A SEARCH FOR LONG-LIVED PARTICLES IN SIGNATURES WITH DISPLACED VERTEX USING NOVEL MACHINE LEARNING TECHNIQUES AT CMS

Ang Li

Tengzhou, Shandong

Bachelor of Science, Shandong University, 2017

A Dissertation submitted to the Graduate Faculty

of the University of Virginia in Candidacy for the Degree of

Doctor of Philosophy

Department of Physics

University of Virginia

May 2023

Chris Neu, Chair

Craig Group

Julian Heeck

Remy Indebetouw

A Search for Long-Lived Particles in Signatures With Displaced Vertex Using Novel Machine Learning Techniques at CMS

Ang Li

(ABSTRACT)

A search for long-lived particles produced in proton-proton collisions at a center-of-mass energy of 13 TeV at the CERN LHC is presented. The search is based on data collected by the CMS experiment in 2016–2018, which corresponds to a total integrated luminosity of 137 fb^{-1}. The search targets final states with at least one displaced vertex and missing transverse momentum. Customized vertex reconstruction and advanced machine learning algorithms are applied to increase the sensitivity of the search. Moreover, the search is designed to be model-independent to be sensitive to a wide range of new physics models. No significant excess over the background–only prediction is observed. For the mean proper decay length in the range 1–100 mm, the search excludes long-lived gluinos in the split supersymmetry model with masses below 1800 or 2000 GeV, depending on the neutralino masses, at a 95% confidence level.

Dedication

To my mom and dad.

Acknowledgments

Chris, thank you so much for being willing to accept me as your student. For all the years we have gone through, you are always supportive and patient. As my supervisor, you gave me the freedom to choose the research topic I am interested in and keep guiding me toward the end of my Ph.D. No words can express my appreciation to you.

Craig, Julian, and Remy, thank you for being willing to be members of my Ph.D. thesis committee. I learned a lot when discussing physics with you. Eugene, thank you very much for the help you provided as the chair of my research review committee.

Ben, thank you for the help you provided at the beginning period of my study.

Bryan, thank you for all the time you spent discussing with me and the great thoughts you suggested. You supported me not only in the research but also in life. Joey, thank you for being so patient and willing to help me all the time. You guided me all the way along the search.

Ritchie, Shaun, Peace, Abby, Bob, Thomas, John, Grace, Matt, Colby, Gracemarie, Becky, Maria, Johnathon, and Zhenyu, thank you for all the valuable discussions and suggestions. It is really a pleasure to work with you.

Slava and Cristian, thank you for all the suggestions and help you provided as the conveners. Your suggestions really made the search more mature and complete.

Adi, Eric, Karlis, Amrutha, and Braden, thank you for the enjoyable experience with the MTD studies.

Everyone in the Physics department at the University of Virginia, thank you for the lovely and supportive environment. Specifically, thank you to Peter and Beth, who

addressed all questions I encountered related to the administrative side of my Ph.D. period.

Thank you to all of my friends at UVA and CERN. Specifically, Chen, Nan, Sheng, and Wei, I still remember the time we spent together preparing for the courses and exams. It is delightful to have such memories in mind and know that we are not alone. Cristiana, you and Bryan are so kind and enthusiastic. I really appreciate the amusement and excitement I get when interacting with you. Yousen and Jackie, thank you for the lovely time we spent together.

Jiawei, it is my fortune to share the fantastic journey with you. You made my life much more enjoyable and exciting. I wish the glee between us lasts forever. I love you.

Mom and Dad, thank you so much for raising me up and giving me all the care and love. I could never go this far without the encouragement and support you provide unconditionally.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Elementary particle physics explores the basic components of the universe and the principles that guide the interactions between them. Generations of particle physicists strive to build theories that explain the universe and conduct experiments to verify these theories. The discovery of the Higgs Boson by the ATLAS and CMS experiment in 2012 heralded the tremendous success of the Standard Model (SM), which describes three of the four fundamental interactions accurately.

Despite the remarkable success of the SM, there are still many mysteries, such as the feebleness of gravity, the nature of dark matter, and the hierarchy problem, that are not explained. To solve these problems, new physics hypotheses Beyond the Standard Model (BSM) have been proposed which experiments then attempt to verify. Searches for new physics have been conducted in experiments in the Large Hadron Collier (LHC) located at the European Organization for Nuclear Research (CERN) but no solid evidence of new physics has yet been found. The lack of observation of evidence for new physics suggests that we should search for unconventional signatures, which are typically harder to reconstruct and analyze.

One category of such unconventional signatures that are possible to be detected by the LHC detector is the displaced vertex signature. Most of the SM particles produced at the LHC decay almost instantly when they are generated, resulting in prompt signatures in the detector. However, many BSM theories predict the existence of

long-lived particles (LLPs), which are likely to travel a significant distance before they decay and create displaced signatures in the detector. Such signatures are generally harder to reconstruct since the detectors are designed to optimally detect prompt signatures.

This thesis presents a search that targets LLPs that result in displaced vertices and missing transverse momentum and uses data taken by the CMS detector. The search utilizes customized reconstruction and a machine learning algorithm to identify displaced vertices and to discriminate signal events from background events efficiently. A data-driven background estimation method is developed based on the vertex reconstruction and machine learning algorithms.

The thesis is structured as follows. Chapter 2 introduces the basic information about the Standard Model and some of the physics models. Chapter 3 presents an overview of the Large Hadron Collider and the Compact Muon Solenoid experiment. Chapter 4 discusses the long-lived particles and the displaced vertex signature. Chapter 5 describes the data and simulated samples. Chapter 6 introduces object selection, trigger requirements, and event preselection. Chapter 7 describes the track selection, vertex reconstruction, vertex selection, and search variables. Chapter 8 describes the machine learning algorithm construction, including the architecture, training, testing, and performance. Chapter 9 summarizes the analysis selection and defines the search regions. Chapter 10 explains the data-driven background estimation procedure. Chapter 11 explains the systematic uncertainties in the signal efficiency and background estimate. Chapter 12 describes the data events in the signal region and statistical interpretation. And Chapter 13 summarizes the thesis.

Chapter 2

Theoretical Foundations

2.1 The Standard Model

The Standard Model (SM) [1] describes our current knowledge of the fundamental particles and their interactions. Specifically, there are four fundamental interactions as far as we know, namely the electromagnetic interaction, the weak interaction, the strong interaction, and gravity. The SM describes three of them, the electromagnetic, weak, and strong interactions.

2.1.1 Particles

A diagram that shows the elementary particles described by the SM is shown in Fig.2.1. Elementary particles have different properties, including their masses, charges, and spins. The spin is a quantum number that describes the intrinsic angular momentum of a particle. Particles with a spin of a half-integer are referred to as fermions while particles with a spin of an integer are referred to as bosons. Regarding the elementary particles in the SM, particles can be divided further into three categories based on the exact value of their spins: $spin - 1/2$, $spin - 1$, and $spin - 0$ particles.

Particles with a spin of 1/2 are the basic components of daily matter. There are three generations of those particles and each generation consists of two quarks and

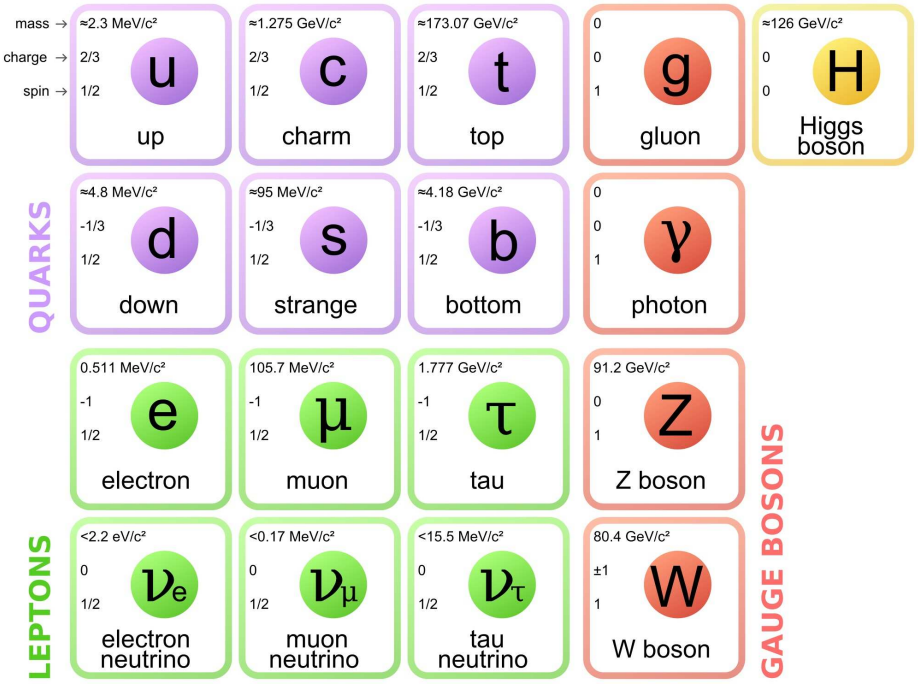

Figure 2.1: A diagram of the Standard Model [2].

two leptons. The first generation of quarks includes up and down quarks; the second generation includes charm and strange quarks; the third generation includes top and bottom quarks. Quarks are the basic components of hadrons: a quark and an antiquark form mesons while three quarks form baryons. Quarks participate in electromagnetic, weak, and strong interactions and they carry electrical charges of either 2/3 or -1/3 and color charges of red, blue, and green. However, due to color confinement, quarks can only be observed in the form of hadrons but cannot be observed as single quarks. Meanwhile, each generation of leptons consists a charged lepton and a neutral lepton: the first generation of leptons includes the electron and electron neutrino; the second generation includes the muon and muon neutrino; the third generation includes the tau and tau neutrino. Electrons, muons, and taus carry an electric charge of -1 and they participate in the electromagnetic and weak interaction. Neutrinos do not carry electric charges and they only participate in the weak

interaction.

Particles with a spin of 1 are the carriers of the interactions. Spin-1 particles include the photon, 8 gluons, two W bosons (W^+ and W^-), and the Z boson. Photons, which are the mediator of the electromagnetic interaction, and gluons, which are the mediator of the strong interaction, are massless. W and Z bosons, which are mediators of the weak interaction, are massive.

There is only one elementary particle in the SM with a spin of 0, namely the Higgs boson. The Higgs field makes the W and Z massive through the Higgs mechanism.

All particles have corresponding anti-particles. A particle and its anti-particle have the same mass and spin, but with opposite charges. Anti-particles can be obtained in the language of quantum field theory by applying charge conjugation on a particle. Anti-particles can annihilate with their corresponding particles, releasing energy in the form of photons.

2.1.2 Interactions

Interactions in the SM, namely the electromagnetic, strong, and weak interaction, are described as quantum fields. The formation of such fields is guided by gauge invariance. In quantum field theory, Lagrangians $\mathcal{L}[\psi(x)]$ are used to describe how fields $\psi(x)$ interact and evolve with time. The symmetry of the physics described by a given Lagrangian is described as $\mathcal{L}[\psi(x)]$ being invariant under transformations of $\psi(x)$ in the symmetry group. The transformation is referred to as global transformation if it does not depend on x. We can further require $\mathcal{L}[\psi(x)]$ to be invariant under a local transformation, which is dependent on x. Doing so introduces a gauge field, which is the mediator of the corresponding interaction.

The electromagnetic interaction is the interaction that acts on electrically charged particles. It is described by quantum electrodynamics (QED). QED is formed by promoting a global $U(1)$ symmetry of a Dirac field to local, thus introducing the photon field as the mediator of electromagnetic interaction. The Lagrangian of QED is

$$\mathcal{L}_{QED} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + i\bar{\psi}\gamma^{\mu}D_{\mu}\psi - m\bar{\psi}\psi \tag{2.1}$$

with a covariant derivative

$$D_{\mu} = \partial_{\mu} - ieA_{\mu} \tag{2.2}$$

where A_{μ} is the new field (photon) introduced by the local symmetry, $F_{\mu\nu}$ is defined as $\partial^{\mu}A^{\nu} - \partial^{\nu}A^{\mu}$, γ_{μ} are the Dirac matrices, and e is the electrical charge.

The strong interaction acts on particles with color charges, which confines quarks into hadrons. Quantum chromodynamics (QCD) describes the strong interaction. By promoting a global $SU(3)$ symmetry to local, the Lagrangian of QCD can be obtained:

$$\mathcal{L}_{QCD} = -\frac{1}{4}F^{\mu\nu a}F^{a}_{\mu\nu} + \psi_f(i\gamma^{\mu}D^{-}_{\mu} - m)\psi_f. \tag{2.3}$$

with a covariance derivative

$$D_{\mu} = \partial_{\mu} - ig_s A^{a}_{\mu}t^{a} \tag{2.4}$$

where $F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f^{abc} A^b_\mu A^c_\nu$, g_s is the coupling constant of the strong interaction, A^a_μ is the new field (gluon) introduced by the local symmetry, f_{abc} and t^a are the structure constants and the representation matrices for the generators of the symmetry group. The index a corresponds to the generators of the $SU(3)$ symmetry group, and the index f corresponds to the quark flavors.

The weak interaction changes the flavor of particles, such as changing electrons to electron neutrinos, and u quarks to d quarks. The weak interaction can be unified with electromagnetic interaction by the electroweak model. The model utilizes the symmetry group of $SU(2) \times U(1)$. That gives rise to four additional boson fields, A^a $(a = 1, 2, 3)$ and B. The covariance derivative is

$$D_\mu \psi = (\partial_\mu - i g A^a_\mu I^a - i g' B_\mu Y)\psi. \tag{2.5}$$

where g and g' are the coupling constants, A^a and B are the boson fields. Linear combinations of the boson fields define the fields of the photon, W, and Z bosons in the SM. When applying the covariant derivative on the Higgs field, the non-zero vacuum expectation value of the Higgs field makes the W and Z massive while keeping the photon massless. I^a is the weak isospin, Y is the hypercharge. The electrical charge of a particle Q can be calculated as $Q = I^3 + Y$.

2.1.3 Open Questions

Although the SM is very successful in describing the fundamental particles and their interactions, there are still many questions that are not answered:

- One of the four fundamental interactions, gravity is not described in the SM;

- Particles described in the SM only consist of about 5% of the universe. Dark matter and dark energy, which compose 26% and 69% of the energy density of the universe, are not explained by the SM;
- Assuming that the universe is created by the Big Bang, the amount of matter and anti-matter is supposed to be the same. However, much more matter particles are observed today compared with anti-matter particles. This imbalance, referred to as the matter-antimatter asymmetry, is not addressed in the SM;
- Neutrino oscillations, –the property of neutrinos that allows them to change flavors as they travel, and which indicates that neutrinos have masses, – is not explained in the SM;
- The SM is affiliated by the hierarchy problem [3], which depicts the huge difference between the vacuum expectation value of the Higgs field v and the Plank scale M_P, as shown in Eq. 2.6;

$$v \sim 250\,\text{GeV} << M_P \sim 10^{19}\,\text{GeV}. \tag{2.6}$$

- It is desired to have an ultimate "theory" that described all the interactions in a single form. But the SM does not unify the strong and electroweak interactions.

All these unsolved questions motivate physics beyond the Standard Model.

2.2 The Physics Beyond the Standard Model

As described in Sec. 2.1.3, the SM is not the ultimate theory. To address the open questions, many physics models Beyond the Standard Model (BSM) are proposed.

One elegant and charming BSM model is Supersymmetry(SUSY) [4].

2.2.1 Supersymmetry

Supersymmetry is the symmetry between bosons and fermions. Specifically, every particle in the SM has its superpartner, referred to as a *sparticle*. The superpartner of a boson is a fermion and vice versa. The masses of the SM particles and their superpartners are supposed to be the same if supersymmetry is unbroken. The fact that no superpartners are observed with masses equal to particle counterpartners shows that the supersymmetry is broken.

As a popular SUSY theory, the Minimal Supersymmetric Standard Model (MSSM) extends the SM to realize supersymmetry. It is composed of the fields of SM particles and their superpartners, in addition to the two-Higgs-doublet field. The MSSM assumes the breaking of supersymmetry at a relatively low-energy scale, which results in superpartners having masses below a few TeV.

Many open questions in the SM are solved by the MSSM theory. In the case of R-parity conservation, which ensures the stableness of the lightest SUSY particle, the lightest SUSY particle can be a candidate for dark matter. The hierarchy problem is solved with the low-energy supersymmetry breaking. Also, the SUSY unifies the gauge coupling constants in the SM at high energy scales. As shown in Fig. 2.2, the gauge coupling converges to the same value at high energy scales in the SUSY while in the SM the coupling constants do not converge.

Although SUSY solves many of the problems elegantly, there is no experimental evidence of the existence of SUSY particles. To explain the lack of evidence of SUSY and solve some new questions brought up by SUSY, a modification of SUSY, called

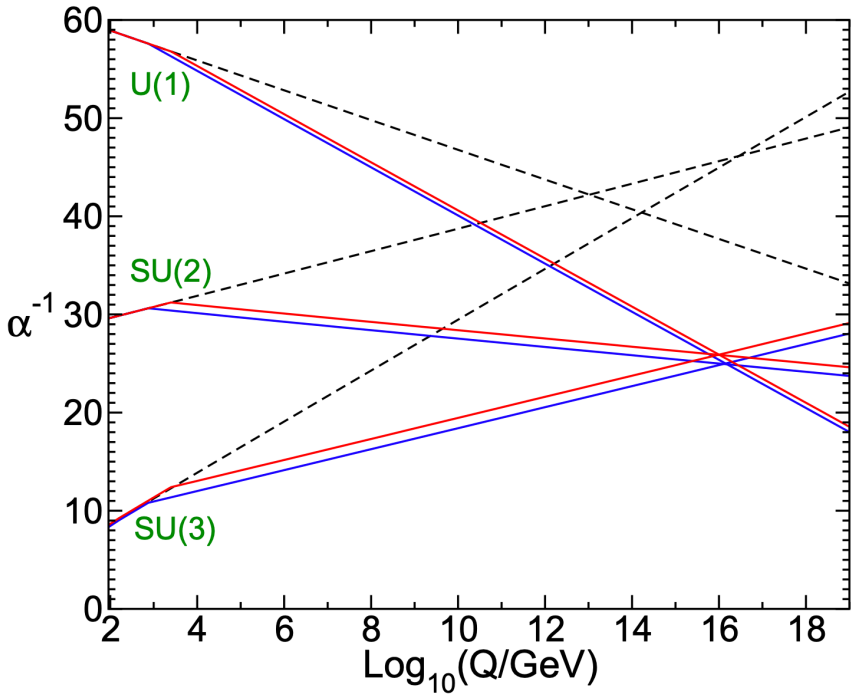

Figure 2.2: The inverse of the SM gauge coupling constant α as a function of energy [4]. The dashed lines show the SM case while the solid lines show the results for the MSSM with different thresholds of the sparticle masses.

split SUSY [5, 6], is proposed. In this model, SUSY breaks at very high energy scales m_S (near 10^9 GeV). The high-energy breaking of SUSY forces the scalars of SUSY to have masses around m_S, except for the two Higgs doublets. It keeps most of the advantages of SUSY, including the potential candidate of dark matter and gauge coupling unification. However, the hierarchy problem cannot be addressed under the assumption of supersymmetry breaking at high-energy scales.

One unique signature predicted by the split SUSY is the long-lived gluino. Gluinos can decay to a quark, an antiquark, and the lightest SUSY particle. The decay is suppressed because it happens through virtual squarks, which have masses close to m_S. This suppressed decay process makes gluinos long-lived, with a lifetime of

$$\tau \sim 3 \times 10^{-2} \text{sec} \left(\frac{m_S}{10^9 \text{ GeV}}\right)^4 \left(\frac{1 \text{ TeV}}{m_{\tilde{g}}}\right)^5, \tag{2.7}$$

where the $m_{\tilde{g}}$ is the gluino mass.

The search presented in this thesis uses the split SUSY model as the benchmark signal model, specifically the pair-produced gluinos that decay to a quark, an antiquark, and a neutralino (lightest SUSY particle). The mass splitting, defined as the mass difference between gluino and neutralino, is often used to describe the kinematics of the decay of gluinos. Different parameters, including the gluino mass, neutralino mass, and the proper decay length ($c\tau$) of the gluino, are explored to cover a wide parameter space.

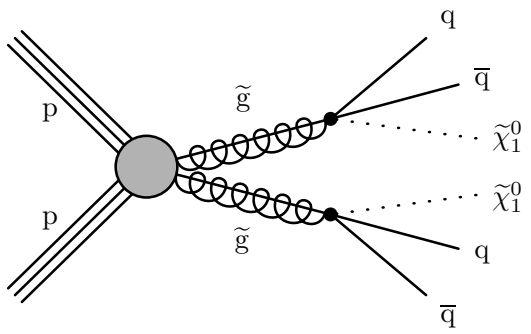

Figure 2.3: The Feynman diagram of the pair-produced long-lived gluinos in split SUSY. Each gluino ($\tilde{g}$) decays into a quark (q), an antiquark ($\bar{q}$), and a neutralino ($\tilde{\chi}_1^0$).

Chapter 3

The Large Hadron Collider and the Compact Muon Solenoid Experiment

3.1 The Large Hadron Collider

The Large Hadron Collier (LHC) [7] is the most powerful collider in the world. It is a 27km ring that is about 100m under the Swiss-French border near Geneva and located at the European Organization for Nuclear Research (CERN). The LHC is designed to accelerate two beams of protons to 7 TeV (or sometimes two beams of heavy ions to 2.8 TeV) and collides them at four interaction points on the LHC ring.

A primary goal of accelerating and colliding particles in the LHC is to explore new physics beyond the Standard Model. The number of events for a given process generated by the LHC collisions is:

$$N_{process} = L\sigma_{process} \tag{3.1}$$

where L is the luminosity of the LHC and $\sigma_{process}$ is the cross section of the process.

The luminosity is defined as:

$$L = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi \varepsilon_n \beta *} F \tag{3.2}$$

where N_b is the number of particles per bunch, n_b is the number of bunches per beam, f_{rev} is the revolution frequency, γ_r is the relativistic gamma factor, ε_n is the normalized transverse beam emittance, $\beta*$ is the beta function at the collision point, and F is the geometric luminosity reduction factor.

The LHC layout is shown in Fig.3.1. Two counter-rotating beams are accelerated and collided in the four interaction points. Each of the interaction points corresponds to an experiment, A Toroidal LHC Apparatus (ATLAS) [8], Compact Muon Solenoid (CMS) [9], LHC-beauty (LHCb) [10], and A Large Ion Collider Experiment (ALICE) [11]. The ATLAS and CMS experiments are high-luminosity experiments, which are general-purpose experiments and target the peak luminosity of $10^{34}\,\mathrm{cm}^{-2}s^{-1}$ for proton collisions. The LHCb is a dedicated experiment to study B-physics and targets the peak luminosity of $2 \times 10^{29}\,\mathrm{cm}^{-2}s^{-1}$ for proton collisions. The ALICE experiment is designed for ion collisions, which targets the peak luminosity of $10^{27}\,\mathrm{cm}^{-2}s^{-1}$ for heavy-ion collisions.

Protons are accelerated and collided in the LHC in bunches. Every beam includes 2808 bunches, with a spacing of 25ns. The beams of protons are accelerated step by step. The CERN accelerator complex is shown in Fig. 3.2. Protons are first accelerated to 50 MeV by the linear accelerator (Linac2). They are then sent to the Proton Synchrotron Booster (PSB) to be accelerated to 1.4 GeV. The Proton Synchrotron (PS) and Super Proton Synchrotron (SPS) will accelerate the protons to 25 GeV, and then 450 GeV. When the protons are accelerated to 450 GeV, they are

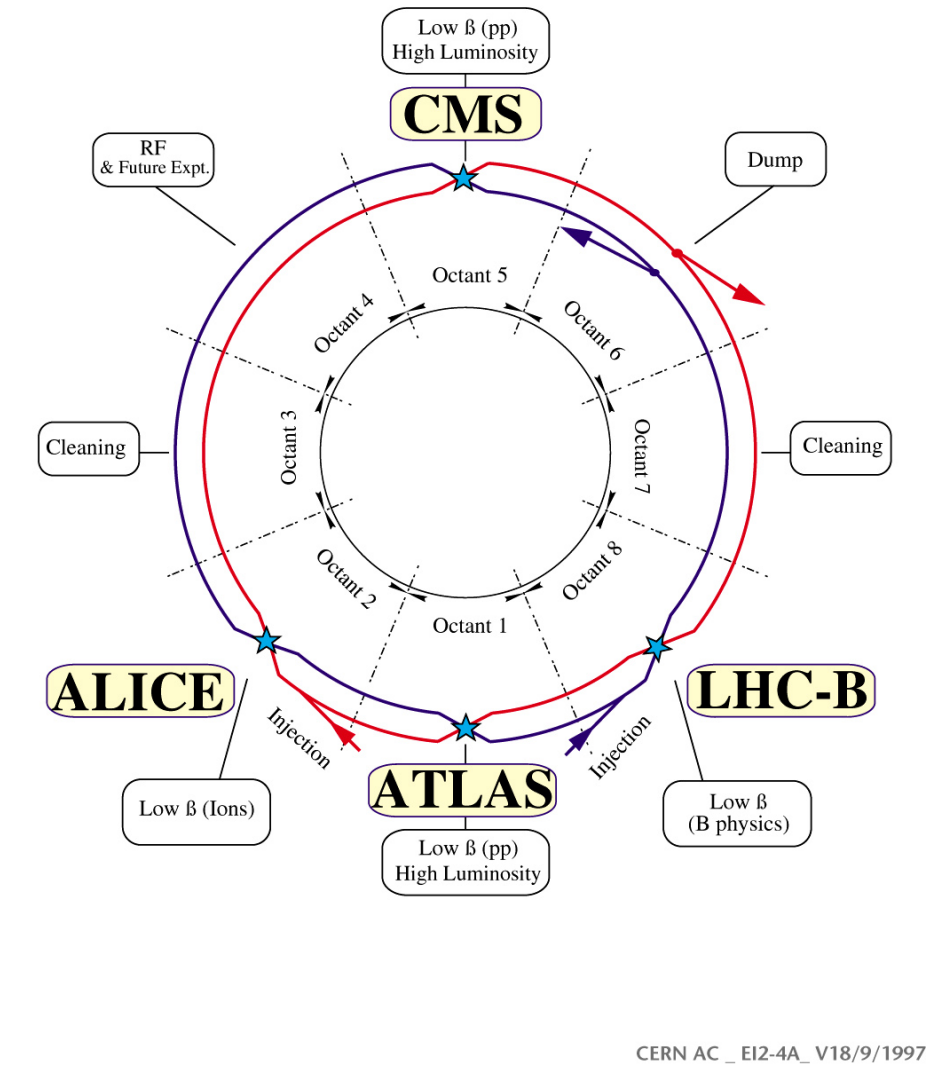

Figure 3.1: The layout of the LHC [12].

injected into the LHC to reach the final desired energy. As described above, the peak energy of protons that LHC can deliver is 7 TeV. The LHC achieves such high energy through the cutting-edge technology of superconducting magnets and radio-frequency cavities.

Given the fact the protons are grouped as bunches during the collisions, more than one collision could happen during a single bunch crossing, which is referred to as *pileup*. Pileup is normally treated as background and it increases with the increase of the instantaneous luminosity.

The search presented in this thesis is based on the data taken by the CMS experiments during the Run 2 data-taking period, which extends from 2016 to 2018, with the center-of-mass energy of 13 TeV. The center-of-mass energy is slightly lower than the designed value (14 TeV) because some of the magnets are not able to sustain the current needed for the 14 TeV energy while maintaining the super-conducting condition [13].

The LHC will undergo a major upgrade after 2025, referred to as high-luminosity LHC (HL-LHC) [14], to achieve a peak luminosity of $5 \times 10^{34}\text{cm}^{-2}\text{s}^{-1}$ and an integrated luminosity of $250\,\text{fb}^{-1}$ per year to achieve $3000\,\text{fb}^{-1}$ in about 12 years of high-luminosity running.

3.2 The Compact Muon Solenoid Experiment

The Compact Muon Solenoid (CMS) experiment [9] is a general-purpose detector that is capable of identifying electrons, muons, photons, and hadrons [16, 17, 18]. The CMS detector is composed of a tracking system, electromagnetic calorimeter

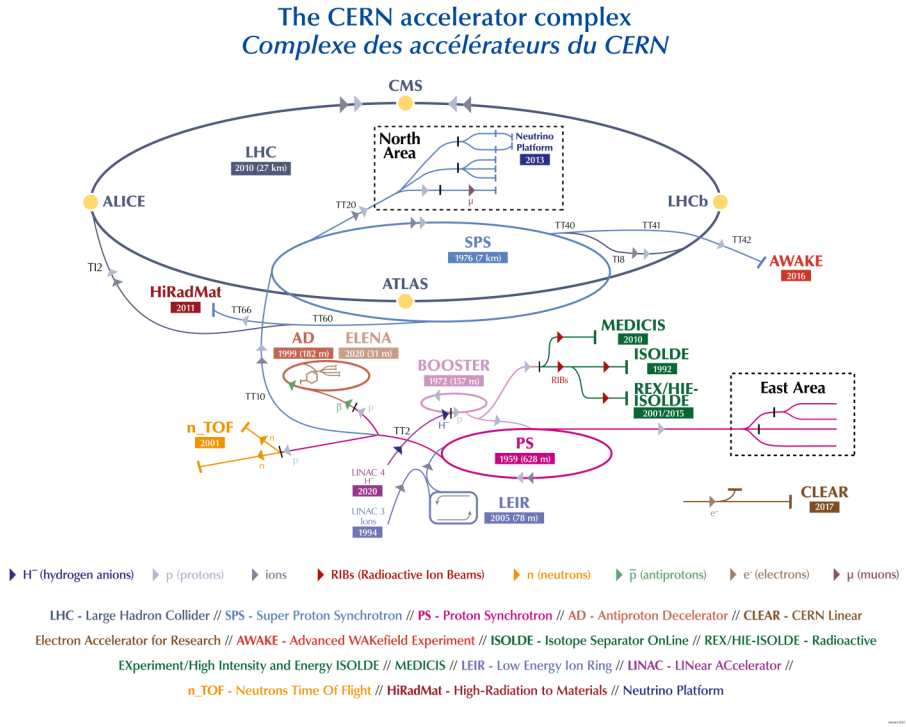

Figure 3.2: The diagram that shows the CERN accelerator complex [15].

(ECAL), hadron calorimeter (HCAL), superconducting solenoid, and muon system. The 3D sketch of the CMS detector is shown in Fig. 3.3.

The superconducting solenoid is 13m long with an inner diameter of 6m so it provides enough space to include the tracking system, ECAL, and HCAL inside the solenoid. The solenoid provides a 3.8T magnetic field to apply a large bending power on charged particles. The trajectory of the bent charged particles will be measured by the tracking system, which is composed of silicon pixel layers and silicon strip layers. Particles that pass through the tracking system will then deposit their energy in the ECAL, which is a homogeneous crystal calorimeter made of lead tungstate ($PbWO_4$) crystals, and HCAL, which is a sampling calorimeter made of brass and scintillator. The materials in the calorimeters, such as $PbWO_4$ crystal and brass, are dense enough so that incoming particles will interact with the material, creating sec-

ondary particles and forming “showers”. The showering process deposits the energy carried by the incoming particles inside the calorimeters. Most particles deposit all of their energy in the calorimeters and thus stop traveling further in the CMS detector. However, muons travel further beyond the calorimeters because they are typically minimum ionizing particles, which do not lose much energy when traveling through materials. To have a better measurement of muons, a muon system is built beyond the superconducting solenoid, based on the returning magnetic field of the solenoid. The muon system is composed of aluminum drift tubes (DT), cathode strip chambers (CSC), and resistive plate chambers (RPC).

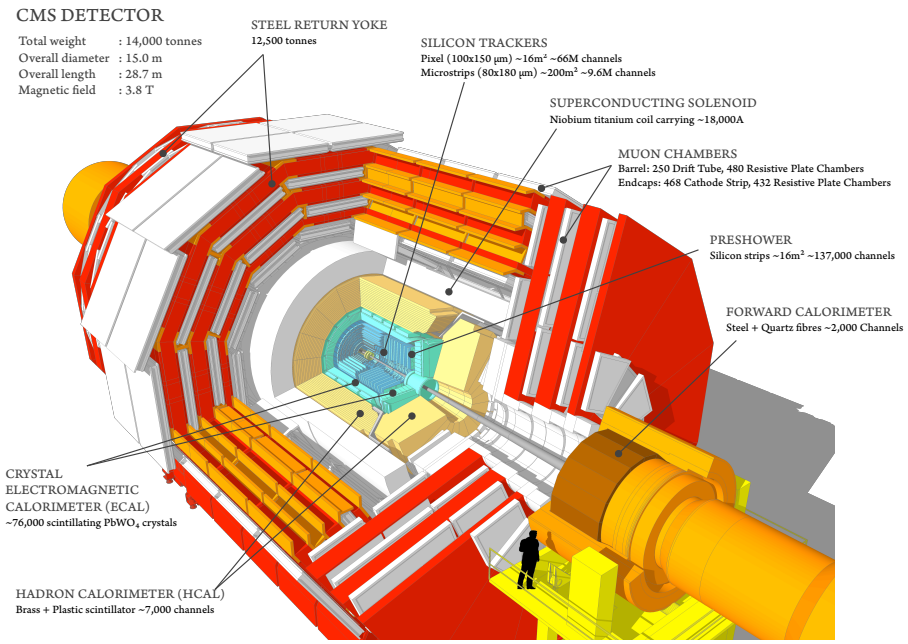

Figure 3.3: The 3D image of the CMS detector [19].

3.2.1 Coordinate System

The CMS uses a right-handed coordinate system. The coordinate system is centered at the nominal collision point at the center of the detector. The x-axis points from the

interaction point to the center of the LHC ring. The y-axis points upwards vertically. The z-axis is oriented along the direction of the beam, with the positive direction defined as counterclockwise when viewed from above. The radial distance in the x-y plane is defined as r. The azimuthal angle ϕ is calculated as the angle from the x-axis in the x-y plane and the polar angle θ is calculated as the angle from the z-axis. Pseudorapidity is then defined based on θ as $\eta = -\ln[\tan(\theta/2])$.

3.2.2 Tracking System

The tracking system of the CMS experiment is designed to provide accurate measurements of the trajectories of charged particles with $|\eta| < 2.5$. A sketch of the cross-section of the CMS tracking system is shown in Fig. 3.4. The CMS tracking system is comprised of two main parts: the silicon pixel tracker and the silicon strip tracker.

The silicon pixel tracker uses n-on-n type silicon sensors as the basic sensor elements while the silicon strip tracker uses single-sided p-on-n type silicon micro-strip sensors as the basic sensor elements. The pixel tracker includes three cylindrical layers of pixel modules as the barrel part, distributed at radii of 4.4, 7.3, and 10.2 cm surrounding the interaction point, and four disks of pixel modules as the endcap part, with two disks on each side of the barrel.

The silicon strip tracker is composed of the Tracker Inner Barrel and Disks (TIB/TID), the Tracker Outer Barrel (TOB), and the Tracker EndCap (TEC). The silicon strip tracker fills the radial region between 20 cm and 116 cm. The TIB/TID includes four barrel layers in the radii range from 20 cm to 55 cm with three disks at each end. The TOB is installed beyond the TIB/TID. It is composed of six barrel layers, cov-

ering the radial space up to 116 cm in the $|z| < 118$ cm region. On the endcap, the TEC, which is composed of nine disks on each side of the barrel, covers the region of 124 cm $< |z| <$ 282 cm and 22.5 cm $< |r| <$ 113.5 cm.

The CMS silicon pixel tracker underwent an upgrade [20] during the year-end technical stop of LHC from December 2016 to April 2017. The goal of the upgrade was to make the CMS pixel tracker perform better for higher rate capability, to increase its tolerance to radioactive damage, and to provide more robust tracking. To achieve the goal, more layers are utilized in the upgraded pixel tracker. It is composed of four concentric barrel layers, distributed at the radii of 29, 68, 109, and 160 mm, and three disks on both ends of the barrel, with a distance of 291, 396, and 516 mm from the center of the detector. A sketch that compares the original and upgraded pixel tracker is shown in Fig. 3.5.

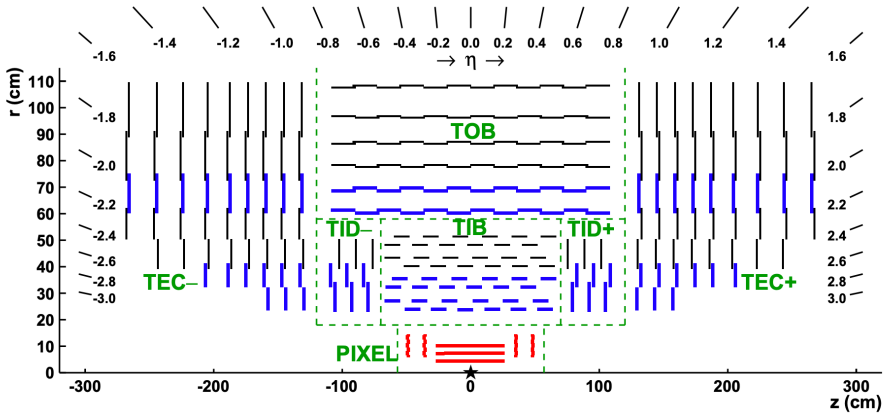

Figure 3.4: The sketch of cross section of the CMS tracking system [18].

3.2.3 Electromagnetic Calorimeter

The CMS electromagnetic calorimeter (ECAL) is designed to measure the energy of particles that interact mainly by the electromagnetic interaction. As a hermetic

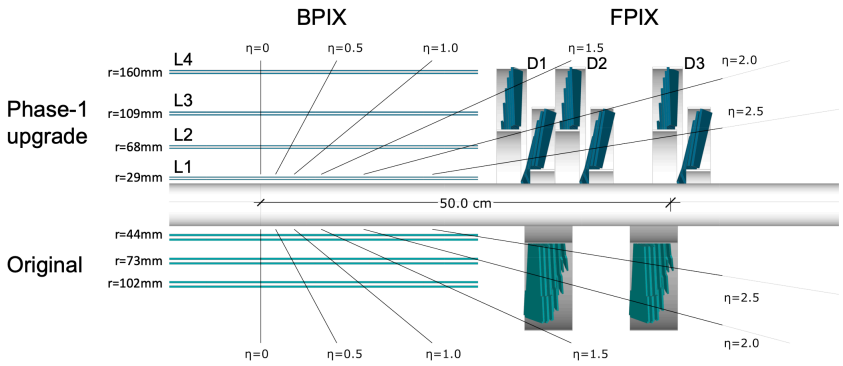

Figure 3.5: Comparison of the CMS pixel tracker layout before and after the upgrade [20]. The BPIX stands for the barrel pixel layers and the FPIX stands for the endcap pixel layers.

homogeneous calorimeter, the ECAL is composed of 61200 lead tungstate ($PbWO_4$) crystals in the barrel part and 7324 crystals in each endcap. The characteristics of the $PbWO_4$ crystals, as shown in Tab. 3.1, make the ECAL fast, have fine granularity, and are tolerant to radiation damage. The high density and small radiation length (X_0) make the ECAL compact; the low Molière radius makes it possible for the ECAL to have fine granularity. The layout of the ECAL is shown in Fig. 3.6.

Table 3.1: Characteristics of the lead tungstate crystals.

Characteristics	value
Density	8.28 g/cm^3
Radiation length	0.89 cm
Molière radius	2.2 cm
Scintillation decay time	within 25ns for 80% of the scintillation light

In the barrel part of the ECAL, there are 360 crystals in ϕ and (2×85) crystals in η, covering the space of $|\eta| < 1.479$ with 61200 crystals in total. The crystals are installed to make their axes have a 3° angle with the vector from the interacting point. A tapering form is designed for the crystal, with a cross-section of $22 \times 22\,\text{mm}^2$ at the front face and $26 \times 26\,\text{mm}^2$ at the rare face. The crystal is 230 mm long, which corresponds to 25.8 X_0. The photodetectors used in the barrel part are the modified

avalanche photodiodes (APDs). Each pair of APDs form a capsule and is attached to a crystal. The APDs operate at a voltage between 340 and 430 V, with a mean gain of 50.

In the endcap part of the ECAL, crystals are grouped into supercrystals (SCs), each with 5×5 crystals arranged in a rectangular grid in the x-y plane. The endcap part consists of 7324 crystals on each side, covering the region of $1.479 < |\eta| < 3.0$. The crystals have a similar shape as those in the barrel part, with the cross-section at the front face of $28.62 \times 28.62\,\mathrm{mm}^2$ and $30 \times 30\,\mathrm{mm}^2$ at the rare face. The length of the crystal is 220 mm, corresponding to 24.7 X_0. The photodetectors used in the endcap part are vacuum phototriodes (VPTs). Each VPT is attached to a crystal. The VPTs operate at a gain of 10 by biasing the dynode with 600V and the anode with 800V.

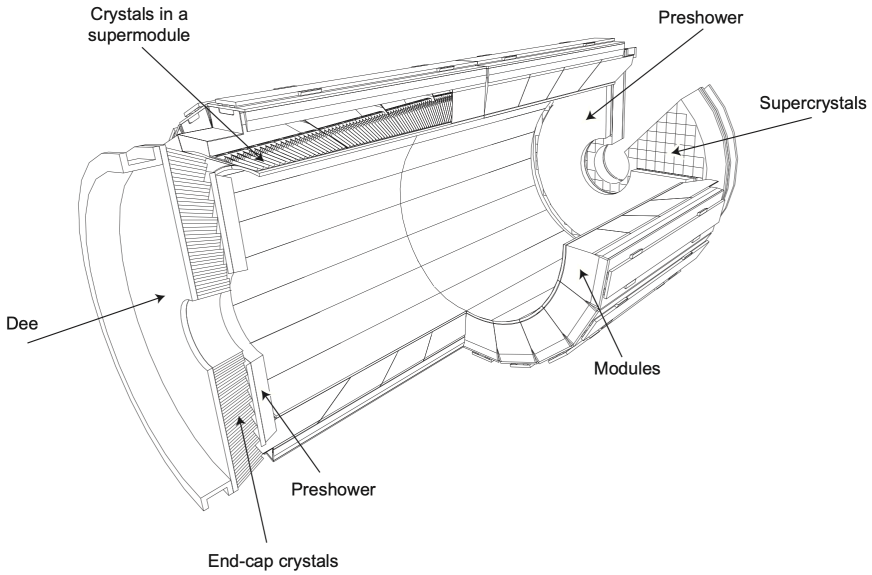

Figure 3.6: The layout of the CMS ECAL [9]. A Dee corresponds to half of the endcap part of the ECAL. The preshower serves to identify neutral pions in the region $1.653 < |\eta| < 2.6$ and identify electrons out of minimum ionizing particles.

3.2.4 Hadron Calorimeter

The CMS hadron calorimeter (HCAL) is composed of four parts: the hadron barrel (HB), endcap (HE), outer (HO), and forward (HF) calorimeters. Figure 3.7 shows the longitudinal view of the CMS detector, with different parts of HCAL labeled.

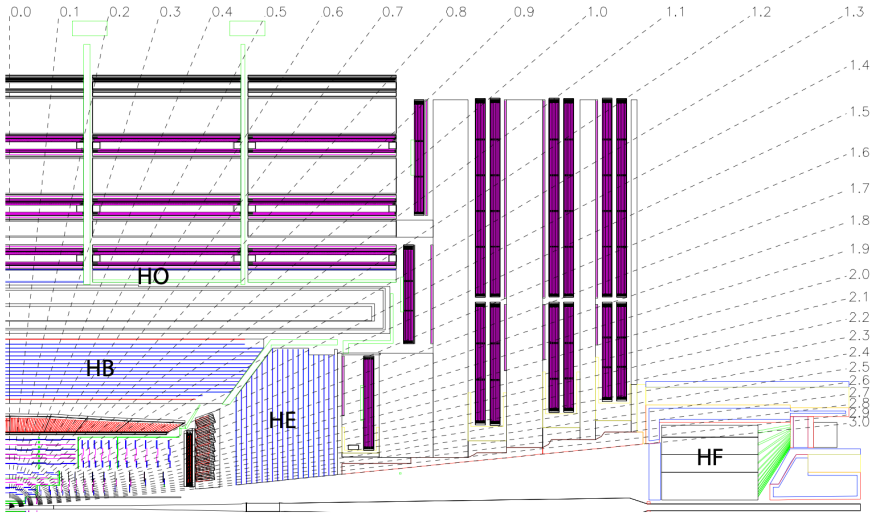

Figure 3.7: The longitudinal view of CMS detector [9]. Positions of different parts of the CMS HCAL are labeled on the diagram.

The HB, HE, and HO are sampling calorimeters, which are composed of alternating layers of absorbers and scintillators. Wavelength shifting fibers and clear fibers are used to guide the scintillated light into the hybrid photodiode (HPD).

The HB, which uses steel and brass as absorbers and tiles as scintillators, covers the region $|\eta| < 1.3$. It has 36 wedges, with 18 wedges distributed in the ϕ direction and 2 wedges in the η direction for each ϕ segment. Each wedge has four segments in ϕ direction and 16 segments in η direction, thus forming a $\Delta\eta \times \Delta\phi \approx 0.087 \times 0.087$ granularity.

The HE shares a similar design to the HB, with brass as absorbers and tiles as

scintillators. It covers the region of $1.3 < |\eta| < 3$ with a granularity of $\Delta\eta \times \Delta\phi \approx 0.087 \times 0.087$ for $|\eta| < 1.6$ and $\Delta\eta \times \Delta\phi \approx 0.17 \times 0.17$ for $|\eta| \geq 1.6$.

The HO, which is designed to measure showers that start late or are not fully deposited in the HB, is placed outside the solenoid. Iron together with the solenoid is used as absorbers and tiles are used as scintillators. The layout of the HO depends on the structure of the iron yoke that returns the magnetic field. The iron yoke is composed of five 2.536 m wide rings, labeled as -2, -1, 0, +1, +2, which correspond to their central z position at -5.342 m, -2.686 m, 0, +2.686 m, +5.342 m. For each ring, the HO is installed as the first sensitive layer. Given the fact that the HB absorber length increase with the increase of η, the HB has the smallest absorber length around the central region with $\eta = 0$. To account for such an effect, the HO has two layers on the central ring 0, distributed before and after the iron absorber. For all the rest of the rings, the HO has only one layer located after the iron absorber.

The HF covers the region $3 < |\eta| < 5$. The high particle flux in that region drives the use of quartz fibers as the active medium to make the HF able to maintain the desired performance under such high radiation. The HF is designed differently compared with other parts of HCAL, the quartz fibers are embedded inside the steel absorber, to capture the Cherenkov light generated by the charged particles. The captured light is read out by photomultiplier tubes. Furthermore, the HF is divided into two sections in the longitudinal direction to distinguish showers generated via electromagnetic and strong nuclear interactions.

3.2.5 Muon System

As indicated by the name of the CMS experiment, measuring muons precisely is one of the major tasks of the CMS detector. The muon system is expected to have good performance on muon identification, momentum measurement, and triggering.

The muons system is composed of three gaseous subdetectors: drift tube (DT) chambers, cathode strip chambers (CSC), and resistive plate chambers (RPC). Determined by the structure of the solenoid, the muon system has one barrel part and two endcap parts.

The DT is installed in the barrel part of the muon system, covering $|\eta| < 1.2$. It is organized into four muon stations, which are concentric cylinders centered at the beam line along the z-axis. The three inner cylinders hold 60 drift chambers while the outer cylinder holds 70 drift chambers. Each chamber consists of groups of layers of rectangular drift cells, which measure the muon tracks in the $r - \phi$ plane or the z positions.

The CSC resides in the endcap part of the muon system and covers the region $0.9 < |\eta| < 2.4$. The CSC chambers are made using seven cathode panels alternating with six anode wire planes. The cathode panels have strips along the r direction in the $x-y$ plane and the anode wire planes have wires along the ϕ direction. The intersection of the strips and wires provides position measurements in the $r - \phi$ plane.

To identify the corresponding bunch crossing a muon track is associated given the high rate and background, the RPC, which is a muon detector capable of precise time of arrival measurement. The RPC is distributed in both the barrel and endcap parts of the muon system, covering the region of η up to 1.6. In the barrel part, the RPC chambers are distributed in the four muons stations, located on the inner side of the

DT chambers or the inner and outer sides of the DT chambers, depending on the stations. In the endcap, the RPC chambers are distributed on the endcap disk.

3.2.6 Trigger

The LHC collides proton bunches with an interval of 25 ns, corresponding to an event rate of 40 MHz. However, the resulting amount of data is too large for every channel from every collision to be stored and processed. To reduce the event rate to an acceptable level, triggers are applied. The triggers are expected to have high efficiencies to physics objects, inclusive selections to include unexpected new phenomena, and satisfy the data recording technology. [21] CMS uses a two-step trigger system, referred to as Level-1 Trigger (L1T) and High-Level Trigger (HLT).

The L1T output limit is 100kHz. It is mostly based on customized and programmable electronics, such as Field Programmable Gate Arrays (FPGA), Application-specific Integrated Circuit (ASIC), and programmable memory lookup tables (LUT). The architecture of the L1T is shown in Fig. 3.8. The L1T is composed of local, regional, and global elements. The L1T takes coarse segmented information from the muon system and calorimeters. The hits, track segments, and energy deposits in the muon system and calorimeters are determined and combined to form the preliminary reconstruction of objects. The Global Trigger takes the reconstructed objects and determines whether to accept the events for further processing by the HLT. The L1T has to deliver the trigger decision within 3.2 μs to be able to process all the events delivered by the LHC in time.

Events that are accepted by the L1T are further processed by the HLT [21], to bring down the event rate to $O(100)$ Hz. Different from the L1T, the HLT is based on

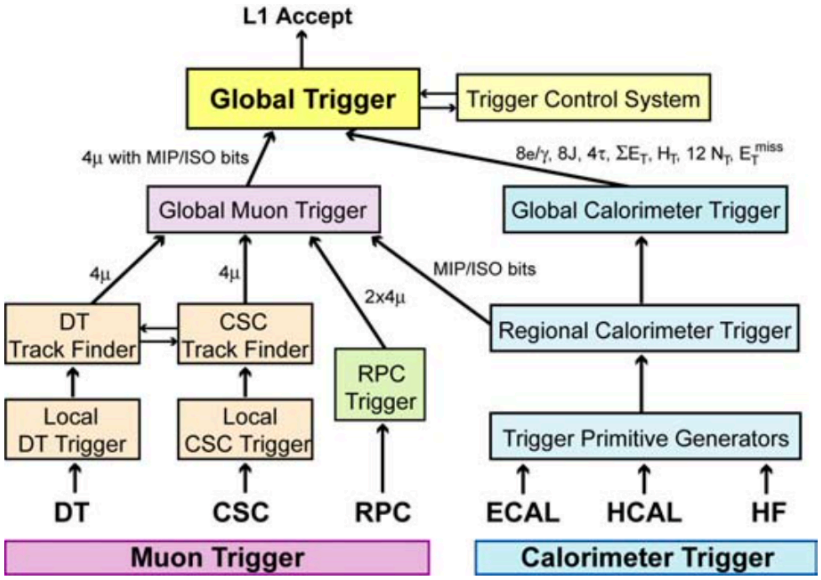

Figure 3.8: The CMS Level-1 Trigger architecture [9].

software with the complete information read out from the electronics. Those features make HLT capable to reconstruct objects more precisely. Also, since the HLT depends mainly on software, it can be updated with offline reconstruction algorithms to achieve better efficiencies as the reconstruction algorithms improve.

3.2.7 Particle Reconstruction

Most particles are expected to be detected by more than one CMS subdetectors. A longitudinal view of the CMS detector with the responses of subdetectors to different particles is shown in Fig. 3.9.

The CMS experiment uses a particle-flow (PF) algorithm to reconstruct physics objects. The PF algorithm correlates information from all subdetectors, including charged-particle trajectory measured by the tracker, clusters of energy deposits mea-

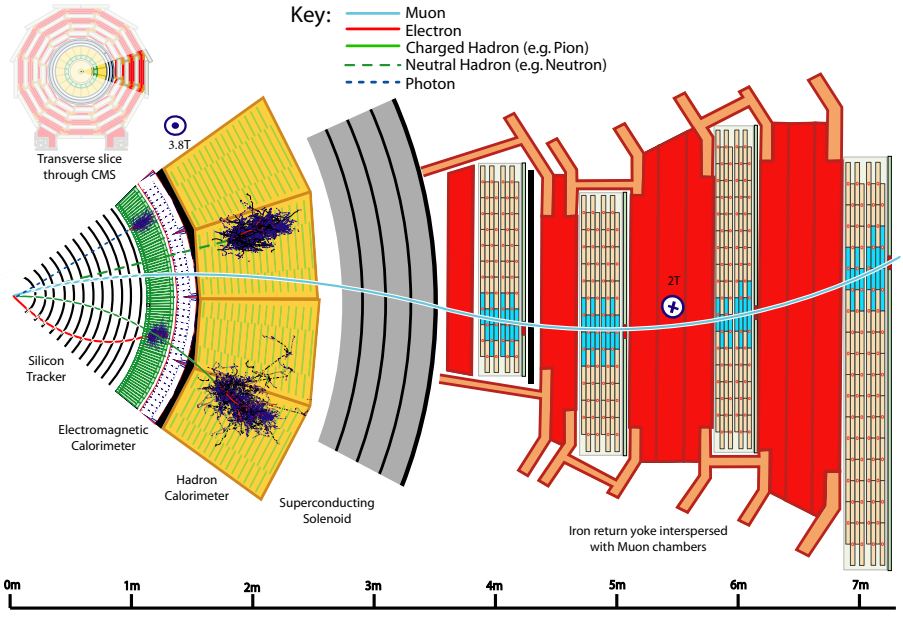

Figure 3.9: The longitudinal view of the CMS detector [22].

sured by the ECAL and HCAL, and measurements from the muon system, to reconstruct every final-state particle. The advantage of the PF algorithms is that the combined information improves the reconstruction significantly compared with the traditional way of reconstruction, which processes information individually for different subdetectors. CMS is suitable to utilize the PF algorithm because of its fine granularity, which enables good separation between different particles to combine information from all subdetectors for individual particles.

The PF algorithm is based on charged-particle trajectories (*tracks*) and calorimeter clusters. The tracks are reconstructed by applying the track finder algorithm on signals (*hits*) measured from the tracking system in several iterations. Each iteration is composed of three major steps: seed trajectory generation with a few hits; pattern recognition, which gathers hits along the seed trajectory from all layers in the tracking system; final fitting to evaluate the characteristics of the charged particle. The track

finder starts with iterations that look for tracks with at least three hits in the pixel layers, followed by iterations that recover tracks that have only one or two hits in the pixel to account for the detector inefficiencies. In the next iterations, displaced tracks, which have no hits measured in the pixel layers, and tracks in the dense core of high-p_{T} jets, which may have several tracks merged as one single track because of the inaccurate position measurement, are reconstructed. Special treatments are applied for tracks originating from electrons and muons. To account for the energy radiated by electrons, a Gaussian-sum filter (GSF) is used to fit the electron tracks and information from the tracking system and ECAL is combined to reconstruct electrons. Muon tracks are reconstructed using the measurements from both the tracking system and the muon system.

The calorimeter clusters are reconstructed separately in different subdetectors. Calorimeter cells with energy above the seed threshold and higher than the energy of all the neighbor cells are used as cluster seeds. Based on the seeds, clusters are grown by merging cells that have energy above the threshold and close to the seeds. A Gaussian-mixture model is applied to the clusters to reconstruct the clusters. Calibrations are applied to obtain accurate measurements of the energy deposited by neutral particles.

Basic elements, namely tracks and clusters, are connected together to form the PF blocks with the *link algorithm* to be further processed in the PF to reconstruct the events. A track-to-cluster link is established when the extrapolated position of the track is within the region of a cluster, or the extrapolated positions of tangents to the GSF tracks fall in the region of an ECAL cluster to account for the photons emitted by electrons. A cluster-to-cluster link is established between different calorimeters when a cluster from the fine granularity calorimeter is within the region of a cluster

from the coarse granularity calorimeter. Track-to-track links are formed by secondary vertices to account for nuclear interactions. Tracks are also linked with measurements from muon detectors for muon reconstruction.

For a given PF block, the PF algorithm identifies and reconstructs different particles in steps. Elements are masked out after each step so no element is reused in the reconstruction. Muons are first reconstructed, using tracks or track segments from both the tracker and the muon detector. Electrons are reconstructed with the information provided by the tracker and the calorimeters, especially the GSF tracks and ECAL clusters. Isolated photons are reconstructed with ECAL clusters without links to GSF tracks. Hadrons and nonisolated photons are then reconstructed using the remaining elements. Charged hadrons, neutral hadrons, and nonisolated photons are reconstructed based on whether the cluster is linked with a track or has compatible energy with the linked track momentum. Hadrons that undergo the material interactions are reconstructed from the selected nuclear-interaction vertices. A post-processing step is performed after all the blocks are processed and all the particles are reconstructed to reduce the possibility of mis-reconstructed $p_{\mathrm{T}}^{\mathrm{miss}}$.

