論文 / 著書情報
Article / Book Information

題目(和文)	ATLAS実験におけるレプトンフレーバー非保存崩壊τ→3μの探索
Title(English)	Search for lepton-flavor-violating decay τ→3μ at the ATLAS experiment
著者(和文)	小林大
Author(English)	Dai Kobayashi
出典(和文)	学位:博士(理学), 学位授与機関:東京工業大学, 報告番号:甲第10396号, 授与年月日:2017年3月26日, 学位の種別:課程博士, 審査員:久世 正弘,陣内 修,柴田 利明,中村 隆司,宗宮 健太郎
Citation(English)	Degree:Doctor (Science), Conferring organization: Tokyo Institute of Technology, Report number:甲第10396号, Conferred date:2017/3/26, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

Search for lepton-flavor-violating decay $\tau \rightarrow 3\mu$ at the ATLAS experiment

Dai Kobayashi

Department of Physics, Tokyo Institute of Technoloogy

March 1, 2017

Acknowledgements

First of all, I would like to show my greatest appreciation to Prof. Masahiro Kuze, my supervisor in MSc and PhD courses. He let me do what I wanted to do, and gave an appropriate help whenever I needed. This challenging research would not have been possible without such flexibility and powerful support. Not only the physics knowledge, but also I could get several wisdoms for the future; how to plan and manage the way to research, to explain my thinking clearly, to manage a group, and to extend one's scope of interests.

Prof. Olga Igonkina and Dr. Noam Tal Hod, who are the leaders of the analysis group, gave insightful comments and suggestions on my research. I owe them very much for the supports. Our sub-group was relatively small in the ATLAS collaboration, therefore, the comments and discussions from them were particularly important.

Also, I would like to express my gratitude to Dr. Masaki Ishitsuka and Prof. Kunihiro Nagano. Dr. Masaki Ishitsuka supported my research as the assistant professor in the Lab by his helpful and constructive comments. Prof. Kunihiro Nagano is one of the big boss of the trigger study in ATLAS. His support was necessary in this study, because the trigger was an important subject. They supported me since I started working for the triggers in ATLAS five years ago.

I want to thank Prof. Osamu Jinnouchi, who is the leader of the ATLAS group in Tokyo Institute of Technology. He was also my supervisor for the bachelor thesis. It was the start of my physics life, therefore, he constructed my important basis.

Discussions and cooperation with Mr. Matteo Bedognetti and Mr. Kohei Shimizu were also grateful contributions to my study. They are the students in the analysis group. I would like to show appreciation for their valuable efforts and comments.

I am deeply grateful to Mr. Kazuki Motohashi, Mr. Shota Suzuki, Mr. Takuto Kunigo, and other students in the ATLAS Japan community, and other members in the Lab, involving non-ATLAS members. The first three members lived in the same house when I stayed around CERN in Switzerland. They helped and encouraged me not only for the research, but also in the life abroad and mental health. Other staffs in the ATLAS Japan community also supported me by helpful comments from many view points, opportunity of discussion, and the financial aid to stay at CERN.

Also, the LHC, and the ATLAS experiment are achieved by the significant efforts by many many people in the collaboration. I do not forget that my thesis is built on them.

This work was supported by Japan Society for the Promotion of Science (JSPS) KAKENHI Grant Number JP00263027. This was significant because it enabled me to concentrate on the physics research.

Finally, I deeply thank family for the support and encouragement including the life before entering this institute.

Abstract

The Standard Model (SM) of particle physics explains the nature of the universe. On the other hand, shortcomings of this framework are also pointed out by several observations in cosmologies and theoretical unnatural assumptions. Several physics models beyond the Standard Model (BSM) are proposed to describe a complementary solution. These models have been investigated, but no evidences have been observed yet. Lepton-Flavor-Violating decay of the charged leptons (cLFV) is one of the important probes to search for the BSM signature because it is strictly suppressed in the SM framework. The cLFV processes in decay of muons and τ leptons have been searched for in several experiments using colliders and muon beams. The $\tau \to 3\mu$ decay is the one of the cLFV channels motivated by several models beyond the SM. No signatures have been observed in this channel, but the sensitivity is worth improving.

Large Hadron Collider (LHC) provided proton-proton collisions with the center-of-mass energy of 8 TeV and 13 TeV in 2012 and after 2015, respectively. This is the highest energy to date, and enables to search for several BSM signatures in higher energy scale. Large number of τ leptons are also produced in the proton-proton collisions. Therefore, this enables to search for not only direct production of new particles, but also rare decays of τ leptons. The ATLAS experiment is one of the experiments running in the LHC. The detector is designed to detect SM particles with large acceptance, therefore the BSM signatures are searched widely.

The search for $\tau \to 3\mu$ signature had not been considered in ATLAS before, but it is possible by using τ leptons from decay of W bosons and heavy flavored (HF) mesons. In this thesis, the first result of the search in ATLAS is described using the data with an integrated luminosity of $19.2\,\mathrm{fb}^{-1}$ in 2012. The evidence of the cLFV signature was not found in this search, and the upper limit for branching ratio of $\tau \to 3\mu$ is set to be 3.67×10^{-7}, focusing on the W-originated signature. This is not the most stringent limit, although the possibility of the search in ATLAS is established. In addition, the prospects in the future run with $100\,\mathrm{fb}^{-1}$ is described. The sensitivity should be improved by the increased number of τ lepton with the higher center-of-mass energy, luminosity and efficiency in Run 2. A new approach using HF-originated τ leptons is also considered in the future analysis. The expected sensitivities of the branching ratio to 3μ decay is estimated using data taken in 2016 to be $4.1 \times 10^{-8}(W)$ and $5.6 \times 10^{-7}(\mathrm{HF})$. These numbers are expected to be further improved by the optimization for each channel and the combination.

Contents

Part I

Introduction to $\tau \to 3\mu$ search in the ATLAS experiment

Chapter 1

Motivation and overview of the search

The frontiers of experimental technique and instruments allow us to investigate the most essential description of nature in elementary particle physics. Most of results from high energy physics experiments support the Standard Model (SM) framework of the particle physics, while there are still several mysteries which cannot be explained. In order to discover the theory complementing such shortcomings of the framework, it is necessary to search for new phenomena Beyond the Standard Model (BSM). High energy experiments can explore the evidence with both direct and indirect approaches. One of the most important probes of anomaly in SM is to search for violation of the flavor conservation. This flavor violation signature has historically led to the physics of mixing of quarks, and is also expected to lead to BSM in decay of leptons. In this chapter, theoretical and historical background are described, focusing on the flavor conservation and violation.

1.1 The Standard Model of Particle Physics

The SM framework describes particles and interactions in the universe [1, 2]. These particles are composed of fermions (Table 1.1) and bosons (Table 1.2). Fermions construct the matter, while bosons play a role in interactions and the Higgs mechanism.

Fermions are separated into quark and lepton sectors. Each sector consists of six kinds of particles which are classified by their charges and flavors. A quark has color charge of the strong interaction (also known as Quantum ChromoDynamics), in addition to charge of electroweak interaction. On the other hand, leptons do not possess color charge and hence do not interact via QCD. Nature of the particles in each sector is similar to each other, nevertheless these are recognized as different particles by the flavor. This flavor is defined as the eigennumber of each kind of particle, and invariance of the initial state and final state of the par-

Table 1.1: Fermions in the SM framework. In the row under the symbol of each particle, the mass which is measured or limited by the experiments is shown.

	generation 1st	 2nd	 3rd	charge	spin
Leptons	e 511 keV	μ 105.7 MeV	τ 1.777 GeV	-1	1/2
	ν_e < 2.2 eV	ν_μ < 0.17 MeV	ν_τ < 15.5 MeV	0	
Quarks	u ~ 2.3 MeV	c ~ 1.275 GeV	t ~ 173.07 GeV	2/3	1/2
	d ~ 4.8 MeV	s ~ 95 MeV	b ~ 4.18 GeV	-1/3	

Table 1.2: Bosons in the SM framework. In the same way as Table 1.1, the mass of each boson is also shown.

		charge	spin	Interaction(mechanism)
Gauge boson	gluon(g) 0	0	1	strong
	photon(γ) 0	0	1	electro-magnetic
	Z 91.2 GeV	0	1	weak
	W 80.4 GeV	±1	1	
Scalar boson	Higgs(h) 125.6 GeV	0	0	Higgs mechanism

ticle interaction holds except for charged current electroweak interaction (quark mixing). In the quark sector, flavor is determined for each particle, i.e. flavor of up, down, strange, charm, bottom, and top are defined. Flavors for the lepton sector are also determined for electron, muon, tau, and each corresponding neutrino. Furthermore, three generations of doublets are defined for the pair of up type-down type and charged-neutral for quarks and leptons, respectively. The quantum number of flavor is inverted by the charge conjugation transformation, hence the anti-particle has the negative number while with the same flavor.

Three types of interactions are represented by the field and bosons propagating between particles. Interactions are described as the exchange of the bosons based on the gauge theory. Therefore, gluon, photon, and weak bosons (W and Z) are called as gauge bosons, and represent QCD, electro-magnetic, and weak interactions, respectively. These bosons are introduced as the particle by quantization of the vector field, which confirms the invariance of the Lagrangian with a transformation of the free property (gauge). In other words, propagation of the force is caused by the conservation of the gauge symmetry. Higgs mechanism introduces the scaler field to express the mass of the particles by the coupling with the boson of the field, called as Higgs boson. In order to accommodate these gauge invariant interactions with fermions and the Higgs mechanism inclusively, the SM Lagrangian is defined as

$$\begin{aligned}\mathcal{L}_{SM} = &-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \\ &+ i\bar{\psi}\gamma^{\mu}D_{\mu}\psi + h.c. \\ &+ \psi_{\mathrm{i}}y_{\mathrm{ij}}\psi_{\mathrm{j}}\phi + h.c. \\ &+ |D_{\mu}\phi|^2 - V(\phi),\end{aligned} \tag{1.1}$$

where F and D are the field strength tensor and the covariant derivative for each gauge field, respectively, and γ is the Dirac matrix for the dimension. y_{ij} and $V(\phi)$ describe the Yukawa coupling matrix and the scalar potential by the Higgs mechanism. Each term of the Lagrangian describes the operation of gauge fields, fermion interactions, fermion masses by the Yukawa coupling, and Higgs field. QCD and Electro-Weak (EW) interaction are described in Section 1.1.1 and 1.1.2, respectively.

1.1.1 Quantum Chromo-Dynamics

For the QCD, color charge of quarks is related to the coupling [3–5]. Color charges, r, g, and b, are described as Dirac fields in the fundamental representation of the gauge group SU(3). Gluons also compose color octant through the exchange of the color of the quarks. Therefore, gluon also has color charges and causes self coupling of the gluons. The Lagrangian of QCD interaction ($\mathcal{L}_{\mathrm{QCD}}$) is introduced as

$$\mathcal{L}_{\mathrm{QCD}} = -\frac{1}{4}G_{\mu\nu}G^{\mu\nu} + \bar{q}(i\gamma^{\mu}D_{\mu} - m)q. \tag{1.2}$$

where q ($\bar{q}$) represents (anti-)quarks. Covariant derivative of the Dirac field (D_{μ}) and gluon field strength tensor ($G_{\mu\nu}$) are given as

$$D_{\mu} = \partial_{\mu} + ig_s\sum_{k=1}^{8}G_{k\mu}\frac{\lambda_k}{2} \tag{1.3}$$

$$G_{\mu\nu} = \partial_{\mu}G_{\nu} - \partial_{\nu}G_{\mu} + g_sG_{\mu} \times G_{\nu}, \tag{1.4}$$

where g_s is the coupling constant of the strong interaction, and λ_k are the 3×3 Hermite matrices to operate the color current. The coupling strength of the interaction is described as a function of the scale of energy transfer q^2 as

$$\begin{aligned}\alpha_s(q^2) &= \frac{\alpha_s(\mu^2)}{1+\frac{7}{4\pi}\alpha_s(\mu^2)\ln\left(\frac{q^2}{\mu^2}\right)} \\ &= \frac{1}{B\ln\left(\frac{q^2}{\Lambda}\right)},\end{aligned} \tag{1.5}$$

where μ^2 is the renormalization scale introduced to avoid divergence of α_s , and transformed to the simple form with B $(= 7/4\pi)$ and Λ $(= \mu^2 \exp(-1/B\alpha_s(\mu^2)))$.

From Equation 1.5, one finds an important feature of the interaction that the coupling strength does not decrease with lower energy i.e. the coupling strength increases at longer distance (called asymptotic freedom in QCD). Quarks and gluons cannot exist independently from other colored particles. Instead, quarks and gluons form hadrons (mesons and baryons) composed of two or three valence quarks to make the sum of the color charge zero. This process is called hadronization, which produces many hadrons in the final state from QCD interaction. Furthermore, QCD is known to be independent of the flavor of the quarks. The particle flavor is conserved in the decay and collision. Therefore the heavy flavored quarks b and c are basically generated as pair with the anti-quark, i.e. $b\bar{b}$ and $c\bar{c}$ by the gluon splitting in the proton-proton collision.

1.1.2 Electro-Weak interaction

First, Quantum ElectroDynamics (QED) is introduced to describe the electro magnetic interaction. The Lagrangian of QED is defined as

$$\mathcal{L}_\mathrm{QED} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\phi}(i\gamma^\mu D_\mu - m)\phi. \tag{1.6}$$

Covariant derivative and field strength tensor of the interaction are given as

$$D_\mu = \partial_\mu + ieA_\mu \tag{1.7}$$

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \tag{1.8}$$

where e is the electric charge and A_μ is the vector field representing electromagnetic propagation.

The coupling strength of the electro-megnetic interaction is defined in a similar form as QCD:

$$\alpha_{em}(q^2) = \frac{\alpha_{em}(\mu^2)}{1-\frac{1}{\pi}\alpha_{em}(\mu^2)\ln\left(\frac{q^2}{\mu^2}\right)}. \tag{1.9}$$

At the value of $\mu = 1\,\mathrm{GeV}$, α_{em} is approximately 1/137 and 1/128 at $q \sim m_Z$. The strength should decrease with distance in this case, unlike QCD.

On the other hand, weak interaction is mediated by the weak bosons, $W^{\pm}$ and Z^0 which represent the charged and neutral current interactions. Historically, this theory was designed to explain the small semi-leptonic decay rate of the hadrons by the relation with the mixing angle θ_c, called the Cabbibo angle. When this scheme was proposed, quark doublet was defined as pair of u and $d\cos\theta_c + s\sin\theta_c$ because only these three types of quarks were known. But in this case, the neutral current of these particles is described to be proportional to the product:

$$u\bar{u} + (d\bar{d}\cos^2\theta_c + s\bar{s}\cos^2\theta_c) + (s\bar{d} + d\bar{s})\sin\theta_c\cos\theta_c, \tag{1.10}$$

where u, d, and s are the wavefunction of each quark. This description conflicted against the fact that flavor changing neutral current (FCNC) was suppressed strongly in the observation of K meson decay. In order to solve the contradiction, charm quark c was introduced and made another pair with $s\cos\theta_c - d\sin\theta_c$. Equation 1.10 were thus replaced by the following form:

$$\begin{aligned} &u\bar{u} + c\bar{c} + (d\bar{d} + s\bar{s})\cos^2\theta_c + (s\bar{s} + d\bar{d})\sin^2\theta_c \\ &+ (s\bar{d} + d\bar{s} - d\bar{s} - s\bar{d})\sin\theta_c\cos\theta_c. \end{aligned} \tag{1.11}$$

This is called as GIM mechanism, which gave a prediction of prohibition of FCNC at the leading order interaction. Cabbibo-Kobayashi-Maskawa theory introduced the six types of quarks, and flavor eigenstates are defined as the mixing of the mass eigenstates of the particles with CKM matrix:

$$\begin{aligned} &\begin{pmatrix} u & c & t \end{pmatrix}_{flavor} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix}_{flavor} \\ &= \begin{pmatrix} u & c & t \end{pmatrix}_{mass} \cdot \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}_{CKM} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix}_{mass}, \end{aligned} \tag{1.12}$$

where $V_{qq'}$ represents the mixing parameter of q and q'. All these quarks have been discovered and their properties are well measured. The absolute value of mixing parameters are determined by several kinds of experiments [6]:

$$\begin{aligned} &\begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix}_{CKM} \\ &= \begin{pmatrix} 0.97417 \pm 0.00021 & 0.2248 \pm 0.0006 & (4.09 \pm 0.39) \times 10^{-3} \\ 0.220 \pm 0.005 & 0.995 \pm 0.016 & (40.5 \pm 1.5) \times 10^{-3} \\ (8.2 \pm 0.6) \times 10^{-3} & (40.0 \pm 2.7) \times 10^{-3} & 1.009 \pm 0.031 \end{pmatrix}. \end{aligned} \tag{1.13}$$

Thus, the flavor physics mechanism is established. This mechanism is also applied to the lepton sector, but doublets for leptons consist of a charged lepton and a corresponding neutrino. Although the mixing structure is different between quark and lepton sector, only the weak interaction can change the flavor of the fermions in the SM.

The Electroweak (EW) description is completed by the unification of these interactions, based on Weinberg-Salam theory [7–9]. This theory introduced SU(2)×U(1) symmetry, which require three bosons ($W_\mu = [W_\mu^{(1)}, W_\mu^{(2)}, W_\mu^{(3)}]$) with $I = 1$ and one boson (B_μ) with $I = 0$ arising from each symmetry. The weak iso-spin I and hypercharge Y are defined by the pair of leptons or quarks, following $Q = I_3 + Y$, where Q is the electric charge. The Lagrangian of the EW description is defined using the weak iso-spin and weak hypercharge current, denoted as J_μ and J_μ^Y:

$$\mathcal{L}_{\mathrm{EW}} = gJ_\mu \cdot W_\mu + g' J_\mu^Y B_\mu. \tag{1.14}$$

g and g' are introduced as coupling constant for each interaction via W and B intermediation, respectively. J_μ^Y represents interaction changing $Y (= Q - I_3)$, therefore can be transformed into J_μ^{em}-$J_\mu^{(3)}$ using electro-magnetic and weak isospin coupling. As a result of the mixing in the third term of the weak current, two charged bosons ($W_\mu^\pm$) and two neutral bosons (A_μ, Z_μ) are realized as

$$W_\mu^\pm = \frac{1}{\sqrt{2}} \left[W_\mu^{(1)} \pm iW_\mu^{(2)}\right] \tag{1.15}$$

$$W_\mu^{(3)} = \frac{gZ_\mu + g'A_\mu}{\sqrt{g^2 + g'^2}} \tag{1.16}$$

$$B_\mu = \frac{-g'Z_\mu + gA_\mu}{\sqrt{g^2 + g'^2}}. \tag{1.17}$$

Hence, Equation 1.14 can be transformed into

$$\begin{aligned} \mathcal{L}_{\mathrm{EW}} = & \frac{g}{\sqrt{2}}(J_\mu^- W_\mu^+ + J_\mu^+ W_\mu^-) \\ & + \frac{g}{\cos\theta_W}(J_\mu^{(3)} - \sin^2\theta_W J_\mu^{em})Z_\mu \\ & + g\sin\theta_W J_\mu^{em} A_\mu, \end{aligned} \tag{1.18}$$

where θ_W is defined as weak mixing angle by $g \sin\theta_W = e$, and $J_\mu^\pm = J_\mu^{(1)} \pm iJ_\mu^{(2)}$. This theory leads to unify the QED and Weak interactions at the high energy scale of approximately 100 GeV.

The masses of the weak bosons are not explained by the gauge interaction mechanism. Higgs mechanism [10, 11] introduced scalar field dealing with the origin of these masses by spontaneous symmetry breaking. The Higgs Lagrangian is generally defined as:

$$\mathcal{L}_{\mathrm{Higgs}} = \frac{1}{2}(\partial_\mu\phi)^2 - \frac{1}{2}\mu^2\phi^2 - \frac{1}{4}\lambda\phi^4, \tag{1.19}$$

where μ is the particle mass and λ is the coupling constant of the 4-boson vertex. In case that μ^2 is negative, the minimum value of the Higgs potential is given at two ϕ_{min} $(-v, +v)$. Therefore, Equation 1.19 is transformed by introducing the σ as $\phi = v + \sigma$:

$$\mathcal{L}_{\mathrm{Higgs}} = \frac{1}{2}(\partial_\mu \phi)^2 - \lambda v^2 \sigma^2 - \left(\lambda v \sigma^3 + \frac{1}{4}\lambda \sigma^4\right) + constant. \tag{1.20}$$

The mass term of this equation is described as

$$m = \sqrt{2\lambda v^2} = \sqrt{-2\mu^2}. \tag{1.21}$$

Thus, it was suggested that the massive weak bosons are given by the Higgs mechanism. By the discovery of the Higgs boson in 2012 [12,13], this theory was established and all particles in the SM framework have been observed.

1.1.3 Shortcomings of the SM framework

The SM is successful at describing the nature of elementary particles, although there are several issues which are not addressed. Lack of the statement for the dark matter (and energy) [14] is one of the most important issues in the SM. Existence of additional matters is pointed out by the astrophysical observations, however such particles are not included in the SM. Therefore, an alternative theory which introduces such ”missing” particles is necessary.

The hierarchy problem of the Higgs mechanism is also a major problem. In the Higgs mechanism, extremely large radiative corrections to the Higgs mass divergence is necessary, considering the observed mass of Higgs boson. This means un-natural cancelation with very high accuracy. Therefore, the reason of the fine tuning is hoped to be explained by an alternative physics model.

Furthermore, the interactions and particles should be unified at a very high energy scale close to the beginning of the universe. This is called the Grand Unification Theory (GUT). The flavor of the fermions is also unified at the energy scale, therefore, this is one of the major motivations of the search for violation of the flavor conservation.

In order to solve these problems of the SM, several BSM models have been proposed and pursued, but the signature has not yet been discovered. It is therefore important to search for the evidence with various complementary approaches.

1.2 Physics in Lepton Flavor Violation

One of the most important probes to search for evidence of BSM physics is Lepton-Flavor-Violation (LFV) signature, which is forbidden strictly in the SM as shown in Section 1.1. Although discovery of neutrino oscillation leads to LFV in neutral sector with a tiny level, LFV in the charged sector (cLFV) has not been observed yet. On the other hand, several BSM physics theories predict visible branching ratios ($\mathcal{BR}$) of cLFV decay modes. Hence, observation of a cLFV decay supports the evidence of BSM directly. In this section, new physics expecting the cLFV signature, and overviews of the searches are described.

1.2.1 Neutrino oscillation

The evidence of neutrino oscillation was discovered by water-Cherenkov experiments [?, ?, ?]. This phenomenon causes the mixing of the flavor of neutrinos, therefore leads to LFV in the neutral sector [15, 16]. In a similar manner to the quark sector, flavor eigenstates are described by the mixing of mass eigenstates:

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = U \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix}, \tag{1.22}$$

where U is the mixing matrix, called Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix [17], and $\nu_{1,2,3}$ are mass eigenstates. The flavor eigenstate can be described as dependent on time t. If the flavor was α at $t = 0$, their time evolution is determined as

$$\begin{aligned} \nu_\alpha(t) &= \cos\theta \nu_1(t) + \sin\theta \nu_2(t) \\ &= \left(\cos^2\theta e^{-iE_1 t} + \sin^2\theta e^{-iE_2 t}\right)\nu_\alpha + \sin\theta\cos\theta\left(e^{-iE_2 t} + e^{-iE_1 t}\right)\nu_\beta, \end{aligned} \tag{1.23}$$

where θ is the mixing parameter with another state assuming only two flavors for simplification. Hence, the flavor changing probability of $\nu_\alpha \to \nu_\beta$ can be expressed as

$$\mathcal{P}(\nu_\alpha \to \nu_\beta) = \sin^2 2\theta \sin^2\left(1.27\frac{\Delta m^2 L}{E}\right), \tag{1.24}$$

where L is the flight length of the neutrino in kilometers, and Δm^2 and E are the mass squared difference and energy of the neutrinos in eV^2 and GeV, respectively. The illustration of the mixing mechanism is shown in Figure 1.1. This shows the flavor of neutrinos can change if they have masses which are different each other. Therefore, it has been indirectly established that at least some neutrinos are massive and cause flavor violation by the observation. cLFV signatures is also allowed with very small order by involving the neutrino oscillation loop with higher order diagrams.

1.2.2 Motivation of the cLFV search

cLFV signature is expected by various BSM physics predictions [19, 20].

The most simplified case is the extension of the SM by heavy gauge bosons (Z' and W') with universal coupling with leptons [21]. This leads to the cLFV decay of leptons, as well as direct production of Z' decaying to different flavor lepton

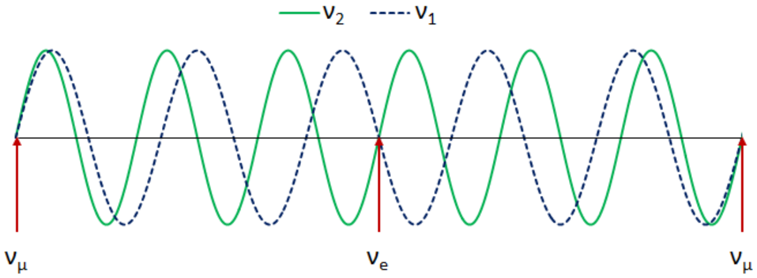


Figure 1.1: Illustration of the oscillation of the neutrino flavor by mixing of mass eigenstates [18].

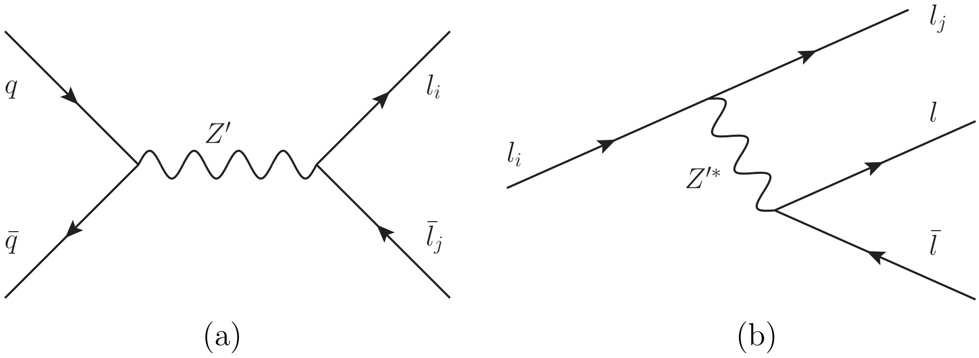
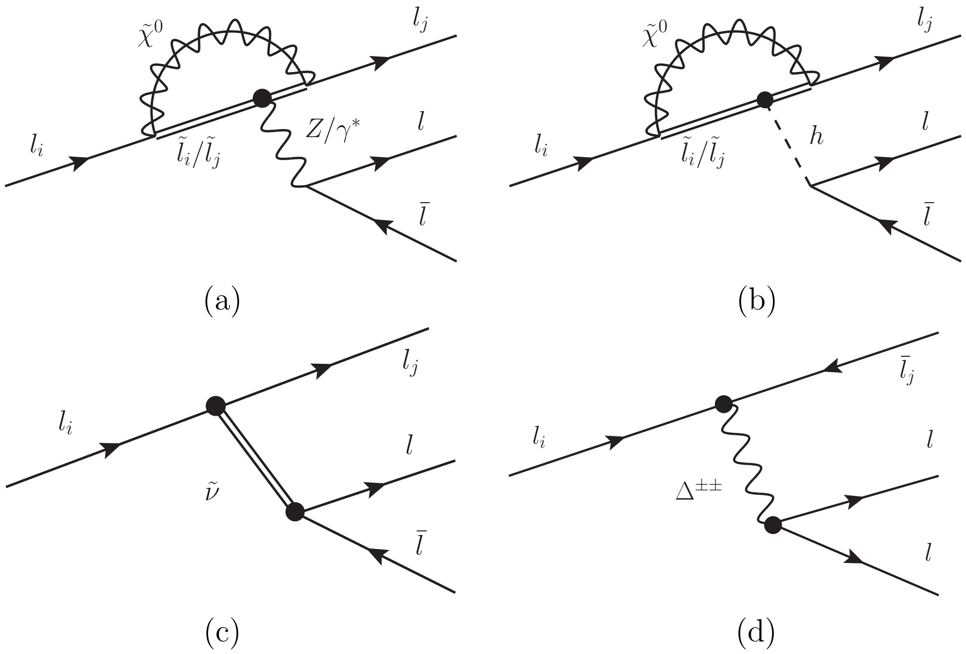


Figure 1.2: (a)Direct and (b)indirect processes by introducing Z'.

pair. Figure 1.2 shows these processes. This extension is introduced by several theoretical motivations; the Left-Right (LR) symmetric, Extra-Dimension, as well as $U(1)'$ group as a simple scale extension of the weak interaction in the SM. The LR symmetric models [22] are simple extensions of the SM. These models are expected to explain the tiny neutrino masses and parity violation of the weak interaction in the SM by introducing $SU(2)_R \times U(1)_{BL}$ in addition. The Extra-Dimension model introduces the other spatial dimension with compactification [23]. This is one of the theories including gravity propagating in the extra dimension. In this model, Z' is predicted by the excitation of the Z boson by the propagation in the additional dimension.

An alternative physics model with SUper-SYmmetry (SUSY) mechanism [24–27] is also highly motivated. The SUSY can explain the shortcomings in the SM (mainly existence of dark matter and hierarchy problems of Higgs mechanism) by introducing SUSY fermion (boson) partners for the SM bosons (fermions). These SUSY particles are called sfermions and bosinos, respectively. Therefore, it was expected that the direct production would be observed in high energy collider experiments, however such signature has not yet been validated. For example, the Minimal Supersymmetric extension of the Standard Model (MSSM) also predicts cLFV processes between muon and tau leptons [28]. In this model, it is supposed that mass matrices of sleptons are tuned to allow cLFV processes



(a) (b)

(c) (d)

Figure 1.3: Diagrams of the cLFV process predicted by the SUSY framework. The simple MSSM description introduces the flavor violation via a charged slepton and a neutralino(a). By introducing Higgs-mediation, R-parity violation, or LR symmetry in addition to the MSSM framework, cLFV processes are given as described in (b), (c) or (d), respectively.

to arise at one loop level by exchanging of neutralinos (charginos) and charged sleptons (sneutrinos) as shown in Figure 1.3(a). Higgs-mediated extension is also proposed in addition to the fundamental MSSM model [29]. The diagram of simplified process is described in Figure 1.3(b). A rare decay mode of Higgs decay into charged leptons with different flavor is another sensitive probe, but this model predicts the cLFV decay of charged leptons too. Therefore, these searches are important and complementary for the verification. There is another possibility of the MSSM expecting cLFV signature to introduce the R-parity violation [30,31], which allows SUSY particles to decay into SM particles. Figure 1.3(c) shows the flavor-changing process via a sneutrino ($\tilde{\nu}$) by the direct coupling with the SM particles. The combination of the MSSM and LR [32] also allow cLFV processes with doubly-charged boson ($\Delta^{\pm\pm}$). A doubly charged boson $\Delta^{\pm\pm}$ is derived from the model, and couples with leptons universally. Figure 1.3(d) describes an example of cLFV process mediated by this model.

Challenge to search for cLFV signature has been carried out by several experiments using direct, and indirect measurements. It is a major approach to search for the cLFV decay of the SM leptons. Electrons cannot decay into lighter visible particles, therefore μ and τ leptons are used as probes in the searches.

The MEG experiment [33] at Paul Scherrer Institute (PSI) is a representative project searching for cLFV decay of muon using $\mu \to e\gamma$ channel. This experiment

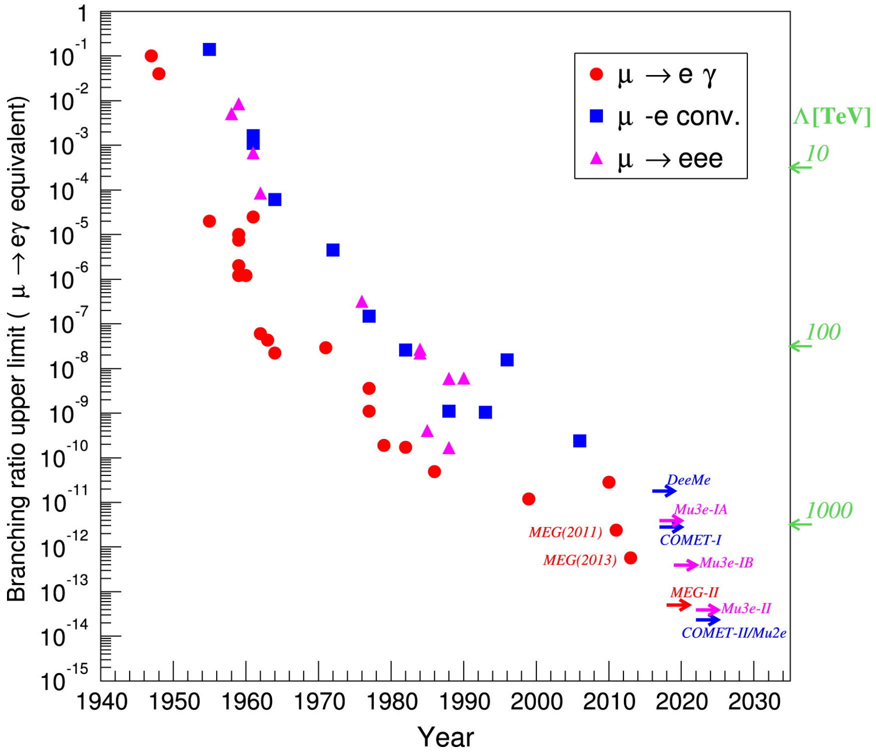


Figure 1.4: Summary plot of the progress and future prospects of cLFV search in muon decay [41].

set stringent limit for the $\mathcal{BR}(\mu \to e\gamma) < 4.2 \times 10^{-13}$ at 90% confidence level, using 7.5×10^{14} stopped muons [34]. This is the best upper limit going over the result by SINDRUM-II experiment [35], however excess from the SM has not been observed. In order to upgrade the sensitivity for the decay, several new experiments are being prepared; DeeMe [36], COMET experiments [37] for $\mu \to e$ conversion process at J-PARC in Japan and Mu2e experiment at Fermilab in USA [38]. Mu3e experiment [39] is also planned to search for $\mu \to eee$, which is another important channel for cLFV. Currently, the most stringent limit for $\mathcal{BR}(\mu \to eee)$ is set by SINDRUM as 10^{-12} [40], which is looser than $\mu \to e\gamma$. Thus, cLFV search in muon decay is eagerly carried out currently as shown in Figure 1.4.

τ decay is also a useful probe for the cLFV evidence. Figure 1.5 shows the summary of the limit of each LFV decay channel. These limits were mainly set by the collider experiments with high luminosity electron-positron collision; the CLEO [42], BaBar [43], and Belle experiments [44, 45]. LHCb also sets limits for three channels [46, 47], using τ leptons produced by pp collision. Currently, the limits for $\mathcal{BR}$ were mostly set by the Belle experiment.

Direct search for a massive boson decaying into different flavor leptons is also a sensitive approach. The target bosons are the SM gauge and scaler bosons (Z, h), and the heavy new particles (Z', H^0, A^0) predicted by the BSM. The AT-

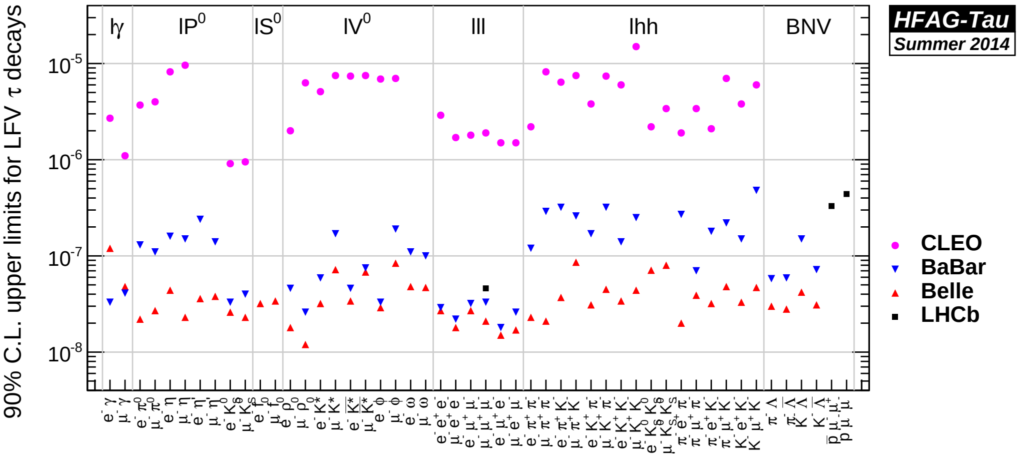


Figure 1.5: Summary of the limits for each cLFV process in τ lepton decay [48].

LAS and CMS experiments searched for $h, Z \to \tau\mu$ signature independently, but no evidence was observed [49–52]. Signature of heavy LFV particles have neither been observed yet [53, 54]. These results show that the cLFV process with a light propagator or large coupling does not exist. Nevertheless, the indirect measurements using huge amount of the SM particles can explore a tiny signature by a very heavy propagator or very small coupling. Hence, searching for the cLFV signature in lepton decay should play an important role at Large Hadron Collider (LHC) era.

1.2.3 Target of $\tau \to 3\mu$ search

$\tau \to 3\mu$ is a probe of the flavor violation between τ and μ using decays of τ leptons. The limits on the branching ratio of the other τ-μ violation channels are strictly set by the $\tau \to \mu\gamma$, $\tau \to \mu\pi^0$, and $X \to \tau\mu$. However, there are several BSM models which can be probed by the $\tau \to \mu ll$ but cannot be done by the others. In addition, it has a characteristic final state which is composed of muons only. Especially for hadron collider experiments, muons are more clearly detectable and identifiable than electrons and hadrons. Therefore, this channel is expected to have the highest feasibility in the $\tau \to \mu ll$ channels. Expected $\mathcal{BR}(\tau \to 3\mu)$ from several BSM physics models are summarized in Table 1.3.

In the SM framework, $\tau \to 3\mu$ can be produced by introducing neutrino oscillation as described in Figure 1.6. The branching ratio of the $\tau \to 3\mu$ decay of this process is estimated using PMNS matrix and the masses of the neutrinos:

$$\begin{aligned}\mathcal{BR}(\tau \to 3\mu)_{SM} &= \frac{3\alpha}{32\pi}\left|\sum_{i=1,3} U^*_{\tau i} U_{\mu i}\frac{\Delta m_{i2}}{M_W^2}\right|^2 \\ &\sim 10^{-54}. \end{aligned} \tag{1.25}$$

Table 1.3: Summary table of the $\mathcal{BR}(\tau \to 3\mu)$ predicted by each BSM physics, as well as only by neutrino oscillation.

Physics Model	Maximal Expected $\mathcal{BR}$
Simplified Z' [21]	10^{-8}
LR-symmetric Z' [22]	10^{-8}
Extra-Dimension Z' [23]	10^{-7}
MSSM [28]	10^{-8}
Higgs-mediated MSSM [29]	10^{-8}
RPV-MSSM [31]	10^{-10}
LR-MSSM [32]	10^{-8}
Netrino Oscillation [55]	10^{-54}

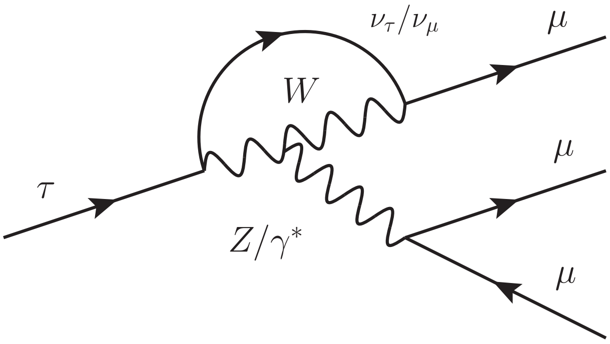


Figure 1.6: Diagram of $\tau \to 3\mu$ process with the neutrino oscillation.

The expected $\mathcal{BR}$ of the target signal is too small to detect significantly when considering the neutrino oscillation only.

On the other hand, the BSM physics models predict experimentally visible values at approximately 10^{-8} level because of the additional cLFV coupling. For example, in the simplified Z' model, the branching ratio is determined as

$$\mathcal{BR}(\tau \to 3\mu)_{Z'} = \frac{\tau_\tau m_\tau}{1536\pi^3 M_{Z'}^4}\left[\left(2g_{R,\mu}^2 + L,\mu^2\right) L,\tau\mu^2 + \left(2g_{L,\mu}^2 + R,\mu^2\right) R,\tau\mu^2\right], \tag{1.26}$$

where τ_τ and m_τ are the life time and mass of the τ lepton. $g_{R(L),i}$ and $g_{R(L),ij}$ represent the left(right)-handed coupling parameter between the same and different flavors, introduced by the Z' mediated interaction:

$$\mathcal{L}_{Z'} = \bar{f}_i\gamma^\mu(g_{L,i}P_L + g_{R,i}P_R)f_i Z'_{mu} + \bar{l}_i\gamma^\mu(g_{L,ij}P_L + g_{R,ij}P_R)l_j Z'_{mu}, \tag{1.27}$$

where f_i and l_i represent the fermions and charged leptons in the SM, respectively. $P_{L(R)}$ is the chiral projector operator for left(right)-handed coupling defined as $(1 \pm \gamma_5)/2$. The first term shows the universal coupling between the same flavor fermions, while the second represents the cLFV term. In this case, $\mathcal{BR}(\tau \to$

3μ) depends on the Z' mass and coupling strength as shown in Eq (1.26). For the assumption of $M_{Z'} = 2000\,\mathrm{GeV}$ and $g_{R(L),\tau\mu} \sim 0.1$, $\mathcal{BR}(\tau \to 3\mu)\sim 10^{-8}$ is predicted.

For an example of the SUSY models, the effective Lagrangian of Higgs-mediated MSSM by the Higgs exchange is described as

$$\mathcal{L}_{Higgs} = -\sqrt{2}G_F m_\mu m_\tau (\mathcal{C}_+ \bar{\mu}\bar{\mu^c} + \mathcal{C}_- \mu^c \mu)(\Delta_R \mu^c \tau + \Delta_L^* \mu \tau^c) + h.c., \tag{1.28}$$

where $\mathcal{C}_\pm$ is determined as $\sim 2\tan^3\beta/m_A^2$ by $\tan\beta$ (fraction of the vacuum potential of A and H) and mass of A. The $\Delta_{L(R)}$ represents the cLFV coupling for the left(right)-handed leptons. The branching ratio can be determined as

$$\mathcal{BR}(\tau \to 3\mu)_{Higgs} \sim 4.0 \times 10^{-5} \left(\frac{\tan\beta}{40}\right)\left(\frac{100\,\mathrm{GeV}}{m_A}\right)\frac{\mathcal{BR}(A \to \mu\tau)}{\mathcal{BR}(A \to \tau\tau)}. \tag{1.29}$$

Hence, this depends on the parameters of A and H. This theory proposes that $\mathcal{BR}(\tau \to 3\mu)\sim 10^{-9}$ at least for the available parameter combination [29].

For the experimental aspect, this signature has never been observed, and the most stringent limit is set as 2.1×10^{-8} [45], as also shown in Figure 1.5. LHCb experiment also set the limit as 4.6×10^{-8}, which is the first result by the hadron collider experiment [47]. However, all these results have not reached the sensitivity of 1.0×10^{-8} yet. Currently, the Belle project is in the process of upgrading to Belle-II with higher luminosity [56], thereby improving the sensitivity for this channel. Therefore, it is also important to search for this channel by the hadron collider, whose feasibility has been established by the LHCb result. At the LHC experiment, it is expected to increase the τ production cross-section by the higher center-of-mass energy, and integrated luminosity in the current run. These approaches for the target signature are important as validation, and possibly improve the search.

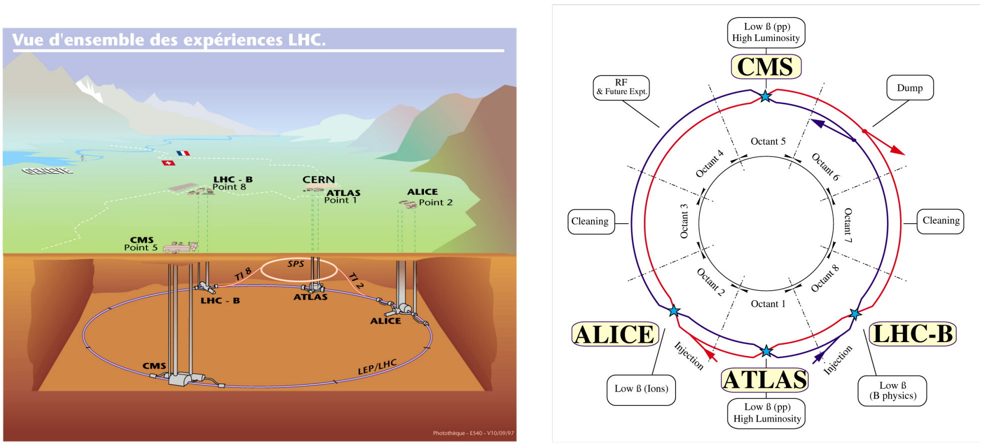
Chapter 2

Experimental Setup

The LHC succeeded in producing the highest center-of-mass energy physics with proton-proton collision, reaching 8 TeV in 2012 and 13 TeV in 2015. This achievement makes it possible to search for new physics at the unexplored high energy scale. To maximize the physics capability of LHC and explore the signature predicted from various new physics models, several different approaches should be taken in the experiments which include direct search for heavy new particles, search for missing particle signatures, and rare decay search of the SM particles. Such require a general detector with large acceptance. A Toroidal LHC ApparatuS (ATLAS) detector is a combined detector of tracker, calorimeter, and muon specific detectors with approximately 4π solid angle coverage. In this chapter, details of the LHC collider and the ATLAS detector are described, as well as the data acquisition system.

2.1 The Large Hadron Collider

LHC is designed to achieve proton-proton collisions with a center-of-mass energy of 14 TeV using two rings of circumference of 26.7 km and strong magnetic field of 8.33 T. This circular collider is located 100 m underground at CERN, as well as Super Proton Synchrotron (SPS), the proton booster accelerator. Proton (or heavy ion) bunches which are boosted by SPS are injected into two circulation rings for opposite direction proton beams in this collider, and collide at four bunch crossing points, where the ATLAS, CMS, LHCb, and ALICE detectors are located. Four detectors were designed and constructed by different collaborations putting focus on each physics subject. The ATLAS and CMS experiments propose to search for several new physics widely by their large solid angle and acceptance of the detector, while LHCb focuses on forward region to study B-physics. The target of these three experiments is to study high energy pp collisions. On the other hand, the ALICE experiment is dedicated to study heavy ion collisions. Figure.2.1 is the overview of the LHC and SPS layout, as well as the experiments at the crossing points.



(a) (b)

Figure 2.1: (a)Overview of the LHC and SPS [57, 58]. They are located between 50 m and 150 m underground. Proton bunches are accelerated by SPS and LHC up-to 480 GeVand 7 TeV as design value, respectively. (b)General schema and experiments at each crossing point. Opposite proton beams are squeezed and focussed for collision at the crossing points of ATLAS, CMS, LHCb and ALICE.

One of the most important factors of collider physics is the instantaneous luminosity of the collision, represented by $\mathcal{L}$. Higher luminosity enables to search for rare processes which could be driven by new physics or help to make precise measurements of the known processes. This parameter is determined by beam parameters as:

$$\mathcal{L} = \frac{N_{\mathrm{b}}^2 n_{\mathrm{b}} f_{\mathrm{rev}} \gamma_{\mathrm{r}}}{4\pi\epsilon_{\mathrm{n}}\beta^*} F, \tag{2.1}$$

where N_{b} is number of particles per bunch, n_{b} number of bunches per beam. f_{rev} is the revolution frequency, γ_{r} the relativistic gamma factor, ϵ_{n} the normalized transverse beam emittance, $\beta*$ the beta function at the collision point, and F the geometric luminosity reduction factor. F is determined by the crossing angle at the interaction point and given by:

$$F = \left(1 + \left(\frac{\theta_{\mathrm{c}}\sigma_{\mathrm{z}}}{2\sigma^*}\right)^2\right)^{-1}, \tag{2.2}$$

where θ_{c} and σ_{z} are the full crossing angle at the interaction point and the RMS bunch length, respectively. In principle, number of interactions is given by the cross section, represented by $\sigma_{\mathrm{process}}$, and integrated luminosity of the data taking period:

$$N_{\mathrm{evt}} = \sigma_{\mathrm{process}} \int \mathcal{L} dt. \tag{2.3}$$

Table 2.1: Parameters of LHC in each year. In the 2012 run, center-of-mass energy was 8 TeV, while it was upgraded up to 13 TeV after the 2015 run. Bunch crossing rate in the Run2 is also higher than Run1 because of higher beam intensity. These values achieved nearly the design parameters already.

