterion for LJ reconstruction from merged clusters associated with separate tracks. This extends the Higgs portal sensitivity limit from 0.25 GeV in previous ATLAS and CMS efforts to 17 MeV. Vector portal is also interpreted for γ_{d}s in the FRVZ and HAHM Higgs portal production. Compared with other vector-portal-only experiments, we have set exclusions to unexplored regions around $\varepsilon \sim 10^{-3}$.

No significant excess of events is observed compared to the expected number of background events, and upper limits are set on the branching fractions of Higgs productions of γ_{d}s for the models of FRVZ and HAHM. The first prompt dark photon exclusion with a limit of 0.7% BR($H \to 2\gamma_{\text{d}} + X$) for the FRVZ model, and 0.07% BR($H \to 2\gamma_{\text{d}}$) for the HAHM model at $m_{\gamma_d} = 17$ MeV is set, which is far below the BR($H \to inv$)= 11%.

References

[1] Ilten Philip et al. Serendipity in dark photon searches. *Journal of High Energy Physics*, 2018(6), June 2018.

[2] ATLAS Collaboration. Combination of searches for invisible decays of the Higgs boson using 139 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment. *Physics Letters B*, 842:137963, July 2023.

[3] ATLAS Collaboration. Search for higgs bosons decaying into new spin-0 or spin-1 particles in four-lepton final states with the atlas detector with 139 fb^{-1} of pp collision data at $\sqrt{s}$ = 13 tev. *Journal of High Energy Physics*, 2022(41), 2022.

[4] ATLAS Collaboration. A search for prompt lepton-jets in pp collisions at $\sqrt{\mathrm{s}} = 8$ tev with the atlas detector. *Journal of High Energy Physics*, 2016(2), February 2016.

[5] CMS Collaboration. A search for pair production of new light bosons decaying into muons in proton-proton collisions at 13 tev. *Physics Letters B*, 796:131–154, September 2019.

[6] Sheldon L. Glashow. Partial-symmetries of weak interactions. *Nuclear Physics*, 22(4):579–588, 1961.

[7] A. Salam and J.C. Ward. Electromagnetic and weak interactions. *Physics Letters*, 13(2):168–171, 1964.

[8] Steven Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19:1264–1266, Nov 1967.

[9] ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716(1):1–29, September 2012.

[10] CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B*, 716(1):30–61, September 2012.

[11] Particle Data Group. Review of particle physics. *Progress of Theoretical and Experimental Physics*, 2022(8):083C01, 2022.

[12] Peter W Higgs. Broken symmetries and the masses of gauge bosons. *Physical Review Letters*, 13(16):508–509, 1964.

[13] Wikipedia contributors. Standard Model of Elementary Particles. https://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg, 2023.

[14] E. Fermi. An attempt of a theory of beta radiation. 1. *Z. Phys.*, 88:161–177, 1934.

[15] Steven Weinberg. A model of leptons. *Physical Review Letters*, 19(21):1264–1266, 1967.

[16] Murray Gell-Mann. The interpretation of the new particles as displaced charge multiplets. *Il Nuovo Cimento*, 4:848–866, 1956.

[17] Kazuhiko Nishijima. Charge independence theory of v particles. *Progress of Theoretical Physics*, 13:285–304, 1955.

[18] F. Englert and R. Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964.

[19] Peter W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964.

[20] C. Amsler et al. Review of Particle Physics. *Physics Letters B*, 667(1):1–6, 2008. Review of Particle Physics.

[21] CDF Collaboration. High-precision measurement of the W boson mass with the CDF II detector. *Science*, 376(6589):170–176, 2022.

[22] Y. Fukuda et al. Evidence for Oscillation of Atmospheric Neutrinos. *Physical Review Letters*, 81:1562–1567, 1998.

[23] D. Walsh et al. 0957 + 561 A, B: twin quasistellar objects or gravitational lens. *Nature*, 279:381–384, 1979.

[24] G. Hinshaw et al. Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results. *The Astrophysical Journal Supplement Series*, 208:19, 2013.

[25] Albert Einstein. The Foundation of the General Theory of Relativity. *Annalen der Physik*, 354(7):769–822, 1916.

[26] Muon g2 Collaboration. Measurement of the Positive Muon Anomalous Magnetic Moment to 0.20 ppm. *Physical Review Letters*, 131(16), October 2023.

[27] T. Aoyama et al. The anomalous magnetic moment of the muon in the Standard Model. *Physics Reports*, 887:1–166, December 2020.

[28] CDF Collaboration. Measurement of the W boson mass with the CDF II detector. *Science*, 376(6589):170–176, 2022.