Chapter 4

Long-Lived Particles and Displaced Vertices

Most of the searches performed in experiments at the LHC target particles that decay promptly. As shown in Fig. 4.1, particles in the SM cover a wide range of lifetimes. However, the lack of new discoveries in recent years motivates new searches to study unexplored phase space, such as Long-Lived Particles (LLPs) [23]. As opposed to prompt decaying particles, LLPs have relatively long lifetimes.

Additionally, many BSM models, such as the split SUSY described in Sec. 2.2.1, predict new LLPs, which are likely to leave unique signatures in the detector, referred to as displaced signatures, as shown in Fig. 4.2. Since the signatures of LLPs from the SM are already studied and understood, the unique signatures of LLPs in the BSM models can be very powerful handles to discover new physics. However, since the experiments at the LHC are designed to be sensitive to prompt particles, the detector structure and the reconstruction algorithm are optimized for prompt signatures. That requires careful treatment to efficiently reconstruct the displaced signatures.

The search presented in this thesis targets displaced vertices. Displaced vertices, as one of the displaced signatures that can be produced by BSM LLPs, are a powerful strategy of looking for new physics. When an LLP is produced in the proton-proton collision, if it travels a detectable distance, decays within the tracker volume, and has

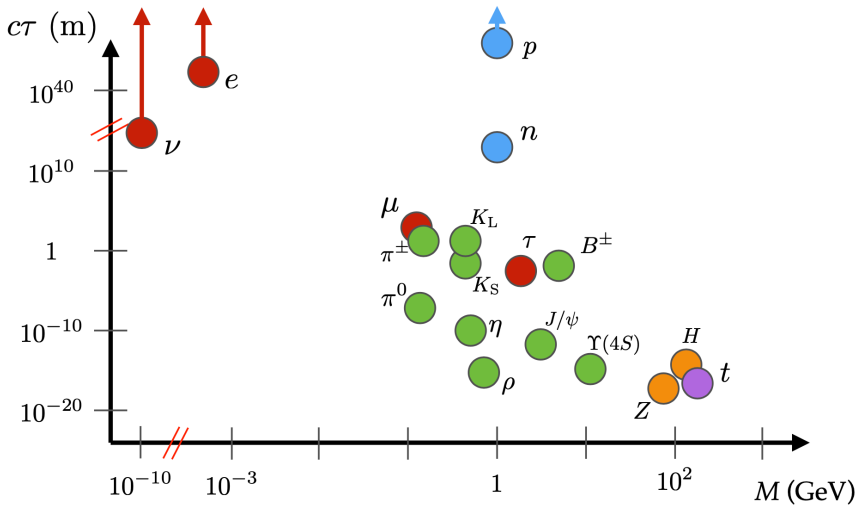

Figure 4.1: Masses of lifetimes of particles in the SM [23]. Different colors represent different types of particles: red corresponds to leptons, blue corresponds to hadrons, green corresponds to mesons, yellow corresponds to bosons, and purple corresponds to quarks.

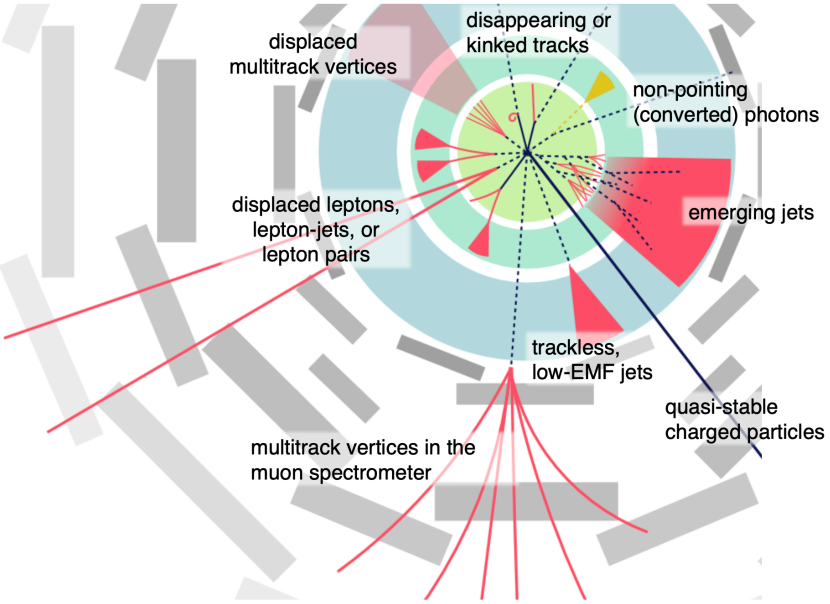

Figure 4.2: Signatures that LLPs predicted by BSM theories can create in detectors [23].

at least part of the decay products as SM particles that are "visible" to the detector, its "visible" decay products will be measured as tracks by the tracker, which converge at the position where the LLP decays. The LLP decay point will then be recognized as a displaced vertex. The signature of displaced vertices is illustrated in Fig. 4.3.

The ATLAS and CMS Collaboration have previously performed searches for LLPs decaying to displaced vertices [24, 25, 26, 27]. The searches performed by the ATLAS Collaboration target LLPs that decay at least 1 mm away from the interaction point, while the searches performed by the CMS Collaboration target heavy LLPs that decay to more energetic final states with a mean proper decay length of at least 100 μm.

The search presented in the thesis goes beyond the previous CMS search for displaced vertices [26] by targeting displaced vertices with transverse missing momentum. It is designed to cover the gaps left by previous searches, which include a wider range of LLP $c\tau$, from 0.1 mm to 1000 mm, and LLPs in "compressed" scenarios, which have most of the energy carried away by "invisible" particles and only 100–200 GeV energy carried by particles that can be detected by the CMS detector.

Given the fact that LLPs are predicted by a wide range of BSM models and the similarity of LLP signatures from different models, it is helpful to design model-independent searches that target a given signature instead of a specific BSM model. Doing so increases the chance of discovering new physics. If the search does not show evidence of new physics, the results can be interpreted into different BSM models to check whether a given model is excluded by the search. The search presented in this thesis is designed to be model-independent to be sensitive to a wide range of potential BSM models that could give rise to the specific signature that the search targets.

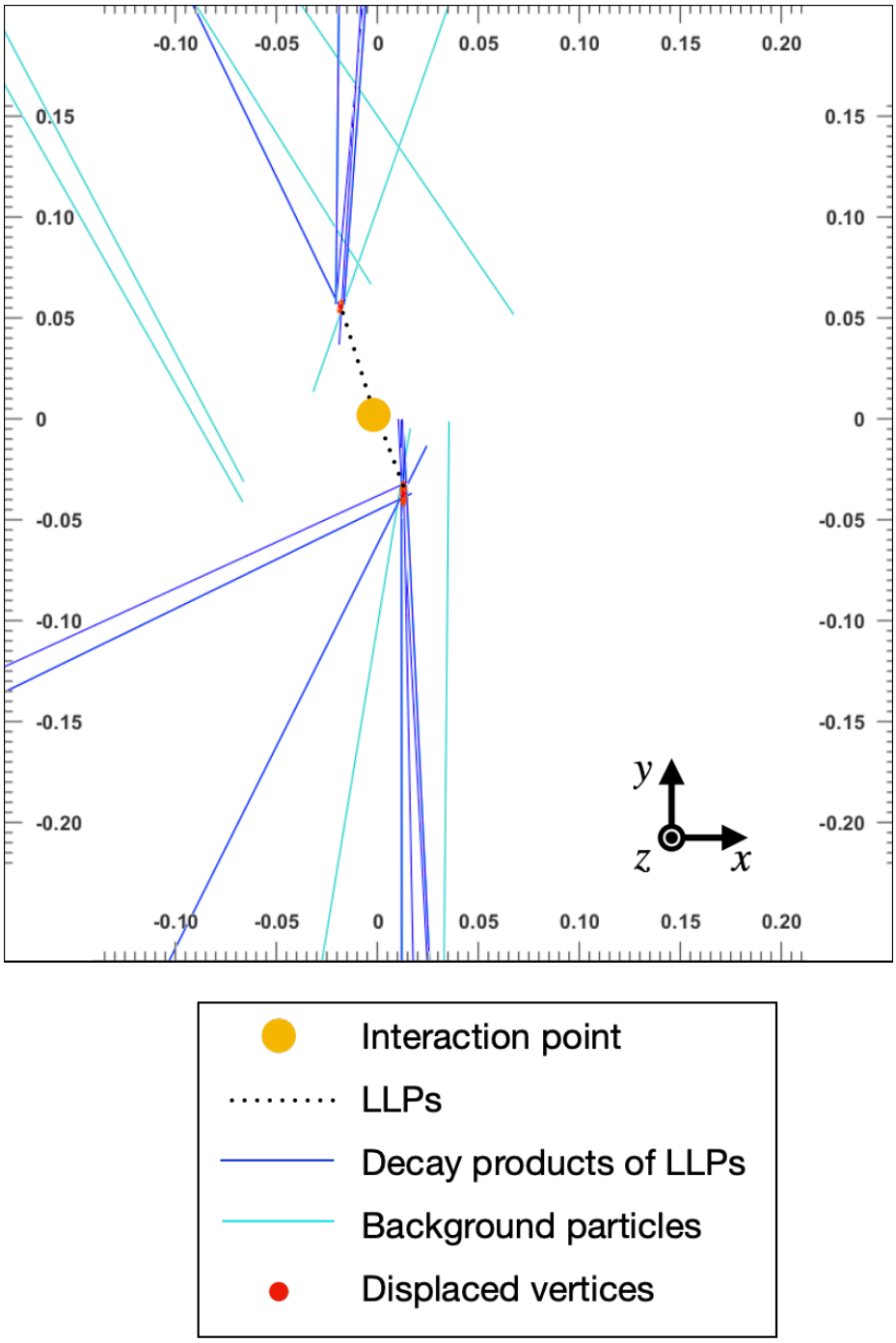

Figure 4.3: One collision with displaced vertices in the detector. A transverse view, with the beam axis (z axis) perpendicular to the paper, is shown. The protons collide at the interaction point (yellow point). Two LLPs (black dashed lines) are produced and travel away from the interaction point. The LLPs then decay at the displaced vertices (red points). The "visible" decay products (blue lines) are measured by the detector. Also present are background particles (cyan lines) that overlap with the decay of LLPs.

Chapter 5

Data Sets and MC Simulation

5.1 Data Sets

This analysis uses proton-proton collision data taken by the CMS experiment in 2016, 2017, and 2018 at a center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ and with a total integrated luminosity of $137\,\text{fb}^{-1}$. Due to different detector conditions, data taken in different years are analyzed individually. Data taken in the same year are split into different eras. In 2016, the data are further divided into two parts are analyzed individually:

- 2016 preVFP: reconstructed with a reconstruction including HIP mitigation (eras B-F, with 31 runs coming from F);

- 2016 postVFP: reconstructed with the default track reconstruction (eras F-H, with 7 runs coming from F).

Data events are triggered by the missing transverse momentum $p_{\text{T}}^{\text{miss}}$, as described in Sec. 6.2. Table 5.1 lists the datasets used in this analysis, with definitions of the data-taking eras. The data events are reconstructed using the most up-to-date reconstruction algorithm, labeled as UltraLegacy (UL).

Table 5.1: Datasets used in this analysis, with the "/MiniAOD" suffix dropped.

Era	Dataset
2016B-v1	/MET/Run2016B-ver1_HIPM_UL2016_MiniAODv2-v2
2016B-v2	/MET/Run2016B-ver2_HIPM_UL2016_MiniAODv2-v2
2016C	/MET/Run2016C-HIPM_UL2016_MiniAODv2-v2
2016D	/MET/Run2016D-HIPM_UL2016_MiniAODv2-v2
2016E	/MET/Run2016E-HIPM_UL2016_MiniAODv2-v2
2016F	/MET/Run2016F-HIPM_UL2016_MiniAODv2-v2
2016F	/MET/Run2016F-UL2016_MiniAODv2-v2
2016G	/MET/Run2016G-UL2016_MiniAODv2-v2
2016H	/MET/Run2016H-UL2016_MiniAODv2-v2
2017B	/MET/Run2017B-UL2017_MiniAODv2-v1
2017C	/MET/Run2017C-UL2017_MiniAODv2-v1
2017D	/MET/Run2017D-UL2017_MiniAODv2-v1
2017E	/MET/Run2017E-UL2017_MiniAODv2-v1
2017F	/MET/Run2017F-UL2017_MiniAODv2-v1
2018A	/MET/Run2018A-UL2018_MiniAODv2-v2
2018B	/MET/Run2018B-UL2018_MiniAODv2-v2
2018C	/MET/Run2018C-UL2018_MiniAODv2-v1
2018D	/MET/Run2018D-UL2018_MiniAODv2-v1

5.2 Simulated Samples

Monte Carlo simulated signal and background events are used to design the search strategy, understand the behavior of signal and background events, and interpret the search results. The simulated events are referred to as "MC" or "simulation" in the remainder of this thesis.

Details of the generation of signal and background events are described below. For both signal and background events, a GEANT4-based[28] package is used to simulate the CMS detector response to generated events. To reproduce the pileup distribution in data, simulated minimum-bias events are superimposed on the hard interaction in simulated events.

5.2.1 Signal

As described in Sec. 2.2, the benchmark signal model used in this search is the pair-produced long-lived gluinos in the split SUSY model.

Using PYTHIA 8.240 [29], signal samples with varying gluino and neutralino masses and gluino $c\tau$ values were produced. The generated mass and $c\tau$ ranges were chosen to target the parameter space that is not excluded by previous prompt or non-prompt searches and focus mainly on compressed split SUSY scenarios. The analysis sensitivity decreases at the low and high gluino lifetimes due to several analysis requirements that target displaced objects and the requirement that displaced vertices are within the beam pipe, respectively. Despite the beam pipe constraint, the search remains sensitive to models with $c\tau$'s up to 1000 mm because particle decay lengths follow an exponentially falling distribution. As a result, the produced masses and $c\tau$'s are summarized in Tab. 5.2.

Table 5.2: Mass of gluino, neutralino, and $c\tau$ of gluino produced. Two neutralino masses are produced for each gluino mass, as shown in the center column. For each gluino and neutralino mass, nine gluino $c\tau$'s are generated, as shown in the right-hand column.

Gluino mass	Neutralino mass	Gluino $c\tau$
1400	(1200, 1300)	
1600	(1400, 1500)	100 μm, 300 μm,
1800	(1600, 1500)	1 mm, 3 mm,
2000	(1800, 1700)	10 mm, 30 mm,
2200	(2000, 1900)	100 mm, 300 mm,
2400	(2200, 2100)	1000 mm
2600	(2400, 2300)	

The cross sections of signal models are determined by gluino-gluino production cross-sections computed at next-to-next-leading-order (NNLO) + next-to-next-leading-logarithm (NNLL) [30] and it is not correlated with gluino lifetime and neutralino mass.

Signal simulations are generated with the NNPDF3.1LO [31] set of parton distribution functions (PDFs) using the CP2 tune [32]. In these generated samples, gluinos are produced in pairs; each gluino is forced to undergo a three-body decay to two (anti-)quarks and a neutralino. The (anti-)quarks that the gluinos decay to are light quarks including u, d, c, and s quarks, with equal branching fraction of 25%. When a colored SUSY particle is long-lived, it could have time to hadronize into an R-hadron, which can lead to material interactions and alter the kinematics of the LLP decay products. Since this search targets displaced vertices within the beam pipe, material interaction is not a concern. The production and decay of R-hadrons are allowed in the simulation, so the effect of R-hadrons on LLP decay products is taken into account.

5.2.2 Background

The dominant background processes include pair production of top quarks and events with jets produced by the strong interaction. Besides the dominant processes, W or Z boson produced in association with jets from the strong interaction, single top quark production, and di-boson production such as $W\ W$, $W\ Z$, and $Z\ Z$ also contribute to background events in this search. The pair-production of top quarks, events with jets produced through the strong interaction with or without W and Z bosons, are generated using MadGraph5_aMC@NLO 2.4.2 [33] at next-to-leading-order (NLO) in quantum chromodynamics (QCD). To match jets from the matrix-element calculation with parton showers, the MLM prescription [34] or FxFx approach [35] are applied. The single top quark production process is generated using POWHEG [36, 37] at NLO in QCD and the di-boson processes are generated using PYTHIA 8. The NNPDF3.1 [38] is used to describe the proton contents in event generation. PYTHIA 8 with the CP5

tune [32] is used to simulate the hadronization and showering.

Different background processes, as described above, are combined together and labeled as background simulation or background MC in studies presented in this thesis.

Chapter 6

Event Selection

6.1 Object Selection

This section defines the selection and reconstruction of objects that are used in the search. The selection and reconstruction of the secondary vertices that serve as the main analysis objects are described in Sec. 7. Data and simulated events are processed using the standard algorithms, provided in the CMS Software CMSSW 10_6_29 for 2016, 2017, and 2018.

Primary vertices are required to consist of at least four tracks, have $|r| < 2$ cm, and $|z| < 24$ cm to be considered as successfully reconstructed. The leading primary vertex is the primary vertex that has the largest summed p_{T}^2 of physics objects associated with it.

Jets are clusters of collimated particles that are likely to come from the hadronization of quarks or gluons. They are formed by clustering PF candidates within a cone with $\Delta R = 0.4$ using the anti-k_{T} algorithm [39, 40]. To avoid including particles from pileup interactions in jets, charged PF candidates are removed if their associated tracks come from pileup primary vertices. The standard set of jet energy corrections (JEC), including L1FastJet, L2Relative, L3Absolute, and L2L3Residual [41] if in data, is applied. PF jets are required to satisfy the "TightLepVeto" PF jet id provided

in the CMSSW, which consists of the following requirements:

- $p_{\mathrm{T}} > 20$ GeV;
- $|\eta| < 2.5$;
- Number of constituents > 1;
- Neutral hadron energy fraction < 0.9;
- Neutral EM energy fraction < 0.9;
- Muon energy fraction < 0.8;
- For $|\eta| < 2.4$:
  - Charged hadron energy fraction > 0;
  - Charged multiplicity > 0;
  - Charged EM energy fraction < 0.80.

To identify heavy-flavor jets, we use the DeepJet/DeepFlavour neural network-based tagger [42] developed by the CMS B-Tagging and Vertexing group to discriminate jets originated from b quarks and light quarks. The jet tagger has several work points, namely "loose", "medium", and "tight", to achieve the mis-tag rate of 10%, 1%, and 0.1%. In addition to the neural network-based jet tagger, jets are required to have p_{T} $> 20\,\mathrm{GeV}$ and $|\eta| < 2.5$ to be considered as heavy-flavor jets. Heavy-flavor jets are not used in the search selection but are used in the studies to understand background behaviors and signal efficiency.

Missing transverse momentum ($\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$), also known as missing transverse energy ($\vec{E}_{\mathrm{T}}^{\,\mathrm{miss}}$ or MET) is calculated as the negative sum of the p_{T} of all reconstructed PF candidates

that originate from the primary vertex in an event. As described above, jets receive the JEC, which could have an effect on the $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$. The Type-1 corrections are applied on the $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ to propagate the JEC to $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$. Also, an xy-shift correction is applied to the $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ to account for the effect of detector misalignment. To reject events with large $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ from sources such as detector noise, a set of $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ filters are applied. In addition to the $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ filters, a fake MET veto, which vetoes events if they have $|MET_{PF} - MET_{CALO}|/MET_{CALO} > 0.6$, is applied to reduce mismeasured $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ originating from the calorimeter.

$\vec{E}_{\mathrm{T\ NoMu}}^{\mathrm{miss}}$ is the negative vector sum of all reconstructed physics objects in an event, except muons. It is calculated by adding the x and y components of muons back into $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$.

A per-event weight is applied to simulated events to ensure consistent pileup distributions between the simulation and data. The pileup weights are calculated as the ratio of pileup distribution between data and the simulation events before any selections are applied. The weights are derived individually for 2016, 2017, and 2018.

6.2 Trigger Selection

Since the search targets the signature of displaced vertices with $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$, triggers relevant to $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ are a natural choice. The trigger used in this analysis is:

- `HLT_PFMETNoMu120_PFMHTNoMu120_IDTight,`

with an online $\vec{E}_{\mathrm{T\ NoMu}}^{\mathrm{miss}}$ threshold at 120 GeV.

Specifically, for the split SUSY benchmark signal samples, the neutralinos produced

by the decay of gluinos, which cannot be detected by the detector, are expected to be the source of $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$. Figure 6.1 shows the $\vec{E}_{\mathrm{T\ NoMu}}^{\mathrm{miss}}$ distribution of split SUSY. The plots show that split SUSY samples with 200 GeV mass splitting are more likely to have large $\vec{E}_{\mathrm{T\ NoMu}}^{\mathrm{miss}}$ compared with 100 GeV mass splitting.

Besides the $\vec{E}_{\mathrm{T\ NoMu}}^{\mathrm{miss}}$ distribution of signal samples, trigger efficiencies of signal samples with different gluino masses, neutralino masses, and $c\tau$'s are compared between different triggers, as shown in Fig. 6.2. The trigger efficiencies shown in the plots are trigger-only efficiencies that no other event selections are applied besides the trigger requirements.

The signal trigger efficiency plot shows that the trigger efficiency is approximately ten times larger than that of other triggers considered in the early stages of analysis design. Fig. 6.2 also shows that the efficiency to pass the trigger is higher with larger gluino-neutralino mass splitting and decreases with increasing gluino $c\tau$. Also, as shown in the plots, for signal samples with the same $c\tau$ and mass splitting, different gluino masses have very similar trigger efficiencies because kinematics are similar for events with the same mass splitting. For small mass splittings, since gluinos are likely to be generated back-to-back, neutralinos are more likely to be back-to-back because most of the gluino momentum is carried away by the neutralino, resulting in low $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$, thus having lower trigger efficiency. For large $c\tau$'s, fewer particles can be well reconstructed because the gluino can decay beyond the beam pipe, resulting in less accurate $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ calculation, thus having lower trigger efficiency.

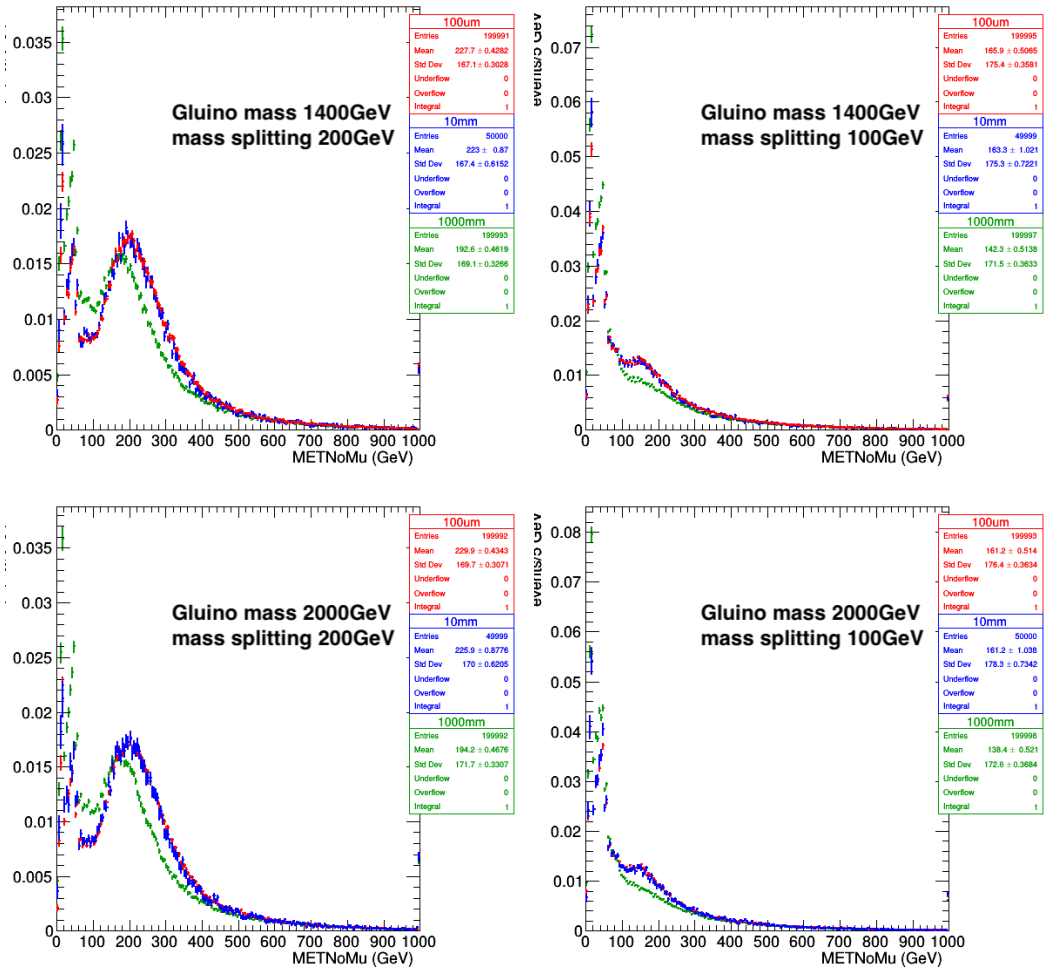

Figure 6.1: $\vec{E}^{\text{miss}}_{\text{T NoMu}}$ distribution of split SUSY signal samples with $c\tau$ as 100 μm (red), 10 mm (blue), and 1000 mm (green) without any selection applied in 2017.
(Top left) Gluino mass 1400 GeV with mass splitting 200 GeV;
(Top right) Gluino mass 1400 GeV with mass splitting 100 GeV;
(Bottom left) Gluino mass 2000 GeV with mass splitting 200 GeV;
(Bottom right) Gluino mass 2000 GeV with mass splitting 100 GeV.

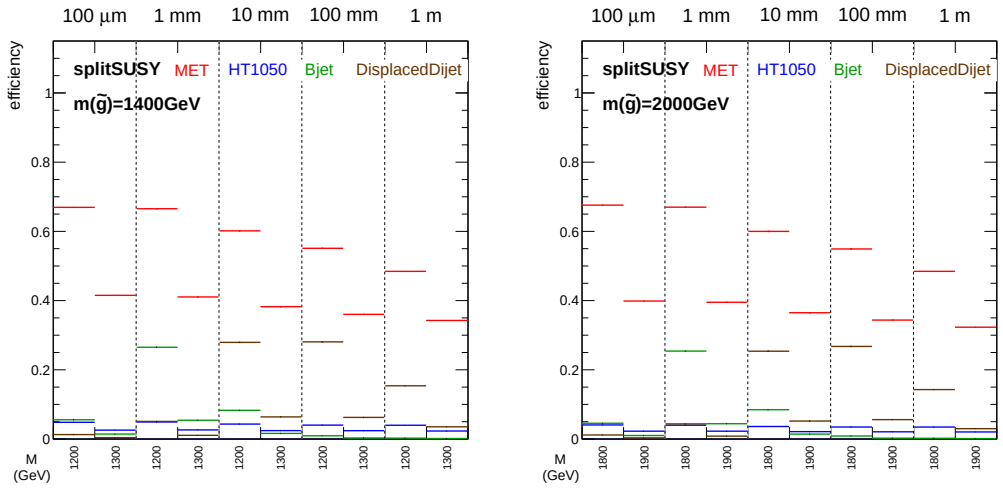

Figure 6.2: Trigger efficiency, i.e. fraction of signal events that pass a given trigger with no other selection applied for samples with 1400 GeV (left) and 2000 GeV (right) gluinos as a function of neutralino mass and lifetime. Red points correspond to the $\vec{E}^{\text{miss}}_{\text{T NoMu}}$ trigger; other points show triggers considered for but not used in this search.

6.2.1 Data-Driven Trigger Efficiency Measurement

The trigger reduces the amount of events recorded in the experiment, which has an effect on the rate of physics processes. It is crucial to understand the trigger efficiency to accurately interpret the search results.

The $\vec{E}^{\text{miss}}_{\text{T NoMu}}$ trigger efficiency is measured with the help of reference triggers. Events that fire a given reference trigger and the corresponding quality requirements are used to measure the trigger efficiency, calculated as the fraction of the number of events that fire the $\vec{E}^{\text{miss}}_{\text{T NoMu}}$ trigger out of all events. A muon-based trigger is used as the reference trigger in this search. To avoid potential bias, an electron-based reference trigger is used as the cross-check. The measurements are conducted on both data and simulated events. The data events are taken from data sets that correspond to the muon and electron-based triggers. The simulated events are taken from lepton-enriched simulated data sets that represent the most relevant processes and have large

cross sections. The simulated data sets are listed in Tab. 6.1.

Table 6.1: Simulated datasets used in the trigger efficiency study.

Process	Dataset name
W+jets, $W \to \ell\nu$	/WJetsToLNu_TuneCP5_13TeV-madgraphMLM-pythia8
Z/γ^*+jets, $Z/\gamma^* \to \ell\ell$, gen. $10 < M_{\ell\ell} < 50\,\text{GeV}$	/DYJetsToLL_M-10to50_TuneCP5_13TeV-madgraphMLM-pythia8
Z/γ^*+jets, $Z/\gamma^* \to \ell\ell$, gen. $M_{\ell\ell} > 50\,\text{GeV}$	/DYJetsToLL_M-50_TuneCP5_13TeV-madgraphMLM-pythia8
$t\bar{t}$	/TTJets_TuneCP5_13TeV-amcatnloFXFX-pythia8

All events used in the trigger efficiency measurements are required to fire the single muon trigger or single electron trigger and satisfy a set of event selections:

- Fire the single muon trigger (`HLT_IsoMu27`) or the single electron trigger (`HLT_Ele35_WPTight_Gsf`);
- Fake MET veto: $|MET_{PF} - MET_{CALO}|/MET_{CALO} < 0.5$;
- Reconstructed $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ passes all MET filters;
- For muon triggered events: leading global muon $p_{\mathrm{T}} > 35$ GeV, $|\eta| < 2.4$, and satisfying tight muon id `CutBasedIdTight`:
  - Global muon and PF muon;
  - Normalized global track $\chi^2 < 10$;
  - Number of inner tracker layers with hits > 5;
  - Number of pixel hits > 0;
  - Number of muon hits > 0 and number of matched muon stations > 1;
  - Muon tracker track transverse impact parameter $\mathrm{d_{xy}} < 2$ mm w.r.t. the primary vertex;
  - Muon tracker track longitudinal distance $\mathrm{d_z} < 5$ mm w.r.t. the primary vertex;

- For electron triggered events: leading global electron $p_{\mathrm{T}} > 38\,\mathrm{GeV}$, $|\eta| < 2.5$, and satisfying tight electron id `cutBasedElectronID-Fall17-94X-V2-tight`, impact parameter cut including d0 and dz are not applied.

These selections help select events with real $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ and well-measured muons or electrons. The cut on lepton p_{T} is selected to make sure all events are in the plateau region of the reference lepton triggers.

The trigger efficiency is plotted as a function of the offline $\vec{E}_{\mathrm{T\ NoMu}}^{\mathrm{miss}}$ in Fig. 6.3 for different data-taking periods. To account for the difference in trigger efficiencies between data and simulation, bin-by-bin correction factors are derived from the data/MC trigger efficiency ratio as a function of $\vec{E}_{\mathrm{T\ NoMu}}^{\mathrm{miss}}$, which are shown in the lower panels of plots in Fig. 6.3. The correction factors are then applied to simulated events.

The trigger efficiency measurements performed on muon-triggered and electron-triggered events shown in Fig. 6.4, and Fig. 6.5 compare the trigger efficiency as a function of $\vec{E}_{\mathrm{T\ NoMu}}^{\mathrm{miss}}$ between muon-triggered and electron-triggered events for data and simulation. The comparisons show consistent behavior for muon-triggered and electron-triggered events. The slight difference in trigger efficiencies measured using muon-triggered and electron-triggered is accounted for in Sec. 11.1.5.

6.3 Event preselection

The event preselection is kept as simple as possible to maintain sensitivity to a wide range of physics models. To ensure that the trigger efficiencies (Sec. 6.2) are well understood and consistent between data and simulation, the $\vec{E}_{\mathrm{T\ NoMu}}^{\mathrm{miss}}$ is required to be greater than 200 GeV. Events with $\vec{E}_{\mathrm{T\ NoMu}}^{\mathrm{miss}} < 200\,\mathrm{GeV}$ are used to train the ML

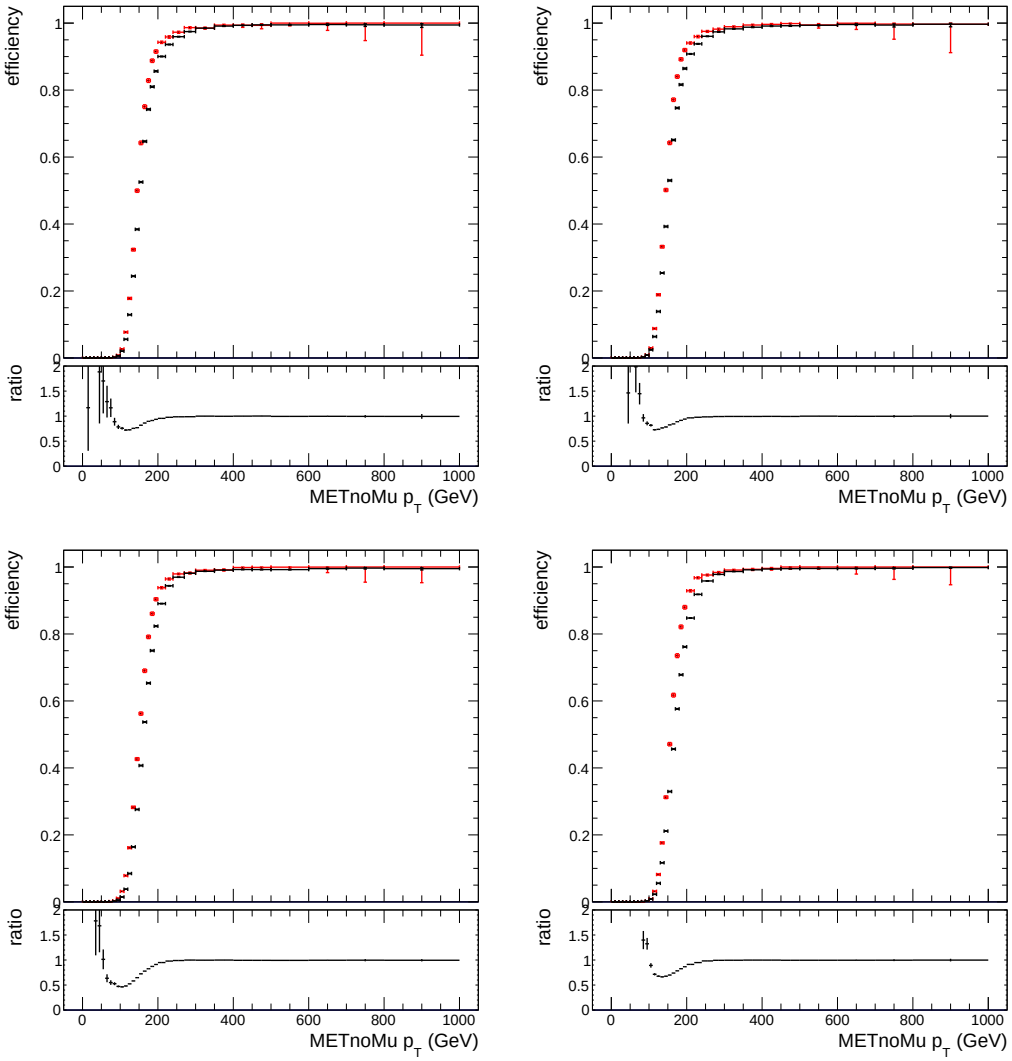

Figure 6.3: `HLT_PFMETNoMu120_PFMHTNoMu120_IDTight` trigger efficiency as a function of $\vec{E}^{\text{miss}}_{\text{T NoMu}}$. Data (black) and background (red) events used in the measurement are taken in 2016 preVFP (top left), 2016 postVFP (top right), 2017 (bottom left), and 2018 (bottom right).

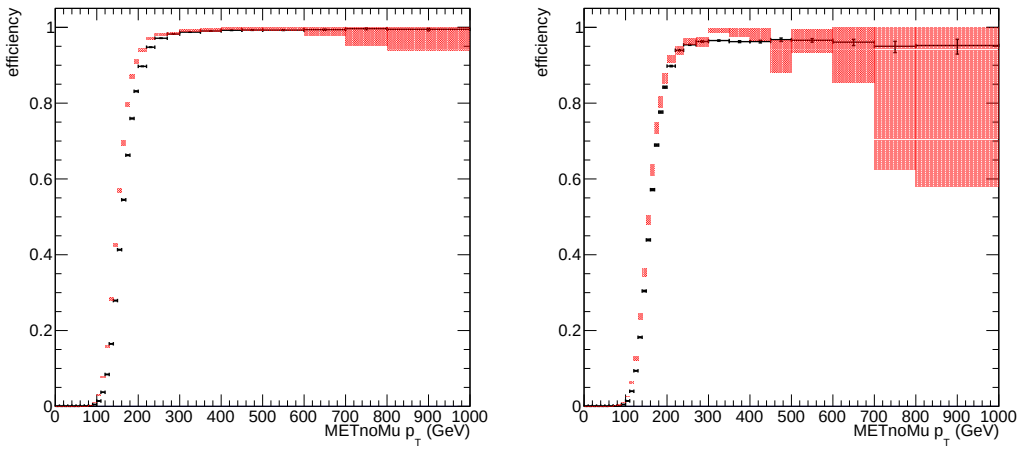

Figure 6.4: `HLT_PFMETNoMu120_PFMHTNoMu120_IDTight` efficiency as a function of offline $\vec{E}^{\text{miss}}_{\text{T NoMu}}$ in muon-triggered (left) and electron-triggered (right) events in the data (black) and simulation (red).

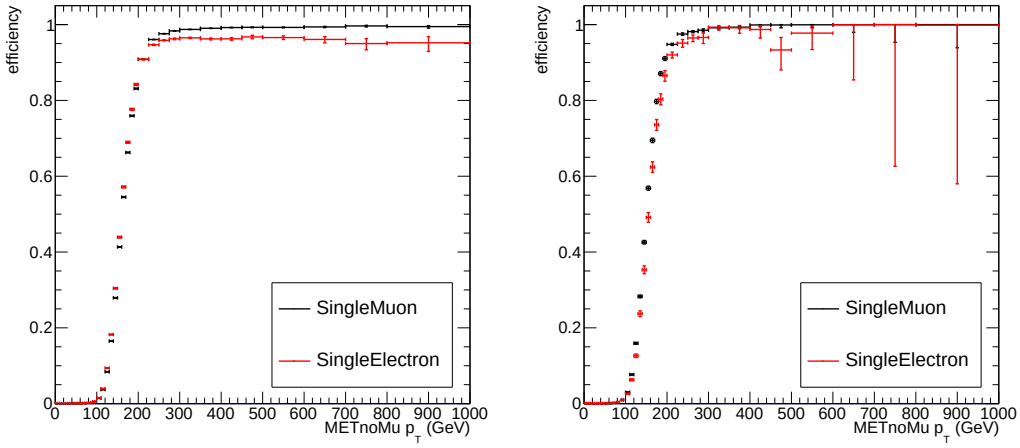

Figure 6.5: `HLT_PFMETNoMu120_PFMHTNoMu120_IDTight` efficiency as a function of offline $\vec{E}^{\text{miss}}_{\text{T NoMu}}$ comparison between muon-triggered (black) and electron-triggered (red) events in the data (left) and simulation (right).

algorithm of the search, as described in Sec. 8.

To pass the event preselection, events must :

- Fire the `HLT_PFMETNoMu120_PFMHTNoMu120_IDTight` trigger;
- Pass the fake MET veto: $|MET_{PF} - MET_{CALO}|/MET_{CALO} < 0.6$;
- Pass the MET filters;
- Have $\vec{E}^{\text{miss}}_{\text{T NoMu}} \geq 200\,\text{GeV}$.

The fake MET veto is not only useful in getting rid of mismeasured $\vec{p}^{\text{miss}}_{\text{T}}$, but also helpful in rejecting background events. A comparison of the fake MET veto distribution between signal and background is shown in Fig. 6.6. For events with reconstructed secondary vertices that have at least 5 tracks, as introduced in Ch. 7, the fake MET veto removes $\sim 25 - 30\%$ of background events. While for signal events with gluino mass of 2000 GeV and $c\tau$ of 1 mm, removing the fake MET veto only helps us gain $\sim 7\%$ for mass splitting as 100 GeV and $\sim 5\%$ for mass splitting as 200 GeV.

The events that pass the event preselection are further analyzed as described in the remainder of the thesis.

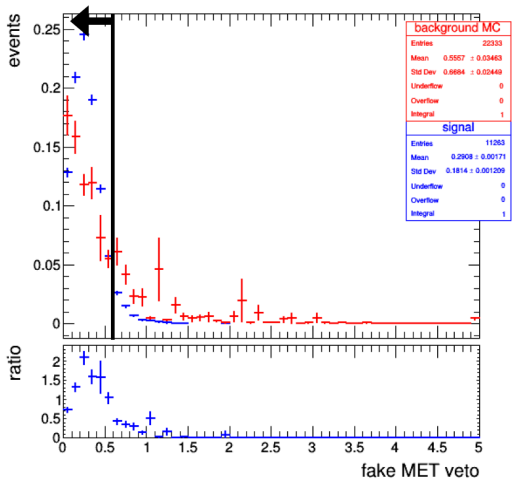

Figure 6.6: Distributions of the fake MET veto value between background simulation (red) and signal (blue) with gluino mass of 2000 GeV, mass splitting of 200 GeV, and $c\tau$ of 1 mm in 2017. The black line and arrow show where are how the fake MET veto is applied. Events in the plot are required to satisfy the preselection, except the fake MET veto.

Chapter 7

Vertex Reconstruction

Reconstructed secondary displaced vertices are the main feature used to identify potential signal events in this search. This chapter introduces how secondary displaced vertices are reconstructed and selected. The vertex reconstruction algorithm is used in [26].

7.1 Track Selection

In each event, tracks measured by the CMS tracker are used to reconstruct secondary displaced vertices. To ensure that tracks used in vertex reconstruction are precisely measured, criteria are imposed on track p_{T}, the number of pixel layers in which a hit is measured for each track (n_{pxl}), the innermost barrel pixel layer that has a measured hit for each track (r_{min}), and the number of strip tracker layers in which a hit is measured for each track (n_{stl}).

Specifically, tracks must satisfy the following requirements:

- $p_{\mathrm{T}} > 1\,\mathrm{GeV}$;
- $r_{min} = 1$, e.g. have a hit in the innermost barrel pixel layer;
- $n_{pxl} \geq 2$;

- $n_{stl} \geq 6$.

These requirements constitute the track *quality* requirements used for the search. Having rejected tracks that are poorly measured, we then select displaced tracks by applying one last requirement on the track transverse impact parameter significance ($|d_{xy}|/\sigma_{d_{xy}}$):

- $|d_{xy}|/\sigma_{d_{xy}} > 4$.

The track variables are shown in Fig. 7.1 for 2017 MC. To ensure the selection is well understood in the data and the simulation, a comparison is performed on data and simulation about track variables and corrections are applied on the simulation to mitigate the differences. The details of the comparisons of track variables between data and simulation are described in Appendix B.1.

7.2 Vertex Reconstruction

The vertex reconstruction algorithm is performed in steps. First, seed vertices are formed by all pairs of tracks that satisfy the above requirements using the Kalman Filter [43] approach. Seed vertices with a χ^2 per degree of freedom (χ^2/dof) less than 5 are kept for the next steps. At this stage, all vertices are composed of only two tracks, and many of them share tracks.

Next, an iterative merging process is applied on vertices that share tracks. Each pair of vertices that share tracks are merged or split depending on the distance between the vertices: if the three-dimensional (3D) distance is smaller than four times its uncertainty, a vertex fit is applied to the complete set of tracks from both vertices.

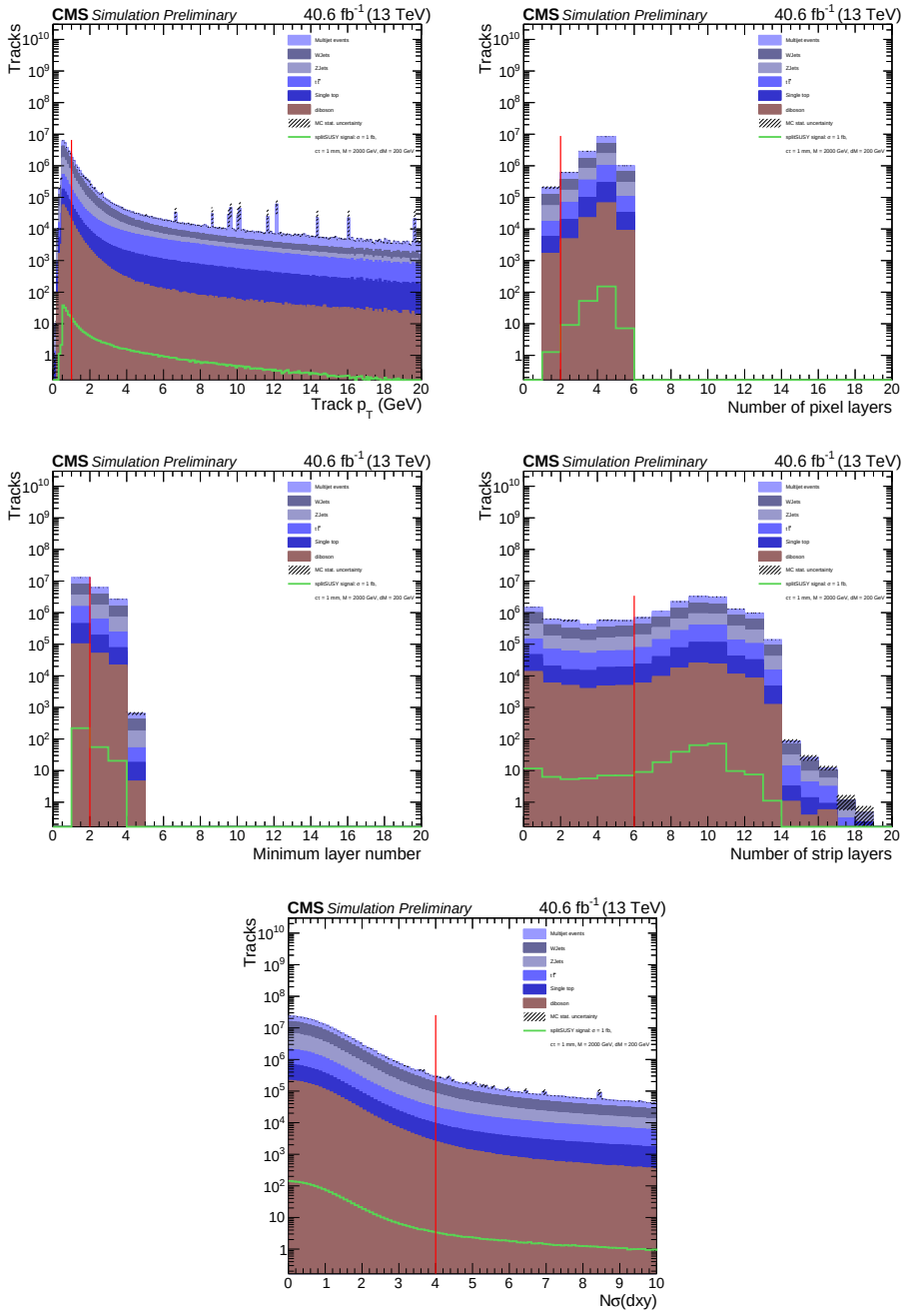

Figure 7.1: Track variable distributions used for track selection. All events in the distributions are from the 2017 simulation and satisfy all event preselection. Tracks shown in the plots satisfy all the track selections except the one shown on the x-axis. The red line labels the selection applied to the shown variable in the plots. In most of the plots, tracks on the right of the red line are taken as tracks that satisfy the selection, while for r_{min}, tracks on the left of the red line are selected.

If the resulting fit satisfies the χ^2 requirement, the two vertices are replaced by the merged vertex. Pairs of vertices that are not merged but share tracks are split, in which case shared tracks are assigned to one of the vertices or removed, depending on the distance between shared tracks and vertices. If the track impact parameters with respect to both vertices are smaller than 1.5 standard deviations, the track is assigned to the vertex with more tracks; if the track impact parameters with respect to both vertices are greater than 5 standard deviations, the track is dropped from both vertices; if none of the conditions above is satisfied, the significance of the track impact parameters with respect to both vertices is compared, and the track is assigned to the vertex with a smaller impact parameter significance. With the newly assigned tracks, all the remaining vertices are refitted and kept if the χ^2 requirement is met.

Following the merging and splitting process, it is found that some background vertices are formed with tracks from pileup primary vertices. To mitigate such an effect, each reconstructed vertex is further processed by iterating over each track included in the vertex and refitting the vertex without the given track. If the refitting causes a shift in the z direction by more than $50\,\mu$m, the track is dropped from the vertex. The iterative process stops when the resulting vertex has only two tracks.

7.3 Vertex Selection

The reconstructed vertices are further selected to discriminate signal and background vertices. Each vertex must:

- Be composed of at least three tracks;

- Be within the beam pipe to suppress vertices from material interactions (i.e. x-y distance of the vertex from the center of the beam pipe must be less than 20 mm in 2016 and 20.9 mm in 2017 and 2018);

- Have $\sigma_{d_{BV}}$, defined as uncertainty in x-y distance of the vertex from the beam axis (d_{BV}), less than 25 μm, to select only well-reconstructed vertices and get rid of vertices from decays of boosted particles.

The beam pipe center, whose positions relative to the geometric center of CMS are (0.124, 0.070) cm in 2016, (0.113, -0.180) cm in 2017, and (0.171, -0.175) cm in 2018, is used when calculating the x-y distance of the vertex from the center of the beam pipe. The variables used for selection, namely the number of tracks per vertex and $\sigma_{d_{BV}}$, are shown in Fig. 7.3 for 2017.

Events with at least one vertex satisfying the above requirements are further analyzed as described in the remainder of this thesis. Specifically, events with vertices that are composed of at least five tracks are used as the potential signal region and events with vertices that are composed of three or four tracks are used as potential control regions.

The vertex reconstruction efficiency as a function of two-dimensional (2D) LLP travel distance is shown in Fig. 7.2 for mass splittings of 100 and 200 GeV. The efficiency is obtained by calculating the ratio between the 2D LLP travel distance distributions. The denominator histogram includes all LLPs in events that satisfy the event pre-selection except offline $\vec{E}^{\text{miss}}_{\text{T NoMu}}$ selection. The numerator histogram includes LLPs that can be matched with reconstructed vertices that satisfy the vertex selection and have at least five tracks. LLPs are matched with vertices by distance: if there is only one vertex in the event, the LLP that is closest to the vertex will be matched; if there

is more than one vertex in the event, both LLPs in the events will be matched.

As shown in the efficiency plots, the efficiencies are low when the LLP decay length is less than 500 μm. Then the efficiencies reach a plateau for larger LLP decay lengths until the LLP decay length is close to the beam pipe, where the efficiency drops again.

Figure 7.3 shows the distributions of vertex variables used in the selection. A double-peak structure of the $\sigma_{d_{BV}}$ is observed. The first peak with lower $\sigma_{d_{BV}}$ corresponds to vertices that come from track random crossing while the second peak with larger $\sigma_{d_{BV}}$ comes from B meson decays. The structure is discussed in detail in Sec. 10.1.

Other vertex variables are checked, such as vertex invariant mass and $\Delta\phi$ between $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ and the vector from the primary vertex and secondary vertex, as shown in Fig. 7.6 and Fig. 7.7 for 2017. For each set of plots, the top two plots include vertices with at least five tracks that satisfy other vertex requirements while the bottom two plots include vertices with at least three tracks that satisfy other vertex requirements to reduce the statistical uncertainty. Comparisons of vertex invariant mass and $\Delta\phi$ between $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ and the vector from the primary vertex and secondary vertex between signal and background MC do not suggest additional selections based on those variables.

Figure 7.4 shows the distribution of the number of reconstructed vertices. Most events have only one reconstructed vertex, even for signal events, due to the low energy and small number of displaced tracks we have in $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$-triggered event, in which a noticeable fraction of energy escapes the detector as $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$. The number of charged particles from the decay of LLP at the generator level is calculated and shown in Fig. 7.5 to confirm the small number of tracks effect in the generator level. Note that generator level information can be sometimes hard to interpret, there are cases where

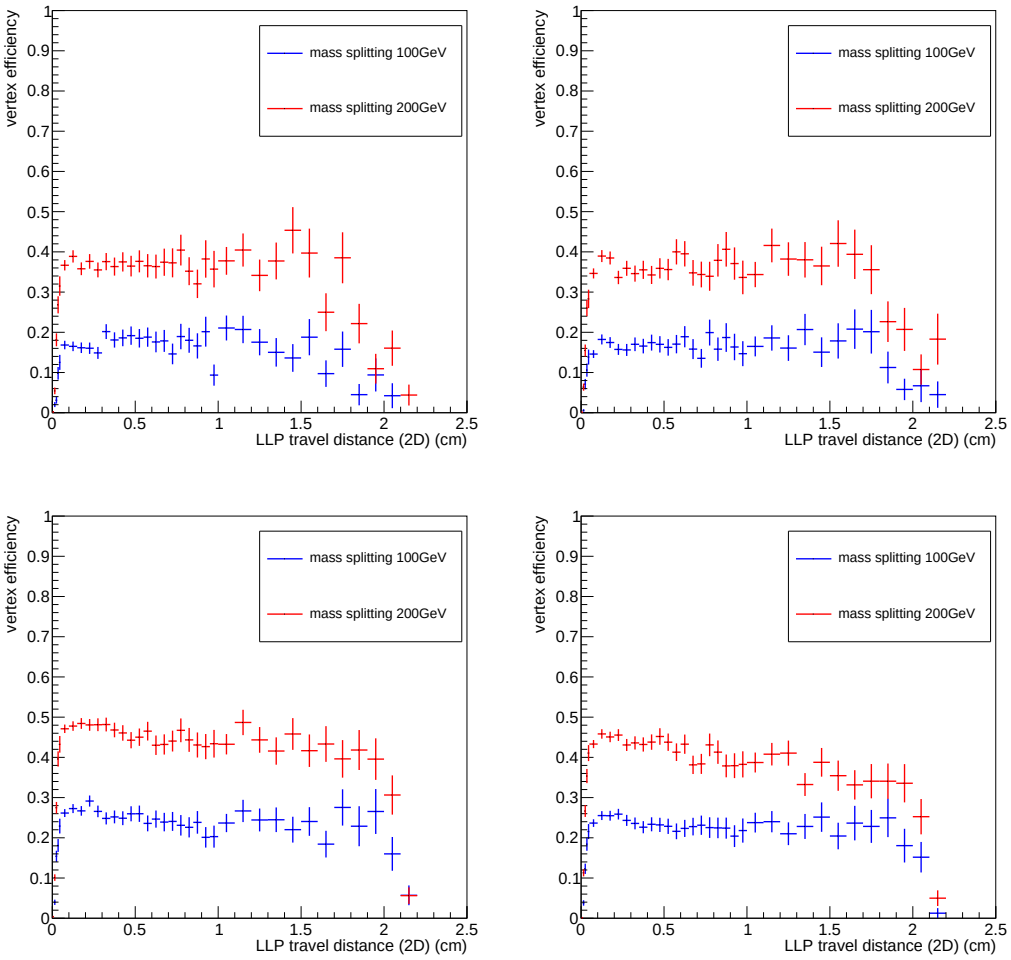

Figure 7.2: Vertex reconstruction efficiency as a function of 2D LLP travel distance. 2016 preVFP (top left), 2016 postVFP (top right), 2017 (bottom left), and 2018 (bottom right) are shown separately.