Parameter	Run1(2012)	Run2(2015)	Run2(2016)	Design value
E_{CM}[TeV]	8	13	13	14
$N_{\mathrm{b}}[10^{11}]$	1.7	1.1	1.1	1.15
$n_{\mathrm{b}}^{\mathrm{max}}$	1374	2244	2076	2808
$\epsilon_{\mathrm{n}}[\mu\mathrm{m}]$	2.5	3.5	3.4	3.75
β^*[cm]	60	80	40	55
$\theta_{\mathrm{c}}[\mu\mathrm{rad}]$	290	290	370	285
$\mathcal{L}[10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}]$	0.77	0.51	1.37	1.0

Hence, amount of the data for physics analysis is determined by $\int \mathcal{L}dt$.

The run period from 2010 to 2012 is called as Run1. However, the center-of-mass energy of pp collision in 2010 and 2011 was set as 7 TeV, while that in 2012 was set as 8 TeV. Therefore, it is difficult to use these data together due to the difference of the energy and the configuration of the data acquisition. In this thesis, the data taken during 2012 is called the Run1 data, excluding those from 2010 to 2011. In this period, the maximum instantaneous luminosity was $\mathcal{L} = 7.7 \times 10^{33}$ cm^{-2}s^{-1} as nearly 80% of the design value. Figure 2.2 shows the evolution of the maximum instantaneous luminosity in 2012. The long shutdown for upgrade toward higher energy and luminosity was finished by the end of 2014, and pp collision with center-of-mass energy of 13 TeV started in 2015 spring. The instantaneous luminosities in 2015 and 2016 are shown in Figure 2.3. Maximum instantaneous luminosity in 2015 was 5.0×10^{33}, which is lower than 2012 run because it was the start-up period of the 13 TeV run. The maximum $\mathcal{L}$ value of 2016 run was up to 1.2×10^{34}, which is around design value. Detailed parameters of the LHC collider in 2012, 2015 and 2016 runs are summarized in Table.2.1. LHC has been running stably, and evolving with the machine.

2.2 The ATLAS Detector

The ATLAS detector is a multi-purpose detector to collect pp collision data at the LHC. This detector is separated into cylindrical (barrel) and disk-shaped (endcap) parts, covering almost 4π rad in solid angle. Its size is 25 m in height and 44 m in length, and the weight is approximately 7000 t. Overview of the ATLAS detector is shown in Figure 2.4.

Both Cartesian and cylindrical coordinate systems with its origin at the Interaction Point (IP) are used in ATLAS. The z-axis of the right-handed coordinate system is the beam direction. The positive x-axis and y-axis of the system are defined as toward the center of the LHC ring and upwards from the IP, respec-

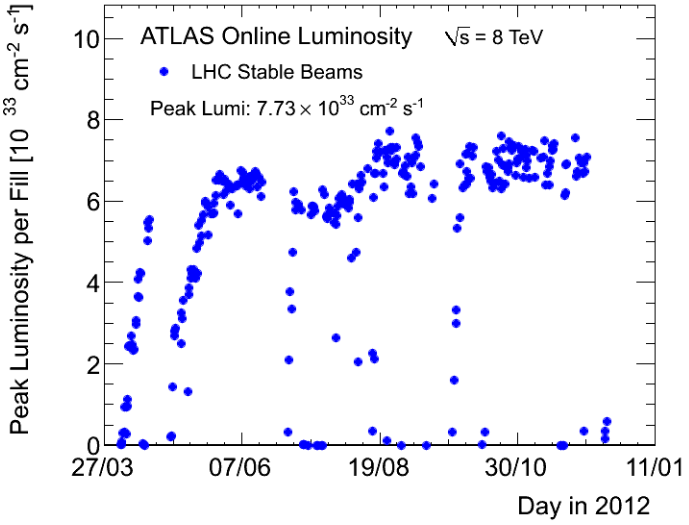
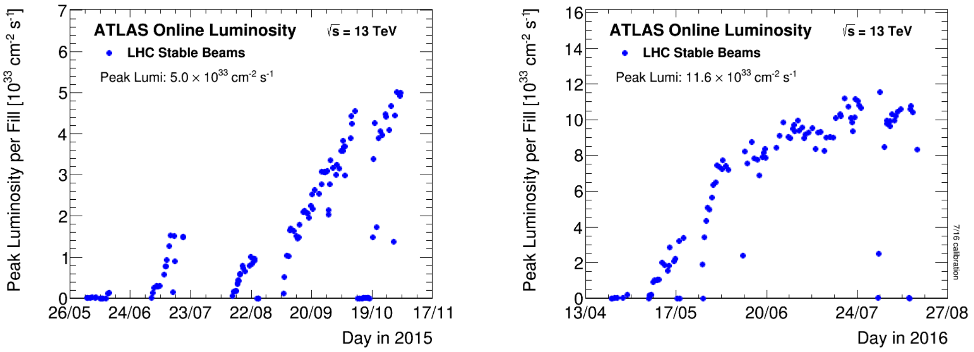


Figure 2.2: The maximum instantaneous luminosity per day recorded by the ATLAS experiment in 2012 [59]. Although the maximum value in the period did not reach the design value, these runs were stable and favorable for data analysis.



(a) (b)

Figure 2.3: The maximum instantaneous luminosity per day recorded by the ATLAS experiment in Run2 [59]. (a)In the 2015 run, instantaneous luminosity was raised gradually since the beginning of the 13 TeV collision. Although the instantaneous luminosity in this year was lower than 2012, this figure shows the improvement of the beam intensity. (b)The maximum value reached 10^{34} in the 2016 run. Even after the achievement of the design value of 1.0×10^{34} cm^{-2}s^{-1}, the luminosity has been further increased thanks to continuous efforts by accelerator working team.

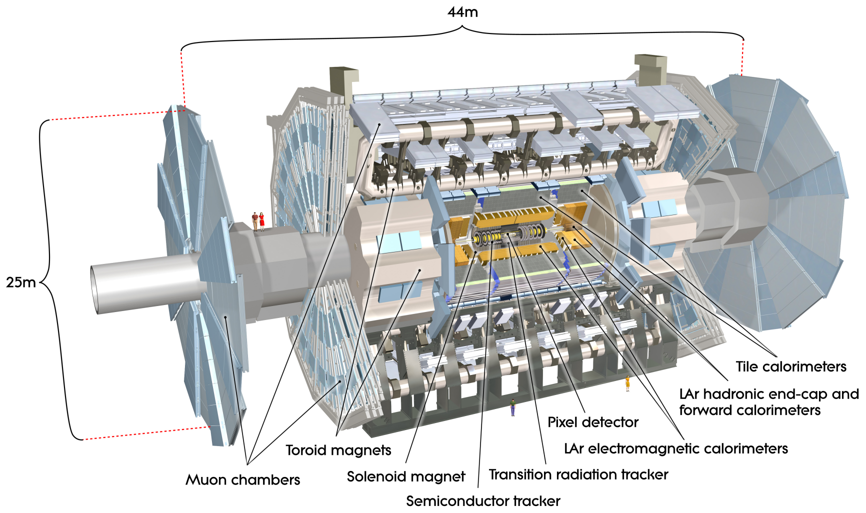


Figure 2.4: Overview of the ATLAS detector [60].

tively. In the cylindrical coordinate system, the beam direction is defined as the longitudinal z-axis and the transverse plane are described with polar axis r and azimuthal coordinate ϕ. The pseudorapidity η is defined using the polar angle θ with respect to the z axis:

$$\eta = -\ln \tan \frac{\theta}{2}. \tag{2.4}$$

Generally, detector location, coverage and particle directions are described using ϕ and η. As well as the position and direction of each particle, the angular distance between two particles is also defined using these angular parameters as:

$$\Delta\eta_{i,j} = |\eta_i - \eta_j| \tag{2.5}$$
$$\Delta\phi_{i,j} = \arccos(\cos(\phi_i - \phi_j)) \tag{2.6}$$
$$\Delta R_{i,j} = \sqrt{\Delta\eta i,j^2 + \Delta\phi i,j^2}. \tag{2.7}$$

Energy and momentum of the detected particles are divided into longitudinal and transverse components. The vector sum of transverse momentum (p_{T}) of emitted particles from IP must be zero because the transverse momentum of the initial partons are always zero. On the other hand, longitudinal momentum of the colliding partons (quarks and gluons) are not fixed but randomly distributed. Hence, transverse parameters of detectable particles and missing energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) of the event assuming transverse momentum balance are used in analyses.

Several specific sub-detector systems compose the ATLAS detector. The innermost sub-detector of ATLAS is the Inner Detector (ID) which measures tracks of charged particles. The ID is enclosed in a superconducting solenoid providing a 2 T axial magnetic field. The track parameters of charged particle and location

Table 2.2: Typical resolution for measured parameters and coverage of each sub-detector.

Sub-detector	Typical Resolution	Coverage in η	
		Precision	Trigger
ID Tracking	$\sigma_{p_\mathrm{T}} = 0.05 \times p_\mathrm{T}(\mathrm{GeV}) \oplus 1\%$	$\lvert\eta\rvert < 2.5$	-
EM Calorimetry	$\sigma_E = 10/\sqrt{E}(\mathrm{GeV}) \oplus 0.7\%$	$\lvert\eta\rvert < 3.2$	$\lvert\eta\rvert < 2.5$
Hadron Calorimetry			
Barrel,Endcap	$\sigma_E = 50/\sqrt{E}(\mathrm{GeV}) \oplus 3\%$	$\lvert\eta\rvert < 3.2$	
Forward	$\sigma_E = 100/\sqrt{E}(\mathrm{GeV}) \oplus 10\%$	$3.1 < \lvert\eta\rvert < 4.9$	
Muon Spectrometry	$\sigma_{p_\mathrm{T}} = 10\%$ at p_T=1 TeV	$\lvert\eta\rvert < 2.7$	$\lvert\eta\rvert < 2.4$

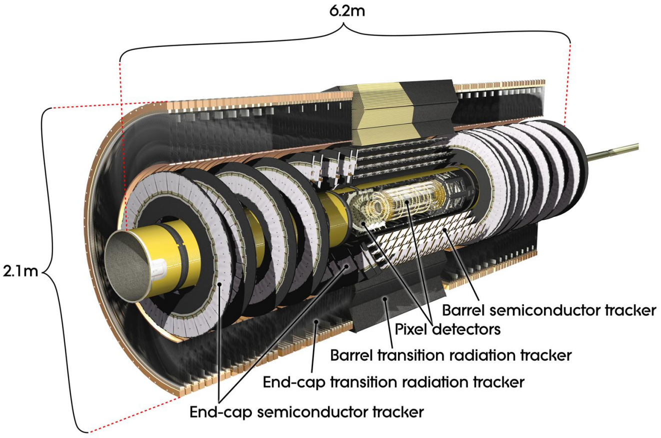


Figure 2.5: Overview of the ATLAS inner detector [61]. Its total length and height are 6.2 m and 2.1 m, respectively.

of the vertices are determined precisely by this detector system. Details of the ID are described in Section 2.2.1. The calorimeters is placed so as to surround the ID. Energies of photons, electrons, and hadrons are measured by the several kinds of calorimeters as described in Section 2.2.2. The outermost part of the detector is a Muon Spectrometer (MS). The MS is located inside these toroidal magnetic fields with bending integral of about 0.5 T, generated by superconducting air-core toroid magnets, one in barrel and two in each endcap. Section 2.2.3 describes details of the MS. The performance and coverage of these sub-detector systems are summarized in Table 2.2.

2.2.1 Inner Detector

The ID system is composed of the Pixel detector (Pixel), the Semi-Conductor Tracker (SCT) and the Transition Radiation Tracker (TRT), as shown in Figure 2.5.

The Pixel is a silicon detector and located at the innermost of the ID. Robustness against radiation and high resolution of the hits are required for the stable

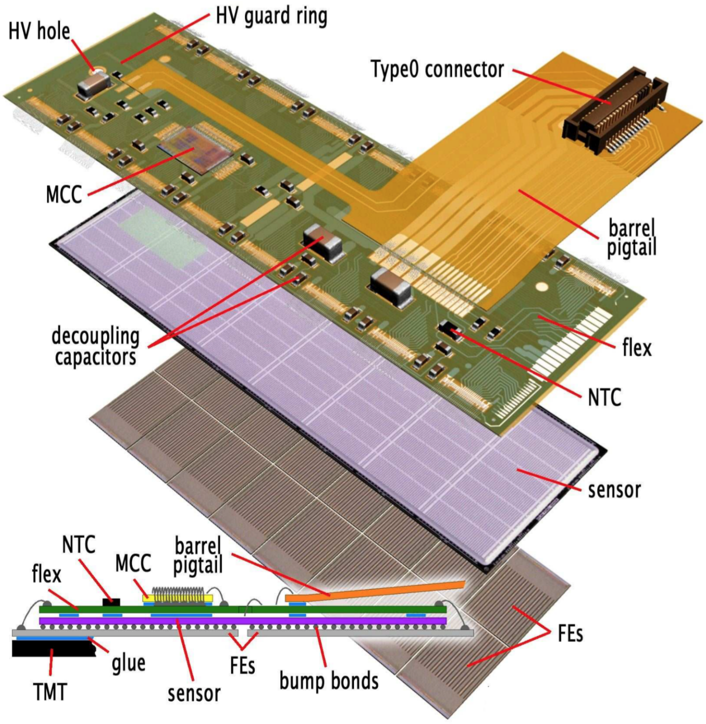


Figure 2.6: The sketch of the ATLAS pixel detector module [62].

and precise measurement of the vertices and track parameters. One Pixel sensor is equipped with a small detection unit of 50×400 $\mu\mathrm{m}^2$ sensitive area with 250 μm thickness, having 46,080 channels per sensor. Hits from typically 500-1000 charged particle tracks on the pixel detector are recorded in each bunch crossing. The design of a detector module and the overall structure are shown in Figure 2.6 and 2.7, respectively. In a heavy radiation environment, it is difficult to reconstruct tracks and vertices correctly because the number of tracks increases. In addition, stable performance of the detector against radiation damage must be kept. The Pixel detector system consists of three layers in 2012. Since 2015 run, Insertable B-Layer (IBL) was installed into the innermost layer of the Pixel to solve the difficulty of the operation and performance in the harder radiation situation with increased beam intensity. The layout of the IBL is shown in Figure 2.8.

The SCT is also a silicon detector surrounding the Pixel layers and covering an acceptance of $|\eta| < 2.5$. The detector system consists of 4088 modules tiling four co-axial cylindrical layers in the barrel region and two end-caps each containing nine disk layers. For the barrel layers, 80 μm pitch micro-strip sensors are employed. Figure 2.9 is a drawing of a barrel SCT module. The four sensors, two each on the top and bottom side, are rotated by ± 20 mrad each other to measure the hit positions both on r-ϕ and z directions. The typical accuracies per module in the barrel are 17 μm ($r\phi$) and 580 μm (z). On the other hand, for the endcap module, there are three types of modules and two sets of sensors

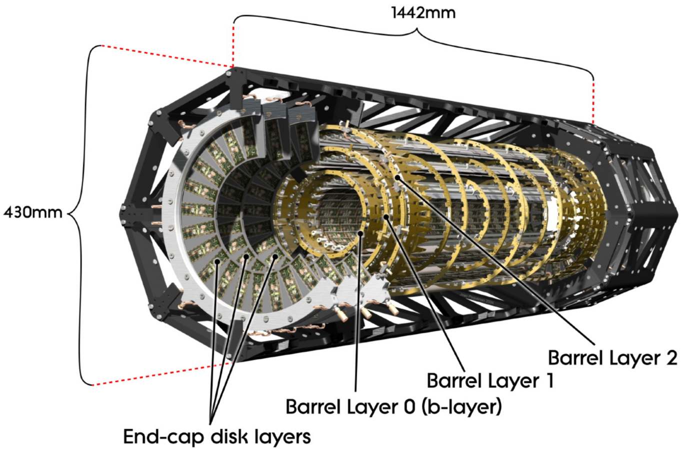


Figure 2.7: Support structure of the pixel detector extracted from Figure 2.5 [63].

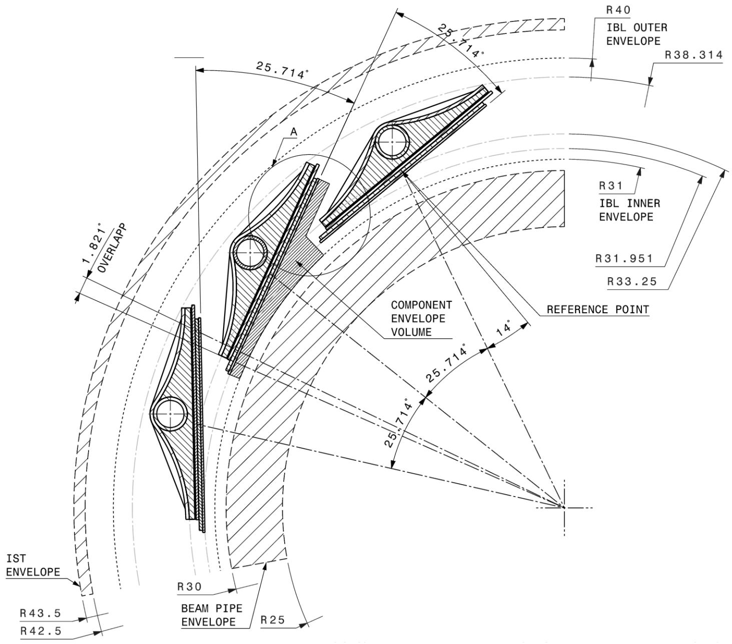


Figure 2.8: r-ϕ view of the IBL layout [63]

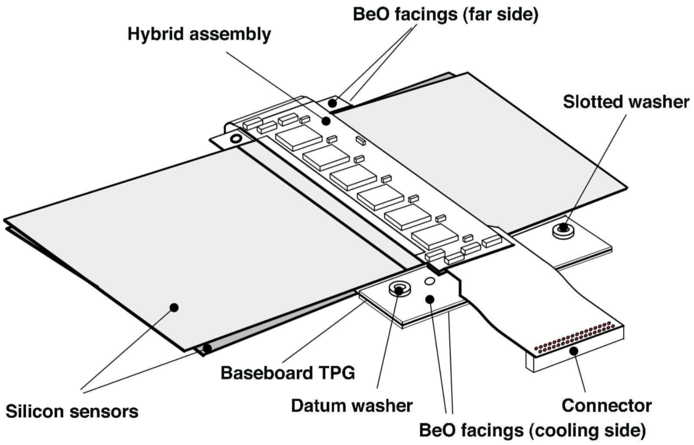


Figure 2.9: Layout of the SCT sensor module [62].

with a relative rotation of ±20 mrad are located to obtain the resolution in r-ϕ and r. The typical accuracies per module in the endcap are also 17 μm ($r\phi$) and 580 μm (r).

The outermost part of the ID is TRT, which is composed of 73 layers of straws with fibers in the barrel and 160 straw planes with foils in the endcap. The overall coverage of the TRT is $|\eta| < 2.0$ (boundary between the barrel and endcap is at $|\eta| = 1.0$). The straw tube is 4 mm diameter polymide-made drift tube, which consists of 35 μm thick multi-layer films and mixed gas of 70% Xe, 27% CO_2 and 3% O_2. The films are made of a 25 μm thick polyimide film coated on one side with a 0.2 μm Al layer, and protected by a 56 μm thick graphite-polyimide layer. Low energy transition radiation photons, whose probability of the emission is proportional to the Lorentz boost factor, are absorbed in the gas. As the Lorentz boost factor depends on the mass of the incident particle, particle types can be discriminated using low and high thresholds in the front-end electronics. The TRT provides the hits on r-ϕ direction with an accuracy of 130 μm per straw.

2.2.2 Calorimeter

The Calorimeter consists of a number of sub-components with full ϕ-symmetry and coverage around the beam axis as depicted in Figure 2.10. These sub-components, called sampling detectors, are made of absorbers and active sensors. The energy of the particles is measured from the energy deposit of electromagnetic or hadronic showers in the active sensors generated by the interaction in absorbers. The Calorimeter covers the range $|\eta| < 4.9$ and consists of several kinds of absorbers and active materials considering the physics processes and the radiation environment. Both of electromagnetic and hadronic calorimeters are placed, to measure the energy of electrons, photons and hadrons. The electromagnetic calorimeters use liquid argon (LAr) as the active detector medium. The hadronic calorimeters employ steel as absorber material and scintillating tiles as

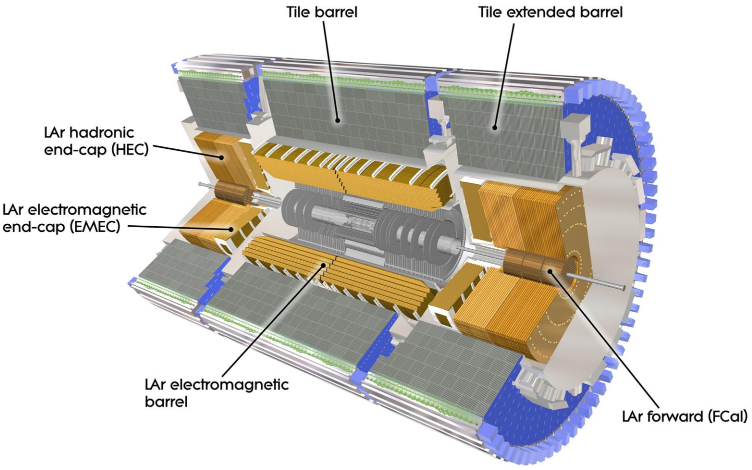


Figure 2.10: Overview of the calorimeter [62]. Three LAr calorimeters for electromagnetic and hadronic interactions and a tile calorimeter for hadronic interactions are placed over a large acceptance up to $|\eta| = 4.9$.

active medium for $|\eta| < 1.7$ and LAr calorimeter for $1.5 < |\eta| < 3.2$. Hence, four kinds of detectors for each region are used in ATLAS: the precision electromagnetic calorimeter (EM), a hadronic end-cap calorimeter (HEC), a forward calorimeter (FCal), and a hadronic tile calorimeter (Tile). The interaction length (λ), determined by the material thickness of the detector, is shown in Figure 2.11 as a function of $|\eta|$. The interaction length of the calorimeter is larger than 10 λ for $|\eta| < 4.9$, which is large enough to keep high energy resolution and to suppress hadronic particles penetrating into muon spectrometer.

The EM calorimeter is made of accordion-shape lead absorber and electronodes as described in Figure 2.12. As shown in Figure 2.10, the combination of the barrel EM detector (covering $|\eta| < 1.475$), and the endcap EM detector (covering $1.375 < |\eta| < 3.2$) achieves its continuous acceptance for $|\eta| < 3.2$ excluding an azimuthal crack around $|\eta| = 0$.

The hadronic calorimeters cover up to $|\eta| = 4.9$ by the Tile placed for $|\eta| < 1.7$, the HEC for $1.5 < |\eta| < 3.2$, and the FCal for $3.1 < |\eta| < 4.9$. The Tile calorimeter is located outside the EM calorimeter and consists of a barrel covering $|\eta| < 1.0$ and two extended-barrel covering $0.8 < |\eta| < 1.7$ as shown in Figure 2.13. The HEC calorimeter is placed behind the endcap EM calorimeter as shown in Figure 2.14. The FCal covers the highest η region of the calorimeters, composed by three layers as shown in Fig.2.14. The benefit of the FCal is not only large coverage of the calorimeter, but also reduction of the radiation background levels in the muon spectrometer.

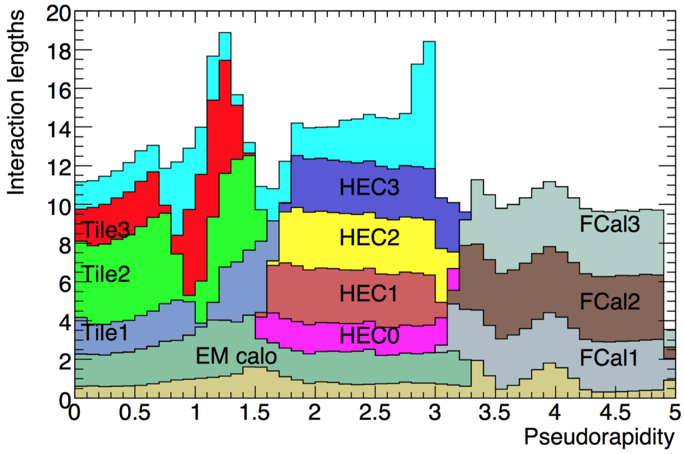


Figure 2.11: Interaction length which represents the amount of materials distributions as function of $|\eta|$ [62]. Over the calorimeter coverage, interaction lengths as large as about 10 are provided by each types of calorimeters. This is important for high resolution and suppression of the radiation background for the muon detectors.

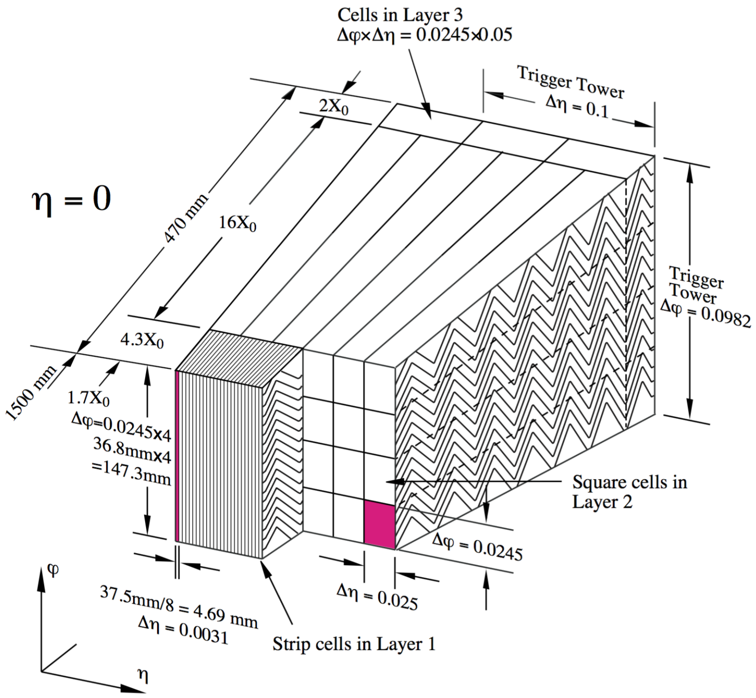


Figure 2.12: Sketch of a accordion-shape barrel module [62]. The granularity in η and ϕ direction of the cells is also shown.

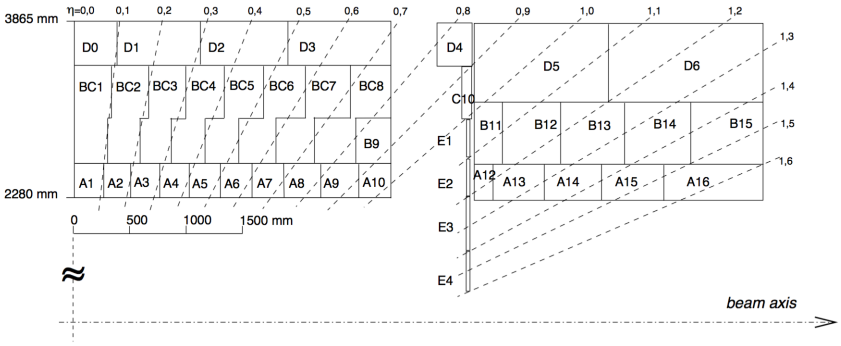


Figure 2.13: Section of the barrel and extended barrel detectors are shown with dashed line of η [62]. The tile calorimeter is symmetric about the interaction point at the origin.

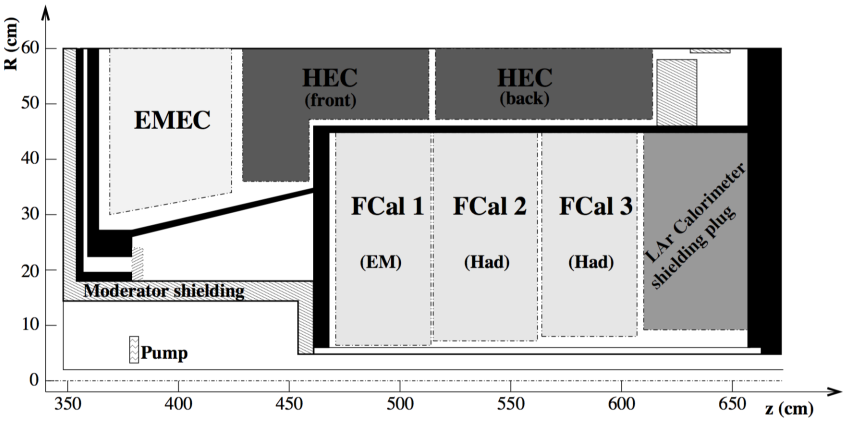
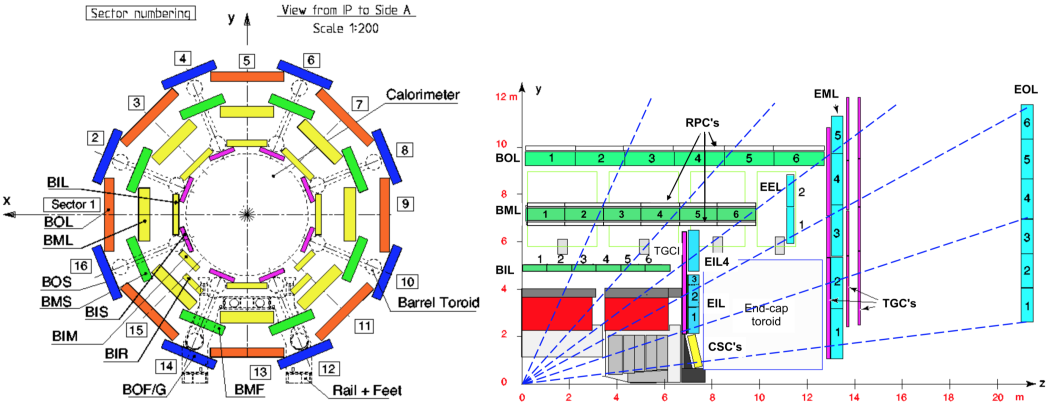


Figure 2.14: Structure of the hadron calorimeters in the endcap region [62]. Two layers of HEC and three layer of FCal are located at the low and high η region, respectively. These detectors are also symmetric about the IP at the origin.



(a) (b)

Figure 2.15: (a) Overview of the barrel detectors as perpendicular to the beam axis and (b) detector location of the detectors in a plane containing the beam axis [62]. Muons with infinite momentum typically traverse three precision and three trigger detector stations even around transition region at $|\eta| = 1.05$. There are eight large and eight small chambers in the barrel region, and chambers around feet of the detector are different and named as large-special and small-special.

2.2.3 Muon Spectrometer

The ATLAS muon detector focuses on the precise momentum measurement using spectrometers. It consists of four kinds of detectors, the Monitored Drift Chamber (MDT), the Cathode Strip Chamber (CSC), the Resistive Plate Chamber (RPC), and the Thin Gap Chamber (TGC). Muons penetrate the ID, calorimeter, and other materials in the detectors, therefore all these spectrometers are placed outside of the calorimeters as shown in Figure 2.15. The region inside and outside $|\eta| = 1.05$ are referred as barrel and endcap of the MS, respectively, and each region is equipped with different detector configuration. At the endcap region, these chambers are tiled like a disk-shape, while cylindrical structure at the barrel region is made by tiling with chambers. The detector construction in both regions are eight-fold rotational symmetric on the basis of the beam direction.

The MDT and CSC covering up to $|\eta| = 2.7$ are used for precision measurement of the track parameters of muons. On the other hand, the RPC at the barrel and TGC at the endcap are specialized for trigger decision described in Section 2.2.4. Table 2.3 shows the properties and typical performances. The response speed and resolution of these detectors are optimized for their usage and amount of radiation background.

The MDT is a set of drift tubes made of anode wire, cathode tube, and gas mixture. Figure 2.16 shows the structure of a barrel chamber composed by drift tube multi-layers and support frame. Four optical alignment rays, two parallel and two diagonal, monitor the internal geometry of the chamber. The gas mixture component is 93% Ar, 7% CO_2 and less than 1000 ppm H_2O. Relatively long drift

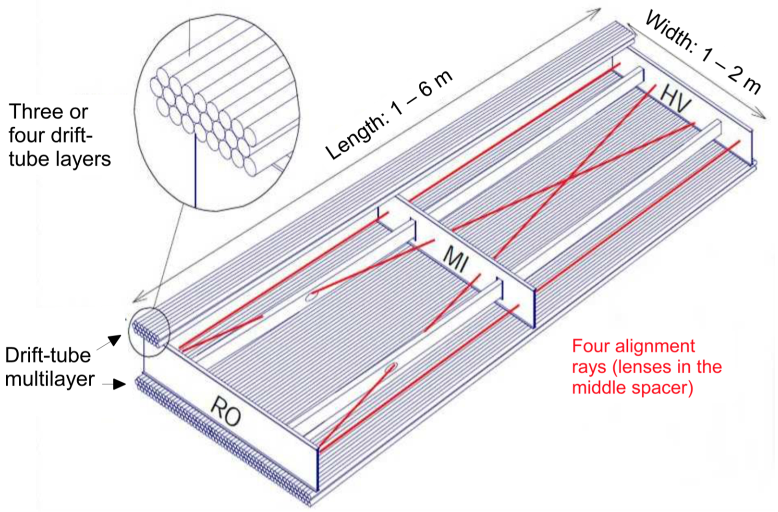


Figure 2.16: Structure of the MDT [62]. It consists of three spacer bars, two multi-layers of three or four drift tube layers, and alignment optimal rays. RO and HV are the readout electronics and high voltage supplies, respectively.

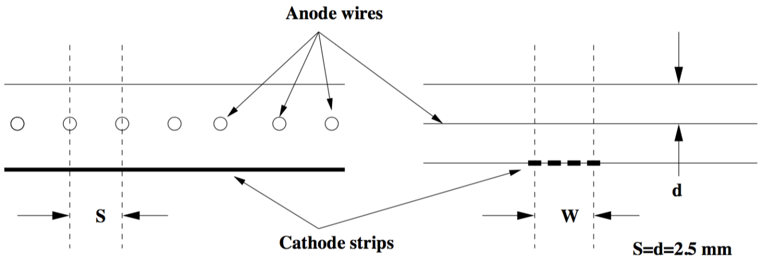


Figure 2.17: Illustration of the CSC mechanism [62]. The left figure is the looking down the wire, while the right is perpendicular view down the readout and intermediate strips.

time (700 ns at the maximum) enables to measure the position with high spacial resolution ($\sim 35\ \mu$m). Three MDT layers are placed at both of barrel and endcap regions and cover the full MS coverage excluding the most inner layer at the endcap region, which is covered by CSC.

The CSC is also a MWPC composed by anode wire, cathode plate, and gas mixture as shown in Figure 2.17. At high $|\eta|$ region above 2.0, the rate of muons and fake background is too high for MDT to be handled. Hence, the innermost layer of the MDT at such high $|\eta|$ region is covered by the CSC, whose response speed is faster than MDT (approximately 7 ns). The CSC is capable of operating with up to 1000 Hz/cm^2 counting rate. The resolution of CSC for the r direction is typically 40 μm.

The RPC consists of two plastic plates and gas mixture of $C_2H_2F_2$: IsoC_4H_{10} : SF_6 = 94.7 : 5.0 : 0.3. The characteristic feature of the RPC is its response

Table 2.3: Parameters of the muon spectrometers in ATLAS. Both of the precision detectors and trigger detectors cover up to $|\eta| = 2.7$.

MDT	Detectable Region	$|\eta| < 2.7$(inner layer:$|\eta| < 2.0$)
	# of chamber	1088(1150)
	# of channel	339000(354000)
	Resolution(z/r)	35 μm
	Usage	Precise measurement
CSC	Detectable Region	$2.0 < |\eta| < 2.7$
	# of chamber	32
	# of chamber	31000
	Resolution(z/r)	40 μm (r)
	Resolution(ϕ)	5 mm
	Response Speed	7 ns
	Usage	Precise measurement
RPC	Detectable Region	$|\eta| < 1.05$
	# of chamber	544(606)
	# of channel	359000(373000)
	Resolution(z/r)	10 mm (z)
	Resolution(ϕ)	10 mm
	Response Speed	1.5 ns
	Usage	Trigger
TGC	Detectable Region	$1.05 < |\eta| < 2.7$
	# of chamber	3588
	# of channel	318000
	Resolution(z/r)	$2-6$ mm (r)
	Resolution(ϕ)	$3-7$ mm
	Response Speed	4 ns
	Usage	Trigger

speed approximately 1.5 ns and its typical efficiency for muons higher than 98.5%. As shown in Figure 2.15, three RPC layers are placed at inside and outside of the middle MDT layer and outside of the outer MDT layer.

The TGC is a multi-wire proportional chamber covering $1.05 < |\eta| < 2.7$. The gas mixture is CO_2 : n-pentane = 55 : 45. The position resolution of TGC is 2-7 mm depending on direction and location, and the response speed is 4 ns.

2.2.4 Trigger system

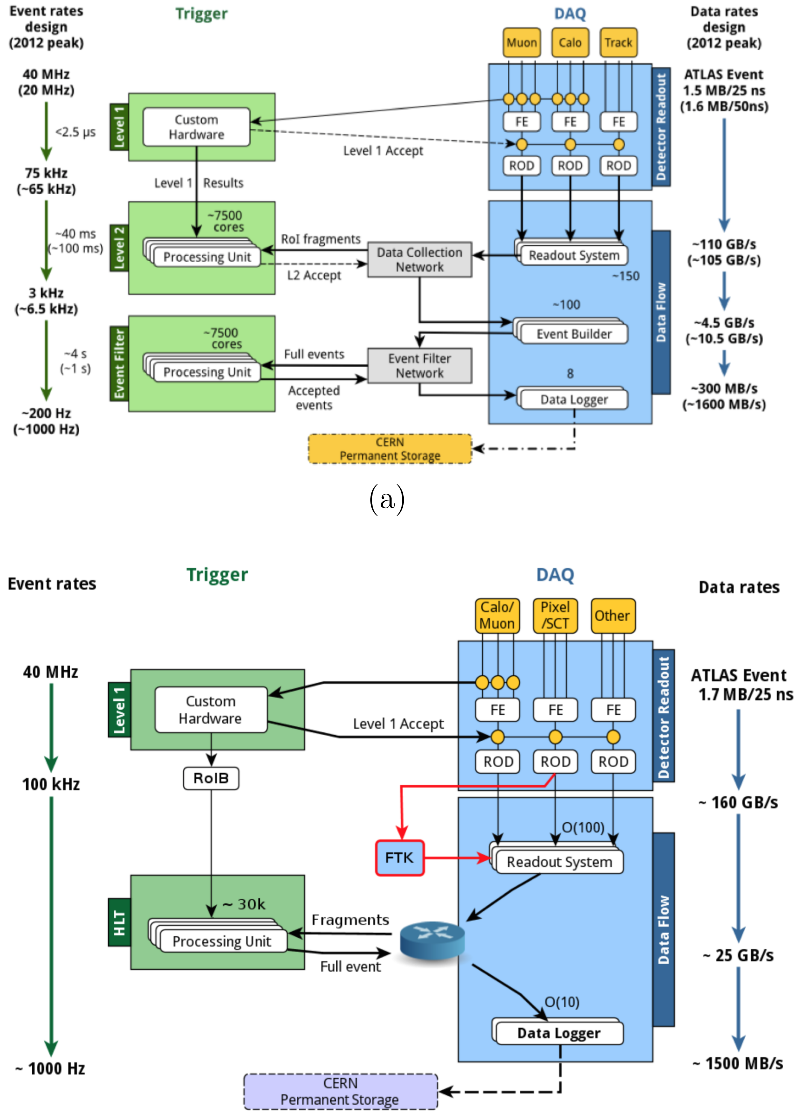
The proton bunches cross every 50 (25) ns at the IP in Run1 (Run2), hence bare event rate is around 20 (40) MHz, which are mostly occupied by the trivial QCD interactions. In the ATLAS experiment, more interesting events having particles with large p_{T} or $E_{\mathrm{T}}^{\mathrm{miss}}$ are selected in running and recorded. This is called trigger for the data acquisition. The trigger flow chart in Run1 and Run2 are shown in

Figure 2.18. The first step of the trigger system is a decision on hardware level named as Level-1 (L1), afterwards the High Level Trigger (HLT) decision works on a software level. In the Run1 trigger structure, HLT is further divided into the Level-2 (L2) and Event Filter (EF) by their requirement for processing speed. On the other hand, these steps are merged into one HLT system for simplification and resource sharing in Run2. This makes the processing time fast and enables to implement flexible algorithms at HLT because of the seamless reconstruction for the sub-detectors such as the calorimeters and the muon spectrometer. Thus, the trigger scheme is changed for the improvement of the processing speed and flexibility against the higher energy and luminosity of the pp collisions in Run2.

The limits of the output rates at each trigger step are upgraded in Run2 as shown in Figure 2.18. The L1 processing is required to work with the same rate as the bunch crossing of 20 (40) MHz, and its output rate must be below 75 (100) kHz in Run1 (Run2) due the upper limit on the HLT input rate. In order to reduce the rate and amount of the data down to writable rate of 200 (1000) Hz, more detailed decision using precise tracking and kinematics is processed on the HLT. A combination of the L1 and HLT is defined as a trigger chain, which determines the total data acquisition efficiency.

There are many kinds of trigger chains using characteristic signal of muons, electrons, photons, jets, and $E_{\mathrm{T}}^{\mathrm{miss}}$. These are running in parallel, and event data are recorded in case that one or more of these are fired. The name of the trigger chain is defined to contain required particle species, p_{T} (or E_{T}) thresholds, and multiplicity. For example, `EF(HLT)_2mu10` is the trigger chain requiring two muons above 10 GeV in Run1 (Run2), and a L1 part is also named like `L1_2MU10` following the same name convention. The name of a combined trigger chain is also defined like `EF_2mu8_xe80`, which require two muons above 8 GeV and large $E_{\mathrm{T}}^{\mathrm{miss}}$ above 80 GeV.

In order to acquire the $\tau \to 3\mu$ event efficiently, the specific feature of the event must be required at the trigger level because these are similar to QCD jet events including multi-muons, whose rate is very high and hence the event data are not acquired by nominal trigger chains. The multi-muon triggers are mainly used for this analysis as the most efficient configuration and a multi-muon+$E_{\mathrm{T}}^{\mathrm{miss}}$ trigger is also used only for W originated $\tau \to 3\mu$ search. In the Run2, dedicated triggers to $\tau \to 3\mu$ search were developed and installed to improve the trigger efficiency as described in Section 9.1.2. The general procedure of muon and $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger decision are described in this section, and their performance will be shown in Section 7.2. Trigger chains with low p_{T} or $E_{\mathrm{T}}^{\mathrm{miss}}$ threshold are difficult to be used for all bunch crossings due to their very high event rate. These chains are used in some parts of collisions by suppression factor, referred to as the Pre-Scale (PS) factor, for validation and performance measurement of the nominal trigger chains.



(a)

(b)

Figure 2.18: Flow chart of the ATLAS triggers system in (a) Run1 and (b) Run2 [64]. Typical requirement for output rates for each step are also shown in left column. These numbers were improved by the hardware and software used in trigger processing. The software decision was divided into two steps in Run1, which were merged into one step in Run2.

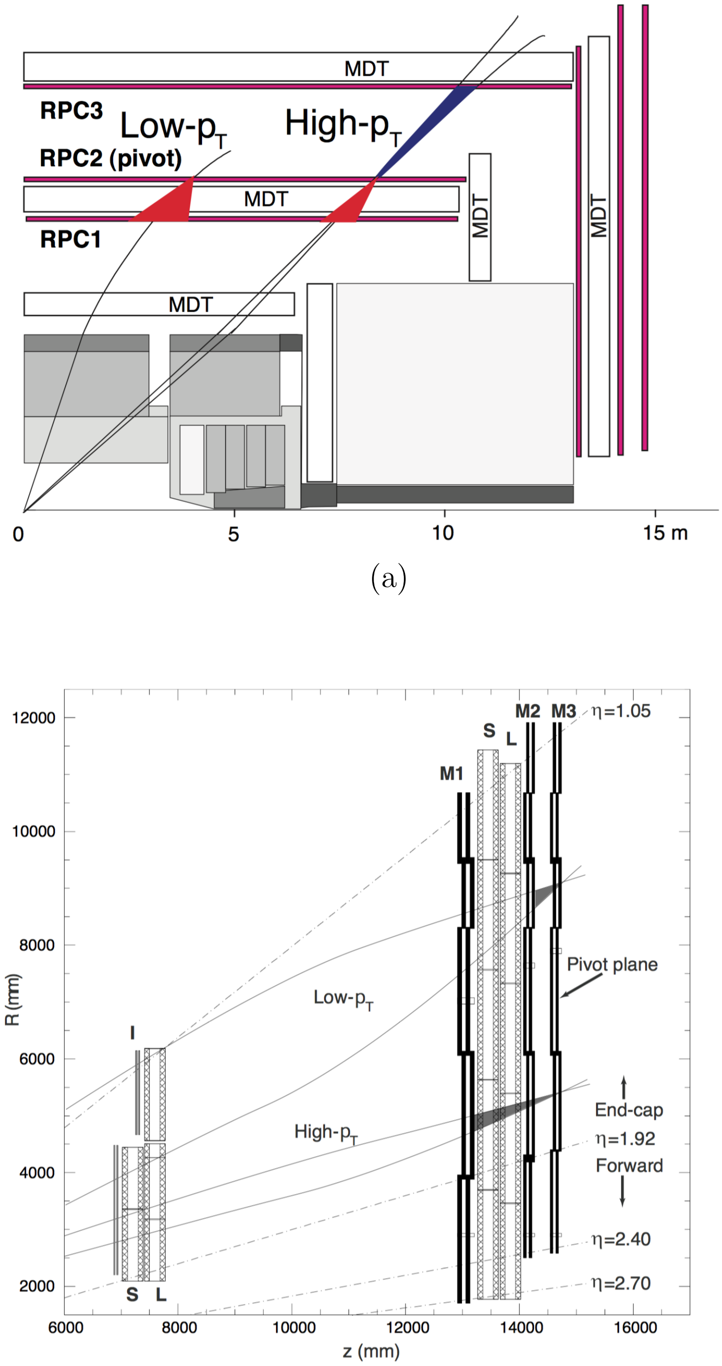
Muon trigger

The L1 muon trigger determines the Region of Interest (RoI) for muons using RPC and TGC. RoI defines the location of the muon track and only the hit information around RoI is read in downstream triggers to minimize computing time. At this level, there is no time for precise tracking and momentum measurement owing to the very tight requirement for the processing speed. Therefore, six variations of the coincidence windows shown in Figure 2.19 are used for determination of muon signature in accordance with the transverse momentum. A coincidence of two or three layers of the trigger detector is required for low and high p_{T} muon decision, respectively. A degree of bend of low p_{T} muon by the toroidal magnetic field is larger than that of high p_{T} muon, hence the difference of hit positions between the layers becomes larger. The width of the coincidence window is defined to contain the bending track of muons above certain p_{T} and therefore represents the p_{T} threshold. The thresholds coincidence windows for p_{T} threshold of 4, 6, 10, 11, 15, and 20 GeV are registered in the L1 trigger system for the decision.

The HLT algorithm uses hits on not only the MDT and CSC but also ID tracks around a RoI determined in the L1. At the beginning of this step, called as TrigL2MuonStandAlone (muonSA), p_{T} of the muons are estimated by the bending angle at the MS layers without the detailed tracking. Hence, looser p_{T} thresholds than the required value at the end of the HLT is applied because of the low resolution of this estimation. In the next step, it is required to have a reconstructed track in the ID with the same direction and momentum, and the combined muon track is defined using the MS and ID track parameters. This step is named as TrigL2MuonCombined (muComb), whose p_{T} resolution is improved by the ID tracking. The thresholds on muComb are tighter than that on the muonSA. TrigMuonEventFilter (muonEF) makes the final decision of the muon trigger based on precise tracking with almost the same resolution as the offline reconstruction algorithms. Although muonSA and muComb work only around RoI, muonEF algorithms can make use of the muons even outside of the RoI. This function allows to recover the inefficiency at the previous steps for the multi-muon triggers. For example, a trigger chain with L1 for one 20 GeV (`L1_MU20`) and HLT for one 20 GeV and additional one 8 GeV muon (`HLT_mu20_mu8noL1`) is more efficient than trigger with `L1_MU20_MU6` due to inefficiency of `L1_MU6` efficiency for the second muon.

$E_{\mathrm{T}}^{\mathrm{miss}}$ trigger

The $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger is not based on the RoI, but defined as one variable per event. $E_{\mathrm{T}}^{\mathrm{miss}}$ on the L1 trigger step is calculated only by the energy deposition in the calorimeter. Therefore, muons are also recognized as $E_{\mathrm{T}}^{\mathrm{miss}}$ at this level. At the L2, refinements can be made to correct using p_{T} of muons, however inefficiency by the mis-reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ at L1 cannot be recovered in this step. The resolution of the $E_{\mathrm{T}}^{\mathrm{miss}}$ is affected by the pile-up events especially in the trigger level, because it is difficult to calculate the energy balance of the explicit interaction point in



(a)

(b)

Figure 2.19: Schema of the coincidence window in (a) barrel and (b) endcap [62]. The required number of layers and the width of coincidence window are set depending on the p_{T}.