[29] P. A. Zyla et al. Review of Particle Physics. *PTEP*, 2020(8):083C01, 2020.

[30] ATLAS Collaboration. Measurement of the W-boson mass in pp collisions at $\sqrt{s} =$ 7 TeV with the ATLAS detector. *The European Physical Journal C*, 78(2), February 2018.

[31] A.J. Krasznahorkay et al. Observation of Anomalous Internal Pair Creation in Be8: A Possible Signature of a Light, Neutral Boson. *Physical Review Letters*, 116(4):042501, 2016.

[32] Jonathan L. et al. Feng. Particle physics models for the 17 MeV anomaly in beryllium nuclear decays. *Phys. Rev. D*, 95:035017, Feb 2017.

[33] HFLAV Collaboration. Preliminary average of R(D) and R(D*) for Winter 2023. https://hflav-eos.web.cern.ch/hflav-eos/semi/winter23_prel/html/RDsDsstar/RDRDs.html, 2023.

[34] S. Aoki et al. Flag review 2019: Flavour lattice averaging group (flag). *The European Physical Journal C*, 80(2), February 2020.

[35] Svjetlana Fajfer et al. On the B to D* tau nu Sensitivity to New Physics. *Physical Review D*, 85(9), May 2012.

[36] Russell Kirk. *Dark Matter Genesis*. PhD thesis, Royal Holloway, U. of London, Royal Holloway, U. of London, 2017.

[37] Matt Reece. Hiding thermal dark matter with leptons. https://www.physicsmatt.com/blog/2017/5/13/paper-explainer-hiding-thermal-dark-matter-with-leptons, 2017.

[38] Search for low-mass dark matter WIMPs with 12 ton-day exposure of DarkSide-50. *Phys. Rev. D*, 107:063001, Mar 2023.

[39] IceCube Collaboration. Evidence for High-Energy Extraterrestrial Neutrinos at the IceCube Detector. *Science*, 342(6161):1242856, 2013.

[40] Super-Kamiokande Collaboration. Search for Supernova Neutrino Bursts at Super-Kamiokande. *The Astrophysical Journal*, 669(1):519–524, 2007.

[41] IceCube Collaboration. Search for GeV-scale dark matter annihilation in the Sun with IceCube DeepCore. *Physical Review D*, 105(6), March 2022.

[42] ATLAS Collaboration. Search for new phenomena in events with an energetic jet and missing transverse momentum in pp collisions at $\sqrt{s}$ =13 TeV with the ATLAS detector. *Phys. Rev. D*, 103(11):112006, 2021.

[43] CMS Collaboration. Search for new particles in events with energetic jets and large missing transverse momentum in proton-proton collisions at $\sqrt{s} = 13$ TeV. 2021.

[44] Curtin et al. Illuminating dark photons with high-energy colliders. *Journal of High Energy Physics*, 2015(2), February 2015.

[45] Batell et al. Probing a Secluded U(1) at B-factories. *Phys. Rev. D*, 79:115008, 2009.

[46] Berryman et al. Searches for decays of new particles in the DUNE Multi-Purpose near Detector. *Journal of High Energy Physics*, 2020, 02 2020.

[47] Iacopo Longarini. *A search for long-lived particles with the ATLAS Experiment*. PhD thesis, Rome U., 2022.

[48] Torben Ferber et al. Dark Higgs bosons at colliders. *Progress in Particle and Nuclear Physics*, 136:104105, March 2024.

[49] Curtin et al. Illuminating dark photons with high-energy colliders. *Journal of High Energy Physics*, 2015(2), February 2015.

[50] Falkowski et al. Hidden Higgs decaying to lepton jets. *Journal of High Energy Physics*, 2010(5), May 2010.

[51] Falkowski et al. Discovering Higgs Boson Decays to Lepton Jets at Hadron Colliders. *Physical Review Letters*, 105(24), December 2010.

[52] ATLAS Collaboration. Search for light long-lived neutral particles that decay to collimated pairs of leptons or light hadrons in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Journal of High Energy Physics*, 2023(6), June 2023.

[53] ATLAS Collaboration. Search for light long-lived neutral particles that decay to collimated pairs of leptons or light hadrons in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Journal of High Energy Physics*, 2023(6), June 2023.

[54] ATLAS Collaboration. Search for light long-lived neutral particles from Higgs boson decays via vector-boson-fusion production from pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, 2023. arXiv:2311.18298 [hep-ex].

[55] Linear accelerator 2. 2012. https://cds.cern.ch/record/1997427.

[56] K. H. Reich. The cern proton synchrotron booster. *IEEE Transactions on Nuclear Science*, 16(3):959–961, 1969.