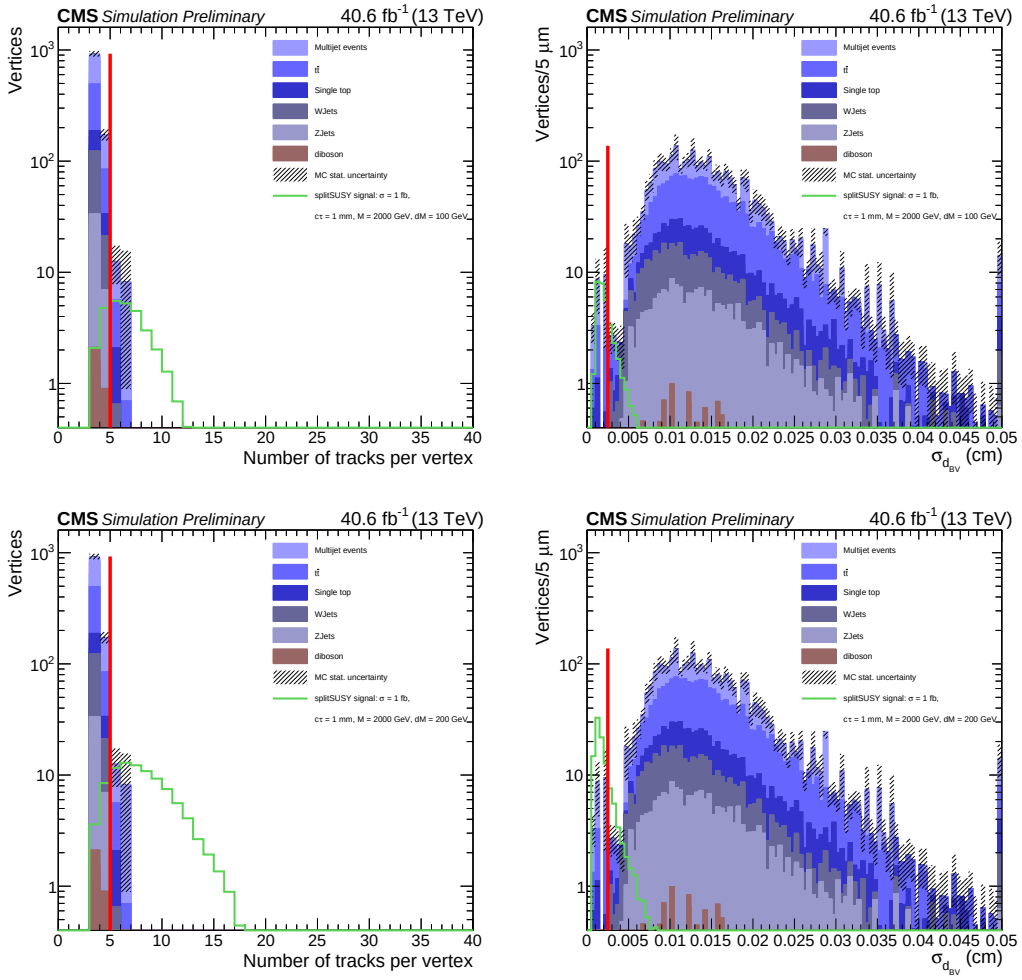

Figure 7.3: Distributions of variables used for vertex selection in the split SUSY signal sample with mass splitting $\Delta M = 100$ GeV (top) and $\Delta M = 200$ GeV (bottom), and background MC. The signals are scaled up by 10 to make a better comparison. All events in the distributions are from the 2017 simulation and satisfy the event preselection. All vertices satisfy the vertex selection except the one that is shown on the x-axis of the plot. The red line labels the selection applied to the shown variable in the plots. For the number of tracks per vertex, vertices on the right of the line are kept, while for $\sigma_{d_{BV}}$, vertices on the left of the line are kept.

quarks have no generator level decay daughters and sometimes quarks end up having all neutral daughters, which results in LLPs that have zero charged decay product in the generator level. Comparisons of vertex variables between data and the simulation and studies to understand the differences are shown in Appendix. B.2.

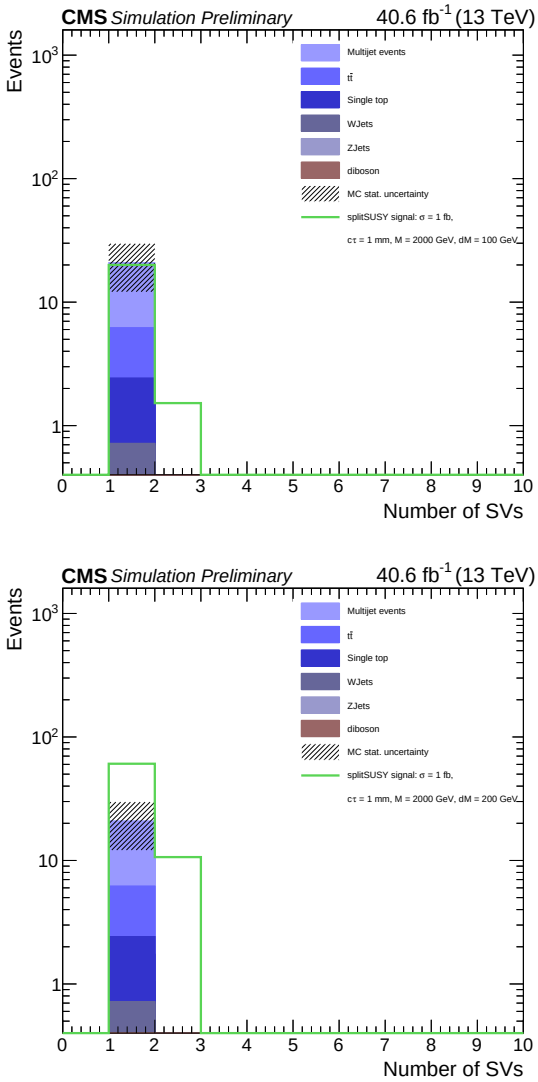

Figure 7.4: Distributions of the number of reconstructed vertices in 2017 split SUSY signal sample with mass splitting $\Delta M = 100\,\text{GeV}$ (left) and $200\,\text{GeV}$ (bottom), and background MC. The signals are scaled up by 10 to make the comparisons easier to visualize. All events are required to pass event preselection and all vertices are required to have at least 5 tracks and satisfy other vertex requirements.

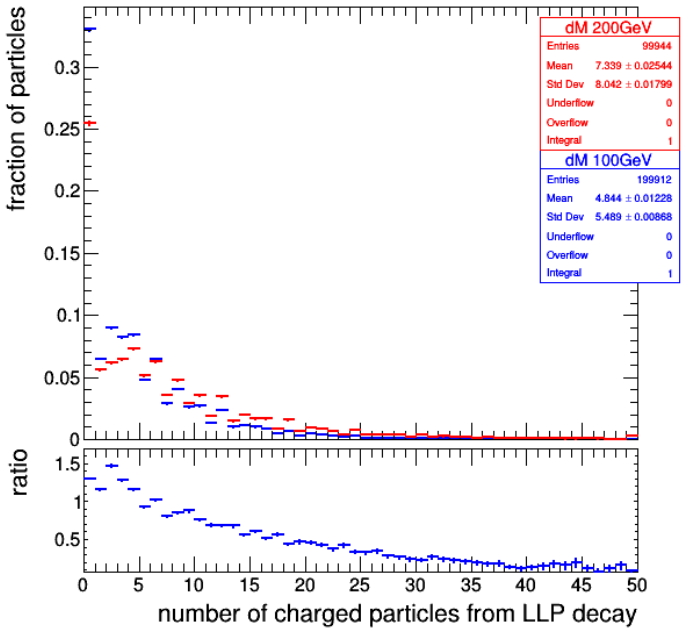

Figure 7.5: Number of charged particles that come from the LLP decay in generator level in 2017 for split SUSY samples with $c\tau$ of 1 mm and gluino mass of 2000 GeV.

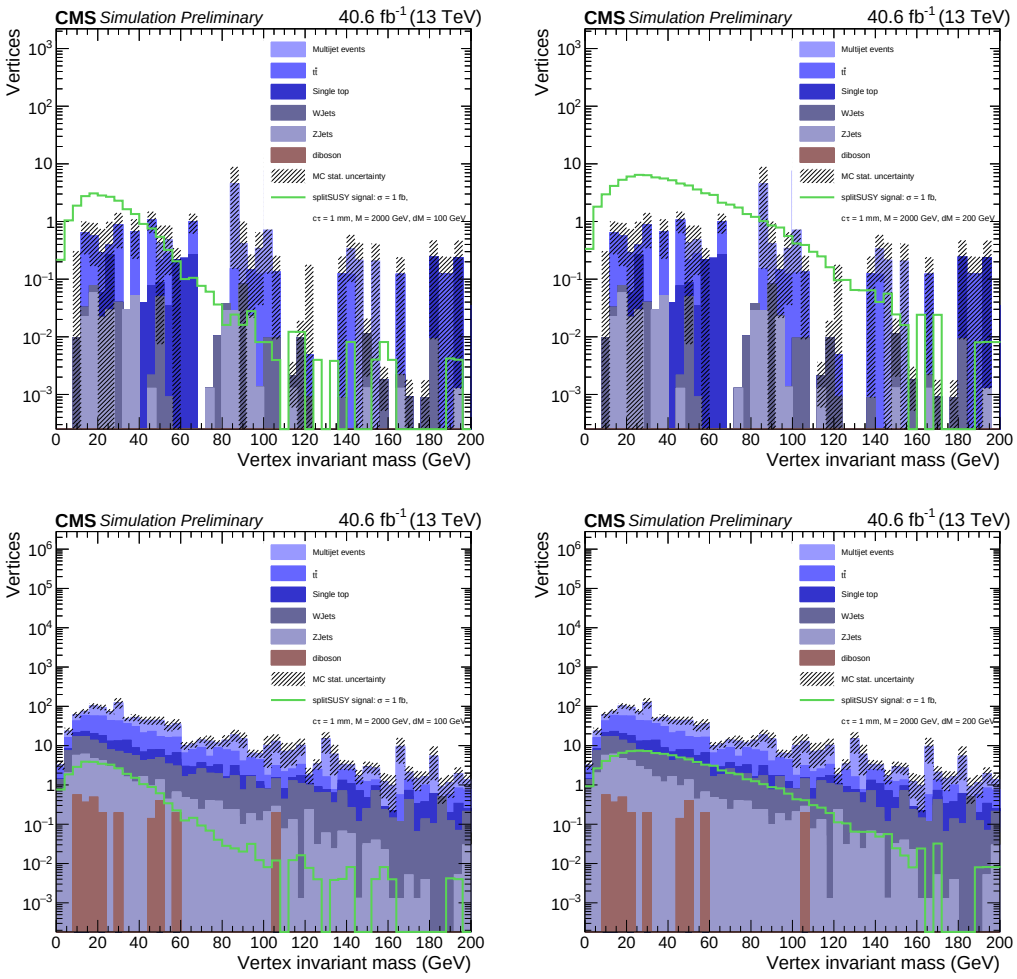

Figure 7.6: Distributions of the invariant mass of reconstructed vertices in 2017 split SUSY signal sample with mass splitting $\Delta M = 100\,\text{GeV}$ (left) and $200\,\text{GeV}$ (right), and background MC. The signals are scaled up by 10 to make a better comparison. All events are required to pass event preselection and all events are required to have at least one vertex with at least five (top) or three (bottom) tracks and satisfy other vertex requirements.

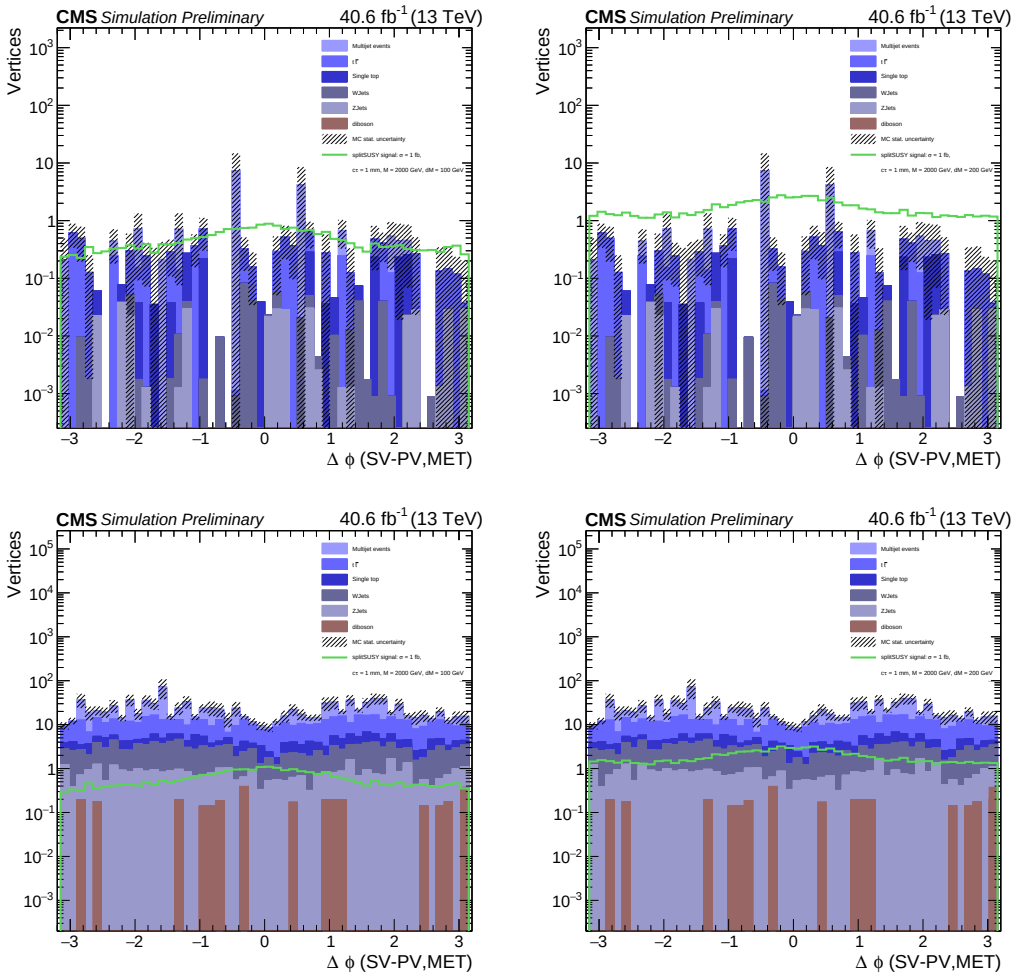

Figure 7.7: Distributions of $\Delta\phi$ between $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ and the vector from the primary vertex to the secondary vertex in 2017 split SUSY signal sample with mass splitting $\Delta M = 100$ GeV (left) and 200 GeV (bottom), and background MC. The signals are scaled up by 10 to make a better comparison. All events are required to pass event preselection and all vertices are required to have at least 5 (top) or 3 (bottom) tracks and satisfy other vertex requirements.

Chapter 8

Machine Learning

Machine Learning (ML) [44] is an approach that is able to make predictions based on its previous experience, given the input features. The ML is constructed as a model with a large number of parameters. It takes input data and calculates the output using parameters in the model. The output is taken as the prediction of the ML. To obtain the desired predictions, ML models undergo a training process, which happens iteratively by fitting the models using the input data.

Based on the method of training the ML models, an ML model can be classified as supervised learning, unsupervised learning, and reinforced learning. For supervised learning algorithms, input features together with the "truth" values, which are the correct output, are fed into the ML during the training process. Supervised ML models are expected to make predictions that are close to the "truth" values. Depending on the output of the ML models, supervised learning algorithms can be categorized as classifiers and regressors. Classifiers predict the class of the given set of input features while regressors predict the values of desired variables based on the input features. For unsupervised learning algorithms, only the input features are provided during the training. Unsupervised ML models are expected to learn the underlying patterns of the input features. Reinforced learning is often used for decision-making. It is able to interact with the environment and thus can explore different possibilities and learn from the successes and failures of the feedback of the environment based

on the decision it makes.

The ML algorithm used in this search is a classifier trained with supervised learning.

8.1 Introduction to Interaction Network

With our event preselection and vertex selection, the number of background events is further reduced by introducing an Interaction Network (IN) [45], an ML algorithm to this search.

INs are a kind of Graph Neural Network (GNN) [46], which is a novel ML algorithm that shows its powerful performance in many different applications. GNNs take graphs, a data structure that is composed of nodes and edges, as input. Nodes are used to represent objects while edges are used to represent relations between different nodes. Special architectures are designed in GNNs to make them able to learn not only the information about objects but also the relations between objects. For INs, nodes are used to describe different physical objects and edges between nodes are used to describe possible interactions between nodes. Calculations are performed to take care of both the object itself and possible interactions between the object and all other objects. With all these features, INs are capable of calculating the interactions for different objects in a multi-body system and thus predicting the final states of all physical objects in the system. Similarly, proton-proton collisions can also be viewed as multi-body interactions. All the tracks reconstructed in the detector are physical objects and interact with each other, so the IN is used to calculate the interactions between tracks and trained as the classifier to predict whether there are LLP decays in a given event.

In this search, the vertex reconstruction is already very effective at reconstructing displaced vertices and distinguishing signal events from background events. To further improve the signal significance, the IN introduced in this search is expected to exploit as much additional information about the displaced vertices as possible but not the information that is already available from the existing vertex reconstruction. As a result, the reconstructed vertex information is not directly used to avoid the IN only learning information from vertex reconstruction. Only tracking information is fed into the IN. In that way, the IN can learn the difference between the topology of tracks in signal and background events and thus classify a given event as signal or background. The IN makes predictions by outputting a score ranging from 0 to 1 for a given event, where 0 corresponds to the background and 1 corresponds to the signal. The IN output score ($MLScore$) is used as a discriminator to distinguish signal events from background events. Tracks are implemented as nodes in the graph while edges are composed of relations between any pairs of tracks. The output of the IN will be used together with variables from reconstructed vertices to improve search efficiency.

The following seven track variables are used in the ML so the spatial and kinematic information of tracks is available to ML: p_{T}, η, ϕ, xy-distance between the beam spot and the closest approach of the track dxy_{BS} and its significance, z distance between beam spot and the closet approach of the track dz_{BS} and its significance.

Comparisons of track variables used in the ML between signal and simulated background events are shown in Fig. 8.1 and 8.2. As shown in the plots, the signal and background distributions are similar. The IN learns the track topology for LLP decays, which is hard to observe directly from the distributions of track variables.

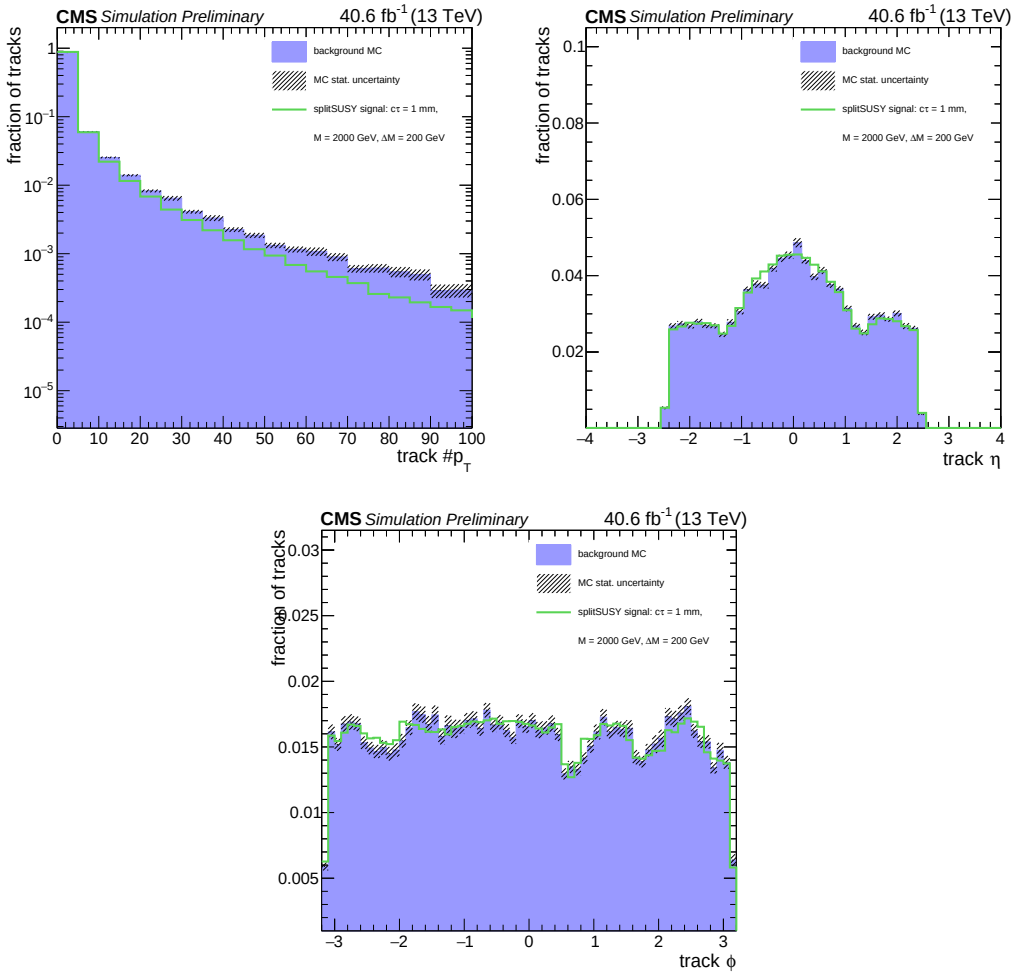

Figure 8.1: Comparisons of track p_{T}, η, and ϕ between signal and background in 2017. The signal used in the comparison is the split SUSY sample with 1 mm $c\tau$, 2000 GeV gluino mass, and 200 GeV mass splitting.

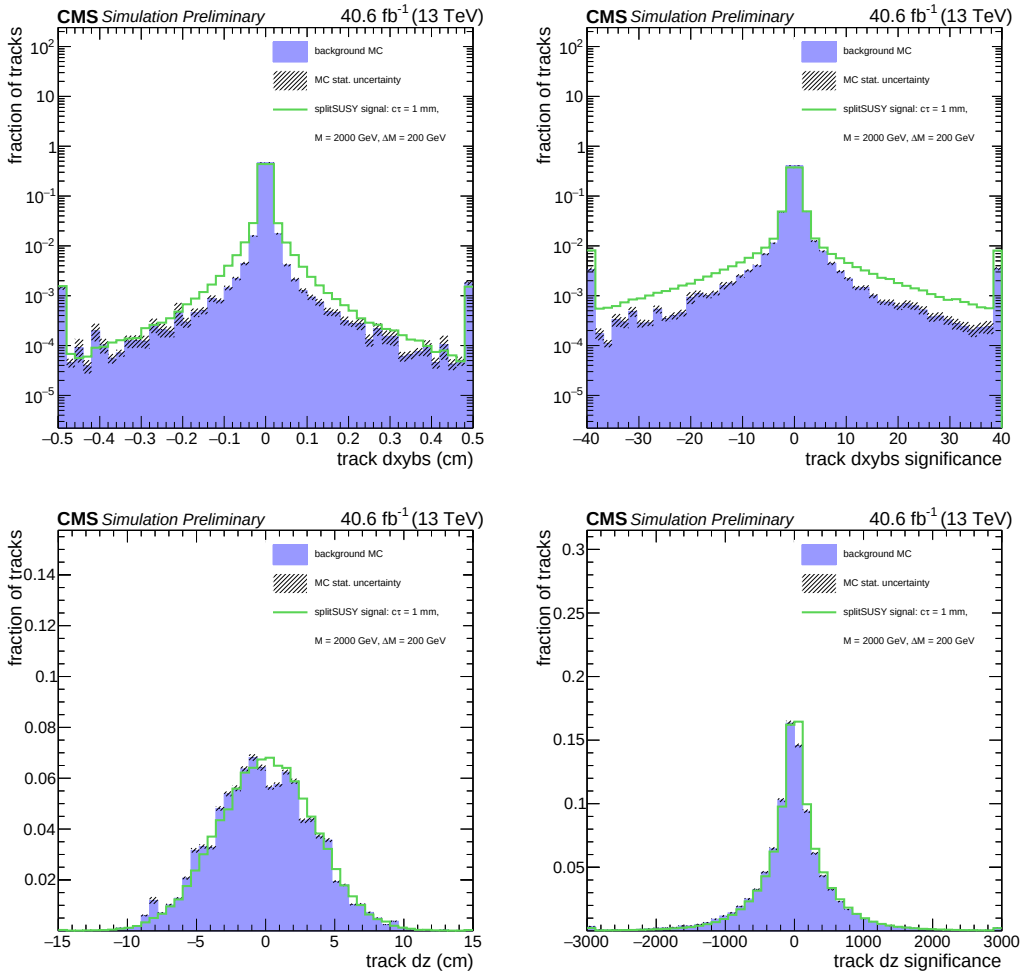

Figure 8.2: Comparisons of track d_{xy}, η, and ϕ between signal and background in 2017. The signal used in the comparison is the split SUSY sample with 1 mm $c\tau$, 2000 GeV gluino mass, and 200 GeV mass splitting.

8.2 Architecture of Interaction Network

As described in the last section, the IN takes the track variables as input. For each track, it calculates the "interactions" the track applies on other tracks and the "interactions" the track receives from other tracks. The applied and received interactions are used to calculate the "effect" of all the "interactions" acted on individual tracks. The IN then combines the original input information of tracks with the "effects" on the tracks to understand how the "interactions" affect the track behaviors and infer the underlying topology of tracks in the events and predict whether an event has tracks from LLP decays.

The detailed architecture of the IN is shown in Figure 8.3 and described as follows to achieve the calculations described above. The standard deep neural network building blocks including multilayer perceptrons (MLP) and matrix operations are used when implementing the IN. Assume we have N events and we define our number of objects, namely tracks, to be fed into ML as N_o, each track has D_s variables that are available to the IN. The graph, which is the input of the IN, includes basic information about tracks and relations between all pairs of tracks. The input graph is constructed using several matrices: variable matrix O (dimension $N \times N_o \times D_s$) that includes the information (such as p_{T}, η, ϕ, etc.) of tracks used in the IN; receive relation matrix R_r (dimension $N \times N_o \times N_r$) that represents the "interactions" a track receives from other tracks; send relation matrix R_s (dimension $N \times N_o \times N_r$) that represents the "interactions" a track sends to other tracks, and a general characteristic relation matrix R_a (dimension $N \times 1 \times N_r$) that represents the general characteristics for all relations. Since the IN calculates relations between all pairs of objects, the number of relations N_r is $N_o \cdot (N_o - 1)$ (not including objects reacting with themselves). All

relations between different tracks are considered relevant so all elements in R_r and R_s except the elements that correspond to the relations between a track and itself are set to 1, and all elements in R_a are set to 1.

The IN calculates the state after tracks receive and send relations with other tracks by $O \cdot R_r$ and $O \cdot R_s$ and combine all those calculated elements together as one whole matrix by $O \cdot R_r - O \cdot R_s$. This matrix is then combined together with R_a to form the full relation matrix. The calculated matrix is passed into an MLP called ϕ_R with five layers with 50 nodes in each layer. The output of ϕ_R, called E, is interpreted as the "effects" of all relations acted on each track. The dimension of matrix E is $N \times N_r \times D_e$ where D_e is a tune-able hyperparameter that represents the dimension of effect for each track. The "effect" matrix E is multiplied with the transpose of matrix R_r to calculate the effect acted on individual tracks. The result is combined together with the original variable matrix O to form a new matrix C containing the original information together with the tracks after the "effects" are applied. The dimension of the new matrix C is $N \times N_o \times (D_s + D_e)$. The matrix C is passed into another MLP with two parts. The first part, labeled as ϕ_O, has two layers with 50 nodes in each layer and the output of the first part of the MLP is a matrix, referred to as P, with dimension $N \times N_o \times D_p$. The second part, labeled as ϕ_{output}, has three layers with 100 nodes in each layer. It takes matrix P as input and outputs a matrix with dimension $N \times 1$, the 1 denotes the calculation results from the MLP for each event, which is a single value. This output is finally processed with a sigmoid function to obtain the final result which corresponds to the final MLScore. The MLScore ranges from 0 to 1, with 0 corresponding to "background-like" and 1 corresponding to "signal-like".

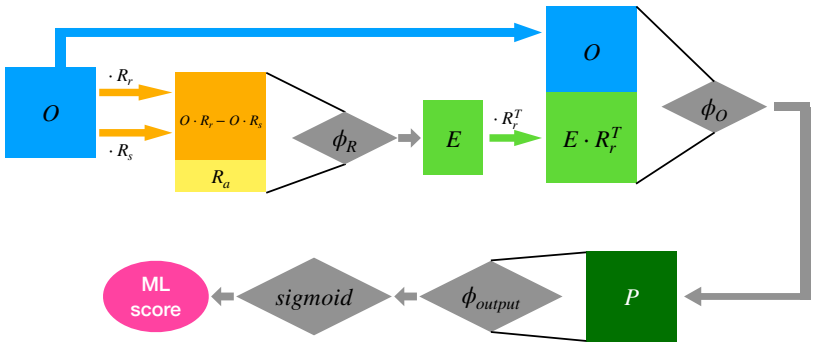

Figure 8.3: The architecture of the IN. The arrows show the flow of data. The rectangular boxes represent matrices and diamonds represent MLPs in the architecture.

8.3 Training and Testing

The IN is computationally expensive because it calculates the "interactions" between each pair of tracks fed into the algorithm. It is impossible to feed all tracks in a given event because the computing resources cannot handle such heavy computations. As a result, we select tracks that satisfy the track quality requirements described in Sec. 7.1 but without the displaced track requirement on $|d_{xy}|/\sigma_{d_{xy}}$. Then the first 50 p_{T}-ordered tracks from the set of tracks selected above in each event are used to feed into the network. If an event has less than 50 tracks, "fake tracks" are made with all variables set to 0 to make all events have the same number of selected tracks to be processed by the algorithm.

We used 2016, 2017, and 2018 data and MC samples mixed together to train and

test the IN. Data events with 3-track and 4-track vertices are used as part of the background events. Background MC samples are combined together according to their cross sections. To have enough background events for training, simulated background events with an effective luminosity of $200000\,\mathrm{pb}^{-1}$ are used. Given the fact that QCD samples with low H_{T} have very large cross sections, it is not feasible to generate enough events to reach $200000\,\mathrm{pb}^{-1}$ luminosity. For such samples, we can only use as many events as possible but cannot strictly choose the number of events according to their cross sections. Signal samples with different lifetimes and masses are combined together as the total sample of signal events used in the training and testing. Given the number of signal samples with different lifetimes and masses, there are much more signal events than background events that can be used to train the IN. To make the number of signal and background events similar in the training, not all signal events are used to train the IN. From each signal sample, we pick at most 250 events in 2016 preVFP and 2016 postVFP, and 500 events in 2017 and 2018 for the training.

Events are divided for training and testing purposes by the following selection:

- Training Selection:
  - Fire $\vec{E}_{\mathrm{T\ NoMu}}^{\mathrm{miss}}$ trigger, pass MET filters and fake MET veto;
  - $80\,\mathrm{GeV} < \vec{E}_{\mathrm{T\ NoMu}}^{\mathrm{miss}} < 200\,\mathrm{GeV}$;
  - Have at least 1 vertex that:
    - Satisfy beam pipe constraints described in Sec. 7.3;
    - Be composed of at least three tracks;
    - Have $\sigma_{d_{BV}} < 25\,\mu\mathrm{m}$.
- Testing Selection:

- Fire $\vec{E}^{\text{miss}}_{\text{T NoMu}}$ trigger, pass MET filters and fake MET veto;
- $\vec{E}^{\text{miss}}_{\text{T NoMu}} > 200\,\text{GeV}$;
- Have at least 1 vertex that:
  - Satisfy beam pipe constraints described in Sec. 7.3;
  - Be composed of at least three tracks for MC and be composed of 3 or 4 tracks for data to avoid unblinding the signal region;
  - Have $\sigma_{d_{BV}} < 25\,\mu\text{m}$.

Given the fact that the ML is trained using events with $\vec{E}^{\text{miss}}_{\text{T NoMu}}$ trigger but low $\vec{E}^{\text{miss}}_{\text{T NoMu}}$, distributions of the magnitude and ϕ of $\vec{E}^{\text{miss}}_{\text{T NoMu}}$ and $\vec{p}^{\text{miss}}_{\text{T}}$ are checked on data and background events used in the training. Only events with 3-track vertices are checked in order to avoid accidentally unblinding the signal region in this analysis. The plots are shown in Fig. 8.4.

Since different variables in data have very different features, different variables can have very different contributions to the IN and make the IN converge slower while training. To make every variable have a similar impact on the IN, data pre-processing is applied. Generally, data pre-processing is done by normalizing all variables by $(x-\mu_x)/\sigma_x$ where x corresponds to a given variable, μ_x corresponds to the mean of x and σ_x corresponds to standard deviation of x. The variable distribution is not necessarily Gaussian to apply this normalization. The main point of this normalization is to make all the input variables have similar value ranges to accelerate the training process. In this search, a special approach is used to achieve the normalization. The IN is trained and tested using a mixture of data and simulation but will be used only on data in the end. Since the simulation is not perfect, data and simulated events can be slightly different because of the mis-modeling of the simulation. To make the

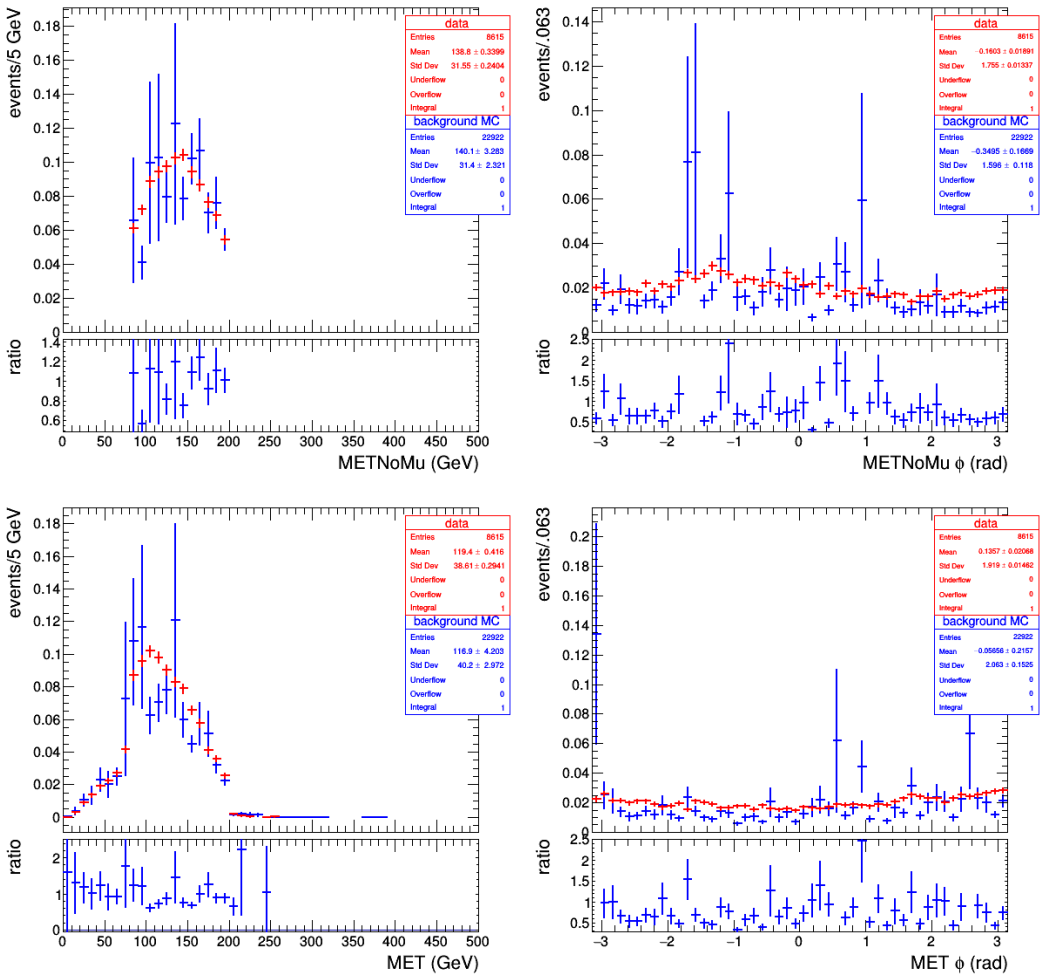

Figure 8.4: Distribution of the magnitude and ϕ of $\vec{E}_{\mathrm{T\ NoMu}}^{\mathrm{miss}}$ and $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ in 2017. Data events are shown in red and simulated background events are shown in blue.

IN perform consistently in data and the simulation, pre-processing is used to make the simulation and data have more similar distributions. However, the differences between distributions in data and simulation come not only from mis-modeling, but also from the lack of events in some background simulation samples, which should not be accounted for in pre-processing. To solve this problem in background simulations, weighted histograms are made for background simulations to account for different number of events, cross section, pileup reweighting, and other corrections recommended by CMS. The pre-processing parameters μ_x and σ_x are extracted from the weighted histograms so those parameters only account for mis-modeling in simulation but not for lack of statistics.

While we train the IN, more signal events are used compared with the background. The different numbers of signal events and background events during training can make the trained model biased in the way that it tends to predict any event as the class with more events during training. To avoid this, weights are applied on signal and background events to make every signal event count as N_{bkg}/N_{sig} events so the signal and background are effectively equal during training.

We used an 85/15 splitting when training the model: 85% events out of the training events are used to update the parameters in the model and 15% events out of the training events are used to check the performance of ML after every iteration. The splitting is done in the following order: At the beginning of every iteration, we shuffle those events and randomly pick 15% of the events for validation and use the remaining 85% events to update the parameters in the model. Doing this makes the validation events different in every iteration to avoid over-training on the validation events.

With the training events prepared, the IN is trained by minimizing the loss function [47] iteration by iteration. As a typical loss function for the binary classifier, bi-

nary cross entropy L_{bce} [47] is used as the loss function. In order to avoid over-fitting, L2 regulation L_{reg} [48] is added to the loss function. In addition to discriminating signal events from background events, the IN is also used for background estimation, as described in Sec. 10. To ensure the background estimation developed based on the IN is valid, the ABCDisCo [49] method is used to decorrelate the output of the IN with the number of tracks in reconstructed vertices for background events. It is done by calculating the distance correlation L_{dcorr} [50] between the ML output and the number of tracks in vertices in the background events, then adding a penalty term based on L_{dcorr} in the loss function of the IN. So the more the output of the IN is correlated with the number of tracks in the vertex, the larger value the loss function outputs. When calculating L_{dcorr}, only one vertex is used per event. Most events have no more than one vertex after reconstruction. If there are more than one vertices reconstructed, the vertex with the highest number of tracks will be selected to use for decorrelation in the IN. Having all those terms for loss function together, the loss function of the IN is $L = L_{bce} + \lambda_{reg} \cdot L_{reg} + \lambda_{dcorr} \cdot L_{dcorr}$, where λ_{reg} and λ_{dcorr} are hyperparameters that control the strength of those restrictions.

In each iteration, the IN runs through all the training events by dividing all events into several batches with N_{batch} events in each batch and update the model parameters after running over each batch using the Adam [51] algorithm to minimize the loss function.

The model that results in the lowest value of the loss function on the validation events is saved as the final trained model. After the training is finished, the saved model is tested on the testing events to evaluate the performance.

Hyperparameters are tuned to ensure the decorrelation requirement is satisfied while having good discriminating power. All values of hyperparameters for the currently

used model are shown in Tab. 8.1.

Table 8.1: Hyperparameters in ML

Hyperparameter	value
λ_{reg}	0.00005
λ_{dcorr}	0.65
learning rate	0.0003
N_o	50
D_s	7
D_e	20
D_p	20
Number of epochs	250
N_{batch}	128

8.4 The IN Performance

The IN performance is evaluated in two aspects: the discriminating power and model independence. Stronger discriminating power helps in improving the sensitivity of the search by separating signal and background events apart, while model independence maintains the capability of the search being able to discover or set limits on not only the benchmark signal models but all other potential signal models that result in the target final state.

The testing events described in Sec. 8.3 are used to test the discriminating power of the IN. As described above, the training and testing events are selected to be orthogonal so all events used for training are not used for testing, which avoids biasing the testing result. Specifically, events with $\vec{E}^{\text{miss}}_{\text{T NoMu}} > 200\,\text{GeV}$ are used to check the IN performance. To evaluate the discriminating power, we compared the distribution of the MLScore between signal and background MC for events with at least one secondary vertex with at least five tracks, as shown in Fig. 8.5. As shown in the plots, the MLScore separates signal from background events in that most background

events have an MLScore of around 0 while most signal events have an MLScore of around 1. To ensure the ML performs consistently between data and the simulation, the MLScore distribution in data and simulated events are compared and consistent behavior is observed. The detailed comparisons are described in Appendix B.3. The obvious differences between the MLScore distributions of background and signal samples show that the ML learns track topology that cannot be easily captured by the secondary vertex reconstruction.

Meanwhile, a small bump around 0.4 is observed in MLScore distributions of signal samples, especially for signal models with large mass splitting and $c\tau$'s. The small bump should result from the restriction we applied to decorrelate the MLScore and the number of tracks in vertices: signal samples with larger mass splitting and lifetimes are more likely to have well-reconstructed vertices with more tracks, thus making the IN give such events smaller MLScores to satisfy the decorrelation restriction.

Overall, signal samples with larger track multiplicity, which are more likely to result in well-reconstructed vertices with more tracks, have more events with an MLScore around 0.4; signal samples with smaller track multiplicity, which are less likely to result in well-reconstructed vertices with more tracks, have fewer events with MLScore around 0.4.

The correlation between MLScore and variables related to vertex reconstruction is studied to understand the behavior of the IN. The correlation between MLScore and the number of tracks per vertex is visualized as the 2D distribution of the number of tracks per vertex vs. MLScore, as shown in Fig. 8.6. The plots show that the number of tracks per vertex stays stable for different MLScore, which proves that applying the ABCDisCo technique makes the MLScore and the number of tracks per vertex decorrelated successfully.

The correlation between MLScore and d_{BV} is studied by comparing 2D MLScore vs. d_{BV} distribution, as shown in Fig. 8.7. For signal events, as d_{BV} goes larger, more events will have MLScore around 0.4, which makes the average MLScore go lower. For background events, as d_{BV} goes larger, the MLScore will go slightly higher, but still much lower than signal events.

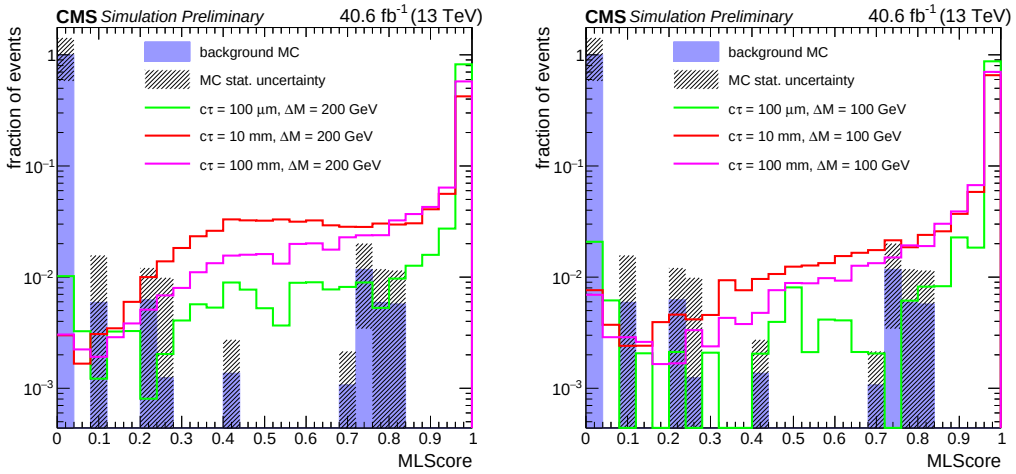

Figure 8.5: MLScore distribution for background MC and signal events with at least one secondary vertex with at least five tracks in 2017. Background (red), split SUSY samples with $c\tau$ as 100 μm (blue), 300 μm (green), and 1 mm (pink) for a given gluino mass and neutralino mass are shown together on the same canvas.
(Left) ML score comparison between background MC and split SUSY sample with gluino mass as 2000 GeV and neutralino mass as 1800 GeV.
(Right) ML score comparison between background MC and split SUSY sample with gluino mass as 2000 GeV and neutralino mass as 1900 GeV.

The model independence of the IN is tested by running the IN on different signal samples and checking the MLScore distribution. Two models are used for the test because their topologies are very different from that of the split SUSY model. The first model is the associated production of exotic Higgs [52] with W boson. The exotic Higgs decays to a pair of scalar particles, which are long-lived and decay to four b quarks or d quarks in total. The mass of the scalar particles (S) used in this test is 55 GeV. The process is symbolized as $WH \rightarrow SS \rightarrow bbbb/dddd$. The second model

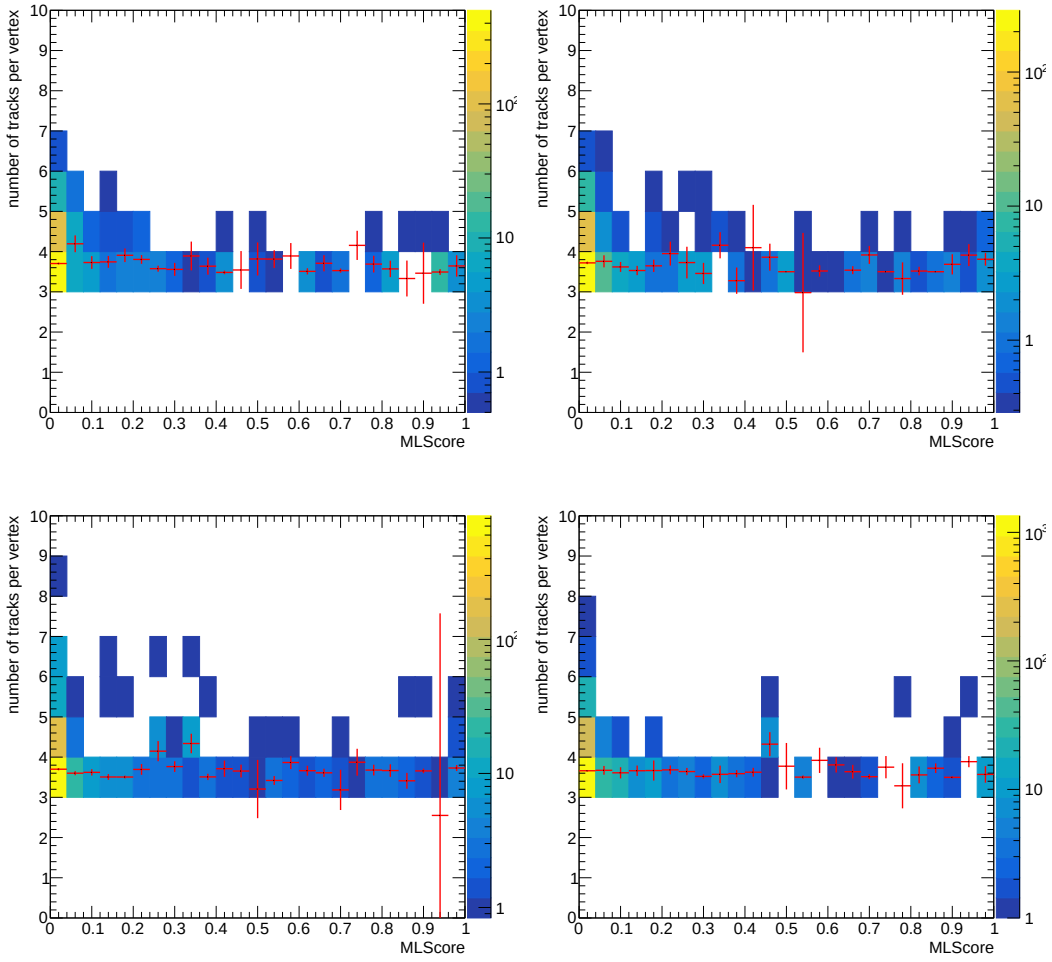

Figure 8.6: Number of tracks per vertex vs. MLScore for background MC in 2016 preVFP (top left), 2016 postVFP (top right), 2017 (bottom left), and 2018 (bottom right). The profile plot (red) that shows the average Number of tracks per vertex as a function of MLScore is overlaid on the 2D plot.

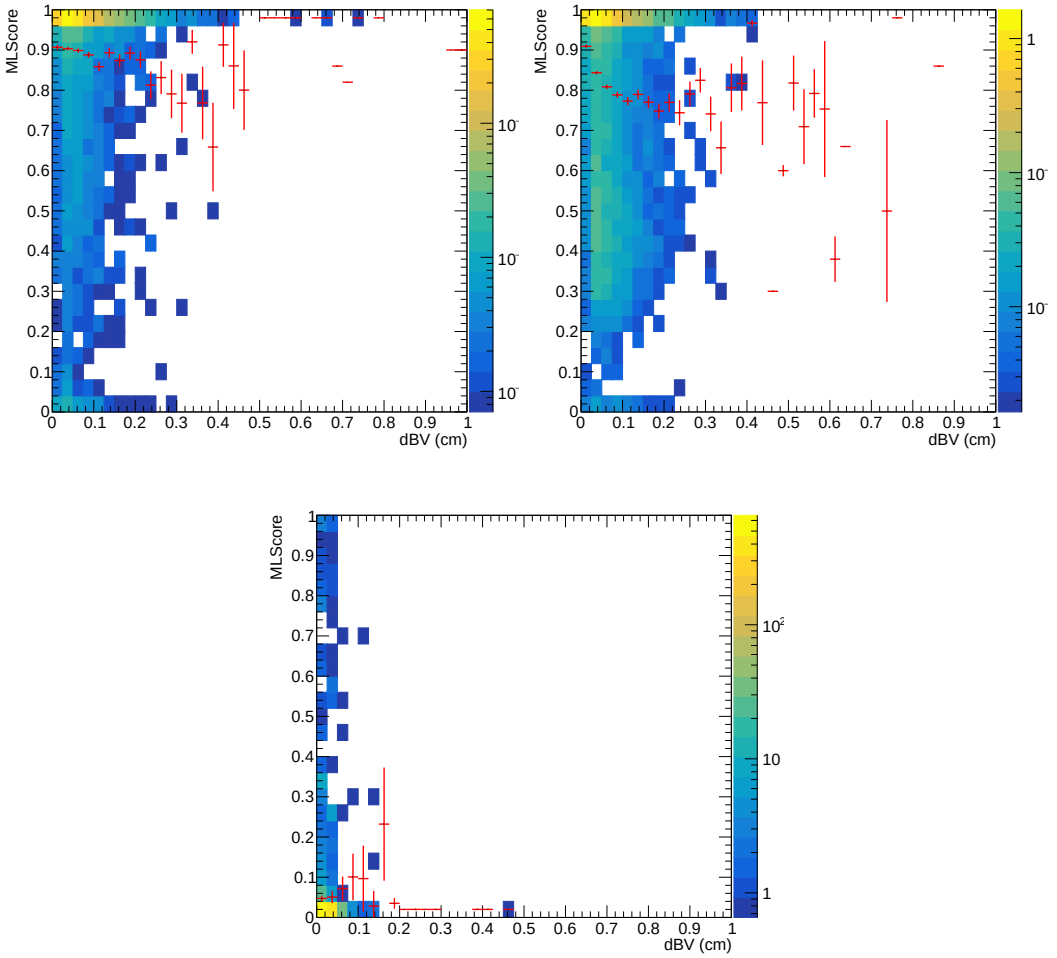

Figure 8.7: MLScore vs. d_{BV} background MC and signal for events with 3, 4, or 5-track vertices in 2017. The profile plot (red) that shows the average MLScore as a function of d_{BV} is overlaid on the 2D plot.
(Top left) Split SUSY samples with 100 GeV mass splitting;
(Top right) Split SUSY samples with 200 GeV mass splitting;
(Bottom) Background simulations.

is the R-parity violated SUSY [53]. In this model, neutralinos ($\tilde{\chi_0}$) are pair produced and long-lived. Each of the neutralinos decays to t, b, and s quarks. The neutralino masses used in this test are 400 GeV and 1600 GeV. The model is symbolized as $\tilde{\chi_0} \to tbs$.

All events used in the test satisfy the testing selection described in 8.3. For some models, the number of events available is limited because of the low trigger efficiency and vertex reconstruction efficiency. The MLScore distributions for models described above are shown in Fig. 8.8 for $WH \to SS \to bbbb/dddd$ and Fig. 8.9 for $\tilde{\chi_0} \to tbs$. As shown in the plots, the MLScore distributions behave consistently as described before: for low track multiplicity signal samples like $WH \to SS \to bbbb/dddd$, most events have MLScore around 1; for high-track multiplicity signal samples like $\tilde{\chi_0} \to tbs$, part of the events have an MLScore of around 0.4, and part of the events have an MLScore of around 1.

As a result, the model independence study shows that the IN behaves consistently for different topologies with displaced vertices resulting from long-lived particles with different masses.

In conclusion, the IN has a strong discriminating power to select signal events out of background events and it is capable to recognize signal models with different topologies. Those features ensure that the IN helps improve the sensitivity and keep the model-independent feature of this search.

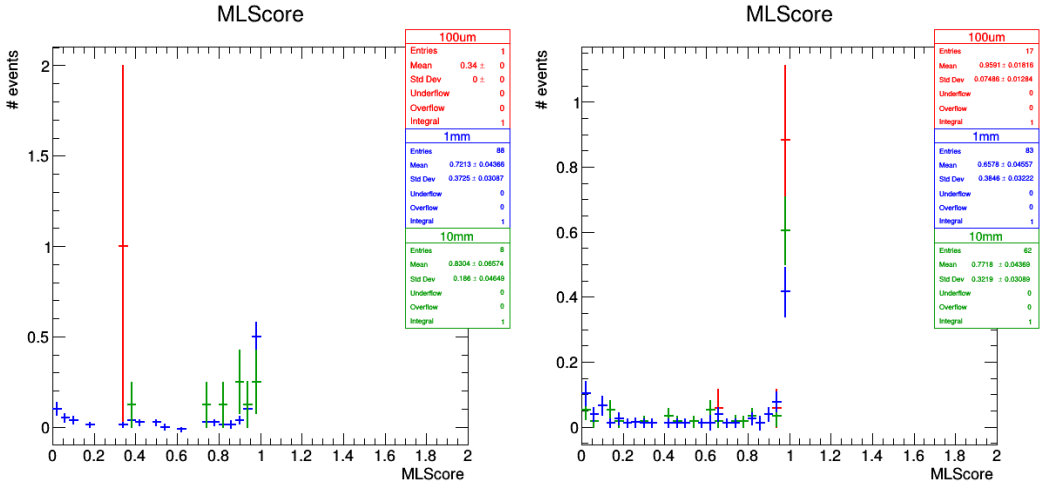

Figure 8.8: MLScore distribution for WH samples in 2017. WH samples with different $c\tau$'s as 100 μm (red), 1 mm (blue), and 10 mm (green) for a given decay channel are shown together on the same canvas.
(Left) MLScore distributions for $WH \rightarrow SS \rightarrow bbbb$.
(Right) MLScore distributions for $WH \rightarrow SS \rightarrow dddd$.

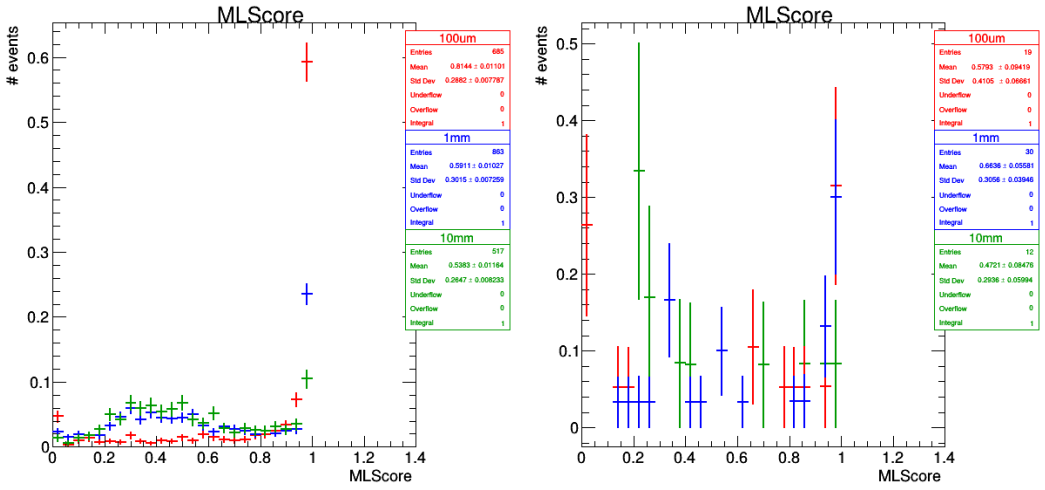

Figure 8.9: MLScore distribution for $\tilde{\chi_0} \rightarrow tbs$ samples in 2017. Different $c\tau$'s of 100 μm (red), 1 mm (blue), and 10 mm (green) for a given decay channel are shown together on the same canvas.
(Left) MLScore distributions for $\tilde{\chi_0}$ mass of 400 GeV.
(Right) MLScore distributions for $\tilde{\chi_0}$ mass of 1600 GeV.