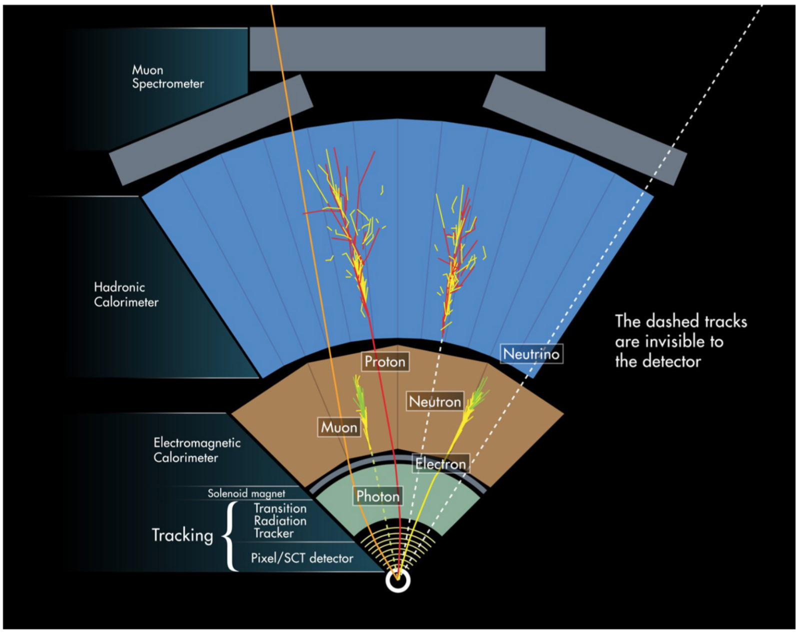


Figure 2.20: Strategy illustration of the particle identification in ATLAS [65]. Charged particles are shown as solid lines, while neutral particles are shown as dashed lines.

case more than one interactions occur in the same bunch crossing. Hence, the threshold for $E_{\mathrm{T}}^{\mathrm{miss}}$ is set at relatively high value at 100 GeV for $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger while lower threshold can be used in combination with other signatures.

2.2.5 Object reconstruction and identification

The general flow of the reconstruction and identification of the objects for physics analysis is described here. The details of the reconstruction process of muons, jets and $E_{\mathrm{T}}^{\mathrm{miss}}$ are shown in a later section for analysis (Section 5.2) because the requirements in these process are optimized and specified for each analysis.

The detectable SM particles, photons, hadrons, electrons, and muons, are reconstructed and identified in the ATLAS detector. $E_{\mathrm{T}}^{\mathrm{miss}}$ is also defined using momentum balance of these detectable particles. Figure 2.20 shows the overview of the methods to detect and identity each kind of particle. Charged hadrons, electrons and muons, are generally reconstructed as tracks in the ID and detected at the hadron calorimeter, EM calorimeter, and the MS, respectively. On the other hand, neutral particles like neutrons and photons are not reconstructed as tracks but detected only in the hadron and EM calorimeter. Neutrinos are neither detected by trackers nor calorimeters, therefore recognized as a component of $E_{\mathrm{T}}^{\mathrm{miss}}$. Vertices of the interaction point and the decay of the particles are also reconstructed in the ID. The vertex is defined as a crossing point of emitted

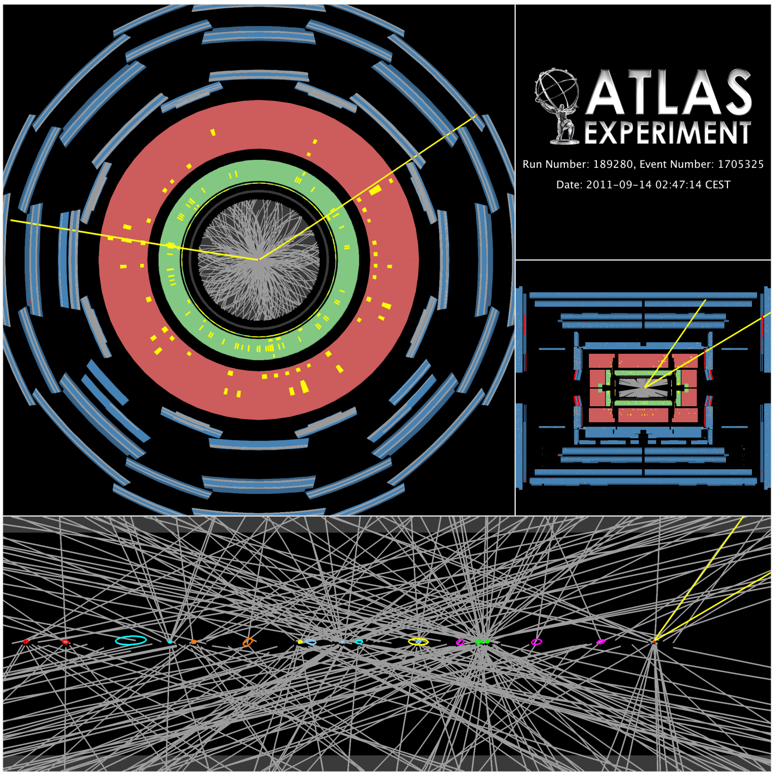


Figure 2.21: Display of the event with large number of tracks with 20 vertices. Left top picture shows the x-y view and right top shows the r-z view of the event. Bottom figure is zoomed around beam line and interaction point in the r-z view.

multi-tracks from the point. This is used to ensure that the tracks are from proton-proton collision as well as to measure the flight length of decaying particle from the distance between the primary and secondary vertices. Even in hard environment with large number of tracks in the same bunch crossing, several vertices can be reconstructed as shown in Figure 2.21.

Chapter 3

Signal Feature and Analysis Strategy

In order to realize an effective search for the $\tau \to 3\mu$ signature, we need to understand the characteristic feature of the signal that discriminates it from background phenomena in the SM. Not only the selection at the analysis level, but the optimization of the trigger requirement is also necessary to collect sufficient signal for the search of rare process. The τ leptons are generated by decays of several kinds of parent particle in ATLAS, and therefore analysis strategy needs to be optimized separately for each signature to maximize the sensitivity. In this chapter, the feature of the target signal and the strategy to search for the signature from each τ production process are described.

3.1 $\tau \to 3\mu$ signature in the ATLAS experiment

One of the most important aspects of this analysis is the statistics of the τ leptons. This directly affects the sensitivity for the search of rare process. The W boson, decaying into τ and ν_τ with Branching Ratio ($\mathcal{BR}$) of approximately 11%, is a significant source of τ production in this analysis. The τ leptons are also emitted by the Z boson decay, however its production cross section is about one-tenth of the W boson[1]. On the other hand, Heavy Flavor (HF) hadrons also decay into τ with associated particles. Hence, the search in the ATLAS experiment focused on the W and HF-originated τ. Typically, the number of the HF-originated τ is 1000 times larger than the W boson whereas τ from W boson decay typically has higher momentum and hence the selection is easier. As shown in Table 3.1, expected numbers of τ leptons are large enough for the target sensitivity of the $\mathcal{BR}$ of 1.0×10^{-9} (Section 1.2) in principle, while the number is suppressed by trigger and selection efficiencies. These signatures should be searched for by different analysis flow owing to the significant difference of their kinematics and the background sources. In this section, the signal feature and analysis flow are

---

[1] and $\mathcal{BR}$to $\tau^-\tau^+$ is even lower (3.3%)

Table 3.1: Sources of τ leptons and their expected numbers (production cross-section) with the typical momentum of emitted τ leptons in the assumption of the luminosity of $100\,\mathrm{fb}^{-1}$ and the c.m.s. energy of 13 TeV.

Production process	Expected number (cross-section)	Typical momentum
$W \to \tau\nu_\tau$	2.0×10^9 $(2.0 \times 10^1$ nb)	~ 40 GeV
$Z \to \tau\tau$	3.0×10^8 $(3.0 \times 10^0$ nb)	~ 45 GeV
B meson $\to \tau + X$	6.0×10^{11} $(6.0 \times 10^3$ nb)	$0 - 20$ GeV
D meson $\to \tau + X$	7.0×10^{12} $(7.0 \times 10^4$ nb)	$0 - 20$ GeV
$D_s \to \tau\nu_\tau$	6.0×10^{12} $(6.0 \times 10^4$ nb)	$0 - 20$ GeV

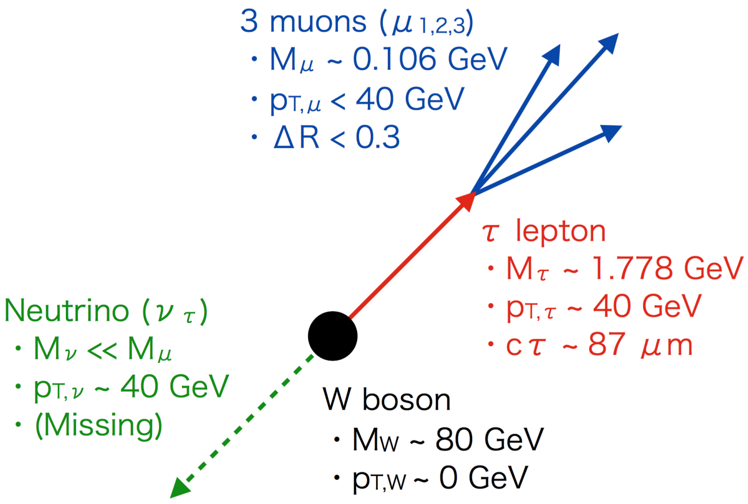


Figure 3.1: Simplified illustration of the W-originated $\tau \to 3\mu$ event, composed by the collimated muons from displaced secondary vertex and large missing energy of neutrino.

explained for the signature via W and HF decay separately.

3.1.1 W-originated signature

The production process of τ leptons via W decay is illustrated in Figure 3.1. This τ lepton is a boosted object with momentum of typically 40 GeV. The muons generated by a decay of such boosted particle are collimated because the mass of τ lepton is significantly smaller than the momentum. Typically, the ΔR (Equation 2.7) of the decay products of the boosted parent particle ($\Delta R_{\mathrm{boost}}$) of particles can be estimated as

$$\Delta R_{\mathrm{boost}} \sim \frac{2 \cdot M_{\mathrm{parent}}}{p_{\mathrm{T,parent}}}, \tag{3.1}$$

where M_{parent} and $p_{\mathrm{T,parent}}$ are the mass and momentum of the parent particle. ΔR between these muons from W-originated τ decay are less than 0.3 and mostly

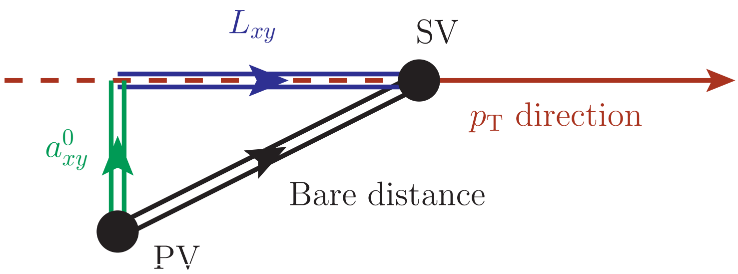


Figure 3.2: Illustration of the L_{xy} and a^0_{xy} definition on the x-y plane. Decay of the long life time particle should have a long L_{xy} and a small a^0_{xy}.

around 0.1.

The flight length of τ lepton is measurable and hence a useful feature to discriminate from accidental combination of three muons. The Secondary Vertex (SV) of the cLFV decay is reconstructed from the muon tracks and displaced from the Primary Vertex (PV) defined as W production point. To account for the difference of the direction from the PV to SV and the momentum vector, the L_{xy} and a^0_{xy} are defined as

$$L_{\mathrm{xy}} = (\overrightarrow{SV} - \overrightarrow{PV}) \cdot \frac{\overrightarrow{p_{\mathrm{T},\tau}}}{|p_{\mathrm{T},\tau}|} \tag{3.2}$$

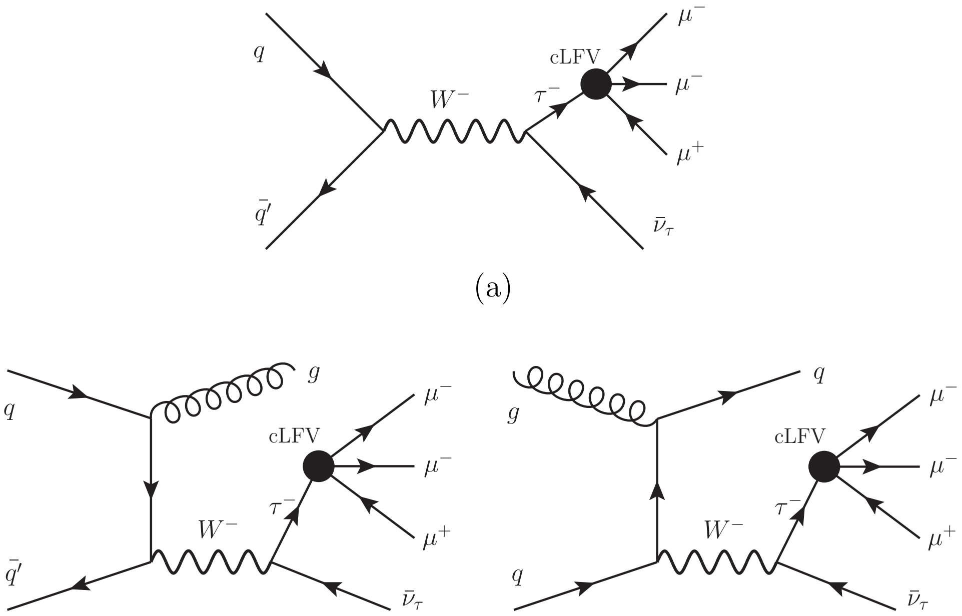
$$a^0_{\mathrm{xy}} = \sqrt{(\overrightarrow{SV} - \overrightarrow{PV})^2 - L^2_{\mathrm{xy}}}, \tag{3.3}$$

where $\overrightarrow{SV}$ and $\overrightarrow{PV}$ are the position vectors of the SV and PV, and $p_{\mathrm{T},\tau}$ is the transverse momentum of the reconstructed τ candidate from the three muons. L_{xy} is the projection of the distance from PV to SV on $x-y$ plane to the parallel direction to p_{T} of the reconstructed particle, whereas a^0_{xy} is the projection to the perpendicular direction to the p_{T} of the particle as shown in Figure 3.2. Therefore, L_{xy} and a^0_{xy} represent the flight length of the particle and the degree of agreement of the p_{T} direction and the distance direction.

Large $E_{\mathrm{T}}^{\mathrm{miss}}$ induced by the missing neutrino is also a characteristic feature of the signal event. This neutrino is emitted backward against the τ lepton, so that the $\Delta R_{\tau,\mathrm{E_T^{miss}}}$ should be close to 3.14. Moreover, the invariant mass reconstructed by the τ lepton and the missing energy of the neutrino ($M_{\tau,\mathrm{E_T^{miss}}}$) should be reconstructed as M_{W}. However, only the transverse mass, denoted as M_{T} can be reconstructed in a pp collision because the longitudinal component of the missing energy cannot be calculated. The transverse mass of a visible particle and $E_{\mathrm{T}}^{\mathrm{miss}}$, denoted as $M_{\mathrm{T,part}}$, is defined as

$$M_{\mathrm{T,part}} = \sqrt{2p_{\mathrm{T,part}}E_{\mathrm{T}}^{\mathrm{miss}}(1 - \cos\Delta\phi_{\mathrm{part,E_T^{miss}}})}, \tag{3.4}$$

where $p_{\mathrm{T,part}}$ and $\Delta\phi_{\mathrm{part,E_T^{miss}}}$ are the p_{T} of the visible particle and $\Delta\phi$ between the visible particle and $E_{\mathrm{T}}^{\mathrm{miss}}$. Since the η component of the missing energy cannot be



(a)

(b) (c)

Figure 3.3: Dominant production channels of the W boson. (a) is a production without any jets. (b) and (c) describe the production associated with a jet. The production cross-section of (a) and (c) are similar to each other, and larger than (b). These associated jets are not collinear to three muons from τ decay kinematically.

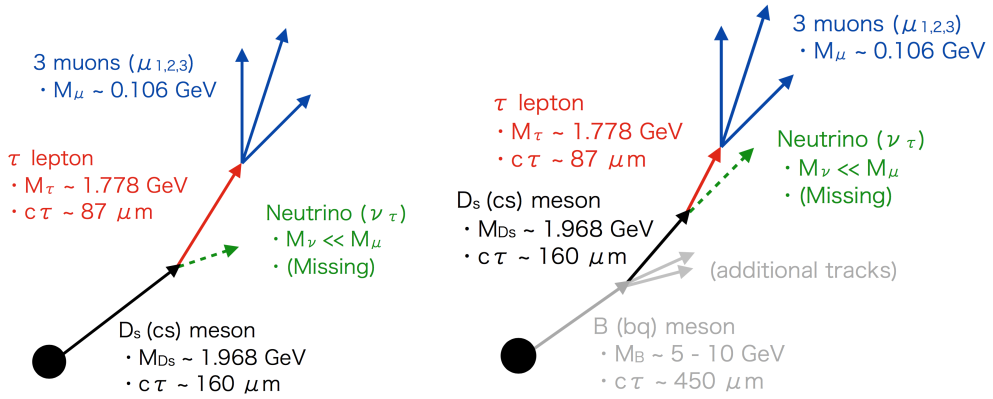
determined, this parameter is used as a representative of the mass reconstructed with the visible and invisible particles. This distribution makes a gentle bump below M_{W} by the projection to the xy plane instead of narrow peak at M_{W}.

An additional important feature of the signal event is that collinear hadronic jets along τ are absent. Although W bosons can also be produced with associated jets as shown in Figure 3.3, momentum of these jets are relatively small and non-collinear to τ. Hence, low jet multiplicity and isolated τ from jets are expected.

In order to calculate $\mathcal{BR}(\tau \to 3\mu)$, at least one relatively well measured or known channel is necessary for the normalization. In the W boson case, Lepton Universality is a good assumption between $W \to \mu\nu$ and $W \to \tau\nu$. Following this assumption, the number of muons from W decay with the same dataset is used for $\mathcal{BR}$ calculation. The $\mathcal{BR}(\tau \to 3\mu)$ in this case is simply defined as

$$\mathcal{BR}(\tau \to 3\mu) = \frac{N_{W\to\tau\nu\to 3\mu\nu}}{N_{W\to\mu\nu}}, \tag{3.5}$$

where $N_{W\to\tau\nu\to 3\mu\nu}$ and $N_{W\to\mu\nu}$ are the bare number of generated signals.



(a) (b)

Figure 3.4: Topology of the τ production from D_s decay. This meson is generated by (a) the hadronization of the prompt charm quark and (b) the decay of heavier charm or bottom hadrons. In (a), secondary vertex of the τ decay is displaced more than W-originated case by the flight length of the D_s. On the other hand in (b), flight length of the parent particle make the distance between SV and PV even longer, although other hadronic particles are around the three muons. The flight length and mass, written for the B meson, are typical values and different depending on the kind of hadron.

3.1.2 HF-originated signature

As shown in Table 3.1, number of τ leptons from B and D meson decays is nearly 1000 times larger than those from W boson decays. Among these mesons, the most important production process is via $D_s \to \tau\nu$, whose $\mathcal{BR}$ is $5.55 \pm 0.24\%$. This simple leptonic decay channel is dominant and the cleanest of the HF-originated signatures. Therefore, a search for $D_s \to \tau\nu \to 3\mu\nu$ is the target in this part.

$D_s^{\pm}$, composed of a charm (anti-charm) quark and an anti-strange (strange) quark, is generated not only by hadronization, but also by decay of heavier charm hadron and bottom hadron as shown in Figure 3.4. Momentum of the τ lepton from decay of D_s is much lower than W-originated case, however L_{xy} (flight length) tends to be longer as D_s decays after $\sim 160\ \mu$m flight. $E_{\mathrm{T}}^{\mathrm{miss}}$ from the D_s decay is not large enough to be recognized as a clear signal. A major advantage of this channel is the large statistics of τ leptons, while the only significant features of the signal kinematics against the background are the characteristic flight lengths of D_s and τ lepton. Hence, it is more important in HF-originated analysis to utilize the SV effectively.

$\mathcal{BR}(\tau \to 3\mu)$ is calculated in the same logic as the W-originated analysis. The normalization channel in this analysis is not $D_s \to \mu\nu$, since the lepton universality does not work in the few GeV scale. Instead, $D_s \to \phi\pi \to \mu\mu\pi$ is used as the alternative normalization channel. $\mathcal{BR}(\tau \to 3\mu)$ is calculated as

$$\mathcal{BR}(\tau \to 3\mu) = \frac{N_{D_s \to \tau\nu \to 3\mu\nu}}{N_{D_s \to \phi\pi \to \mu\mu\pi}} \cdot \frac{\mathcal{BR}(D_s \to \phi\pi \to \mu\mu\pi)}{\mathcal{BR}(D_s \to \tau\nu)}, \tag{3.6}$$

where $\mathcal{BR}(D_s \to \phi\pi \to \mu\mu\pi)$ and $\mathcal{BR}(D_s \to \tau\nu)$ were determined precisely by the previous experiments. Therefore, $\mathcal{BR}(\tau \to 3\mu)$ is determined by the observed numbers of $D_s \to \tau\nu \to 3\mu\nu$ and $D_s \to \phi\pi \to \mu\mu\pi$ events, which are denoted by $N_{D_s \to \tau\nu \to 3\mu\nu}$ and $N_{D_s \to \phi\pi \to \mu\mu\pi}$, respectively. This normalization channel is also used to ensure the robustness of this analysis. For example, a long flight length and a small χ^2 of the SV fitting were expected from the simulation of $D_s \to \tau\nu \to 3\mu\nu$ signals, although the reliability of these parameters should be confirmed by the data. The $D_s \to \phi\pi \to \mu\mu\pi$ events provide such validation due to their similar kinematics. In addition, the efficiencies of the trigger and reconstruction should also be similar to $D_s \to \tau\nu \to 3\mu\nu$ events. Hence, this channel is useful for this analysis.

3.1.3 Background components

To optimize the signal selection and evaluate the sensitivity, background components must be understood precisely. The Standard-Model events which contain three muons are candidates for the background of this analysis. These can be background in the case the invariant mass of the three muon system is in the signal window.

Most multi-muon events are generated by decays of heavy-flavor hadrons, while the decay products of charged light flavor mesons composed of u, d, and s quarks do not include multi-muons due to their lighter masses. For example, as shown in Figure 3.5(a), $D_s \to \phi\mu\nu \to 3\mu\nu$ decay ($\mathcal{BR}$= 7.1×10^{-6}) can produce three muons with a small invariant mass around τ lepton. These three muons, generated at the vertices away from PV, hence can mimic the signal. While this background component has an opposite-sign muon pair from the ϕ resonance decay, it has smaller $E_{\mathrm{T}}^{\mathrm{miss}}$ than $W \to \tau\nu \to 3\mu\nu$ events, and generally has a collinear hadron jet in the three muon direction. Therefore, the contamination of this component can be suppressed by the requirements on these features.

The second dominant background component is a combinatorial one shown in Figure 3.5(b). This occurs when muons from different sources are recognized as a signal coming from a shared origin, accidentally. Unlike the multi-muon events from HF hadron decay, they tend to have a short flight length and poor probability of the SV fitting. Figure 3.5(c) shows another candidate of the background involving a mis-identified muon (fake muon). Although muon identification by the hits on the MS is almost 100% reliable, muons from π and K meson can be recognized as a prompt muon. In this case, track parameters in the ID and MS are mismatched due to the kink of the decay. As a result of such kink tracks, the track reconstruction probability is poor for the fake muons. Hence, these can be

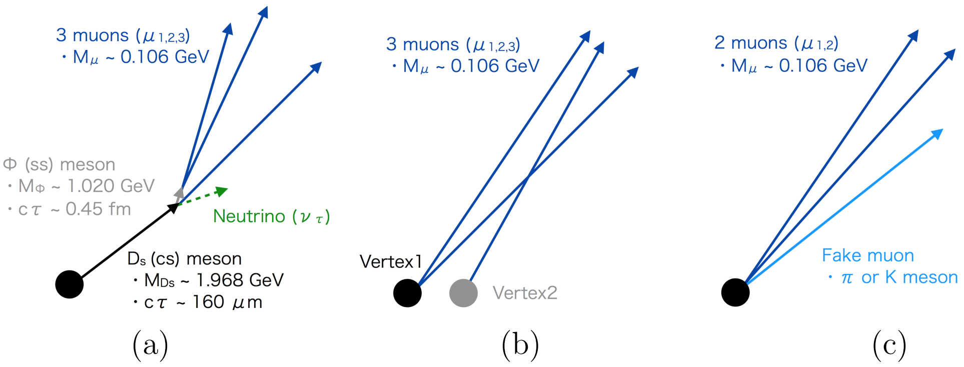


Figure 3.5: Dominant processes of the background events from (a) heavy flavored meson decay, (b) combinatorial, and (c) fake muons.

suppressed by the track quality selection based on their detector hits and by the fitting probability.

3.2 Analysis strategy in the experiment

This analysis aims to show the first result and the possibility of its extension at the ATLAS experiment. The strategies of the trigger for the data acquisition, the event selections, and the evaluation of the branching ratio are considered first. In summary, the analysis procedure consists of a search for $\tau \to 3\mu$ signature and a measurement of the number of generated τ leptons. However, in order to evaluate the significance of the signal or to set an appropriate limit, experimental and technical aspects should also be considered in the actual analysis, especially for the trigger and reconstruction in the ATLAS experiment. First, p_{T} thresholds of muon triggers are set sufficiently high to acquire all the generated $\tau \to 3\mu$ signatures. The way to discriminate the signature from the huge number of the HF background (Section 3.1.3) is also an important issue. Hence, the effective analysis channel in the run depends largely on the trigger configuration and background cross-section of the run.

3.2.1 Experimental challenge

Although experiments at the LHC have a merit of a huge number of τ leptons produced in pp collisions for the search, several challenges still exist in the experiment in practice.

A large number of background events with multi-muons by the QCD interaction are the most essential problem at the hadron colliders. Generally, these events are suppressed by the requirements for the muon momentum thresholds at the trigger level. However it is difficult to avoid critical efficiency loss for the target signature especially for the HF-originated signature, with low p_{T} muons by the online selection. The inefficiency is also caused by the case that muons are too close in space to resolve them as separate muons. This issue is crucial for this search because three muons from τ decay are always highly collimated by the high momentum thresholds for muons. Although these issues are more significant at the trigger level because of the low resolution of momentum and position due to limited computing time, reconstruction also has slight inefficiency for the collimated muons. Therefore, the efficiencies of the trigger and reconstruction should be carefully checked in this analysis. For the trigger study, a dedicated trigger for the target signal has been developed and used in ATLAS. The efficiency of the multi-muon trigger could be estimated by the product of single-muon trigger efficiencies in principle. However, this estimation was not suitable for this analysis because of additional inefficiency of close-by muons. Therefore, it is critical in this study to evaluate inefficiency of such collimated muon pairs.

The similarity of the background events by the QCD to the target signal is another important issue. Simple cut-based selections for the muons and reconstructed τ candidates were not enough to discriminate target events from the backgrounds while keeping high signal efficiency. This is because the signal and background have different shapes of the measurable kinematic variables while they are not clearly separated. In order to improve the signal and background separation, the Multi-Variate-Analysis (MVA) selection, which can evaluate the signal likelihood of an event using several variables comprehensively, was employed. Especially for HF $\to \tau \to 3\mu$ analysis, this technique is essential because the HF $\to \tau \to 3\mu$ signature does not have large $E_{\mathrm{T}}^{\mathrm{miss}}$ and is not isolated from hadron jets. This MVA technique is also effective for the $D_s \to \phi\pi \to \mu\mu\pi$ analysis and thus employed.

3.2.2 Targets in Run1 and Run2

In the Run1 period, none of the dedicated triggers were implemented. Since the HF $\to \tau \to 3\mu$ signature was hard to be triggered effectively without a dedicated trigger, the search requires more complicated analysis to discriminate signal from background. By this reason, $W \to \tau\nu \to 3\mu\nu$ was the main target in the Run1 analysis. The analysis procedure and results in Run1 are described in Part II.

On the other hand, dedicated triggers were developed and implemented in the Run2 period. Accordingly, HF $\to \tau \to 3\mu$ analysis could be carried out from this period by referring to the $W \to \tau\nu \to 3\mu\nu$ analysis procedure in Run1. The prospects for $W \to \tau\nu \to 3\mu\nu$ and HF $\to \tau \to 3\mu$ analyses with the 2016 data are described in Part III.

Part II

Analysis using Run1 data with $\sqrt{s} = 8\,\mathrm{TeV}$

Chapter 4

Data and Monte-Carlo Simulation

In Run1 data taken in 2012, W-originated $\tau \to 3\mu$ was the target of the search. The details of the acquired data in pp collision with $\sqrt{s} = 8\,\text{TeV}$ and triggers used for data taking are described in the first section of this chapter. Simulation samples were also generated to estimate the number of signal, to validate the distribution of the background, and to evaluate the difference of the simulation with the data. Details of the simulation samples are shown in the second section.

4.1 Data in pp collisions with $\sqrt{s} = 8\,\text{TeV}$

4.1.1 Status of 2012 run

LHC was running at the center-of-mass energy of 8 TeV in the 2012 period. The amount of data is represented by the time-integrated luminosity which is the sum of instantaneous luminosity in the bunch crossing. The instantaneous luminosity is defined by cross-section (σ_{inel}) and rate (R_{inel}) of the inelastic scattering events as

$$\mathcal{L} = \frac{R_{\text{inel}}}{\sigma_{\text{inel}}}. \tag{4.1}$$

These values of the inelastic interactions are measured by the van der Meer (vdM) method [66] using dedicated detectors [67]. Figure 4.1 shows the time evolution of the integrated luminosity in 2012. The bare luminosity of the pp collisions reached $22.8\,\text{fb}^{-1}$, of which $21.3\,\text{fb}^{-1}$ of data was recorded by the ATLAS trigger system. This efficiency loss is caused by dead or stopped time of the ATLAS detector system. On the other hand, the actual amount of analyzed data is approximately $20.3\,\text{fb}^{-1}$ due to requirement of stable performance of the detector system. The uncertainty of this luminosity determination is estimated to be ±2.8% during this year [67].

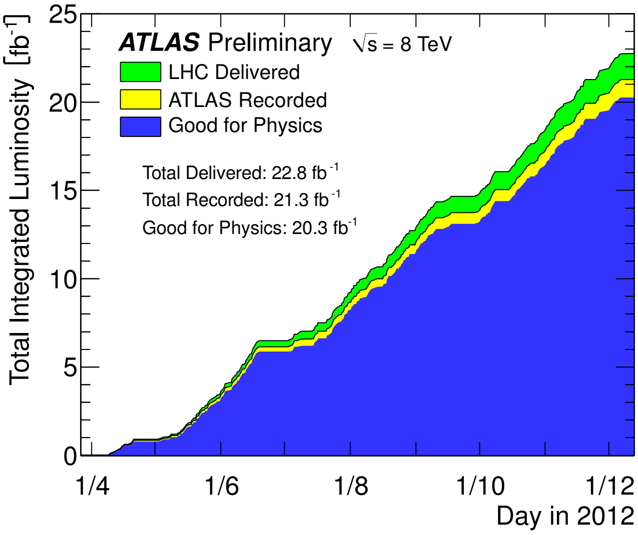


Figure 4.1: Evolution of the integrated luminosity in 2012 run as a function of the date [68]. The ATLAS experiment recorded 21.3 fb^{-1} in the delivered 22.8 fb^{-1} collision.

4.1.2 Trigger

As mentioned in Section 3.2.2, the target signature consists of three muons and large $E_{\mathrm{T}}^{\mathrm{miss}}$. These six trigger chains with multi-muon (and $E_{\mathrm{T}}^{\mathrm{miss}}$) requirement were employed as efficient chains for this analysis: `HLT_2mu13`, `HLT_3mu4T`, `HLT_3mu6_msonly`, `HLT_mu18_tight_mu8_EFFS`, `HLT_mu18_tight_2mu4_EFFS`, and `HLT_2mu8_EFxe80_tclcw`. The events which are fired by at least one of these chains are derived from whole acquired data, and used in the analysis procedure. Hence, the efficiency has to be evaluated as the inclusive efficiency for combination of these, which will be described in Section 7.2. In this period, the dedicated triggers using the topological feature of the signal were not developed. Therefore, only the chains designed with simple requirement for muons and $E_{\mathrm{T}}^{\mathrm{miss}}$ without any relations between these were used in 2012.

4.2 Monte-carlo simulation

The Monte-Carlo (MC) simulation for the $W \to \tau\nu \to 3\mu\nu$ signature is necessary for determination of the event selection efficiency for the search. In addition to the signal sample, several simulation samples for the validation are also important to check the robustness of the analysis procedure. As the first step of the simulation, the target events are generated using the proper configuration of physics models and Parton Distribution Function (PDF) set. These are used as the inputs for the ATLAS detector simulation based on GEANT4 package [69, 70]. This output is derived in the same formats as that of the actual data taking. The physics objects are reconstructed in the same way as for the data, based on the response of detector simulation. Therefore, the simulated samples can be compared with the data directly, although the difference in the performance must be evaluated.

4.2.1 Signal sample of $W \to \tau\nu \to 3\mu\nu$

A signal sample of $W \to \tau\nu \to 3\mu\nu$ is produced by a generator for parton-level (Pythia8 [71]) with the configuration as AU2 [72] and the PDF set of MSTW2008LO [73]. Hadronization of the quarks(gluons) and the hadron decay are also simulated by Pythia8 package, based on the leading order calculation. The $\mathcal{BR}$ of the $\tau \to 3\mu$ decay is assumed to be 100%[1] without any specific BSM physics model in the event generation in order to search this generally. Finally, the cross-section of the signature is determined by the number of $W \to \tau\nu$ events and $\mathcal{BR}$ of the τ decay. The production cross-section of $W \to \tau\nu$ is confirmed by the data (Section 6.1), and $\mathcal{BR}(\tau \to 3\mu)$ is considered as a free parameter representing the sensitivity for the $\tau \to 3\mu$ signature. The number of generated events of this sample was 10^5.

4.2.2 Background samples

In order to emulate the background events from the heavy flavor hadron decays, events composed of muons from the parton-level initial states of $b\bar{b}$ and $c\bar{c}$ were generated independently. For the $b\bar{b}$ background, two samples are prepared; $b\bar{b}$ originated events composed of at least one muon whose p_{T} is above 15 GeV, and at least two muons whose p_{T}'s are above 4 GeV. On the other hand, $c\bar{c}$ background events with one muon with $p_{\mathrm{T}} > 15$ GeV are also generated. These three samples are used only for validation of the background distribution shapes, but not for the direct estimation of the number of background events. This is because it is difficult to simulate all the components of backgrounds, especially the accidental background explained in Section 3.1.3. Therefore background rate is estimated by the sideband data in the invariant mass distribution. The Pythia8 tuned for heavy flavor physics (Pythia8B) with AU2 tune and the MSTW2008LO PDF set is used for the generation of the events in the same way as the signal sample.

In order to validate the background contribution of multi-jet and Z, W associated events, inclusive multi-jet samples, $W \to \mu\nu + jets$ and $Z \to \mu\mu + jets$ were also produced. However these events have negligible contribution to low invariant mass events to be considered as background. Hence, these were discarded from the background estimation and validation. The background samples are listed in Table 4.1.

---

[1] The SM decays of τ leptons were not simulated since contamination of these decays to the signal selection is negligible

Table 4.1: Summary of the cross-sections of the background samples from the heavy flavor hadron decays.

Description	σ_{prod}[nb]
$\mathrm{b\bar{b}}, p_{\mathrm{T},\mu 1} > 15\,\mathrm{GeV}$	1.99×10^5
$\mathrm{b\bar{b}}, p_{\mathrm{T},\mu 1,2} > 4, 4\,\mathrm{GeV}$	1.15×10^5
$\mathrm{c\bar{c}}, p_{\mathrm{T},\mu 1} > 15\,\mathrm{GeV}$	8.0×10^4

Table 4.2: Summary of the cross-sections of the $W \to \tau\nu + jets$ events. These samples are used for the variation of the estimation of number of the $W \to \tau\nu$ events

Description	σ_{prod}[nb]
$W \to \tau\nu + jets$ with $N_{partons} = 0$	8.036
$W \to \tau\nu + jets$ with $N_{partons} = 1$	1.579
$W \to \tau\nu + jets$ with $N_{partons} = 2$	4.778×10^{-1}
$W \to \tau\nu + jets$ with $N_{partons} = 3$	1.3401×10^{-1}
$W \to \tau\nu + jets$ with $N_{partons} = 4$	3.526×10^{-2}
$W \to \tau\nu + jets$ with $N_{partons} = 5$	1.065×10^{-2}
Sum of $N_{partons} \leq 5$	10.27

4.2.3 Normalization samples

The $W \to \tau\nu + jets$ signature is also produced for estimation of the number of τ. This sample is generated using different parton-shower generator (ALPGEN [74] and JIMMY [75]), instead of PYTHIA8 for the signal sample. A higher order correction is applied in the calculation of this configuration. Therefore, these samples are suitable with the better accuracy of cross-section of W boson production than PYTHIA8 with the leading order calculation. These samples are separately produced by the number of associated partons in the final state, noted as $N_{partons}$. Cross-sections of $W \to \tau\nu$ events are listed in Table 4.2. The number of $W \to \tau\nu$ events is estimated by the theoretical extrapolation of the cross-section measurements at 7 TeV as described in Section 6.2. These samples are used for the variation of the estimation.

4.2.4 $J/\psi \to \mu\mu$ simulation for efficiency measurement

A large number of $J/\psi \to \mu\mu$ events are also needed to estimate the efficiency of the trigger and reconstruction. This sample is produced by PYTHIA8B in the same configuration as the heavy-flavor-based background sample generation. First, two types of samples are produced with the threshold of $p_{\mathrm{T},\mu 1} > 4\,\mathrm{GeV}$ & $p_{\mathrm{T},\mu 2} > 4\,\mathrm{GeV}$ and $p_{\mathrm{T},\mu 1} > 15\,\mathrm{GeV}$ & $p_{\mathrm{T},\mu 2} > 2.5\,\mathrm{GeV}$, respectively. This is because boosted $J/\psi \to \mu\mu$ events including at least one high p_{T} muon are used in the efficiency measurement (Section 7.1.2). The latter sample contains more events with such high p_{T} muon used for measurement of the efficiency.

The secondary J/ψ events, which is emitted from the B hadron decays, are not included in this sample.

Chapter 5

Event Selection

As described in Section 3.1.1, the characteristic feature of the signal is collimated and isolated three muons and large $E_{\mathrm{T}}^{\mathrm{miss}}$. First, physics objects are defined and reconstructed. The baseline requirement for each object is applied for reduction of fakes. The combinations of variables are also determined using each object after baseline selection. In order to maximize the sensitivity, MVA technique is employed after the loose selection for each physics object and combined variable is calculated. Hence, this event selection procedure consists of three steps: object selection; pre-selection using each object and the combination variables for MVA; and final MVA selection. In this chapter, details of these selections are explained.

5.1 Region definition

The object reconstructed by three muons, which is defined as the triplet, is the candidate of τ lepton in this analysis. The most characteristic feature is the invariant mass of the triplet (M_{triplet}). Events are divided into several samples by the M_{triplet} and MVA tuning is optimized for each region. These ranges of the regions are set depending on the expected resolution of the $\tau \to 3\mu$ signal by the simulation shown in Figure 5.1.

The region for counting the signal events for the search is defined as the Signal Region (SR), whose range is set as twice of σ_{mass} and the center is the τ lepton mass. The Sideband region (SB) and the training region for MVA are defined for the validation of the background distribution and optimization of the event selection. In order to minimize the signal contamination in the SB and the training region, the region of three times of σ_{mass} is blinded, which is larger than the SR with the same center value. The SB is mainly used for validation and comparison with data with similar mass to the signal, therefore defined by loose window.

The training region was used for MVA training, described in Section 5.3. In this chapter, data distribution in the SB region is shown together with the signal and background simulation in order to evaluate the discriminating power of the

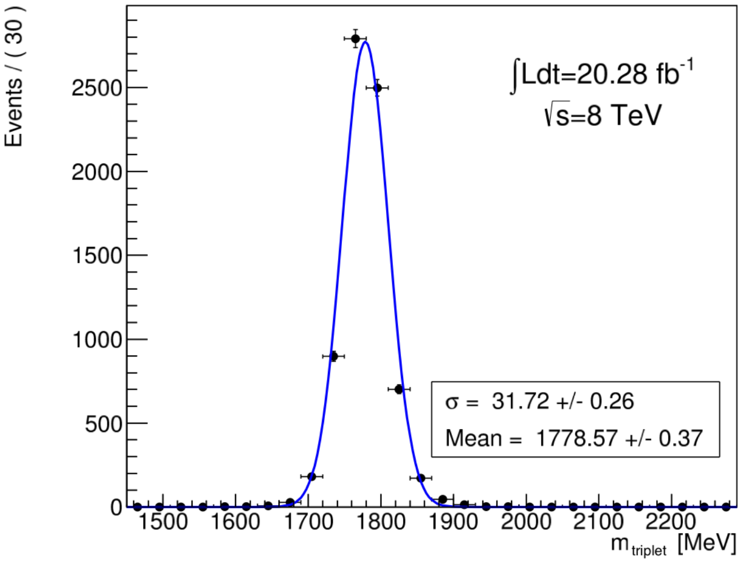


Figure 5.1: Invariant mass distribution of the triplet of the signal simulation. The mean value and the width (σ_{mass}) are estimated to be 1778 MeV and 31.72 MeV by the Gaussian fitting (blue line), respectively.

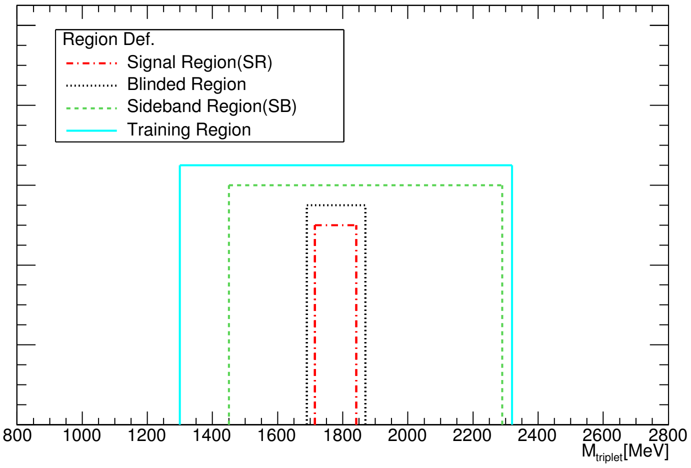


Figure 5.2: Definition of each region; signal, blind, sideband, and MVA training, as the invariant mass limitation. The reconstruction procedure and resolution of the mass is described in Section 5.2.5 and 5.3.

selection. The definitions of these regions are

- Signal Region (SR) : $[1715, 1842]$ MeV
- Blinded Region : $[1690, 1870]$ MeV
- Sideband Region (SB): $[1450, 2290]$ MeV (excluding the blinded region)
- Training for MVA : $[1300, 2320]$ MeV (excluding the blinded region)

as also shown in Figure 5.2.

5.2 Signal reconstruction and selection

Several physics objects, i.e. ID tracks, muons, hadron jets, and $E_{\mathrm{T}}^{\mathrm{miss}}$ are used in this analysis procedure. These physics objects are reconstructed based on the baseline procedure in ATLAS, while several requirements are optimized for this analysis. Especially for tracks and muons, requirements for the hit qualities and reconstructed parameters are set to maximize the efficiency for the specific kinematics with collimated and low momentum muons of the signal.

The combination variables of these reconstructed objects are also defined. The kinematic feature of the triplet, which is reconstructed as the candidate of τ lepton decaying into three muons, is also useful for event selection. It is important for event selection to consider the correlation between the objects. The definitions, reconstruction procedures, and requirements for these variables are described in this section.

5.2.1 ID Track and vertex

Charged particles are reconstructed as tracks using hits at the ID. In the first step of tracking, track candidates are reconstructed by hits on Pixel and SCT with Kalman filter [76] assuming the origin is the IP. Using TRT hits around the extended track candidates, the track parameters are determined minimizing χ^2 of the fitting with associated hits of all the detector systems. The track parameter set consists of impact parameters (d_0, z_0), track direction angles (θ, ϕ), and inverted momentum multiplied by the charge (q/p). d_0 and z_0 represent transverse and longitudinal projection of the distance from reference point to the perigee point of the track, respectively. The θ, ϕ, and momentum are also calculated at the perigee point. The covariance error matrix of these parameters are also calculated. p_T and η distributions of the signal simulation, HF-originated background and inclusive muons of the sideband region are shown in Figure 5.3. These distributions for the signal and background simulations do not have differences for discrimination, and the thresholds for these are determined by the acceptance of the detector geometry and the reconstruction. Also, the SB data shows a similar distribution to the background simulation. This means that the most background events consist of the HF-originated events.

All tracks with high enough momentum above 400 MeV are reconstructed, but there is large contamination arising from fake tracks. Fake tracks are mostly due to accidental coincidence of noise hits or hits by independent tracks. Following requirements for the number of detector hits, η region, and momentum are additionally applied for these tracks in order to reduce these fake tracks.

- If a hit on the inner most Pixel layer (B-layer) is expected, the number of hits on the B-layer should be larger than or equal to 1
- $N_\mathrm{hits,Pixel} \geq 2$
- $N_\mathrm{hits,SCT} \geq 5$
- $N_\mathrm{holes,Pixel+SCT} \leq 2$
- $N_\mathrm{hits,TRT} \geq 6$ and $N_\mathrm{outhits,TRT}/N_\mathrm{hits,TRT} < 90\%$ ($|\eta| < 1.9$)
  $N_\mathrm{hits,TRT} \geq 6$ or $N_\mathrm{outhits,TRT}/N_\mathrm{hits,TRT} < 90\%$ ($|\eta| \geq 1.9$)
- $|\eta| < 2.5$
- p_T for leading, 2nd leading, and 3rd leading tracks should be larger than 5.5 GeV, 3.5 GeV, and 2.5 GeV, respectively

$N_\mathrm{hits,detector}$ represents the number of hits at the corresponding detector. The number of dead sensors on the track path is also involved in that of hits in each detector. The number of detector holes ($N_\mathrm{holes,detector}$) is defined by that of layers with the track path at insensitive position. The hits on the edge of the TRT on the track path are defined as the outlier hits on the TRT, whose number is denoted as $N_{\mathrm{outhits,TRT}}$, used for the quality of the TRT tracking. These are required for ID tracks associated with muon, and are used for a track quality selection in the analysis procedure. The requirements for η and p_{T} of the muon tracks are determined by considering the discrimination from the fake tracks.