[57] Edouard Regenstreif. *The CERN Proton Synchrotron*. CERN Yellow Reports: Monographs. CERN, Geneva, 1962.

[58] The Super Proton Synchrotron. 2012. https://cds.cern.ch/record/1997188.

[59] CERN. Run 3 first collisions, 2023. https://atlas.cern/Updates/Press-Statement/Run3-first-collisions.

[60] ATLAS Collaboration. The atlas experiment at the cern large hadron collider. *Journal of Instrumentation*, 3:S08003, 2008.

[61] CMS Collaboration. The cms experiment at the cern lhc. *Journal of Instrumentation*, 3:S08004, 2008.

[62] ALICE Collaboration. The alice experiment at the cern lhc. *Journal of Instrumentation*, 3:S08002, 2008.

[63] LHCb Collaboration. The lhcb detector at the lhc. *Journal of Instrumentation*, 3:S08005, 2008.

[64] Esma Mobs. The CERN accelerator complex - August 2018. Complexe des accélérateurs du CERN - Août 2018. 2018. General Photo.

[65] Ruth Pöttgen. Search for Dark Matter in events with a highly energetic jet and missing transverse momentum in proton-proton collisions at $\sqrt{s}$ = 8 TeV with the ATLAS Detector, 2016. Presented 08 May 2015.

[66] Brüning Oliver Sim et al. *LHC Design Report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2004.

[67] ATLAS Collaboration. Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC. 6 2019.

[68] ATLAS Collaboration. Luminosity determination in pp collisions at $\sqrt{s} = 13$ tev using the atlas detector at the lhc. *The European Physical Journal C*, 83(10), October 2023.

[69] CERN. Public atlas luminosity results for run-2 of the lhc. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2#Luminosity_Plots_for_multiple_Ru.

[70] Joao Pequenao. Computer generated image of the whole ATLAS detector. https://cds.cern.ch/record/1095924, 2008.

[71] Izaak Neutelings. CMS coordinate system, 2023. https://tikz.net/axis3d_cms/.

[72] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008.

[73] ATLAS Collaboration. ATLAS inner detector: Technical design report. Vol. 1. 4 1997.

[74] Lydia Audrey Beresford. The atlas experiment. In *Searches for Dijet Resonances*, Springer Theses, pages 23–38. Springer, Cham, 2018.

[75] ATLAS Collaboration. ATLAS liquid argon calorimeter: Technical design report. 12 1996.

[76] ATLAS Collaboration. Calorimeter. https://atlas.cern/Discover/Detector/Calorimeter.

[77] ATLAS Collaboration. ATLAS muon spectrometer: Technical design report. 6 1997.

[78] ATLAS Collaboration. ATLAS level-1 trigger: Technical Design Report. 6 1998.

[79] ATLAS Collaboration. Performance of electron and photon triggers in ATLAS during LHC Run 2. *Eur. Phys. J. C*, 80:47, 2020.

[80] David J. Gross and Frank Wilczek. Ultraviolet behavior of non-abelian gauge theories. *Physical Review Letters*, 30(26):1343–1346, 1973.

[81] J. Alwall et al. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *Journal of High Energy Physics*, 2014(7):079, 2014.

[82] S. Alioli et al. A general framework for implementing nlo calculations in shower monte carlo programs: the powheg box. *Journal of High Energy Physics*, 2010(6):043, 2010.

[83] Torbjörn Sjöstrand et al. An Introduction to PYTHIA 8.2. *Comput. Phys. Commun.*, 191:159–177, 2015.

[84] J. Bellm et al. Herwig 7.0/Herwig++ 3.0 release note. *Eur. Phys. J. C*, 76:196, 2016.

[85] T. Gleisberg et al. Event generation with SHERPA 2.2. *Journal of High Energy Physics*, 2009:007, 2009.

[86] GEANT4 Collaboration. Recent developments in geant4. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 835:186–225, 2016.

[87] ATLAS Collaboration. Atlas track reconstruction. https://atlassoftwaredocs.web.cern.ch/trackingTutorial/idoverview/.

[88] ATLAS Collaboration. Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC run 2. Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC run 2. *Eur. Phys. J. C*, 77(10):673, 2017.

[89] ATLAS Collaboration. Track Reconstruction Performance of the ATLAS Inner Detector at $\sqrt{s} = 13$~TeV. 7 2015.

[90] ATLAS Collaboration. Reconstruction of primary vertices at the atlas experiment in run 1 proton–proton collisions at the lhc. *The European Physical Journal C*, 77(5), May 2017.

[91] ATLAS Collaboration. Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data. *Journal of Instrumentation*, 14(12):P12006–P12006, December 2019.