8.5 The IN Selection

To pick a reasonable selection on the MLScore for the search, different selection thresholds ranging from 0.15 to 0.25 are applied on events with all preselections[1] and vertex requirements applied to check the effect of varying MLScore selection, as shown in Tab. 8.2, 8.3, and 8.4 for selection thresholds at 0.15, 0.20, and 0.25. The uncertainties in the table are purely statistical uncertainties.

The tables show that loosening the ML selection to 0.15 makes the number of background events increase by about 30% or 40% for events with vertices that have exactly 3 tracks (labeled as 3-track) and events with vertices that have exactly 4 tracks (labeled as 4-track) events, which might result in too many background events in the search. Tightening the ML selection can only help decrease the number of background events by about 15% in 3-track events but 0% in 4-track events, which may not be very helpful in further reducing background events. In the meanwhile, tightening the selection threshold could also result in the possibility of cutting off an obvious fraction of heavy mass LLP signal samples, such as the $\tilde{\chi}_0 \rightarrow tbs$ signal samples shown in Fig 8.9. As a result, to avoid including too many background events while maintaining sensitivity to a wide range of signal models, an MLScore of 0.2 is used to separate signal and background events.

ML selection = 0.15	4-track	3-track
Measured high ML events	17.00 ± 4.12	88.00 ± 9.38
Measured low ML events	399.00 ± 19.98	2072.00 ± 45.52

Table 8.2: Number of events in 2017 data with ML selection at 0.15.

[1] the fake MET veto threshold used in this study is 0.5, which is slightly different from the one used in the search 0.6

ML selection = 0.20	4-track	3-track
Measured high ML events	12.00 ± 3.46	66.00 ± 8.12
Measured low ML events	404.00 ± 20.10	2094.00 ± 45.76

Table 8.3: Number of events in 2017 data with ML selection at 0.20.

ML selection = 0.25	4-track	3-track
Measured high ML events	12.00 ± 3.46	55.00 ± 7.42
Measured low ML events	404.00 ± 20.10	2105.00 ± 45.88

Table 8.4: Number of events in 2017 data with ML selection at 0.25.

Chapter 9

Analysis Selection

This search is a counting experiment. A set of selections are applied to get rid of background events while maintaining most of the signal events. The region formed by applying all the selections in the search is referred to as the *signal region*, which is most likely to include signal events. The search strategy is then to predict the number of events based on the assumption that there are no BSM physics processes present and compare the observed and predicted number of events in the signal region to determine whether the observation and prediction match with each other. *Validation region* and *control regions* are used to verify the search strategy, predict the number of events in the signal region, and study the behavior of background events. Those regions are defined by applying almost all the selections but with one of the selections reversed or relaxed.

The analysis selection is defined by integrating the event preselection described in Sec. 6.3 with vertex reconstruction described in Ch. 7. Events must:

- Fire the `HLT_PFMETNoMu120_PFMHTNoMu120_IDTight` trigger;
- Pass Fake MET veto: $|MET_{PF} - MET_{CALO}|/MET_{CALO} < 0.6$;
- Have reconstructed $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ pass all MET filters;
- Have $\vec{E}_{\mathrm{T\ NoMu}}^{\mathrm{miss}} > 200\,\mathrm{GeV}$;

- Have at least 1 reconstructed vertices that satisfy:
  - at least 3 tracks;
  - x-y distance of the vertex from the (year-dependent) beam pipe center of less than 20 mm for 2016 and 20.9mm for 2017 and 2018;
  - uncertainty in d_{BV} of less than 25 μm.

Figure 9.1 shows the distribution of $\vec{E}^{\text{miss}}_{\text{T NoMu}}$ after the analysis selection is applied.

With the addition of the ML described in Ch. 8 developed in the search as the final piece of the selection, the signal region, validation region, and control regions are defined based on the number of tracks in the leading reconstructed vertex (nTracks/SV) and the output of ML (MLScore), as shown in Fig. 9.2:

- Signal Region
  - A (5-track highML): nTracks/SV$\geq$5, MLScore$\geq$0.2
- Validation Region
  - C (4-track highML): nTracks/SV=4, MLScore$\geq$0.2
- Control Regions
  - B (5-track lowML): nTracks/SV$\geq$5, MLScore$<$0.2
  - D (4-track lowML): nTracks/SV=4, MLScore$<$0.2
  - E (3-track highML): nTracks/SV=3, MLScore$\geq$0.2
  - F (3-track lowML): nTracks/SV=3, MLScore$<$0.2

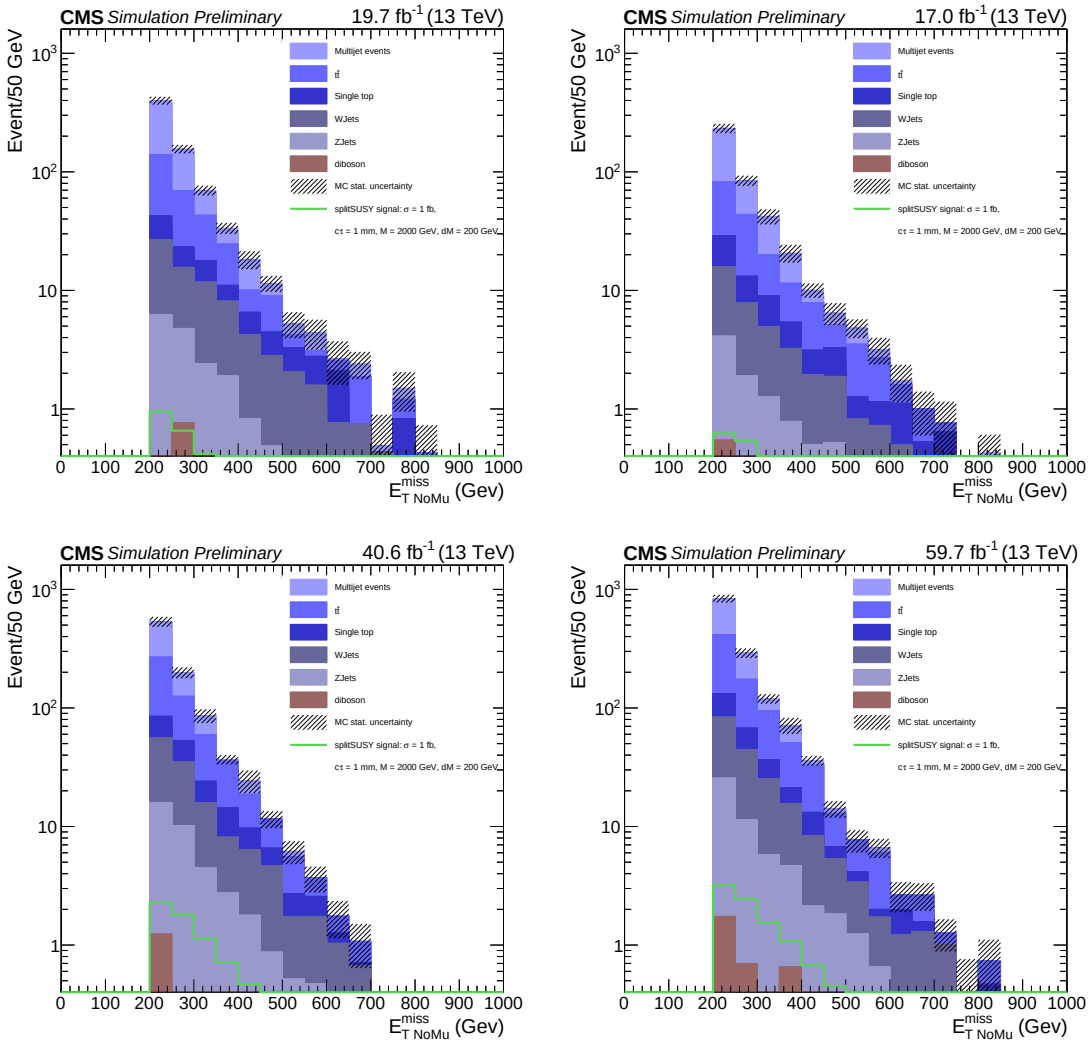

Figure 9.1: Comparisons of distributions of $\vec{E}^{\text{miss}}_{\text{T NoMu}}$ for events that pass the analysis selection between simulated background events and simulated split SUSY signal events with gluino $c\tau = 1\,\text{mm}$, $M = 2000\,\text{GeV}$, neutralino $M = 1800\,\text{GeV}$, and production cross section 1 fb in 2016 preVFP (top left), 2016 postVFP (top right), 2017 (bottom left), and 2018 (bottom right).

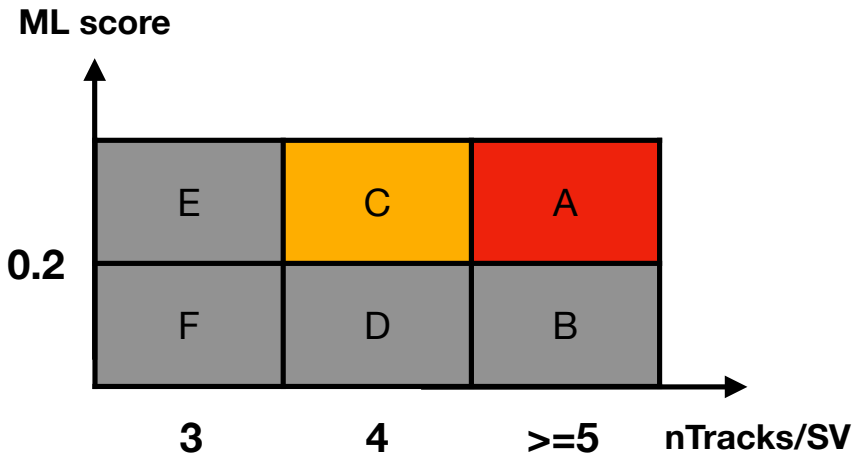

Figure 9.2: Signal region, validation regions, and control regions are shown in the cartoon as red, yellow, and grey blocks.

The cutflow plots and tables that show the number of signal and background MC events after each single event selection for 2016 preVFP, 2016 postVFP, 2017, 2018, and Run 2 are shown in Fig. 9.3 and Tab. 9.1, 9.2, 9.3, 9.4, 9.5.

	background MC	signal $c\tau$=1 mm	
		Δm=100 GeV	Δm=200 GeV
Trigger & at least one 3 track vertices	75409.06	3.24	7.39
fake MET veto & MET filters	52163.56	2.66	6.47
$\vec{E}^{\text{miss}}_{\text{T NoMu}}$ >200 GeV	12094.78	1.84	4.78
Vertices with $\sigma_{d_{BV}}$ <25um	696.08	0.85	3.03
Vertices with at least 5 tracks	12.88	0.67	2.69
MLScore>0.2	0.73	0.66	2.67

Table 9.1: Number of events after each selection for background MC and signal with $m_{\tilde{g}}$ of 2000 GeV, $c\tau$ of 1 mm, and different mass splittings in 2016 preVFP.

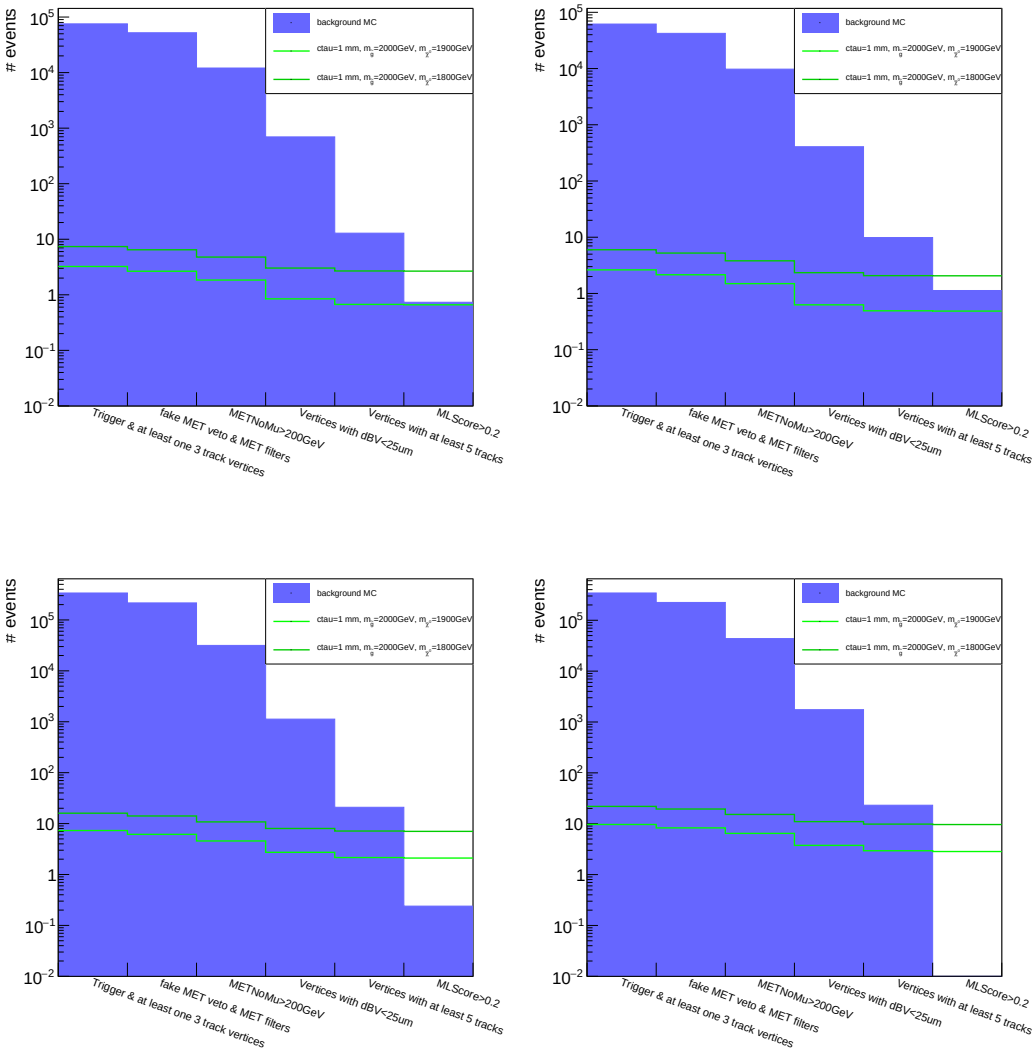

Figure 9.3: Cutflow plots that show the number of signal and background events after each step of event selection. The signal used in the plots is split SUSY samples with gluino mass 2000 GeV and $c\tau$ 1 mm. In each plot, mass splitting as 100 GeV and 200 GeV are shown together.
(Top left) 2016 preVFP;
(Top right) 2016 postVFP;
(Bottom left) 2017;
(Bottom right) 2018.

	background MC	signal $c\tau$=1 mm	
		Δm=100 GeV	Δm=200 GeV
Trigger & at least one 3 track vertices	61847.41	2.63	5.99
fake MET veto & MET filters	42247.89	2.15	5.23
$\vec{E}^{\text{miss}}_{\text{T NoMu}}$ >200 GeV	9761.40	1.50	3.81
Vertices with $\sigma_{d_{BV}}$ <25um	407.24	0.63	2.34
Vertices with at least 5 tracks	9.87	0.49	2.07
MLScore>0.2	1.13	0.49	2.06

Table 9.2: Number of events after each selection for background MC and signal with $m_{\tilde{g}}$ of 2000 GeV, $c\tau$ of 1 mm, and different mass splittings in 2016 postVFP.

	background MC	signal $c\tau$=1 mm	
		Δm=100 GeV	Δm=200 GeV
Trigger & at least one 3 track vertices	338357.00	7.30	16.00
fake MET veto & MET filters	216081.84	6.15	14.11
$\vec{E}^{\text{miss}}_{\text{T NoMu}}$ >200 GeV	31611.00	4.56	10.79
Vertices with $\sigma_{d_{BV}}$ <25um	1127.27	2.73	8.01
Vertices with at least 5 tracks	20.92	2.15	7.14
MLScore>0.2	0.24	2.10	7.03

Table 9.3: Number of events after each selection for background MC and signal with $m_{\tilde{g}}$ of 2000 GeV, $c\tau$ of 1 mm, and different mass splittings in 2017.

	background MC	signal $c\tau$=1 mm	
		Δm=100 GeV	Δm=200 GeV
Trigger & at least one 3 track vertices	343956.03	9.67	21.91
fake MET veto & MET filters	224057.13	8.28	19.45
$\vec{E}^{\text{miss}}_{\text{T NoMu}}$ >200 GeV	43643.50	6.46	15.23
Vertices with $\sigma_{d_{BV}}$ <25um	1747.35	3.75	11.00
Vertices with at least 5 tracks	23.12	2.95	9.83
MLScore>0.2	-0.07	2.84	9.64

Table 9.4: Number of events after each selection for background MC and signal with $m_{\tilde{g}}$ of 2000 GeV, $c\tau$ of 1 mm, and different mass splittings in 2018.

	background MC	signal $c\tau$=1 mm	
		Δm=100 GeV	Δm=200 GeV
Trigger & at least one 3 track vertices	819569.51	22.84	51.29
fake MET veto & MET filters	534550.42	19.24	45.25
$\vec{E}^{\text{miss}}_{\text{T NoMu}}$ >200 GeV	97110.67	14.36	34.61
Vertices with $\sigma_{d_{BV}}$ <25um	3977.94	7.95	24.39
Vertices with at least 5 tracks	66.78	6.26	21.73
MLScore>0.2	2.03	6.09	21.40

Table 9.5: Number of events after each selection for background MC and signal with $m_{\tilde{g}}$ of 2000 GeV, $c\tau$ of 1 mm, and different mass splittings in Run 2.

Chapter 10

Background Estimation

10.1 Source of Background

There are two main categories of background vertices:

- B decay vertices: vertices reconstructed from B meson decays;
- Track random crossing vertices: vertices reconstructed from irrelevant tracks crossing through each other by coincidence.

Figure 10.1 compares the $\sigma_{d_{BV}}$ and $\Delta\phi$ of the vector connecting from the beam spot to the vertex and b-jets between the two categories. B-jets are selected as described in Sec. 6.1 using the tight working point ($tight\ b-jets$). As shown in the plots, vertices with $\sigma_{d_{BV}} \geq 25\,\mu$m tend to have $\Delta\phi$ around 0 with tight b-jets, which means that they overlap with b-jets and thus are dominated by B decay vertices. For vertices with $\sigma_{d_{BV}} < 25\,\mu$m, the vertex $\Delta\phi$ to tight b-jets distribution shows that they are mostly not overlapping with tight b-jets so they are dominated by track random crossing vertices.

The reason for such a phenomenon is that B mesons are generally boosted so their decay products are more collimated compared to track random crossing vertices, which results in a higher $\sigma_{d_{BV}}$.

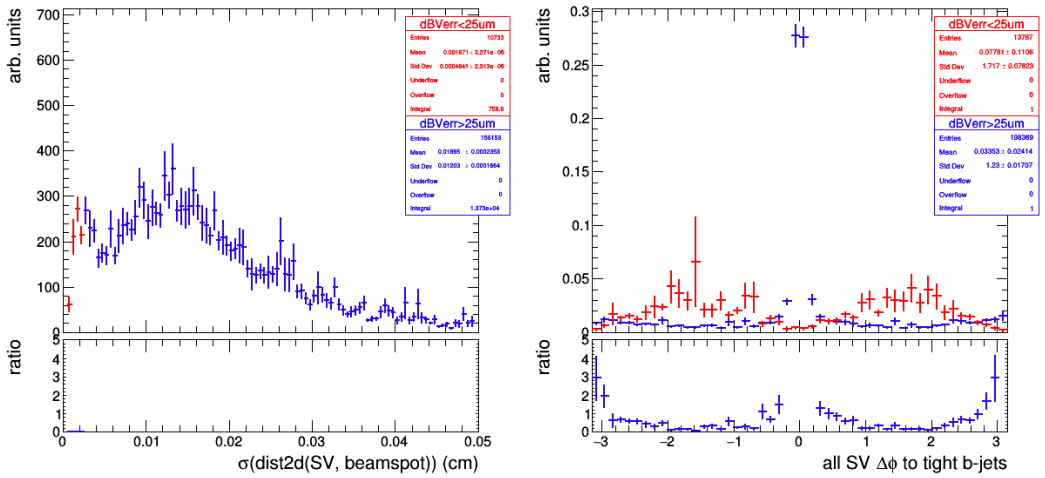

Figure 10.1: Comparison between vertices with $\sigma_{d_{BV}}$ smaller and larger than 25 μm shows different types of background vertices for 2017.

10.2 Background Estimation Method

The search targets displaced vertices, which is a unique signature that might be mis-modeled in the simulation. To avoid potential uncertainty introduced by the mis-modeling, a data-driven background estimation method that does not depend on the simulation is developed based on the ABCD method. The background estimation method uses signal region, validation, and control regions defined by nTracks/SV and MLScore, as described in Sec. 9.

ABCD method requires the two discriminants that define the search regions to be statistically uncorrelated with each other. The uncorrelation is checked by comparing the MLScore distribution between events with nTracks/SV as 3, 4, ≥ 5 for background simulation, as shown in Fig. 10.2. To see the distribution better for events with low MLScores, the distributions in Fig. 10.2 are zoomed in MLScore range (0,0.4), as shown in Fig. 10.3. All the plots used fine binning in the small MLScore region and coarse binning in the large MLScore region because the statistical uncertainty

in the large MLScore region is large. The comparisons of MLScore distributions for events with different nTracks/SV shown in Fig. 10.2 and nTracks/SV vs. MLScore 2D distributions and profiles shown in Fig. 8.6 prove that MLScore and nTracks/SV are statistically uncorrelated.

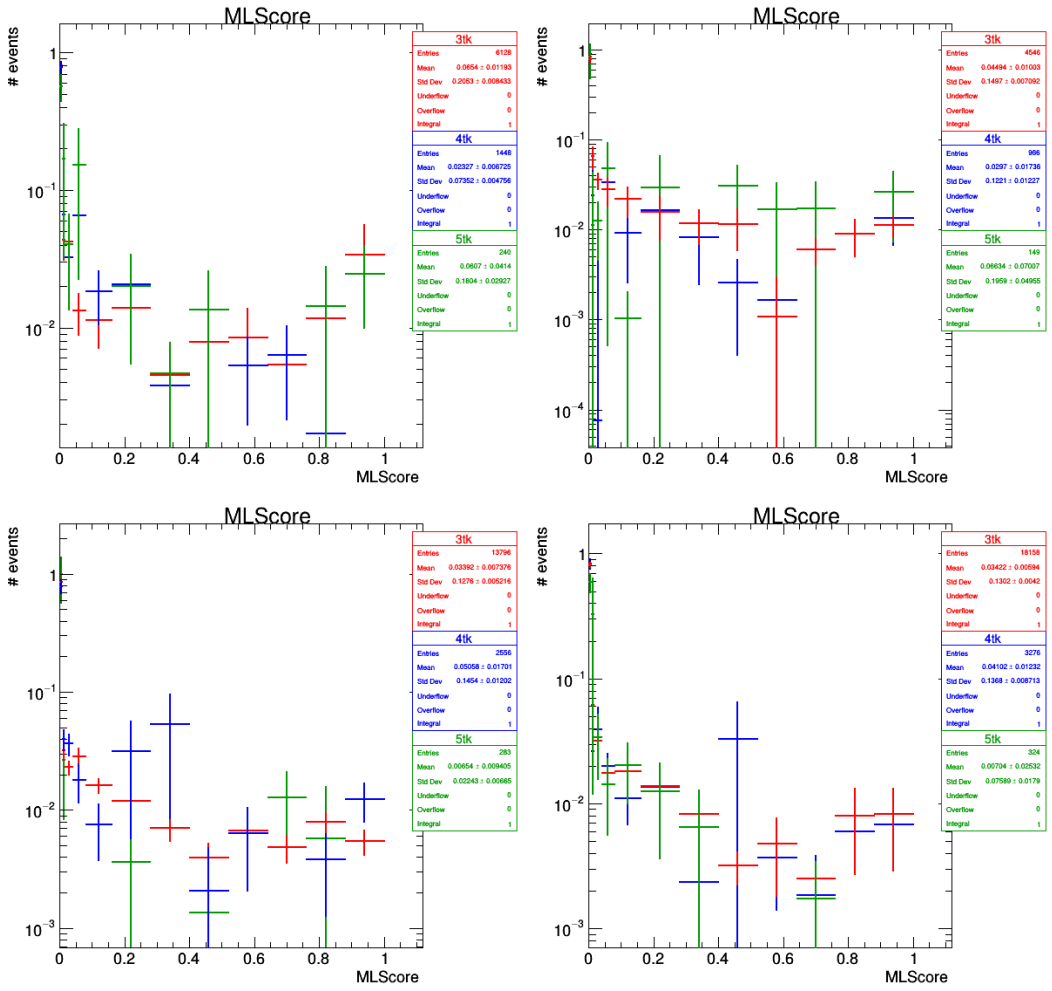

Figure 10.2: MLScore distribution for simulated background events with 3-track (red), 4-track (blue), and ≥5-track (green) vertices in 2016 preVFP, 2016 postVFP, 2017, and 2018.

When the two discriminators are independent of each other, the number of background events in the signal region and validation region is predicted based on the number of events measured in 3-track regions by $N^A_{bkg} = \frac{N^B_{bkg} \cdot N^E_{bkg}}{N^F_{bkg}}$ and $N^C_{bkg} = \frac{N^D_{bkg} \cdot N^E_{bkg}}{N^F_{bkg}}$.

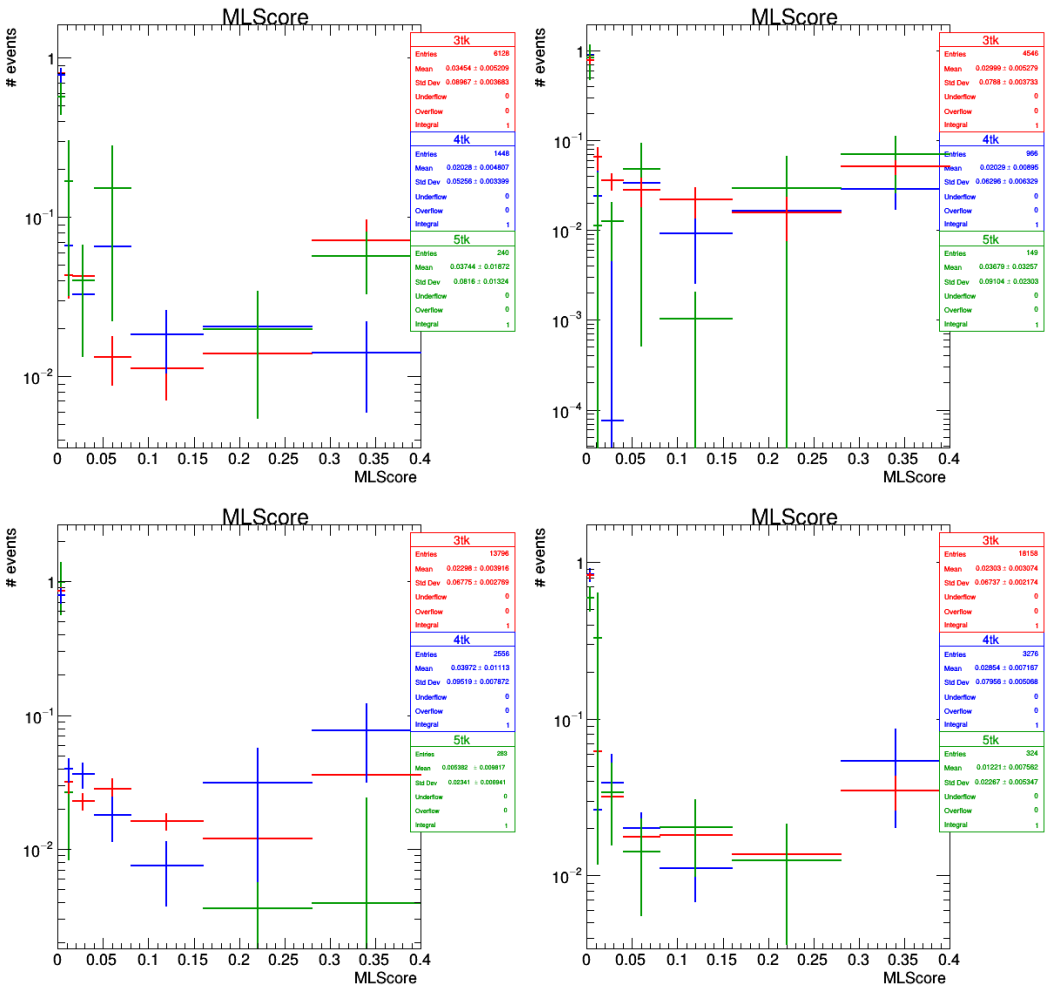

Figure 10.3: MLScore distribution that are zoomed in range (0,0.4) for simulated background events with 3-track (red), 4-track (blue), and ≥5-track (green) vertices in 2016 preVFP, 2016 postVFP, 2017, and 2018.

10.3 Signal Contamination

In the ABCD method, the calculation of expected background events in the signal region is based on the assumption that the control regions are dominated by background events. A signal contamination check is performed to ensure the validation and control regions are dominated by background events.

Two different split SUSY samples with the same $c\tau$ from 100 μm to 1000 mm but different gluino masses and different cross sections in 2017 are used to perform the check:

- Gluino mass 2000 GeV with cross section 1.01 fb^{-1}, which represents the target sensitivity of this analysis
- Gluino mass 1400 GeV with cross section 28.4 fb^{-1}, which represents the largest unexcluded cross-section

The sample that contributes most in the control region B ($\geq$5-track with low ML) is reported in Table 10.1, 10.2, 10.3, 10.4, and 10.5.

Number of events	data	$m(\tilde{g}) = 2000$ GeV	$m(\tilde{g}) = 1400$ GeV
A(5-track highML)	3.00 $\pm$ 1.73	0.506 $\pm$ 0.010	17.091 $\pm$ 0.310
B(5-track lowML)	56.00 $\pm$ 7.48	0.010 $\pm$ 0.001	0.235 $\pm$ 0.036
C(4-track highML)	12.00 $\pm$ 3.46	0.177 $\pm$ 0.006	4.935 $\pm$ 0.166
D(4-track lowML)	456.00 $\pm$ 21.35	0.008 $\pm$ 0.001	0.296 $\pm$ 0.041
E(3-track highML)	75.00 $\pm$ 8.66	0.099 $\pm$ 0.004	3.115 $\pm$ 0.132
F(3-track lowML)	2311.00 $\pm$ 48.07	0.022 $\pm$ 0.002	0.565 $\pm$ 0.056

Table 10.1: Signal contribution in signal, validation, and control regions in 2017 for 100 μm $c\tau$.

As shown in the tables, for the target cross section, the signal events contribute less than one event in validation and control regions, which shows that those regions are

Number of events	data	$m(\tilde{g}) = 2000\,\text{GeV}$	$m(\tilde{g}) = 1400\,\text{GeV}$
A(5-track highML)	3.00 ± 1.73	7.444 ± 0.077	207.446 ± 2.158
B(5-track lowML)	56.00 ± 7.48	0.112 ± 0.009	3.727 ± 0.290
C(4-track highML)	12.00 ± 3.46	0.627 ± 0.022	16.601 ± 0.610
D(4-track lowML)	456.00 ± 21.35	0.017 ± 0.004	0.493 ± 0.105
E(3-track highML)	75.00 ± 8.66	0.261 ± 0.014	6.443 ± 0.380
F(3-track lowML)	2311.00 ± 48.07	0.009 ± 0.003	0.503 ± 0.106

Table 10.2: Signal contribution in signal, validation, and control regions in 2017 for $1\,\text{mm}$ $c\tau$.

Number of events	data	$m(\tilde{g}) = 2000\,\text{GeV}$	$m(\tilde{g}) = 1400\,\text{GeV}$
A(5-track highML)	3.00 ± 1.73	10.576 ± 0.092	287.655 ± 2.539
B(5-track lowML)	56.00 ± 7.48	0.180 ± 0.012	5.521 ± 0.353
C(4-track highML)	12.00 ± 3.46	0.475 ± 0.020	13.283 ± 0.546
D(4-track lowML)	456.00 ± 21.35	0.004 ± 0.002	0.225 ± 0.071
E(3-track highML)	75.00 ± 8.66	0.129 ± 0.010	3.853 ± 0.295
F(3-track lowML)	2311.00 ± 48.07	0.004 ± 0.002	0.080 ± 0.041

Table 10.3: Signal contribution in signal, validation, and control regions in 2017 for $10\,\text{mm}$ $c\tau$.

Number of events	data	$m(\tilde{g}) = 2000\,\text{GeV}$	$m(\tilde{g}) = 1400\,\text{GeV}$
A(5-track highML)	3.00 ± 1.73	5.202 ± 0.065	135.079 ± 1.739
B(5-track lowML)	56.00 ± 7.48	0.071 ± 0.008	1.457 ± 0.181
C(4-track highML)	12.00 ± 3.46	0.362 ± 0.017	9.212 ± 0.454
D(4-track lowML)	456.00 ± 21.35	0.006 ± 0.002	0.068 ± 0.039
E(3-track highML)	75.00 ± 8.66	0.104 ± 0.009	2.618 ± 0.242
F(3-track lowML)	2311.00 ± 48.07	0.002 ± 0.001	0.046 ± 0.032

Table 10.4: Signal contribution in signal, validation, and control regions in 2017 for $100\,\text{mm}$ $c\tau$.

Number of events	data	$m(\tilde{g}) = 2000\,\text{GeV}$	$m(\tilde{g}) = 1400\,\text{GeV}$
A(5-track highML)	blinded	0.719 ± 0.012	17.788 ± 0.315
B(5-track lowML)	56.00 ± 7.48	0.012 ± 0.002	0.302 ± 0.041
C(4-track highML)	12.00 ± 3.46	0.054 ± 0.003	1.380 ± 0.088
D(4-track lowML)	456.00 ± 21.35	0.002 ± 0.001	0.017 ± 0.010
E(3-track highML)	75.00 ± 8.66	0.017 ± 0.002	0.460 ± 0.050
F(3-track lowML)	2311.00 ± 48.07	0.002 ± 0.001	0.006 ± 0.006

Table 10.5: Signal contribution in signal, validation, and control regions in 2017 for $1000\,\text{mm}$ $c\tau$.

dominated by background events. For the largest unexcluded cross section, there is some signal contamination in validation E and control region B, but the huge number of events in the signal region will lead to an obvious discovery.

10.4 Closure Test

To check whether the ABCD background estimation method works for this search, closure tests are performed in the validation region C for data events by comparing the prediction and observation of the number of events in the validation region C. The closure tests results are shown in Tab. 10.6, 10.7, 10.8, 10.9, and 10.10 for 2016 preVFP, 2016 postVFP, 2017, 2018, and the full Run 2 added together. The results show that the ABCD closure test works nicely for data in all four data-taking campaigns.

As a result, a purely data-driven background estimation method is developed and the closure test for validation regions works well in data for all years.

	4-track	3-track
Predicted high ML events	5.16 ± 1.17	N/A
Observed high ML events	5.00 ± 2.24	25.00 ± 5.00
Observed low ML events	109.00 ± 10.44	528.00 ± 22.98

Table 10.6: Closure test results for data in 2016 preVFP.

	4-track	3-track
Predicted high ML events	3.12 ± 0.91	N/A
Observed high ML events	5.00 ± 2.24	14.00 ± 3.74
Observed low ML events	83.00 ± 9.11	373.00 ± 19.31

Table 10.7: Closure test results for data in 2016 postVFP.

	4-track	3-track
Predicted high ML events	14.80 ± 1.87	N/A
Observed high ML events	12.00 ± 3.46	75.00 ± 8.66
Observed low ML events	456.00 ± 21.35	2311.00 ± 48.07

Table 10.8: Closure test results for data in 2017.

	4-track	3-track
Predicted high ML events	17.94 ± 2.06	N/A
Observed high ML events	16.00 ± 4.00	89.00 ± 9.43
Observed low ML events	628.00 ± 25.06	3115.00 ± 55.81

Table 10.9: Closure test results for data in 2018.

	4-track	3-track
Predicted high ML events	40.94 ± 3.14	N/A
Observed high ML events	38.00 ± 6.16	203.00 ± 14.25
Observed low ML events	1276.00 ± 35.72	6327.00 ± 79.54

Table 10.10: Closure test results for data in Run 2.

Chapter 11

Systematic Uncertainties

Systematic uncertainties are used in the search to account for factors that could possibly have an impact on signal efficiency or background estimation. Generally, systematic uncertainties are evaluated individually for different factors. For each factor, its variation is propagated to estimate the resulting change in the signal efficiency.

11.1 Correction Factors and Systematic Uncertainties Related to Signal Distribution and Efficiency

11.1.1 Track Reconstruction

The dominant sources of uncertainty in signal efficiency arise from the differences in the reconstruction of tracks and vertices in data and simulation. The study relevant to track reconstruction is described in this section and the study of vertex reconstruction is described in Sec. 11.1.2.

Track reconstruction efficiency is studied by reconstructing vertices from K_S^0 decays. The procedure is modified based on the study described in [26]. The study focuses on K_S^0's that decay to π^- and π^+, resulting in two tracks. The track reconstruction

efficiency difference between data and simulation will then be inferred from the K_S^0 reconstruction efficiency.

The events used in this study are required to satisfy the event preselection described in Sec. 6.3. Every pair of tracks in the events is used to reconstruct vertices and vertices that satisfy $\chi^2/\text{dof} < 7$ are kept. To get rid of most background vertices, vertices are further required to:

- Have $c\tau > 268\,\mu$m;
- Have $\rho > 0.268\,$cm;
- Have the cosine of the 2D angle between fitted candidate momentum and flight direction from primary vertex > 0.99975;
- Be within the beam pipe.

Vertices that satisfy the selection are divided into two categories depending on the invariant mass:

- Vertices with mass in the range [0.490, 0.505] GeV: used as "signal region" that most of the K0 candidates reside in;
- Vertices with mass out of range [0.490, 0.505] GeV: used as side band to estimate the background in the "signal region".

The invariant mass distribution of K_S^0 vertices is shown in Fig. 11.1. To get rid of background vertices, a linear fit is performed using the distribution in the side band. The fit result is used to calculate the expected number of background vertices in the K_S^0 signal region. The number of vertices that come from K_S^0's is calculated from the

difference between the total number of vertices and expected background vertices in the K_S^0 signal region. The purity of K_S^0 vertices, calculated as the ratio of the number of vertices from K_S^0 decays and the total number of vertices, are reported in Tab. 11.1. The similarity of the purity in data and simulation shows that data and simulated events used in the study behave consistently.

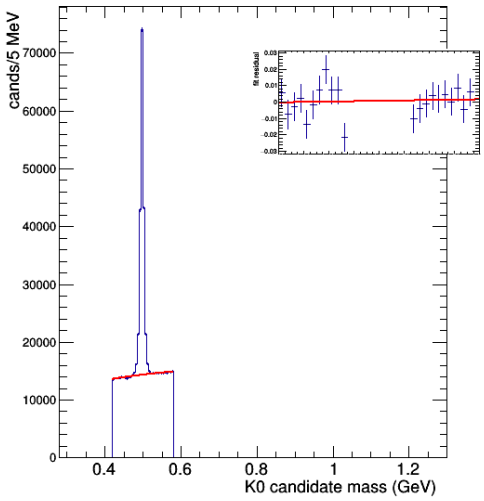

Figure 11.1: The invariant mass distribution of reconstructed vertices in 2017. The red line shows the fit result of background vertices using the side band.

year	data	simulation
2016 preVFP	0.69	0.68
2016 postVFP	0.66	0.67
2017	0.73	0.74
2018	0.73	0.74

Table 11.1: The purity of K_S^0 vertices, calculated as the ratio of the number of vertices from K_S^0 decays and the total number of vertices in the K_S^0 signal region.

Distributions in the K_S^0 signal region are compared between data and simulation, as shown in Fig. 11.2. The plots show that tracks and reconstructed K_S^0 vertices are generally less energetic in data.

Since this study is designed to only account for the track reconstruction efficiency, dif-

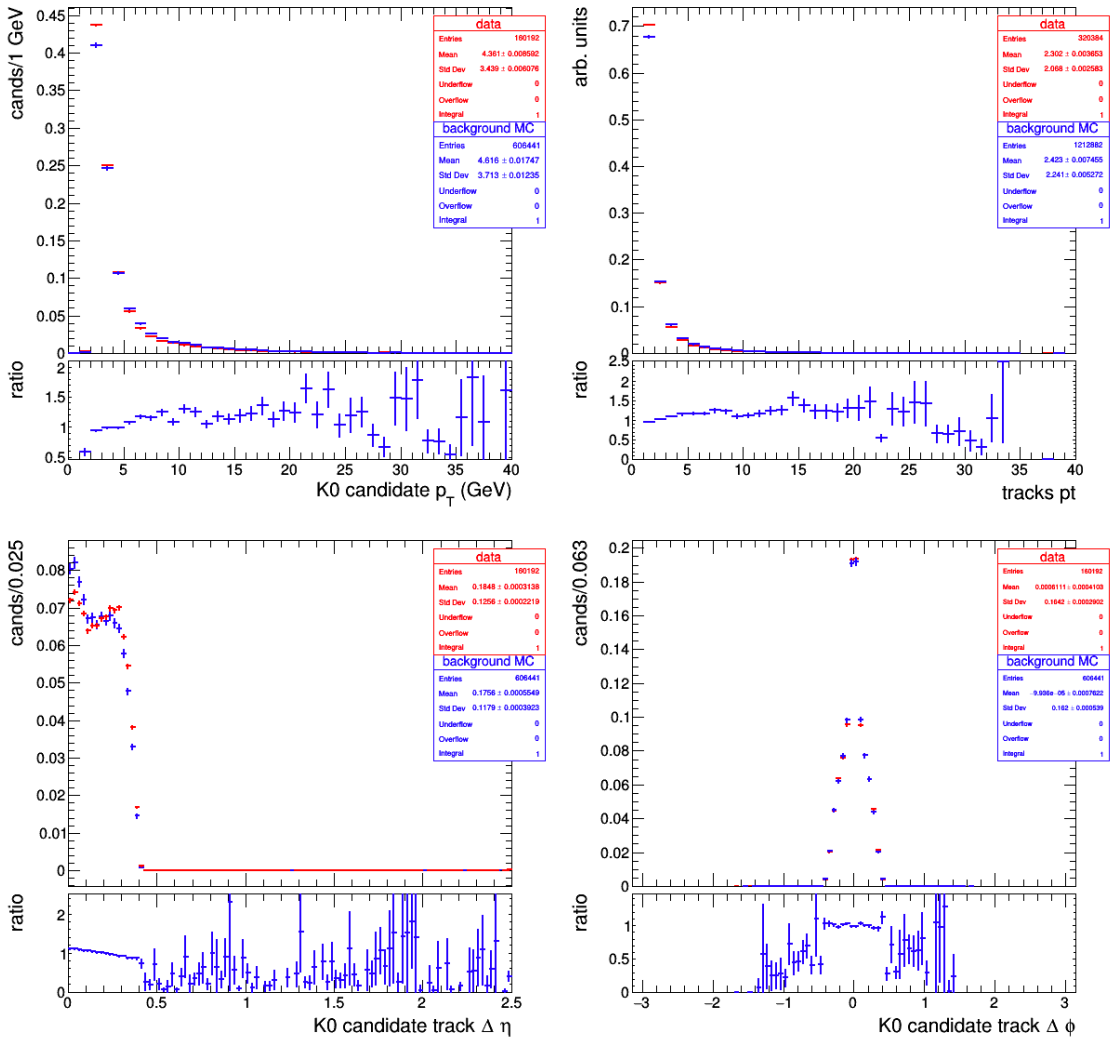

Figure 11.2: Comparisons of kinematic variables of vertices in 2017 between data (red) and simulation (blue) before corrections are applied.

ferences between data and simulation that are not relevant to the track reconstruction efficiency, such as the p_{T} of tracks and reconstructed K_S^0 vertices, should be mitigated to get a more precise result. The difference in track p_{T} is mitigated by applying correction weights on simulated events. Weights are calculated as the data/MC ratio of K_S^0 vertex track p_{T} distribution in the signal region. To avoid the effect of the mis-modeling of background events in the simulation, the distribution is obtained by subtracting the background distribution from the distribution of the K_S^0 signal region. The background distribution is obtained by normalizing the distribution of the side band to the number of expected background vertices in the K_S^0 signal region. The K_S^0 track p_{T} distribution in data and simulation are compared in Fig. 11.3. Distributions after the weights are applied are shown in Fig. 11.4. As shown in the plots, applying the corrections mitigates most of the inconsistency of the K_S^0 vertices between data and simulation.

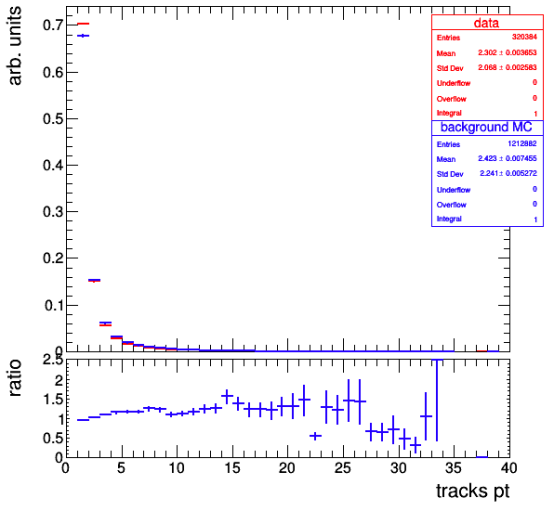

Figure 11.3: Comparison of K_S^0 vertex track p_{T} between data and MC in 2017. The derived correction weights are shown in the lower panel.

The transverse distance between the K_S^0 vertices and the beam spot distribution is compared between data and simulation and is used to infer the track reconstruc-

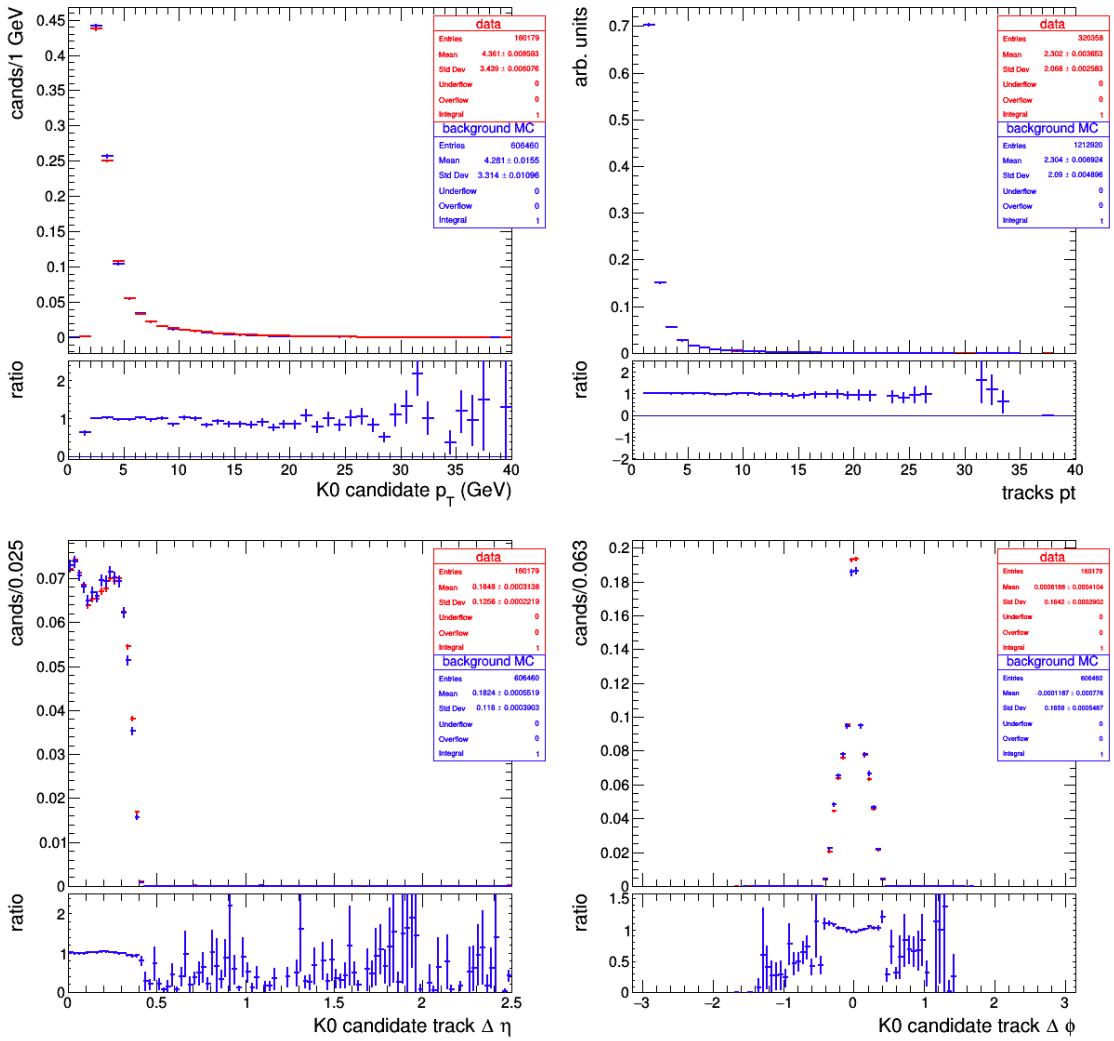

Figure 11.4: Comparisons of kinematic variables of vertices in 2017 between data (red) and simulation (blue) after corrections are applied.

tion efficiency difference, as shown in Fig. 11.5. The distributions are obtained by subtracting normalized background distributions from K_S^0 signal region distributions. As shown in the plots, the region of [0.2,0.4] cm is the "turn-on" region, with a low number of vertices; the region of [0.4,1.6] cm is the plateau of the distribution, which shows that the K_S^0 reconstruction efficiency is almost a constant as a function of distance in that region; the region of [1.6,2.0] cm is the "turn-off" region, the number of vertices start to decrease because of the beam pipe exclusion.

MC distributions are normalized in the way that MC and data have the same number of vertices in the [0.5,0.8] cm region. This region is selected to be relatively close to the beam spot and not in the "turn-on" region to reduce the inaccuracy resulting from large statistical uncertainty.

The data and MC behave slightly differently in the "turn-off" region. That results from different positions of the beam pipe center and the beam spot in data and MC. This discrepancy is not relevant to track reconstruction efficiency but has an effect on K_S^0 vertex reconstruction efficiency because of the beam pipe constraints. The effect of this discrepancy is accounted for in the vertex reconstruction efficiency study, as described in Sec. 11.1.2. As a result, regions that are close to the beam pipe are excluded to mitigate the effect.

Finally, the histograms are rebinned to reduce statistical uncertainties and are shown in Fig. 11.6. For each data-taking year, the bin that has the largest difference between data and MC is reported in Tab. 11.2.

Since the K_S^0 vertices are composed of exactly two tracks, the K_S^0 vertex reconstruction efficiency is proportional to the square of track reconstruction efficiency. The track reconstruction efficiency can be inferred as the square root of the K_S^0 reconstruction

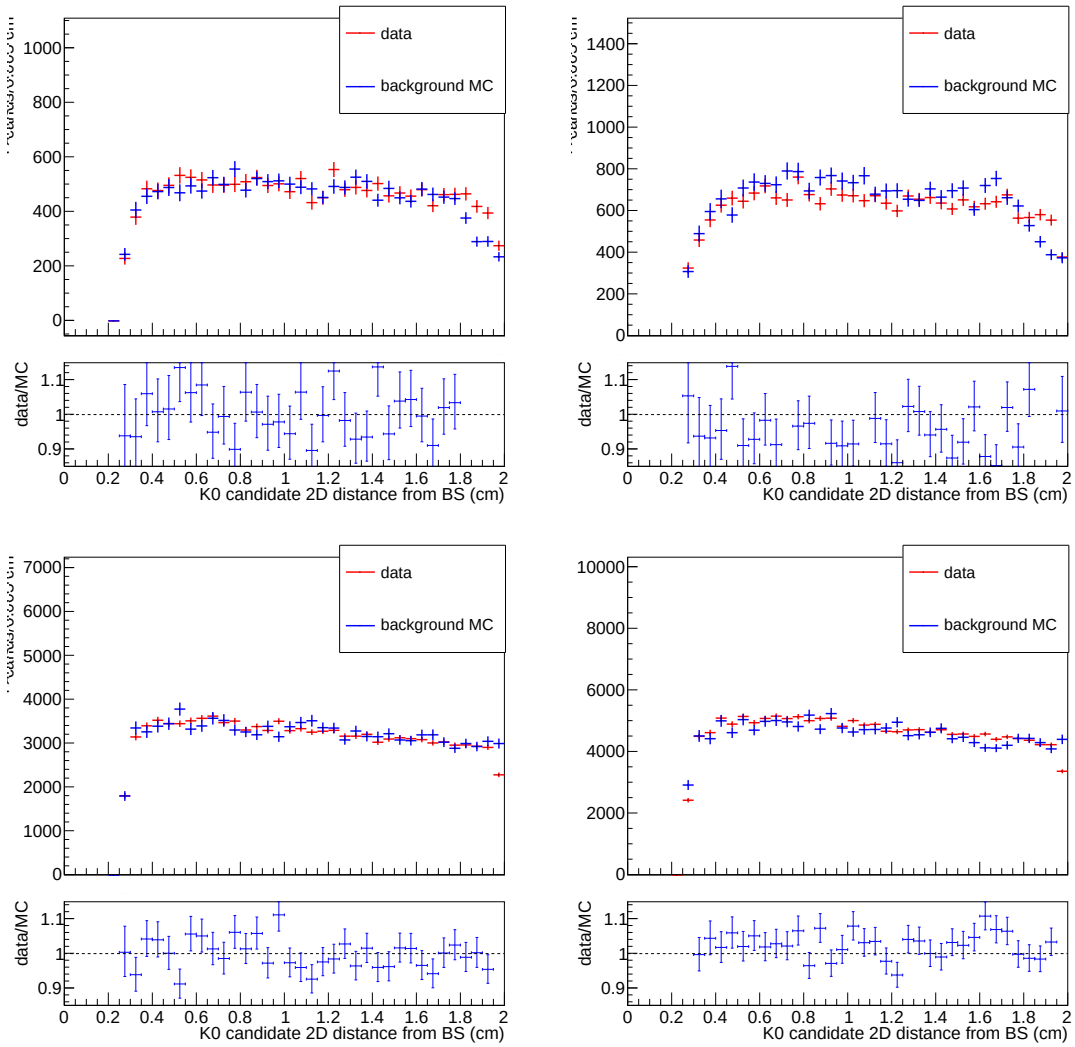

Figure 11.5: Transverse distance from K_0 vertices and the beam spot in data and MC.

year	data/MC
2016 preVFP	0.9746 ± 0.0268
2016 postVFP	1.0803 ± 0.0477
2017	0.9753 ± 0.0170
2018	0.9712 ± 0.0206

Table 11.2: Maximum of Data/MC ratio of K_S^0 vertex reconstruction efficiency.