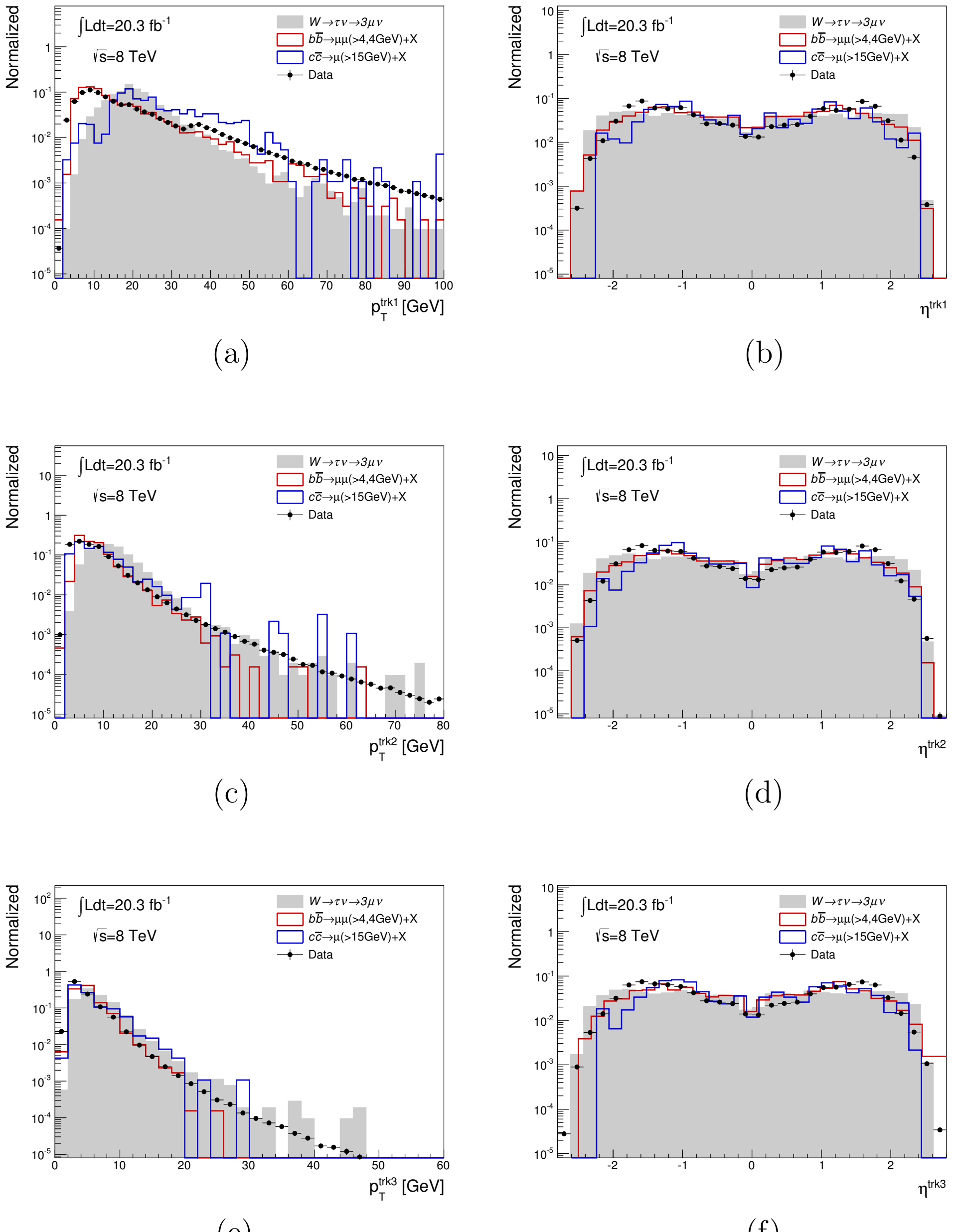
The interaction vertices are also reconstructed as a crossing point of the ID tracks. Several vertices are reconstructed in an event. This is processed at the same time of the tracking with Kalman filter. The vertex with the largest scalar sum of p_{T} of the associated tracks of among all reconstructed vertices is defined as the PV of the event. The SV is reconstructed by fitting using three tracks of signal muon with the same procedure as in the PV reconstruction. Then, the PV is also re-fitted after subtraction of the tracks used for the SV reconstruction. The track parameters of the signal muons are also re-defined with the associated vertex as the reference point, i.e. re-defined the PV or the SV. The SV is defined as the decay point of three muon decay, while the PV as the generated point of the initial particles. These vertices are used for the evaluation of the flight of the triplet using L_{xy} and a^{0}_{xy} shown in Section 3.1.1. Furthermore, the number of tracks which are used for the PV reconstruction is used to represent the strength of the interaction, and the probability of the SV is used to establish whether the three tracks cross the point.

5.2.2 Muons

There are two types of reconstruction procedure for muons. The basic procedure of muon reconstruction is performed by hits on the ID and precision detector of the MS, i.e. MDT and CSC. This muon, called combined muon, is determined as a combination of the ID track and MS stand-alone track, which is reconstructed using only MDT and CSC. In the MS stand-alone track reconstruction, segments on the detector layers which is defined as a set of the hits are used for fitting. This fitting procedure is based on an assumption that a track is emitted from the center of the detector, considering the curvature and energy deposit in the calorimeter. The precision of the ID tracks is much better than the MS stand-alone track, therefore the combined muon precision is improved by the ID track with the good identification of the muon by the MS stand-alone track. The efficiency of the muon reconstruction is understood as a combination of those at ID and MS, and the measurement procedure is described in Section 7.3.

Another reconstruction procedure is performed with the ID track and MS detector hits without tracking. In this procedure, at least one segment around extended ID track is required instead of the MS stand-alone track. This muon track, called as Muon-Associated-Track-Particle (TPa), shows worse separation from fake muons, however efficiency loss at the MS stand-alone track reconstruction is suppressed. The efficiency of TPa was also measured in Section 7.3, as well as the combined muon.

The quality selections for muons are defined separately for these types of muons. The common selection for the ID tracks are shown below:



(a) (b)

(c) (d)

(e) (f)

Figure 5.3: p_{T} and η distribution of each track in the reconstructed triplet of $W \to \tau\nu \to 3\mu\nu$ signal (gray filled), HF-originated background simulation (red and blue lines), and the sideband data (black points). Figures in the left side show p_{T} distributions for (a) leading muon, (c) 2nd leading muon, and (e) 3rd leading muon, respectively. In the right side, η distributions for (b) leading muon, (d) 2nd leading muon, and (f) 3rd leading muon are shown.

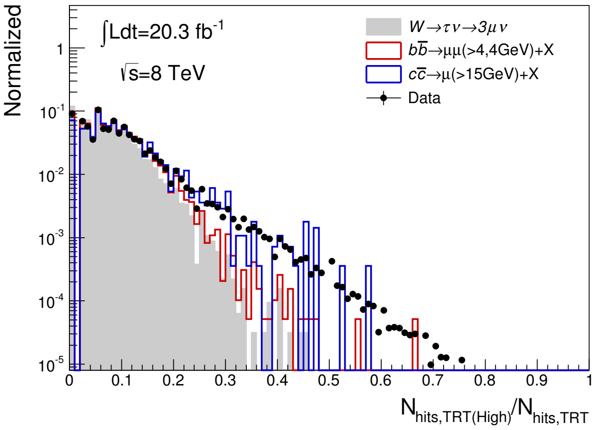


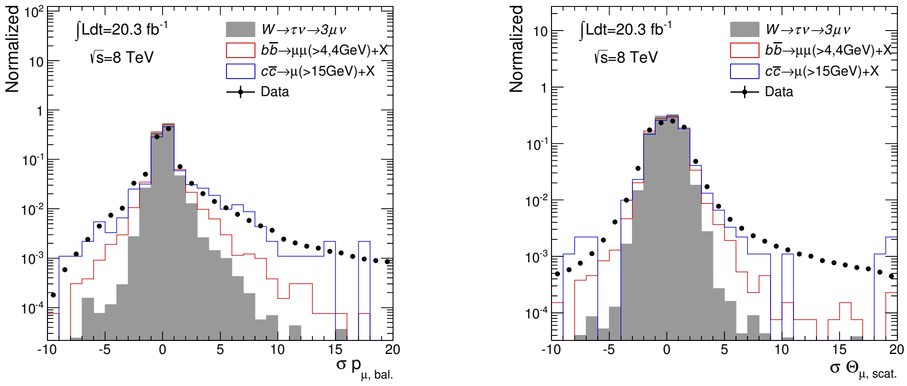
Figure 5.4: Ratio of number of hits with the high energy deposition to the total number of hits for muons in the reconstructed triplet of $W \to \tau\nu \to 3\mu\nu$ signal (gray filled), HF-originated background simulation (red and blue lines), and the sideband data (black points). Most of hits by muons from signal events are below 0.35.

- Track quality selection in Section 5.2.1
- Ratio of the number of TRT hits with high energy deposition to that of all the hits associated to the track should be smaller than or equal to 0.35

The hits with high energy deposit at TRT are specific for the electron as shown in Section 2.2.1. Hence, electron tracks are rejected by requirements of low ratio of the high TRT hits, as shown in Figure 5.4.

Specific selections are applied for combined muons using the number of hits in the muon detectors and parameters rerated to MS stand-alone tracks. The muon momentum balance significance ($\sigma p_{\mu,\text{bal.}}$) is defined as the difference between the momentum measured by ID and MS, concerning the weight by accuracy. In principle, ID and MS momentum should match with each other for muons, i.e. should sharply peak around zero. The tail of this distribution is made by the kink track by the pion or K meson. The muon scattering angle significance ($\sigma\Theta_{\mu,scat.}$) represents the difference of the track directions reconstructed in ID and MS. Therefore, this tail is also understood by the kink tracks. Selections are applied for rejection of the fake muons by pions and K mesons. Figure 5.5 shows the muon momentum balance significance and the muon scattering angle significance. Figure 5.6 shows the distribution of number of hits on precision detector (MDT and CSC) and number of trigger layers (RPC and TGC) with hits for each muon. The selections for muons are commonly set to make the fraction of signal muons higher than fake-like muons in data and HF background simulation. The following selection criteria are applied for the combined muons:

- $-3 < \sigma p_{\mu,\text{bal.}} < +4$



(a) (b)

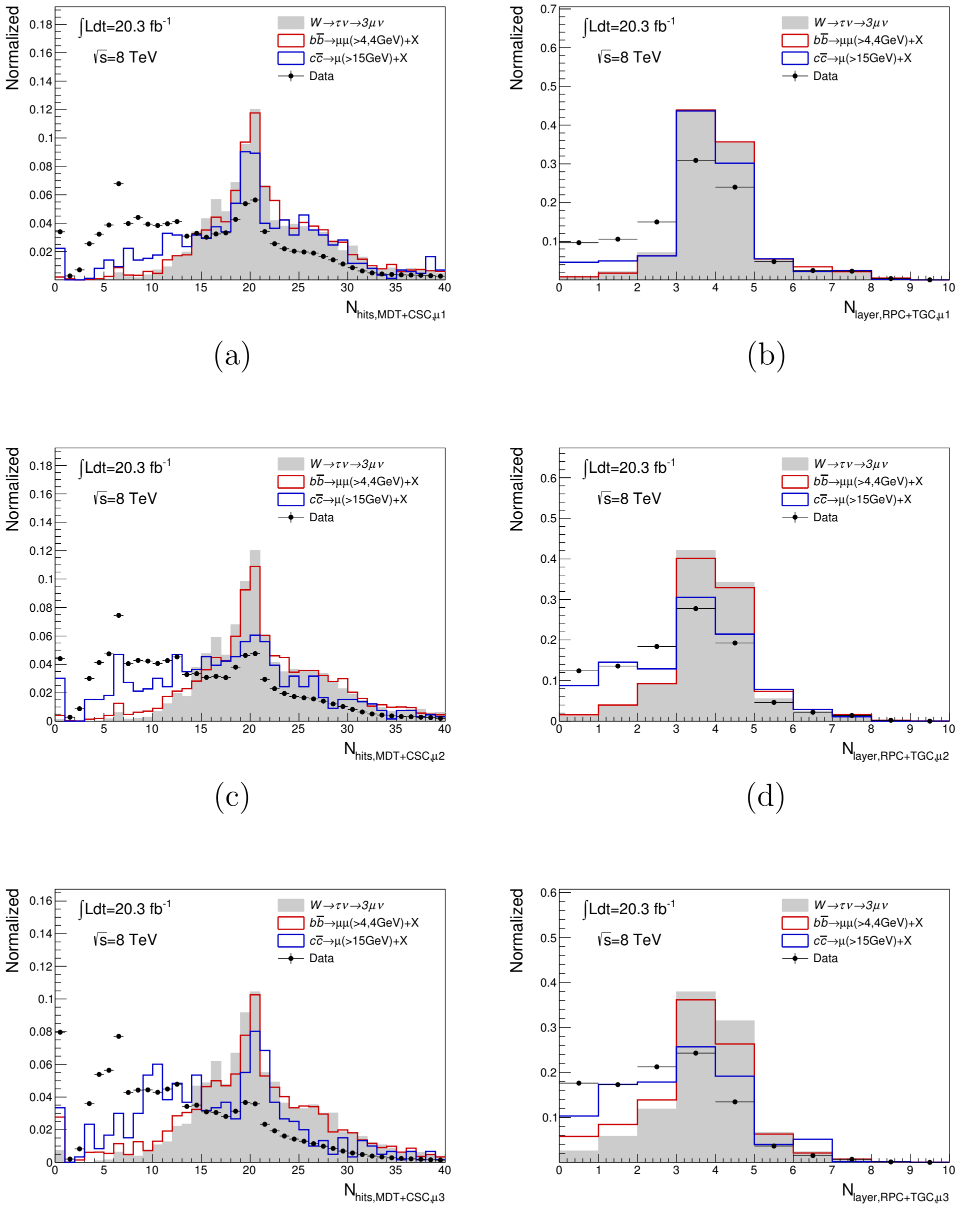
Figure 5.5: (a) Muon momentum balance significance and (b) muon scattering angle significance for the simulation and the data. $W \to \tau\nu \to 3\mu\nu$ signal, HF-originated background simulation, and the sideband data are shown as a gray filled histogram, red and blue lines, and black points, respectively.

- $|\sigma\Theta_{\mu,\mathrm{scat.}}| < 4$
- $N_{hits,MDT+CSC} \geq 12$
- $N_{hits,RPC+TGC} \geq 2$

For TPa muons, tighter selections are applied for the hits on the MS:

- $N_{hits,MDT+CSC} \geq 13$
- $N_{hits,RPC+TGC} \geq 3$
- The track must have either TGC or RPC hits (η or ϕ) to exclude shared tracks in TGC-RPC transition region.

The probability of track (p-value$_{trk}$) is determined by the χ^2 fitting of hits on the ID track. The quality of the ID tracking is more important because reconstruction procedure is based on ID tracking for the TPa muons. The muon sample with small p-value$_{trk}$ is dominated by a lot of fake tracks as shown as the discrepancy of the sideband data and HF-originated background simulation in Figure 5.7. The condition on p-value$_{trk}$ is not required for each track, but for the product of the three tracks in order to evaluate as a whole event. Figure 5.8 shows the distribution of the number of hits in the precision detector and trigger layers. The peak around 12 in precision hits distribution is validated for each muon in data, furthermore the number of hits tends to be smaller than combined muons. The requirement for this parameter is set to be larger than 13 in order to remove fakes. The number of trigger layers in the data is not significantly different from combined muon case. While the most of signal muons has 4 associated trigger



(a) (b)

(c) (d)

(e) (f)

Figure 5.6: Number of hits on the MDT and CSC for (a) leading muon, (c) 2nd leading muon, and (e) 3rd leading muon, as well as a number of TGC and RPC layers with hit for each muon in (b), (d), and (f). Signal muons had a large number of the hits on precision and trigger detectors, while data and background simulation involved muons with few number of hits. $W \to \tau\nu \to 3\mu\nu$ signal (gray filled), HF-originated background simulation (red and blue lines), and the sideband data (black points) are shown respectively in each plot.

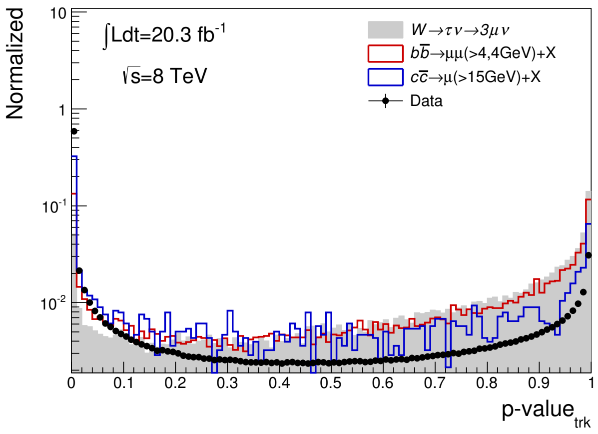


Figure 5.7: The probability of ID tracking for TPa muons in the data and simulation. Signal muons (gray filled) showed higher value than the sideband data (black points) and background simulations (red and blue lines), while muons with small value also remained in data.

layers, the requirement is set to be 3 for combined muons. However, there are shared hits at both TGC and RPC around the transition region for TPa muon as shown in Figure 5.9. In order to suppress the fake tracks by the accidental hits on across both trigger detectors, hits on trigger layer are required to be at either TGC or RPC.

5.2.3 Hadronic jet

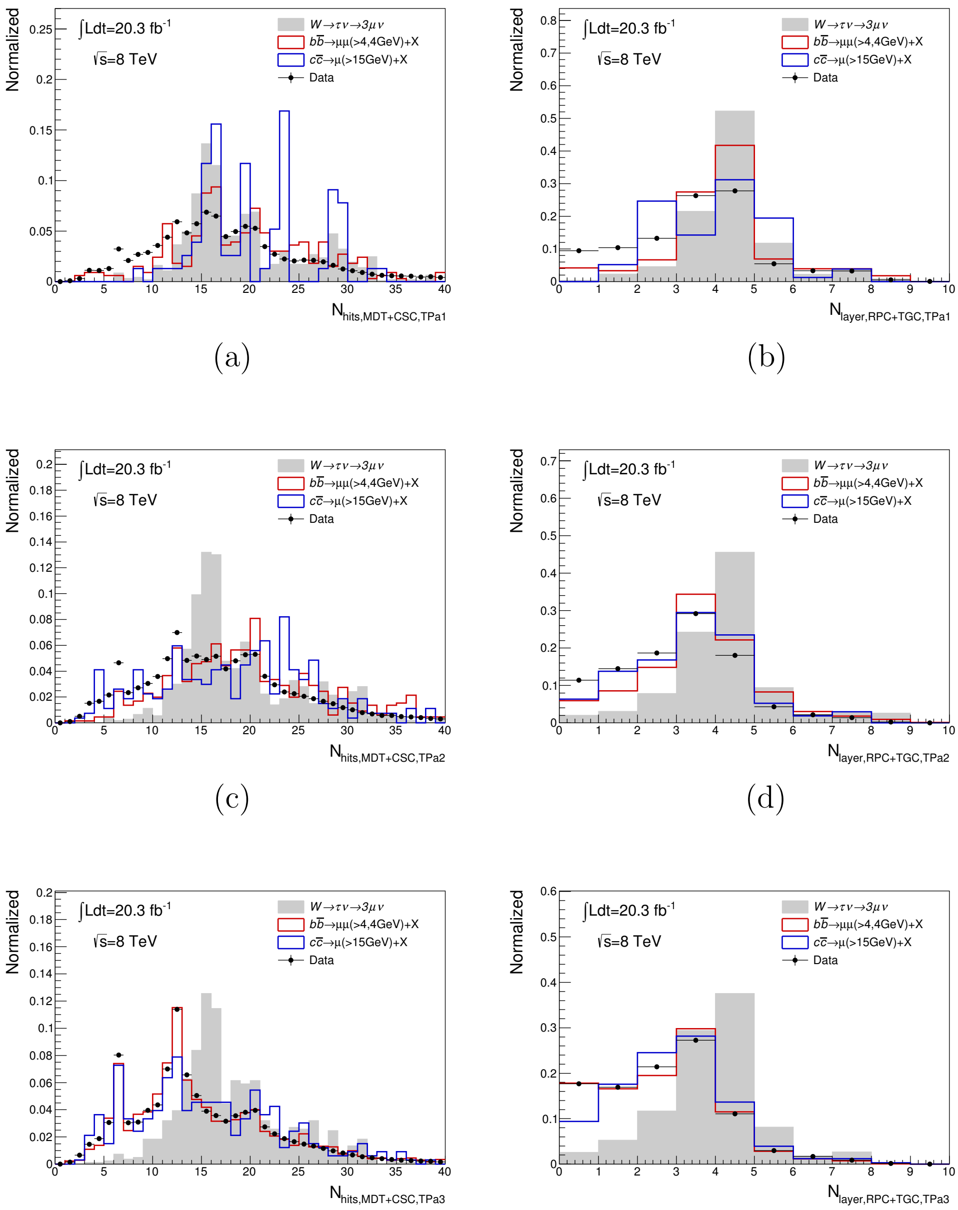
Hadron jets are reconstructed using energy deposition in the calorimeter and associated ID tracks. In this analysis, this object is used for rejection of contamination of the background originating from QCD jets. Also, missing energy is calculated based on reconstructed jets as shown in Figure 5.2.4.

For the first step of the jet reconstruction, initial clusters of the energy deposition in the calorimeter are defined using Topo-cluster algorithm [77] with local cluster weighting (LCW) calibration [78]. These clusters are merged by the anti-k_T algorithm [79], and used as an anti-k_T jet in this analysis. In this algorithm, distances d_{ij} between clusters i and j, and d_{iB} between cluster i and the beam (B) are introduced as

$$d_{ij} = min\left(\frac{1}{k_{T,i}^2}, \frac{1}{k_{T,j}^2}\right)\frac{\Delta_{ij}}{R^2} \tag{5.1}$$

$$d_{iB} = \frac{1}{k_{T,i}^2}, \tag{5.2}$$

where Δ_{ij} is a ΔR like 2-dimensional angular distance defined as $((y_i - y_j)^2 + (\phi_i - \phi_j)^2)^{1/2}$, and $k_{T,i}$, y_i, and ϕ_i are the transverse momentum, rapidity and azimuthal

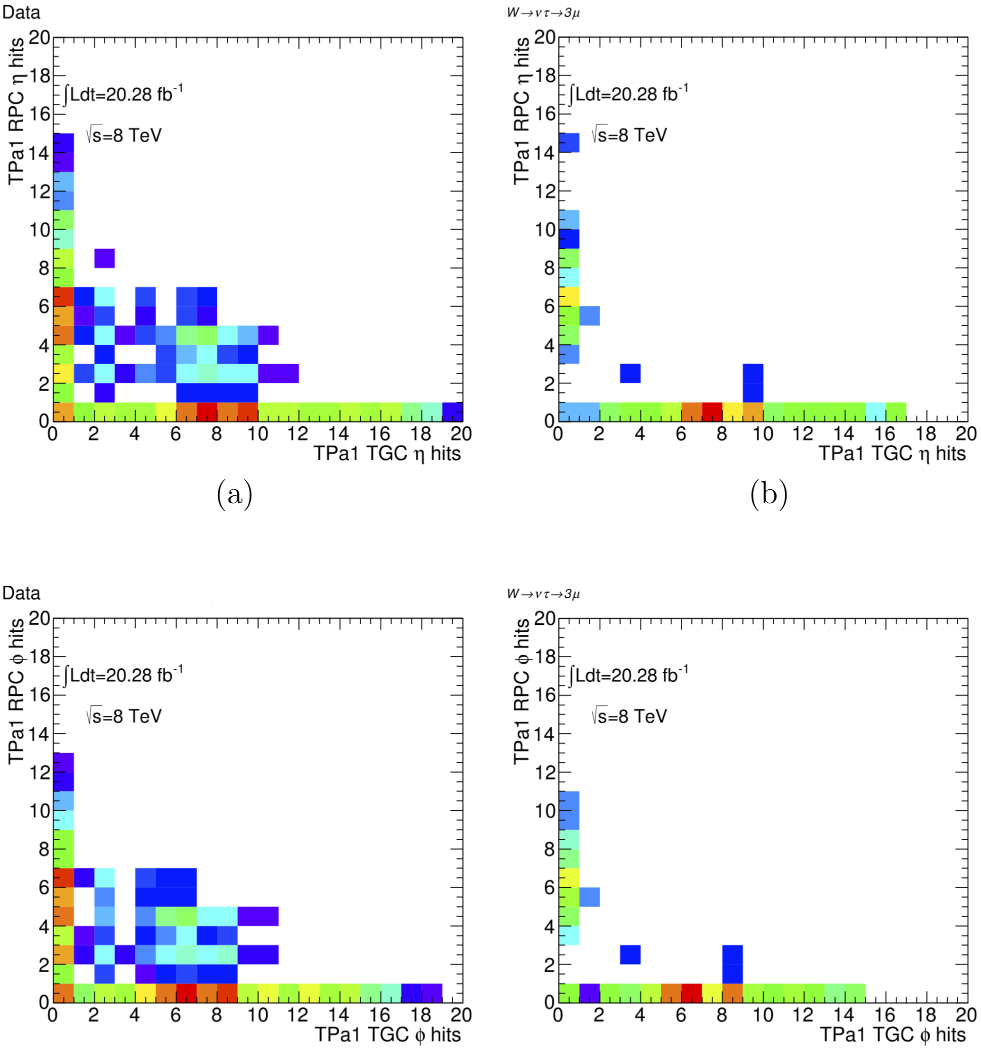


(a) (b)

(c) (d)

(e) (f)

Figure 5.8: Number of hits on the MDT and CSC for (a) leading muon, (c) 2nd leading muon, and (e) 3rd leading muon, as well as number of TGC and RPC layers with hit for each muon in (b), (d), and (f), respectively. TPa muons are selected in these pictures. Signal muons had a large number of the hits on precision and trigger detectors, while data and background simulation involved muons with few number of hits. $W \to \tau\nu \to 3\mu\nu$ signal (gray filled), HF-originated background simulation (red and blue lines), and the sideband data (black points) are shown respectively in each plot.



(a) (b)

(c) (d)

Figure 5.9: 2-dimensional distribution of number of hits on η (a,b) and ϕ (c,d) layers of TGC and RPC. The sideband data are shown in left side, i.e. (a) and (c), and $W \to \tau\nu \to 3\mu\nu$ signal are shown in right side, i.e. (b) and (d). Although hits on trigger plane should be either on TGC or RPC typically, muons could be reconstructed using hits on both plane around transition region of these trigger chambers. Such muons in transition region involve a lot of fakes, therefore should be removed.

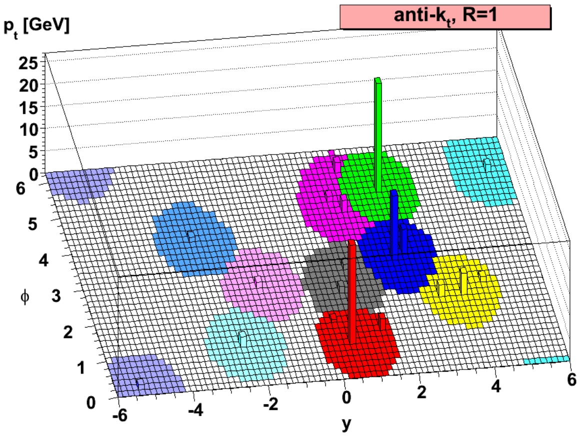


Figure 5.10: Reconstructed anti-k_T jets with $R = 1.0$ of a parton level event by Herwig [80] with soft particles, discriminated by their color [79]. The height of a bin represents the p_{T} at the cells, and cells filled by the same color are merged into a jet.

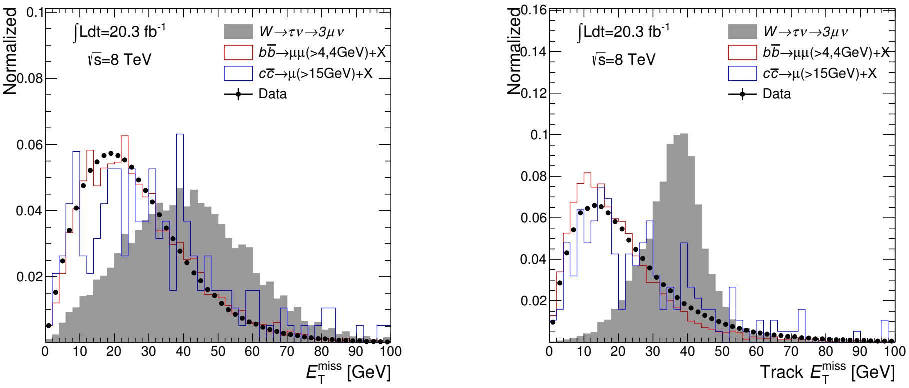
angle of cluster i, respectively. R is the cone size of the jet reconstruction and takes a role as the threshold of the distance between clusters. Identifying the smallest of the distances, merged cluster parameters (k_T, y, ϕ) are re-defined using initial cluster i and j removed from the candidates. If d_{iB} is the smallest candidate, i-th cluster is defined as a jet and removed from the candidates for merging. This merging procedure is repeated until no cluster candidates for merging are left. Then, the reconstructed jet is made by simply accumulating all the soft particles within R, resulting in a conical jet as shown in Figure 5.10. In this analysis, anti-k_T jets with $R = 0.4$ is used as hadron jets.

5.2.4 $E_{\mathrm{T}}^{\mathrm{miss}}$

The total transverse energy in an event is calculated using the vector sum of the reconstructed particles on x-y plane, which is denoted as $\vec{E}_T$.

$$\vec{E}_T = \vec{E}_T^{jet} + \vec{E}_T^{e} + \vec{E}_T^{\mu} + \vec{E}_T^{soft}, \tag{5.3}$$

where $\vec{E}_T^{jet,e,\mu,soft}$ are the 2-dimensional vector sum of each kind of particles (jet, e, μ), and topo-clusters or tracks within $|\eta| < 4.9$ which are not associated with any reconstructed objects are defined as soft terms, noted as soft. This calculation uses electrons, which is not used as an object in the analysis procedure. The electrons are reconstructed by the combination of tracks at ID and energy deposit in the EM calorimeter. The identification, discriminating from photons and charged hadrons, are determined by the shower shape in the calorimater,



(a) (b)

Figure 5.11: (a) Calo$E_{\mathrm{T}}^{\mathrm{miss}}$ and (b) Track$E_{\mathrm{T}}^{\mathrm{miss}}$ distribution of $W \to \tau\nu \to 3\mu\nu$ signal (gray filled), HF-originated background simulation (red and blue lines), and the sideband data (black points). A large $E_{\mathrm{T}}^{\mathrm{miss}}$ of the signal events is generated by the neutrino from W decay.

matching with tracks and balance of energy and momentum, and the transition radiation at TRT. The inversion of x and y components of the $\vec{E}_{\mathrm{T}}$ is defined as the $E_{\mathrm{T}}^{\mathrm{miss}}$:

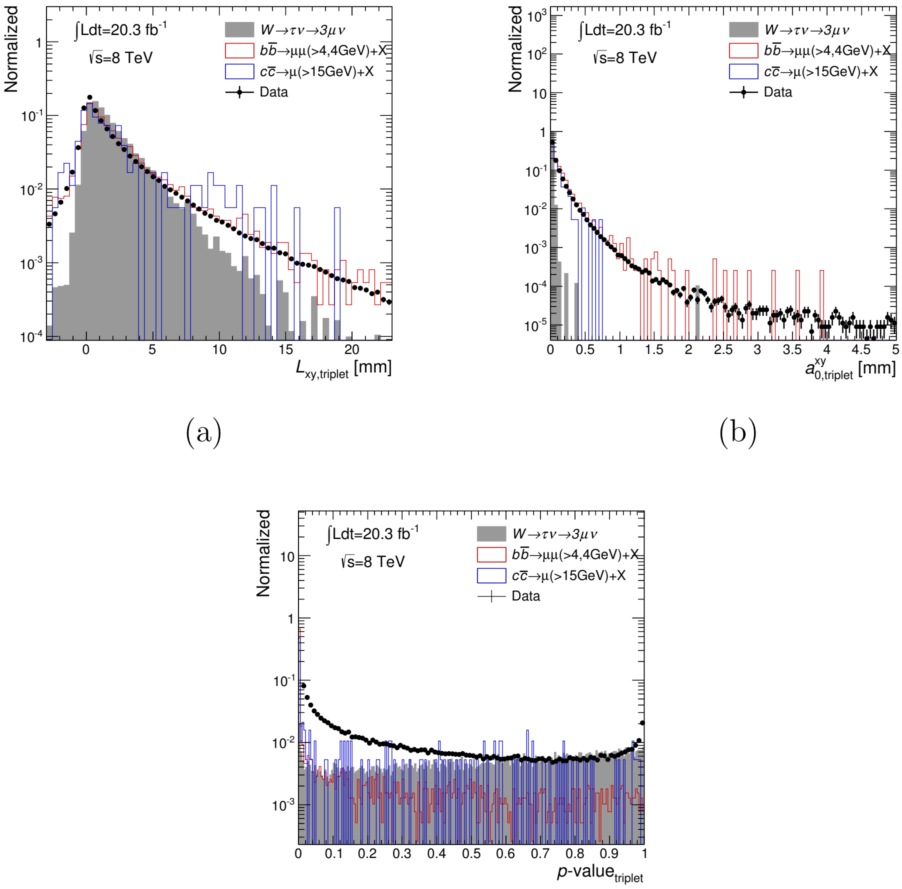
$$E_{\mathrm{T}}^{\mathrm{miss}} = |\vec{E}_{\mathrm{T}}|,\ \ \vec{E}_{\mathrm{T}}^{miss} = -\vec{E}_{\mathrm{T}}. \tag{5.4}$$

Two types of $E_{\mathrm{T}}^{\mathrm{miss}}$ variables with the different methods of the soft term calculation are used to ensure the significance of $E_{\mathrm{T}}^{\mathrm{miss}}$; one of these is based on the energy deposit at the calorimeter (Calo$E_{\mathrm{T}}^{\mathrm{miss}}$) and another is based on the tracks at the ID (Track$E_{\mathrm{T}}^{\mathrm{miss}}$). In this analysis, Track$E_{\mathrm{T}}^{\mathrm{miss}}$ is particularly useful because there should be the three muon tracks which balance with the missing energy. Furthermore, Track$E_{\mathrm{T}}^{\mathrm{miss}}$ accounts for the tracks of mis-reconstructed muons with the MS requirements failed while the Calo$E_{\mathrm{T}}^{\mathrm{miss}}$ does not. These reasons also lead to a better resolution of Track$E_{\mathrm{T}}^{\mathrm{miss}}$ for the missing energy calculated by the three muons than Calo$E_{\mathrm{T}}^{\mathrm{miss}}$. Calo$E_{\mathrm{T}}^{\mathrm{miss}}$ is mainly used for validation of the effect of using clusters by the comparison with Track$E_{\mathrm{T}}^{\mathrm{miss}}$. The Calo$E_{\mathrm{T}}^{\mathrm{miss}}$ and Track$E_{\mathrm{T}}^{\mathrm{miss}}$ distributions of the signal and background simulation, and the data in the sideband range are shown in Figure 5.11 with quality selection for the each reconstructed muon. As the loose selection criteria, $10 < E_{\mathrm{T}}^{\mathrm{miss}} < 250\,\text{GeV}$ is required for both of Calo$E_{\mathrm{T}}^{\mathrm{miss}}$ and Track$E_{\mathrm{T}}^{\mathrm{miss}}$.

5.2.5 Triplet properties and related physics variables

A triplet consists of three combined muons or the combination of two combined muons and a TPa muon as the candidate of the τ lepton in this analysis. The properties of the triplet are used for discrimination from background events, as well as for region definition and the final discriminant. One of the characteristic feature of the signature is the significant life-time of τ lepton. This can be used as flight length on the x-y plane, denoted as L_{xy} as explained in Section 3.1.1. The degree of pointing of the L_{xy} is also defined as a^0_{xy}, which is also a characteristic feature. For the calculation of L_{xy} and a^0_{xy}, the SV is reconstructed with three muons in a triplet in addition, and also the PV is also recalculated subtracting the tracks which are used in the SV reconstruction. The distributions of the L_{xy} and a^0_{xy} are shown in Figure 5.12(a), (b), respectively. The L_{xy} distribution of the signal events shows a tail above 0, but the sideband data and the HF-originated background simulations show longer tails. This is because of the long life time of the HF mesons. Therefore, this is not effective to discriminate from HF-originated background, however it can be used for rejection of the background from other sources, i.e. accidental events and fakes. The a^0_{xy} distribution of the signal events tend to be concentrated around zero, while the background events have tails due to the other emitted particles by the HF meson decays, especially for B meson. The flight length of the $W \to \tau\nu \to 3\mu\nu$ events is generated by the τ, therefore it can be used for the discrimination from the HF-background. In order to consider the accuracy of the measurement of these parameters, their significance (σL_{xy} and σa^0_{xy}) are also defined by dividing by their own uncertainties. The probability of the SV is defined by the χ^2 for number degree of freedom ($ndof$) of the vertex fitting. High probability means that these three tracks come from one point spatially, which is useful for discrimination from the background events by the accidental matching of irrelevant tracks. Hence, in the comparison of the probability distributions shown in Figure 5.12(c), the sideband data cannot be understood by only the HF-originated background due to accidental events. The requirement of higher probability of the SV is expected to be effective to reduce such backgrounds. The triplet kinematic variables, p_{T}, η, ϕ, and invariant mass, are reconstructed by the re-fitted tracks, which are corrected by the SV fitting process. The invariant mass of the signal triplet should be the mass of τ, therefore used as the final discriminant in the analysis. Using these variables of the triplet, the loose selection, also noted as the pre-selection (hereafter loose selection criteria), is set as

- Product of the three track probabilities in the triplet as $\mathcal{P}_{trk} = \prod_{i=1}^{3} p$-value$_{\mathrm{i-thtrack}}$ should be larger than 10^{-9}
- $-10 < L_{\mathrm{xy}}/\sigma L_{\mathrm{xy}} < 50$
- $a^0_{\mathrm{xy}}/\sigma a^0_{\mathrm{xy}} < 25$
- p-value of the SV reconstruction should be larger than 0.01
- $p_{\mathrm{T,triplet}} > 10\,\mathrm{GeV}$
- Absolute value of the total charge of the three muons is equal to 1
- $M_{\mathrm{triplet}} < 4.0\,\mathrm{GeV}$ (a loose mass window for triplet mass)



(a) (b)

(c)

Figure 5.12: (a) $L_{\rm xy}$ and (b) $a^0_{\rm xy}$ distribution which are reconstructed by the PV and the re-fitted SV, as well as (c) probability of the SV reconstruction. $W \to \tau\nu \to 3\mu\nu$ signal (gray filled), HF-originated background simulation (red and blue lines), and the sideband data (black points) are shown as blue filled, red and green meshed, and black points, respectively. Large number of HF background also comes from displaced vertex, however these tend to be less pointing and with low probability of the SV.

$M_{\mathrm{T,triplet,E_T^{miss}}}$ and $\Delta\phi_{\mathrm{triplet,E_T^{miss}}}$ are defined using the triplet and $E_{\mathrm{T}}^{\mathrm{miss}}$ (both Calo and Track-based) as shown in Figure 5.13. These are important to enhance the signal from $W \to \tau\nu$ as shown in Section 3.1.1. The requirement of $M_{\mathrm{T,triplet,E_T^{miss}}} > 20\,\mathrm{GeV}$ are applied for the event as the loose selection for both of Calo$E_{\mathrm{T}}^{\mathrm{miss}}$ and Track$E_{\mathrm{T}}^{\mathrm{miss}}$. These variables are also used in the tight selection described in Section 5.3.

The sum of the transverse momenta of visible particles $\vec{H}_{\mathrm{T}}$ is also defined as

$$\vec{H}_{\mathrm{T}} = \vec{p}_{\mathrm{T,triplet}} + \sum_{i=1}^{4} \vec{p}_{\mathrm{T,jet\ i}}, \tag{5.5}$$

where the $\vec{p}_{\mathrm{T,triplet}}$ and $\vec{p}_{\mathrm{T,jet\ i}}$ are the momentum vector of the triplet and i-th leading hadron jet, respectively. Up to four jets with large p_{T} above $30\,\mathrm{GeV}$ are taken into account. This variable is used to ensure the significant $E_{\mathrm{T}}^{\mathrm{miss}}$ from W decay by calculating $M_{\mathrm{T,H_T}}$ using $\vec{H}_{\mathrm{T}}$ and $E_{\mathrm{T}}^{\mathrm{miss}}$. Figure 5.14 shows the distribution of the $\vec{H}_{\mathrm{T}}$ and $M_{\mathrm{T,H_T}}$ after the selection of muons. $M_{\mathrm{T,H_T}}$ is also required to be larger than $30\,\mathrm{GeV}$ as well as $M_{\mathrm{T,triplet}}$.

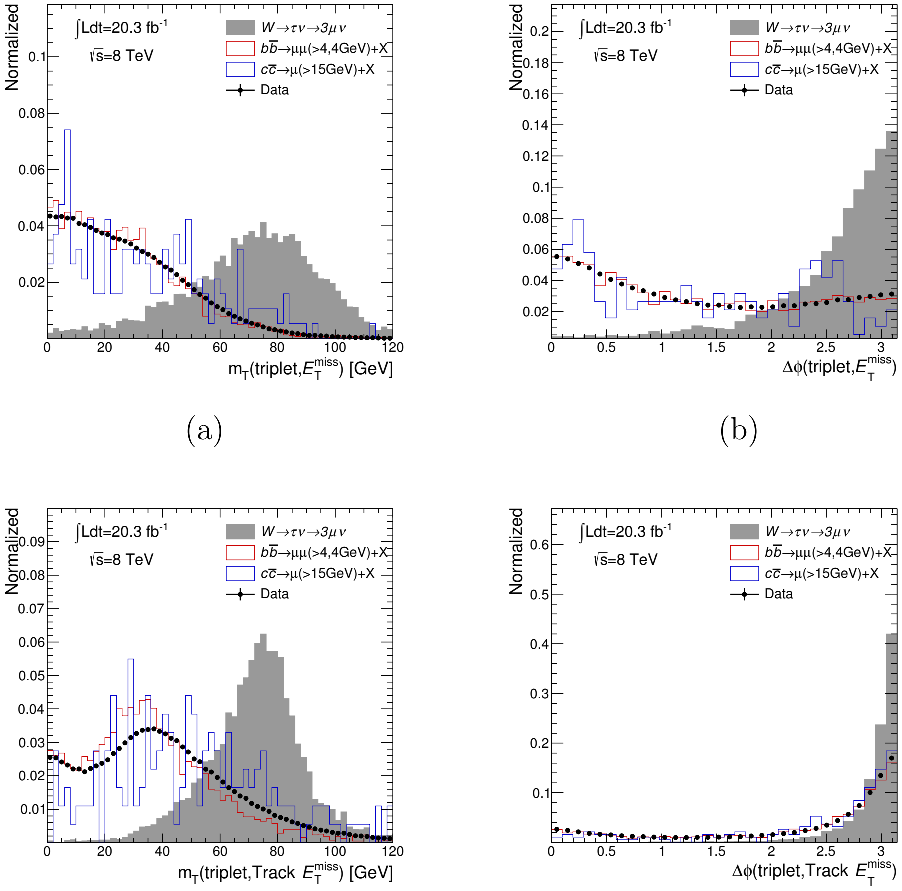
The track isolation of triplet is defined as the sum of the track p_{T} (with $500\,\mathrm{MeV}$or above) in a cone excluding the three tracks constructing triplet

$$\Sigma p_{\mathrm{T,track}} = \sum_{i=0}^{\Delta R_{\mathrm{triplet,track}} < cone} p_{\mathrm{T,track\ i}} \tag{5.6}$$

$$cone = \Delta R_{\mathrm{triplet}}^{max} + margin, \tag{5.7}$$

where $\Delta R_{\mathrm{triplet,track}}$ is ΔR between the triplet and the track. The cone size is determined by the fixed margin of 0.2 (and 0.3) and $\Delta R_{\mathrm{triplet}}^{max}$, which is the maximum of the ΔR of all pairs of three tracks. The $\Delta R_{\mathrm{triplet}}^{max}$ depends on the degree of boost of the triplet, i.e. $p_{\mathrm{T,triplet}}$. Other objects in the cone also tends to be close to highly boosted triplet, which mostly come from boosted object decays. Therefore, isolation requirements should be set for $\Sigma p_{\mathrm{T,track}}/p_{\mathrm{T,triplet}}$, denoted as $\mathrm{Isolation_{triplet,marginsize}}$, considering the degree of collimation. Figure 5.15 shows the distribution of the isolation value $\mathrm{Iso_{triplet}}$ with the margin size of 0.2 and 0.3. A triplet in a hadron jet has a large track isolation value, therefore the requirement of low $\mathrm{Iso_{triplet}}$ is useful to reduce background events. The loose requirement for the isolation for triplet is set as $\mathrm{Iso_{triplet,0.2(0.3)}} < 0.3(1.0)$.

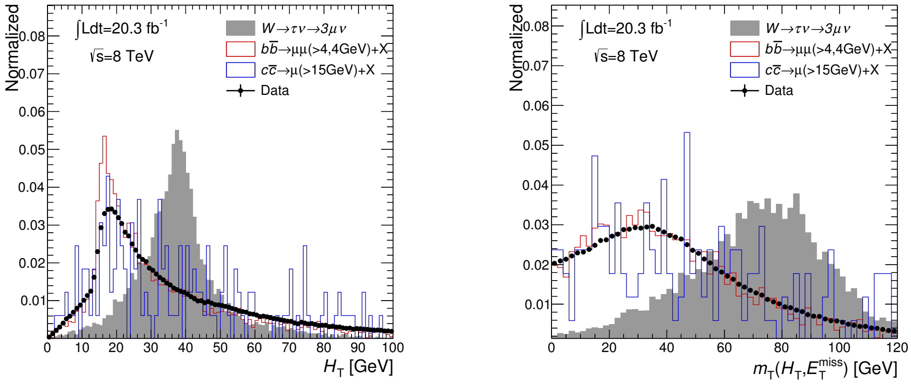
The events in the loose selection criteria are used for the training and optimization of the MVA (will be shown in Section 5.3), while the tight selection criteria is also defined to use together with the selection based on the MVA. In this selection, three muon-pairs in the triplet are also used. Two of them are opposite-sign (OS1 and OS2) pairs, and another is same-sign (SS) pair. OS1 is defined as reconstructed by two muons with higher p_{T} than muons of OS2. In particular, the invariant masses of the OS1 and OS2 (M_{OS1}, M_{OS2}) are used to reject the light resonances decaying into two muons, ϕ and ω mesons, whose



(a) (b)

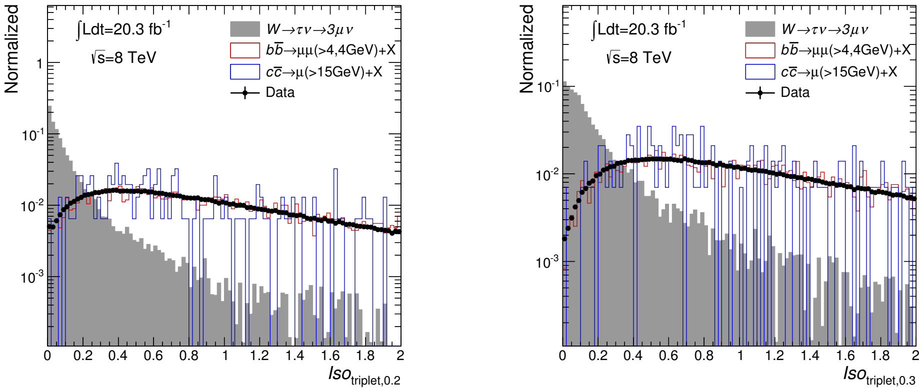
(c) (d)

Figure 5.13: (a) $M_{\mathrm{T,triplet,TrackE_T^{miss}}}$ and (b) $\Delta\phi_{\mathrm{triplet,TrackE_T^{miss}}}$ distribution defined by the triplet and Track$E_{\mathrm{T}}^{\mathrm{miss}}$ in the event, as well as (c) $M_{\mathrm{T,triplet,CaloE_T^{miss}}}$ and (d) $\Delta\phi_{\mathrm{triplet,CaloE_T^{miss}}}$ distribution using Calo$E_{\mathrm{T}}^{\mathrm{miss}}$ instead of Track$E_{\mathrm{T}}^{\mathrm{miss}}$. $W \to \tau\nu \to 3\mu\nu$ signal (gray filled) shows a large $M_{\mathrm{T,triplet}}$ and opposite direction $E_{\mathrm{T}}^{\mathrm{miss}}$ against the triplet in x-y plane, while HF-originated background simulation (red and blue lines), and the sideband data (black points) show similar smooth distributions.



(a) (b)

Figure 5.14: (a) $\vec{H}_\mathrm{T}$ and (b) $M_{\mathrm{T},\mathrm{H_T}}$ distributions of $W \to \tau\nu \to 3\mu\nu$ signal (gray filled), while HF-originated background simulation (red and blue lines), and the sideband data (black points). The small cutoff around $\vec{H}_\mathrm{T} = 20\,\mathrm{GeV}$ is due to the loose threshold for $p_{\mathrm{T,triplet}}$ by $20\,\mathrm{GeV}$.



(a) (b)

Figure 5.15: $\mathrm{Iso}_\mathrm{triplet}$ distributions with the margin size of (a) 0.2 and (b) 0.3 for $W \to \tau\nu \to 3\mu\nu$ signal (gray filled), HF-originated background simulation (red and blue lines), and the sideband data (black points). Signal triplets concentrate around zero, while backgrounds are distributed broadly.

masses (M_ϕ, M_ω) are 1020 MeV and 782 MeV, respectively. The tight selection requirements are shown as follows:

- $\mathcal{P}_{trk} = \prod_{i=1}^{3} p\text{-value}_{\rm i-thtrack}$ should be larger than 8.0×10^{-9}
- $1 < L_{\rm xy}/\sigma L_{\rm xy} < 50$
- $a^0_{\rm xy}/\sigma a^0_{\rm xy} < 25$
- p-value of the SV reconstruction should be larger than 0.2
- $M_{\rm T,triplet,Calo(Track)E_T^{miss}} > 45.0\,\text{GeV}$
- $\Delta\phi H_{\rm T}$,Calo(Track)$E_{\rm T}^{\rm miss} > 2$
- $M_{\rm OS1}$, $M_{\rm OS2}$, and $M_{\rm SS}$ (invariant mass of the SS) should be larger than 300 MeV
- If either of the $p_{\rm T,triplet}$ or the Calo(Track)$E_{\rm T}^{\rm miss}$ is lower than 35.0 GeV, the differences of $M_{\rm OS1(OS2)}$ from M_ϕ and M_ω should be larger than 50 MeV
- In order to reject the $D_s \to \phi\pi \to \mu\mu\pi$ originated background events, the difference of $M_{\rm OS1(OS2)}$ from M_ϕ should be smaller than 40 MeV, if the difference of $M_{\rm triplet}$ from $M_{\rm D_s}$ (1968 MeV) is larger than 100 MeV

5.3 Multi-Variate Analysis procedure

The main event selection procedure of this analysis was based on a MVA technique. This is an effective method especially for the case without any clear discriminant variables from background events. Considering the several aspects of the signal events, it is expected to give a better separation than applying a simple threshold for each variable. On the other hand, the degree of "signal-like" for an event, called as score of the event, was determined by the training using background and signal samples. The score depends on the way of training, the number of sample events, and quality of samples for background. Furthermore, an appropriate set of variables for MVA score calculation is also important not only for higher selection efficiency, but also avoiding the bias by enhancing the triplet with mass of τ accidentally.