[92] ATLAS Collaboration. Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung. 5 2012.

[93] A. Hoecker et al. Tmva - toolkit for multivariate data analysis, 2009.

[94] ATLAS Collaboration. Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13$ TeV. *Eur. Phys. J. C*, 81(7):578, 2021.

[95] ATLAS Collaboration. Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13$ TeV. *Eur. Phys. J. C*, 81(7):578, 2021.

[96] ATLAS Collaboration. Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s}$ =13 TeV. *Eur. Phys. J. C*, 76(5):292, 2016.

[97] ATLAS Collaboration. Jet reconstruction and performance using particle flow with the ATLAS Detector. *European Physical Journal C*, 77:466, 2017.

[98] Matteo Cacciari et al. The anti-k_t jet clustering algorithm. *Journal of High Energy Physics*, 2008(04):063, 2008.

[99] S. Catani et al. Longitudinally invariant k_t clustering algorithms for hadron hadron collisions. *Nuclear Physics B*, 406:187–224, 1993.

[100] M. Wobisch and T. Wengler. The cambridge jet algorithm: Features and applications. In *Monte Carlo Generators for HERA Physics, Proceedings*, pages 270–279, 1998.

[101] ATLAS Collaboration. Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *The European Physical Journal C*, 81(8), August 2021.

[102] ATLAS Collaboration. Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *The European Physical Journal C*, 81(8), August 2021.

[103] ATLAS Collaboration. Performance of jet substructure techniques for large-R jets in proton-proton collisions at 7 TeV using the ATLAS detector. *Journal of High Energy Physics*, 2013(76):1–48, 2013.

[104] ATLAS Collaboration. Luminosity determination in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ using the ATLAS detector at the LHC. *Eur. Phys. J. C*, 83:982, 2023.

[105] E.W.N. Glover and J.J. van der Bij. Higgs boson pair production via gluon fusion. *Nuclear Physics B*, 309(2):282–294, 1988.

[106] CERN. Cern yellow report page at 13 tev. https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageAt13TeV.

[107] R. D. Ball, V. Bertone, S. Carrazza, C. S. Deans, L. Del Debbio, S. Forte, A. Guffanti, N. P. Hartland, J. I. Latorre, J. Rojo, and L. Rottoli. Parton distributions for the lhc run ii. *Journal of High Energy Physics*, 2015(4):40, 2015.

[108] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. Pythia 6.4 physics and manual. *Journal of High Energy Physics*, 2006(05):026, 2006.

[109] Peter Z. Skands. Tuning Monte-Carlo generators: The Perugia tunes. *Physical Review D*, 82(7), October 2010.

[110] H.-L. Lai et al. New parton distributions for collider physics. *Physical Review D*, 82:074024, 2010.

[111] J. et al. Pumplin. New generation of parton distributions with uncertainties from global qcd analysis. *Journal of High Energy Physics*, 2002(07):012, 2002.

[112] Yu.L. Dokshitzer et al. Better jet clustering algorithms. *Journal of High Energy Physics*, 1997(08):001, sep 1997.

[113] Suyong Choi and Hayoung Oh. Improved extrapolation methods of data-driven background estimation in high-energy physics, 2021.

[114] ATLAS Collaboration. Electron and photon energy calibration with the ATLAS detector using LHC Run-2 data. *Journal of Instrumentation*, 19(02):P02009, February 2024.

[115] G. Avoni et al. The new lucid-2 detector for luminosity measurement and monitoring in atlas. *Journal of Instrumentation*, 13(07):P07017, jul 2018.

[116] Börner, Daniela. A method for the construction of strongly reduced representations of ATLAS experimental uncertainties and the application thereof to the jet energy scale. *EPJ Web Conf.*, 137:11003, 2017.

[117] V. N. Gribov and L. N. Lipatov. Deep inelastic e p scattering in perturbation theory. *Sov. J. Nucl. Phys.*, 15:438–450, 1972.

[118] Yuri L. Dokshitzer. Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics. *Sov. Phys. JETP*, 46:641–653, 1977.

[119] Guido Altarelli and G. Parisi. Asymptotic Freedom in Parton Language. *Nucl. Phys. B*, 126:298–318, 1977.

[120] Christopher S. Deans. Progress in the NNPDF global analysis, 2013.

[121] A L Read. Presentation of search results: the cls technique. *Journal of Physics G: Nuclear and Particle Physics*, 28(10):2693, sep 2002.

[122] Cowan Glen et al. Asymptotic formulae for likelihood-based tests of new physics. *The European Physical Journal C*, 71(2), February 2011.