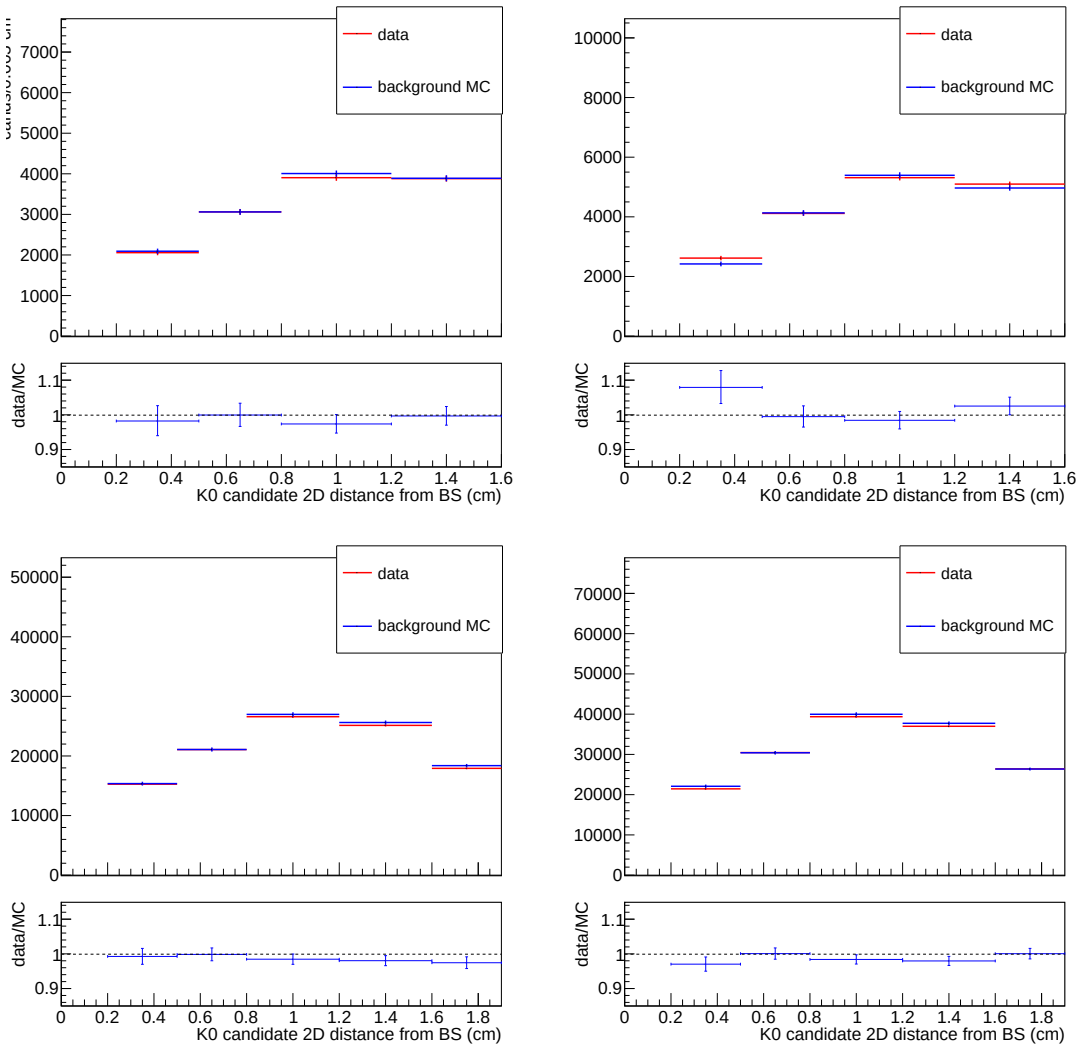

Figure 11.6: Transverse distance from K_0 vertices and the beam spot in data and MC.

efficiency. The track reconstruction efficiency has an impact on signal efficiency because the number of tracks in the displaced vertices is used as one of the discriminants to define the signal region. Specifically, the signal region requires vertices to have at least five tracks. In cases where a vertex has exactly five tracks, failure to reconstruct any track will cause the vertex to drop out of the signal region, and the impact of track reconstruction on such vertices is calculated as the track reconstruction efficiency to the power of five. On the other hand, for vertices with more than five tracks, the failure to reconstruct any of the tracks will not necessarily cause the vertex to be excluded from the signal region, as long as there are at least five tracks remaining, so the impact of track reconstruction efficiency is negligible on those vertices. Systematic uncertainty associated with the track reconstruction efficiency is assigned as the difference between the data/MC ratio of the efficiency of reconstructing five tracks and the unity. The systematic uncertainty for different years is reported in Tab. 11.3.

year	uncertainty
2016 preVFP	0.0632 ± 0.0645
2016 postVFP	0.2130 ± 0.1339
2017	0.0606 ± 0.0409
2018	0.0705 ± 0.0493

Table 11.3: Systematic uncertainty associated with track reconstruction efficiency.

11.1.2 Vertex Reconstruction

With the track reconstruction difference between data and simulation accounted for in Sec. 11.1.1, the vertex reconstruction is studied assuming that track reconstruction is consistent in data and simulation. The study is performed by applying the vertex reconstruction algorithm described in Ch. 7 on artificial "moved" vertices, which are signal-like vertices made by artificially displacing real background tracks away from

their original positions.

The procedure of this study is developed based on [26] and organized as follows: events are selected to satisfy the preselection cuts outlined in Sec. 6.3 and to have a well-measured primary vertex. Within these events, jets with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and at least three matched particle-flow candidate tracks are randomly selected to be moved away from their original positions. In order to match the qualities of these selected jets to those arising from the signal, which can originate from light quarks and b quarks, the number of light jets n_l and b-tagged jets n_b are specified in the jet selection based on the signal model. For the split SUSY benchmark signal model, since the long-lived gluino decays to two light quarks and a neutralino, n_l is set to be 2 and n_b is set to be 0.

Next, the flight direction of the artificial signal-like vertex is determined by taking the vector sum of the chosen jets' momenta. However, for the split SUSY benchmark signal model, the flight direction is supposed to be defined by the vector sum of the chosen jets that represent the quarks and an "invisible particle" that represents the neutralino. To account for the effect of the invisible particle, the flight direction is smeared by 2.0 radians in θ and ϕ. The relatively large smearing is designed to ensure that we have enough artificial "moved" vertices to mimic cases in signal when jet momentum vector sum has large angular differences with the LLP flight direction. The angular difference is corrected accordingly for different signal samples, as shown in the bottom two plots in Fig. 11.11 and 11.13. The flight distance is then generated by sampling from an exponential distribution with the $c\tau$ of $300\,\mu$m and 10 mm. Events with a $c\tau$ of $300\,\mu$m are rescaled to signal samples with a $c\tau$ of $100\,\mu$m and $300\,\mu$m. Events with a $c\tau$ of 10 mm are rescaled to signal samples with the $c\tau$ of $\{1\,\mathrm{mm}, 10\,\mathrm{mm}, 30\,\mathrm{mm}, 100\,\mathrm{mm}, 300\,\mathrm{mm}, 1000\,\mathrm{mm}\}$ individually. The rescaling is

based on the LLP travel distance distributions. The xy-projection of flight distance is required to be less than 2 cm, which is approximately the fiducial cut used in the search region. The 3D displacement vector of the artificial vertex is formed by adding the flight vector to the location of the primary vertex.

All tracks associated with the selected jets are moved by adding the displacement vector to their impact parameters. The vertex reconstruction algorithm is applied on moved tracks together with non-moved tracks. To be considered as successfully reconstructed, vertices are required to have at least five tracks, pass the vertex quality cuts described in Ch. 7, and be within 200 μm of the generated positions of the artificial vertices. The requirement of the distance between reconstructed and generated vertex positions to be less than 200 μm reduces the possibility of taking the background vertices as successfully reconstructed vertices while keeping at least 95% of vertices for signal samples with different $c\tau$'s and masses. Figure 11.7 shows the 3D distance between reconstructed vertices and their closest generator level LLP decay positions. The vertex reconstruction efficiency is calculated as the number of successfully reconstructed vertices divided by the total number of artificial "moved" vertices.

To check whether the artificial "moved" vertices are similar enough to signal vertices, background events with artificial "moved" vertices are compared with signal events, as shown in Fig.11.8. Comparisons show that the angular distributions of moved jets and the number of tracks in each jet are not very similar to the signal. The inconsistency is expected given the very different topologies of generic Standard Model background and exotic LLPs. Based on the comparisons, a set of signal-mocking selections on the artificially moved jets are applied to mitigate the differences. It is found that the highest p_{T} jet in the event tends to result in $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ that points in the opposite

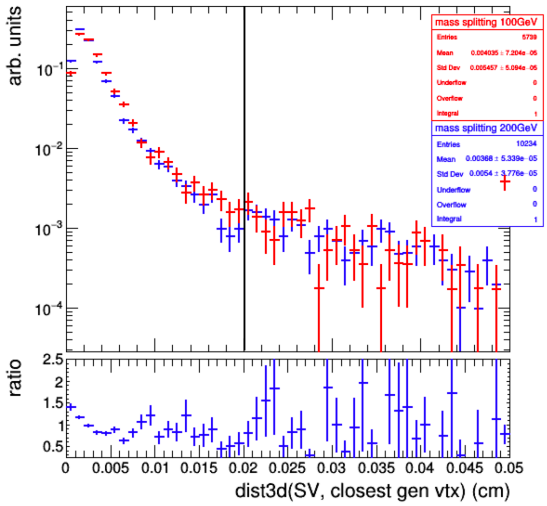

Figure 11.7: The 3D distance between reconstructed vertices with at least five tracks and its closest LLP decay position for split SUSY signal samples with 1 mm $c\tau$, 2000 GeV gluino mass, and mass splitting as 100 GeV (red) and 200 GeV (blue) in 2017. The black line shows the selection at 200 μm. Histograms are normalized to unity.

direction, as the jet and $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ recoil off of one another. As this is quite different from the signal topology where stable neutralinos and low p_{T} jets are the decay products of a heavy gluino, we require that the highest p_{T} jet not be considered as part of this artificial displacement procedure, and only consider "moving" the other jets in the event. Besides, it is also required that the angular distance between the two moved jets to satisfy $\Delta R > 0.4$, and the $\Delta\phi$ ($\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$, moved jet) > 1.0 for each of the two moved jets. This selection is helpful in mimicking the kinematic relations between the two jets from LLP decays. In addition, to make the kinematic relations of the two "moved" jets similar to the signal, we order all jets in an event by decreasing p_{T}, and assign to them an index representing their p_{T} ordering. A 2D jet index distribution for split SUSY samples with 2000 GeV gluino mass, 200 GeV mass splitting, and 10 mm $c\tau$ is shown in Fig. 11.9. The plot shows that less than 1% of LLP decays in the signal sample have jet indices that differ by more than five. According to the result,

we require that the index difference between the harder artificially-moved jet and the softer moved jet to be within five.

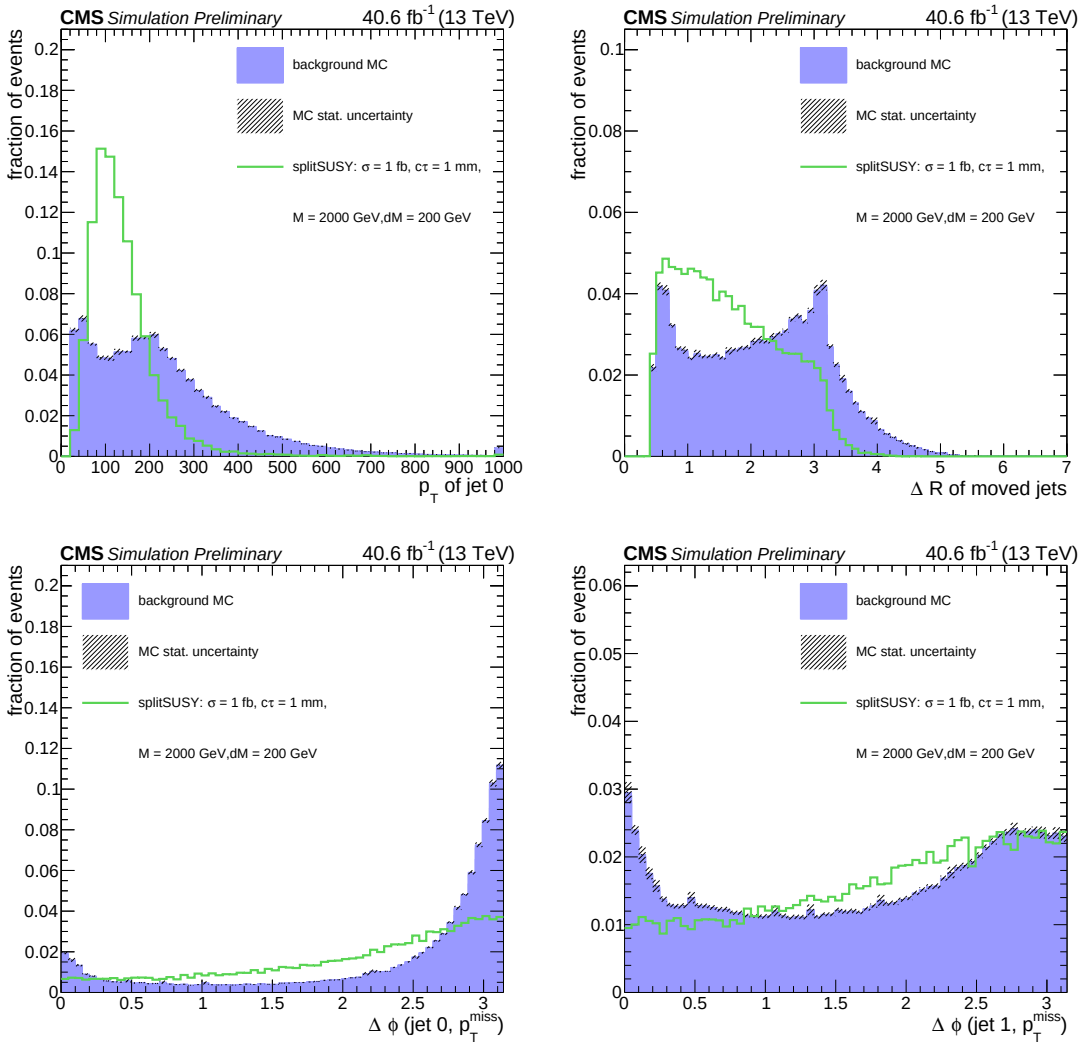

Figure 11.8: Comparisons of variables between background MC and signal samples with 200 GeV mass splitting and 10 mm $c\tau$ before the "signal-mocking cuts" are applied. Signal and background histograms are normalized to the unity.

With the selections applied, there are still some differences between the artificially moved jets and the signal. Figure 11.10 and 11.11 show some of the differences after the selections described above are applied in the 2017 simulation. The $\Delta\eta$ and $\Delta\phi$ between jet momentum vector sum and the chosen flight direction of "moved" vertex

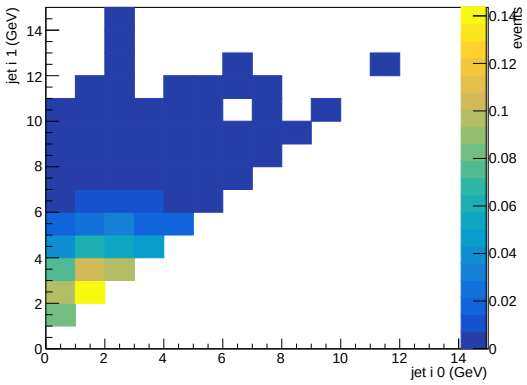

Figure 11.9: Jet index 2D distribution for signal sample with 200 GeV mass splitting and 10 mm $c\tau$ before the "signal-mocking cuts" are applied. The histogram is normalized to the unity.

distributions for background MC are relatively flat due to the large radian smearing we set when we determine the flight direction. As a result, further corrections are needed to mitigate the differences and reliably use this procedure to approximate the signal vertices.

The variables that are corrected are: the number of tracks in both moved jets, η of both moved jets, $\Delta\eta$ and $\Delta\phi$ between the 4-vector momentum vector sum of both moved jets and the chosen flight direction of artificial "moved" vertices. The jet p_{T} is not corrected because correcting jet p_{T} makes the jet number of tracks not match and vice versa.

Weights for corrections are derived from the ratio between signal MC and artificial "moved" vertices in background MC distributions. For the number of tracks and η of moved jets, weights are derived from the ratio of 2D histograms such as η of moved jet 0 vs. η of moved jet 1. Doing that accounts for not only the single jet differences but also the correlations between both moved jets. For the correction for $\Delta\eta$ and $\Delta\phi$ between the vector sum of p_{T} of both moved jets and the artificial flight distance, 2D

histograms are also used to account for the correlation between $\Delta\eta$ and $\Delta\phi$. Finally, the lab-frame decay lengths are not the same as the $c\tau$ due to the different momentum carried by LLPs. The xy-projection of flight distance is corrected. One example of the weight we use to correct the artificial "moved" vertices is shown in Appendix D. The corrections are applied on both artificially moved background MC and data events.

Different weights are derived for different mass splittings and $c\tau$'s. For split SUSY samples with the same mass splitting and $c\tau$, the event topology and kinematics are found to be similar for different gluino masses. However, the signal efficiency is still correlated with gluino masses, especially for signal samples with 100 μm $c\tau$. To understand whether the data/MC efficiency ratio change with different gluino masses, a check is performed to measure the data/MC vertex reconstruction efficiency ratio for different gluino masses, namely 1400 GeV, 2000 GeV, and 2600 GeV with the mass splitting of 100 GeV and 200 GeV. The $c\tau$'s considered in the check are 100 μm, 10 mm, and 1000 mm to cover a wide range of lifetimes. The data and MC used in this check are from 2017. The results are shown in Tab. 11.4. The data/MC efficiency ratio measured from different gluino masses with the same mass splitting and $c\tau$ does not differ by the order of percent. The difference is covered by the systematic uncertainty associated with the study, which is described in the last paragraph of this section. As a result, to save time and computation resources, split SUSY samples with a gluino mass of 2000 GeV are used to derive the weights.

The comparisons between artificially moved background MC and signal are shown in Fig. 11.12 and 11.13 for 2017.

With all the selections and corrections in place, the vertex reconstruction efficiencies for artificially moved background MC and data are calculated. The comparisons of vertex reconstruction efficiency between background MC and data are shown in

$m_{\tilde{g}}$ (GeV)	Δm (GeV)	100 μm	10 mm	1000 mm
1400	100	0.7577 $\pm$ 0.0289	0.8286 $\pm$ 0.0254	0.7346 $\pm$ 0.0977
2000	100	0.7894 $\pm$ 0.0500	0.8177 $\pm$ 0.0163	0.7935 $\pm$ 0.0265
2600	100	0.7959 $\pm$ 0.0482	0.8079 $\pm$ 0.0176	0.7853 $\pm$ 0.0241
1400	200	0.7655 $\pm$ 0.0241	0.8450 $\pm$ 0.0150	0.8174 $\pm$ 0.0192
2000	200	0.7899 $\pm$ 0.0323	0.8362 $\pm$ 0.0121	0.8249 $\pm$ 0.0155
2600	200	0.7699 $\pm$ 0.0292	0.8379 $\pm$ 0.0136	0.8132 $\pm$ 0.0156

Table 11.4: The data/MC ratio of vertex reconstruction efficiency for gluino mass of 1400 GeV, 2000 GeV, and 2600 GeV with different mass splittings and $c\tau$'s.

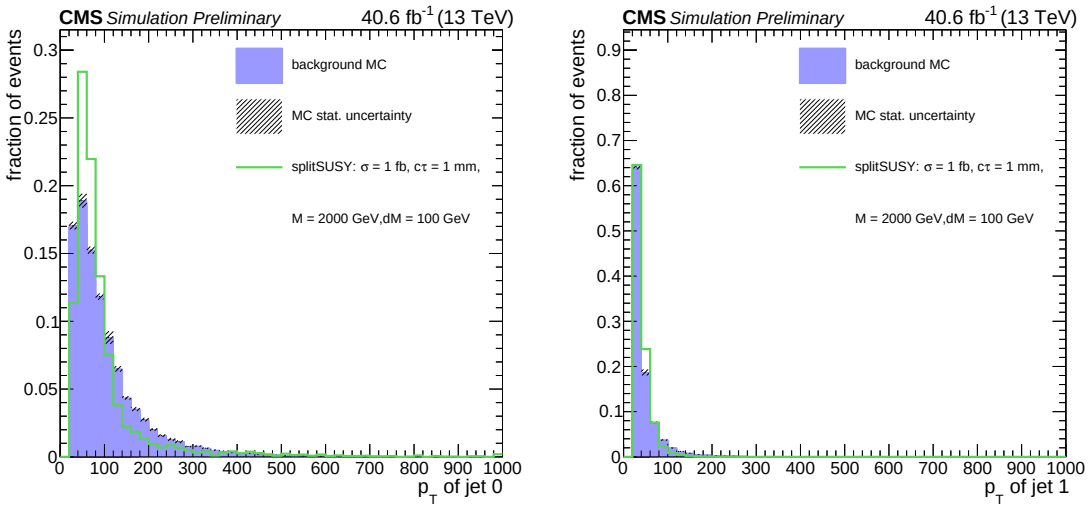

Figure 11.10: Comparisons of jet p_{T} between background MC and signal samples with 100 GeV mass splitting and 10 mm $c\tau$ after the "signal-mocking cuts" are applied but before corrections are applied. Signal and background histograms are normalized to the unity.

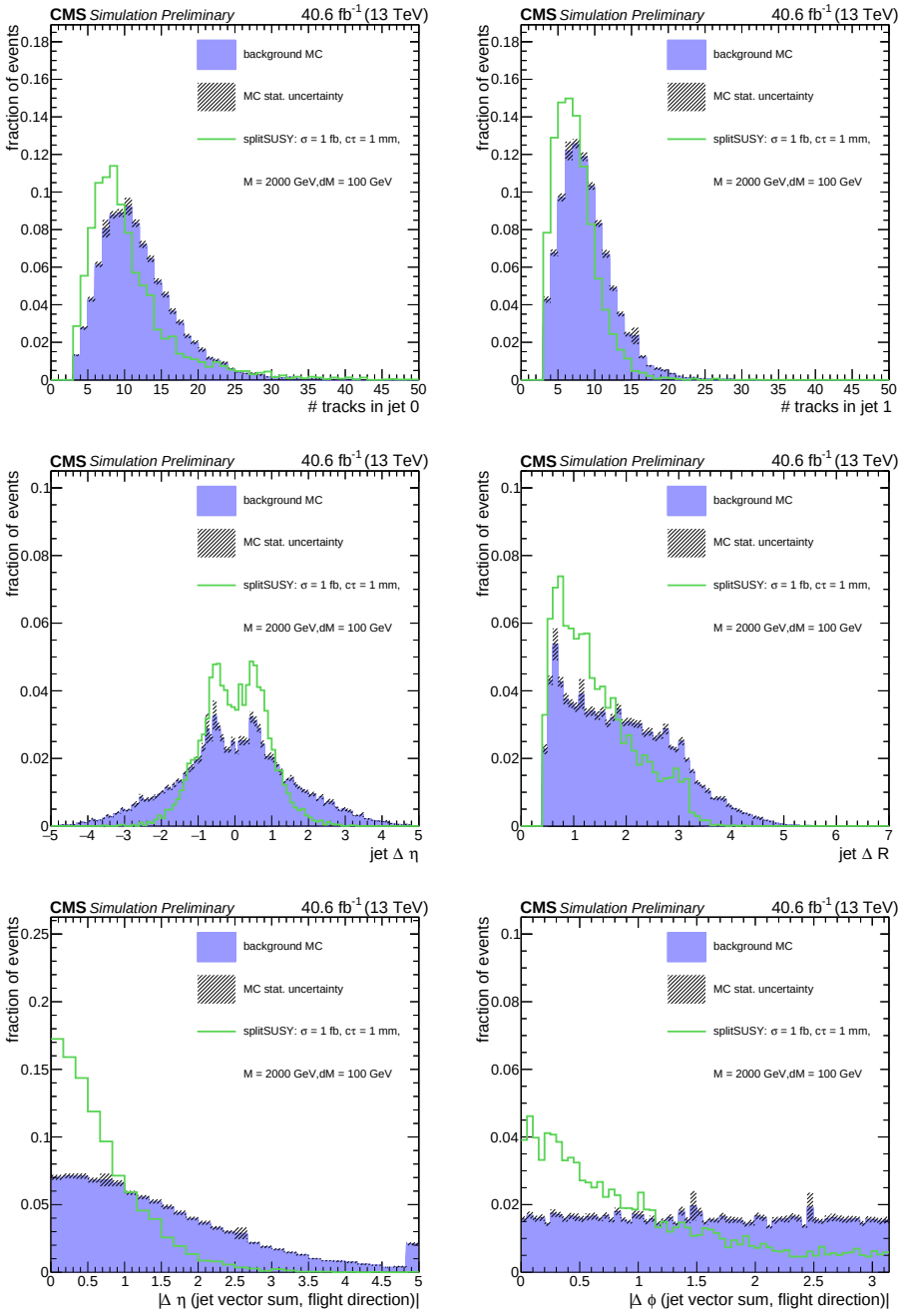

Figure 11.11: Comparisons between background MC and signal samples with 100 GeV mass splitting and 10 mm $c\tau$ after the "signal-mocking cuts" are applied but before corrections are applied. Signal and background histograms are normalized to the unity.

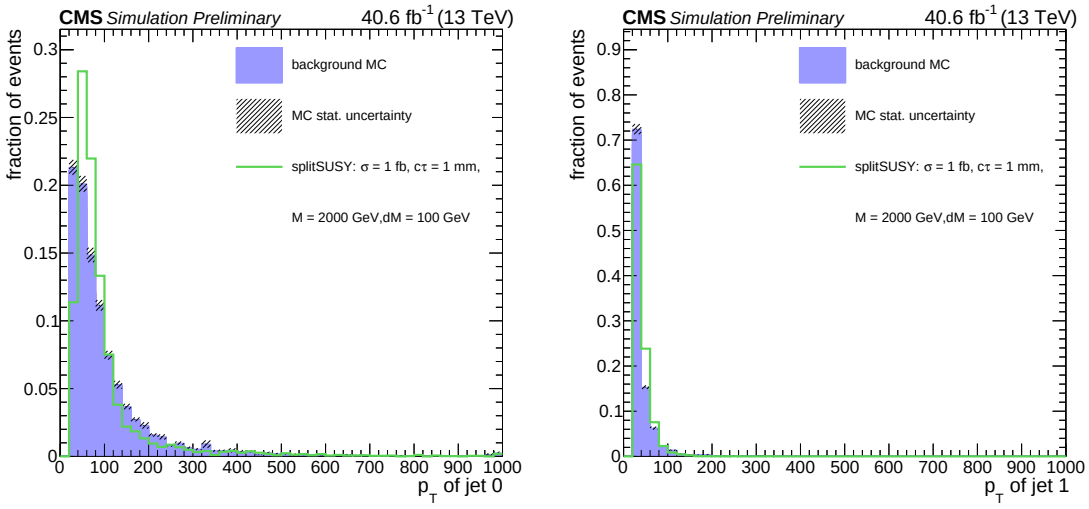

Figure 11.12: Comparisons of jet p_{T} between background MC and signal samples with 100 GeV mass splitting and 10 mm $c\tau$ after the "signal-mocking cuts" are applied and after corrections are applied. Signal and background histograms are normalized to the unity.

Fig. 11.14, 11.15, 11.16, and 11.17 for 2016 preVFP, 2016 postVFP, 2017, and 2018. The difference in vertex reconstruction efficiency between data and MC mainly comes from different vertex reconstruction efficiency as a function of the number of displaced tracks selected for vertex reconstruction, which is shown in Fig. 11.18.

The vertex reconstruction efficiencies measured for data and MC show that the ratio between data and MC does not depend on the $c\tau$. This motivates using one universal ratio to represent all lifetimes for a given mass splitting. As a result, for a given year and mass splitting, a universal efficiency correction scale factor is derived using the 10 mm sample data and MC efficiency ratio. The results for different years and mass splittings are shown in Tab. 11.5. The scale factors are applied to the signal event yields to reflect the expected number of signal events to be observed in the data events. The systematic uncertainty associated with this scaling procedure is calculated as the difference between the MC and data efficiency ratio and the unity.

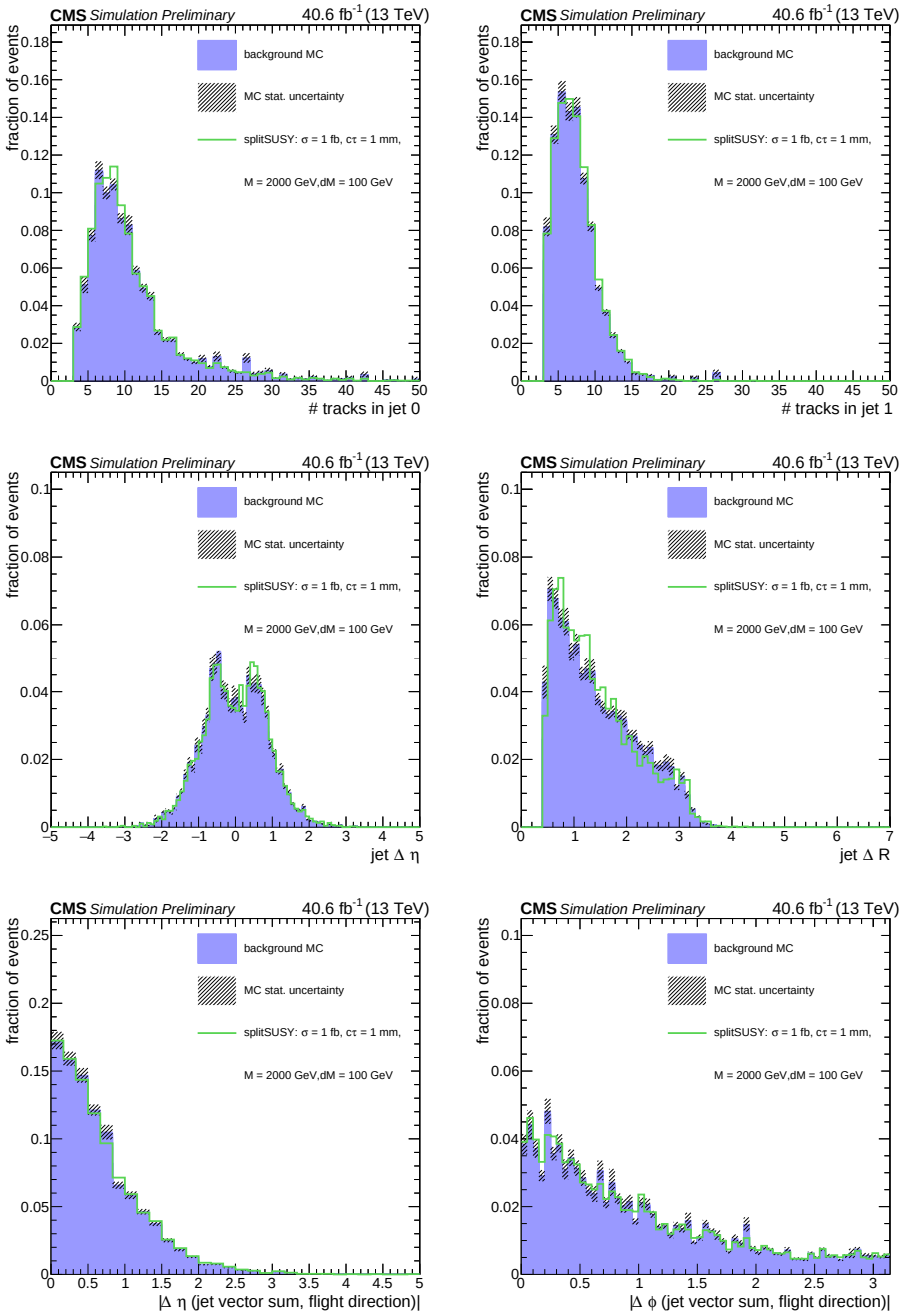

Figure 11.13: Comparisons between background MC and signal samples with 100 GeV mass splitting and 10 mm $c\tau$ after the "signal-mocking cuts" and corrections are both applied. Signal and background histograms are normalized to the unity.

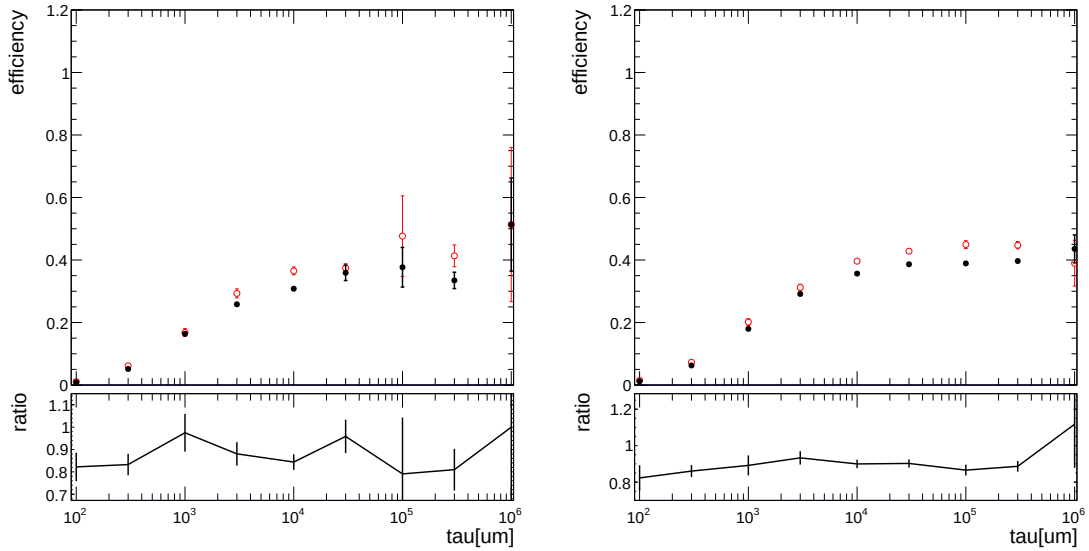

Figure 11.14: Vertex reconstruction efficiency for data (black) and background MC (red) in 2016 preVFP as a function of $c\tau$. Mass splittings of 100 GeV (left) and 200 GeV (right) are shown separately.

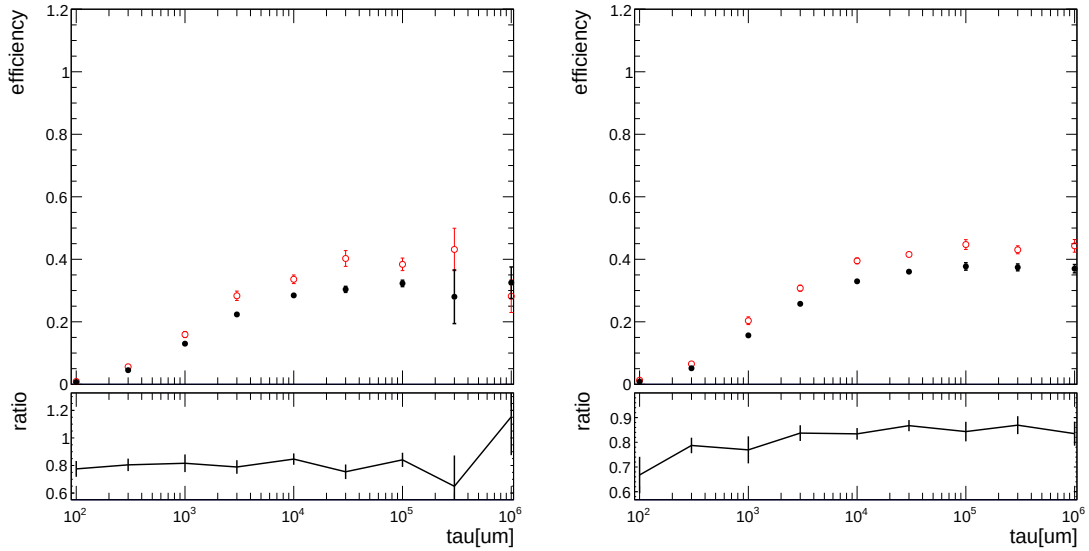

Figure 11.15: Vertex reconstruction efficiency for data (black) and background MC (red) in 2016 postVFP as a function of $c\tau$. Mass splittings of 100 GeV (left) and 200 GeV (right) are shown separately.

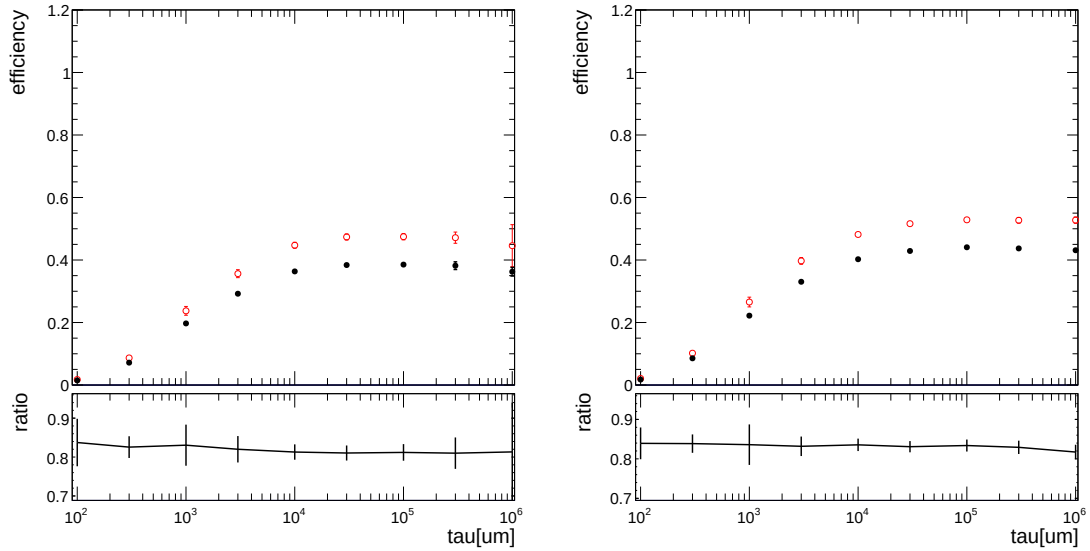

Figure 11.16: Vertex reconstruction efficiency for data (black) and background MC (red) in 2017 as a function of $c\tau$. Mass splittings of 100 GeV (left) and 200 GeV (right) are shown separately.

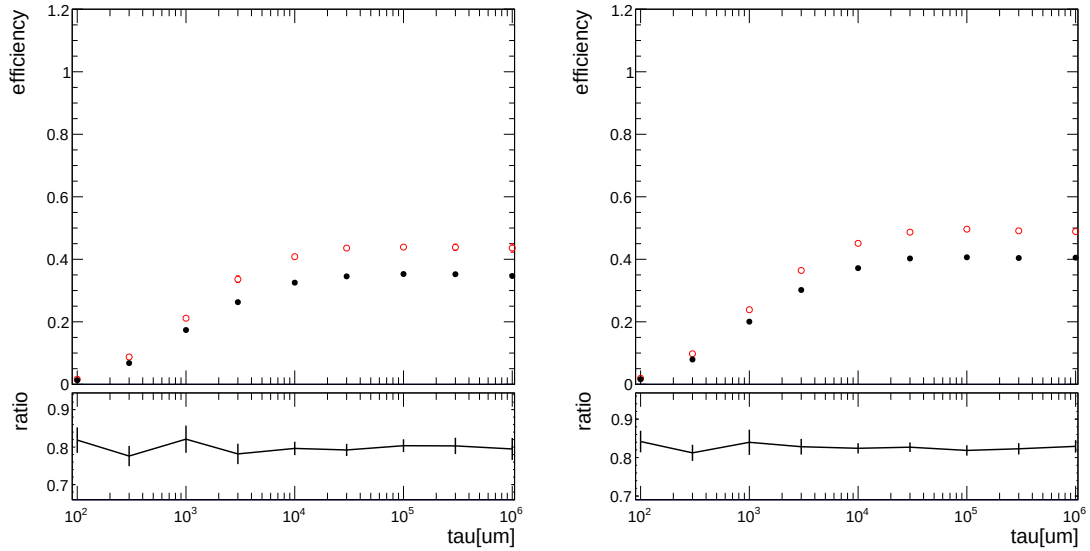

Figure 11.17: Vertex reconstruction efficiency for data (black) and background MC (red) in 2018 as a function of $c\tau$. Mass splittings of 100 GeV (left) and 200 GeV (right) are shown separately.

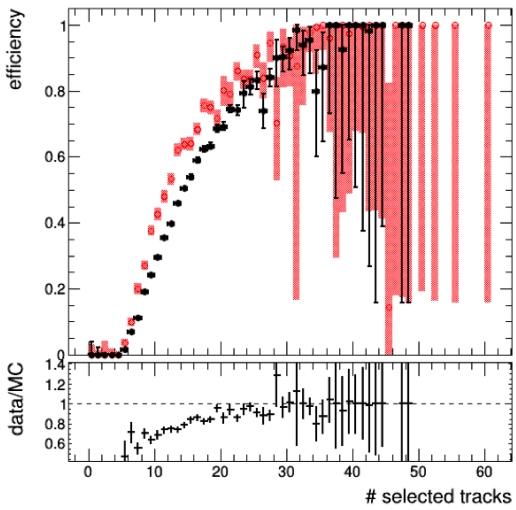

Figure 11.18: Vertex reconstruction efficiency for data (black) and background MC (red) in 2017 as a function of the number of selected displaced tracks. Corrections for the signal sample with gluino mass of 2000 GeV, mass splitting of 200 GeV, and $c\tau$ of 10 mm are applied to both data and background MC.

year	$\Delta m = 100$ GeV	$\Delta m = 200$ GeV
2016 preVFP	0.8532 ± 0.0288	0.9062 ± 0.0173
2016 postVFP	0.8042 ± 0.0309	0.8338 ± 0.0203
2017	0.8155 ± 0.0182	0.8385 ± 0.0129
2018	0.8210 ± 0.0137	0.8321 ± 0.0103

Table 11.5: Scale factor derived from ratios of vertex reconstruction efficiency between data and MC.

11.1.3 ML Tagging Efficiency

This section accounts for the differences between data and MC in ML tagging efficiency. The artificial displaced vertex procedure described in Sec. 11.1.2 is also used to evaluate the ML tagging efficiency. The IN is applied to background events with artificial "moved" vertices. The resulting MLScore is checked and the ML tagging efficiency is calculated for data and MC. Unlike what is done in the vertex reconstruction efficiency study, which uses events with 300 μm and 10 mm $c\tau$'s, the ML tagging efficiency study uses only events with 10 mm $c\tau$. Results of other lifetimes are derived by reweighting the 10 mm events to those lifetimes.

The artificial "moved" vertices in the background events are obtained following the same procedure, selections, and corrections described in Sec. 11.1.2. Since the IN takes track information as input, comparisons of track variables used in the ML between background MC events with artificial "moved" vertices and signal events, shown in Fig. 11.19 and 11.20, are performed to ensure that the artificial "moved" vertices are good proxies of signal vertices in the track level.

The artificial "moved" events first undergo the vertex reconstruction process and events that have successfully reconstructed displaced vertices are passed to the IN.

After applying the IN to the artificially moved events, the MLScore distribution comparison between artificially moved data and background MC events is shown in Fig. 11.21. As shown in the plot, most of the artificially moved events have MLScores around 1, which are tagged as signal events. The fact that the IN recognizes the artificially moved events with only one displaced vertex as signal events further shows that the IN is model-independent and robust.

The ML tagging efficiency is calculated as the ratio of the number of events with

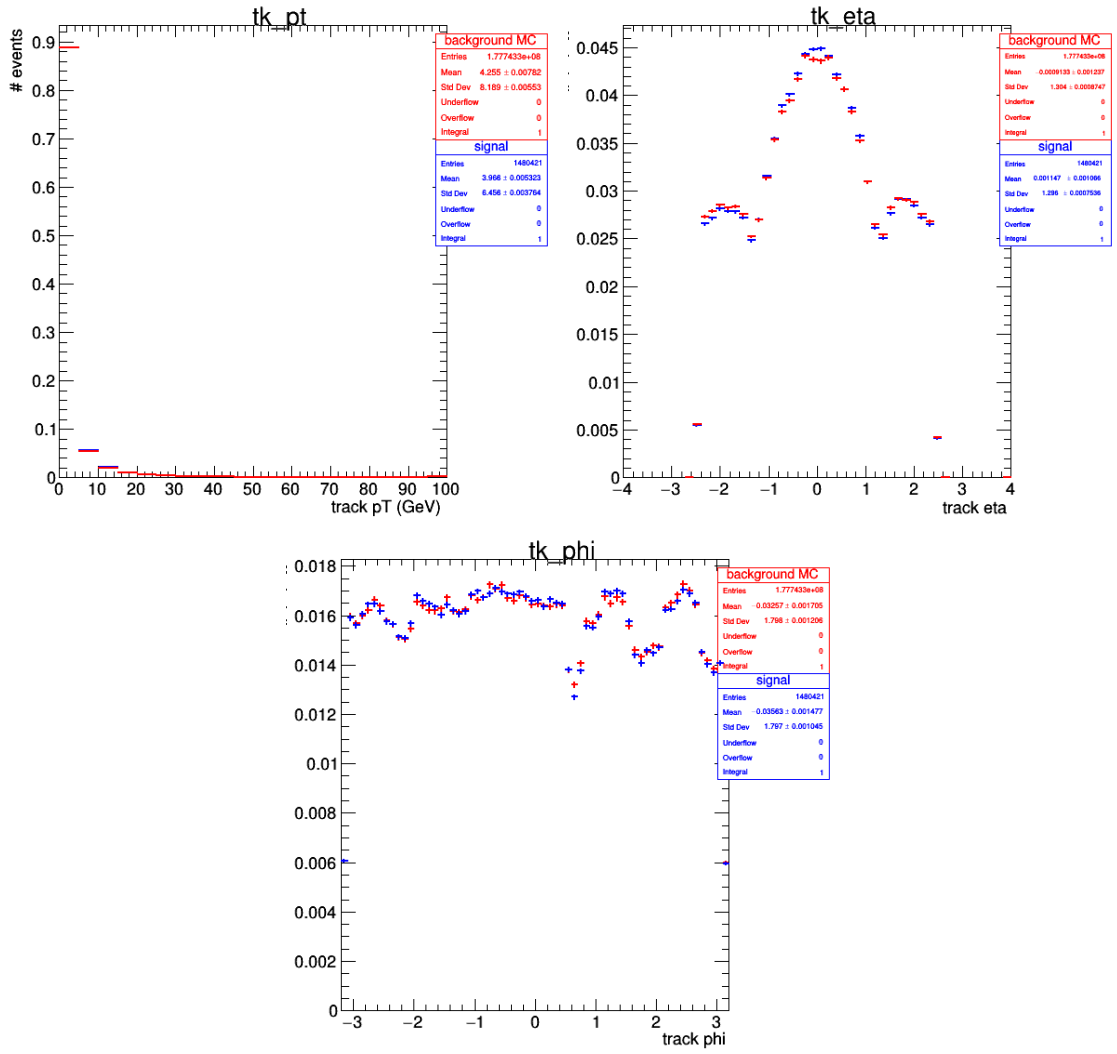

Figure 11.19: Comparisons for track kinematic variables used in the IN between background MC and signal samples with 200 GeV mass splitting and 10 mm $c\tau$ after the "signal-mocking cuts" and corrections are both applied. Signal and background histograms are normalized to the unity.

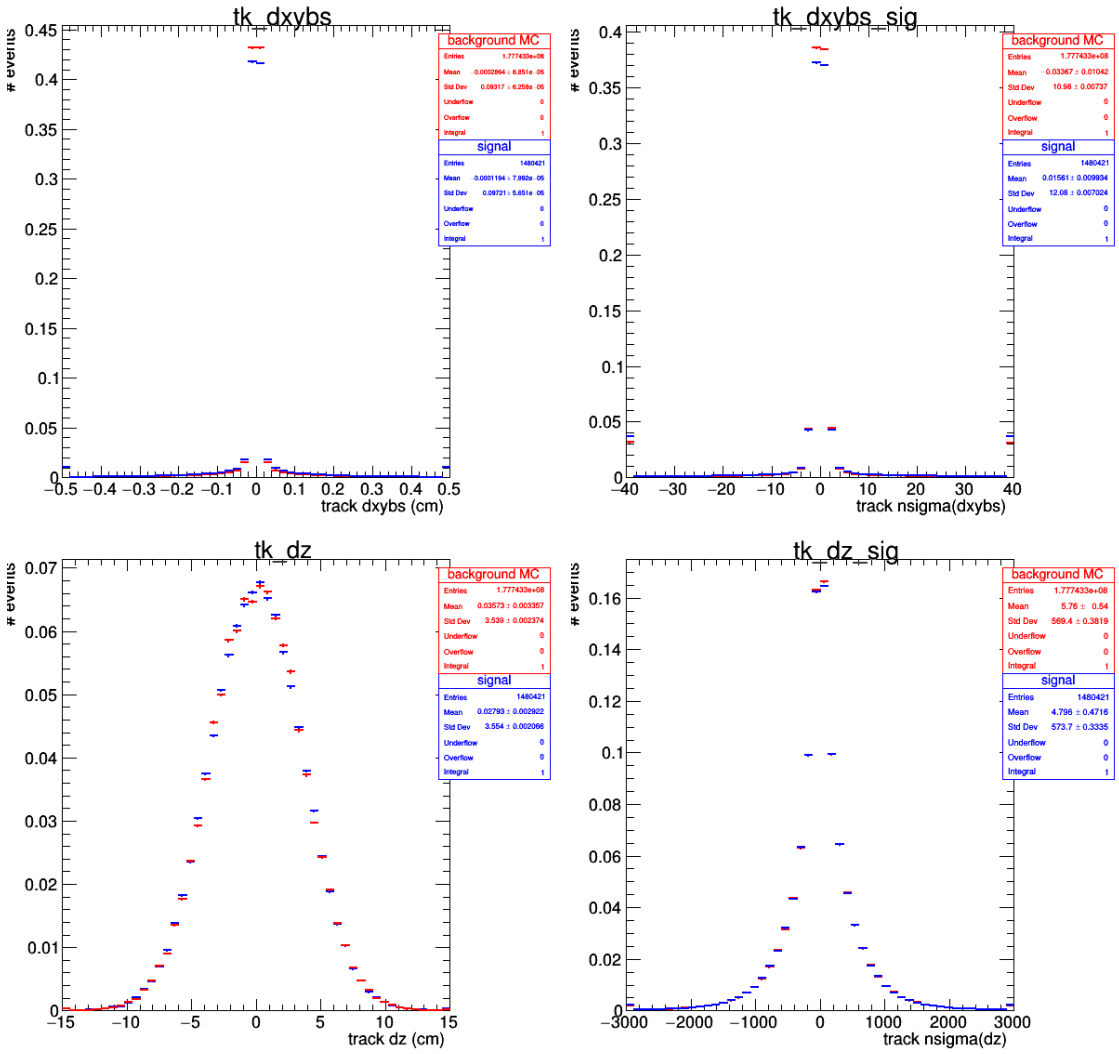

Figure 11.20: Comparisons for track displacement variables used in the IN between background MC and signal samples with 200 GeV mass splitting and 10 mm $c\tau$ after the "signal-mocking cuts" and corrections are both applied. Signal and background histograms are normalized to the unity.

MLScore > 2 and all events that are fed into the IN. The efficiencies are shown in Fig. 11.22, 11.23, 11.24, and 11.25 for 2016 preVFP, 2016 postVFP, 2017, and 2018, as a function of the $c\tau$. The ratios of ML tagging efficiency between data and MC does not depend on lifetimes, which implies that we can use a universal data/MC ratio to represent all lifetimes.

The data/MC ML tagging efficiency ratios are shown in Tab. 11.6. Since the efficiency ratios are very close to the unity, no scale factor is used for ML tagging efficiency. Systematic uncertainty associated with the ML tagging efficiency is calculated as the difference between the MC and data efficiency ratio and the unity.

year	$\Delta m = 100\,\text{GeV}$	$\Delta m = 200\,\text{GeV}$
2016 preVFP	1.0049 ± 0.0397	0.9943 ± 0.0224
2016 postVFP	1.0005 ± 0.0460	0.9990 ± 0.0256
2017	0.9743 ± 0.0252	0.9726 ± 0.0170
2018	0.9849 ± 0.0177	0.9838 ± 0.0138

Table 11.6: Data/MC ratios of ML tagging efficiency.

11.1.4 Fake MET Veto Efficiency

To get rid of events with fake MET, the fake MET veto, calculated as $|MET_{PF} - MET_{CALO}|/MET_{CALO} < 0.6$, is applied in the event selection. The fake MET veto results in a signal event yield loss of 5–8%. To account for the potentially inaccurate modeling of the fake MET veto in the simulation, a systematic uncertainty is calculated as the difference between the fake MET veto efficiency from the unity. The values of this uncertainty are reported in Tab. 11.7.

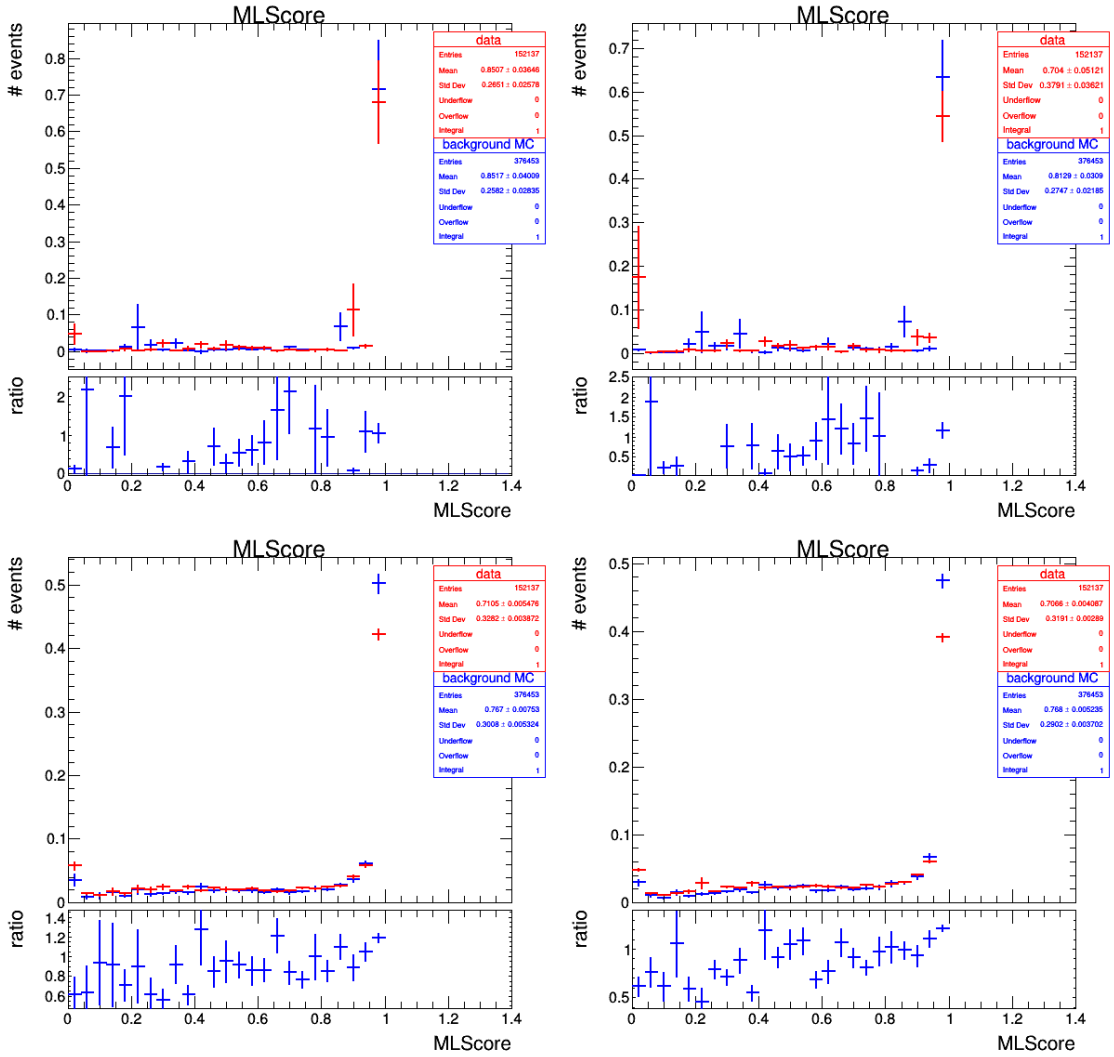

Figure 11.21: ML score distribution in 2017 for artificially moved data and background MC that mimics signal sample with 100 (left) and 200 (right) GeV mass splitting and 0.1 (up) and 10 (down) mm $c\tau$.

year	$\Delta m = 100\,\text{GeV}$	$\Delta m = 200\,\text{GeV}$
2016 preVFP	0.069 ± 0.033	0.048 ± 0.019
2016 postVFP	0.072 ± 0.033	0.052 ± 0.018
2017	0.075 ± 0.019	0.055 ± 0.011
2018	0.066 ± 0.020	0.051 ± 0.012

Table 11.7: Systematic uncertainty of the modeling of fake MET veto.