5.3.1 Configuration and options

The MVA selection strategy is not for each variable separately, but a score of an event is determined using all variables inclusively and totally. Then, the score calculation is based on the "training" step using the signal (S) and background (B) event samples. Boosted Decision Tree (BDT) [81], constructed by the decision tree with variables considering the weight of each, was employed in the score determination.

In the basic decision tree method even without boost, the best separation value into one side having mostly signal and the other mostly background for each variable is decided first. Variable giving the best separation in all the variables is used as the the first selection of the decision tree, and other variables are also ordered by their separation values. Figure 5.16 shows a simple case with three variables: A; B; and C. The input events are separated into two sides of the signal-like (S-side) and background-like (B-side) by each variable. These two parts are called "sons" of the variable. If the contamination of background (signal) events in the S-side (B-side) was large, the next leading variable is used for further separation. The selection point for each variable should be optimized to accept the signal samples as S-side and the background samples as B-side. The purity of the selection, which is used for the optimization, is given as:

$$P = \frac{\sum_s W_s}{\sum_s W_s + \sum_b W_b}, \tag{5.8}$$

where W_s and W_b are weights for each event in signal and background samples, respectively. These weights are set to be constant,1.0, initially. Therefore, P value should be worse by the contamination of background events in the S-side. The weight is re-obtained to represent the importance for the mis-classified event, i.e. an weight for a signal (background) event which is classified as the B-side (S-side) by the decision tree is set higher. The selection of each variable is re-optimized using these weights. This step is called "Boost" of the decision tree. For this boost process, the goodness of the selection is defined as:

$$\text{Gini} = \left(\sum_{i=1}^{n} W_i\right) P(1-P), \tag{5.9}$$

where n is number of events, and W_i is weight of the event. In order to minimize the sum of Gini for sons, the criterion for the optimization is also defined as:

$$\text{Criterion} = \text{Gini}_{\text{branch}} - \text{Gini}_{\text{S-side}} - \text{Gini}_{\text{B-side}}. \tag{5.10}$$

By minimizing this criterion, the mis-classified events are enhanced in the "boosted" optimization process. After the boost, the score is determined based on the P value. The score for the highest P value should be set as 1.0, and lowest one should be set as -1.0. Therefore, the tight selection for the score can keep purity of the signal higher, although the signal efficiency also should be lower. For usage of the BDT method, the depth of the decision tree, the number of boost steps, and the limitation of the purity for a variable should be set considering the number of events in S and B samples.

In this analysis, S sample is signal simulation, while B sample is extracted from data in the training region. The triplet mass distribution with loose selection and

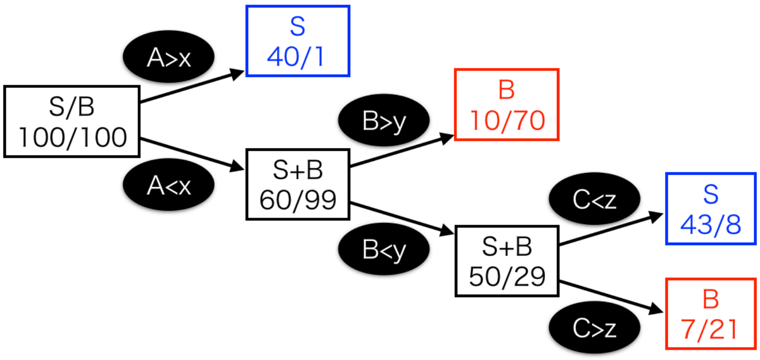


Figure 5.16: Brief illustration of the BDT flow. Events in each sample are classified into signal-like (blue boxes) or background-like (red boxes), however the classified events with loose selection (black boxes) is further classified by the next leading variable. The final terminals are called leaf of the tree, i.e. blue and red boxes. On the other hand, the black boxes are called branches of the tree. A, B, and C represent the variables used for the selection, and x, y, and z are the requirement for each variable.

regions defined in Section 5.1 are shown in Figure 5.17. The shape in the sideband and training region is distributed gently enough to use as the same component in the whole region including SR. Finally, S and B samples are defined by splitting randomly into training sample and validation sample of the selection efficiency, to ensure there is no overtraining. Number of events for S and B samples which are used for training are 6400 and 3500, respectively. The options of the BDT method is to set the number of variables, as well as the amount of events.

5.3.2 Choice of variables

16 variables of tracks, triplet, $E_{\mathrm{T}}^{\mathrm{miss}}$, and their combination were inclusively used for the training and score calculation of the BDT as follows:

- Product of the three track probabilities in the triplet as $\mathcal{P}_{trk} = \prod_{i=1}^{3} p_{\mathrm{i-thtrack}}$
- Number of tracks which were used for the PV reconstruction N_{trks}^{PV}.
- Probability of the SV re-fitting p-value$_{triplet}$.
- Significance of L_{xy} $(L_{\mathrm{xy}}/\Delta L_{\mathrm{xy}})$
- Significance of a_{xy}^{0} $(a_{\mathrm{xy}}^{0}/\Delta a_{\mathrm{xy}}^{0})$
- Transverse momentum of the triplet $p_{\mathrm{T,triplet}}$
- Isolation variable with the ΔR margin of 0.2 $\mathrm{Iso}_{triplet,0.2}$
- Vector sum of the visible energy H_{T}

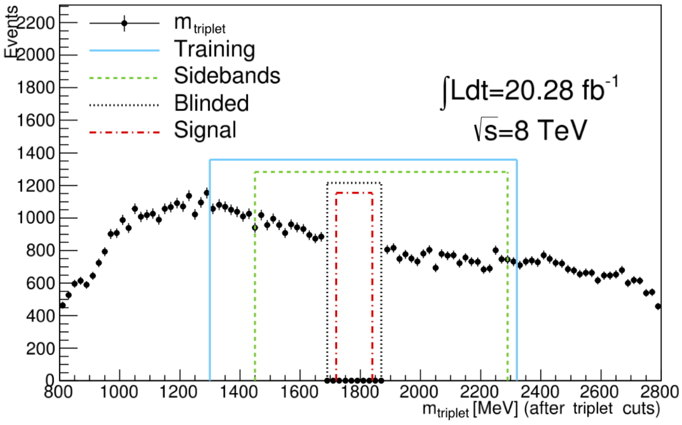


Figure 5.17: Triplet mass distribution with region definition for background sample (data). The events within the blinded region is not used for the optimization of the selections.

- Calo-based missing transverse energy Calo$E_{\mathrm{T}}^{\mathrm{miss}}$
- Transverse mass reconstructed by the triplet and Calo$E_{\mathrm{T}}^{\mathrm{miss}}$ $M_{\mathrm{T,triplet,CaloE_{T}^{miss}}}$
- Degree of agreement of the triplet momentum and missing energy $p_{\mathrm{T,triplet}}/\mathrm{Calo}E_{\mathrm{T}}^{\mathrm{miss}}-1$
- Angular difference of the triplet and Calo$E_{\mathrm{T}}^{\mathrm{miss}}$ on the transverse plane $\Delta\phi_{\mathrm{triplet,CaloE_{T}^{miss}}}$
- Track-based missing transverse energy Track$E_{\mathrm{T}}^{\mathrm{miss}}$
- Transverse mass reconstructed by the triplet and Track$E_{\mathrm{T}}^{\mathrm{miss}}$ $M_{\mathrm{T,triplet,TrackE_{T}^{miss}}}$
- Degree of agreement of the triplet momentum and missing energy $p_{\mathrm{T,triplet}}/\mathrm{Track}E_{\mathrm{T}}^{\mathrm{miss}}-1$
- Angular difference of the Calo$E_{\mathrm{T}}^{\mathrm{miss}}$ and Track$E_{\mathrm{T}}^{\mathrm{miss}}$ on the transverse plane $\Delta\phi_{\mathrm{CaloE_{T}^{miss},TrackE_{T}^{miss}}}$

The distributions of these variables are shown in the next section (Section 5.3.3) with the loose requirement for the BDT score in order to compare between with and without selections.

5.3.3 MVA response and selection

MVA score is calculated for the validation sample for signal and background. Figure 5.18 shows the response distributions of the BDT with the loose selection criteria, and after the tight selection criteria and BDT score $> x_0$ requirement.

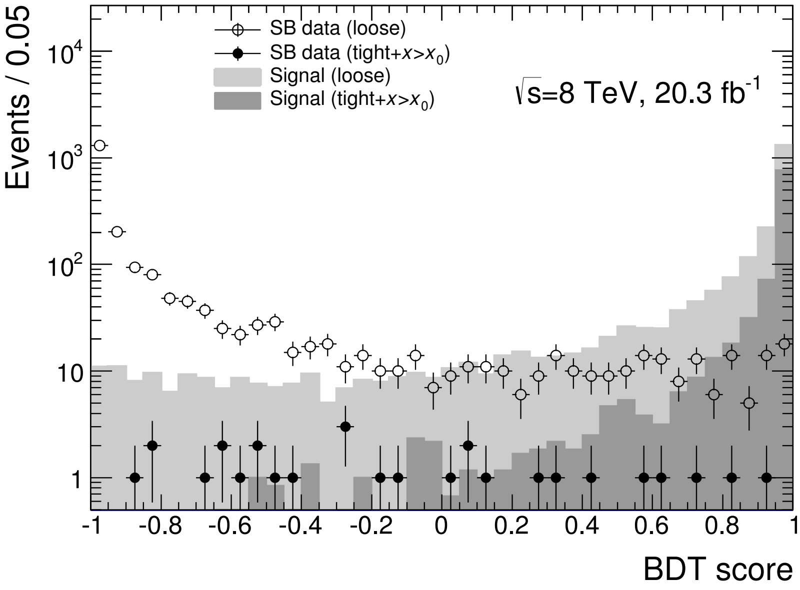
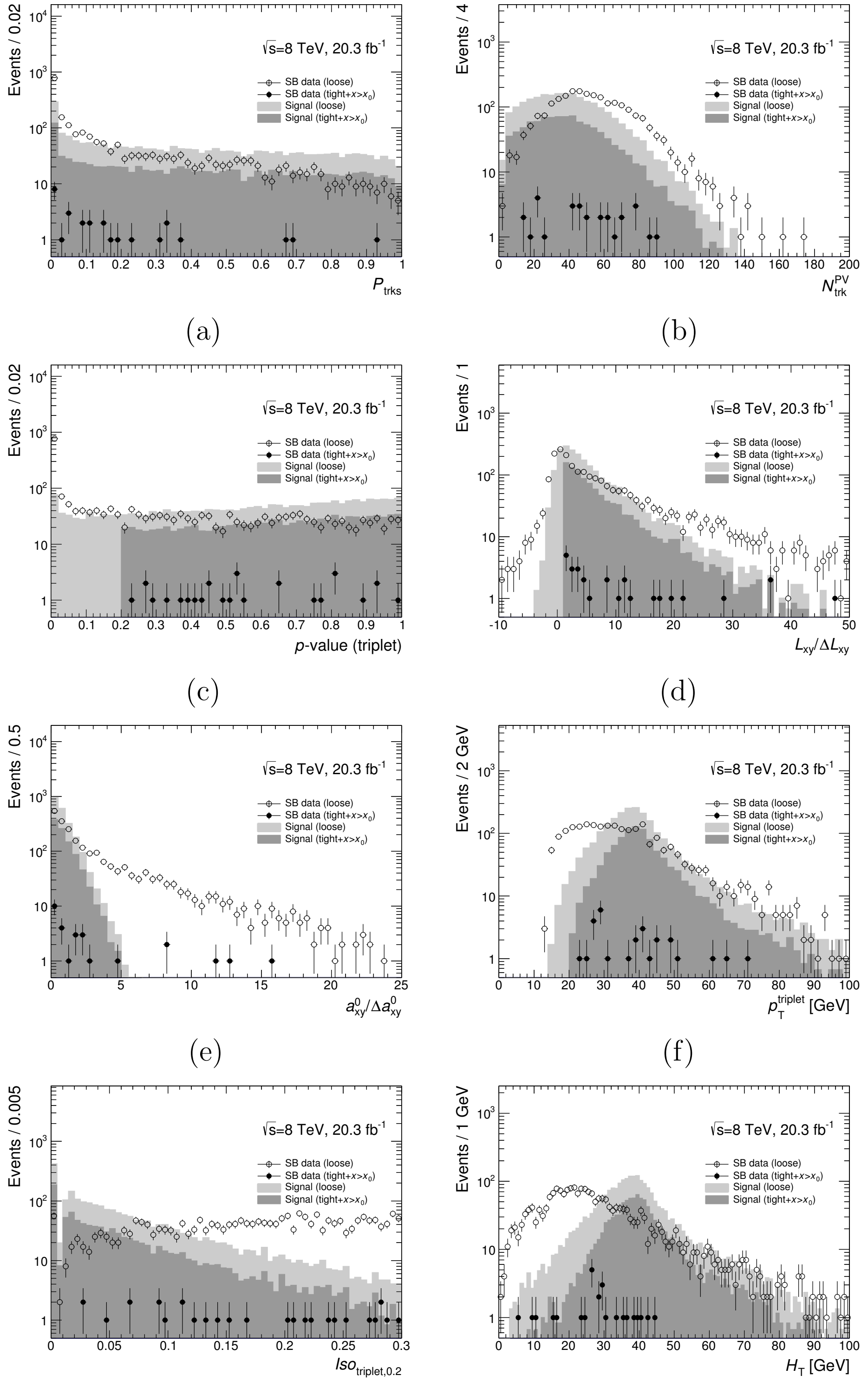


Figure 5.18: Distribution of the BDT score. The signal simulation (and the sideband data) after the loose selection criteria is shown as light gray filled histograms (black hollow circles) and those after the tight selection criteria with x_0 requirement for the BDT response is shown as dark gray filled histograms (black solid circles).

Few events remained in the sideband region after the tight selection even without a BDT requirement.

Based on the distribution shape of the BDT response for signal and background samples, the lower limit was set in order to define the tight selected SR. The loose threshold for the validation and estimation of the background contamination (Section 6.1) is defined as x_0, and the tight one for the signal selection is defined as x_1. x_0 is fixed to be -0.9, and distributions for the variables are shown in Figure 5.19 and 5.20 after the x_0 requirement with the tight selection criteria, as well as after the loose selection criteria only. x_1 should be optimized considering the number of background events in the region of $x_1 <$ BDT response < 1.0. On the other hand, it is difficult to estimate the number of background by the counting in the region due to few events with the tight selection. Hence, $x_1 = 0.9$ is initially set in this step and re-optimized with proper background estimation procedure (Section 6.1).



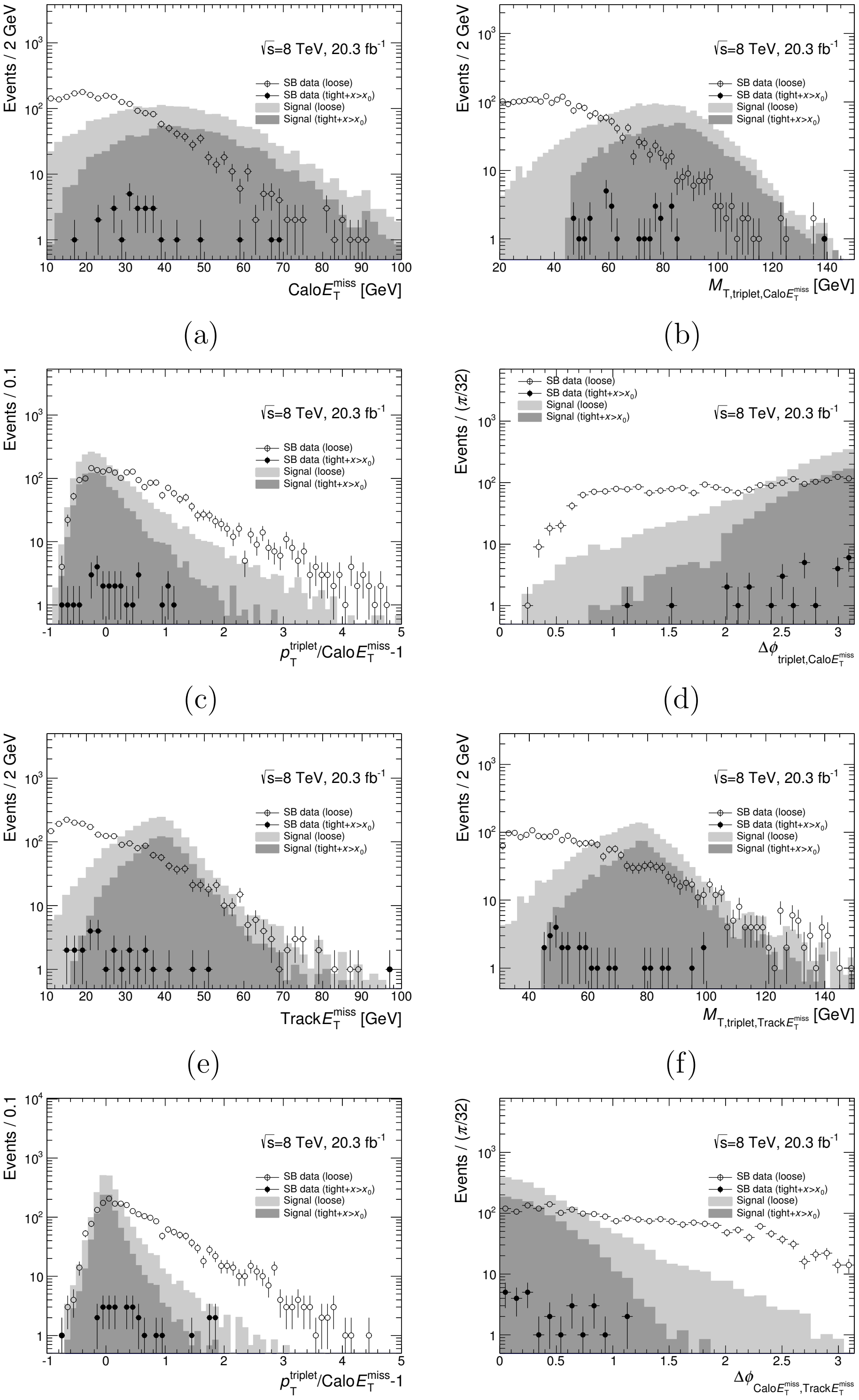
(a) (b)

(c) (d)

(e) (f)

(g) (h)

Figure 5.19: Distributions of (a) $\mathcal{P}_{trk}$, (b) N^{PV}_{trks}, (c) p-value$_{triplet}$, (d) $L_{\mathrm{xy}}/\Delta L_{\mathrm{xy}}$, (e) $a^0_{\mathrm{xy}}/\Delta a^0_{\mathrm{xy}}$, (f) $p_{\mathrm{T,triplet}}$, (g) Iso$_{triplet,0.2}$, (h) H_{T}. The signal simulation (and the sideband data) after the loose selection criteria is shown as light gray filled histograms (black hollow circles) and those after the tight selection criteria with x_0 requirement for the BDT response is shown as dark gray filled histograms (black solid circles).



(a) (b) (c) (d) (e) (f) (g) (h)

Figure 5.20: Distributions of (a) Calo$E_{\mathrm{T}}^{\mathrm{miss}}$, (b) $M_{\mathrm{T,triplet,CaloE_T^{miss}}}$, (c) $p_{\mathrm{T,triplet}}/\mathrm{Calo}E_{\mathrm{T}}^{\mathrm{miss}}-1$, (d) $\Delta\phi_{\mathrm{triplet,CaloE_T^{miss}}}$, (e) Track$E_{\mathrm{T}}^{\mathrm{miss}}$, (f) $M_{\mathrm{T,triplet,TrackE_T^{miss}}}$, (g) $p_{\mathrm{T,triplet}}/\mathrm{Track}E_{\mathrm{T}}^{\mathrm{miss}}-1$, (h) $\Delta\phi_{\mathrm{CaloE_T^{miss},TrackE_T^{miss}}}$. The signal simulation (and the sideband data) after the loose selection criteria is shown as light gray filled histograms (black hollow circles) and those after the tight selection criteria with x_0 requirement for the BDT response is shown as dark gray filled histograms (black solid circles).

Chapter 6

Background Estimation and Normalization

Not only the search for signal events in the data, but also other measurements are necessary to exploit a sensitive result. The number of background events is important to evaluate the significance of the remaining signal in the signal region. It is also important to estimate the number of $W \to \tau\nu$ events in the 8 TeV pp collision for the calculation of $\mathcal{BR}(\tau \to 3\mu)$. In this chapter, the procedure to estimate these numbers are described.

6.1 Background estimation

The background components are difficult to split into several obvious sources correctly because these are similar to each other and kinds of considerable sources are too many to estimate the cross section of each. The background components in the SB region are assumed to be the same as in the SR region, therefore the background is extracted from the SB distribution of data in this analysis. On the other hand, two issues need to be considered, that is (1) events in the SB region are used for MVA training, (2) number of events are not enough to make precise estimation. In order to resolve the problem, it is essential to use the MVA score distribution of SB events not used for training, i.e. "validation" side of the splitting in Section 5.3.1. Furthermore, amount of background events is one of the important factors for sensitivity of the signal. As explained in the MVA selection (Section 5.3.3), the tight SR range is re-optimized depending on the background estimation.

6.1.1 Extrapolation from loose MVA selection

First, two control regions, denoted as CR_0 and CR_1, are defined as the SB region with loose and tight selections for the BDT response ($x_0 < \mathrm{BDT} < x_1$ and $x_1 < \mathrm{BDT} < 1.0$), respectively. SR_0 and SR_1 were also defined with the same thresholds in the SR mass window. Ideally, background contamination in SR_1

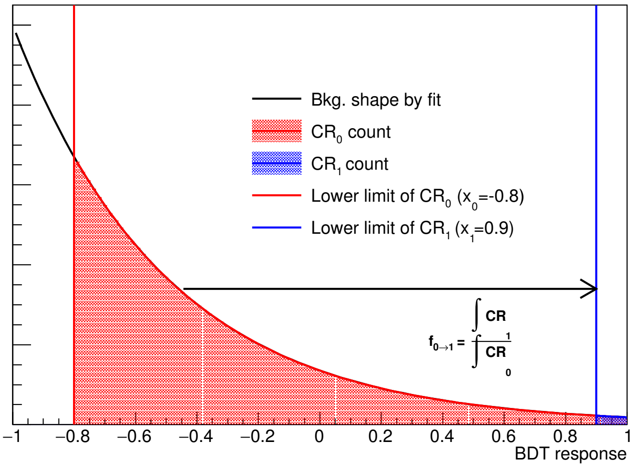


Figure 6.1: Overview of the strategy of extrapolation for background distribution. It makes the estimation of background robust to use the fitting in mass distribution of CR$_0$ instead of CR$_1$, which has only few events. $f_{0\to1}$ is determined for extrapolation from SR$_0$ (extracted from CR$_0$ by fitting the triplet mass distribution) to SR$_1$ by counts in CR$_0$ and CR$_1$ using BDT response shape by fitting. In this figure, x_0 and x_1 are set as -0.8 and 0.9 in this figure.

should be extracted from the CR$_1$ by fitting of the triplet mass distribution. But in this analysis, in order to resolve the lack of statistics in CR$_1$, it was estimated by extrapolation from SR$_0$ to SR$_1$ with a transfer factor ($f_{0\to1}$) defined as the fraction of CR$_1$ in CR$_0$ using the shape of BDT response. Figure 6.1 describes the procedure overview and region definition of CR.

Background contamination in the SR$_0$ is determined by fitting of CR$_0$ as the SB region. BDT response and the triplet mass distributions after x_0 requirement (fixed to be -0.9 in Section 5.3.3) are fitted by, respectively,

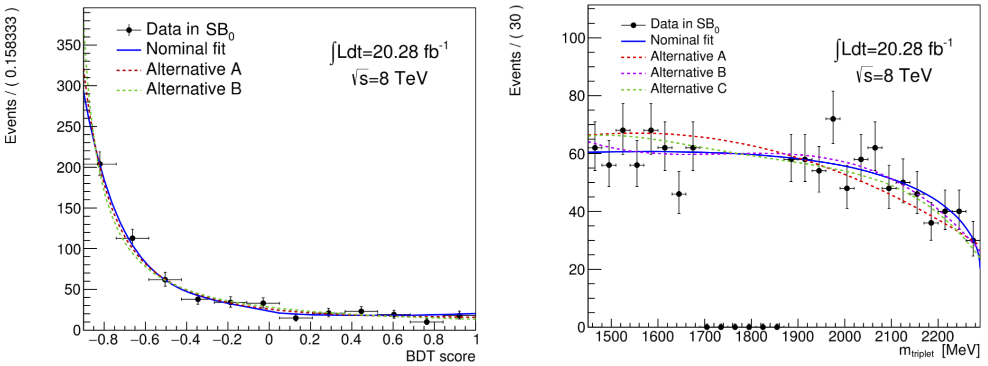
$$\mathcal{B} = p_0 + p_1 \exp(x p_2) + p_3 |x|^{p_4} \tag{6.1}$$

$$\begin{aligned} \mathcal{SB} =& N_{\text{Bernstein}} \times \text{Bernstein}(y|p_0, p_1, p_2) && (6.2)\\ &+ N_{\text{Argus}} \times \text{Argus}(y|m_0 = M_{\text{triplet}^{max}}, \kappa), && (6.3) \end{aligned}$$

where p_i's are free parameters for fit, and x and m_0 are the BDT response and the triplet mass, respectively. Berstein and Argus functions in the sideband fit of mass distribution are described in [82, 83]. Although $\mathcal{B}$ and $\mathcal{SB}$ are used as nominal functions, it is necessary to evaluate fluctuation of the fitting by using alternative functions. Alternative functions of $\mathcal{B}$ are

$$\mathcal{B}_{Alt.A} = p_0 + \frac{p_1}{(x+1)^{p_2}} + p_3 (x+1)^{p_4} \tag{6.4}$$

$$\mathcal{B}_{Alt.B} = p_0 + p_1 \times (p_2)^x. \tag{6.5}$$



(a) (b)

Figure 6.2: Fitting results for (a) BDT response and (b) triplet mass by the nominal and alternative functions. Black markers show distributions of SB data, and blue solid lines show fit by the nominal function. The alternative functions drawn as dashed lines.

Alternative functions of $\mathcal{SB}$ are

$$\mathcal{SB}_{Alt.A} = \text{Chebychev}(y|p_0, p_1) \tag{6.6}$$
$$\mathcal{SB}_{Alt.B} = \text{Chebychev}(y|p_0, p_1, p_2) \tag{6.7}$$
$$\mathcal{SB}_{Alt.C} = \text{Chebychev}(y|p_0, p_1, p_2, p_3), \tag{6.8}$$

where Chebychev$(y|p_0, ..., p_n)$ are Chebychev polynomials of first kind with the triplet mass (x) for n=1, 2, 3. Figure 6.2 shows the fitting results for BDT response and triplet mass distribution with nominal and alternative functions. The difference between nominal function and alternative functions was quoted as the systematic uncertainty of the fit.

On the other hand, the estimated value should fluctuate by the cutoff of x_0. This is considered as the uncertainty of the background modeling. Therefore, the variation of responses for x_0 from -0.9 to -0.8 is confirmed with a step of 0.01. The SB range definition can also fluctuate the estimation of the background. The lower and upper limits of the SB definition are changed by the combination of $[1300, 1330, \cdots, 1540]$ and $[2260, 2230, \cdots, 2020]$. The maximal differences for each x_1 requirement for the signal region are evaluated as shown in Figure 6.3. The total uncertainty of the background estimation is determined to be [45-80]% depending on the x_1 value.

6.1.2 Number of background and SR determination

The number of background events change as a function of x_1, as well as the number of signal events. Therefore, the threshold of x_1 should be optimized, considering

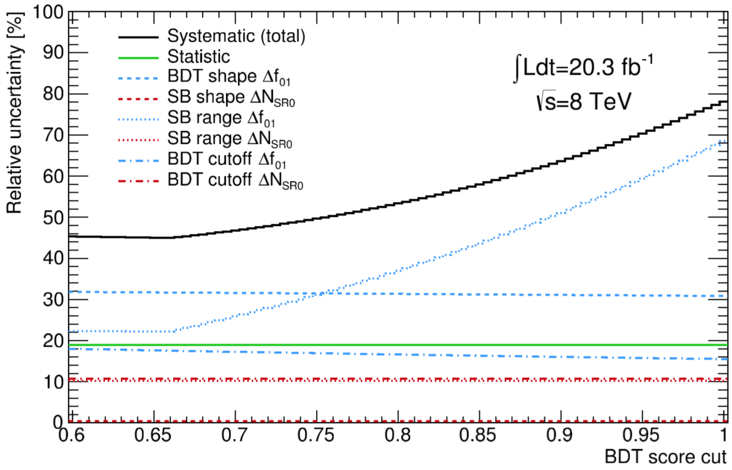


Figure 6.3: Uncertainty components for the variation of the x_0 (BDT cutoff) and SB range definition. The fluctuation of fit functions Δf for the variations and nominal configuration are considered as the uncertainties. The difference of the sideband events (N_{SR0}) is also considered. The total uncertainty of the background modeling is given by the combination of the statistical uncertainty and all components.

the expected signal ($N_{\mathrm{exp}}^{SR1,sig}$) and background ($N_{\mathrm{exp}}^{SR1,bkg}$) with the value. Not only the number of background events in SR1, but also the expected sensitivity was determined at this step. In order to optimize x_1, the exclusion limit with the assumption of no signal event is used as the criteria. The expected limit was estimated with the statistical and systematic uncertainty from background modeling and simulated cross-section of $W \to \tau\nu$. Figure 6.4 shows the expected exclusion limit as a function of x_1, which is set to be 0.933 to maximize the sensitivity. Hereafter, systematic uncertainty is estimated for this selection.

6.2 Estimation of $N_{W\to\tau\nu}$

To calculate the upper limit of the $\mathcal{BR}(\tau \to 3\mu)$, the number of observed events is normalized by the generated number of $W \to \tau\nu \to 3\mu\nu$ events ($\mathrm{N}_{W\to\tau\nu}$). Precise estimation of the cross-section of $W \to \tau\nu$ is necessary to set the limit on $\mathcal{BR}(\tau \to 3\mu)$. In the mass scale of the W boson, the difference of the masses of charged leptons can be assumed to be negligible. Therefore, the $\mathcal{BR}(W \to \tau\nu)$ is the same as that of $W \to e\nu$ and $W \to \mu\nu$. The average of these modes is denoted as $W \to l\nu$ inclusively. The cross-section of $W \to l\nu$ had been measured using the $W \to e\nu$ and $W \to \mu\nu$ modes in ATLAS at $\sqrt{s} = 7\,\mathrm{TeV}$ [84] as:

$$\sigma_{W\to l\nu}^{7\,\mathrm{TeV}} = 10.207 \pm 0.021(stat) \pm 0.121(syst) \pm 0.347(lumi) \pm 0.164(acc)\ \mathrm{nb}, \tag{6.9}$$

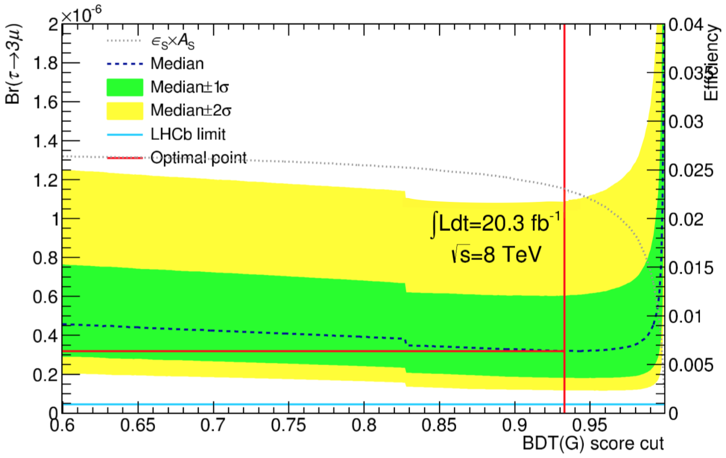


Figure 6.4: Sensitivity for $\mathcal{BR}(\tau \to 3\mu)$ as a function of BDT selection value (x_1). Yellow and green bands represent the significance range of 1 and 2σ for the sensitivity. Black dotted line shows the signal efficiency, which is related with the number of signal events. Blue and red solid lines describe the limits from LHCb experiment and expected by ATLAS.

where the uncertainties by *lumi* and *acc* are obtained by the luminosity measurement and detector acceptance, respectively. This is in agreement with the NNLO theoretical calculation of 10.46 ± 0.52 nb shown in Table 4.2. $\sigma^{7\,\text{TeV}}_{W\to l\nu}$ was estimated using this measured cross-section with scaling to account for the difference between 7 TeV and 8 TeV, based on the theoretical estimation. The number was determined as:

$$\begin{aligned} \sigma^{8\,\text{TeV}}_{W\to l\nu} &= \sigma^{7\,\text{TeV}}_{W\to l\nu}\frac{\sigma^{8\,\text{TeV},theory}_{W\to l\nu}}{\sigma^{7\,\text{TeV},theory}_{W\to l\nu}} \\ &= 11.89 \pm 0.24 \text{ nb.} \end{aligned} \tag{6.10}$$

This estimation is based on the measurement in the same experiment, therefore the systematic errors by experimental setup cancel. Furthermore, the largest uncertainty of the theoretical NNLO calculation is due to PDF, which also cancel in the ratio. With these reasons, systematic uncertainties were not considered in the cross-section, $\sigma^{8\,\text{TeV}}_{W\to l\nu}$. The total number of the $W \to \tau\nu$ events with $\mathcal{L}= 20.28\,\text{fb}^{-1}$ were given as:

$$\begin{aligned} \text{N}^{8\,\text{TeV}}_{W\to\tau\nu} &= \sigma^{8\,\text{TeV}}_{W\to l\nu} \times 20.28 \times 10^{6} \\ &= (2.41 \pm 0.08) \times 10^{8}. \end{aligned} \tag{6.11}$$

Chapter 7

Efficiency Measurement

The efficiency measurement is one of the most important subjects in the analysis. First, it is difficult to collect highly collimated low p_{T} muons for efficiency measurement. In order to ensure the robustness of the analysis, it is essential to establish a method to acquire enough amount of low p_{T} muons. In addition, highly collimated muons are hard to be recognized separately. The efficiency loss by this issue is called as close-by-muon inefficiency, and known to be significant at the trigger level and seen even at the reconstruction level if the opening angle of muons are very small. In order to validate the signal efficiency from the simulation, it is necessary to understand the close-by-muon inefficiency and compare between the data and simulation. The method of efficiency measurement, called Tag&Probe, and the efficiency evaluation at the trigger and reconstruction level are described in this chapter.

7.1 Tag&Probe method

In order to measure the efficiency of the target signature, bias in the efficiency measurement must be considered as the data themselves are acquired by the trigger requirement. The Tag&Probe method is used for the measurement of single muon trigger efficiency using pairs of muon from resonance decays to avoid such bias [85]. By requiring a muon in the pair to be triggered by a single muon trigger, the other muon is used for efficiency measurement. The trigger efficiency measurement for the signal event in this analysis is also based on this method using J/ψ meson decay.

7.1.1 Using $Z \to \mu\mu$ event

As the simple case, Tag&Probe procedure using $Z \to \mu\mu$ is described. As the first step of the method, the event was required to be acquired by a single muon trigger with 24 GeV threshold, named `EF_mu24i_tight`. This trigger chain also requires the isolation of the muon with no or a small energy deposit in calorimeter cells

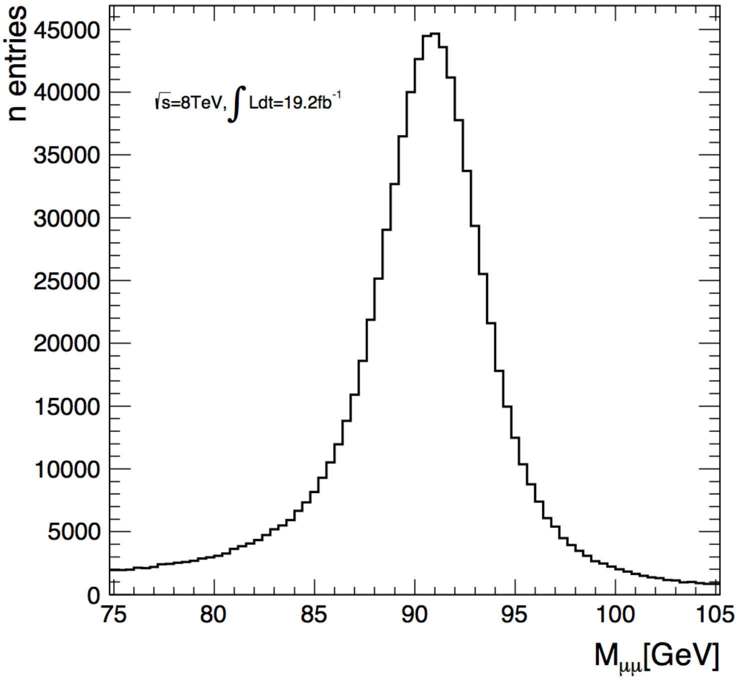


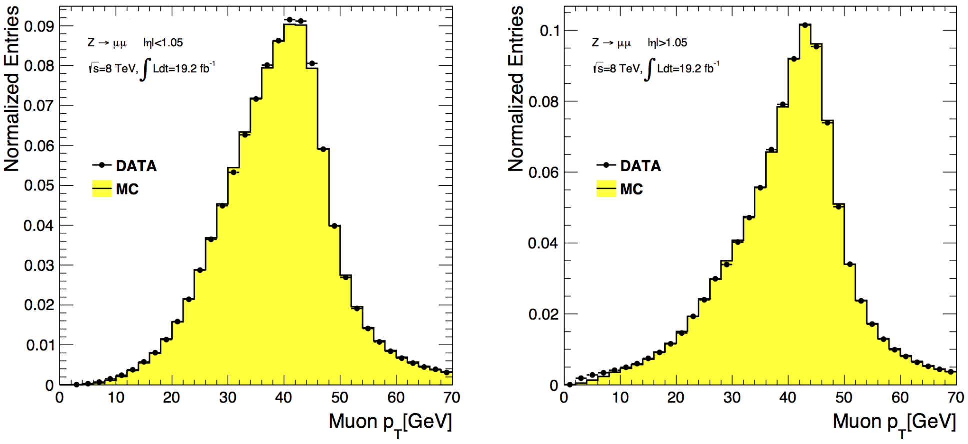
Figure 7.1: Invariant mass distribution of the data around Z mass.

around the muon direction. It is required that at least one opposite sign muon pair passes the following selections.

- Each muon has enough number of hits at each sub-detector of the ID
- $p_{\mathrm{T},\mu}>$ 10 GeV and $|\eta_\mu| < 2.5$ for each muon
- $80 < M_{\mu1,\mu2} < 100$ GeV for muon pair

The third cut selects muon pairs from Z boson decay, whose mass distribution is shown in Figure 7.1 peaks at 91 GeV. Among two muons, the one which fired `EF_mu24i_tight` trigger chain is defined as the tag muon, while the other muon is defined as the probe muon, which can be used for efficiency measurement without trigger bias. Distributions of the probe muon of $Z \to \mu\mu$ Tag&Probe is shown in Figure 7.2.

The $Z \to \mu\mu$ events have large statistics of high p_T muons above 10 GeV, while smaller statistics at low p_T below 10 GeV. Therefore, this method is not suitable for the analysis composed mostly at low p_T muons. This method was used only for validation of efficiency at the high p_T region. An alternative method is thus developed for low p_T muons from the decays of lighter mass resonances.



(a) (b)

Figure 7.2: p_{T} distributions of probe muons which are defined by Z Tag&Probe for the data (black points) and simulation (yellow filled histogram). The distributions at the barrel region and endcap region are shown in (a) and (b), respectively.

7.1.2 Using boosted $J/\psi \to \mu\mu$

As an alternative resonance decaying into muon pair with lower p_{T}, $J/\psi \to \mu\mu$ is the most suitable candidate. However, it was difficult to require the trigger with high p_{T} threshold and isolation for tag muon, because p_{T} of the tag muon is also low, typically at a few GeV. Single low p_{T} muon triggers, for example `EF_mu4`, are heavily prescaled due to high rate. Thus, high p_{T}, at least 20 GeV, is desirable for the tag muon. In the muon pair from decay of J/ψ with approximately 20 GeV at least one muon should have high p_{T} while another muon may have p_{T} of a few GeV. Hence, the boosted $J/\psi \to \mu\mu$ events enable to measure the efficiency for low p_{T} muons. This method is named as boosted $J/\psi \to \mu\mu$ Tag&Probe, whose procedure flow is almost the same as $Z \to \mu\mu$ Tag&Probe while the mass window is optimized for J/ψ as $2.8 < M_{\mathrm{J}/\psi} < 3.3$ GeV based on Figure 7.3. Figure 7.4 shows the p_{T} distributions of the probe muon selected by this method.

There are many muon pairs close to each other within a cone of $\Delta R_{\mu1,\mu2} < 0.3$ as shown in Figure 7.5. Therefore, the close-by-muon inefficiency can be evaluated by the probe muons. For this analysis, Correlation Factor (CF) is measured as a function of ΔR between tag and probe muons. CF represents the additional inefficiency for close-by-muon. Efficiency of muon at small ΔR is defined as $\mathrm{CF}_{\mathrm{tag,probe}} \times \epsilon_{probe}$.

These two important aspects for the analysis were measured precisely using boosted $J/\psi \to \mu\mu$ Tag&Probe.

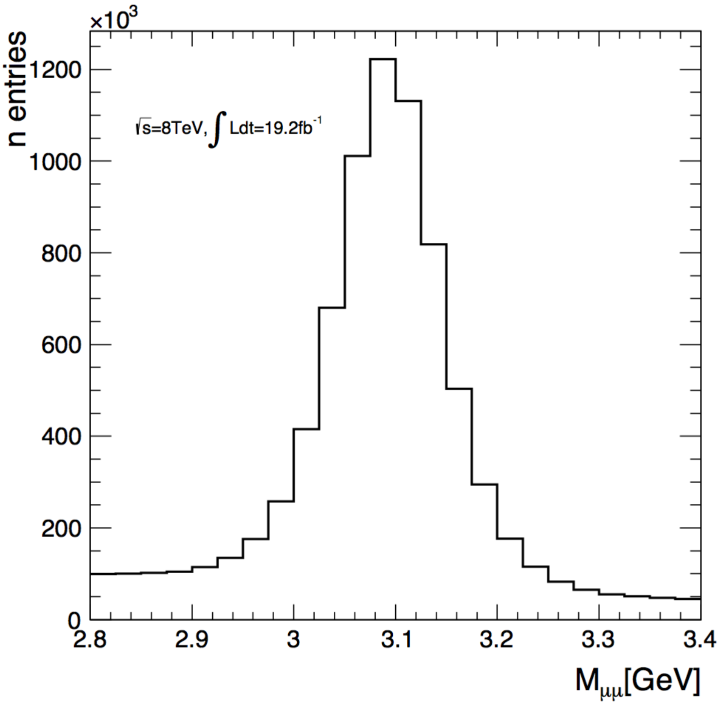
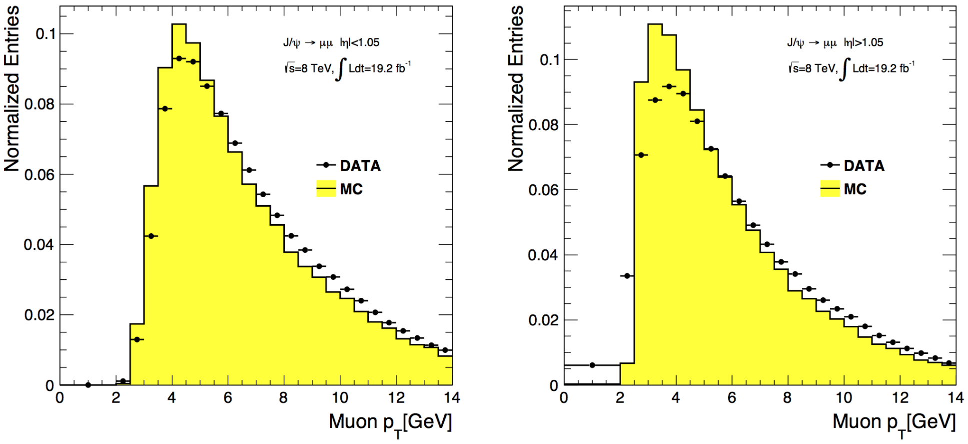


Figure 7.3: Invariant mass distribution of the data around J/ψ mass.



(a) (b)

Figure 7.4: p_{T} distributions of probe muons which are defined by Z Tag&Probe for the data (black points) and simulation (yellow filled histogram). The distributions at the barrel region and endcap region are shown in (a) and (b), respectively.

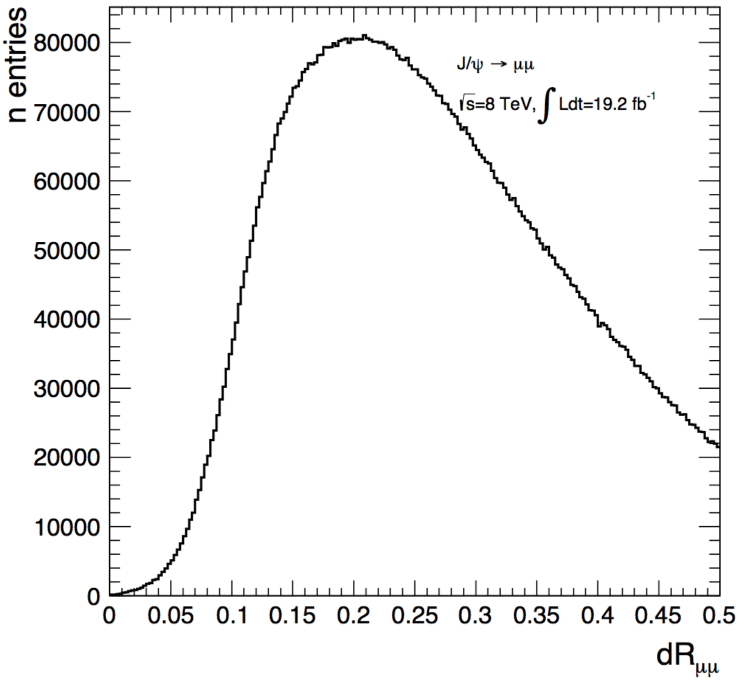


Figure 7.5: Distribution of ΔR between the tag and probe muons in the data.

7.2 Trigger efficiency measurement

It is not possible in practice to measure the efficiency with respect to true muon events using data, but the relative efficiency of the data to that of the simulation can be evaluated by the efficiency with respect to the reconstructed muon. Therefore, the trigger efficiency (ϵ_{trig}) of this measurement is defined as

$$\epsilon_{trig} = \frac{\text{Number of trigger passed and reconstructed muon}}{\text{Number of reconstructed muons}}. \tag{7.1}$$

Several trigger chains with multi-muon requirement were used in this analysis as shown in Section 4.1.2. Hence, efficiency of each trigger chain and the combination with *or* condition, as the difference between the data and simulation, need to be evaluated as well as systematic uncertainties. As the first step, the efficiencies of the single muon triggers are measured as a function of p_{T}, η, and ϕ of the reconstructed muons. The CF for muon pair was also defined as a function of $\Delta R_{\mu 1,\mu 2}$. The efficiency of the multi-muon trigger can be emulated by the single muon trigger efficiency and the CF, measured by the data and simulation. As the last step, combination of these triggers with *or* condition was estimated for the target signature for both data and simulation. In this section, this procedure is described in detail.

7.2.1 Single muon trigger with low p_{T} threshold

The efficiency of the low p_{T} threshold muon trigger was estimated by boosted $J/\psi \to \mu\mu$ Tag&Probe method. In this method, two types of J/ψ specific trigger chains were used to acquire the data.