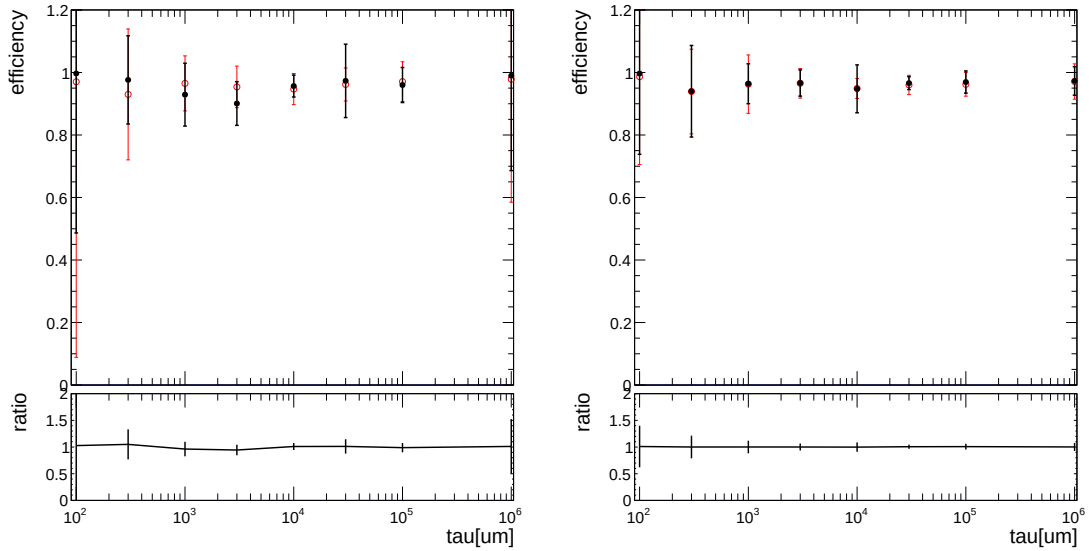

Figure 11.22: ML tagging efficiency for data (black) and background MC (red) in 2016 preVFP as a function of $c\tau$. Mass splittings as 100 GeV (left) and 200 GeV (right) are shown separately.

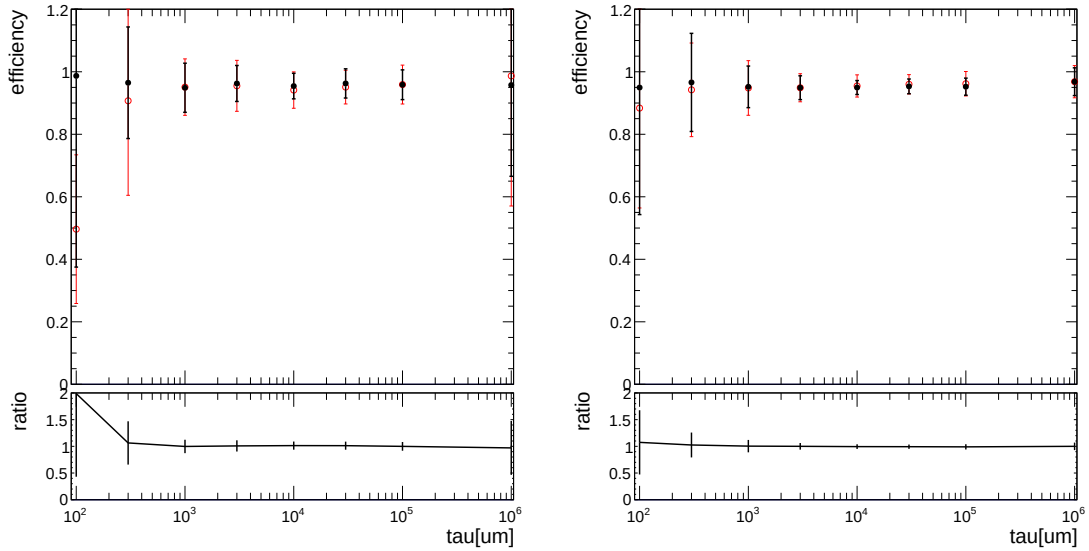

Figure 11.23: ML tagging efficiency for data (black) and background MC (red) in 2016 postVFP as a function of $c\tau$. Mass splittings as 100 GeV (left) and 200 GeV (right) are shown separately.

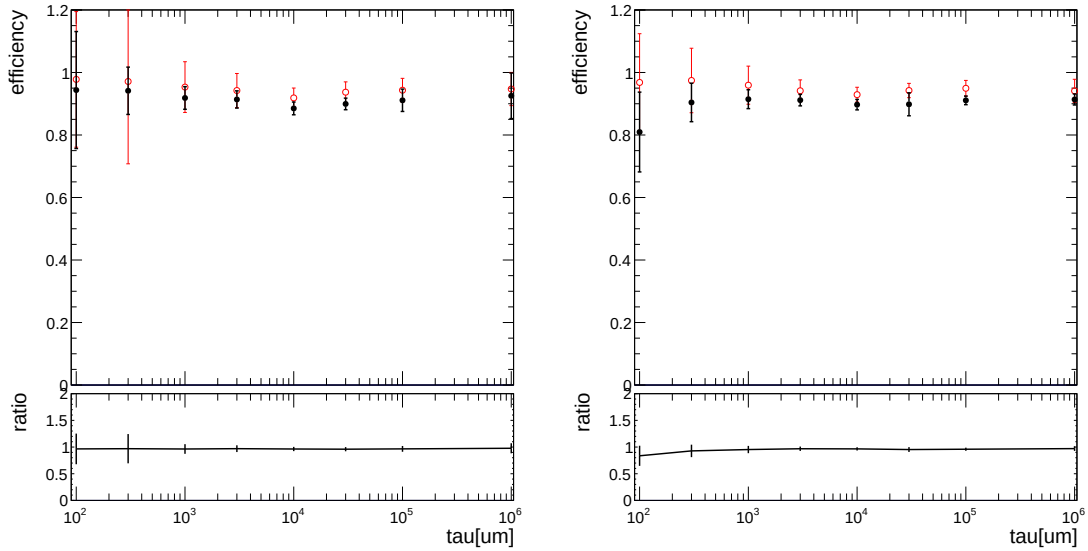

Figure 11.24: ML tagging efficiency for data (black) and background MC (red) in 2017 as a function of $c\tau$. Mass splittings as 100 GeV (left) and 200 GeV (right) are shown separately.

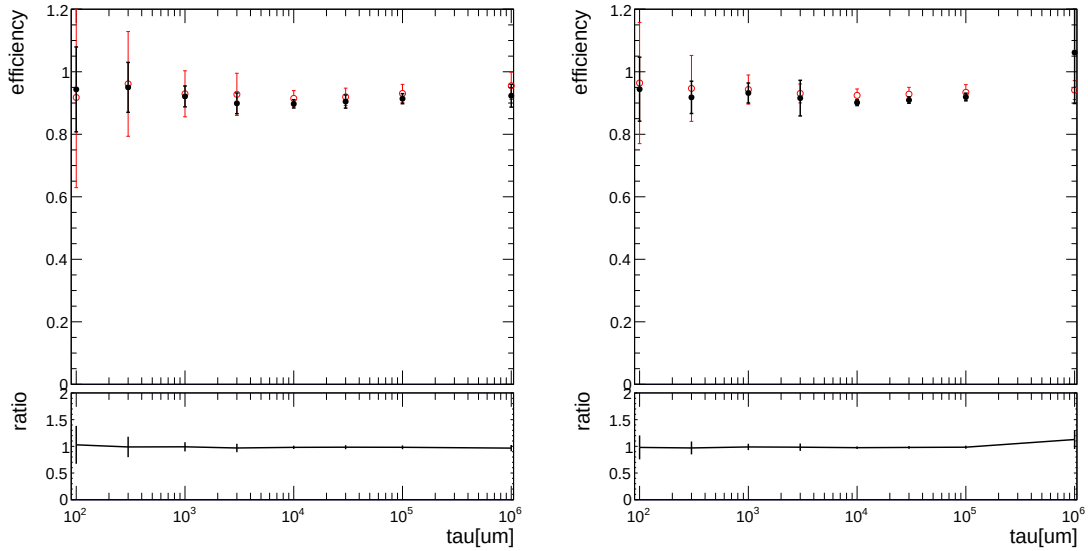

Figure 11.25: ML tagging efficiency for data (black) and background MC (red) in 2018 as a function of $c\tau$. Mass splittings as 100 GeV (left) and 200 GeV (right) are shown separately.

11.1.5 Trigger Efficiency

The trigger efficiency is described in Sec. 6.2. Differences in the measured efficiency of the trigger between data and simulation are used to scale the signal efficiency. The relative difference of event yield in the signal region due to the statistical uncertainty of the scale factor is used as systematic uncertainty of trigger efficiency. The value of the uncertainty is 0.7–1.3%. In addition, given the slightly different trigger efficiency measurements from muon-triggered and electron-triggered events, the difference between trigger efficiencies measured from muon-triggered and electron-triggered events is also taken as the systematic uncertainty. The value of this uncertainty is 1.0–2.1%.

11.1.6 Pileup

To study the influence of the pileup distribution on this search, different event weights derived from varied cross section of proton-proton collision are applied to the simulated events. The difference between signal yield before and after shifting the pileup weight is taken as the systematic uncertainty associated with pileup. The value of the systematic uncertainty associated with pileup is 3.2–5.5%.

11.1.7 Integrated luminosity

The uncertainty in the integrated luminosity is 1.2% in 2016 preVFP, 1.2% in 2016 postVFP, 2.3% in 2017, and 2.5% in 2018.

11.1.8 Issues in Run 2 data

During the data-taking period from 2016 to 2018, there were some issues with the operation of the CMS detector, which had effects on the data. Some of those issues are not simulated in the simulation, which introduces additional differences between data and simulation, thus becoming additional sources of systematic uncertainty in this search. Special treatments are applied to mitigate those effects, as described below.

EE L1 prefiring (2016 + 2017) During the data-taking in 2016 and 2017, the ECAL experienced a gradual timing shift, which made the L1T EE trigger rate affected in the high η range. A check is performed by rerunning the search on 2016 and 2017 signal MC by removing events with jets with $p_{\mathrm{T}} > 100\,\mathrm{GeV}$ and $2.25 < |\eta| < 3.0$. The signal MC used is split SUSY samples with $c\tau$ of 10 mm, gluino mass of 2000 GeV, and mass splittings of 100 and 200 GeV. The number of expected events in the signal region change in the range of 1.69–3.63%. To mitigate the effect, correction weights provided centrally by the CMS experiment are applied in the MC. Weights are shifted up and down to determine the uncertainty, which is 0.5–1.0%.

HEM 15/16 failure (2018) Following the power interruptions generated by false fire alarms on Saturday, June 30th, negative endcap HCAL sectors HEM15 and HEM16 could no longer be operated until the end of the 2018 run. The first regular physics run affected is 319077 and a 40-degree section (η from -3.0 to -1.3 and ϕ from -1.57 to -0.87) in HCAL is effectively off. We tried mitigating the effect by vetoing events that have jets in the problematic region after the first affected run in 2018. The ϕ distribution of $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ before and after the mitigation is shown in Fig. 11.26. It

shows that the mitigation removed the excess of events in the problematic region.

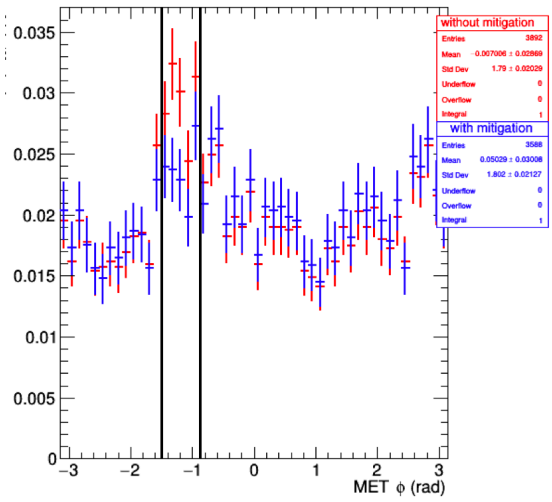

Figure 11.26: Comparison of $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ ϕ before and after the HEM issue mitigation. The problematic region is labeled between the two black solid lines.

A summary of systematic uncertainties on signal efficiency is shown in Tab. 11.8.

Table 11.8: Summary for systematic uncertainties that can affect the signal yield.

Systematic uncertainty (%)	2016 preVFP	2016 postVFP	2017	2018
Track reconstruction	6.32	21.30	6.06	7.05
Vertex reconstruction	9.38-14.68	16.62-19.58	16.15-18.45	16.79-17.90
ML tagging	0.49-0.57	0.05-0.10	2.74-2.57	1.51-1.62
fake MET veto efficiency	4.8-6.9	5.2-7.2	5.5-7.5	5.1-6.6
Trigger efficiency (scale factor)	1.30	1.30	0.74	0.66
Trigger efficiency (electron&muon)	2.10	1.72	1.00	1.25
Pileup	5.5	5.0	3.5	3.2
Integrated luminosity	1.2	1.2	2.3	2.5
L1 EE prefiring	0.50	0.10	0.82	0

11.2 Systematic Uncertainties Related to Background Estimation

11.2.1 ABCD Method

The background estimation method, as described in Sec. 10, uses 3-track control regions to predict the number of background events in the signal region. Besides that, 4-track control regions can also be used to predict the number of background events in the signal region. Ideally, if the two variables that define the search regions, namely nTracks/SV and MLScore, are totally decorrelated, the prediction using 3-track and 4-track control regions are supposed to be exactly the same.

The number of predicted background events in the signal region using 3-track and 4-track control regions are reported in Tab. 11.9. The results show that the 3-track and 4-track-based predictions are consistent with statistical uncertainty. So systematic uncertainty associated with the background estimation method is not assigned.

Table 11.9: Number of the predicted number of background events in the signal region using 3-track and 4-track based control regions.

	2016 preVFP	2016 postVFP	2017	2018
3-track based	0.90 ± 0.28	0.26 ± 0.12	1.82 ± 0.32	2.00 ± 0.32
4-track based	0.87 ± 0.45	0.42 ± 0.25	1.47 ± 0.47	1.78 ± 0.50

Chapter 12

Statistical Interpretation and Results

12.1 Observed Events in the Signal Region

With the signal region unblinded, the observed events compared with the prediction in the search regions are shown in Tab. 12.1, 12.2, 12.3, and 12.4. The predicted and observed numbers of events in total are reported in Tab. 12.5. The uncertainty in the prediction only represents the uncertainty resulting from the statistical uncertainty of the observed number of events used in the background estimation. In total, 9 events are observed in the signal region while the prediction of background estimation is 4.98 ± 0.55 events.

Table 12.1: Event yield for data in 2016 preVFP.

	5-track	4-track	3-track
Predicted high ML events	0.90 ± 0.28	5.16 ± 1.17	N/A
Observed high ML events	2.00	5.00	25.00
Observed low ML events	19.00	109.00	528.00

Distributions of data events in the signal region are shown in Fig. 12.1. The distributions show that all of the events have only one displaced vertex and each vertex has five tracks, with relatively low d_{BV}.

Table 12.2: Event yield for data in 2016 postVFP.

	5-track	4-track	3-track
Predicted high ML events	0.26 ± 0.12	3.12 ± 0.91	N/A
Observed high ML events	0.00	5.00	14.00
Observed low ML events	7.00	83.00	373.00

Table 12.3: Event yield for data in 2017.

	5-track	4-track	3-track
Predicted high ML events	1.82 ± 0.32	14.80 ± 1.87	N/A
Observed high ML events	3.00	12.00	75.00
Observed low ML events	56.00	456.00	2311.00

Table 12.4: Event yield for data in 2018.

	5-track	4-track	3-track
Predicted high ML events	2.00 ± 0.32	17.94 ± 2.06	N/A
Observed high ML events	4.00	16.00	89.00
Observed low ML events	70.00	628.00	3115.00

Table 12.5: Event yield for data in Run 2. The predicted numbers of events are calculated as the sum of the predictions from all individual data-taking years.

	5-track	4-track	3-track
Predicted high ML events	4.98 ± 0.55	41.02 ± 3.15	N/A
Observed high ML events	9.00	38.00	203.00
Observed low ML events	152.00	1276.00	6327.00

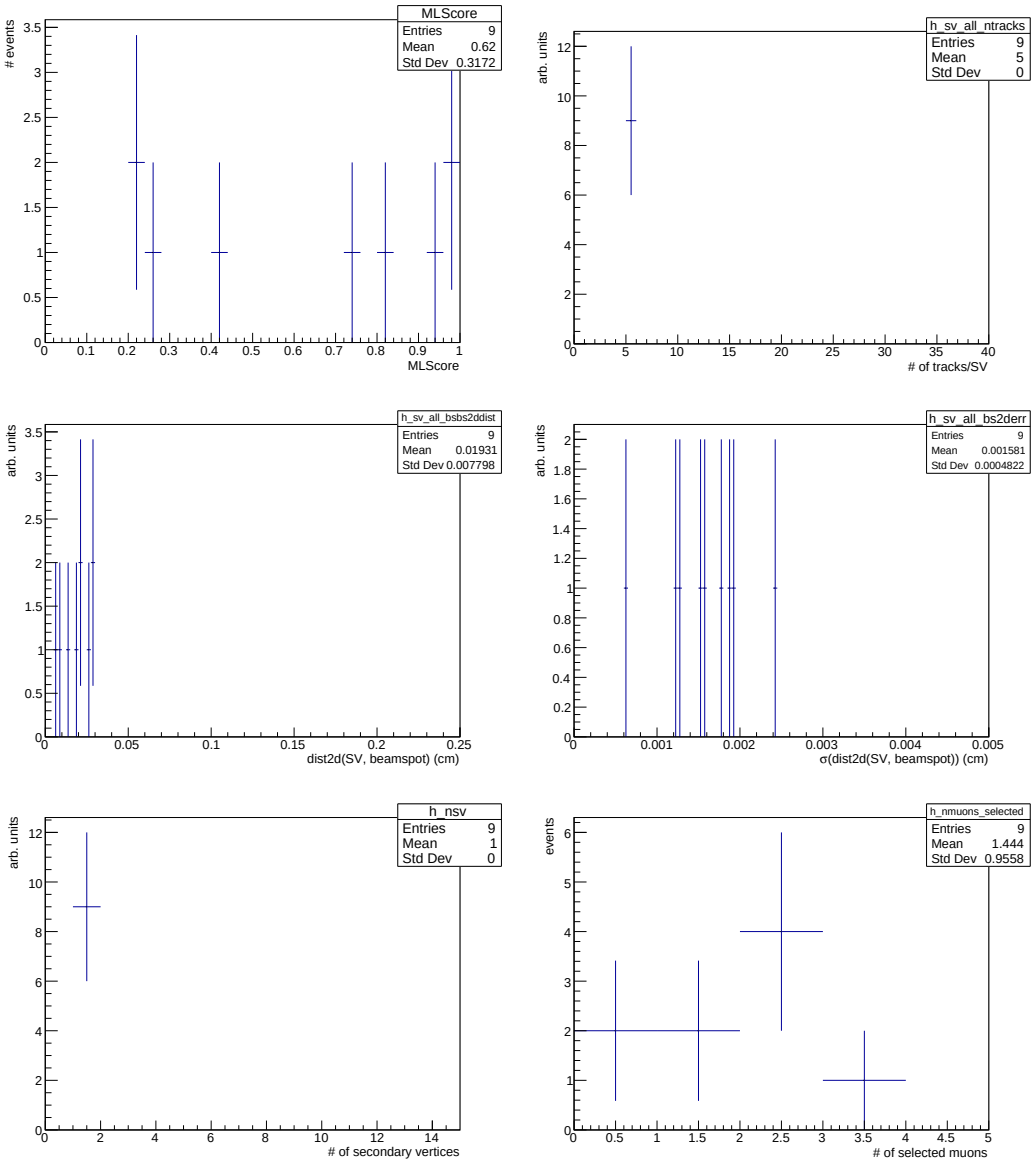

Figure 12.1: Distributions of data events in the signal region with 2016 preVFP, 2016 postVFP, 2017, and 2018 summed together.

12.2 Upper Limits on Signal Cross Section

Since no significant excess is observed in this search, the 95% confidence level (CL) upper limit on the product of production cross section and branching ratio of the given decay for signal samples considered in this search is estimated[1]. The estimation is performed simultaneously in all search regions, including the signal region, validation region, and control regions described in Ch.9.

The resulting upper limits are then compared with the theoretical prediction of the production cross section to exclude points of the parameter space of signal models in this search.

The upper limits of long-lived gluino as a function of gluino mass for different $c\tau$ values are shown in Fig. 12.2, 12.3, and 12.4. The upper limits as a function of $c\tau$ is shown in Fig. 12.5.

A comparison of the observed upper limits of this search with the ATLAS 2016 results [24] and the CMS 2016 results [54] for split SUSY samples with gluino mass of 1400 GeV and neutralino mass of 1300 GeV is shown in Fig. 12.6. This search covers low $c\tau$ such as 100 and 300 μm, which is not included in the ATLAS results. Also, this search achieves better sensitivity for $c\tau$ in the range from 1 mm to 100 mm.

In conclusion, considering the split SUSY benchmark signal, for 100 GeV mass splitting, the search excludes gluinos with masses below 1800 GeV within the $c\tau$ range of 1–100 mm; for 200 GeV mass splitting, the search excludes gluinos with masses below 2000 GeV within the $c\tau$ range of 1–100 mm.

[1] The estimation is performed using the HybridNew statistical method in the "Combine" package

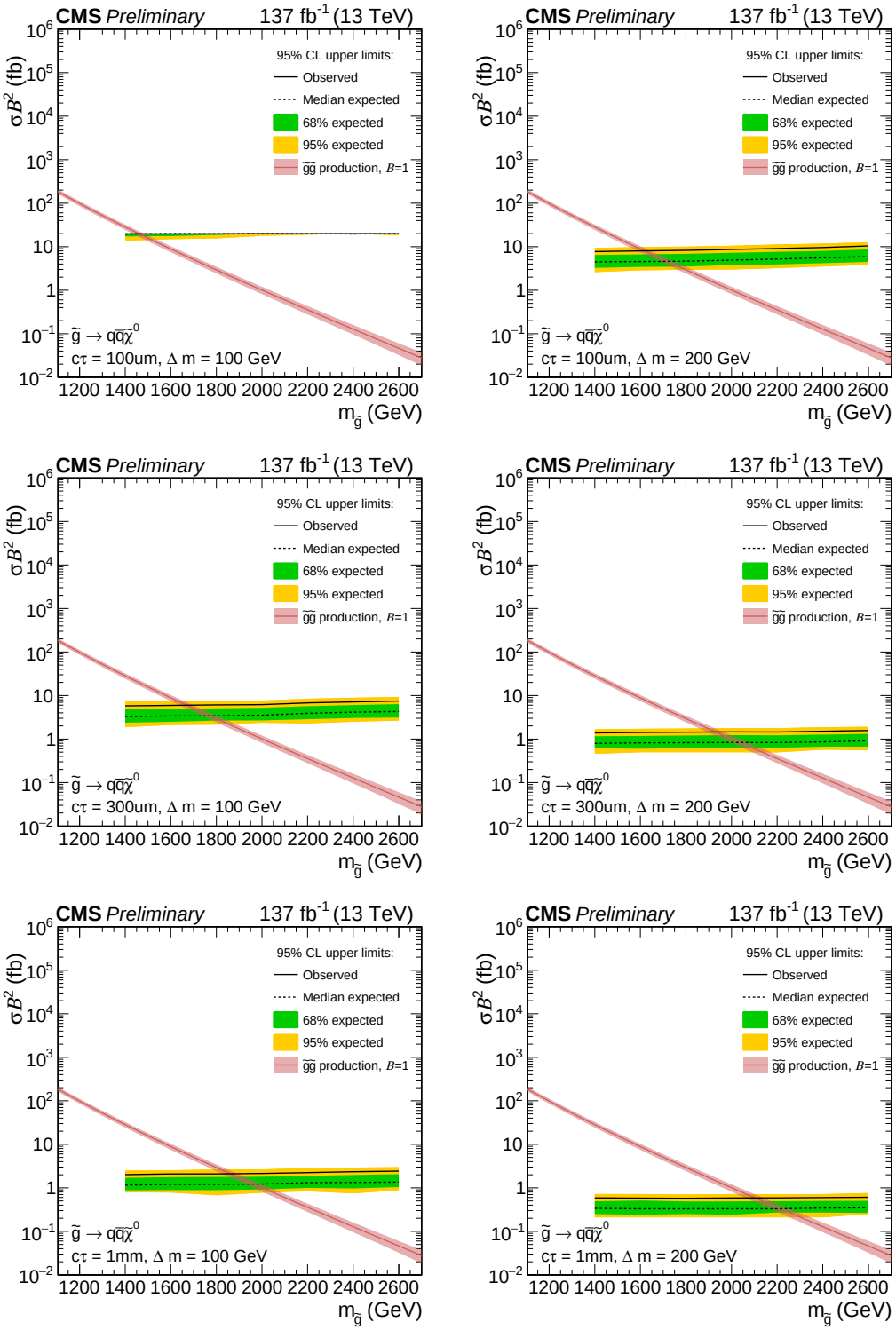

Figure 12.2: The 95% CL upper limits on cross section times branching fraction squared in Run 2 for split SUSY long-lived gluinos, as a function of gluino mass. Different $c\tau$ values of 0.1 mm (top), 0.3 mm (middle), and 1 mm (bottom) and mass splittings of 100 GeV (left), 200 GeV (right) are shown.

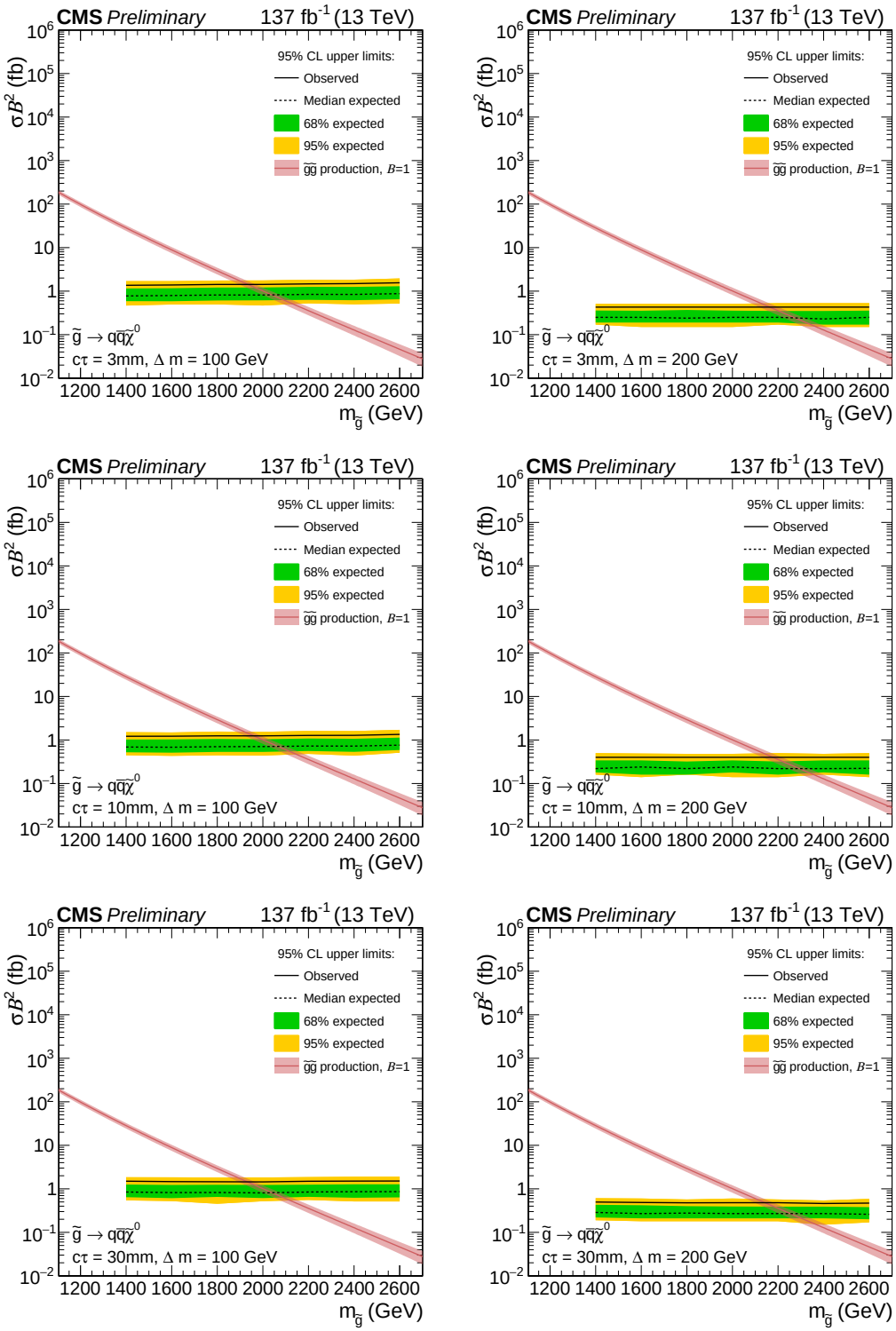

Figure 12.3: The 95% CL upper limits on cross section times branching fraction squared in Run 2 for split SUSY long-lived gluinos, as a function of gluino mass. Different $c\tau$ values of 0.1 mm (top), 0.3 mm (middle), and 1 mm (bottom) and mass splittings of 100 GeV (left), 200 GeV (right) are shown.

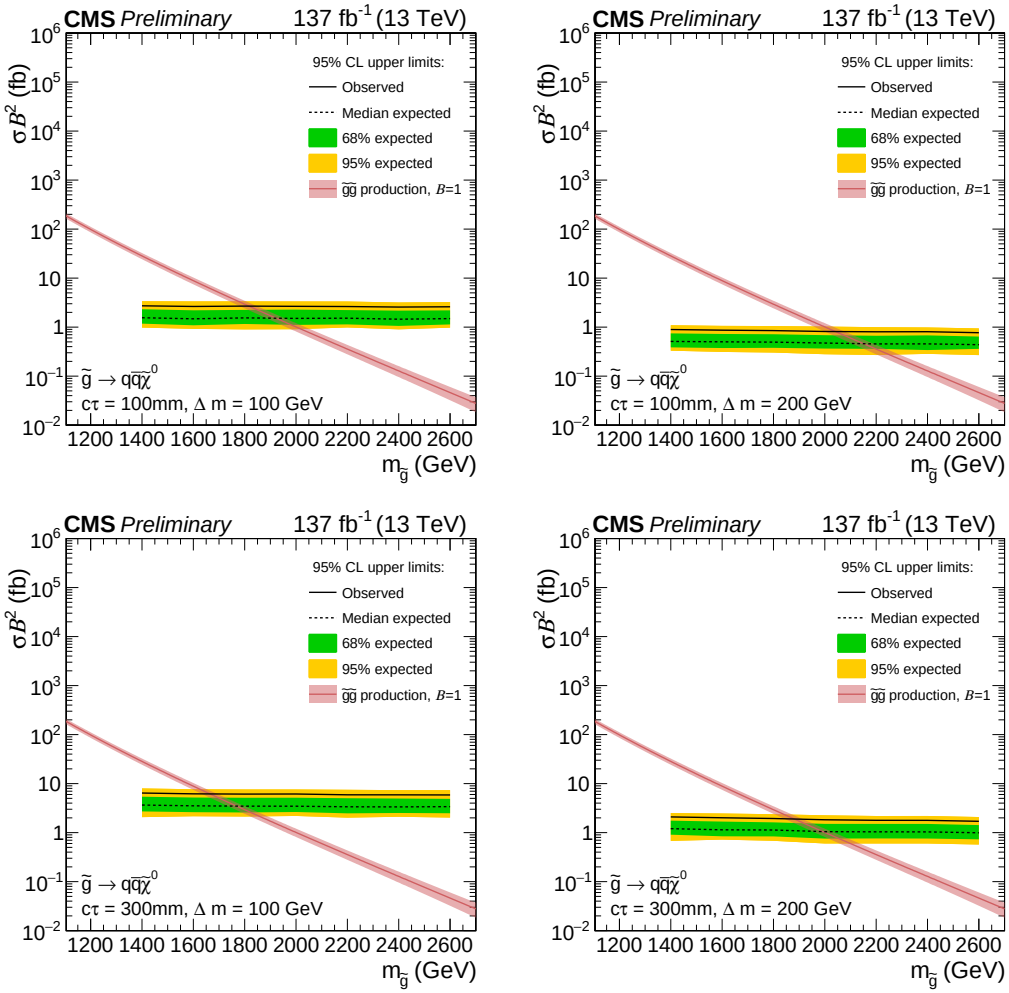

Figure 12.4: The 95% CL upper limits on cross section times branching fraction squared in Run 2 for split SUSY long-lived gluinos, as a function of gluino mass. Different $c\tau$ values of 0.1 mm (top), 0.3 mm (middle), and 1 mm (bottom) and mass splittings of 100 GeV (left), 200 GeV (right) are shown.

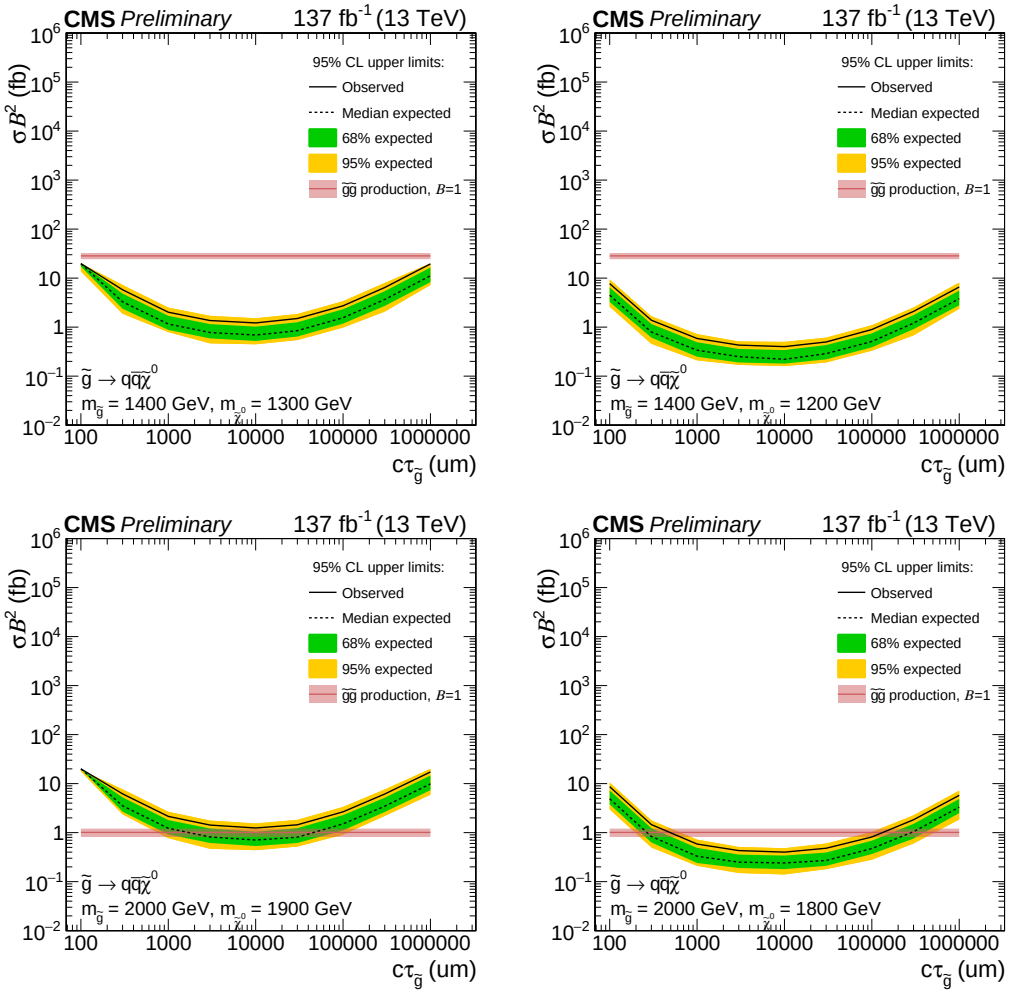

Figure 12.5: The 95% CL upper limits on cross section times branching fraction squared in Run 2 for split SUSY long-lived gluinos, as a function of gluino $c\tau$. Different gluino masses of 1400 GeV (top) and 2000 GeV (bottom) and mass splittings of 100 GeV (left), 200 GeV (right) are shown.

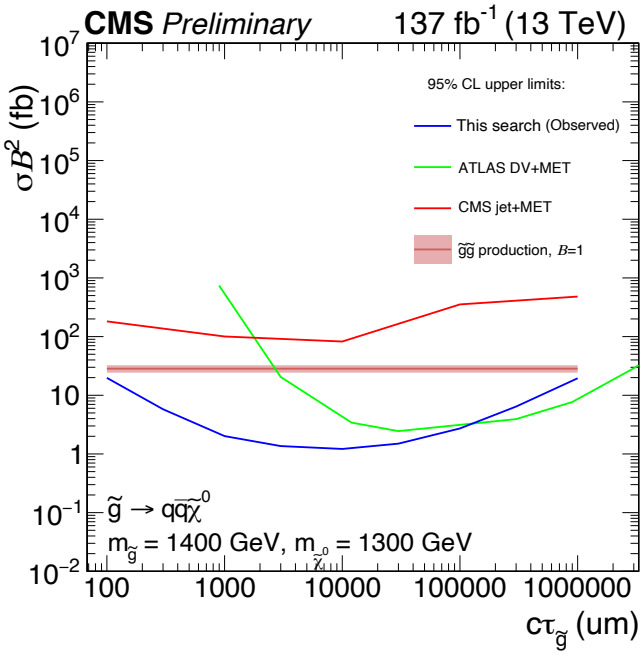

Figure 12.6: Comparison of observed 95% CL upper limits on cross section times branching fraction squared in Run 2 for split SUSY (blue), between the expected results of this analysis, the ATLAS 2016 results (green), and the CMS results (red).

Chapter 13

Conclusion

A search for long-lived particles (LLPs) that result in at least one displaced vertex and missing transverse momentum is presented in this thesis. The search utilizes proton-proton collision data with a center-of-mass energy of 13 TeV taken in the CMS experiment at the LHC during Run 2 (from 2016 to 2018), which corresponds to an integrated luminosity of $137\,\mathrm{fb}^{-1}$. This search explores a new final state, which consists of at least one displaced vertex inside the beam pipe of the CMS detector and missing transverse momentum. Only requiring one displaced vertex makes the search sensitive to LLPs with a wider $c\tau$ range. However, this makes it harder to discriminate signal events from background events. To increase the search sensitivity, customized vertex reconstruction and advanced machine learning algorithms based on the Graph Neural Network are applied. A data-driven background estimation method is developed with vertex reconstruction and ML algorithms. This search achieves a world-leading sensitivity. The 95% confidence level upper limits are set for the split SUSY benchmark signal model. In the $c\tau$ range of 1-100 mm, gluinos with masses below 1800 GeV are excluded for 100 GeV mass splitting, and gluinos with masses below 2000 GeV are excluded for 200 GeV mass splitting. The search results will become public soon, once the review process within the CMS experiment is finished.

With the model-independent design of this search, it is not only able to provide clues

to the benchmark split SUSY model, but also able to help reveal any other BSM model that could produce the targeted final state containing the displaced vertex and missing transverse momentum. Given the unclear direction of new physics, the capability of detecting a wide range of potential signal models provided by this search is an attractive feature of the approach.

Although this search brings new sensitivity, there are still some limitations. The exclusion of vertices beyond the beam pipe makes the search sensitivity decrease in the large $c\tau$ region. The decorrelation between the machine learning and the vertex reconstruction makes the machine learning tagging efficiency slightly worse for signal events with more tracks.

To further improve the search in the future, the beam pipe constraint on displaced vertices can be removed and a material map veto can be applied to get rid of background vertices from material interactions. Also, the machine learning can be modified to recognize tracks that are likely to result from LLP decays and the selected tracks can be further used to reconstruct displaced vertices. This procedure is expected to identify more tracks from LLP decays, and thus improve the vertex reconstruction efficiency. What's more, the high-luminosity LHC, which will result in much more data and bring the addition of a dedicated timing detector, will increase the search sensitivity even more. Specifically, with the help of the timing detector, the timing information will be available for tracks, which can be helpful in removing pileup tracks. The timing information of tracks can also be propagated to displaced vertices and the velocity of the LLPs can be estimated with the timing information and the distance from the proton-proton interaction point to the displaced vertices.

Bibliography

[1] Michael E. Peskin. *Concepts of Elementary Particle Physics*. Oxford Master Series in Physics. Oxford University Press, Sept. 2019. ISBN: 978-0-19-881218-0, 978-0-19-881219-7. DOI: 10.1093/oso/9780198812180.001.0001.

[2] A. B. Arbuzov. "Quantum Field Theory and the Electroweak Standard Model". In: (2017). Ed. by Martijn Mulders and Giulia Zanderighi, pp. 1–34. DOI: 10.23730/CYRSP-2017-004.1. arXiv: 1801.05670 [hep-ph].

[3] Leonard Susskind. "THE GAUGE HIERARCHY PROBLEM, TECHNICOLOR, SUPERSYMMETRY, AND ALL THAT." In: *Phys. Rept.* 104 (1984). Ed. by Anne Mosher et al., pp. 181–193. DOI: 10.1016/0370-1573(84)90208-4.

[4] Stephen P. Martin. "A Supersymmetry primer". In: *Adv. Ser. Direct. High Energy Phys.* 18 (1998). Ed. by Gordon L. Kane, pp. 1–98. DOI: 10.1142/9789812839657_0001. arXiv: hep-ph/9709356.

[5] Nima Arkani-Hamed and Savas Dimopoulos. "Supersymmetric unification without low energy supersymmetry and signatures for fine-tuning at the LHC". In: *JHEP* 06 (2005), p. 073. DOI: 10.1088/1126-6708/2005/06/073. arXiv: hep-th/0405159.

[6] G. F. Giudice and A. Romanino. "Split supersymmetry". In: *Nucl. Phys. B* 699 (2004). [Erratum: Nucl.Phys.B 706, 487–487 (2005)], pp. 65–89. DOI: 10.1016/j.nuclphysb.2004.08.001. arXiv: hep-ph/0406088.

[7] "LHC Machine". In: *JINST* 3 (2008). Ed. by Lyndon Evans and Philip Bryant, S08001. DOI: 10.1088/1748-0221/3/08/S08001.

[8] G. Aad et al. “The ATLAS Experiment at the CERN Large Hadron Collider”. In: *JINST* 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003.

[9] S. Chatrchyan et al. “The CMS Experiment at the CERN LHC”. In: *JINST* 3 (2008), S08004. DOI: 10.1088/1748-0221/3/08/S08004.

[10] A. Augusto Alves Jr. et al. “The LHCb Detector at the LHC”. In: *JINST* 3 (2008), S08005. DOI: 10.1088/1748-0221/3/08/S08005.

[11] K. Aamodt et al. “The ALICE experiment at the CERN LHC”. In: *JINST* 3 (2008), S08002. DOI: 10.1088/1748-0221/3/08/S08002.

[12] Jean-Luc Caron. “LHC layout. Schema general du LHC”. AC Collection. Legacy of AC. Pictures from 1992 to 2002. 1997. URL: https://cds.cern.ch/record/841573.

[13] “Restarting the LHC: Why 13 Tev?” In: (2014). URL: https://cds.cern.ch/record/1998739.

[14] I. Zurbano Fernandez et al. “High-Luminosity Large Hadron Collider (HL-LHC): Technical design report”. In: 10/2020 (Dec. 2020). Ed. by I. Béjar Alonso et al. DOI: 10.23731/CYRM-2020-0010.

[15] Ewa Lopienska. “The CERN accelerator complex, layout in 2022. Complexe des accélérateurs du CERN en janvier 2022”. In: (2022). General Photo. URL: https://cds.cern.ch/record/2800984.

[16] Albert M Sirunyan et al. “Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC”. In: *JINST* 16 (2021), P05014. DOI: 10.1088/1748-0221/16/05/P05014. arXiv: 2012.06888 [hep-ex].

[17] A. M. Sirunyan et al. "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *JINST* 13 (2018), P06015. DOI: 10.1088/1748-0221/13/06/P06015. arXiv: 1804.04528 [physics.ins-det].

[18] Serguei Chatrchyan et al. "Description and performance of track and primary-vertex reconstruction with the CMS tracker". In: *JINST* 9.10 (2014), P10009. DOI: 10.1088/1748-0221/9/10/P10009. arXiv: 1405.6569 [physics.ins-det].

[19] Tai Sakuma. "3D SketchUp images of the CMS detector (120918)". In: (2016). URL: https://cds.cern.ch/record/2628527.

[20] W. Adam et al. "The CMS Phase-1 Pixel Detector Upgrade". In: *JINST* 16.02 (2021), P02027. DOI: 10.1088/1748-0221/16/02/P02027. arXiv: 2012.14304 [physics.ins-det].

[21] W. Adam et al. "The CMS high level trigger". In: *Eur. Phys. J. C* 46 (2006), pp. 605–667. DOI: 10.1140/epjc/s2006-02495-8. arXiv: hep-ex/0512077.

[22] A. M. Sirunyan et al. "Particle-flow reconstruction and global event description with the CMS detector". In: *JINST* 12.10 (2017), P10003. DOI: 10.1088/1748-0221/12/10/P10003. arXiv: 1706.04965 [physics.ins-det].

[23] Juliette Alimena et al. "Searching for long-lived particles beyond the Standard Model at the Large Hadron Collider". In: *J. Phys. G* 47.9 (2020), p. 090501. DOI: 10.1088/1361-6471/ab4574. arXiv: 1903.04497 [hep-ex].

[24] Morad Aaboud et al. "Search for long-lived, massive particles in events with displaced vertices and missing transverse momentum in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector". In: *Phys. Rev. D* 97.5 (2018), p. 052012. DOI: 10.1103/PhysRevD.97.052012. arXiv: 1710.04901 [hep-ex].

[25] Georges Aad et al. "Search for long-lived, massive particles in events with a displaced vertex and a muon with large impact parameter in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Rev. D* 102.3 (2020), p. 032006. DOI: 10.1103/PhysRevD.102.032006. arXiv: 2003.11956 [hep-ex].

[26] Albert M Sirunyan et al. "Search for long-lived particles decaying to jets with displaced vertices in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *Phys. Rev. D* 104.5 (2021), p. 052011. DOI: 10.1103/PhysRevD.104.052011. arXiv: 2104.13474 [hep-ex].

[27] Albert M Sirunyan et al. "Search for long-lived particles using displaced jets in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *Phys. Rev. D* 104.1 (2021), p. 012015. DOI: 10.1103/PhysRevD.104.012015. arXiv: 2012.01581 [hep-ex].

[28] S. Agostinelli et al. "GEANT4–a simulation toolkit". In: *Nucl. Instrum. Meth. A* 506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.

[29] Torbjörn Sjöstrand et al. "An introduction to PYTHIA 8.2". In: *Comput. Phys. Commun.* 191 (2015), pp. 159–177. DOI: 10.1016/j.cpc.2015.01.024. arXiv: 1410.3012 [hep-ph].

[30] Christoph Borschensky et al. "Squark and gluino production cross sections in pp collisions at $\sqrt{s} = 13$, 14, 33 and 100 TeV". In: *Eur. Phys. J. C* 74 (2014), p. 3174. DOI: 10.1140/epjc/s10052-014-3174-y. arXiv: 1407.5066 [hep-ph].

[31] Richard D. Ball et al. "Parton distributions from high-precision collider data". In: *Eur. Phys. J. C* 77.10 (2017), p. 663. DOI: 10.1140/epjc/s10052-017-5199-5. arXiv: 1706.00428 [hep-ph].

[32] Albert M Sirunyan et al. "Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements". In: *Eur. Phys. J. C*

80.1 (2020), p. 4. DOI: 10.1140/epjc/s10052-019-7499-4. arXiv: 1903.12179 [hep-ex].

[33] J. Alwall et al. "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations". In: *JHEP* 07 (2014), p. 079. DOI: 10.1007/JHEP07(2014)079. arXiv: 1405.0301 [hep-ph].

[34] Johan Alwall et al. "Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions". In: *Eur. Phys. J. C* 53 (2008), pp. 473–500. DOI: 10.1140/epjc/s10052-007-0490-5. arXiv: 0706.2569 [hep-ph].

[35] Rikkert Frederix and Stefano Frixione. "Merging meets matching in MC@NLO". In: *JHEP* 12 (2012), p. 061. DOI: 10.1007/JHEP12(2012)061. arXiv: 1209.6215 [hep-ph].

[36] Stefano Frixione, Paolo Nason, and Carlo Oleari. "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method". In: *JHEP* 11 (2007), p. 070. DOI: 10.1088/1126-6708/2007/11/070. arXiv: 0709.2092 [hep-ph].

[37] Simone Alioli et al. "NLO single-top production matched with shower in POWHEG: s- and t-channel contributions". In: *JHEP* 09 (2009). [Erratum: JHEP 02, 011 (2010)], p. 111. DOI: 10.1088/1126-6708/2009/09/111. arXiv: 0907.4076 [hep-ph].

[38] Richard D. Ball et al. "Parton distributions for the LHC Run II". In: *JHEP* 04 (2015), p. 040. DOI: 10.1007/JHEP04(2015)040. arXiv: 1410.8849 [hep-ph].

[39] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "The anti-k_{T} jet clustering algorithm". In: *JHEP* 04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ex].

[40] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "FastJet user manual". In: *Eur. Phys. J. C* 72 (2012), p. 1896. DOI: 10.1140/epjc/s10052-012-1896-2. arXiv: 1111.6097 [hep-ph].

[41] Vardan Khachatryan et al. "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV". In: *JINST* 12.02 (2017), P02014. DOI: 10.1088/1748-0221/12/02/P02014. arXiv: 1607.03663 [hep-ex].

[42] Emil Bols et al. "Jet Flavour Classification Using DeepJet". In: *JINST* 15.12 (2020), P12012. DOI: 10.1088/1748-0221/15/12/P12012. arXiv: 2008.10519 [hep-ex].

[43] R. Fruhwirth. "Application of Kalman filtering to track and vertex fitting". In: *Nucl. Instrum. Meth. A* 262 (1987), pp. 444–450. DOI: 10.1016/0168-9002(87)90887-4.

[44] Christian Janiesch, Patrick Zschech, and Kai Heinrich. "Machine Learning and Deep Learning". In: *Electronic Markets* 31.3 (2021), pp. 685–695. DOI: 10.1007/s12525-021-00475-2.

[45] Peter W. Battaglia et al. *Interaction Networks for Learning about Objects, Relations and Physics*. 2016. arXiv: 1612.00222 [cs.AI].

[46] Jie Zhou et al. "Graph neural networks: A review of methods and applications". In: *AI Open* 1 (2020), pp. 57–81.

[47] Katarzyna Janocha and Wojciech Marian Czarnecki. *On Loss Functions for Deep Neural Networks in Classification*. 2017. DOI: 10.48550/ARXIV.1702.05659. URL: https://arxiv.org/abs/1702.05659.

[48] Andrew Y. Ng. "Feature Selection, L1 vs. L2 Regularization, and Rotational Invariance". In: *Proceedings of the Twenty-First International Conference on Machine Learning*. ICML '04. Banff, Alberta, Canada: Association for Computing Machinery, 2004, p. 78. ISBN: 1581138385. DOI: 10.1145/1015330.1015435. URL: https://doi.org/10.1145/1015330.1015435.

[49] Gregor Kasieczka et al. "Automating the ABCD method with machine learning". In: *Physical Review D* 103.3 (Feb. 2021). ISSN: 2470-0029. DOI: 10.1103/physrevd.103.035021. URL: http://dx.doi.org/10.1103/PhysRevD.103.035021.

[50] Gregor Kasieczka and David Shih. "Robust Jet Classifiers through Distance Correlation". In: *Phys. Rev. Lett.* 125.12 (2020), p. 122001. DOI: 10.1103/PhysRevLett.125.122001. arXiv: 2001.05310 [hep-ph].

[51] Diederik P. Kingma and Jimmy Ba. *Adam: A Method for Stochastic Optimization*. 2014. DOI: 10.48550/ARXIV.1412.6980. URL: https://arxiv.org/abs/1412.6980.

[52] David Curtin et al. "Exotic decays of the 125 GeV Higgs boson". In: *Phys. Rev. D* 90.7 (2014), p. 075004. DOI: 10.1103/PhysRevD.90.075004. arXiv: 1312.4992 [hep-ph].

[53] Csaba Csaki, Yuval Grossman, and Ben Heidenreich. "MFV SUSY: A Natural Theory for R-Parity Violation". In: *Phys. Rev. D* 85 (2012), p. 095009. DOI: 10.1103/PhysRevD.85.095009. arXiv: 1111.1239 [hep-ph].

[54] Albert M Sirunyan et al. "Search for natural and split supersymmetry in proton-proton collisions at $\sqrt{s} = 13$ TeV in final states with jets and missing transverse momentum". In: *JHEP* 05 (2018), p. 025. DOI: 10.1007/JHEP05(2018)025. arXiv: 1802.02110 [hep-ex].

Appendices

Appendix A

Background Estimation Closure Test on Simulation

The closure test in MC is done using slightly different events selection. The only difference is the fake MET veto. In the nominal analysis, we require $|MET_{PF} - MET_{CALO}|/MET_{CALO} < 0.6$, while for this test, we required $|MET_{PF} - MET_{CALO}|/MET_{CALO} < 0.5$ because the selection was not fully determined when the check is performed.

The closure test results are shown in Table A.1, A.2, A.3, A.4, and A.5 for 2016 preVFP, 2016 postVFP, 2017, 2018, and all Run 2 added together.

However, given the low number of events, it is hard to draw conclusion about the closure. Given the large uncertainties we observed in MC and the fact that the ABCD method is purely data-driven, the differences observed in simulation is not a concern.

Table A.1: Closure test results for background MC in 2016 preVFP.

	5-track	4-track	3-track
Predicted high ML events	1.08 ± 0.42	10.14 ± 3.25	N/A
Measured high ML events	0.73 ± 0.31	3.09 ± 1.08	45.64 ± 13.74
Measured low ML events	12.14 ± 2.94	114.46 ± 10.70	515.40 ± 29.55

Table A.2: Closure test results for background MC in 2016 postVFP.

	5-track	4-track	3-track
Predicted high ML events	0.59 ± 0.27	4.40 ± 1.11	N/A
Measured high ML events	1.13 ± 0.54	2.52 ± 0.84	20.75 ± 4.01
Measured low ML events	8.74 ± 3.49	64.90 ± 9.55	306.13 ± 21.52

Table A.3: Closure test results for background MC in 2017.

	5-track	4-track	3-track
Predicted high ML events	0.91 ± 0.40	6.77 ± 1.08	N/A
Measured high ML events	0.24 ± 0.48	18.57 ± 8.96	38.62 ± 3.45
Measured low ML events	20.68 ± 8.76	154.71 ± 18.18	882.01 ± 53.10

Table A.4: Closure test results for background MC in 2018.

	5-track	4-track	3-track
Predicted high ML events	0.98 ± 0.39	9.11 ± 2.18	N/A
Measured high ML events	-0.07 ± 0.32	13.91 ± 7.69	59.45 ± 12.91
Measured low ML events	23.19 ± 7.71	215.95 ± 19.22	1409.62 ± 65.78

Table A.5: Closure test results for background MC in Run 2.

	5-track	4-track	3-track
Predicted high ML events	3.42 ± 0.78	29.06 ± 3.90	N/A
Measured high ML events	2.03 ± 0.85	38.09 ± 11.89	164.45 ± 19.58
Measured low ML events	64.76 ± 12.53	550.02 ± 30.09	3113.16 ± 92.11

Appendix B

Data/Simulation Comparison

With the search strategy defined, data and simulated events are compared to understand the potential differences between data and simulation. Also, since the search strategy is largely dependent on the simulation, comparing data with simulation ensures that the search strategy works in data.

B.1 Track Variables

It is noticed that track uncertainty distributions between data and simulation are different. This section compares the differences and describes a method to correct the differences. All the variables and uncertainties relevant to the displacement studied in this section are calculated with respect to the beam spot. Most of the plots in this section use 2017 data and MC as examples. All events studied in this section are required to pass the event preselection defined in Sec. 6.3[1] and the events do not necessarily have a reconstructed secondary vertex. All the tracks are required to satisfy the track quality selection defined in Sec. 7.1. Track displacement, namely $|d_{xy}|/\sigma_{d_{xy}}$, is not required for this study. The detailed plots for 2016, 2017, and 2018 are shown in Appendix C.

Comparisons are made between 10% of data and a subset of background simulation

[1]The fake MET veto threshold is 0.5 instead of 0.6 in this study

samples:

- $t\bar{t}$
- QCD, $500 < H_{\mathrm{T}} < 700\,\mathrm{GeV}$
- QCD, $700 < H_{\mathrm{T}} < 1000\,\mathrm{GeV}$
- QCD, $1000 < H_{\mathrm{T}} < 1500\,\mathrm{GeV}$
- QCD, $1500 < H_{\mathrm{T}} < 2000\,\mathrm{GeV}$
- QCD, $H_{\mathrm{T}} > 2000\,\mathrm{GeV}$

Figure B.1, B.2, and B.3 show the comparison of track p_{T}, absolute impact parameter $|d_{xy}|$, and impact parameter uncertainty $\sigma_{d_{xy}}$ distributions between data and simulation for tracks with $|\eta| < 1.5$. The track p_{T} and absolute impact parameter distributions show that tracks in data are generally softer and more displaced. The impact parameter uncertainty plots show that the simulation tends to overestimate the track impact parameter uncertainty for 2016 but underestimate that for 2017 and 2018. Track impact parameter uncertainty $\sigma_{d_{xy}}$ is used during vertex reconstruction to decide whether to assign tracks to vertices. Because of that, different distributions of $\sigma_{d_{xy}}$ can make vertices reconstructed from different sets of tracks and have different vertex uncertainties. As a result, rescaling track $\sigma_{d_{xy}}$ to make the distribution similar in data and simulation can be helpful to make reconstructed vertices behave in a consistent way in data and simulation. The rescaling process is developed based on the procedure in [26].

The idea of rescaling $\sigma_{d_{xy}}$ is applying scale factors to $\sigma_{d_{xy}}$ for simulated samples so data and simulation have similar values. Since the tracks reconstructed in the barrel

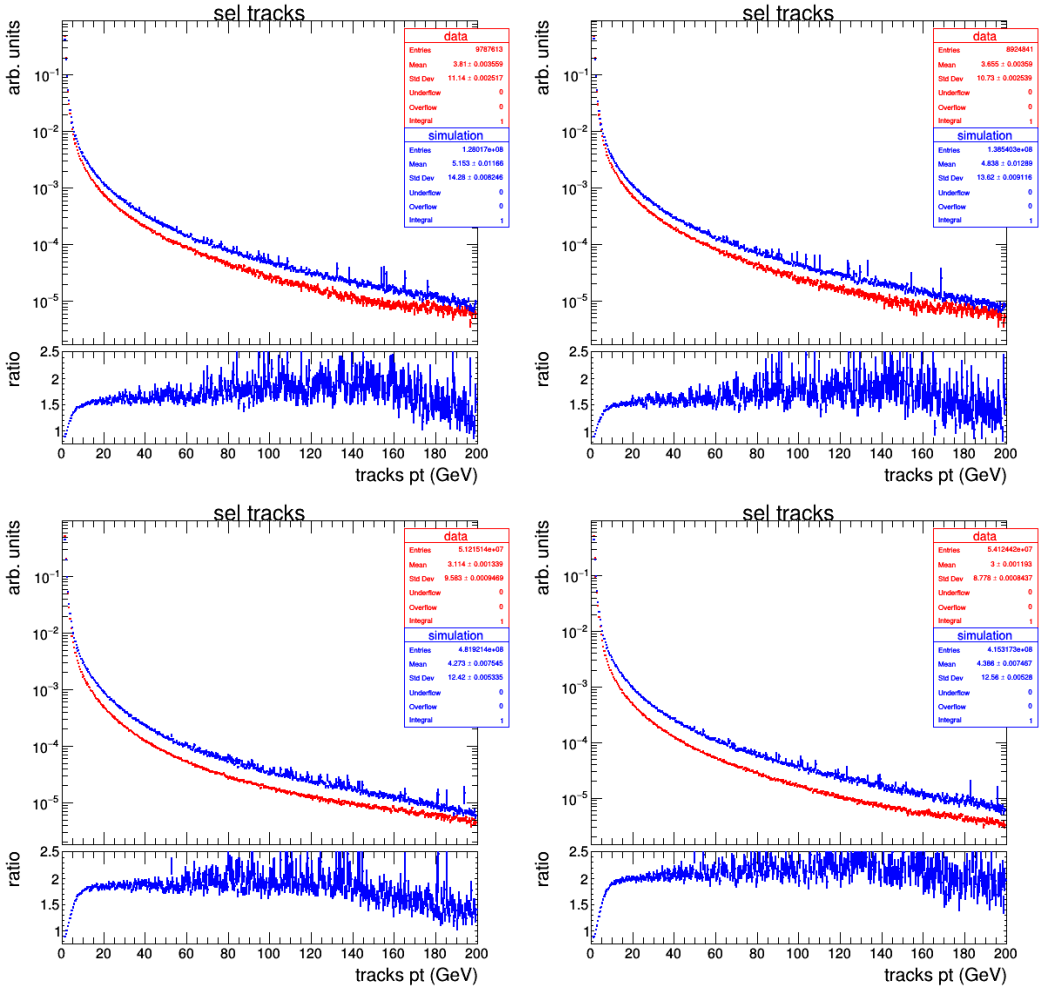

Figure B.1: Comparisons of track p_{T} for tracks with $|\eta| < 1.5$ between data and simulation.
(Top left) Tracks in 2016 preVFP.
(Top right) Tracks in 2016 postVFP.
(Bottom left) Tracks in 2017.
(Bottom right) Tracks in 2018.

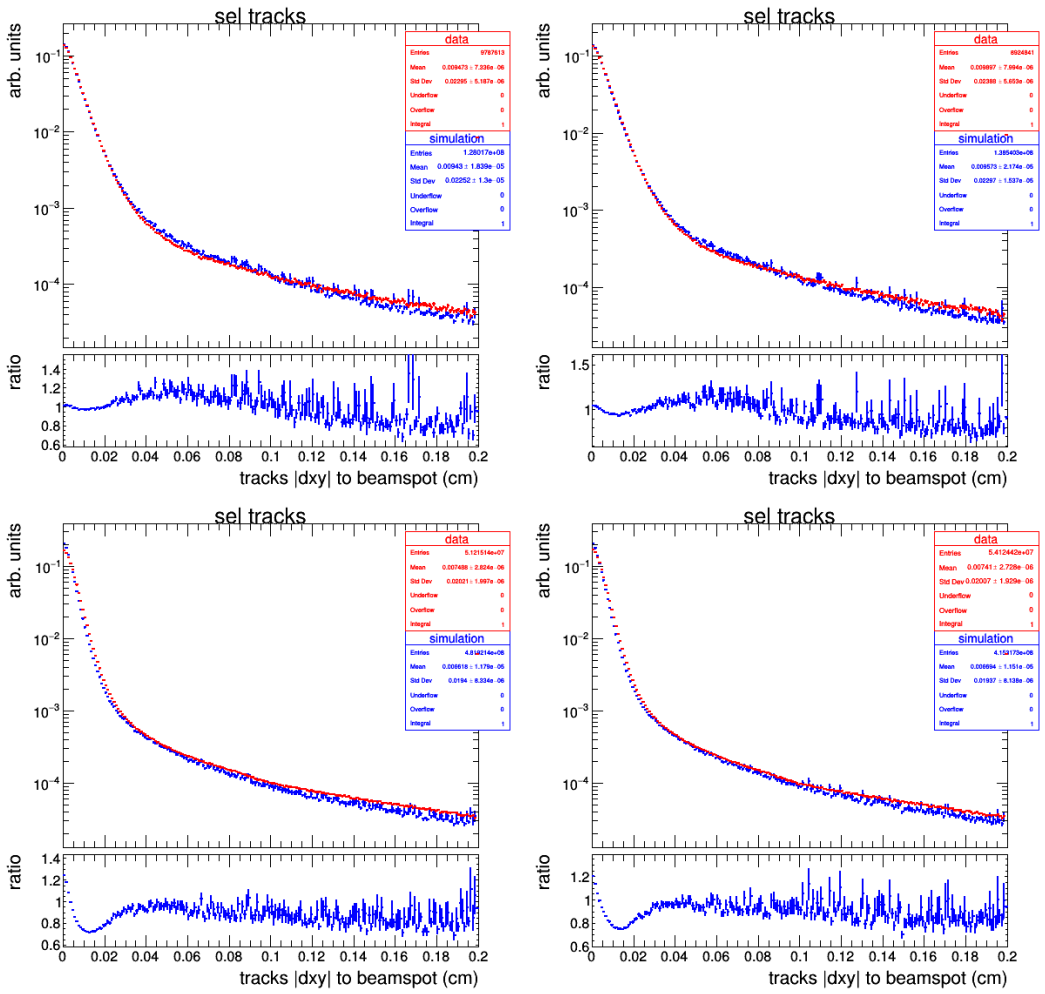

Figure B.2: Comparisons of track absolute impact parameter for tracks with $|\eta| < 1.5$ between data and simulation.
(Top left) Tracks in 2016 preVFP.
(Top right) Tracks in 2016 postVFP.
(Bottom left) Tracks in 2017.
(Bottom right) Tracks in 2018.

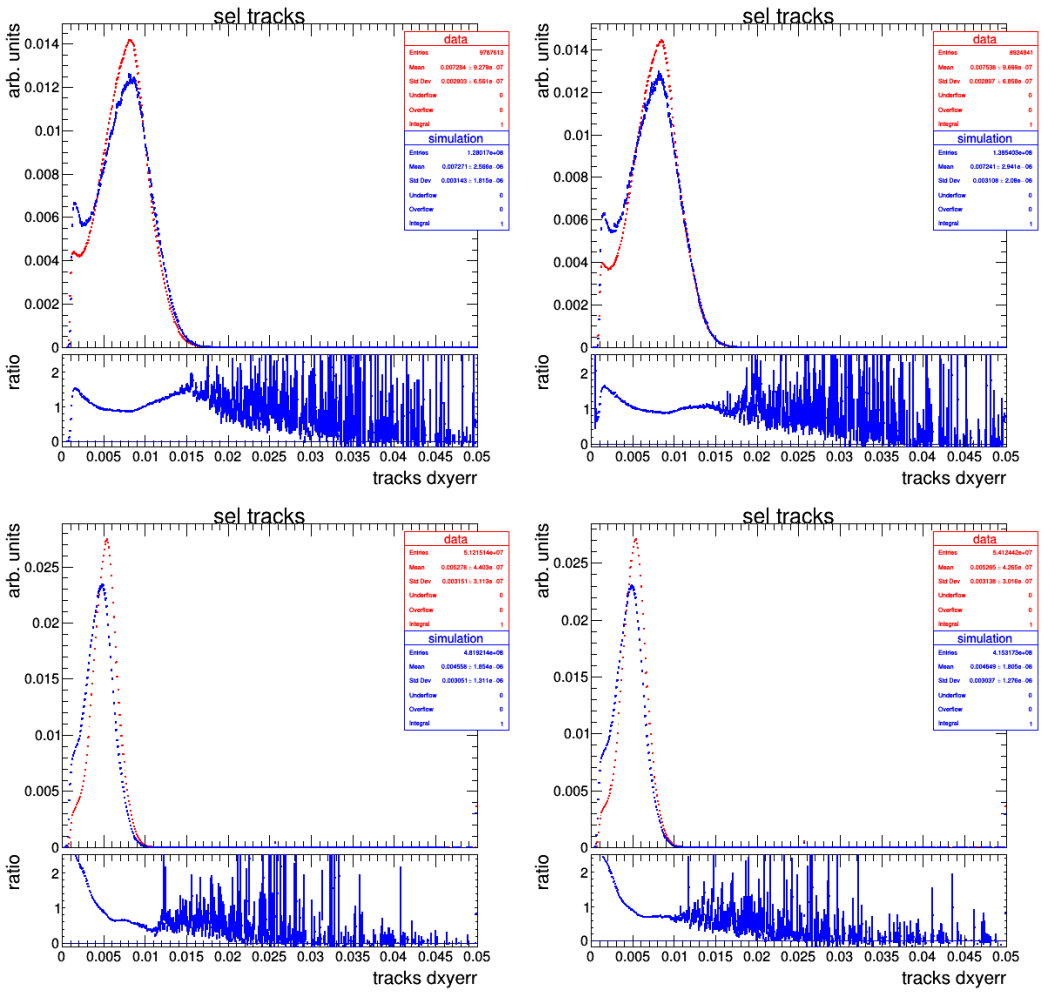

Figure B.3: Comparisons of track impact parameter uncertainties for tracks with $|\eta| < 1.5$ between data and simulation. Units are cm.
(Top left) Tracks in 2016 preVFP.
(Top right) Tracks in 2016 postVFP.
(Bottom left) Tracks in 2017.
(Bottom right) Tracks in 2018.

can be different compared with tracks reconstructed in the endcap, tracks are treated individually in two $|\eta|$ bins divided at 1.5. The track measurement in CMS depends on p_{T} of tracks. Normally tracks with higher p_{T} have smaller uncertainties. Based on that, for each η range, the rescaling is made as a function of track p_{T}. Profiling is used to study the relationship between $\sigma_{d_{xy}}$ and p_{T}, and the data/simulation ratio of the profiles is used to derive the scale factors for tracks. Different eras during the same year can be different so the comparison between data and simulation is done separately for every single data-taking era. The profile and data/simulation ratio that shows the track mean impact parameter as a function of p_{T} is shown in Fig. B.4 for 2017.

With the data/simulation ratio plot, scale factors can be derived by fitting functions to the ratio plot. Figure B.5 shows an example of the fit for 2017B data. When corrections are applied on a given simulated track for a specific data-taking era, the original track $\sigma_{d_{xy}}$ is scaled by the scale factors that are calculated using the fitted function.

Figure B.6 shows the $\sigma_{d_{xy}}$ for tracks in 2017B after the simulation is corrected. The mean $\sigma_{d_{xy}}$ profile shows that the relationship between $\sigma_{d_{xy}}$ and p_{T} is consistent between data and simulation. The remaining difference between the $\sigma_{d_{xy}}$ distribution results from the difference between track p_{T} distribution, as shown in Fig. B.1. Figure B.7 shows the data/MC ratio of mean $\sigma_{d_{xy}}$ as a function of track p_{T} after the corrections are applied. The ratios are close to 1, which proves that the correction works as expected.

Besides $\sigma_{d_{xy}}$, there are other track covariance matrix elements that behave differently between data and simulation. Although other covariance matrix elements do not directly participate in vertex variable calculations, they can make the calculation of

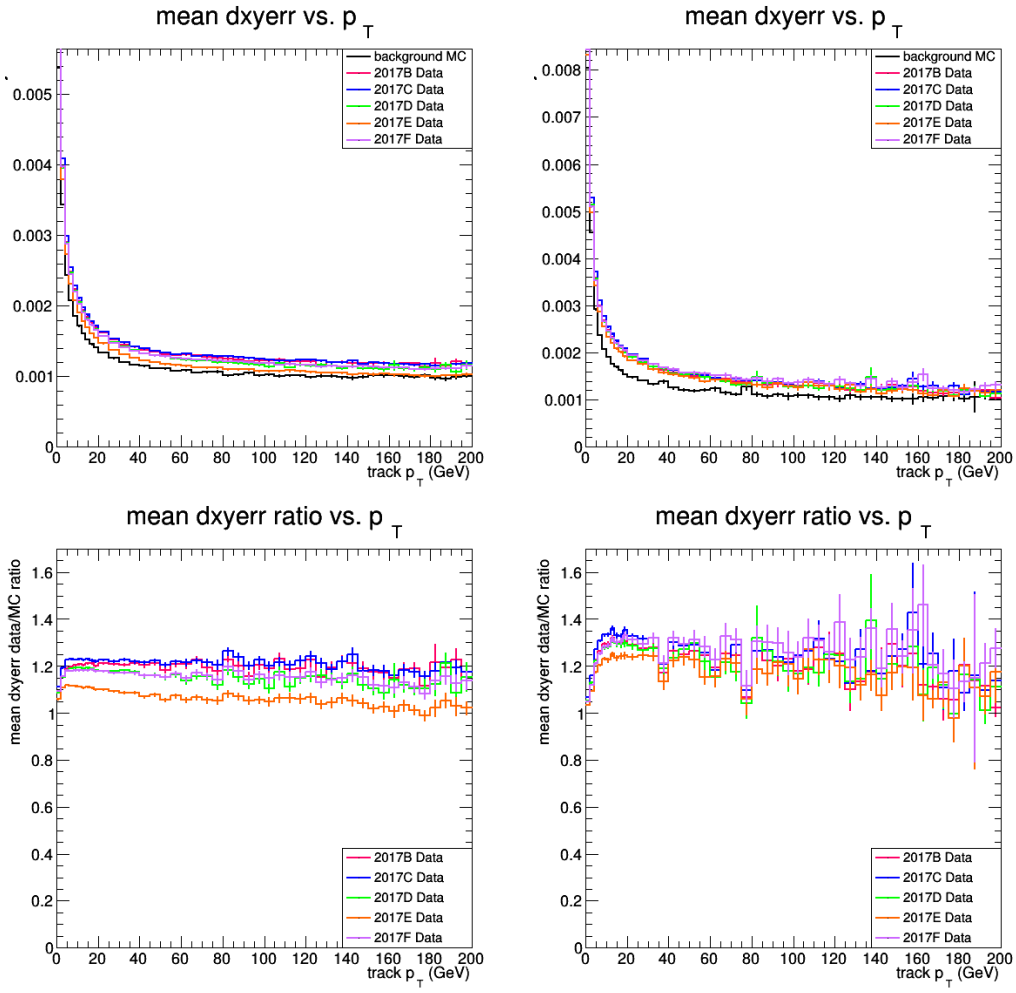

Figure B.4: Mean $\sigma_{d_{xy}}$ profiles (top) and mean $\sigma_{d_{xy}}$ data / MC ratios (bottom) for $\eta < 1.5$ (left) and $\eta > 1.5$ (right) in 2017.

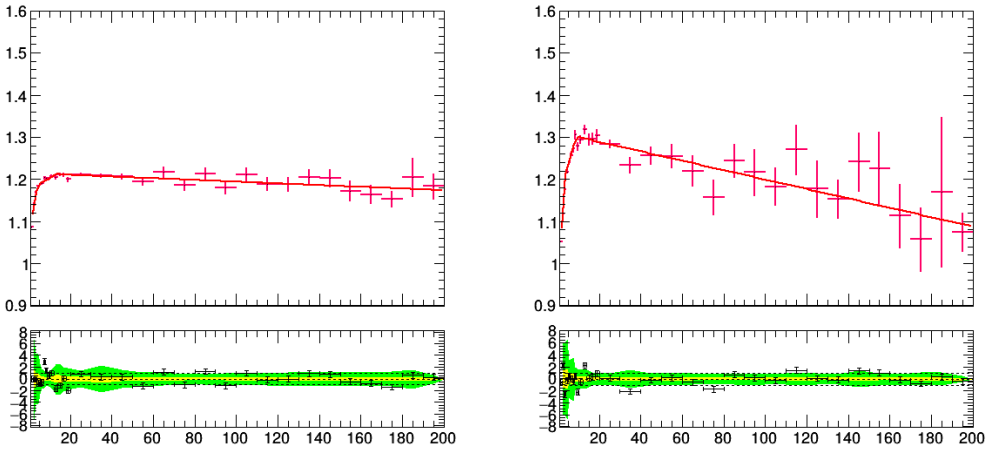

Figure B.5: Fit applied to the 2017B era data/simulation $\sigma_{d_{xy}}$ ratio vs. p_{T} curve for $|\eta| < 1.5$ (left) and $|\eta| > 1.5$ (right). The graphs below the fits show the residuals of the actual values to the fitted curve, i.e. their difference divided by the uncertainty.

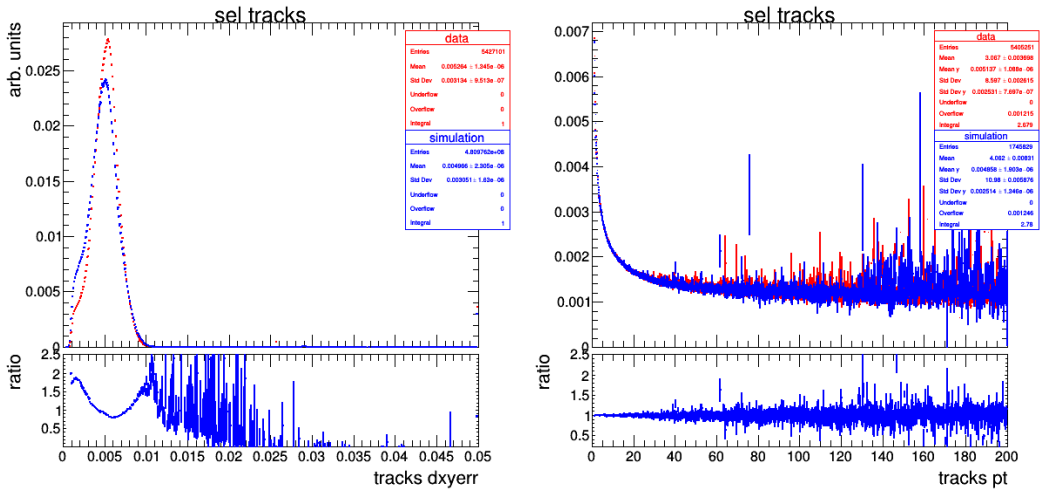

Figure B.6: Distribution of $\sigma_{d_{xy}}$ (left) and mean $\sigma_{d_{xy}}$ profile after the scaling applied on tracks for 2017B. Units are cm.

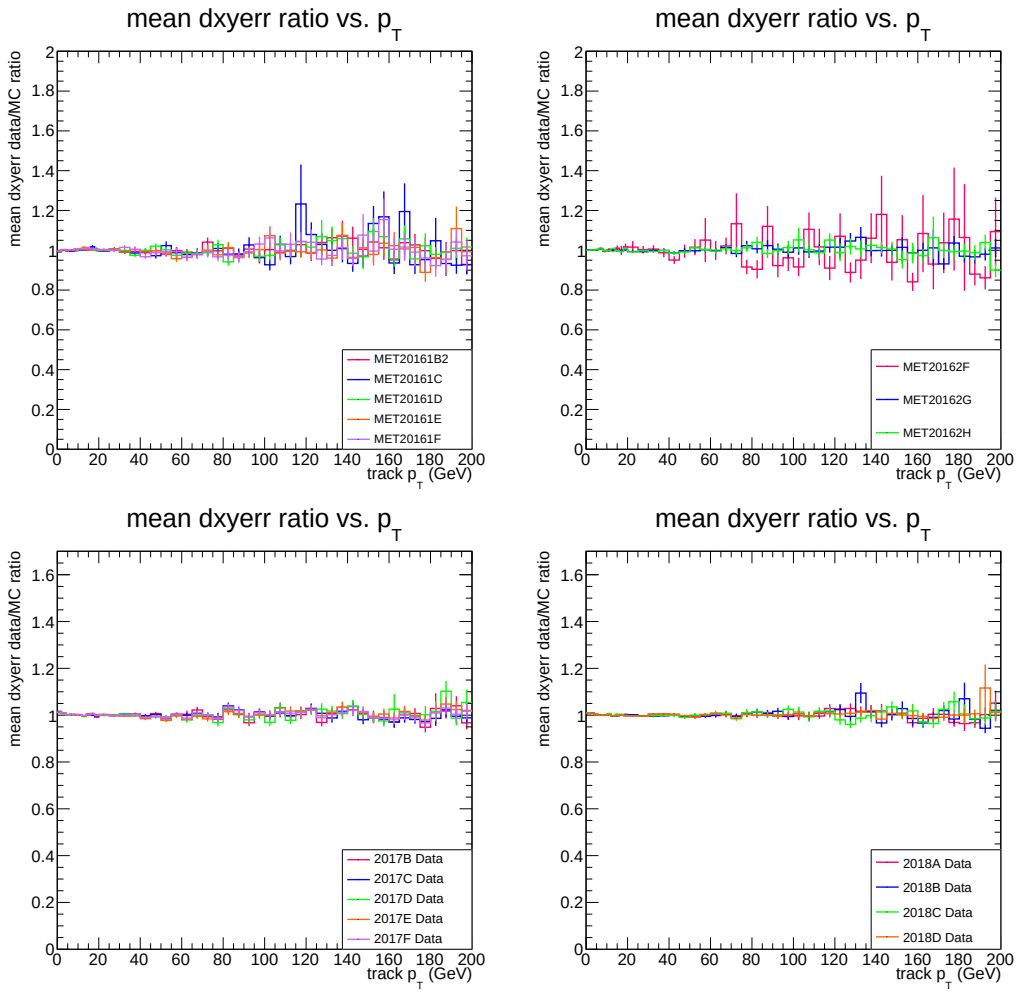

Figure B.7: Mean $\sigma_{d_{xy}}$ data / MC ratios for $|\eta| < 1.5$ in 2016 preVFP (top left), 2016 postVFP (top right), 2017 (bottom left), and 2018 (bottom right).

vertex position uncertainties or different sets of tracks assigned to the vertex different, which will make the vertex variables different in the end. As a result, corrections on other track covariance matrix elements, namely $\sigma_{d_{sz}}$ and the cross-term $\sigma_{d_{xy}} - \sigma_{d_{sz}}$, are worth considering. Figure B.8 shows the difference of $\sigma_{d_{sz}}$ and $\sigma_{d_{xy}} - \sigma_{d_{sz}}$ distribution between simulation and data.

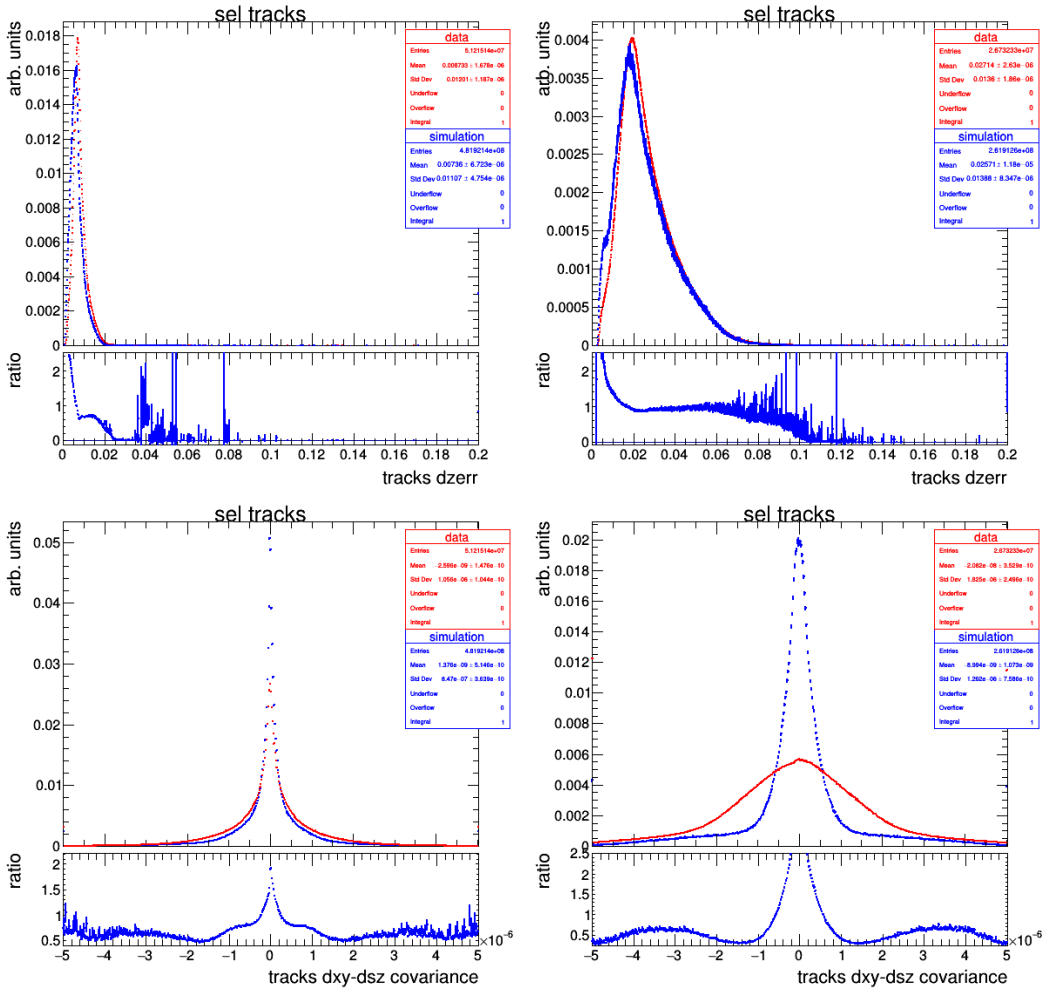

Figure B.8: Comparison of $\sigma_{d_{sz}}$ (top) and $\sigma_{d_{xy}} - \sigma_{d_{sz}}$ (bottom) for $\eta < 1.5$ (left) and $\eta > 1.5$ (right) between simulation and 2017 data. Units are cm.

To rescale for track covariance matrix elements, the same procedure can be used. The mean profile of $\sigma_{d_{sz}}$ is compared between data and MC, as shown in Fig. B.9.

Correctional fit functions will be derived from the data/simulation ratio plot for each data-taking era as shown in Figure B.10.

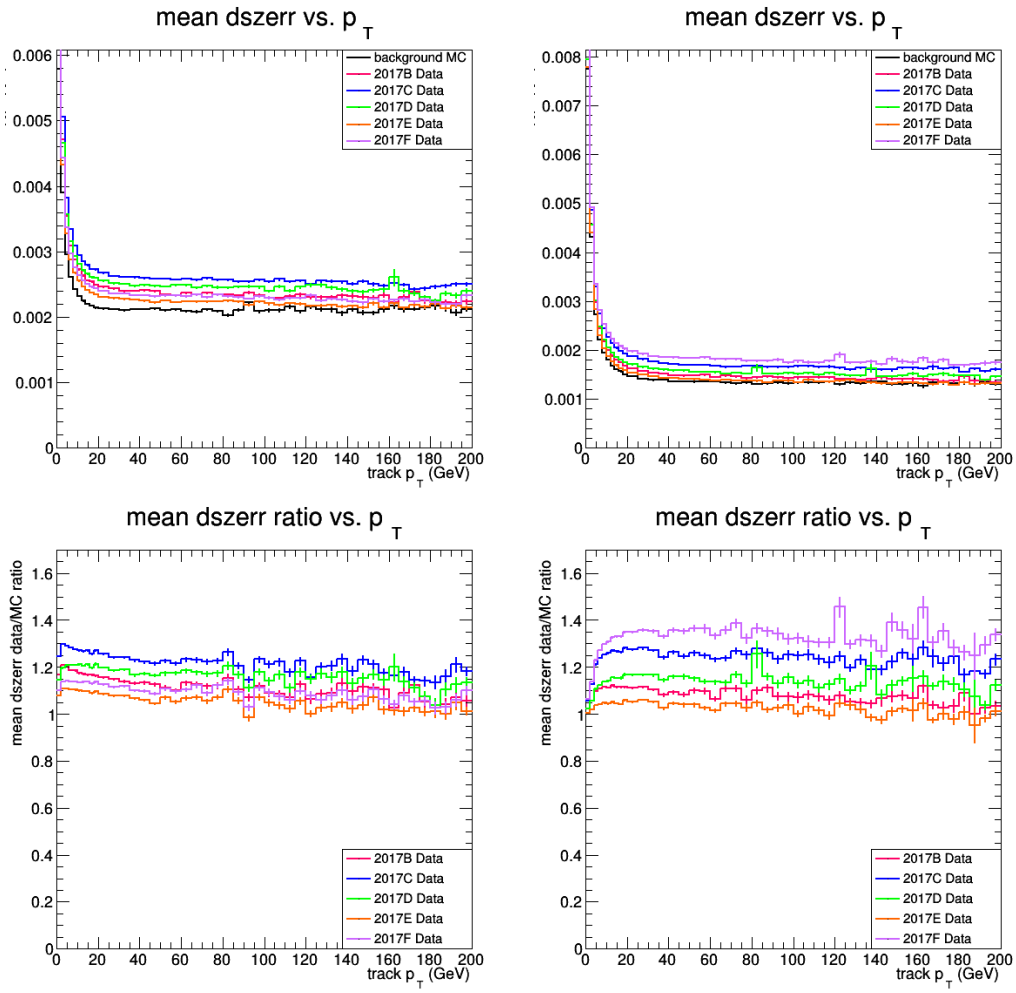

Figure B.9: Mean $\sigma_{d_{xy}}$ profiles (top) and mean $\sigma_{d_{sz}}$ data / MC ratios (bottom) for $\eta < 1.5$ (left) and $\eta > 1.5$ (right) in 2017.

The $\sigma_{d_{xy}} - \sigma_{d_{sz}}$, as the last correction we are going to make on the track impact parameter uncertainty, has been already rescaled when $\sigma_{d_{xy}}$ and $\sigma_{d_{sz}}$ are rescaled because they are correlated. The $\sigma_{d_{xy}} - \sigma_{d_{sz}}$ is further rescaled after applying rescaling on $\sigma_{d_{xy}}$ and $\sigma_{d_{sz}}$ to make the final rescaled $\sigma_{d_{xy}} - \sigma_{d_{sz}}$ similar in data and simulation. The same procedure is applied, Fig. B.11 shows the mean profile of $\sigma_{d_{xy}} - \sigma_{d_{sz}}$ and

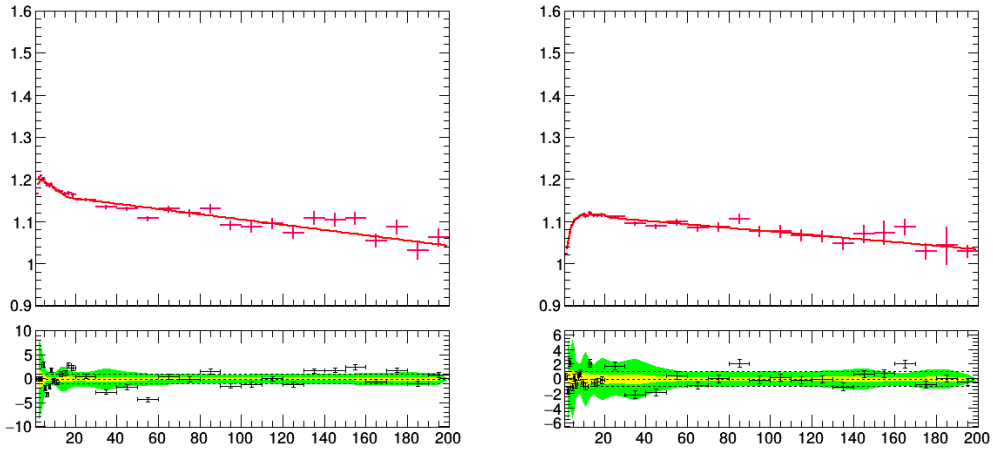

Figure B.10: Fit applied to the 2017B era data/simulation $\sigma_{d_{sz}}$ ratio vs. p_{T} curve for $|\eta| < 1.5$ (left) and $|\eta| > 1.5$ (right). The x-axis shows track p_{T} in GeV.

the data/simulation ratio plot. Figure B.12 shows the correctional fit function for 2017B data.

In conclusion, the track covariance matrix elements, especially $\sigma_{d_{xy}}$, $\sigma_{d_{sz}}$, and $\sigma_{d_{xy}} - \sigma_{d_{sz}}$ distributions are different between data and simulation. To correct for this, the mean profile of those elements as a function of p_{T} is made and a function is fitted on the data/simulation ratio plot of the profile plots. With the fitted functions, the matrix elements can be rescaled to make the simulation describe data better.

The track rescaling is applied to both signal and background simulation. Tracks are rescaled before they are used in this search so tracks used in track selection, vertex reconstruction, and ML are all rescaled.

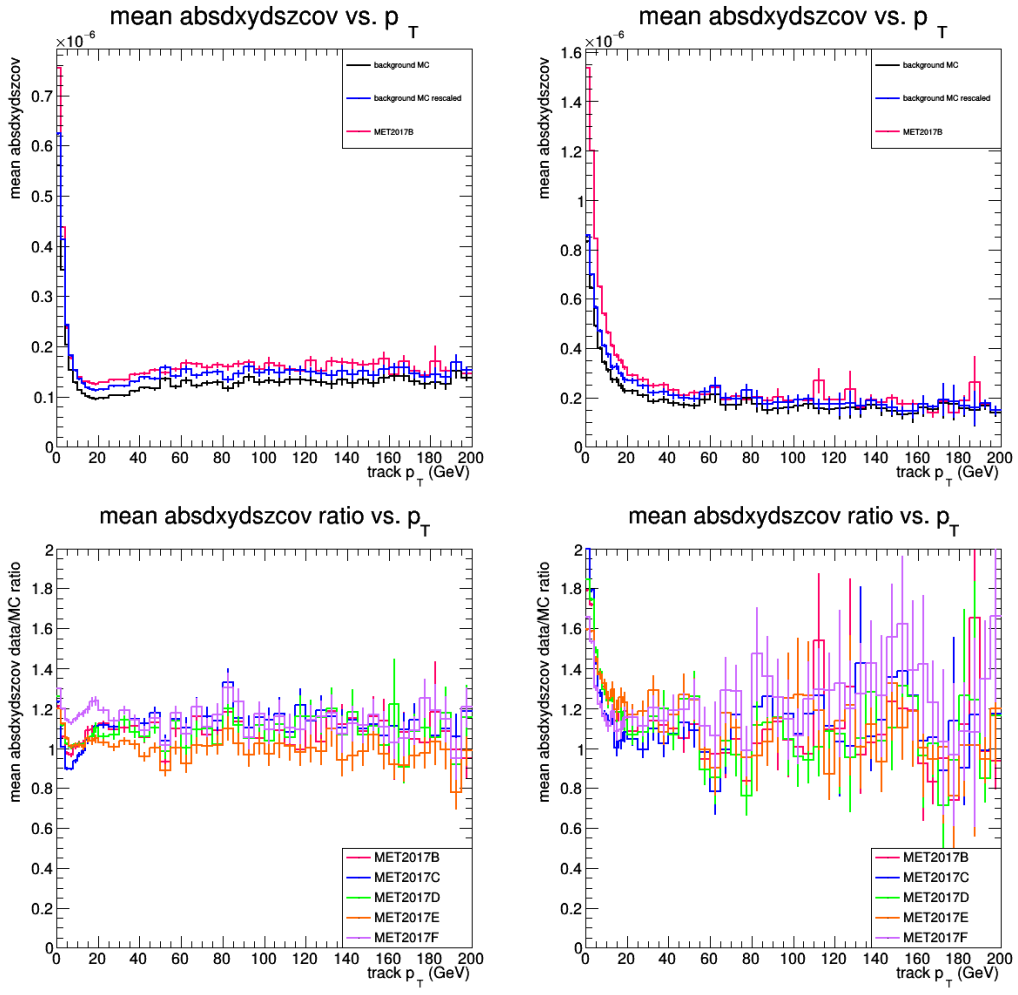

Figure B.11: Mean $\sigma_{d_{xy}}$ profiles (top) and mean $\sigma_{d_{xy}} - \sigma_{d_{sz}}$ data / MC ratios (bottom) for $\eta < 1.5$ (left) and $\eta > 1.5$ (right) in 2017B.

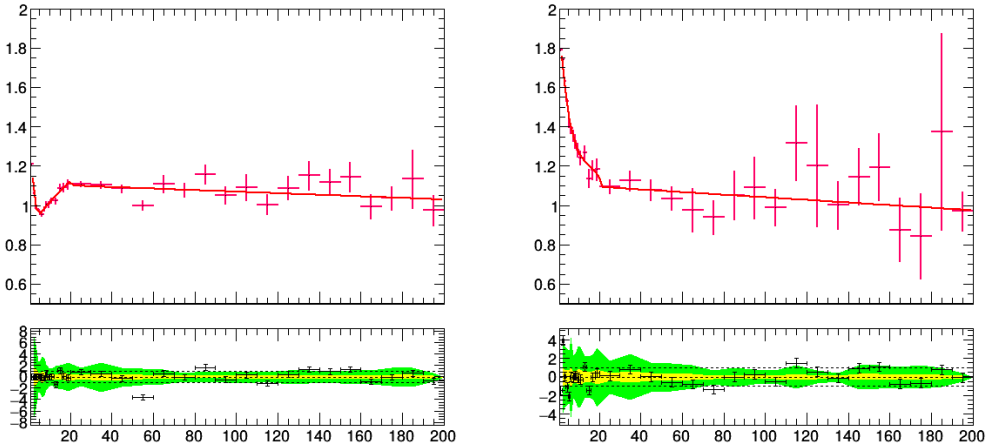

Figure B.12: Fit applied to the 2017B era data/simulation $\sigma_{d_{xy}} - \sigma_{d_{sz}}$ ratio vs. p_{T} curve for $|\eta| < 1.5$ (left) and $|\eta| > 1.5$ (right).

B.2 Vertex Variables

We compared data and MC in control regions of data taken in 2016, 2017, and 2018. The number of events with 3-track vertices that satisfy vertex selection described in Ch. 7 predicted by simulation and observed in data is compared as summarized in Tab. B.1. To quantify the agreement of event yield in simulation and data, the ratio of the prediction in MC and observation in data is calculated. We expect the simulation/data ratio to be 1 when the simulation matches the data perfectly. According to the table, the number of events in simulation and data matches in 2016 while do not match in 2017 and 2018.

Table B.1: Number of events in 3-track CR. 100% data is reported in the table.

year	Prediction by simulation	Observation in data	Simulation / Data
2016 preVFP	565.773 ± 32.68	484.00 ± 22.00	1.17 ± 0.09
2016 postVFP	330.21 ± 21.95	344.00 ± 18.57	0.96 ± 0.08
2017	929.36 ± 53.36	2162.00 ± 46.50	0.43 ± 0.03
2018	1494.51 ± 68.04	3178.00 ± 56.37	0.47 ± 0.02

We used 10% of data evenly spread among the data-taking eras in 2016, 2017, and 2018 to study the data and MC differences. To understand the difference described above, different distributions are compared between data and simulation; comparisons reveal that the difference mainly comes from the different distribution of vertex $\sigma_{d_{BV}}$, as shown in Fig. B.13. The comparison shows that data has much more vertices with $\sigma_{d_{BV}} < 0.005\,\text{cm}$ compared with simulation. The reasons for the discrepancy might come from 2 possible sources. First, data has more low-p_{T} tracks compared with the simulation, as shown in Fig. B.14. Second, the number of tracks from tight b-jets in vertices $N_{btracks}$, calculated by looping over all tracks in a given secondary vertex and counting the number of tracks that are included in a tight b-jet, differ in simulation and data, as shown in Figure B.15. The two factors The impacts of those factors on vertices are discussed below.

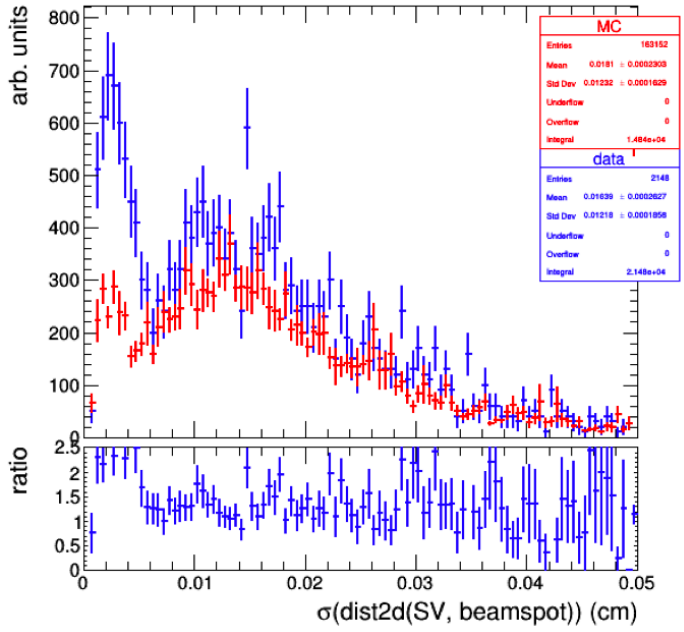

Figure B.13: Distribution of $\sigma_{d_{BV}}$, the red histogram corresponds to simulation and the blue histogram corresponds to data.

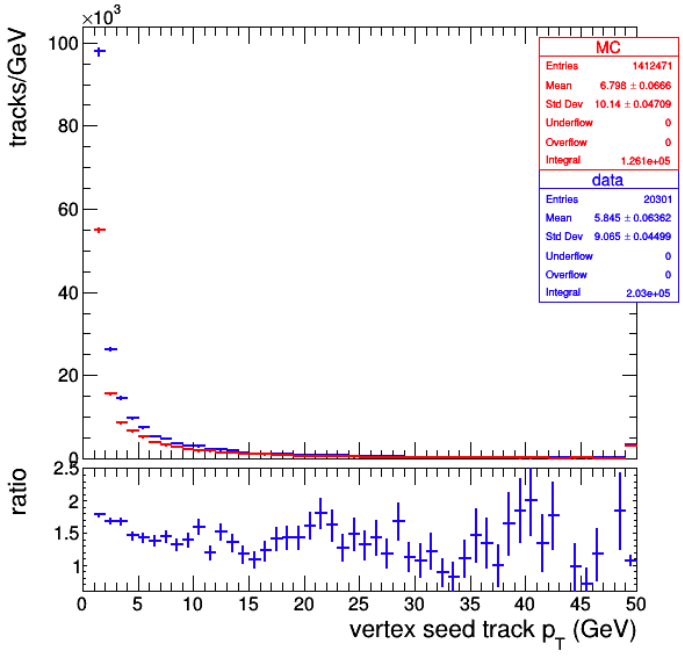

Figure B.14: Distribution of vertex seed track p_{T}, the red histogram corresponds to simulation and the blue histogram corresponds to data.

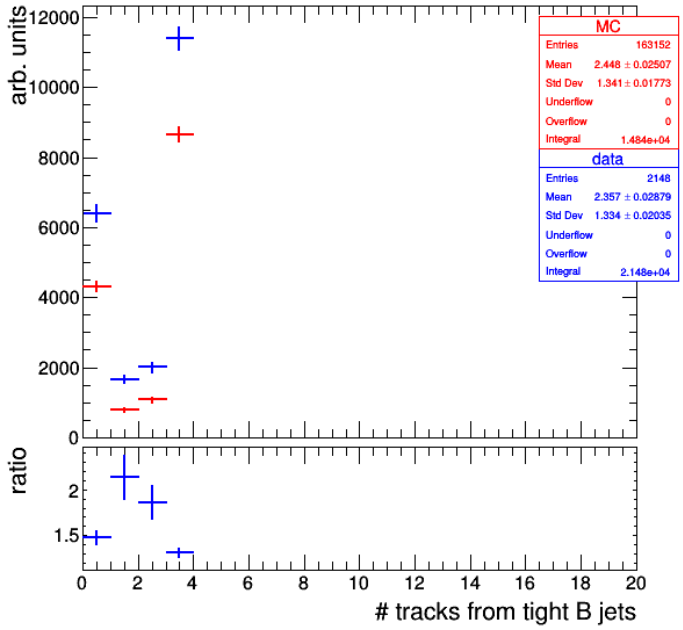

Figure B.15: Distribution of $N_{btracks}$, the red histogram corresponds to simulation and the blue histogram corresponds to data.

B.2.1 Impact of Low-p_{T} Tracks

The first factor is that data has more low-p_{T} tracks. To understand whether low-p_{T} tracks make vertex $\sigma_{d_{BV}}$ distribution in data and simulation different, we tried removing low-p_{T} tracks in vertex reconstruction and comparing data and simulation after the reconstruction. As discussed in Ch. 7, the vertex reconstruction is performed using seed tracks with $p_{\mathrm{T}} > 1\,\mathrm{GeV}$. According to the vertex seed track p_{T} distribution in Fig. B.14, data and simulation are similar when tracks have $p_{\mathrm{T}} > 5\,\mathrm{GeV}$. As a result, tracks with $p_{\mathrm{T}} < 5\,\mathrm{GeV}$ are temporarily removed in vertex reconstruction to study the effect of low-p_{T} tracks. Figure B.16 compares the $\sigma_{d_{BV}}$ distribution after the vertex seed track p_{T} selection is tightened in vertex reconstruction. As shown in the plot, the difference between data and simulation at $\sigma_{d_{BV}} < 0.005\,\mathrm{cm}$ is still obvious. Also, the number of events in the simulation/data ratio is 0.43, which is still far from the unity. These comparisons show that tightening the seed track p_{T} selection does not mitigate the difference observed between data and simulation.

In the meanwhile, comparing the $\sigma_{d_{BV}}$ distribution before and after tightening the seed track selection, as shown in Fig. B.17, shows that the shape of $\sigma_{d_{BV}}$ changes in the way that the first peak shifts to the left and the second peak shift to the right after the seed track selection is tightened. As discussed in Sec. 10.1, the two peaks in $\sigma_{d_{BV}}$ distribution correspond to two different types of background vertices: the peak at smaller $\sigma_{d_{BV}}$ is dominated by vertices from track random crossing while the peak at larger $\sigma_{d_{BV}}$ is dominated by vertices from B meson decays. The shift of peaks can then be interpreted as that after the seed track p_{T} criteria is tightened, the $\sigma_{d_{BV}}$ of track random crossing vertices become smaller while that of vertices from B meson decays become larger. To understand this effect, we study the relation between low-p_{T} track and different types of vertices one by one. We define vertices with $\sigma_{d_{BV}} < 25\,\mu\mathrm{m}$ as

track random crossing vertices and vertices with $\sigma_{d_{BV}} > 25\,\mu$m as B decay vertices.

The study is done using the original reconstructed vertices with seed track $p_{\mathrm{T}} > 1\,\mathrm{GeV}$ selection. To quantify the relation of vertex track p_{T} and $\sigma_{d_{BV}}$, profiles are made to show how $\sigma_{d_{BV}}$ changes as a function of minimum track p_{T} in the vertex ($min(p_{\mathrm{T}}^{track})$) for track random crossing vertices and B meson decay vertices in Fig. B.18. The left plot corresponds to track random crossing vertices and it shows that as the track p_{T} in the vertex gets larger, the $\sigma_{d_{BV}}$ gets smaller. The right plot corresponds to B decay vertices and it shows that as the track p_{T} in the vertex gets larger, the $\sigma_{d_{BV}}$ gets larger. The profile plots are consistent with what the observation described above.

To investigate the reason for such behaviors, vertices are divided into two categories:

- Soft vertices: vertices with $min(p_{\mathrm{T}}^{track}) < 5\,\mathrm{GeV}$
- Hard vertices: vertices with $min(p_{\mathrm{T}}^{track}) \geq 5\,\mathrm{GeV}$

For track random crossing vertices, since softer tracks usually have larger uncertainties and tracks are assigned to vertices by the significance of the transverse distance between vertex and tracks, larger track uncertainty makes a track be assigned to a vertex that has a relatively large transverse distance to the track. That will make the uncertainty of the vertex larger. This can be proved by performing an "n-1" refit to vertices and checking how the transverse position of those vertices changes. The "n-1" refit is done by iterating over each track in the vertex and refitting the vertex with the given track removed. If the explanation is true, soft vertices should shift more than hard vertices after the "n-1" refit. Figure B.19 shows the maximum shift of transverse position during the "n-1" refit for vertices. In the plot, the soft vertices shift more than the hard vertices, which proves the correctness of the hypothesis.

For B decay vertices, soft vertices have lower p_{T} and thus are generally less boosted. That makes tracks from soft B decay have larger open angles, resulting in smaller $\sigma_{d_{BV}}$. This can be proved by examining generator level b quark p_{T} and the $\Delta\phi$ between track in B decay vertices. If the hypothesis is true, the generator level b quark p_{T} should be smaller for soft vertices and $\Delta\phi$ between pairs of tracks should be larger for soft vertices. Figure B.20 shows that the generator level b quark p_{T} for soft vertices are smaller than hard vertices and Fig. B.21 shows that tracks from soft vertices have larger open angles than hard vertices, which proves the correctness of the hypothesis.

In conclusion, removing low-p_{T} tracks in the vertex reconstruction does not solve the discrepancy, so this study does not lead to any correction in this search.

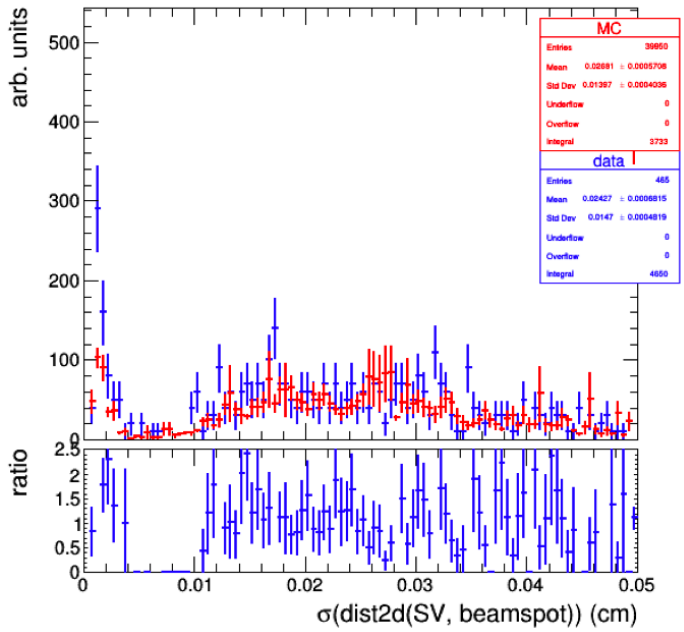

Figure B.16: Distribution of $\sigma_{d_{BV}}$ after applying seed track $p_{\mathrm{T}} > 5GeV$ selection in 2017.

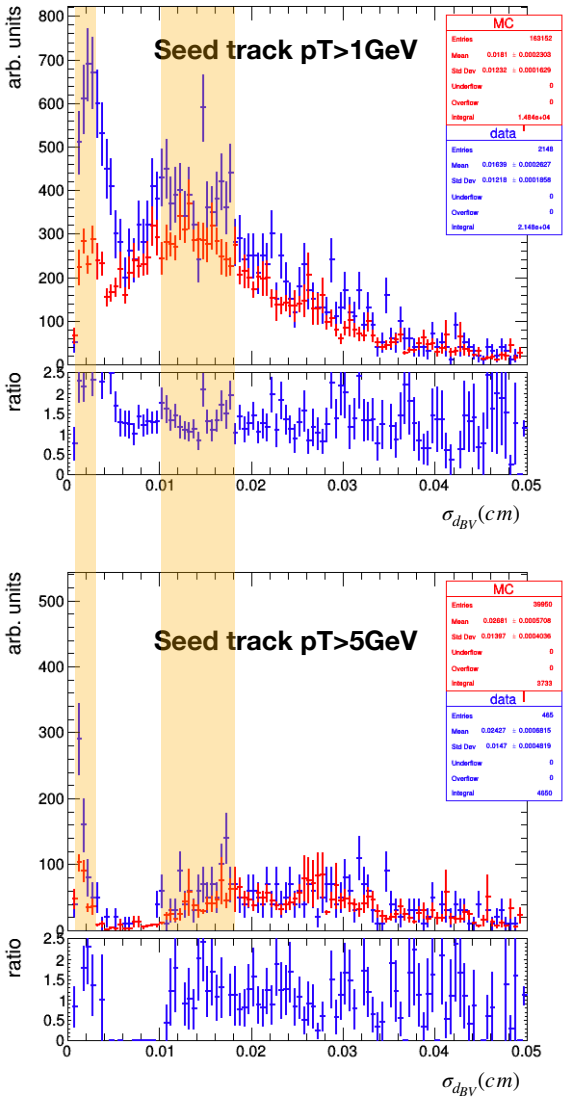

Figure B.17: Distribution of $\sigma_{d_{BV}}$ before and after applying seed track $p_{\mathrm{T}} > 5\,\mathrm{GeV}$ selection in 2017.

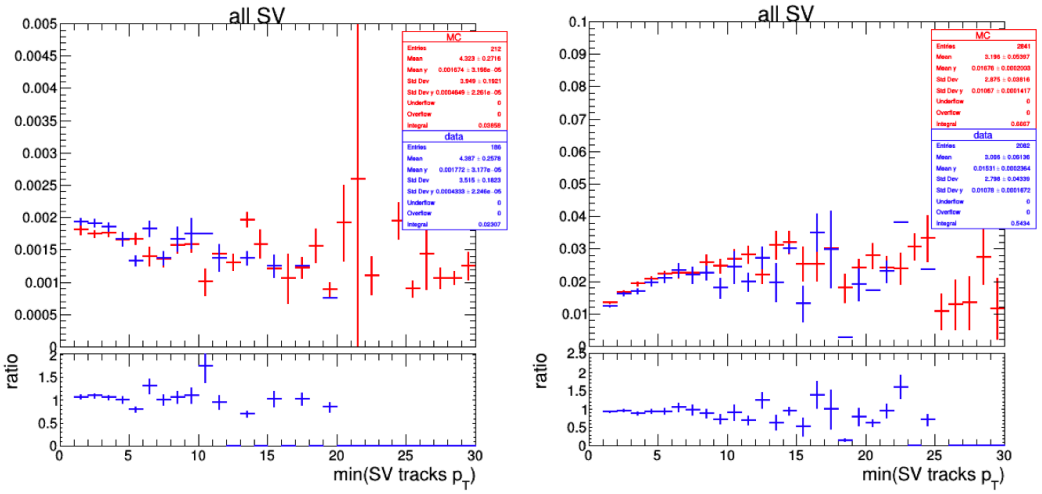

Figure B.18: Profile of $\sigma_{d_{BV}}$ vs. $min(p_{\mathrm{T}}^{track})$. The plot on the left is made with track random crossing vertices while the plot on the right is made with B decay vertices.