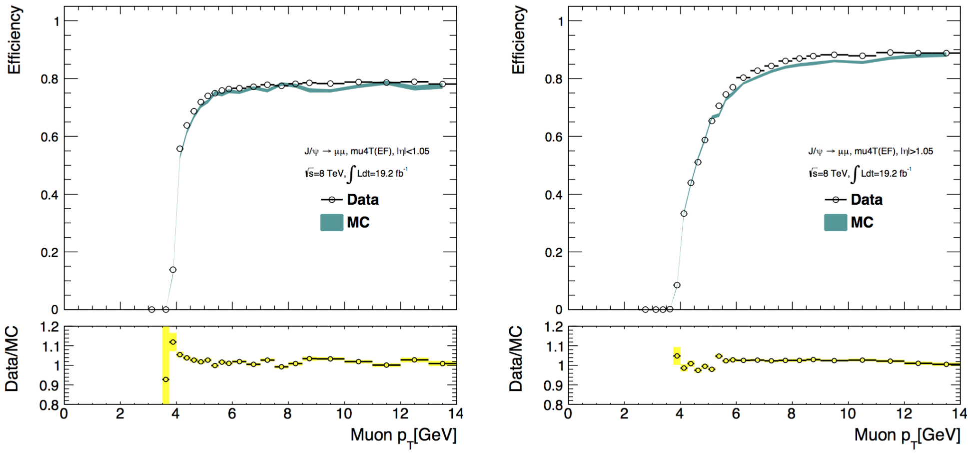
- `EF_mu18_tight_Jpsi_EFFS`
- `EF_mu18_tight_L2_2mu4T_Jpsimumu`

`EF_mu18_tight_Jpsi_EFFS` is designed based on `EF_mu18_tight`, which is the single muon trigger with the threshold for p_{T} of 18 GeV. But the other muon is required at the EF level and the invariant mass of these two muons is required to be around the J/ψ mass (from 2.5 GeV to 4.0 GeV). Therefore, efficiencies at L1 and L2 steps can be measured by requiring tag muon to pass `EF_mu18_tight`. On the other hand, `EF_mu18_tight_L2_2mu4T_Jpsimumu` required two muons with the invariant mass around J/ψ mass at the L2 step. Only EF efficiency of the target trigger can be measured by the triggered event by the same way as the former case. The total efficiency of a trigger chain was calculated as a product of the L2 efficiency with respect to the probe muon and the EF efficiency with respect to the L2 passed probe muon. Thus, using the combination of independent measurements, the threshold for tag muon can be kept lower without isolation requirement like `EF_mu24i_tight` in Z Tag&Probe. In this case, $\Delta R_{\mu1,\mu2} > 0.25$ was required for the independent efficiency measurement from close-by muon.

Single muon trigger component needed for the emulation of the efficiency of the target trigger chain were; `HLT_mu4`, `HLT_mu6`, `HLT_mu10`, `HLT_mu18`, and `HLT_mu24`. The trigger efficiency at the barrel and endcap regions are clearly different due to the different muon spectrometer acceptance and configuration, therefore the efficiency should be evaluated separately. The efficiency of the `HLT_mu4` at the barrel and endcap regions are shown as a function of p_{T} of the probe muon in Figure 7.6 for both data and simulation. The simulated $J/\psi \to \mu\mu$ events were filtered by the p_{T} cuts on both muons as $p_{\mathrm{T},\mu1} > 15$ GeV and $p_{\mathrm{T},\mu2} > 2.5$ GeV at generator level. The difference of the plateau value and shape of the turn-on curve between the barrel and endcap are clear, while the data and simulation agree well in both regions.

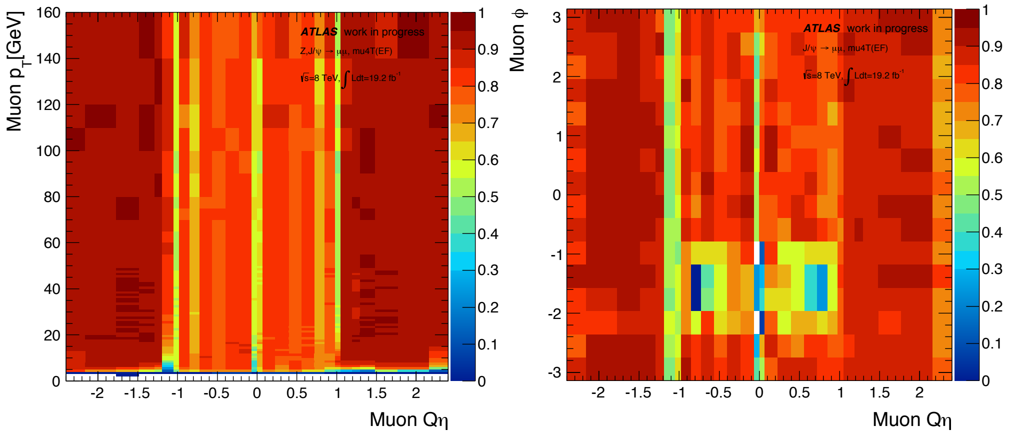
The efficiency depends on the detector geometry of the muon spectrometer. Figure 7.7 shows the `HLT_mu4` efficiency maps on the $\mathrm{Q}\times\eta - p_{\mathrm{T}}$ and $\mathrm{Q}\times\eta - \phi$ planes, where Q is the charge of the probe muon. There are inefficient area at $\eta \sim 0$, $|\eta| \sim 1.0$, and $-2 < \phi < -1$, which are the crack region of the barrel, transition region of the barrel and endcap, and feet region of the ATLAS detector, respectively. Using the $\mathrm{Q}\times\eta$ instead of η, dependence on the bending direction by the toroidal magnetic field could be evaluated in these figures. Those features of the muon trigger efficiency necessitate the precise parametrization for $\mathrm{Q}\times\eta$ and ϕ as well as p_{T} of the muon.

In addition, the EF efficiency without L1 nor L2 requirements is necessary for the estimation of the Full-Scan trigger, for example `EF_mu18_tight_2mu4_EFFS`.



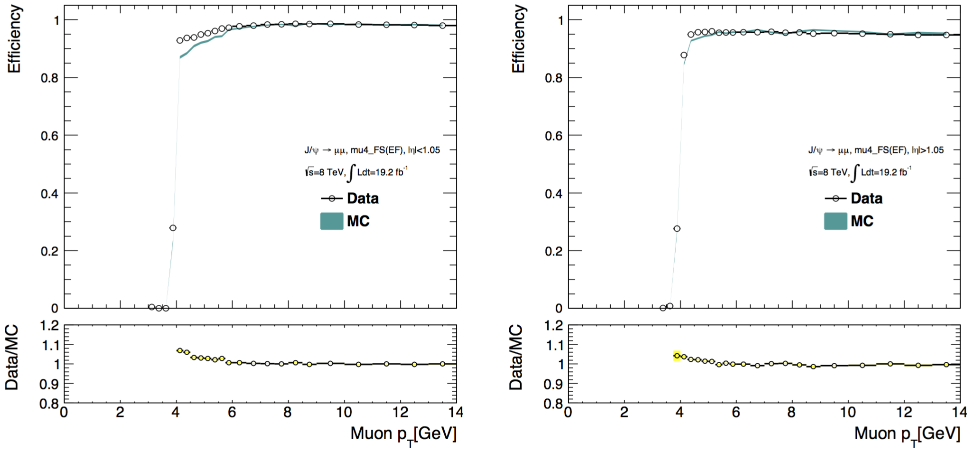
(a) (b)

Figure 7.6: p_{T} dependence of `HLT_mu4` efficiency is shown for (a) barrel region and (b) endcap region. The plateau values of the trigger at the barrel and endcap are 80% and 90%, respectively.



(a) (b)

Figure 7.7: Efficiency map of the `HLT_mu4` as a function of (a) Q×η−p_{T} and (b) Q×η−ϕ. For high p_{T} muons of above 20 GeV, efficiency dependency was symmetric for ϕ because the degree of bend in the magnetic field was negligible with respect to geometrical difference. The specific detector construction around the middle of the detector, and feet of the detector, and transition region of the barrel and endcap are also clearly seen as inefficient regions.



(a) (b)

Figure 7.8: p_{T} dependency of `mu4_FS` trigger efficiency is shown for the (a) barrel region and (b) endcap region.

The specific part of the trigger is noted as `mu4_FS`, whose efficiency was measured by the latter method of the boosted J/ψ Tag&Probe. Although this efficiency was measured with respect to L2 passed events, the difference is negligible because L2 requirement is much looser than the requirement of the offline reconstruction. Figure 7.8 shows the efficiency of `mu4_FS` is approximately 95% in both regions. Although these efficiencies barely depend on the Q×η and ϕ compared to the nominal efficiency, they are also parametrized with the same variables of muons.

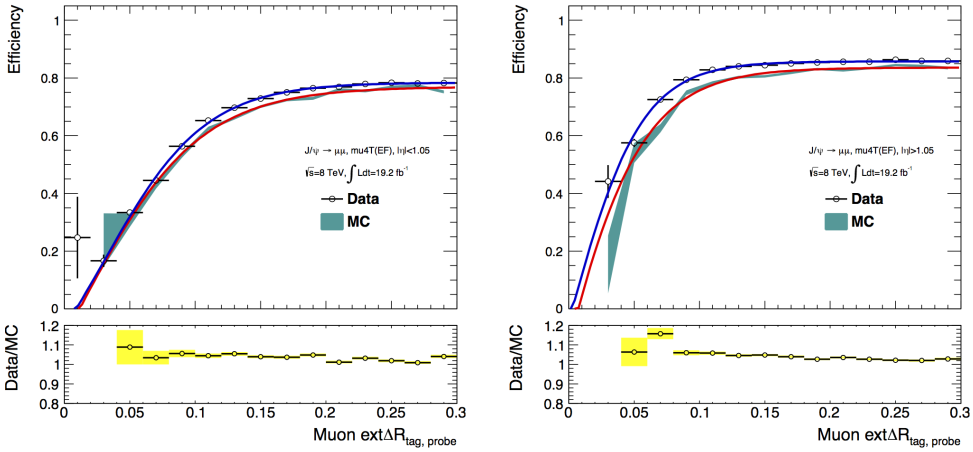
7.2.2 ΔR dependence of trigger efficiency

The efficiency of the single muon trigger is usually measured with an isolation cut which requires that there is no nearly muon around the probe muon. On the other hand, in order to measure the CF for muon pairs as a function of $\Delta R_{\mu1,\mu2}$, ΔR requirement for the single muon trigger efficiency was removed. When the tag and probe muons were close to each other, the efficiency for the probe muon was not independent from the tag muon and defined as

$$\epsilon'_{trig} = \epsilon_{trig} \times CF(\Delta R_{\mu1,\mu2}). \tag{7.2}$$

ΔR dependence of the $\epsilon'_{\texttt{HLT\_mu4}}$ was validated below 0.2, and stable above 0.25 as shown in Figure 7.9. This correlation is understood by the size of the RoI given by the L1 trigger. Muons in the same RoI region are recognized as one muon at the L1 trigger. The difference between the barrel and endcap regions and the end point of the upturn are explained by the typical RoI sizes at the barrel and endcap regions which are about 0.15 and 0.1, respectively. The shape of the ΔR dependence is determined by the fit with a Fermi function $F(\Delta R_{\mu1,\mu2})$ as

$$F(\Delta R_{\mu1,\mu2}) = \frac{A}{\exp(-B \times \Delta R_{\mu1,\mu2} + C) + D} + E, \tag{7.3}$$



(a) (b)

Figure 7.9: ΔR dependence between tag and probe muons is shown for (a) barrel region and (b) endcap region. The fitting functions for the data and simulation are shown as the blue and red lines, respectively.

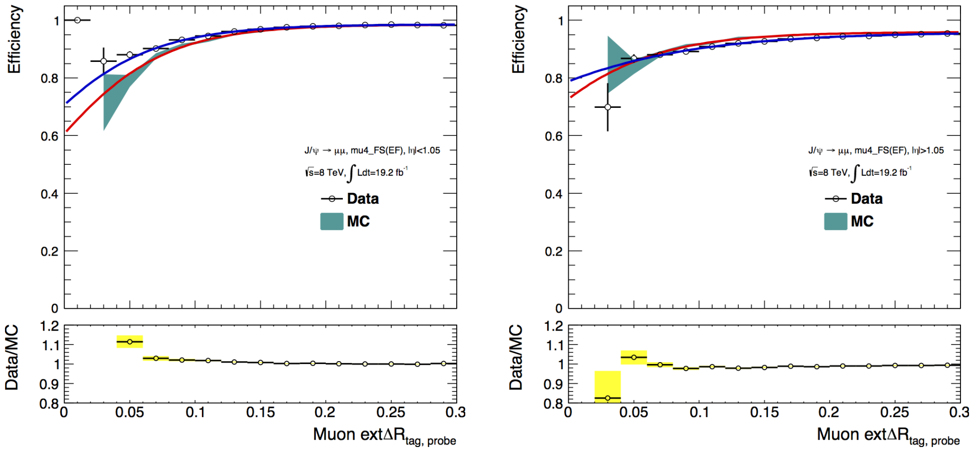
where five free parameters, A, B, C, D, and E, were used in the fit. Normalizing the efficiency at plateau to 1.0, the CF is defined as

$$CF(\Delta R_{\mu 1,\mu 2}) = \frac{D}{A + D \times E} F(\Delta R_{\mu 1,\mu 2}). \tag{7.4}$$

In order to evaluate the difference of the simulation from the data, this factor should be measured with simulation and data separately for the endcap and barrel regions. Moreover, these variations for the FS case must be determined separately as shown in Figure 7.10. In the end, eight CFs are determined for combination of (normal, FS) × (data, simulation) × (barrel, endcap) as shown in Table 7.1.

Table 7.1: Fitting results of the ΔR dependence each combination

Type	Sample	Region	A	B	C	D	E
Normal	Data	Barrel	1.00	26.0	0.539	- 0.816	-0.446
		Endcap	1.24	36.0	-0.410	0.682	-0.970
	Simulation	Barrel	1.14	22.1	-0.0547	0.738	-0.775
		Endcap	1.43	29.0	-1.24	0.526	-1.87
FS	Data	Barrel	0.836	22.1	0.863	0.860	-0.0280
		Endcap	0.820	12.4	-1.63	0.875	-0.0220
	Simulation	Barrel	0.826	22.1	-0.572	0.863	-0.0283
		Endcap	0.820	19.5	-1.24	0.874	-0.0221



(a) (b)

Figure 7.10: ΔR dependency between tag and probe muons of the full scan (FS) trigger case is shown for (a) barrel region and (b) endcap region. The fitting functions for the data and simulation are shown as the blue and red lines, respectively.

7.2.3 Efficiency estimation of multi-muon trigger

Combining the efficiency of the single muon trigger and CF for close-by two muons measured by the boosted J/ψ Tag&Probe, the multi-muon trigger efficiency can be emulated. If there are two reconstructed muons in an event, the emulated trigger efficiency for the event is defined as

$$\epsilon_{\texttt{2muX},\mu1,\mu2} = \epsilon_{\texttt{muX},\mu1} \times \epsilon_{\texttt{muX},\mu2} \times CF(\Delta R_{\mu1,\mu2}), \tag{7.5}$$

where $\epsilon_{\texttt{muX},\mu}$ represents the `HLT_muX` efficiency for p_{T}, η, and ϕ of each muon. This equation form with three muons is complicated because many patterns of the combination of the passed muons. Also, the calculation of the `HLT_2muX`, `HLT_3muX` and Full-Scan triggers were determined in different ways due to the same reason. All these equations are written in Appendix. In this method, $\epsilon_{\texttt{Trigger},\mu1,\mu2,\mu3}$ is determined for the three muon candidate in each event, but not the efficiency for the signal events. The total signal efficiency $\epsilon_{\texttt{Trigger},repro}$ can be estimated by the sum of efficiency for each event, instead of the fraction of the trigger-passed events in the generated events which is used in the analysis procedure ($\epsilon_{\texttt{Trigger},trigmatch}$). These definitions can be described as

$$\epsilon_{\texttt{Trigger},repro} = \frac{1}{N}\sum_{i=1}^{N} \epsilon^{i}_{\texttt{Trigger},\mu1,\mu2,\mu3} \tag{7.6}$$

$$\epsilon_{\texttt{Trigger},trigmatch} = \frac{1}{N}\sum_{i=1}^{N} 1(passed)\ or\ 0(not\ passed), \tag{7.7}$$

where N represents of the number of target events. The former efficiency is based on the "measured" efficiency by both data and simulation, while the latter

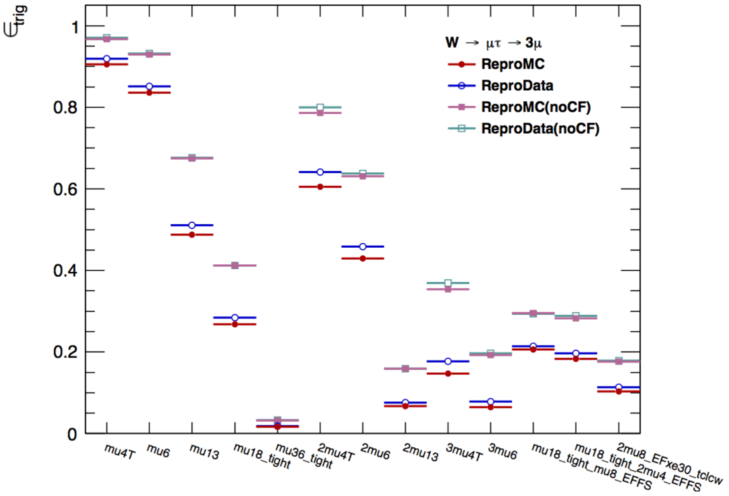
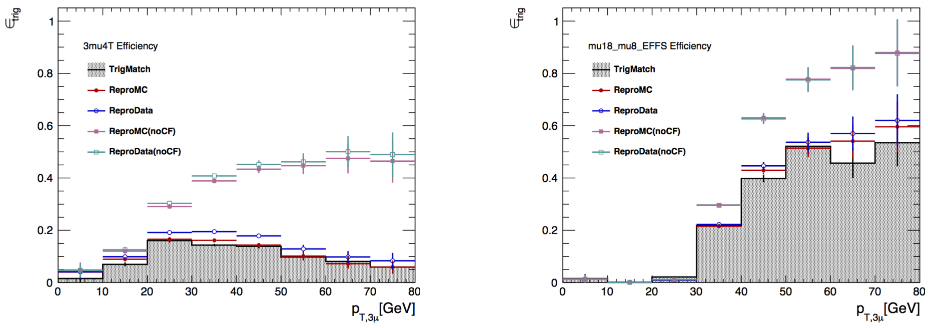


Figure 7.11: The emulated efficiency of several kinds of trigger chains for W originated signal. Red and blue histograms show the efficiency considering the CF from the data and simulation, respectively. The results without CF are shown as cyan and pink for the data and simulation, respectively.

one is estimated by the simulation. Therefore, the difference between the data and simulation can be evaluated by the comparison between the data-based and simulation-based $\epsilon_{\texttt{Trigger},repro}$. $\epsilon_{\texttt{Trigger},trigmatch}$ is also used for validation of the method by comparing with the simulation-based $\epsilon_{\texttt{Trigger},repro}$. Figure 7.11 shows the emulated efficiency for the various trigger chains for the data and simulation, and the results without CF consideration are also shown for the confirmation of the close-by muon inefficiency. Effective correction for the multi-muon trigger chains are clear in the plot. Efficiencies of single muon trigger chains are also suppressed by CF which can be understood by close-by muon inefficiency at L1 for the components. The EFFS trigger is less affected by the CF like a single muon trigger, nevertheless these required multi-muons. This is because the close-by inefficiency for the EFFS is less effective as shown in Figure 7.10.

As an alternative validation of the CF, $p_{\rm T}$ dependence of the emulated efficiency was also calculated as shown in Figure 7.12. The CF affects the efficiency heavily because three muons composing high $p_{\rm T}$ candidates is highly boosted. In Figure 7.12, emulated trigger efficiency is also compared with the true efficiency by the simulation as a fraction of trigger passed events. The comparison between emulated efficiency with simulation templates and the true efficiency ensures the accuracy of this method.

For the analysis, the estimation of the inclusive efficiency with all trigger chains was necessary. The inclusive efficiency was estimated using the scale factor (SF) defined as the ratio of the emulated efficiency by the data to that of the



(a) (b)

Figure 7.12: The efficiencies of (a) `HLT_3mu4T` and (b) `HLT_mu18_mu8_EFFS` are shown as a function of p_{T} of triplets. Black filled histograms show the true efficiency given by the simulation, and other colored histograms are emulated efficiency with the same color convention of the Figure 7.11. A good agreement of emulation considering the CF with the efficiency in the simulation indicates the good accuracy of the method.

simulation. When an event passed several triggers, the event is weighted by the average of the SF for them. Figure 7.13 describes the efficiency for these triggers and inclusive total efficiency. The efficiency and SF for both of each trigger chain and inclusive values are summarized in Table 7.2. Difference of the inclusive efficiencies with and without SF was considered as the systematic uncertainty in the trigger efficiency of 11%.

Table 7.2: Summary table of the signal efficiency and SF for each trigger chain. Total efficiency and SF are also shown.

Trigger chain	ϵ_{trig}[%]	SF$_{trig}$
`HLT_2mu13`	9.1	1.13±0.18
`HLT_3mu4T`	16.3	1.20±0.10
`HLT_3mu6_msonly`	6.9	1.20±0.16
`HLT_mu18_tight_mu8_EFFS`	21.2	1.04±0.01
`HLT_mu18_tight_2mu4_EFFS`	18.6	1.07±0.06
`HLT_2mu8_EFxe80_tclcw`	12.4	1.09±0.09
Inclusive	42.3	1.11±0.09

7.3 Reconstruction efficiency measurement

Efficiency of the muon reconstruction was also measured using boosted J/ψ Tag&Probe. The reconstruction efficiency can be simply defined as a ratio of the

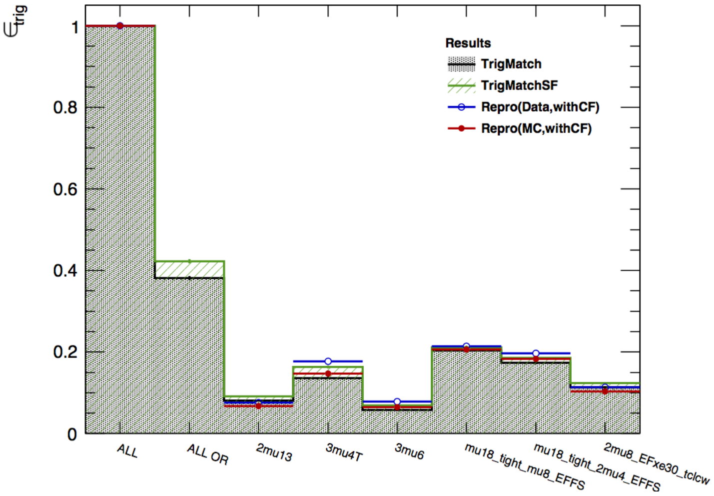


Figure 7.13: The efficiency of each trigger and inclusive usage. The emulated efficiency for each trigger is shown with templates by data (blue) and simulation (red). The efficiency with the simple trigger requirement is shown as black for all triggers and inclusive. In order to evaluate the inclusive SF for OR'ed product of the next to the first bin, the green histogram is the simple efficiency weighted by the SF.

reconstructed by the ID and MS to the truth particles. However, the efficiency of muon reconstruction at the ID is nearly 100% and truth particles cannot be used for the data. Therefore, as an approximation, the reconstruction efficiency(ϵ_{reco}) is

$$\epsilon_{reco} = \frac{\text{Number of reconstructed muon defined as at MS and ID}}{\text{Number of reconstructed muons at ID}}. \tag{7.8}$$

For this reason, J/ψ candidates in this measurement were reconstructed with the pair of a muon and a track without MS requirement. This track was required to pass the same selection of the quality as in the event selection. The event trigger was `HLT_mu24_tight`, which is rather similar to Z Tag&Probe but without isolation requirement. This setup is different from the trigger case because they cannot be required to have two muons at any trigger levels.

7.3.1 Measurement of the efficiency and ΔR dependence

The probe of the method is a track reconstructed at least in the ID. There are large number of background events in this case with non-muon ID track as the probe as shown in Figure 7.14. Therefore, the background contamination should be subtracted in order to extract the number of muon tracks for the reconstruction efficiency calculation. The subtraction was executed considering the p_{T} (ΔR) dependence of the purity for the measurement of p_{T} (ΔR) dependence of the efficiency. Mass distribution of each p_{T} (or ΔR) bin was independently fitted with a function composed by a Gaussian and an exponential. The Gaussian component was treated as the muon from J/ψ, while the exponential component was considered as background including non-muon tracks and subtracted as non-J/ψ. This fitting procedure is also applied for the numerator which is required to match with muons. Mass distributions for each p_{T} bin of the denominator (ID track) and numerator (muon track) are shown in Figure 7.15 together with their fitting results. (The binning of the p_{T} distribution is the same as shown in Figure 7.16.) The subtracted p_{T} distribution is shown in Figure 7.16. The reconstruction efficiency was calculated as the ratio defined in Equation 7.8 after subtraction of background in denominator and numerator. On the other hand, there was no underlying background event in the $J/\psi \to \mu\mu$ simulation sample, therefore this subtraction procedure was not applied for the efficiency calculation using the simulation. Figure 7.17 shows the reconstruction efficiency of muons as a function of muon p_{T} and ΔR between two muons.

7.3.2 Uncertainty of the close-by muon reconstruction

The dominant aspect of the reconstruction efficiency in our case is the uncertainty for the collimated muons with small ΔR rather than dependence of the muon p_{T}. Therefore, the difference between the data and simulation at small ΔR should be considered in this analysis since the maximum ΔR of three muons of target

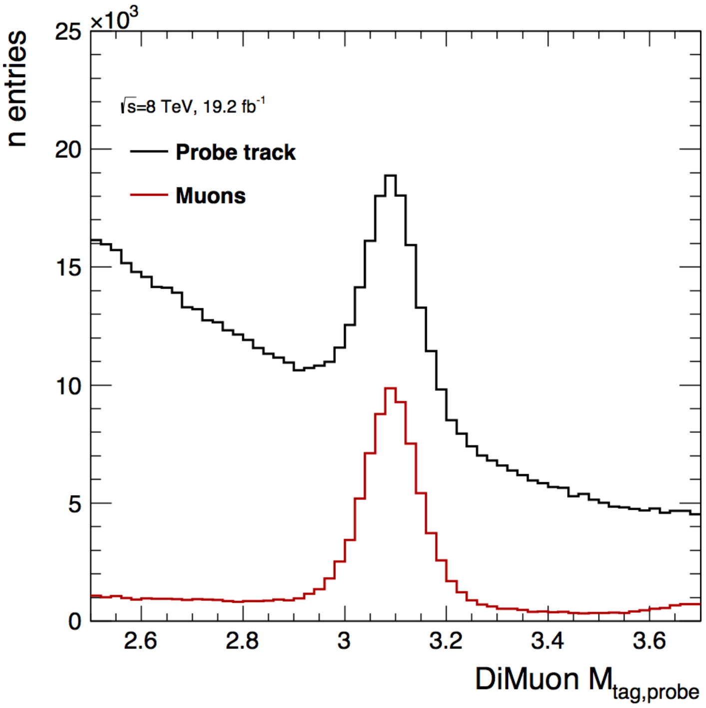


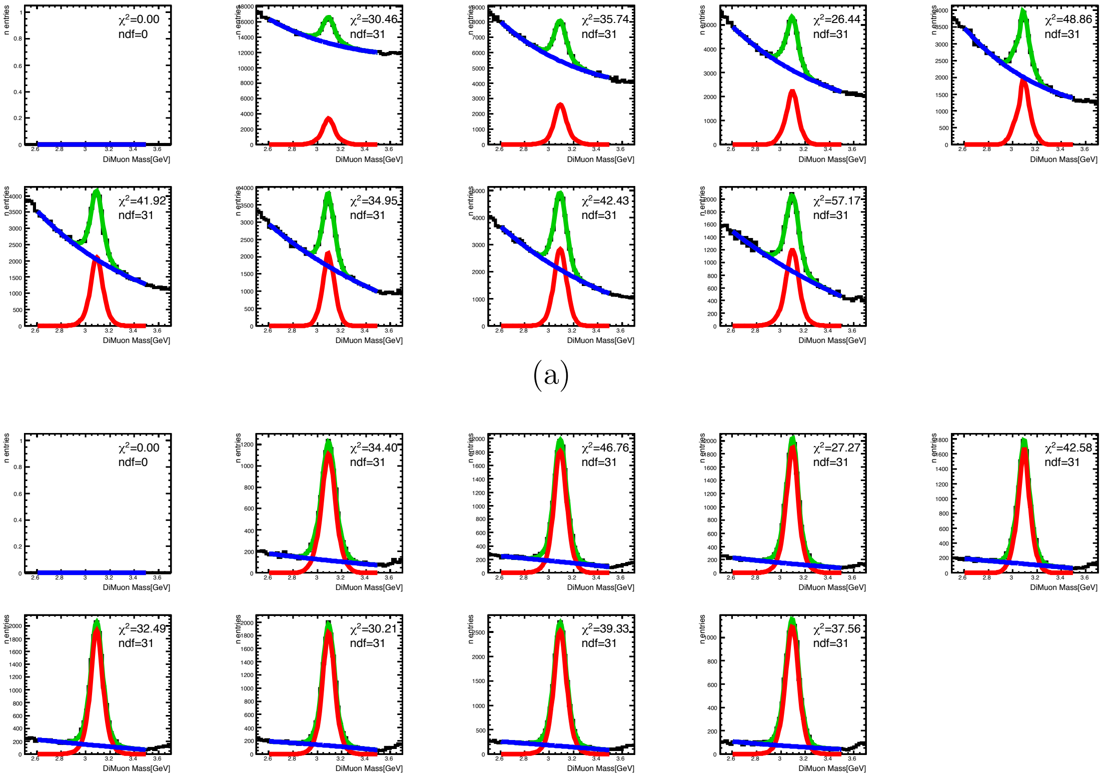
Figure 7.14: Mass distributions with the probe ID tracks and muon tracks as black and red, respectively. It is clearly seen that the probe tracks include large number of background by non-muon tracks.

signal is 0.3 at most as shown in Section 5.2. Therefore, the fluctuation for the collimated muons was estimated using the first three bins of Figure 7.17(b) as

$$\sigma_{reco} = 1.0 - \frac{\sum_{i=1}^{3} R_i/\sigma_i}{\sum_{i=1}^{3} 1.0/\sigma_i} \tag{7.9}$$

$$= 0.08 \pm 0.05, \tag{7.10}$$

where R_i and σ_i are the ratio of efficiency and statistical error of the i-th bin, respectively. However, statistical uncertainties for each bin of the data and simulation are relatively large. Therefore, the systematic uncertainty for this analysis is set as the maximum value of the fluctuation $8 + 5 = 13\%$ conservatively.



(a)

(b)

Figure 7.15: Mass distribution in each p_{T} region for (a) the denominator and (b) numerator of the efficiency calculation. The fitting results are shown as red, and signal and background components are also shown as green and blue, respectively. The first bin shows for the range from 0 GeV to 2.5 GeV, therefore the no entries are shown at the both regions.

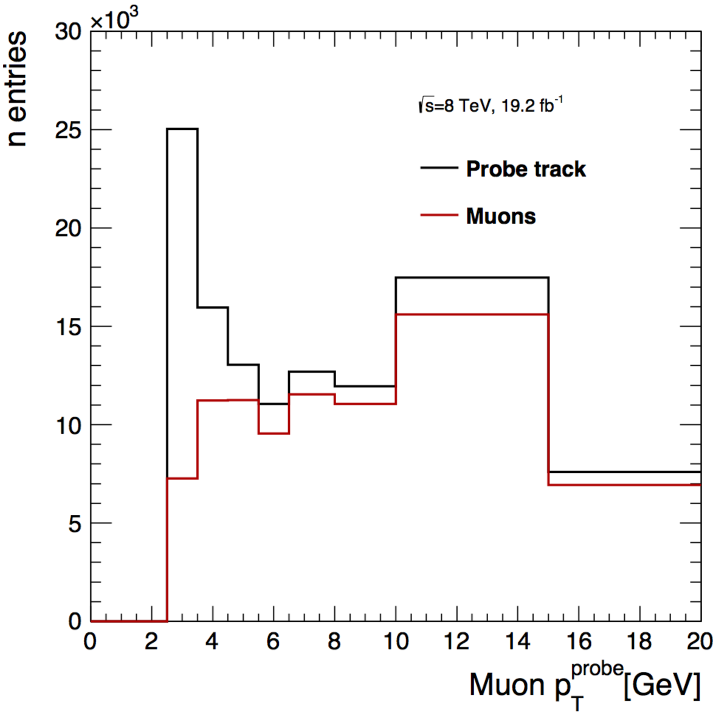
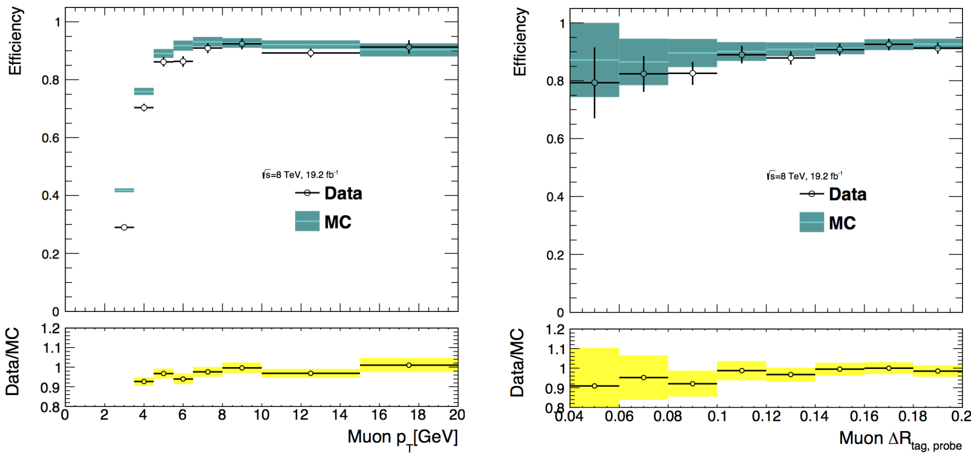


Figure 7.16: p_{T} distribution of the probe tracks of this measurement. Black and red histograms show the inclusive probe tracks and the probe tracks which is matched with muons. The bin ranges of low and high p_{T} are set to be different in order to make the entries in each bin comparable.



(a)

(b)

Figure 7.17: (a) p_{T} and (b) ΔR dependence of the reconstruction efficiency after the bin-by-bin subtraction. Results of data and simulation are shown as black points and green lines, and the ratio of these are also shown in the bottom panel as black points with yellow error bands. These show a good agreement, however the difference between data and simulation was taken as uncertainty in the next section.

Chapter 8

Result

In this chapter, the final result is described after all event selections and evaluation of the uncertainty. The result of the search is shown as $\mathcal{BR}(\tau \to 3\mu)$ by the statistical evaluation. The expected sensitivity with or without the signal events in the signal region and the results for this analysis will be shown, as well as the method for the determination.

8.1 Expected numbers for the $\mathcal{BR}(\tau \to 3\mu)$

$\mathcal{BR}(\tau \to 3\mu)$ is calculated by three numbers; acceptance times efficiency $(A \times \epsilon)$, number of observed events N_{obs}, and number of $W \to \tau\nu$ events $N_{W\to\tau\nu}$:

$$\mathcal{BR}(\tau \to 3\mu) = \frac{N_{obs}}{(A \times \epsilon) N_{W\to\tau\nu}} \tag{8.1}$$

$A \times \epsilon$ and $N_{W\to\tau\nu}$ are determined in Chapter 7 and Chapter 6. Therefore, the expected limit can be set for the expected number of events in the signal region (SR). Number of background events N_{bkg} are used for the evaluation of the signal strength of the observed events. These numbers are summarized in Table 8.1.

8.2 Summary of systematic uncertainties

Systematic uncertainties in $A \times \epsilon$, $N_{W\to\tau\nu}$ and $N_{bkg}(SB)$ are evaluated separately and summarized in Table 8.2.

Table 8.1: $A \times \epsilon$, $N_{W\to\tau\nu}$, and N_{bkg} in the SR and SB determined by the measurement in the previous sections.

$A \times \epsilon$	2.31%
$N_{W\to\tau\nu}$	2.41×10^8
$N_{bkg}(SR)$	0.193
$N_{bkg}(SB)$	0.726

Table 8.2: Summary table of the systematic uncertainties.

Source	Uncertainty[%]
$A \times \epsilon$	
Reconstruction efficiency	13
Trigger efficiency	11
Signal modeling	4
Jet energy	2.1
$N_{W \to \tau\nu}$	
Extrapolation from 7 TeV	3.9
N_{bkg}	
Background modeling	68

The first category summarizes the systematic uncertainties in $A \times \epsilon$. This comes from the efficiency and resolution of the measurements for each physics object. Uncertainties associated with the efficiency was estimated by the difference between the data and simulation as described in Chapter 7 to be 11% and 13% for the trigger and reconstruction, respectively. The uncertainty in the measurement of hadron jets and $E_{\mathrm{T}}^{\mathrm{miss}}$ was evalulated to be small ($\sim 2.1\%$). This estimation is based on the measurement of Jet Energy Scale (JES) and Jet Energy Resolution (JER) [86]. JES is obtained by the energy calibration of the hadron jet based on the energy deposit at the calorimeters. Biases of the energy scale and resolution were evaluated from comparison of the data with the simulation and assigned as the systematic uncertainties on the JES and JER.

The modeling of the simulation is also considered as a source of systematic uncertainty. Because the selection efficiency for the signal can be estimated only by the simulation, $W \to \mu\nu$ event candidates are used in order to validate it with the data in comparison with the simulation. As the SV for the event cannot be defined differently from $W \to \tau\nu \to 3\mu\nu$ signature, it is not possible to follow the full analysis procedure within an independent validation sample. Instead, it is alternated by the BDT selection without SV-based variables. The uncertainty on the signal modeling is determined to be 4% using the difference of the selection response between $W \to \mu\nu$ and $W \to \tau\nu \to 3\mu\nu$ signals.

The uncertainties in the $N_{W \to \tau\nu}$ and N_{bkg} are described in the previous sections. The uncertainty associated with the $N_{W \to \tau\nu}$ determination was estimated to be 3.9% based in Section 6.2, considering the precision of luminosity measurement for each year. As shown in Section 6.1, the uncertainty by the background modeling is 68% which is the dominant component.

Using the numbers and their uncertainties, the upper limit for $\mathcal{BR}(\tau \to 3\mu)$ is set as follows

$$(A \times \epsilon) \cdot N_{W \to \tau\nu} = (5.567 \times 10^6) \pm 18\%_{syst.}, \tag{8.2}$$

$$N_{bkg}(SR) = 0.193 \pm 19\%_{stat.} \pm 68\%_{syst.}. \tag{8.3}$$

Then, 6 (11) events in the SR is needed for the observation of the excess with significance of 3 (5)σ, which means that $\mathcal{BR}(\tau \to 3\mu) \sim 1.0(1.9) \times 10^{6}$. On the other hand, the expected limit when N_{obs} is the equivalent to $N_{bkg}(SR) = 0.193$ is given as 3.94×10^{-7} at the 90% CL[1] by the CLs method [87].

8.3 Result of the $W \to \tau\nu \to 3\mu\nu$ search

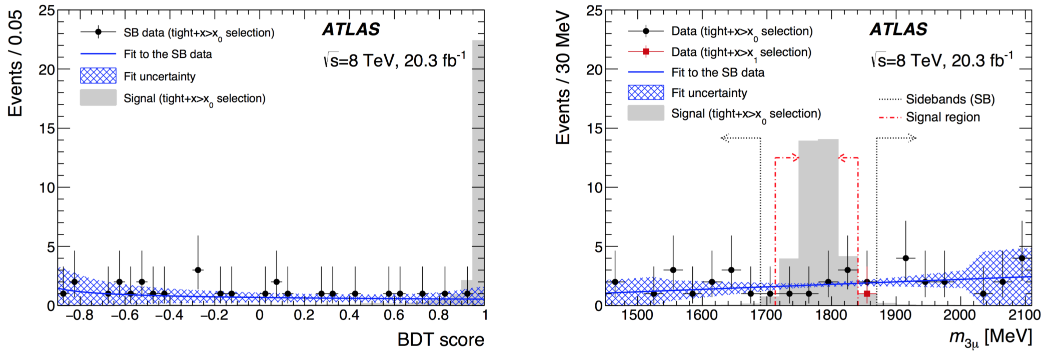
In order to derive the result, N_{obs} is obtained by the selection with BDT score of $x_1 = 0.933$ as described in Section 6.1.2. The result with loose selection with $x_0 = -0.9$ is also shown as the validation of the sideband and the BDT shape fit. Figure 8.1 shows the distributions of BDT score and invariant masses with the cut values of x_0 and x_1. Finally, no events were found in the SR after BDT cut, however one event candidate remained in the crack region between the SB and the SR [88]. Although this is not the candidate of the target signal because this region is not used for the signal counting by the definition, the event display is shown in Figure 8.2. This event is composed of collimated three muons isolated from hadron jets and large $E_{\mathrm{T}}^{\mathrm{miss}}$. This event is presumed to be from a decay of heavy flavor hadron into a very small number of energetic non-muon tracks with a particle not detected in the opposite-side (which could be neutrinos). The emulation of such case is difficult as it is a quite rare situation. The result, considering the event in the crack, shows no evidence for $\tau \to 3\mu$ signature. The upper limit at 90% CL is set to be, with $N_{\mathrm{obs}} = 0$,

$$\mathcal{BR}(\tau \to 3\mu) < 3.76 \times 10^{-7}. \tag{8.4}$$

This is the first result of the search for $\tau \to 3\mu$ at the ATLAS experiment, using the new approach with the characteristic signature of the boosted three muons.

---

[1] Usually, upper limits are set at 95% CL in ATLAS. However, the limit for the $\tau \to 3\mu$ search in the past is set at 90% CL, hence, this benchmark is employed in this search.



(a) (b)

Figure 8.1: (a) Sideband fit result for the BDT score distribution, and (b) the result as the triplet mass distribution. In the left figure, the distribution and fit result with the uncertainty are shown after the loose selection by $x_0 = -0.9$. The final results show the both selection by x_0 and x_1 as black and red markers, respectively. There are no events in the SR, while one event in the crack region between the SR and SB. Therefore, this shows no evidence of the existence of $\tau \to 3\mu$ signature.

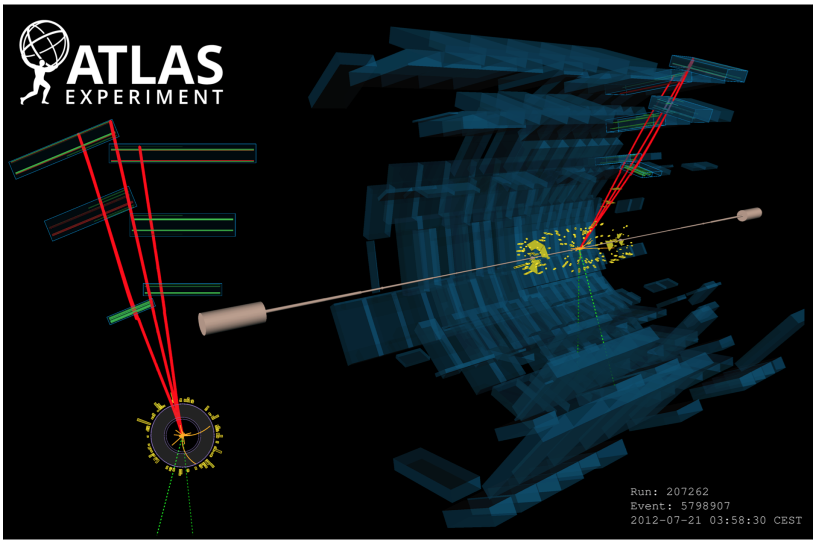


Figure 8.2: Event display of the event in the crack region. This event is composed by three muons (red lines), large $E_{\mathrm{T}}^{\mathrm{miss}}$ (green dashed lines) for Calo and Track, and small number of non-muon tracks. This is similar to the target signature, however this can be understood as an accidental event by a decay of heavy flavor hadron with the mis-reconstructed or undetectable particles. Although it is difficult to simulate such rare case, the number is in the uncertainty range.

Part III

Prospect of the sensitivity in Run2 with $\sqrt{s} = 13\,\mathrm{TeV}$

Chapter 9

Data and Monte-Carlo Simulation

The ATLAS experiment plans to collect data of integrated luminosity of $100\,\mathrm{fb}^{-1}$ at the center-of-mass energy of 13-14 TeV in 2015-2018 (Run2). For the $\tau \to 3\mu$ search, a significant improvement is expected with the higher luminosity and energy than those in 2012. Especially for the HF originated signature, which is associated with the QCD interaction, the production rate will increase dramatically. Therefore, the sensitivity should be further improved by an inclusive search with W and HF-originated signatures. The status of the data and simulation for the Run2 analysis is described in this chapter.

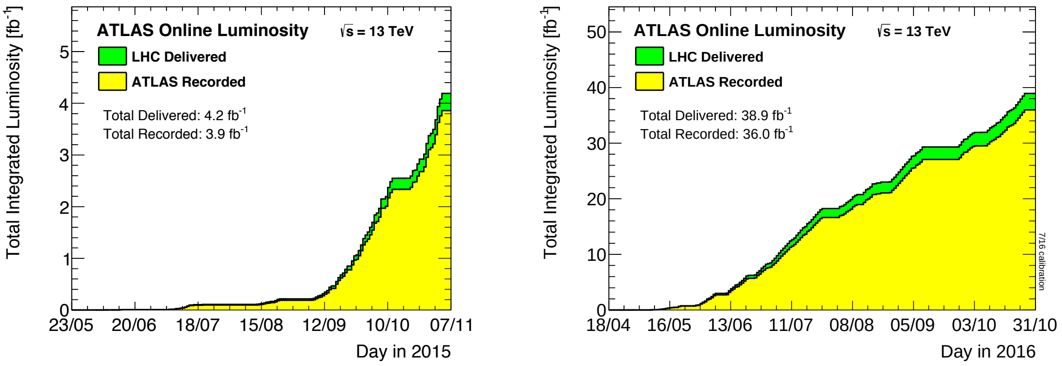
9.1 Dataset in pp collision with $\sqrt{s} = 13\,\mathrm{TeV}$

In 2015 and 2016, LHC was running with a center-of-mass energy of 13 TeV. Prospects for the sensitivity in Run2 can be estimated using these runs. The integrated luminosity of the data in 2016 is ten times larger than that in 2015 as shown in Figure 9.1.

The signal efficiency is expected to be improved by the additional dedicated trigger chains for the search of $\tau \to 3\mu$. These chains have been implemented since 2015, however, configurations of triggers are different run-by-run in 2015, while it was relatively stable in 2016. Because of the integrated luminosity and trigger stability, only the data in 2016 are used for the sensitivity estimation of the whole data in Run2.

9.1.1 Status in 2016 run

Figure 9.1 shows the recorded integrated luminosity by the ATLAS detector in 2016, which reached $36.0\,\mathrm{fb}^{-1}$. After the data quality requirement in a similar way to 2012 run, the luminosity for the analysis is obtained to be $33.2\,\mathrm{fb}^{-1}$. This is larger than 2012 run, therefore the sensitivity should be improved. However,



(a) (b)

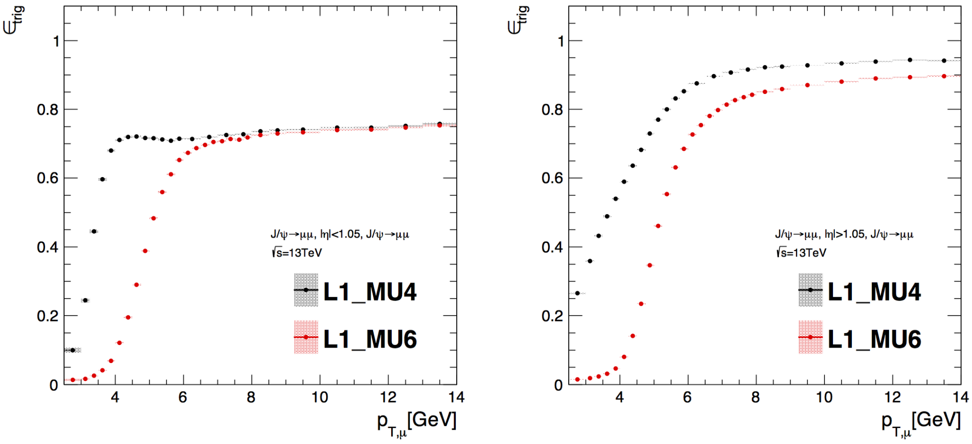
Figure 9.1: Evolution of the integrated luminosity in (a) 2015 and (b) 2016 as a function of date [59].

the instantaneous luminosity has also increased. The peak value in the 2016 was 1.37×10^{34} cm^{-2}s^{-1}, while that in the 2012 run was 0.77×10^{34} cm^{-2}s^{-1} as shown in Figure 2.3 and 2.2, respectively. Therefore, it was difficult to use the same trigger menu as that of 2012 run due to the higher rates. Thresholds for muon trigger chains are tuned in order to suppress the rates as follows:

- `EF_mu18_mu8_EFFS` → `HLT_mu22_mu8noL1`
- `EF_mu18_2mu4_EFFS` → `HLT_mu22_2mu4noL1`
- `EF_3mu4T` → `HLT_mu6_2mu4`

The signal efficiency of the combination of general trigger chains decreased from 2012 by these tighter thresholds. Furthermore, the combination chain with $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement is discarded because the threshold for $E_{\mathrm{T}}^{\mathrm{miss}}$ at the trigger level is also tightened to 80 GeV (while it was 40 GeV). With these reasons, it was essential to develop dedicated trigger chains in order to avoid efficiency loss. The dedicated chains are described in Section 9.1.2.