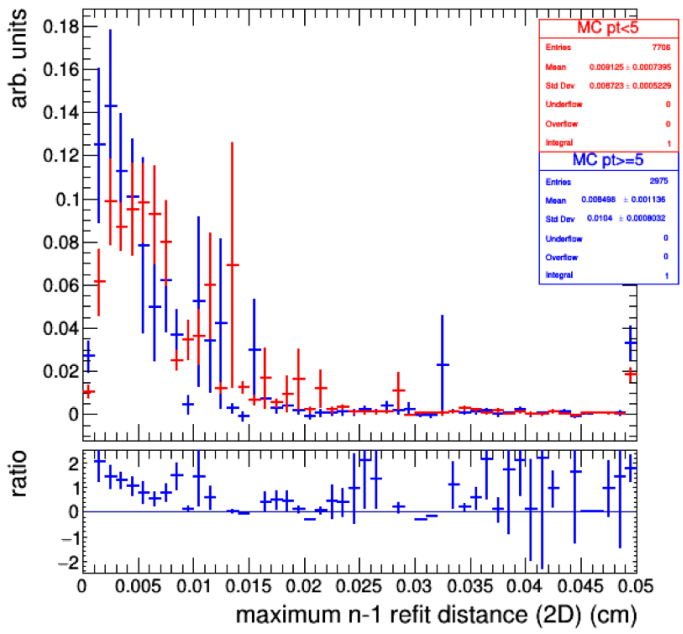

Figure B.19: Maximum of the transverse distance between "n-1" refitted vertices and the original vertex for track random crossing vertices. Soft vertices shift more than hard vertices after the "n-1" refit.

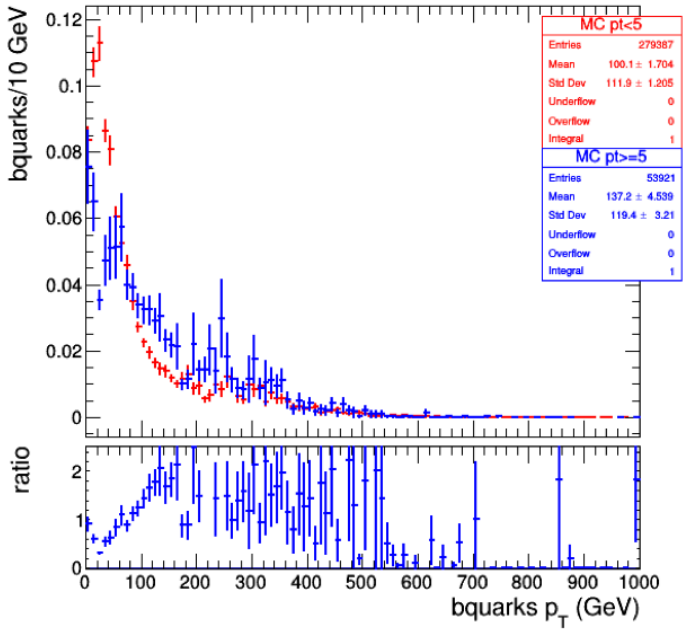

Figure B.20: Generator level b quark p_{T} for B decay vertices. Soft vertices correspond to softer b quarks, and thus are less boosted than hard vertices.

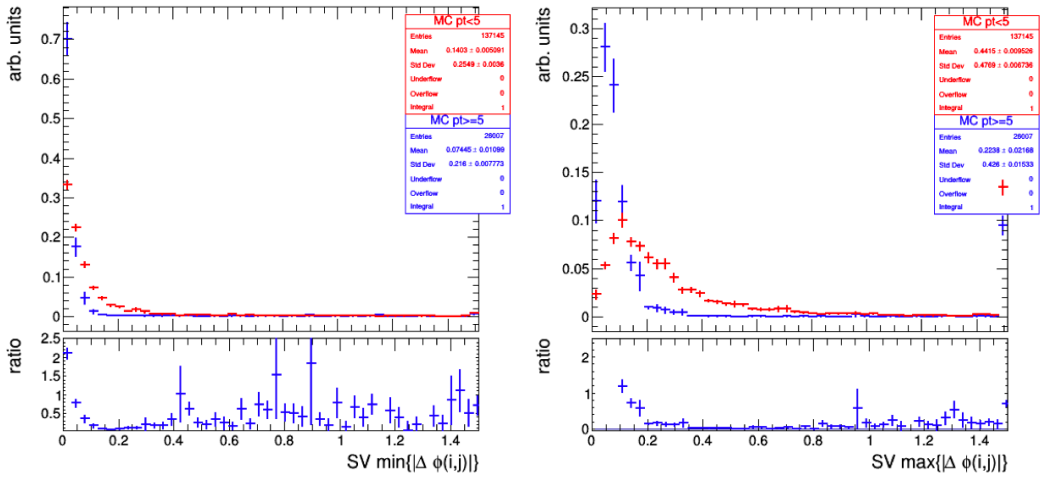

Figure B.21: Minimum (left) and maximum (right) track $\Delta\phi$ between pairs of tracks in B decay vertices. Soft vertices have larger $\Delta\phi$ between pairs of tracks, which shows that soft vertices tracks have larger opening angles than hard vertices.

B.2.2 Number of Tracks from Tight B-jets in Vertices

The second possible reason for the data vs. MC control region yield differences is that data and simulation have different $N_{btracks}$ distribution, as shown in Fig. B.22, B.23, B.24, B.25 for 2016 preVFP, 2016 postVFP, 2017, and 2018.

The plots show that data and simulation have different distributions but the $\sigma_{d_{BV}}$ selection efficiency as a function of $N_{btracks}$ are similar for data and simulation. That means the $\sigma_{d_{BV}}$ distribution can be similar for data and simulation for vertices with a given $N_{btracks}$. This motivates applying a correction on $N_{btracks}$ to make the distribution in data and simulation the same. Correcting for $N_{btracks}$ is helpful in checking whether the different number of events between data and simulation results from the different distribution of $N_{btracks}$.

Weights are derived from the data/simulation ratio plot of $N_{btracks}$ distribution and each vertex has a weight value. Since the correction is used to correct for the number of events, the correction needs to be at the event-level instead of the vertex-level. Since most events in this search have only one vertex, the vertex-level weights can be directly used as event-level weights. For events with more than one vertices, the vertex with the most number of tracks in the event is used to derive the weight.

After the correction is applied, the $N_{btracks}$ and $\sigma_{d_{BV}}$ distributions are compared between data and simulation, as shown in Fig. B.26. The plots show that the correction made $N_{btracks}$ distributions the same for data and simulation successfully, and the $\sigma_{d_{BV}}$ distributions in data and simulation look much closer after the correction is applied. The number of events predicted by the simulation after $N_{btracks}$ correction applied is compared with the number of events observed in data, as shown in Tab. B.2. Comparison of the simulation/data ratios between Tab. B.1 and Tab. B.2 show that

the number of events predicted by simulation are much closer to data after the $N_{btracks}$ correction.

Table B.2: Number of events in 3-track CR after $N_{btracks}$ correction. 100% data is reported in the table.

year	Prediction by simulation	Observation in data	Simulation / Data
2016 preVFP	672.43 ± 39.04	484.00 ± 22.00	1.39 ± 0.10
2016 postVFP	383.73 ± 21.95	344.00 ± 18.57	1.12 ± 0.09
2017	1742.85 ± 104.22	2162.00 ± 46.50	0.81 ± 0.05
2018	2909.57 ± 131.43	3178.00 ± 56.37	0.92 ± 0.04

This concludes that the different $\sigma_{d_{BV}}$ distribution between data and simulation mainly comes from the different $N_{btracks}$ distribution. Correcting for $N_{btracks}$ distribution in simulation solves the vast majority of the difference. However, the reason for different $N_{btracks}$ distributions between data and simulation is not fully understood and the efficiency difference is still present when there are zero tracks from tight B jets in Fig. B.22, B.23, B.24, B.25. A potential reason is the low number of events in some of our background MC samples (such as QCD with HT below 500GeV) or some modeling differences (e.g. b quark fragmentation). Also, the fact that different numbers of events are observed in 2017 and 2018 but not in 2016 implies that the difference could be caused by the different tracking detectors used in 2016 and 2017/2018. In the simulation, we tend to overestimate track d_{xy} uncertainties in 2016 and underestimate that in 2017 and 2018, as shown in Fig. C.1 - Fig. C.4 in the thesis, which show the d_{xy} uncertainty data/MC ratio is below 1 for 2016 and above 1 for 2017 and 2018. That difference might make the track selection and vertex reconstruction different.

In conclusion, it is found that correcting for $N_{btracks}$ could solve part of the discrepancy of the number of events between data and background MC in 2017 and 2018.

However, the discrepancy corresponds to the reconstruction of background vertices, which is mainly composed of random track crossings or tracks crossing b-jets. These background vertices are very different from the signal vertices. So the discrepancy in background vertices is not able to represent the modeling of reconstructing signal vertices. The modeling and efficiency of reconstructing signal vertices are studied by the artificial displaced vertex study, as described in Sec. 11.1.2. Besides the fact that the discrepancy is not relevant to the signal efficiency of this search, the background estimation is also not affected because the background estimation method is purely data-driven. As a result, this study does not lead to any correction.

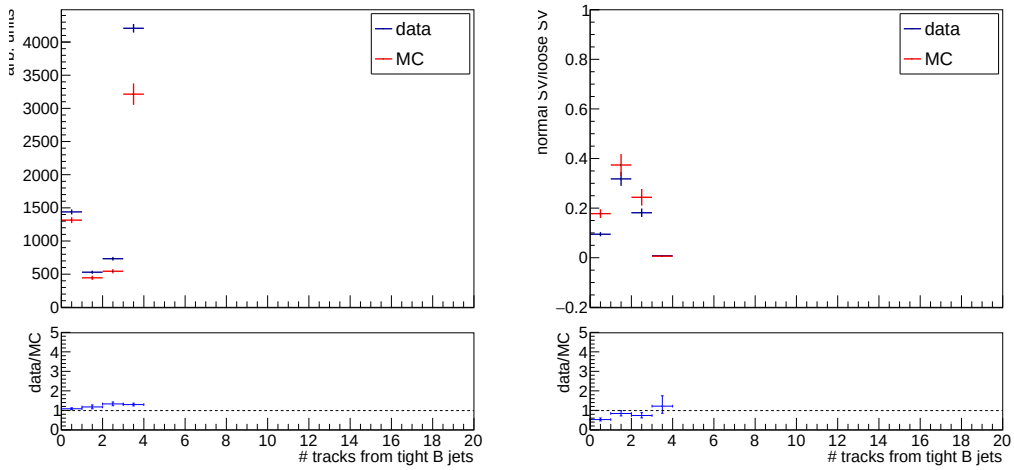

Figure B.22: Plot on the left is $N_{btracks}$ distribution and plot on the right is the $\sigma_{d_{BV}}$ selection efficiency as a function of $N_{btracks}$ for 2016 preVFP.

B.3 Machine Learning

The MLScore distribution is compared between all data and background in 2016 preVFP, 2016 postVFP, 2017, and 2018 in 3-track vertex control regions and 4-track validation region, as shown in Fig. B.27 and B.28. The comparisons show consistent

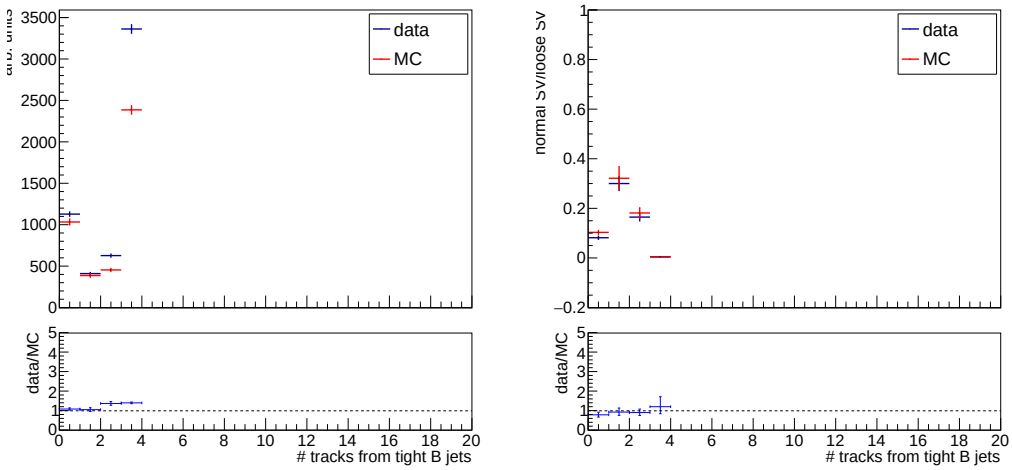

Figure B.23: Plot on the left is $N_{btracks}$ distribution and plot on the right is the $\sigma_{d_{BV}}$ selection efficiency as a function of $N_{btracks}$ for 2016 postVFP.

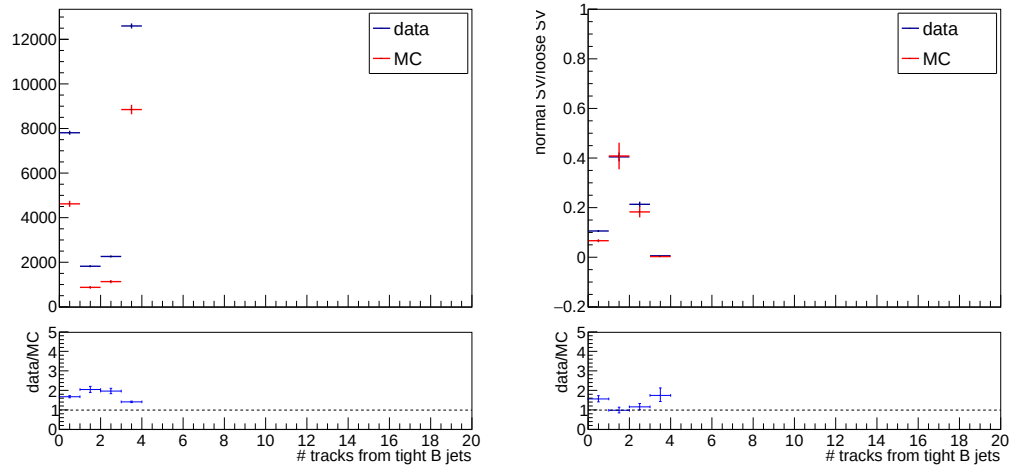

Figure B.24: Plot on the left is $N_{btracks}$ distribution and plot on the right is the $\sigma_{d_{BV}}$ selection efficiency as a function of $N_{btracks}$ for 2017.

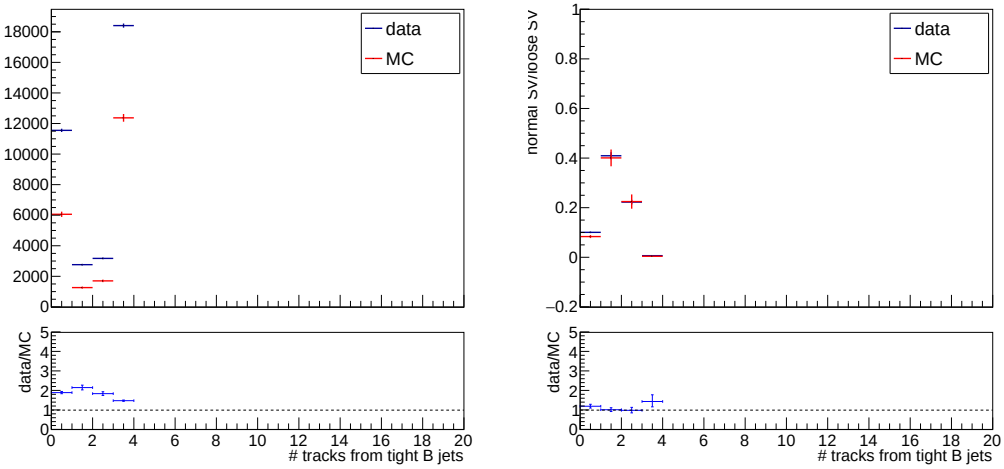

Figure B.25: Plot on the left is $N_{btracks}$ distribution and plot on the right is the $\sigma_{d_{BV}}$ selection efficiency as a function of $N_{btracks}$ for 2018.

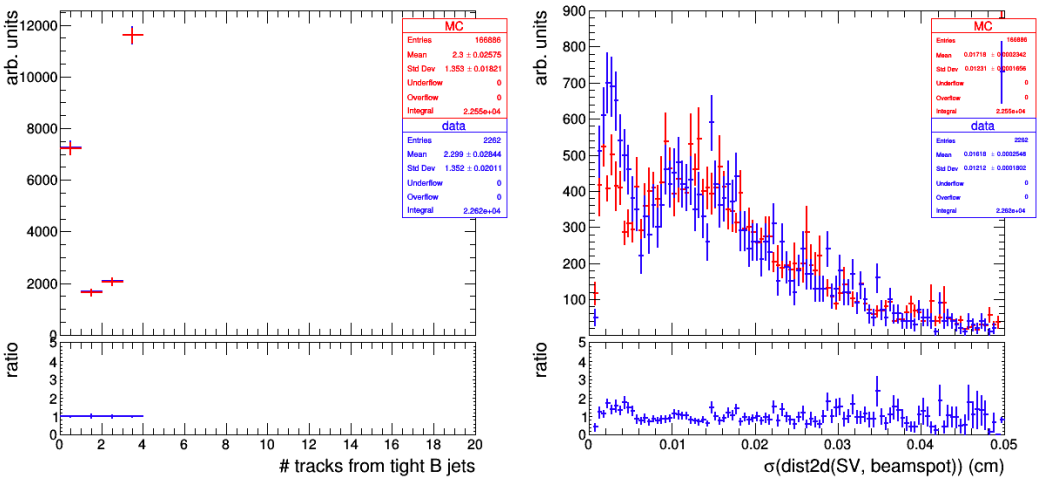

Figure B.26: Plot on the left is $N_{btracks}$ distribution after the correction is applied and plot on the right is the $\sigma_{d_{BV}}$ distribution after the correction applied.

behavior between background simulation and data.

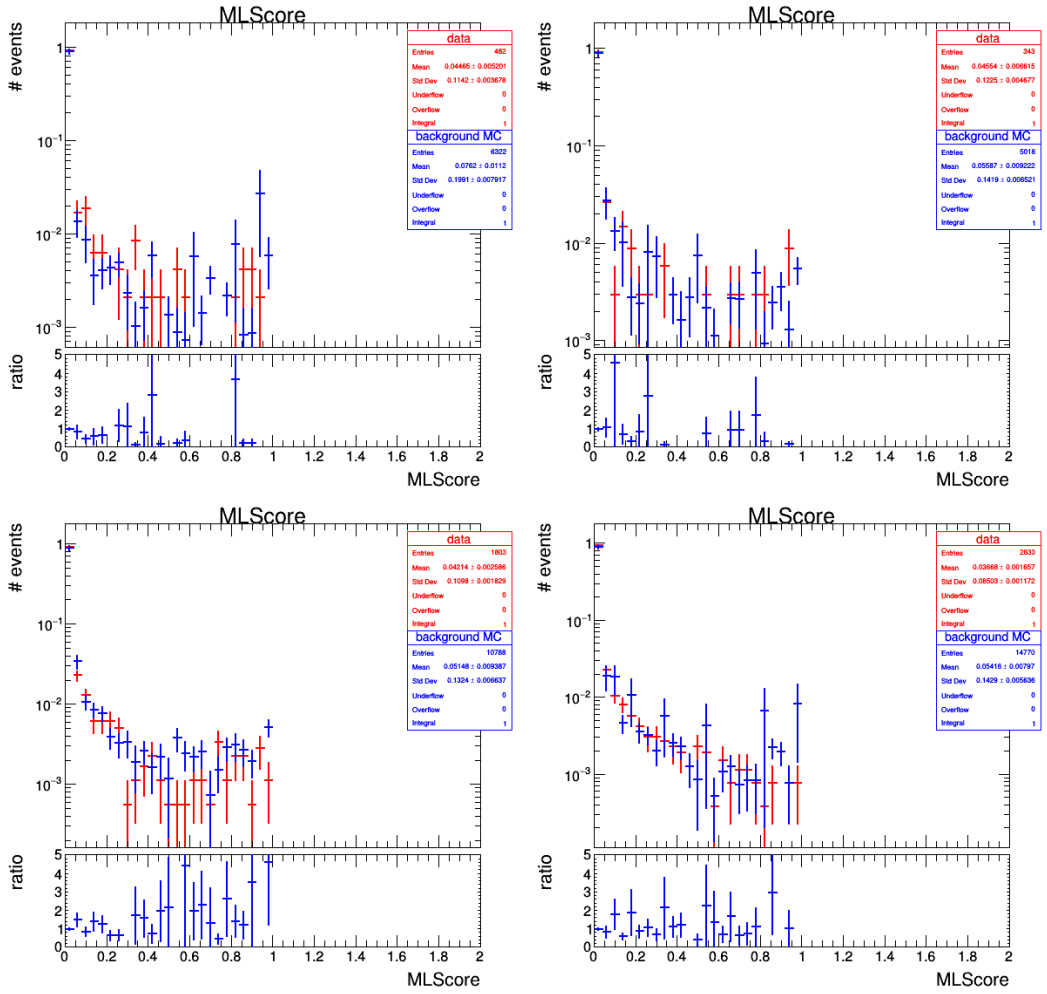

Figure B.27: MLScore distribution comparison between background simulation and data in 3-track control regions. Histograms are all normalized to 1.
(Top left) 2016 preVFP.
(Top right) 2016 postVFP.
(Bottom left) 2017.
(Bottom right) 2018.

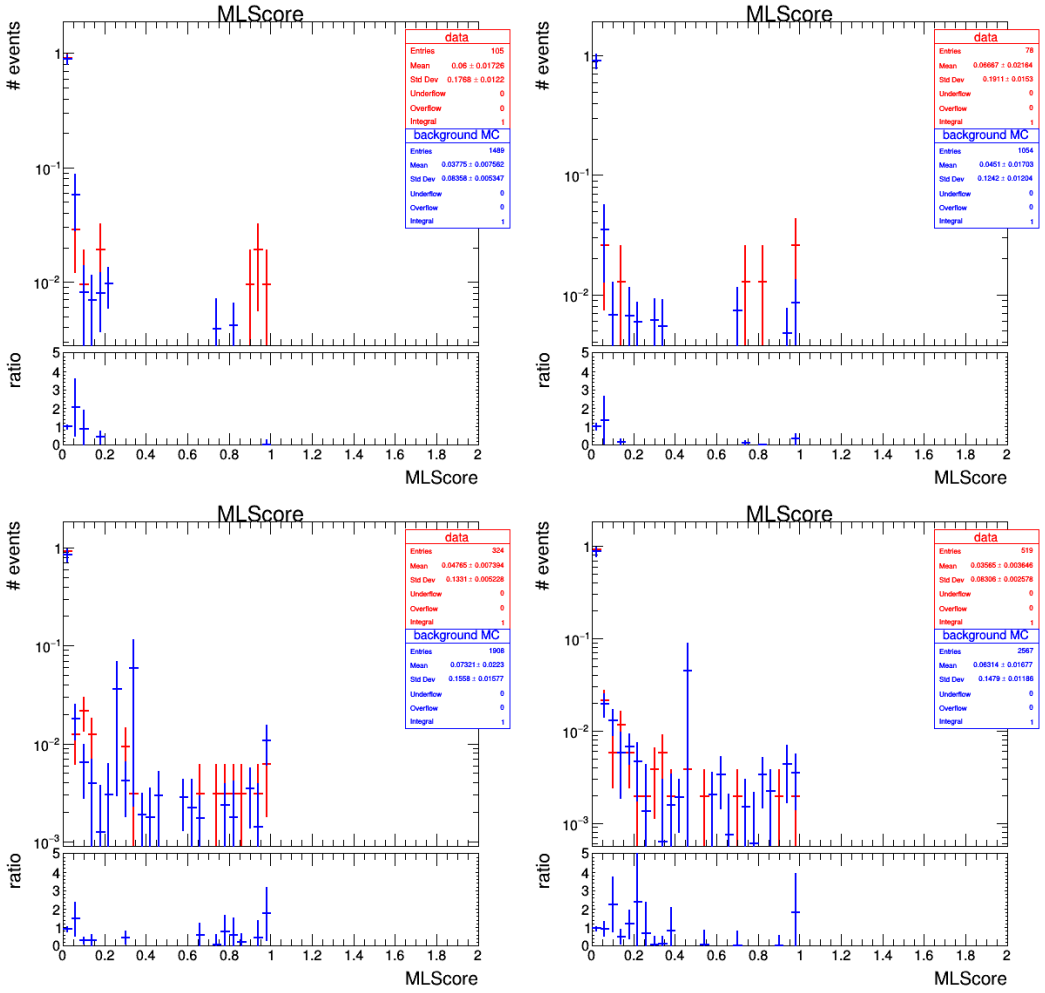

Figure B.28: MLScore distribution comparison between background simulation and data in 4-track validation regions. Histograms are all normalized to 1. (Top left) 2016 preVFP; (Top right) 2016 postVFP; (Bottom left) 2017; (Bottom right) 2018.

Appendix C

Track Rescaling

The detailed plots are documented here. Figure C.1, C.2, C.3, C.4 shows the profile of mean $\sigma_{d_{xy}}$ as a function of track p_{T} in 2016 preVFP, 2016 postVFP, 2017 and 2018.

Figure C.1, C.2, C.3, C.4 shows the profile of mean $\sigma_{d_{sz}}$ as a function of track p_{T} in 2016 preVFP, 2016 postVFP, 2017 and 2018.

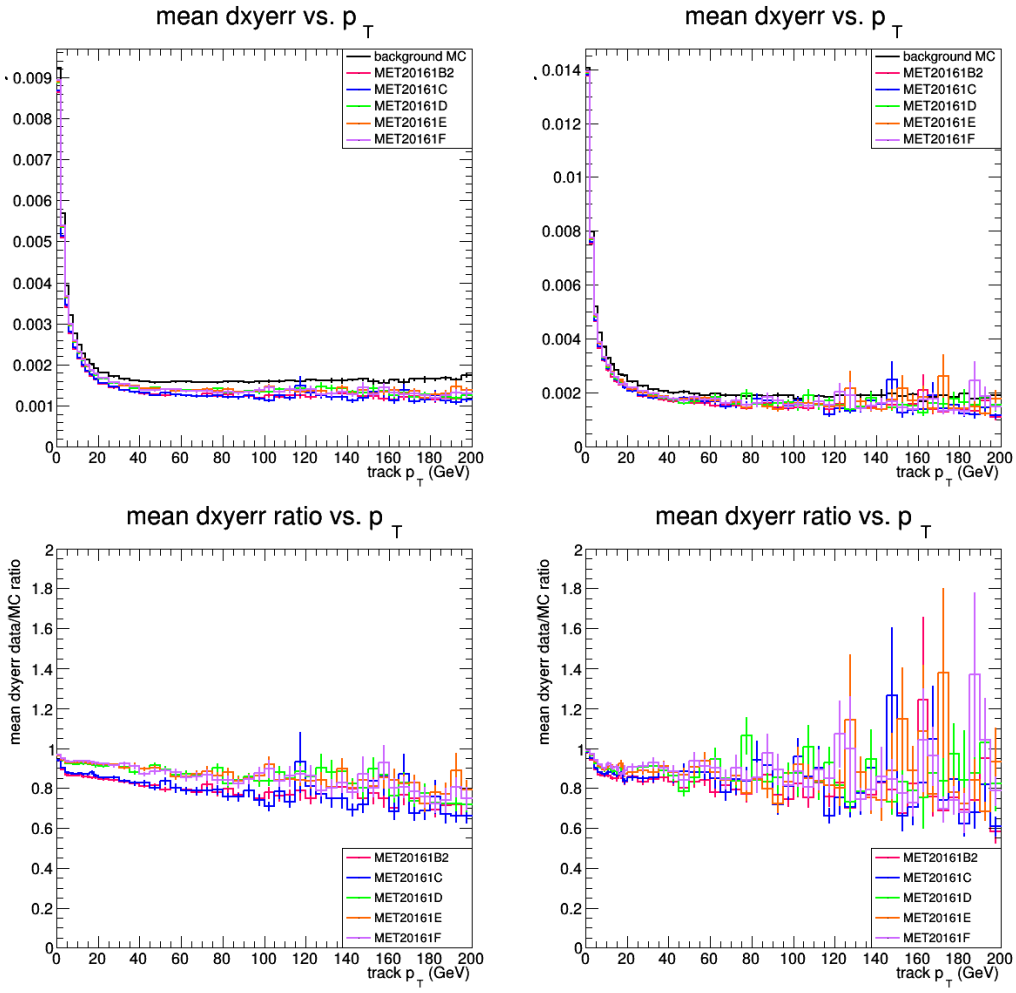

Figure C.1: Mean $\sigma_{d_{xy}}$ profiles (top) and mean $\sigma_{d_{xy}}$ data / MC ratios (bottom) for $\eta < 1.5$ (left) and $\eta > 1.5$ (right) in 2016 preVFP.

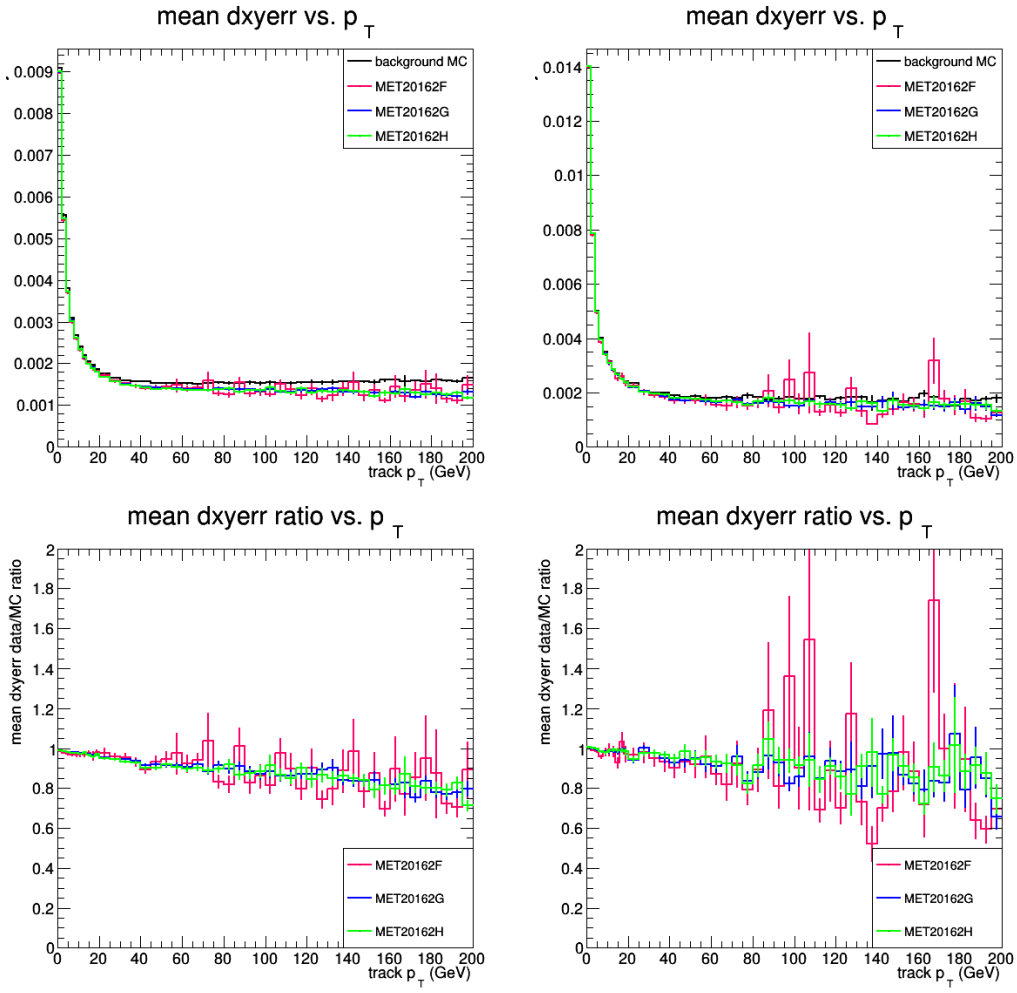

Figure C.2: Mean $\sigma_{d_{xy}}$ profiles (top) and mean $\sigma_{d_{xy}}$ data / MC ratios (bottom) for $\eta < 1.5$ (left) and $\eta > 1.5$ (right) in 2016 postVFP.

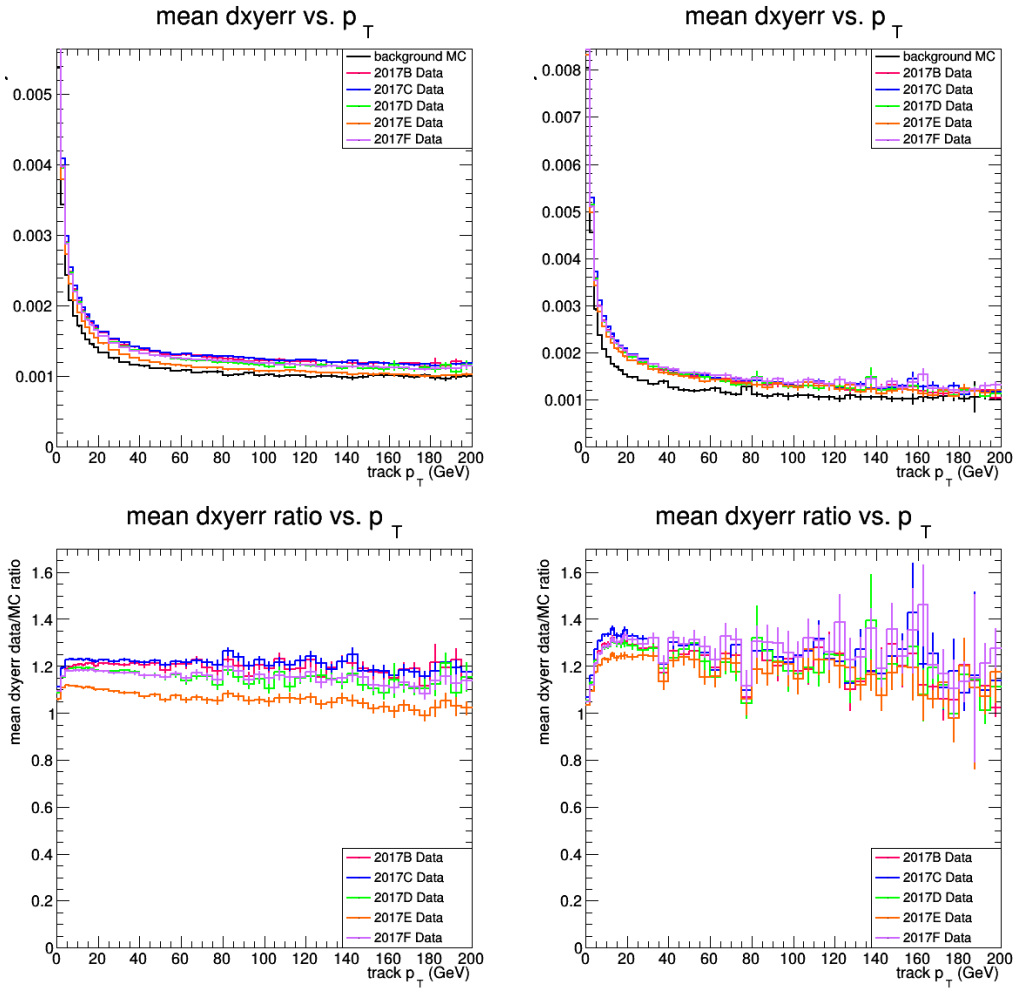

Figure C.3: Mean $\sigma_{d_{xy}}$ profiles (top) and mean $\sigma_{d_{xy}}$ data / MC ratios (bottom) for $\eta < 1.5$ (left) and $\eta > 1.5$ (right) in 2017.

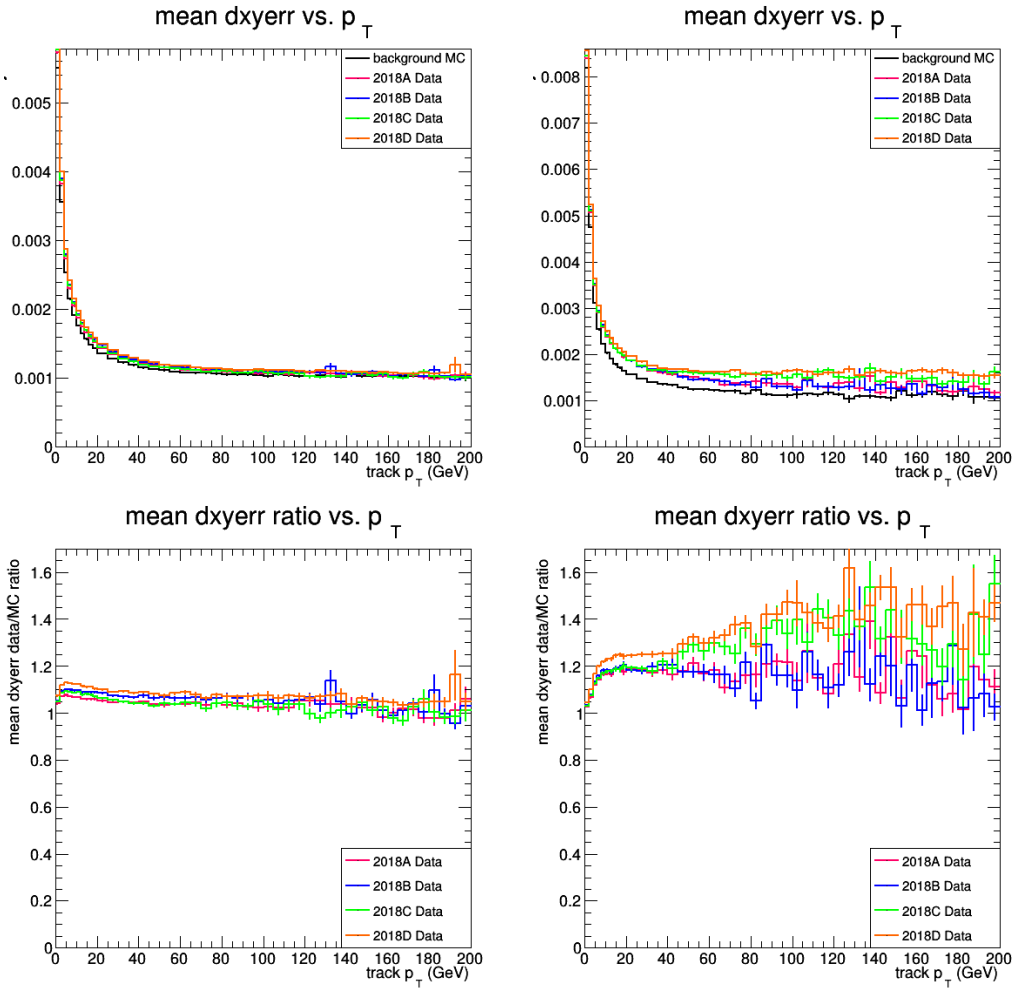

Figure C.4: Mean $\sigma_{d_{xy}}$ profiles (top) and mean $\sigma_{d_{xy}}$ data / MC ratios (bottom) for $\eta < 1.5$ (left) and $\eta > 1.5$ (right) in 2018.

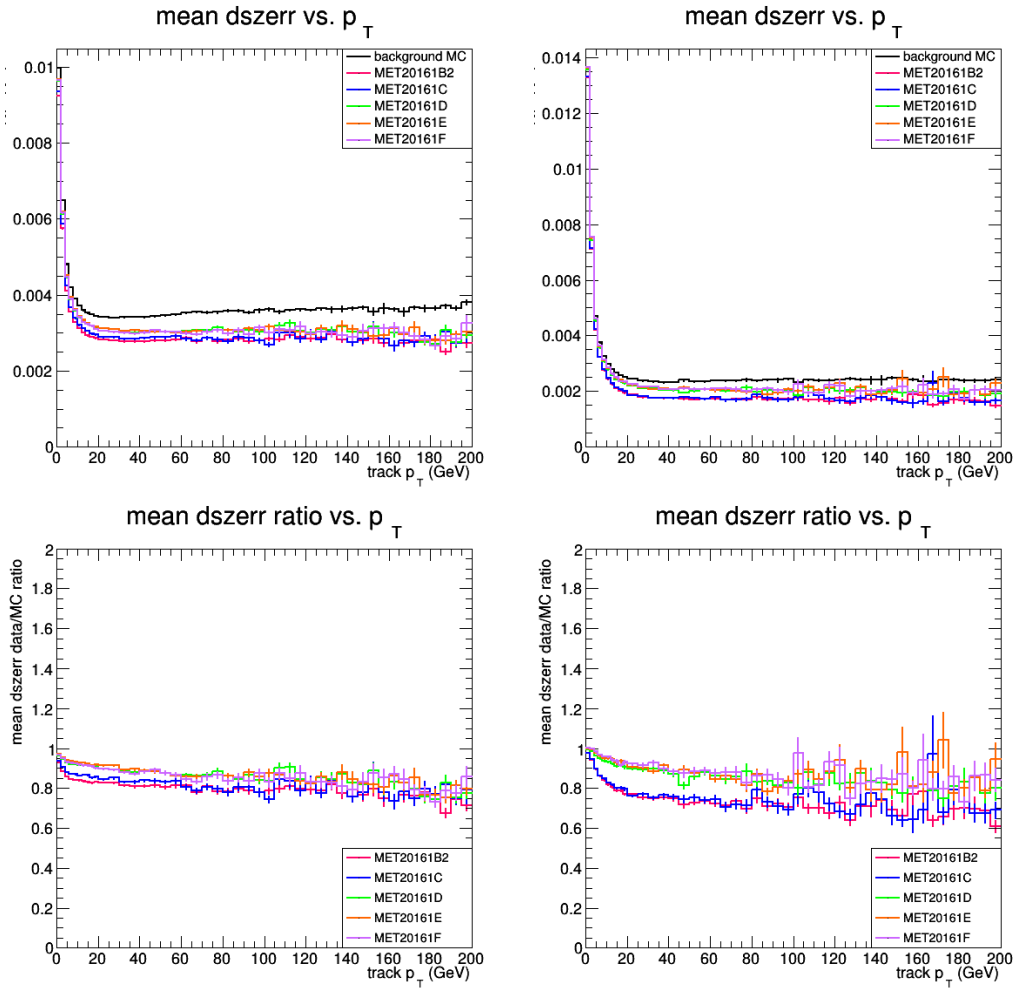

Figure C.5: Mean $\sigma_{d_{xy}}$ profiles (top) and mean $\sigma_{d_{sz}}$ data / MC ratios (bottom) for $\eta < 1.5$ (left) and $\eta > 1.5$ (right) in 2016 preVFP.

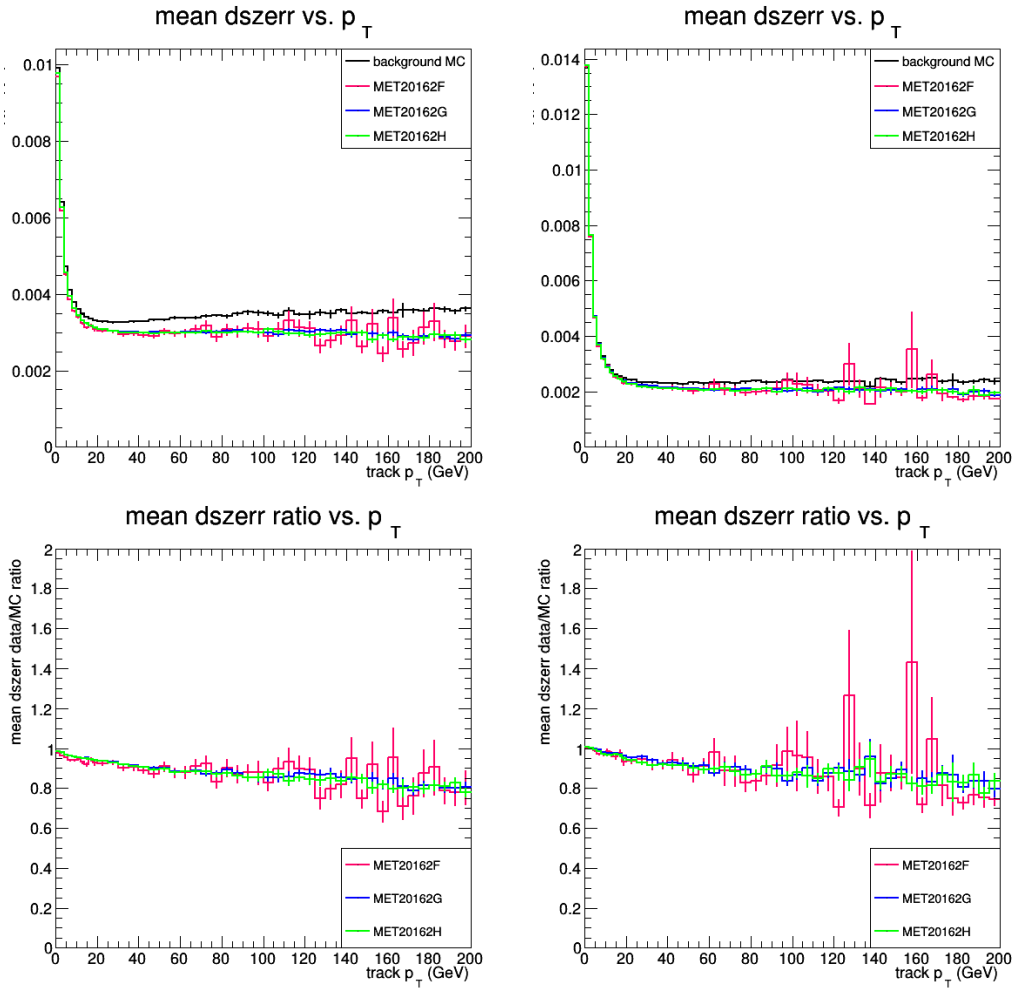

Figure C.6: Mean $\sigma_{d_{xy}}$ profiles (top) and mean $\sigma_{d_{sz}}$ data / MC ratios (bottom) for $\eta < 1.5$ (left) and $\eta > 1.5$ (right) in 2016 postVFP.

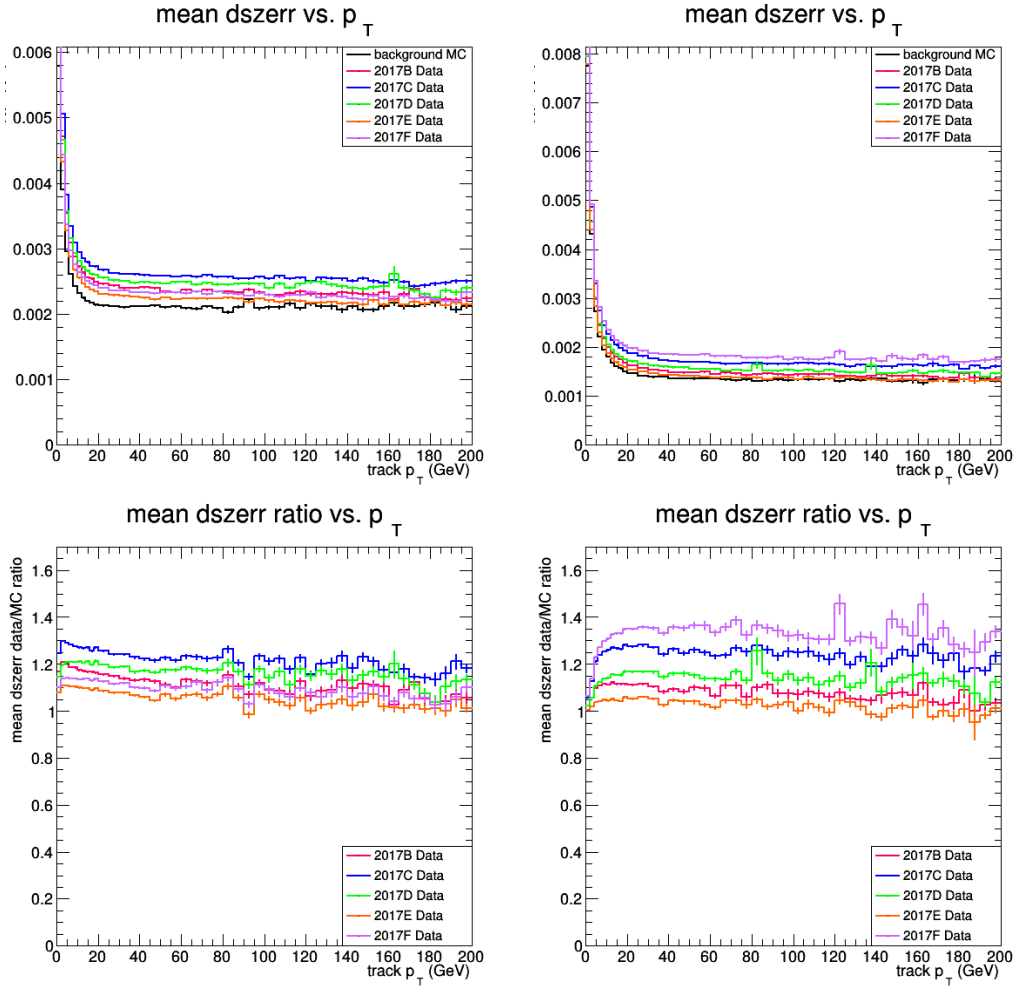

Figure C.7: Mean $\sigma_{d_{xy}}$ profiles (top) and mean $\sigma_{d_{sz}}$ data / MC ratios (bottom) for $\eta < 1.5$ (left) and $\eta > 1.5$ (right) in 2017.

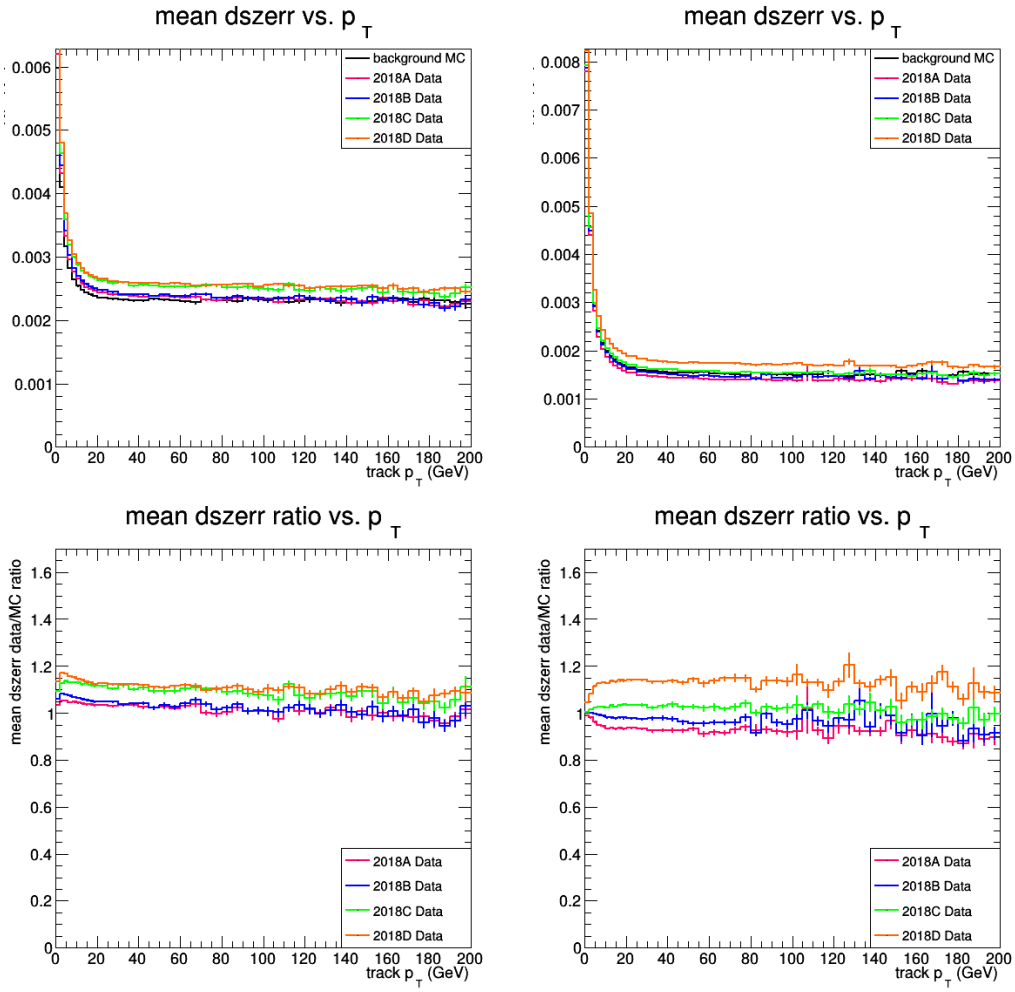

Figure C.8: Mean $\sigma_{d_{xy}}$ profiles (top) and mean $\sigma_{d_{sz}}$ data / MC ratios (bottom) for $\eta < 1.5$ (left) and $\eta > 1.5$ (right) in 2018.

Appendix D

Weights Applied on Artificial Displaced Vertices

The weights used for correcting artificial "moved" vertices to make them similar with signal samples with 10 mm $c\tau$ as 100 GeV mass splitting in 2017 are shown in Figure D.1.

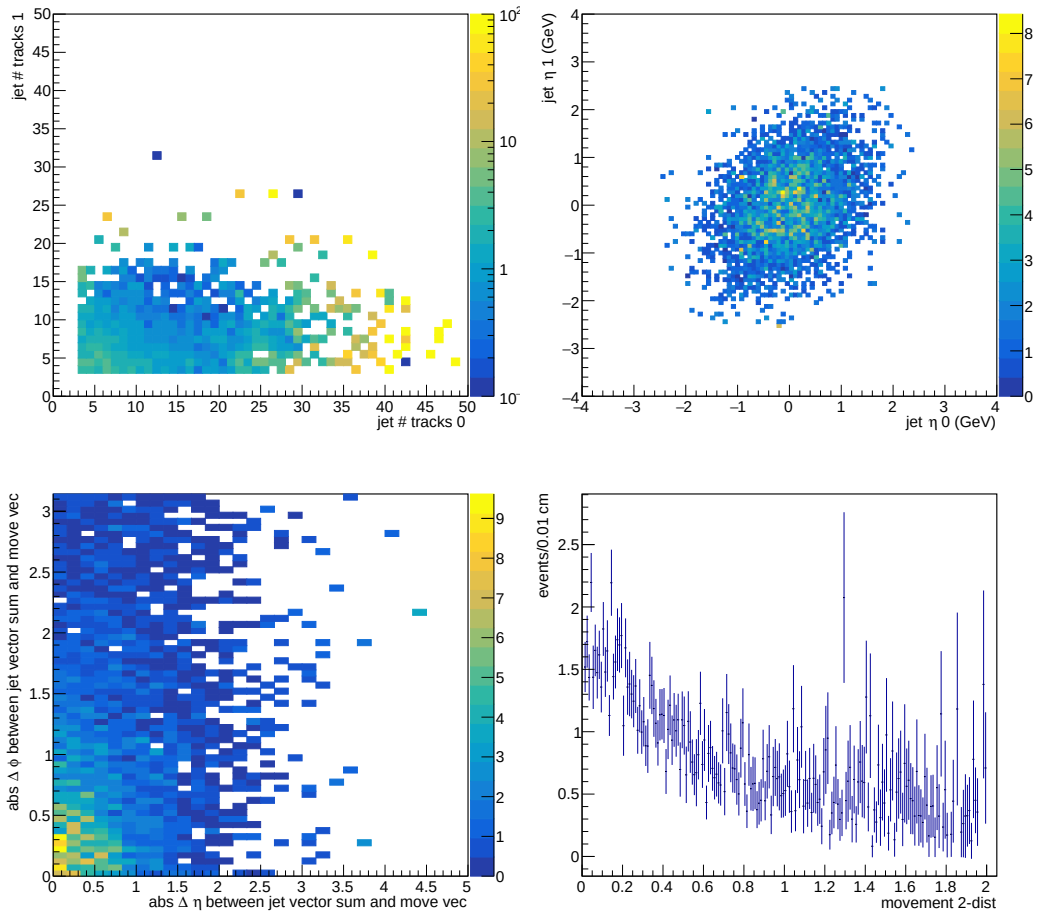

Figure D.1: The weights applied on artificial "moved" vertices that serve as mock-ups for signal samples with 10 mm $c\tau$ as 100 GeV mass splitting in 2017.