In this run, other improvements are involved for precision of the tracking and reconstruction efficiency. The installation of IBL detector at the innermost layer of ID is the effective upgrade for the target signature. The precision of the tracking and vertexing are improved by the additional layer. The significance of L_{xy} and a^0_{xy} are determined with higher precision for the SV of $\tau \to 3\mu$ signature. This is more effective for HF-originated signature, for which requirement of displaced SV is effective because of the long life-time of the B and D mesons. In addition, an alternative reconstruction algorithm for muons has been developed to keep the efficiency flat for highly collimated muons. For the analysis with 2012 data, reconstruction efficiency decreased for highly collimated signature as shown in Section 7.3. Therefore, this improvement is also effective for the analysis.



(a) (b)

Figure 9.2: p_T dependence of `L1_MU4` and `L1_MU6` for (a) the barrel and (b) endcap regions. Although `L1_MU6` shows lower efficiency below 6 GeV, approximately 40% of events with 4-6 GeVcan be taken. Therefore, it can be a solution for rate suppression to require `L1_MU6` while 4 GeV threshold is applied at the HLT.

9.1.2 Implementation of dedicated triggers

In order to improve the trigger efficiency, both inefficiencies by close-by muons and p_T thresholds should be considered. As described in Section 7.2, the close-by muon inefficiency is mainly caused by L1 due to the size of RoI. Therefore, the trigger efficiency loss can be suppressed by reducing the number of muon RoI required at the L1. For example, for `HLT_mu11_2mu4` requiring three muons at the HLT, `L1_MU11_MU4` is employed at the upstream of the chain. The additional muon at HLT is then required nearby the L1 triggered muon. This algorithm is called narrow scan (`nscan`), focusing on the collimated muons at the most downstream of the HLT. However, considering the maximal luminosity assumption of 1.5×10^{34} cm^{-2}s^{-1}, the expected rate of `L1_MU11_MU4` is still too high. Therefore, instead of `L1_MU4`, `L1_MU6` is employed for the second leg of the L1 requirement. Although the threshold for coincidence of `L1_MU6` is tighter than `L1_MU4`, `L1_MU6` can acquire muons with below 6 GeV as shown in Figure 9.2.

In addition, mass window for the three (or two) muons is set around the τ lepton mass at the trigger level. This requirement is applied as the range from 1000 MeV to 2700 MeV at the HLT because of its higher precision of the parameter measurement than that of the L1. The mass range is set wider than the mass resolution for the τ lepton mass in order to record the sideband region inclusively. This can suppress the signature from $J/\psi \to \mu\mu$ and other heavier hadrons decaying to multi-muons. `bTau` represents this specific mass window.

Thus, new trigger chains are developed to have lower threshold for each muon, while inefficiency for the close-by muon is minimized. The implemented trigger chains in this run are shown below:

- `HLT_mu11_2mu4noL1_nscan_L1_MU11_MU6_bTau`
- `HLT_mu11_mu6noL1_nscan_L1_MU11_MU6_bTau`
- `HLT_mu20_mu6noL1_nscan_L1_MU20_bTau`

These are used inclusively with the following nominal multi-muon triggers:

- `HLT_mu6_2mu4`
- `HLT_2mu14`
- `HLT_mu22_2mu4noL1`

This trigger setup is used for both searches of W and HF-originated signatures.

However, `HLT_mu11_mu6noL1_nscan_L1_MU11_MU6_bTau` and `HLT_mu20_mu6noL1_nscan_L1_MU20_bTau` were implemented in the middle of the run period. Therefore, this combination can be used after the implementation, and the integrated luminosity for the configuration is determined to be 18.9 fb^{-1}.

9.2 Monte-carlo simulation

Monte-Carlo simulation samples for both analyses of $W \to \tau\nu \to 3\mu\nu$ and HF $\to \tau \to 3\mu$ searches are generated. $W \to \tau\nu \to 3\mu\nu$ samples are generated based on the same configuration as in the 2012 analysis with the modification of the center-of-mass energy, PDF set, and additional decay model by EvtGen package [89, 90]. EvtGen achieves to simulate more correct kinematics of resonance decays (especially for B mesons) by using decay amplitude instead of the probability. Therefore, this is employed for the decay process instead of Pythia8, which is still used for the parton level simulation. HF $\to \tau \to 3\mu$ analysis needs the samples for normalization channel in addition to signal samples. These samples are produced in the same configuration as the $W \to \tau\nu \to 3\mu\nu$ signature.

9.2.1 Samples for $W \to \tau\nu \to 3\mu\nu$ search

The signal samples are generated by Pythia8 and EvtGen with NNPDF [91] of version 2.3. NNPDF is a PDF set calculated by neural network technique based on measurements in several experiments, including measurements of QCD jets in the LHC experiments. The samples for W^+ and W^- are generated with the production cross-section calculated independently. This separation is introduced to consider the charge difference of the W boson production, and to avoid making two $\tau \to 3\mu$ signatures in an event when two opposite-sign τ leptons are generated. Another τ lepton can be emitted from decays of mesons in the $W \to \tau\nu$ decay. Although its contamination is tiny, both τ leptons are forced to decay into three muons in this case. Such events are hard to be produced, considering the small $\mathcal{BR}$ of $\tau \to 3\mu$ events. Therefore, it is suppressed by the requirement for the charge of

Table 9.1: Configuration and cross-section of the $W \to \tau\nu \to 3\mu\nu$ signature with each charge of W.

Description	σ_{prod}[nb]	Eff. $\times$ $\mathcal{BR}(W \to \tau\nu)$	σ_{tot}[nb]
$W \to \tau\nu \to 3\mu\nu$, $p_{\text{T},\tau}$> 10 GeV			
W$^-$, $p_{\text{T},\mu1,2,3} > 4, 2, 1$ GeV	2.1300×10^2	4.03×10^{-2}	8.58
W$^+$, $p_{\text{T},\mu1,2,3} > 4, 2, 1$ GeV	1.6155×10^2	8.33×10^{-2}	13.5

τ because the production rate of same-sign τ leptons should be smaller than the opposite-sign ones. These samples are filtered by $p_{\text{T},\mu1} > 4$ GeV, $p_{\text{T},\mu2} > 2$ GeV, and $p_{\text{T},\mu3} > 1$ GeV, considering the thresholds at trigger level. $|\eta|$ for each muon is limited to be less than 3.0 by the detector acceptance. $p_{\text{T},\tau}$ is also restricted to be higher than 10 GeV. These threshold tunings are set mainly for HF-originated signature with many low p_{T} muons, and the effect for W-originated signature is not significant. The cross-section of new $W \to \tau\nu \to 3\mu\nu$ samples are summarized in Table 9.1.

9.2.2 Samples for HF $\to \tau \to 3\mu$ search

These are additional samples to the previous analysis procedure in 2012. The combination of Pythia8B and EvtGen is employed as the event generator with the same PDF set as $W \to \tau\nu \to 3\mu\nu$ samples. As described in Section 3.1.2, the main process of the search is $D_s \to \tau\nu \to 3\mu\nu$ because of large branching ratio of $D_s \to \tau\nu$ decay ($\mathcal{BR}(D_s \to \tau\nu \to 3\mu\nu) = 5.55 \pm 0.24\%$). $D_s \to \phi\pi \to \mu\mu\pi$ events are also generated in order to use as the normalization process to measure the cross-section of D_s meson production. On the other hand, it is difficult to discriminate the $D_s \to \tau\nu \to 3\mu\nu$ signature from other HF-processes (for example, $B \to \tau + X \to 3\mu + X$). Therefore, inclusive samples are also generated, which are required to have $b\bar{b}$ or $c\bar{c}$ at the parton level process by Pythia8B. The signal feature of the inclusive sample should be similar to $D_s \to \tau\nu \to 3\mu\nu$. Hence, the target of this search can be extended to inclusive HF-orginated signature by factorization using the fraction of D_s, after validation of the similarity between these simulated events.

These events include a large number of low p_{T} muons which are less efficient to be reconstructed and triggered. In order to keep actual number of events in the analysis procedure high, the filter for muons and a pion was introduced considering thresholds at the reconstruction and trigger levels. As shown in Section 9.1, the lowest unprescaled threshold combination was $p_{\text{T},\mu1} > 10$ GeV and $p_{\text{T},\mu2} > 6$ GeV for di-muon, and $p_{\text{T},\mu1} > 6$ GeV, $p_{\text{T},\mu2} > 4$ GeV, and $p_{\text{T},\mu3} > 4$ GeV for three muons. In order to keep a margin for the generator level thresholds, the requirement for HF $\to \tau \to 3\mu$ is set to be: $p_{\text{T},\mu1} > 4$ GeV; $p_{\text{T},\mu2} > 2$ GeV; and $p_{\text{T},\mu3} > 1$ GeV, while the requirement for $D_s \to \phi\pi \to \mu\mu\pi$ is set to be $p_{\text{T},\mu1} > 8$ GeV; $p_{\text{T},\mu2} > 4$ GeV; and $p_{\text{T},\pi} > 1$ GeV.

Table 9.2 is the summary of the signal and normalization samples for HF →

Table 9.2: Configuration and cross-section for the search of HF $\to \tau \to 3\mu$. The signal samples from decay of D_s mesons are used as the main target, however inclusive HF-originated signal samples are also generated. The normalization samples ($D_s \to \phi\pi \to \mu\mu\pi$) are prepared focusing on the D_s-based analysis. $D^\pm \to \phi\pi \to \mu\mu\pi$ samples are used for the measurement of the $D_s \to \phi\pi \to \mu\mu\pi$ due to its close invariant mass to τ. The production cross-sections (σ_{prod}) of these samples show values before the hadronization, therefore, these are the same values. The efficiency of the selection at the production (Eff.) and the fixed branching ratio ($\mathcal{BR}$) for each sample is also shown. The total cross-section of each sample (σ_{tot}) is estimated by the product of these factors.

Description	σ_{prod}[nb]	Eff.	$\mathcal{BR}$	σ_{tot}[nb]
$D_s \to \tau\nu \to 3\mu\nu$, $p_{\text{T},\text{D}_\text{s}} > 10\,\text{GeV}$				
$D_s^-, p_{\text{T},\mu1,2,3} > 4, 2, 1\,\text{GeV}$	4.14×10^5	8.02×10^{-3}	5.5×10^{-2}	1.83×10^2
$D_s^+, p_{\text{T},\mu1,2,3} > 4, 2, 1\,\text{GeV}$	4.14×10^5	8.10×10^{-3}	5.5×10^{-2}	1.84×10^5
HF $\to \tau \to 3\mu$, $p_{\text{T},\tau} > 10\,\text{GeV}$				
$\tau^-, p_{\text{T},\mu1,2,3} > 4, 2, 1\,\text{GeV}$	4.14×10^5	8.08×10^{-4}	1.0(fixed)	3.34×10^2
$\tau^+, p_{\text{T},\mu1,2,3} > 4, 2, 1\,\text{GeV}$	4.14×10^5	8.03×10^{-4}	1.0(fixed)	3.32×10^2
$D_s \to \phi\pi \to \mu\mu\pi$, $p_{\text{T},\text{D}_\text{s}} > 10\,\text{GeV}$				
$D_s^-, p_{\text{T},\mu1,2,\pi} > 8, 4, 1\,\text{GeV}$	4.14×10^5	7.71×10^{-4}	1.28×10^{-5}	4.09×10^{-3}
$D_s^+, p_{\text{T},\mu1,2,\pi} > 8, 4, 1\,\text{GeV}$	4.14×10^5	7.85×10^{-4}	1.28×10^{-5}	4.16×10^{-3}
$D^\pm \to \phi\pi \to \mu\mu\pi$, $p_{\text{T},\text{D}} > 10\,\text{GeV}$				
$D^-, p_{\text{T},\mu1,2,\pi} > 8, 4, 1\,\text{GeV}$	4.14×10^5	1.97×10^{-3}	2.03×10^{-6}	1.66×10^{-3}
$D^+, p_{\text{T},\mu1,2,\pi} > 8, 4, 1\,\text{GeV}$	4.14×10^5	2.00×10^{-3}	2.03×10^{-6}	1.68×10^{-3}

$\tau \to 3\mu$ search.

Chapter 10

Improvement of the W-originated Analysis

The analysis procedure of W-originated $\tau \rightarrow 3\mu$ search is basically the same as that for 2012 data described in Part II, however sensitivity should improve by the increasing efficiency and number of the τ leptons. This benefit is mainly from implementation of dedicated triggers. On the other hand, optional reconstruction algorithm is also developed to avoid the close-by muon inefficiency at the reconstruction. These improvements can be estimated by the scale of acceptance times efficiency ($A \times \epsilon$) in $\mathcal{BR}$ calculation with respect to Run1. The factor of efficiency improvement is determined by the data and simulation in 2016. Not only the integrated luminosity but the production cross-section also increases in Run2 with higher proton beam energy. In this chapter, the impact of efficiency improvement and the sensitivity prospect in Run2 are described.

10.1 Improvement of the signal efficiency

By the measurement in 2012, it was clarified that the trigger efficiency was low ($42.3 \pm 4.6\%$) and this indicates possible improvements in Run2 if the efficiency can be increased. The dedicated triggers for $\tau \rightarrow 3\mu$ search were developed as described in Section 9.1 for this purpose and employed in ATLAS trigger system since 2016. The impact of this improvement is estimated as a factor, called improvement factor, for the $A \times \epsilon$ in the analysis procedure in 2012. Evaluation of the reconstruction efficiency is also described and involved in the improvement factor.

10.1.1 Trigger efficiency

Trigger efficiency is evaluated with respect to three muon reconstructed events with the same quality selection as in 2012. The selection criteria for each muon is a bit different from that in 2012 analysis considering improved detector performance and reconstruction algorithms. However, this difference is negligible for the trigger study.

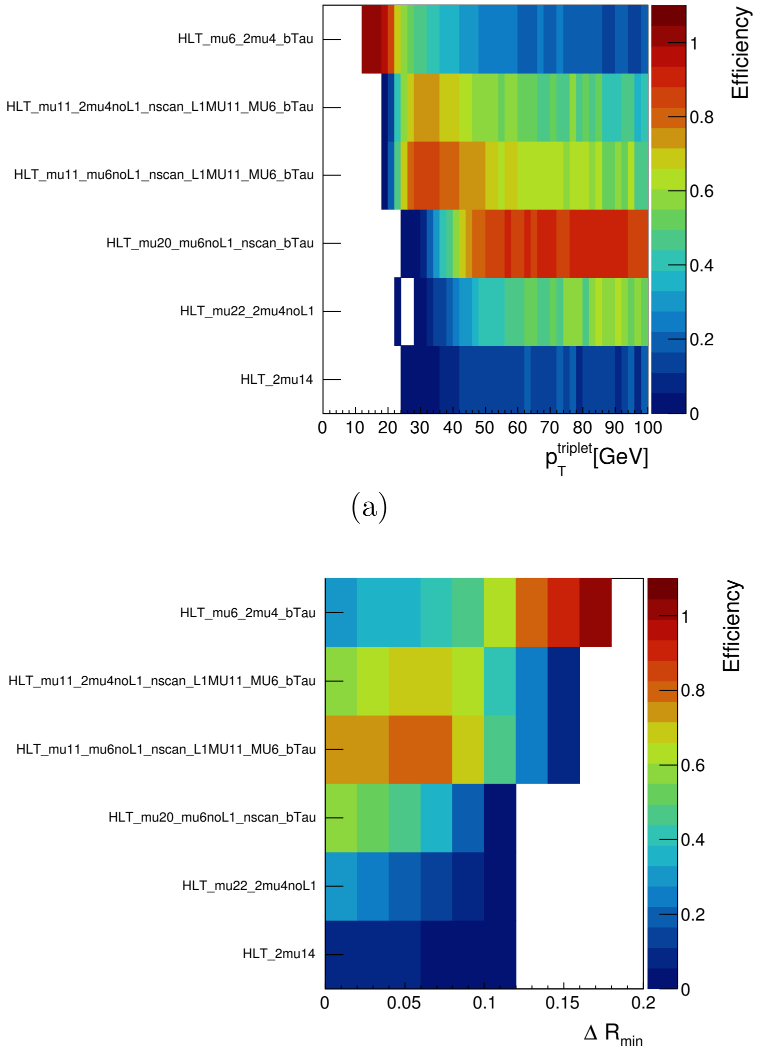
The efficiencies of the nominal muon trigger chains and dedicated trigger chains are shown in Figure 10.1 as a function of p_T and ΔR_min of the triplets, respectively. `HLT_mu6_2mu4_bTau` is efficient for triplet with p_T below 20 GeV, while it is inefficient for highly boosted signature due to requirement of three muons at L1. This effect is also found in ΔR_min dependence. `HLT_mu11_2mu4noL1_nscan_L1MU11_MU6_bTau` and `HLT_mu11_mu6noL1_nscan_L1MU11_MU6_bTau` are less affected by the close-by muon inefficiency because only two muons are required at L1. `HLT_mu20_mu6noL1_nscan_bTau` requires only one muon at L1, therefore this chain is most efficient for highly boosted signature. However, these triggers with smaller number of muon requirement at L1 is applied with high p_T threshold. `HLT_2mu14` is a simple di-muon trigger. The efficiency is lower than other chains, however the threshold for the leading muon p_T is 10 GeV, which is lower than `L1_MU11_MU6`-based triggers. Therefore, this chain has unique efficient region for leading muon p_T around 10 GeV, although has large overlap region. Inefficiency of these chains can be compensated each other by using them inclusively.

Figure 10.2 shows the expected efficiencies of each chain and their combination. In this figure, efficiency of each chain is calculated independently, while the overlap of these chains is considered for the inclusive efficiency calculation. The inclusive efficiency is estimated to be 64.3%, which is higher than that in 2012. Hence, the improvement factor by additional trigger chains ($\mathcal{F}_\mathrm{trig}$) is found to be $64.3/42.3 = 1.5$.

10.1.2 Reconstruction efficiency

The close-by muon inefficiency at the reconstruction is caused by shared-hits in the muon spectrometer. If two muons share hits, one muon can be reconstructed as a muon while the other is identified as an ID track due to missing hits at muon spectrometer by overlap removal. Approximately 10% of the signal events consist of two muons and one ID track in the reconstruction algorithm used in 2012. Therefore, the requirement for the overlap removal should be loosened to improve the efficiency. In the nominal overlap removal requirement, all hits used for a muon are removed from the candidates for the other muons. This selection is modified to allow for having shared-hits up to 10% of all hits which are used for the reconstruction.

Figure 10.3 shows the performance difference between 2012 and 2016 for $W \to \tau\nu \to 3\mu\nu$ signature. Events with only two reconstructed muons in 2012, which are not used in analysis, are recovered in 2016 with another reconstructed muon. Number of events with four or more muons did not increase, hence it is validated that this improvement of the efficiency does not raise rates of fake muons by misidentification of an ID track. The improvement factor for reconstruction ($\mathcal{F}_\mathrm{reco}$) is obtained to be 1.08 from the ratio of the number of three muon events with the new overlap removal to that with the old one.



(a)

(b)

Figure 10.1: (a) p_{T} and (b) ΔR_{min} dependence of the efficiency for each trigger chain. It is validated that the combination makes efficiency higher than only using one chain. It is found in these figures that inefficient region of each chain is different each other, and can be recovered by the combination.

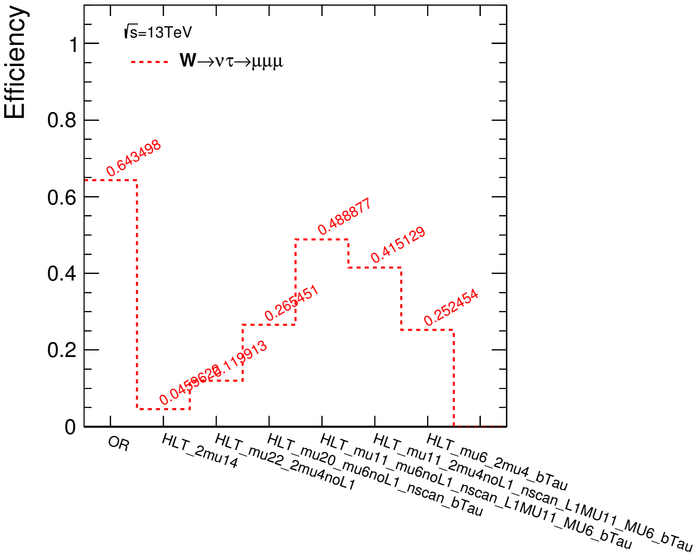
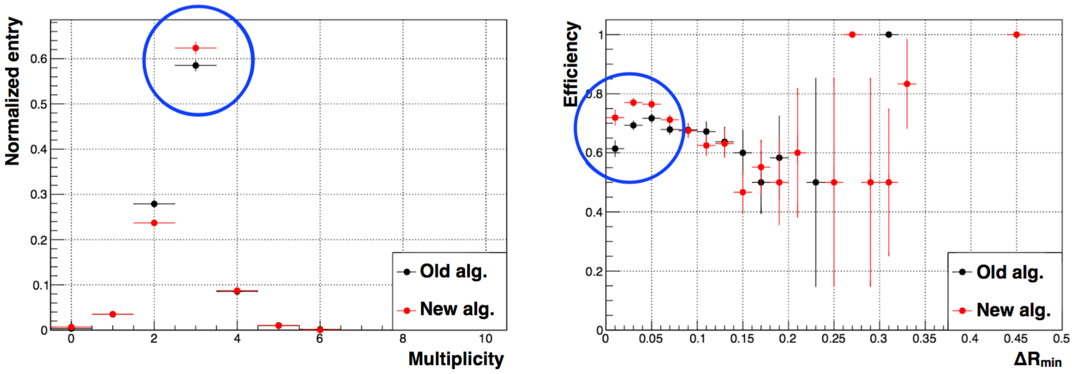


Figure 10.2: Efficiencies of each chain and their combination (denoted as ”OR”).



(a) (b)

Figure 10.3: (a) Muon multiplicity distribution of the $W \to \tau\nu \to 3\mu\nu$ signal events and (b) p_{T} dependence of the reconstruction efficiency. Black and red markers represent old and new algorithms used for the reconstruction, respectively. By recovering the close-by muon inefficiency, number of events with three muons has increased.

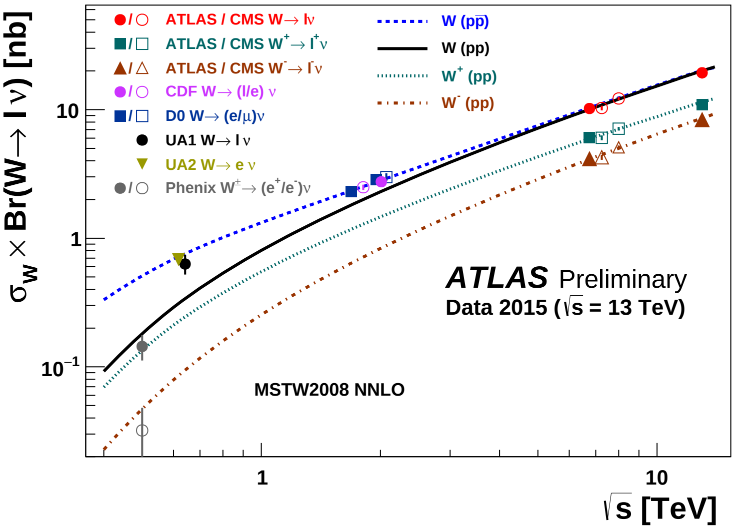


Figure 10.4: Cross-section times $\mathcal{BR}(W \to l\nu)$ as a function of center-of-mass energy [92]. Green and brown dashed lines represent the theoretical predictions for W^+ and W^-, respectively. The total cross-section which is measured by pp ($p\bar{p}$) collider experiment is shown as a black solid (blue dashed) line. Markers represent the results by the experiments with the same color convention as the prediction lines.

10.2 Expected sensitivity in Run2

Considering the improvement of the efficiency, the signal yield of the data in 2016 can be estimated. In addition, the production cross-section of $W \to \tau\nu$ increases. The expected limit for $\mathcal{BR}(\tau \to 3\mu)$ with $100\,\mathrm{fb}^{-1}$ of whole Run2 is estimated by using these improvement factors and luminosity evolution.

10.2.1 Extrapolation of number of $W \to \tau\nu$

Production cross-section of $W^\pm$ at the center-of-mass energy of 13 TeV is measured in the ATLAS experiment [92], using $W \to l\nu$ mode. This result shows good agreement with the expected value by theoretical extrapolation from the measured values in 7 TeV and 8 TeV as shown in Figure 10.4. According to this result, production cross-sections of $W^- \to l^-\bar{\nu}$, $W^+ \to l^+\nu$ and the sum of these are determined as

$$\sigma_{W^-\to l^-\nu} = 8.38 \pm 0.02(stat.) \pm 0.35(syst.) \pm 0.75(lumi.)\mathrm{nb}, \tag{10.1}$$

$$\sigma_{W^+\to l^+\nu} = 10.96 \pm 0.02(stat.) \pm 0.35(syst.) \pm 0.99(lumi.)\mathrm{nb}, \tag{10.2}$$

$$\sigma_{W^\pm\to l^\pm\nu} = 19.35 \pm 0.02(stat.) \pm 0.76(syst.) \pm 1.74(lumi.)\mathrm{nb}. \tag{10.3}$$

Comparing this to the production cross-section in 8 TeV collision, which is shown in Section 6.2, the improvement factor for number of $W \to \tau\nu$ events ($\mathcal{F}_{W\to\tau\nu}$) is determined as

$$\mathcal{F}_{W\to\tau\nu} = \frac{\sigma^{13\,\text{TeV}}_{W\to\tau\nu}}{\sigma^{8\,\text{TeV}}_{W\to\tau\nu}} = \frac{19.35}{11.89} \tag{10.4}$$

$$= 1.63. \tag{10.5}$$

10.2.2 Sensitivity prospect for $W \to \tau\nu \to 3\mu\nu$ search

As explained in previous sections, $A \times \epsilon$ and number of $W \to \tau\nu$ ($N_{W\to\tau\nu}$) have increased in Run2 analysis by the improvement factors. By multiplying $\mathcal{F}_{\text{trig}}$ and $\mathcal{F}_{\text{reco}}$ to $A\times\epsilon$ in 2012 analysis (Section 7.2, **??**), $A\times\epsilon$ in Run2 analysis is obtained as:

$$A \times \epsilon = 2.32 \times \mathcal{F}_{\text{trig}} \times \mathcal{F}_{\text{reco}} \tag{10.6}$$

$$= 3.74\%. \tag{10.7}$$

$N_{W\to\tau\nu}$, which was 2.41 $\times 10^8$ in Run1 (Section 6.2), is estimated to increase in Run2 by $\mathcal{F}_{W\to\tau\nu}$ and evolution of the integrated luminosity from 19.2 fb^{-1} to 100 fb^{-1} as

$$N_{W\to\tau\nu} = (2.41 \times 10^8) \times \mathcal{F}_{W\to\tau\nu} \times \frac{\mathcal{L}_{\text{Run2}}}{\mathcal{L}_{2012}} \tag{10.8}$$

$$= 1.96 \times 10^9. \tag{10.9}$$

On the other hand, number of background events (N_{bkg}), which was estimated to be 0.193 in Run1 in Section 6.1.2, should also increase with the larger integrated luminosity in Run2. Therefore, assuming the accidental background is the dominant source, N_{bkg} is estimated by the ratio of the integrated luminosity of Run2 to that of 2012:

$$N_{\text{bkg}} = 0.193 \times \frac{\mathcal{L}_{\text{Run2}}}{\mathcal{L}_{2012}} = 0.965. \tag{10.10}$$

In fact, it is not certain that the number of background events increase linearly. Therefore, although the relative uncertainty is expected to be reduced by the increasing sideband events, the uncertainty of the estimation is set to be the same as that in 2012 conservatively. These numbers are summarized in Table 10.1.

The sensitivity prospect of the search without any changes in the analysis procedure from that in 2012 can be estimated by these improved parameters. In this estimation, systematic uncertainties are fixed as the same values as those in 2012. For the discovery with the significance of 3σ (5σ), 7 (11) events in the SR is necessary to remain, which mean $\mathcal{BR}(\tau \to 3\mu)\sim 8.2(14.0) \times 10^{-8}$. But if the number of observed events are fewer for the significant observation of the signal, the limits with 90% CL are expected[1] to follow the pseudo number of observed

---

[1] Usually for the rare decay searches, a single event sensitivity ($1/(A \times \epsilon \times N_{W\to\tau\nu})$) is used as a benchmark. In this case, this value is estimated to be 1.4×10^{-8}.

Table 10.1: $A \times \epsilon$, $N_{W\to\tau\nu}$, and $N_{\rm bkg}$ in the SR with their systematic uncertainties. The right row shows the improvement factor for each element including the integrated luminosity evolution.

	Expected in Run2	Improvement factor $\mathcal{F}$
$A \times \epsilon$	$3.74\% \pm 0.67$ ($syst. = 18\%$)	1.5 (trig.) × 1.08 (reco.)
$N_{W\to\tau\nu}$	$(1.96 \pm 0.076) \times 10^9$ ($syst. = 3.9\%$)	1.63 ($\sigma_{W\to\tau\nu}$) × 5.0 (lumi.)
$N_{bkg}(SR)$	0.965 ± 0.652 ($syst. = 68\%$)	5.0 (lumi.)

events ($N_{\rm obs}$) of 0, 1, and 2.

$$\mathcal{BR}(\tau \to 3\mu)_{N_{\rm obs}=0} < 2.4 \times 10^{-8} \tag{10.11}$$

$$\mathcal{BR}(\tau \to 3\mu)_{N_{\rm obs}=1} < 4.1 \times 10^{-8} \tag{10.12}$$

$$\mathcal{BR}(\tau \to 3\mu)_{N_{\rm obs}=2} < 5.8 \times 10^{-8} \tag{10.13}$$

These limits are obtained by CLs method used in the analysis in 2012 (Section 8.3). If no signal events are observed in the signal region, the most probable value of $N_{\rm obs}$ should be 1.0 because $N_{\rm bkg}$ is estimated to be 0.965. Therefore, the sensitivity will be improved to 4.1×10^{-8} without any upgrades in the analysis procedure. This shows approximately ten times better than 2012 result, and more stringent than that by LHCb [47] (4.6×10^{-8}). However, it is still looser than the limit by Belle [45] (2.1×10^{-8}). On the other hand, this limit can be further improved by the reduction of the systematic uncertainties associated with the background estimation and efficiency measurement to around 10^{-8}, at the level predicted new physics beyond the SM.

Chapter 11

Prospect of the HF-originated Analysis

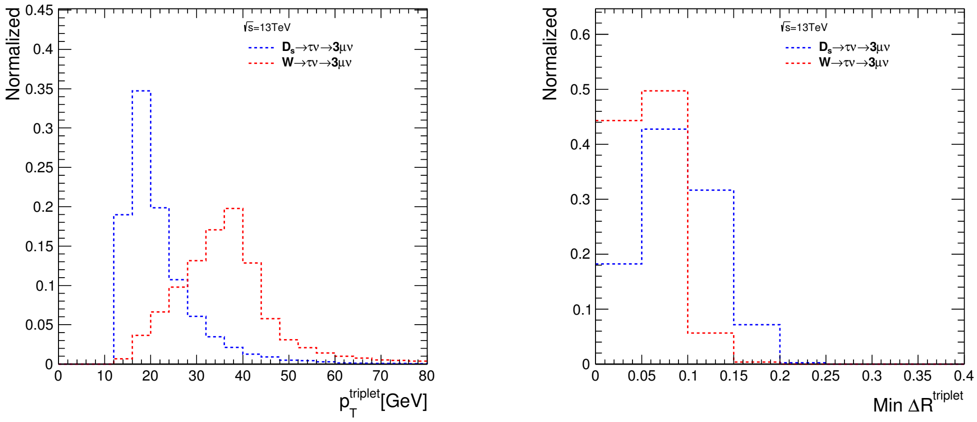
Search for HF-originated signature can be considered in Run2 thanks to the improvement of the efficiency and the large production cross-section of heavy flavored mesons. The analysis procedure is designed based on that of $W \to \tau\nu \to 3\mu\nu$ search. However, HF events do not have a large $E_{\mathrm{T}}^{\mathrm{miss}}$ which is a characteristic feature of $W \to \tau\nu \to 3\mu\nu$ signature. Furthermore, τ leptons from decay of HF hadrons are less boosted and less isolated from hadron jets than those from W boson decay. Therefore, the analysis strategy should be modified as suitable for this search as described in Section 3.2.

In order to evaluate the sensitivity, number of HF-originated τ leptons needs to be estimated. Although lepton universality can be assumed in $W \to l\nu$ mode, it cannot be made for the HF-originated leptons due to the lower mass scale. Hence, focusing on $D_s \to \tau\nu \to 3\mu\nu$ signature, number of D_s mesons (N_{Ds}) is used for $\mathcal{BR}(\tau \to 3\mu)$ calculation. N_{Ds} is obtained by the measurement of the $D_s \to \phi\pi \to \mu\mu\pi$ signature.

In this chapter, the event selections for the $D_s \to \tau\nu \to 3\mu\nu$ search and the $D_s \to \phi\pi \to \mu\mu\pi$ measurement are described. Extension from $D_s \to \tau\nu \to 3\mu\nu$ search to inclusive HF $\to \tau \to 3\mu$ search is also discussed using the similarity of the HF-originated signatures. Finally, the possibility of the approach and expectation are shown.

11.1 Event selection for $D_s \to \tau\nu \to 3\mu\nu$

The trigger efficiency and analysis procedure for the HF-originated signature are different from $W \to \tau\nu \to 3\mu\nu$ signature because of lower p_{T} of τ leptons and similarity to background events. The trigger response for $D_s \to \tau\nu \to 3\mu\nu$ signature is first checked to ensure the difference from W-originated one. The analysis procedure is designed based on that of $W \to \tau\nu \to 3\mu\nu$ search. The muon reconstruction improvement enables to merge two types of muons (combined muons



(a) (b)

Figure 11.1: (a) p_{T} and (b) minimum ΔR distributions of the signal events. $W \to \tau\nu \to 3\mu\nu$ sample (red dashed line) is also shown for comparison with $D_s \to \tau\nu \to 3\mu\nu$ sample (blue dashed line).

and TPa muons in Run1) into an alternative combined muon inclusively. However, quality selections for tracks and muons are not different. Also, this signature does not have characteristic features in hadron jets and $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution. Therefore, the event selection criteria should be optimized by only the combined muons and their combination.

11.1.1 Trigger efficiency

The same trigger combination as $W \to \tau\nu \to 3\mu\nu$ in Run2 is used for this analysis. As previously described, p_{T} of HF-originated τ lepton is lower, and the minimum ΔR of three muons is larger than that of W-originated signal as shown in Figure 11.1. Therefore, the trigger efficiency is lower due to p_{T} thresholds at trigger level, although the close-by muon inefficiency also less affects the efficiency. It is expected that lower threshold triggers are effective even if these require three muons at L1.

Figure 11.2 shows the response of each trigger chain and their combination. `HLT_mu6_2mu4` is the most effective chain, and `HLT_mu11`-based triggers compensate the inefficiency due to third muon requirement. The other chains also help to acquire highly boosted signature. This is validated as a reasonable difference of trigger efficiency from $W \to \tau\nu \to 3\mu\nu$ case (Figure 10.2). The inclusive efficiency is estimated as 25% by the signal simulation. The systematic uncertainty of the trigger efficiency is assumed to be 11%, which is the same as that of $W \to \tau\nu \to 3\mu\nu$ search determined in the measurement in 2012.

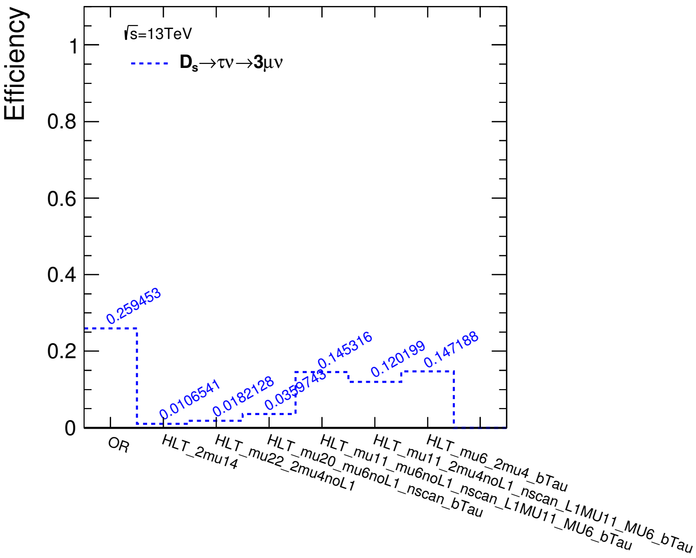


Figure 11.2: Efficiency of each trigger chain and the combination. The inclusive efficiency is lower than that of $W \to \tau\nu \to 3\mu\nu$ case, although it is above 25%.

11.1.2 Object selections

In this analysis, only muons are used for the event selection. Although the combination of these muons are used, selection criteria for each muon are described first. Requirements for the properties of hits at the ID and MS are set as follows based on those of $W \to \tau\nu \to 3\mu\nu$ search (meanings of the variables are described in Section 5.2):

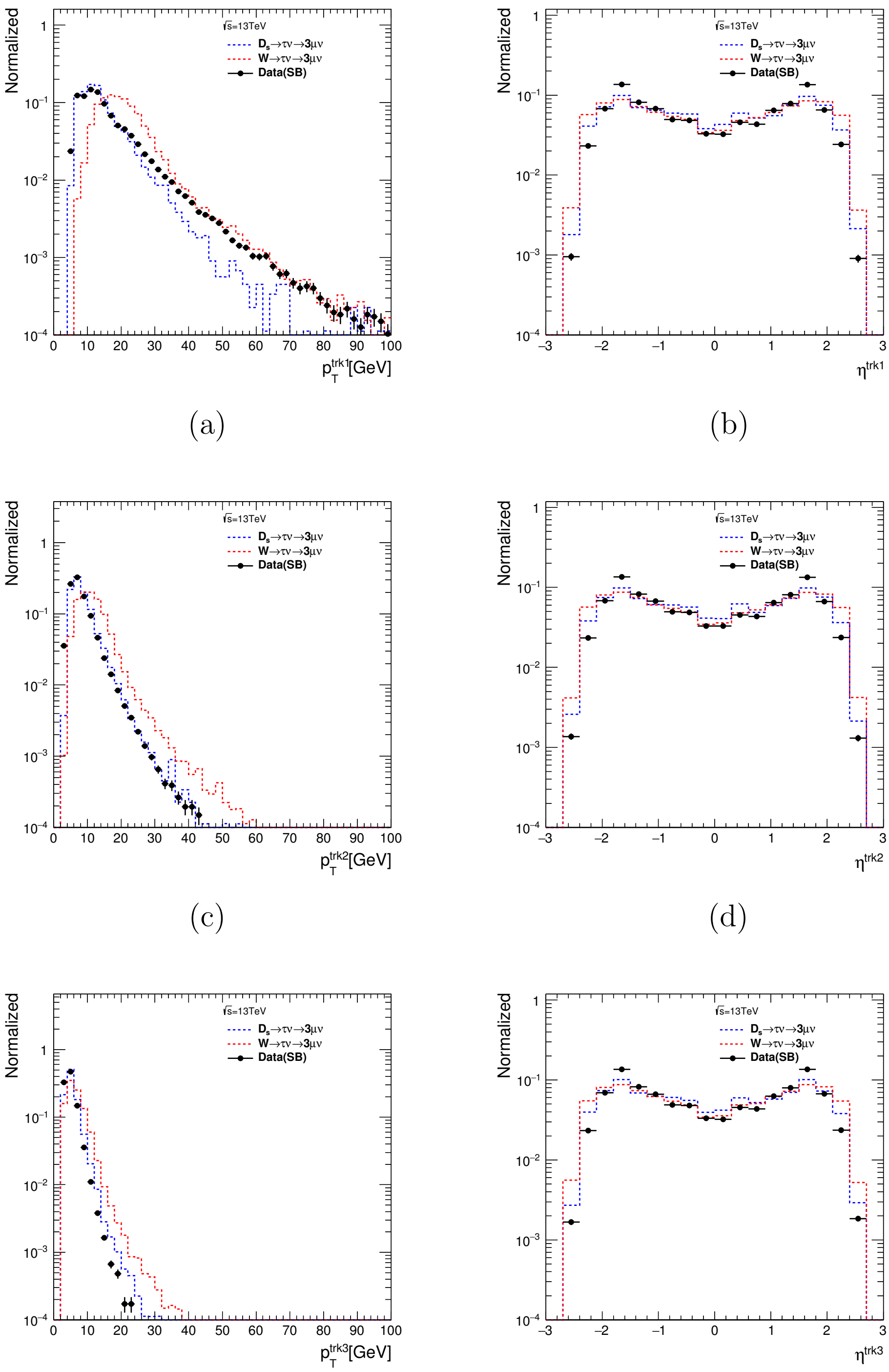
- If a hit at the innermost Pixel layer (B-layer or IBL) is expected, number of the hits should be larger than or equal to 1
- $N_{\text{hits,Pixel}} \geq 2$
- $N_{\text{hits,SCT}} \geq 5$
- $N_{\text{holes,Pixel+SCT}} \leq 2$
- $N_{\text{hits,TRT}} \geq 6$ and $N_{\text{outhits,TRT}}/N_{\text{hits,TRT}} < 90\%$ $(|\eta| < 1.9)$
  $N_{\text{hits,TRT}} \geq 6$ or $N_{\text{outhits,TRT}}/N_{\text{hits,TRT}} < 90\%$ $(|\eta| \geq 1.9)$
- $N_{\text{hits,MDT+CSC}} \geq 12$
- $N_{\text{hits,RPC+TGC}} \geq 2$
- Ratio of the TRT hits with high energy deposition to all the associated hits on TRT should be larger than or equal to 0.35

On the other hand, track parameters and qualities of the muons are also validated. These variables are defined based on those described in Section 5.2. Selections for p_{T} and η are set by the limits of the reconstruction and the acceptance of the detector. Figure 11.3 shows p_{T} and η distributions. In these figures, distributions of the sideband data, whose region definition is described in Figure 5.2, are also shown. These plots are shown after looser selections for p_{T} and η are applied at the data preparation level in order to consider the thresholds at generator level for comparison with data. The p_{T} thresholds are set as $p_{\mathrm{T},\mu1} > 4\,\mathrm{GeV}$, $p_{\mathrm{T},\mu2} > 2\,\mathrm{GeV}$, and $p_{\mathrm{T},\mu3} > 2\,\mathrm{GeV}$. The η threshold is set to be $|\eta| < 2.7$ for all muons at the loose selection. The muon momentum balance significance ($\sigma p_{\mathrm{bal.}}$) and the muon scattering angle significance ($\sigma\Theta_{\mathrm{scat.}}$) are introduced for evaluation of the "kink" of the muon track using the difference of tracks parameters at the ID and MS as shown in Figure 11.4. The requirements for these kink variables are set to reduce fake muons from decay of π and K mesons. In addition, a quality of each track is represented by the probability (p-value) of the track fitting. The track p-value distributions of these tracks are shown in Figure 11.5 and used as input to MVA procedure. Quality selections for muon tracks are set as follows:

- $-3 < \sigma p_{\mathrm{bal.}} < +4$
- $\sigma\Theta_{\mathrm{scat.}} < 4$
- $|\eta| < 2.5$
- p_{T} for leading, 2nd leading, and 3rd leading tracks should be larger than 5.5 GeV, 3.5 GeV, and 2.5 GeV, respectively

These requirements are the same as those of $W \to \tau\nu \to 3\mu\nu$ analysis, because no significant difference is found between $D_s \to \tau\nu \to 3\mu\nu$ and $W \to \tau\nu \to 3\mu\nu$ samples. Therefore, it can be assumed that the muon reconstruction efficiency is the same as that described in Sections 7.3 and 10.1.2.

On the other hand, selections for the SV reconstructed by these signal muons are important to discriminate the signal from the background events. Figure 11.6 shows the distributions of L_{xy} and its significance of the signal events, the sideband data, and W-originated signal events for a reference. It is expected that $D_s \to \tau\nu \to 3\mu\nu$ signature has longer life-time, which is a characteristic feature, than $W \to \tau\nu \to 3\mu\nu$ signature. The distributions of a^0_{xy} (and its significance) and provability of the SV are shown in Figure 11.7 and 11.8, respectively. The significance of a^0_{xy} is a discriminant in W-originated analysis, but it is less useful in HF-originated. This is because D_s mesons can be generated from decays of B mesons with other associated particles, while those from W decays are produced at interaction point directly. A threshold for probability of the SV is applied to be larger than 10^{-4} at the data preparation level in order to keep the size of the data small without losing the signal. These are similar to those of $W \to \tau\nu \to 3\mu\nu$ signature, and used for discrimination from background events. Because of the

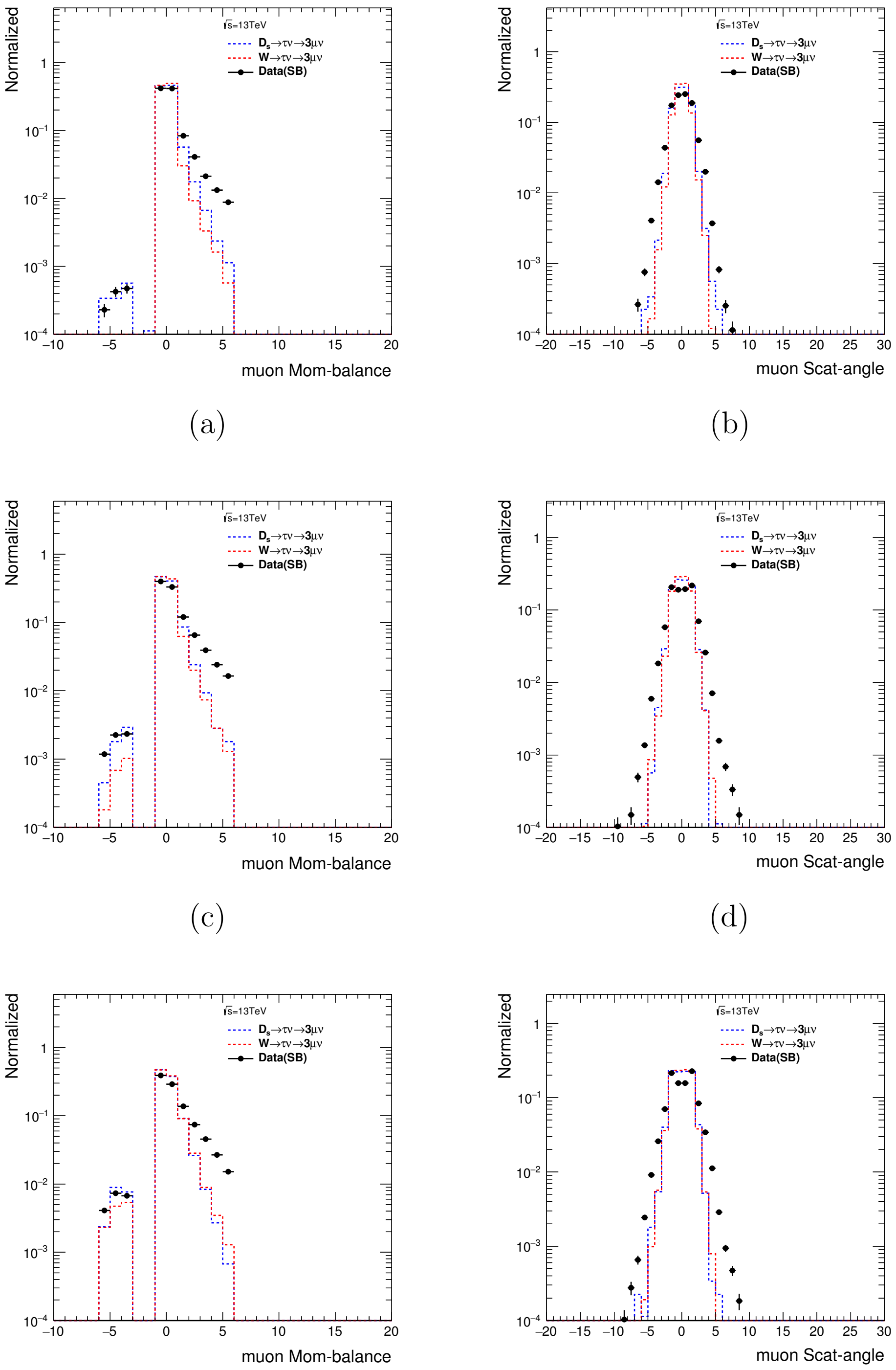


(a) (b)

(c) (d)

(e) (f)

Figure 11.3: p_T and η distributions for $D_s \to \tau\nu \to 3\mu\nu$ events (blue dashed line), $W \to \tau\nu \to 3\mu\nu$ events (red dashed line), and the sideband data (black solid line). (a), (c), and (e) show p_T distributions of the leading muon, 2nd leading muon, and 3rd leading muon, respectively. (b), (d), and (f) show η distributions of the leading muon, 2nd leading muon, and 3rd leading muon, respectively. Loose selections are applied for p_T and η in order to consider the thresholds at generator level.

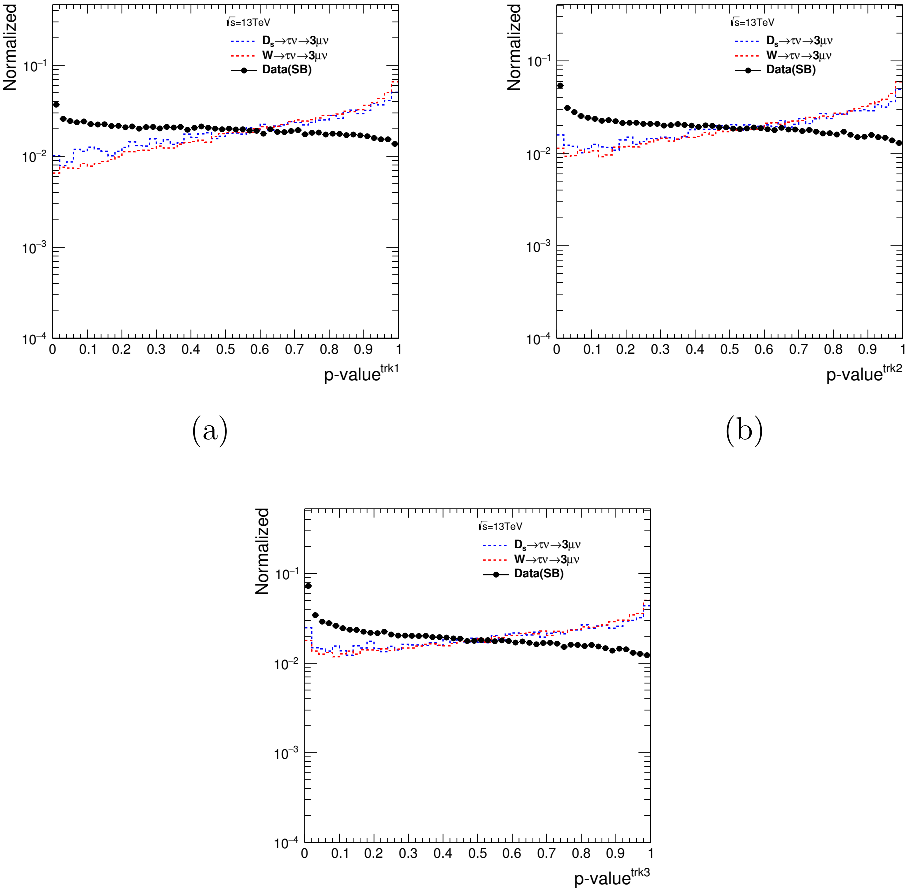


(a) (b)

(c) (d)

(e) (f)

Figure 11.4: Distributions of (a,c,e) muon momentum balance and (b,d,f) muon scattering angle. Blue dashed line, red dashed line, and black solid line represent $D_s \to \tau\nu \to 3\mu\nu$ events, $W \to \tau\nu \to 3\mu\nu$ events, and the sideband data, respectively. The first (a,b), second (c,d), and third (e,f) raws represent the leading muon, 2nd leading muon, and 3rd leading muon, respectively.

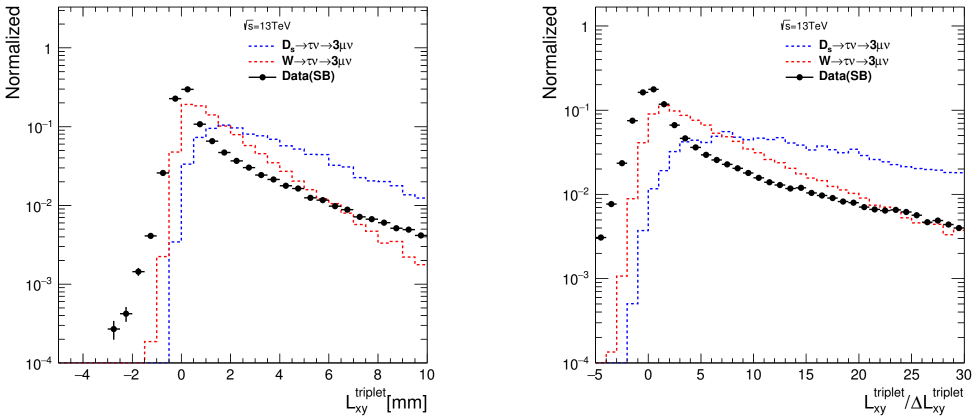


(a)

(b)

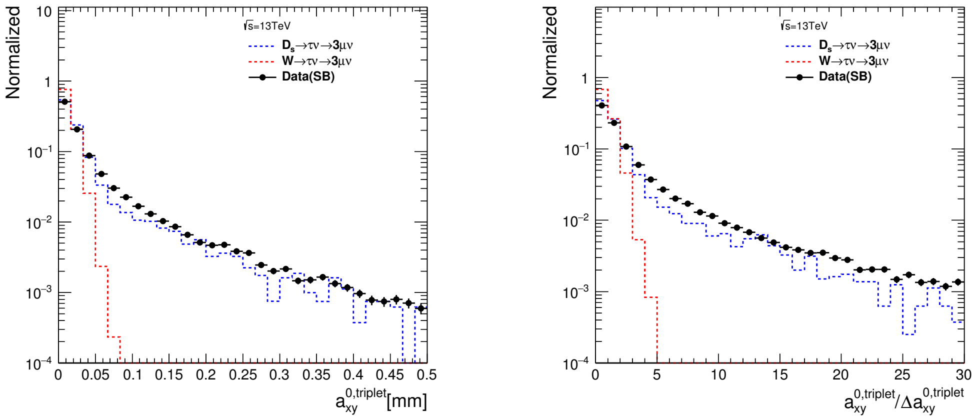
(c)

Figure 11.5: Distributions of the track p-value for (a) leading, (b) 2nd leading, and (c) 3rd leading muon. Blue dashed line, red dashed line, and black solid line represent $D_s \to \tau\nu \to 3\mu\nu$ events, $W \to \tau\nu \to 3\mu\nu$ events, and the sideband data, respectively.



(a) (b)

Figure 11.6: (a) L_{xy} and (b) its significance distributions of the signal events. $D_s \to \tau\nu \to 3\mu\nu$ signal, $W \to \tau\nu \to 3\mu\nu$ signal and the sideband data are shown as blue dashed line, red dashed line, and black solid line, respectively. $D_s \to \tau\nu \to 3\mu\nu$ events show larger value because of the additional flight length of D mesons.



(a) (b)

Figure 11.7: (a) a^0_{xy} and (b) its significance distributions of the signal events. The distribution of $D_s \to \tau\nu \to 3\mu\nu$ (blue dashed line) is not sharply like that of $W \to \tau\nu \to 3\mu\nu$ (red dashed line), and rather similar to that of the sideband data (black solid line).

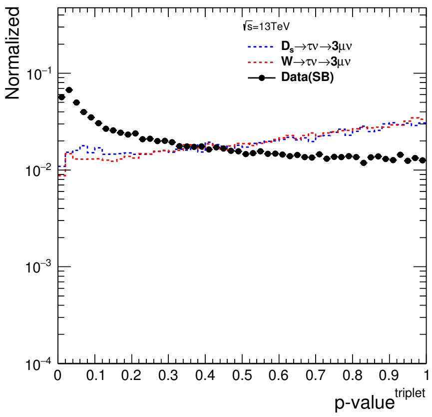


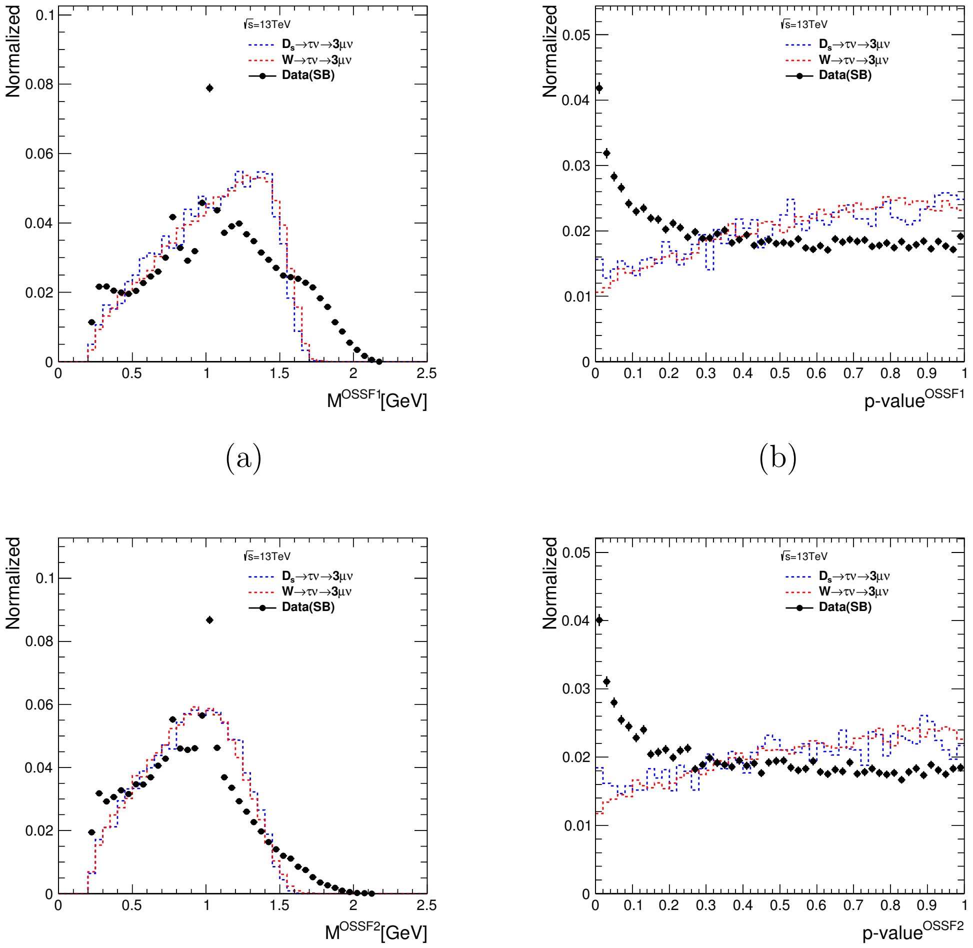
Figure 11.8: Distribution of probability of the SV for $D_s \rightarrow \tau\nu \rightarrow 3\mu\nu$ events (blue dashed line), $W \rightarrow \tau\nu \rightarrow 3\mu\nu$ events (red dashed line), and the sideband data (black solid line).

significant difference of L_{xy}, a^0_{xy}, and p_{T}triplet, the pre-selection requirements for MVA selections are set as follows:

- Probability of the SV should be larger than 10^{-4} (at the data preparation level)
- $0 < L_{\mathrm{xy}}$
- No limitation for a^0_{xy}
- No limitation for $p_{\mathrm{T,triplet}}$
- Absolute value of the total charge of the three muons is equal to 1

$E_{\mathrm{T}}^{\mathrm{miss}}$ associated variables, jet associated variables, and the isolation variable used in the $W \rightarrow \tau\nu \rightarrow 3\mu\nu$ analysis, are discarded from the selection for the $D_s \rightarrow \tau\nu \rightarrow 3\mu\nu$ search.

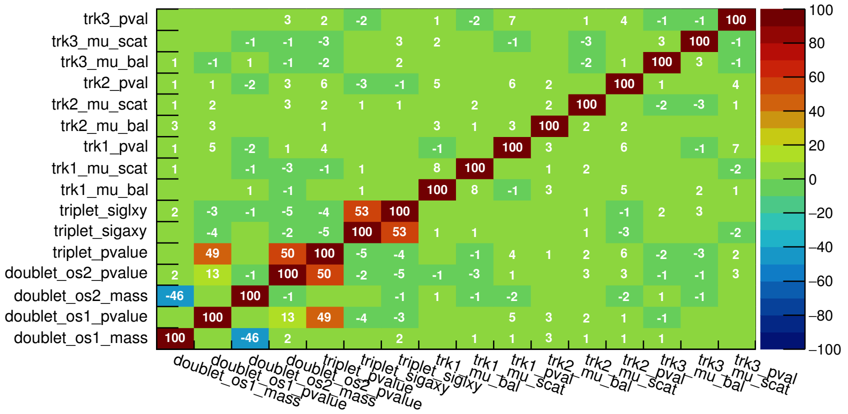
In addition, invariant masses of opposite-sign muon pairs in a triplet (OS1 and OS2) are defined in order to reject di-muons from decay of light mesons. This rejection is more important for this analysis because HF-originated background events cannot be reduced by the requirements for $E_{\mathrm{T}}^{\mathrm{miss}}$, and hadron jets. However, the vetoes on their masses also reduce large number of signal events. Therefore, the masses of OS1(OS2) and probability of the $\mathrm{SV}_{OS1(OS2)}$ are used for the MVA selection, while any pre-selections are not applied for these variables. The definition of the order of OS1 and OS2 are defined in Section 5.2. The distributions of $M_{\mathrm{OS1(OS2)}}$ and $\mathrm{SV}_{OS1(OS2)}$ are shown in Figure 11.9. The peaks of the resonances are found in the mass distributions.



(a) (b)

(c) (d)

Figure 11.9: Distributions of invariant mass and probability of the SV for each opposite-sign muons. (a) and (c) shows the invariant masses for OS1 and OS2, respectively. On the other hand, (b) and (d) shows the probabilities for OS1 and OS2, respectively. The peaks of $\rho \to \mu\mu$ ($M_\rho \sim 770\,\mathrm{MeV}$), $\omega \to \mu\mu$ ($M_\omega \sim 782\,\mathrm{MeV}$), and $\phi \to \mu\mu$ ($M_\phi \sim 1020\,\mathrm{MeV}$) can be seen in the distribution of the sideband data (black solid line), while the distributions of the signal samples (red and blue dashed lines) are smooth without any peaks.



	doublet_os1_mass	doublet_os1_pvalue	doublet_os2_mass	doublet_os2_pvalue	triplet_pvalue	triplet_sigaxy	triplet_siglxy	trk1_mu_bal	trk1_mu_scat	trk1_pval	trk2_mu_bal	trk2_mu_scat	trk2_pval	trk3_mu_bal	trk3_mu_scat	trk3_pval
trk3_pval				3	2	-2		1	-2	7		1	4	-1	-1	100
trk3_mu_scat			-1	-1	-3		3	2		-1		-3		3	100	-1
trk3_mu_bal	1	-1	1	-1	-2		2					-2	1	100	3	-1
trk2_pval	1	1	-2	3	6	-3	-1	5		6	2		100	1		4
trk2_mu_scat	1	2		3	2	1	1		2		2	100		-2	-3	1
trk2_mu_bal	3	3			1			3	1	3	100	2	2			
trk1_pval	1	5	-2	1	4			-1		100	3		6		-1	7
trk1_mu_scat	1		-1	-3	-1	1		8	100		1	2				-2
trk1_mu_bal			1	-1		1		100	8	-1	3		5		2	1
triplet_siglxy	2	-3	-1	-5	-4	53	100					1	-1	2	3	
triplet_sigaxy		-4		-2	-5	100	53	1	1			1	-3			-2
triplet_pvalue		49		50	100	-5	-4		-1	4	1	2	6	-2	-3	2
doublet_os2_pvalue	2	13	-1	100	50	-2	-5	-1	-3	1		3	3	-1	-1	3
doublet_os2_mass	-46		100	-1			-1	1	-1	-2			-2	1	-1	
doublet_os1_pvalue		100		13	49	-4	-3			5	3	2	1	-1		
doublet_os1_mass	100		-46	2			2		1	1	3	1	1	1		

Figure 11.10: Correlation matrix (in %) of the variables used in BDT. The SV-based variables of the triplet, OS1, and OS2 are correlated but with approximately 50%, which means the discriminate power independently.

11.1.3 MVA selection

A MVA technique is employed also in the search for the HF-originated signature. The BDT method is used with the same configuration as that in the 2012 analysis, while the set of the variables is modified. The variables which are used for the training and the calculation of the score are chosen to be:

- Probability of the SV.
- Significance of $L_{\rm xy}$ and $a^0_{\rm xy}$.
- Invariant masses of the two pairs of opposite-sign muons in the triplet ($M_{\rm OS1}$ and $M_{\rm OS2}$)
- Provabilities of the SV with two pair of opposite-sign muons in the triplet (SV$_{OS1}$ and SV$_{OS2}$)
- Probability of each muon track in the triplet (p-value$_{\mu1,\mu2,\mu3}$)
- Muon momentum balance and muon scattering angle for each muon.

The correlation matrix is shown in Figure 11.10, and the responses for the signal and the sideband data are shown in Figure 11.11. Loose selection criteria were defined in the analysis procedure in 2012 in order to estimate the number of background events. However, the number of sideband events is expected to be larger than the $W \to \tau\nu \to 3\mu\nu$ search due to the looser separation of the BDT. Therefore, the sideband region with the tight criteria (x_1), which is set to be 0.95 for now, is used for the estimation of background events.

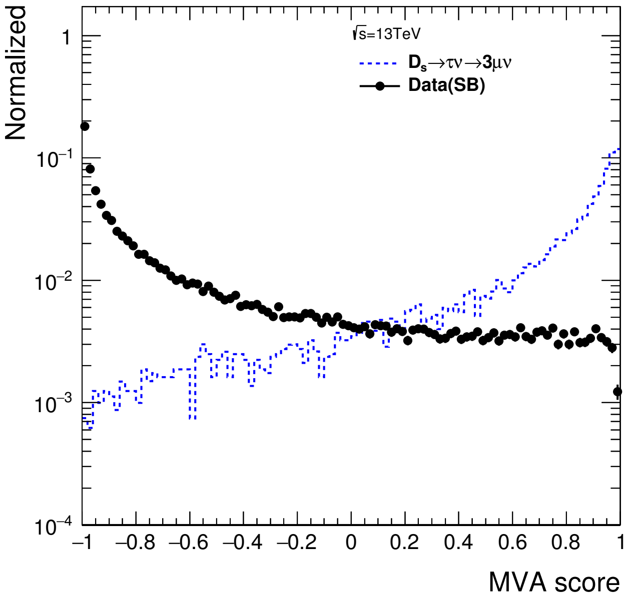


Figure 11.11: BDT response distributions for the signal sample (blue dashed line) and the sideband data (black solid line).

11.1.4 Background estimation

The triplet mass distributions of the sideband data and the signal are shown in Figure 11.12 after the tight selection criteria (BDT score $> x_1$) is applied. In order to determine the shape of the background events, the sideband distribution is fitted by a cubic polynomial. Number of background events in the signal region is estimated to be 55.6 by the integration of the sideband fit function within the range. This number is obtained for the integrated luminosity of 18.9 fb^{-1}, therefore it should be multiplied for the prospect for the 100 fb^{-1}. The systematic uncertainty of the background estimation was set to be 68% from the variation of the fitting functions in the $W \to \tau\nu \to 3\mu\nu$ analysis, and the value is conservatively kept to be the same for HF-originated signal search. The number of signal events is also obtained after the all selections. The selection efficiency, including the acceptance, the trigger requirements, and the event selections, is estimated to be 0.36%, which is approximately one-tenth of $W \to \tau\nu \to 3\mu\nu$ case (3.74%). This is estimated only by the signal simulation.

11.2 Measurement of $D_s \to \phi\pi \to \mu\mu\pi$

$D_s \to \phi\pi \to \mu\mu\pi$ events are used for normalization of the number of D_s mesons and validation of the reconstruction of three collimated tracks. There are two muons in the final state, therefore trigger chains requiring three muons cannot be used for this analysis. In addition, in order to simplify the kinematics mixture,

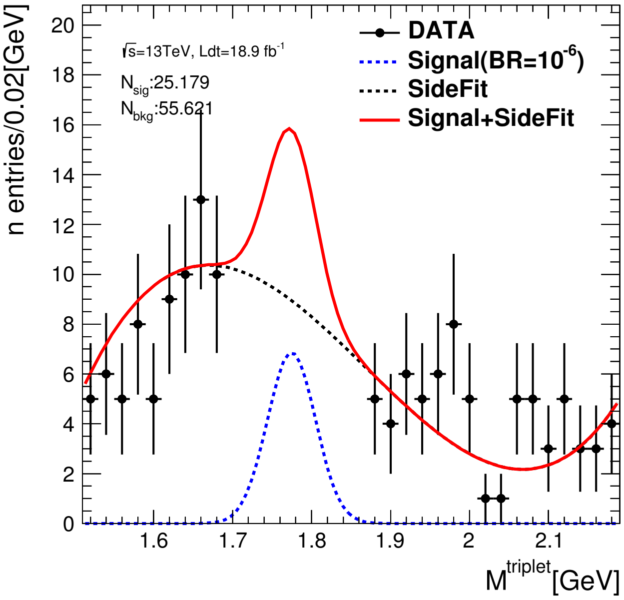


Figure 11.12: Invariant mass distribution of the triplet. $\mathcal{BR}(\tau \to 3\mu)$ in the $D_s \to \tau\nu \to 3\mu\nu$ (blue dashed line) is assumed to be 10^{-6}. The sideband data (black solid line) is fitted by a cubic polynomial (red solid line).

only `HLT_mu11_mu6noL1_nscan_L1MU11_MU6` is used in this measurement. Fraction of data set with integrated luminosity of $14.8\,\mathrm{fb}^{-1}$ is used in this measurement. The basic procedure of the event selections are the same as that of $D_s \to \tau\nu \to 3\mu\nu$ search but an opposite-sign muon pair and a non-muon track are required instead of three muons. Hence, the selections for an opposite-sign muon pair and a non-muon track should be used instead of two opposite-sign muon pairs and only muons.

The production cross-section of D_s which is determined by this measurement is used for the $\mathcal{BR}(\tau \to 3\mu)$ calculation. However, the difference between the data and the simulation also should be validated. Comparison of the L_{xy} distribution between the data and simulation is important for validation of the fraction of direct and indirect production processes, i.e. $B \to D_s + X$. The differences of the cross-section and the fraction between the data and simulation are evaluated.

11.2.1 Event selection for 2μ and one ID track

The quality requirements for muons are the same as in Section 11.1. Therefore, only the requirements for the non-muon track is described in this section. First, ID tracks which is used for the reconstructed muons without the quality selection are rejected. The remaining tracks are defined as non-muon tracks, and required to pass the following selections:

- If a hit at the innermost Pixel layer (B-layer or IBL) is expected, the number of hits should be larger than or equal to 1
- $N_{\mathrm{hits,Pixel}} \geq 2$
- $N_{\mathrm{hits,SCT}} \geq 5$
- $N_{\mathrm{holes,Pixel+SCT}} \leq 2$
- $N_{\mathrm{hits,TRT}} \geq 6$ and $N_{\mathrm{outhits,TRT}}/N_{\mathrm{hits,TRT}} < 90\%$ ($|\eta| < 1.9$)
  $N_{\mathrm{hits,TRT}} \geq 6$ or $N_{\mathrm{outhits,TRT}}/N_{\mathrm{hits,TRT}} < 90\%$ ($|\eta| \geq 1.9$)
- Ratio of the TRT hits with high energy deposition to all the associated hits on TRT should be larger than or equal to 0.35

These selections are applied only for hits at the ID. After the quality selections for muons and non-muon tracks are applied, their p_{T} and η are also limited by the detector acceptance. These selections are set as follows:

- $-3 < \sigma p_{\mathrm{bal.}} < +4$ (only for muons)
- $\sigma\Theta_{\mathrm{scat.}} < 4$ (only for muons)
- $|\eta| < 2.5$
- p_{T} for leading, 2nd leading muons, and non-muon track should be larger than 10.5 GeV, 5.5 GeV, and 2.5 GeV, respectively

The thresholds for the leading and 2nd leading muons are tightened, considering the thresholds at the trigger level.

In this measurement, a cut-based selection is employed because the statistics is enough for measurement. The signal selection criteria are defined as below:

- p-value$_{\mathrm{triplet}} > 10^{-3}$.
- $L_{\mathrm{xy}}/\Delta L_{\mathrm{xy}} > 3.0$ and $a^0_{\mathrm{xy}}/\Delta a^0_{\mathrm{xy}} < 2.5$.
- Invariant mass of the opposite-sign muons within the ϕ mass range from 1000 MeV to 1040 MeV.

The distribution of p-value of the triplet SV is shown in Figure 11.13. In the same manner as the $\tau \to 3\mu$ analysis procedure, the sideband region is defined to be [1600,1750] MeVand [2150,2400] MeV. It distributes broadly, therefore a loose threshold is applied in order to reject bad SVs with small probability. L_{xy}, a^0_{xy}, and their significances are shown in Figure 11.14. Large L_{xy} and small a^0_{xy} are expected in the signal events, therefore the tight selections are applied for the significance. The invariant mass peak of $\phi \to \mu\mu$, shown in Figure 11.15, is an important characteristic feature. Hence, the window is required for the invariant mass of opposite-sign muons.

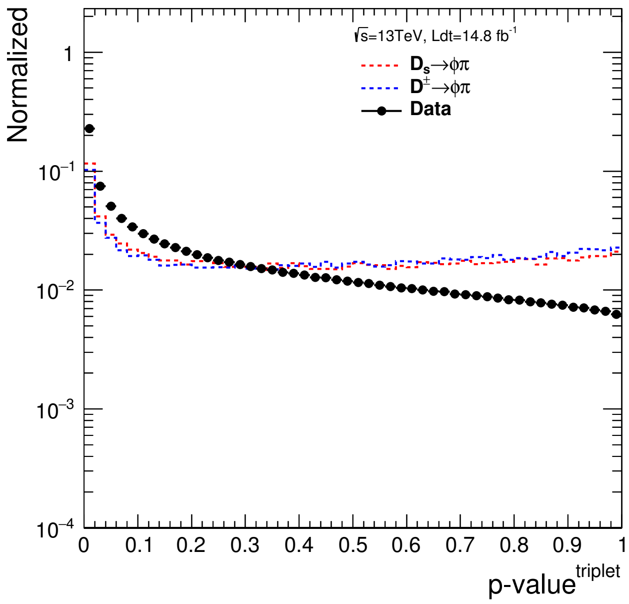


Figure 11.13: Distribution of the probability of triplet SV. Red and blue dashed lines represent the simulation samples of $D_s \to \phi\pi \to \mu\mu\pi$ and $D^{\pm} \to \phi\pi \to \mu\mu\pi$, respectively. The black points represent the data.

11.2.2 Validation of the D_s production

Figure 11.16 shows the invariant mass distribution of the triplet after the signal selection criteria are applied. In this figure, peaks of $D^{\pm} \to \phi\pi \to \mu\mu\pi$ and $D_s \to \phi\pi \to \mu\mu\pi$ are found. Numbers of events of these signals are determined by fitting the distribution with two Gaussian for signals and an exponential for the background component:

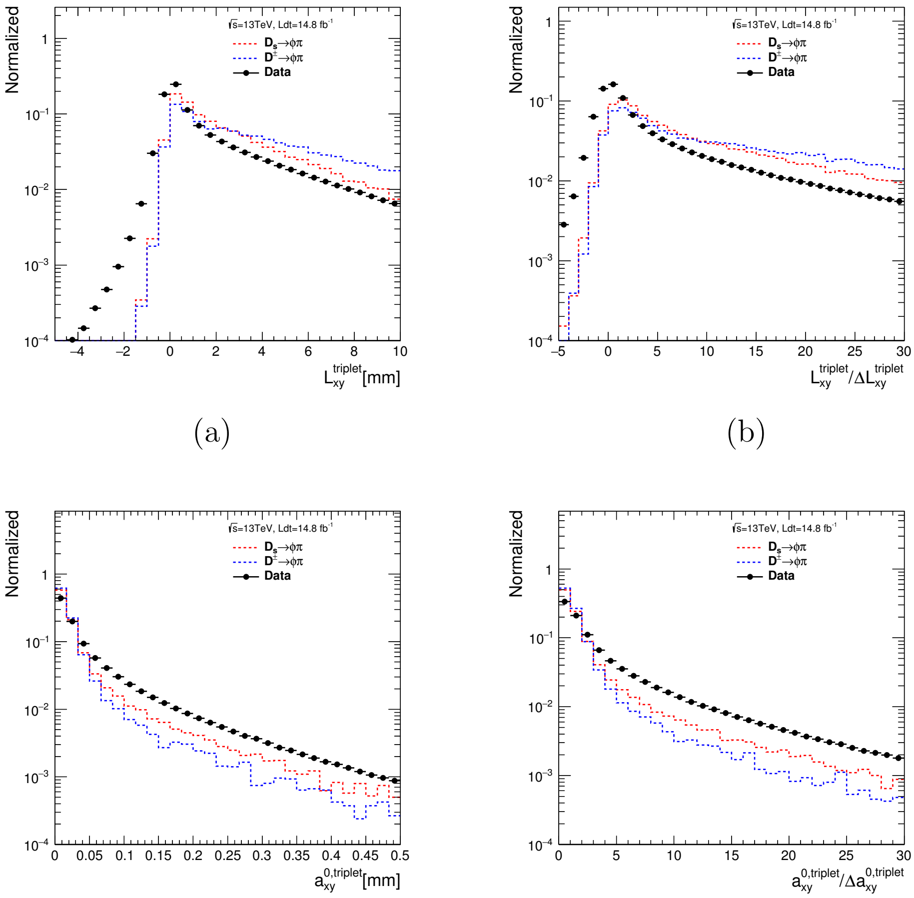
$$N_{D_s \to \phi\pi \to \mu\mu\pi,obs} = 3170 \pm 74(stat), \tag{11.1}$$

$$N_{D^{\pm} \to \phi\pi \to \mu\mu\pi,obs} = 824 \pm 58(stat). \tag{11.2}$$

The uncertainties for these numbers are obtained by the parameters in fitting function. The inclusive selection efficiency ($A \times \epsilon_{D_s \to \phi\pi \to \mu\mu\pi}$) is estimated to be 2.4% by the simulation. The dominant systematic uncertainty of the efficiency is assumed to originate from the trigger efficiency. This is estimated to be 8.5%, which is smaller than $\tau \to 3\mu$ case because it used only one chain, in the same method as that of the measurement in 2012.

The production cross-section of D_s (σ_{D_s}) is calculated by $N_{D_s \to \phi\pi \to \mu\mu\pi,obs}$ using the selection efficiency and the integrated luminosity ($\mathcal{L}dt$):

$$\begin{aligned}\sigma_{D_s} &= \frac{N_{D_s \to \phi\pi \to \mu\mu\pi,obs}}{(A \times \epsilon_{D_s \to \phi\pi \to \mu\mu\pi} \times \mathcal{BR}(D_s \to \phi\pi \to \mu\mu\pi) \times \mathcal{L}dt} \\ &= (6.70 \pm 0.78) \times 10^2 \text{nb},\end{aligned} \tag{11.3}$$



(a)

(b)

(c)

Figure 11.14: Distributions of L_{xy}, a^0_{xy}, and their significances. Red and blue dashed lines represent the simulation samples of $D_s \to \phi\pi \to \mu\mu\pi$ and $D^\pm \to \phi\pi \to \mu\mu\pi$, respectively. The black points represent the data.

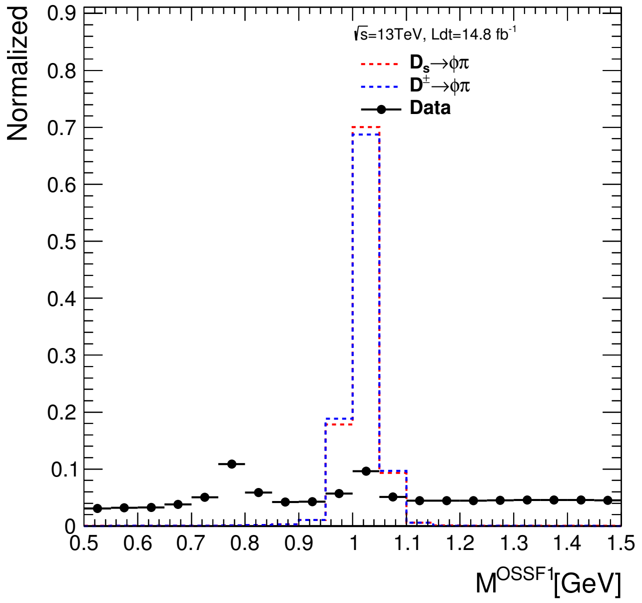


Figure 11.15: Invariant mass distribution of the opposite-sign muon pair. $D_s \rightarrow \phi\pi \rightarrow \mu\mu\pi$ (red dashed line) and $D^{\pm} \rightarrow \phi\pi \rightarrow \mu\mu\pi$ (blue dashed line) simulation samples have a peak of the $\phi \rightarrow \mu\mu$ process around 1020 MeV. Even in the data (black points), two peaks by $\rho/\omega \rightarrow \mu\mu$ and $\phi \rightarrow \mu\mu$ processes are found around 770 MeVand 1020 MeV, respectively.

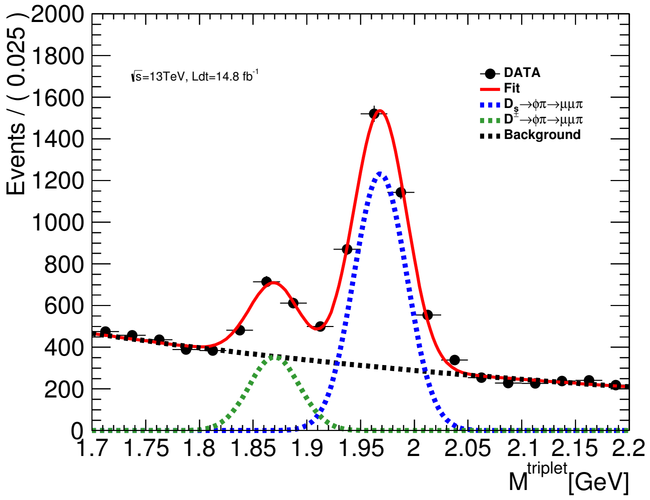
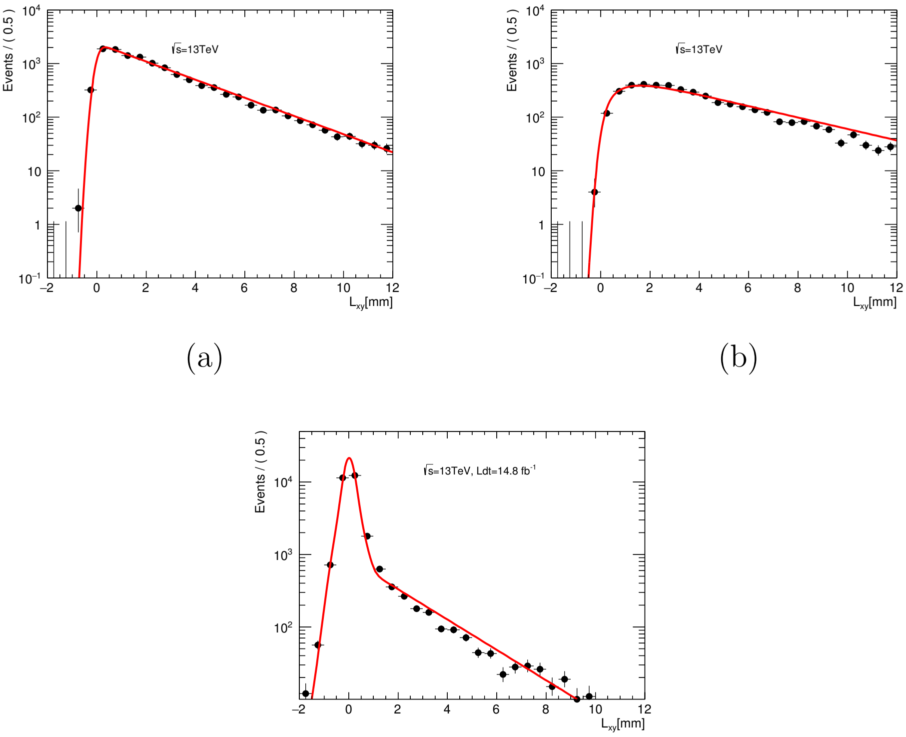


Figure 11.16: Invariant mass distribution which is fitted by the two Gaussian and an exponential functions. Two peaks are found around 1870 MeV and 1970 MeV in the data (black points). These signatures represent the $D^{\pm} \rightarrow \phi\pi \rightarrow \mu\mu\pi$ and $D_s \rightarrow \phi\pi \rightarrow \mu\mu\pi$ processes, respectively. The two Gaussian components are shown as a green and a blue dashed line, respectively.



(a) (b)

(c)

Figure 11.17: L_{xy} distribution of (a) the direct production, (b) indirect production of the simulation sample, and (c) the sideband data. These distributions are used as templates for the fitting of the events in the D_s mass range.

where $\mathcal{BR}(D_s \to \phi\pi \to \mu\mu\pi)$ is well measured by several experiments and determined to be $(1.63 \pm 0.12) \times 10^{-6}$ [93]. The uncertainty of the branching ratio is also considered as a systematic uncertainty of the measurement. As the production cross-section of D_s mesons is measured using a similar final state to that of the $D_s \to \tau\nu \to 3\mu\nu$ signature, the comparison with the simulation gives the validation of the expected number of $D_s \to \tau\nu$ signature. The production cross-section of the signal simulation by PYTHIA8B with leading order calculation is estimated to be 6.44×10^2 nb. This value shows a good agreement with that from the data within the uncertainty. The uncertainty of the N_{D_s} measurement is not set by the difference between the data and simulation, but set to be 11% as the same as that for the measurement conservatively.

In this measurement, the indirect production process from B meson decay is inclusively measured. This process is also involved in the $D_s \to \tau\nu \to 3\mu\nu$ search, therefore the fraction of this process should be validated. This is evaluated by the fitting of the L_{xy} distribution which is sensitive to the additional lifetime by a B meson. Figure 11.17 shows the L_{xy} distributions of direct and indirect processes, as well as the sideband data. The distribution in the signal region can

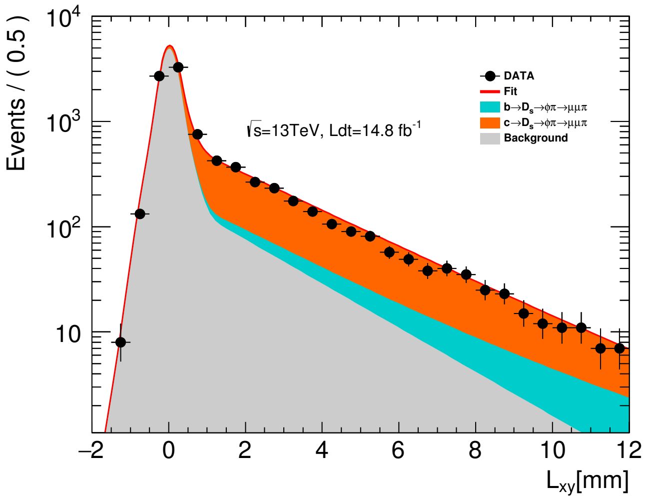


Figure 11.18: Fitting result of the $L_{\rm xy}$ distribution (black points) in the tight mass range. Blue and orange filled components show the direct and indirect components, and a gray component shows the background distribution. The total fit result ($f_{tot}(L_{\rm xy})$ in Equation 11.4) is shown as a red solid line.

be understood by the mixture of these three contributions:

$$f_{tot}(L_{\rm xy}) = S_{sigA} \cdot f_{sigA}(L_{\rm xy}) + S_{sigB} \cdot f_{sigB}(L_{\rm xy}) + B \cdot f_{bkg}(L_{\rm xy}), \tag{11.4}$$

where S_{sigA}, S_{sigB}, and B are the number of events of the direct production, the indirect production, and the sideband data, respectively. Total number of signal (S_{tot}) is defined as sum of the direct and indirect production, i.e. $S_{tot} = S_{sigA} + S_{sigB}$. Then, S_{tot} and B are determined and fixed by fitting of the invariant mass distributions. $f_{sigA}(L_{\rm xy}), f_{sigB}(L_{\rm xy})$, and $f_{bkg}(L_{\rm xy})$ are the fixed template distribution of each process shown in Figure 11.17. Therefore, the S_{sigB}/S_{tot}, which represents the fraction of the indirect process, is obtained by the fitting with the Equation 11.4. In this procedure, the tighter mass range [1940 MeV − 1990 MeV] is required in order to reject the $D^{\pm} \to \phi\pi \to \mu\mu\pi$ contamination. Figure 11.18 shows the fitting result of the $L_{\rm xy}$ distribution, and the fraction is determined to be:

$$S_{sigB}/S_{tot}(obs.) = 14.7 \pm 2.8(stat.) \pm 1.8(syst.)\%. \tag{11.5}$$

The systematic uncertainty of the fraction is determined from the error of fitting parameters. On the other hand, the fraction in the simulation is 14.8%, in good agreement with the measurement. This confirms the mixture of indirect production in simulation is appropriate. This precision is sufficient to validate the signal modeling with negligible sensitivity to the mixture of indirect process.

11.3 Extension to inclusive search of $\mathrm{HF} \to \tau \to 3\mu$

In the $D_s \to \tau\nu \to 3\mu\nu$ search described in the previous section, the contamination of the signals from the decay of B mesons is not considered. Therefore, the other processes also should be included in the estimation of number of signal events. It should be a good assumption to scale by the fraction of the $D_s \to \tau\nu \to 3\mu\nu$ in the all $\mathrm{HF} \to \tau \to 3\mu$ processes because the feature of $\mathrm{HF} \to \tau \to 3\mu$ events is expected to be similar to that of $D_s \to \tau\nu \to 3\mu\nu$. This assumption is validated by the comparison between $D_s \to \tau\nu \to 3\mu\nu$ signal and inclusive $\mathrm{HF} \to \tau \to 3\mu$ events. The fraction is estimated by using the simulation samples, and the difference of the selections efficiency with the same requirements is considered as the systematic uncertainty of the extension to the inclusive $\mathrm{HF} \to \tau \to 3\mu$ search.

11.3.1 Validation of the similarity

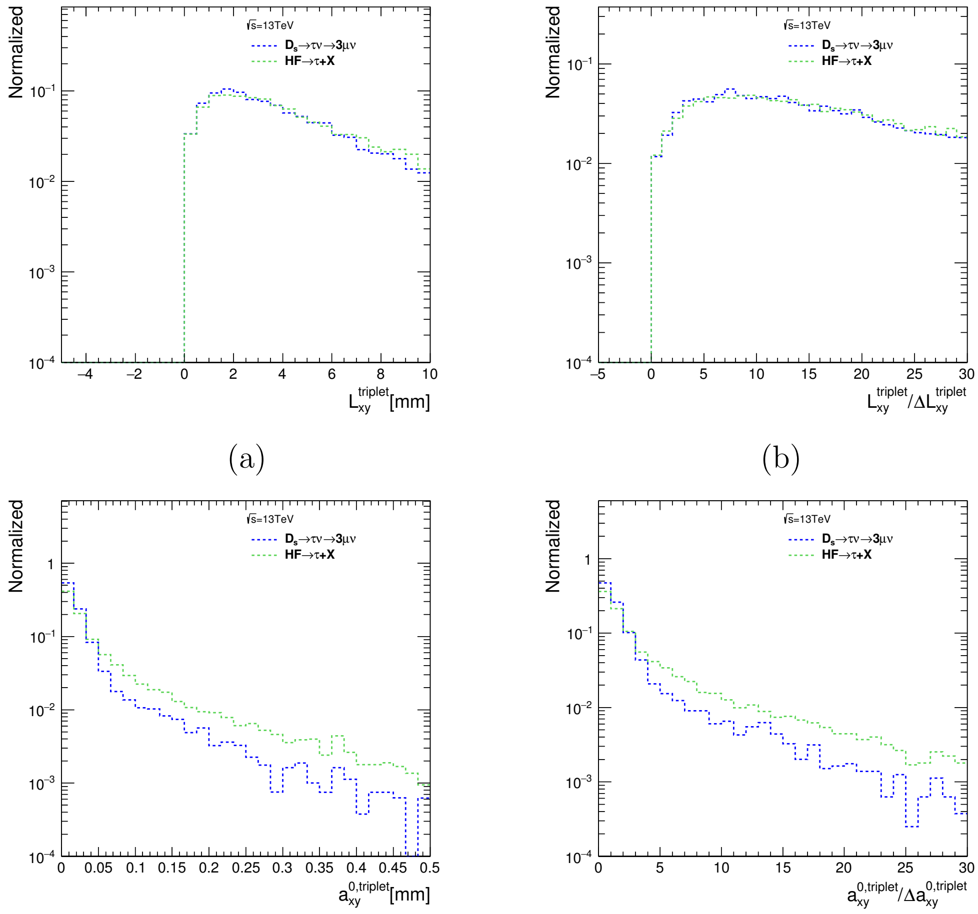
First, the similarity of $D_s \to \tau\nu \to 3\mu\nu$ and $\mathrm{HF} \to \tau \to 3\mu$ signal features needs to be validated in order to ensure that this assumption is appropriate. Figure 11.19 shows the comparison of the L_{xy}, a^0_{xy}, and their significance distributions. The L_{xy} distributions show a good similarity, while the a^0_{xy} distributions show a difference. This difference can be understood to be caused by the large multiplicity of the other emitted particles from B decay. For example, in the $B^\pm \to D^0\tau^\pm\nu$ process, p_{T} of the τ lepton is not on the direction of the flight length of the $B^\pm$ by the momentum of the D^0 meson. The probabilities of the SV also show good agreement as shown in Figure 11.20. Therefore, the signal selection efficiency can be affected by the a^0_{xy} difference. However, this affects the difference of the response of the BDT trained for $D_s \to \tau\nu \to 3\mu\nu$ signature as shown in Figure 11.21.

11.3.2 Fraction of the source of τ lepton

The fraction of $D_s \to \tau\nu \to 3\mu\nu$ is determined by the expected number of events after the BDT selection. Figure 11.22 shows the triplet mass distributions of inclusive $\mathrm{HF} \to \tau \to 3\mu$. The expected numbers in the SR with $\mathcal{BR}(\tau \to 3\mu)= 10^{-6}$ are estimated to be 42.8, while that of the $D_s \to \tau\nu \to 3\mu\nu$ signature is 25.2. Therefore, the fraction ($F_{Ds/HF}$) is set to be (1/1.69) and the expected number of HF-originated τ leptons are expected by the following equation:

$$N_{HF\to\tau+X} = \frac{N_{Ds} \times \mathcal{BR}(D_s \to \tau\nu)}{F_{Ds/HF}}. \tag{11.6}$$

The uncertainty of the extension is estimated by the efficiency difference after the event selection in the $D_s \to \tau\nu \to 3\mu\nu$ search. The selection efficiency for the



(a) (b)

(c) (d)

Figure 11.19: (a) L_{xy} and (c) a^0_{xy} distributions and (b,d) their significances for the $D_s \to \tau\nu \to 3\mu\nu$ signal (blue dashed line) and the inclusive HF $\to \tau \to 3\mu$ sample (green dashed line).

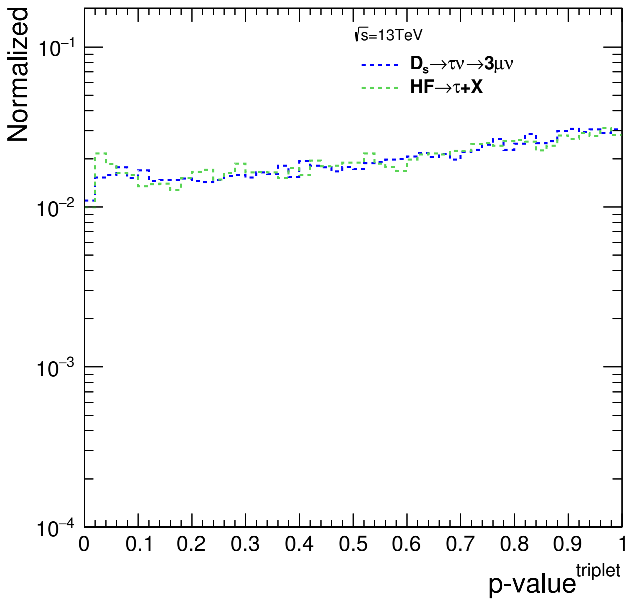


Figure 11.20: Distributions of p-value of the SV for the $D_s \to \tau\nu \to 3\mu\nu$ signal (blue dashed line) and the inclusive HF $\to \tau \to 3\mu$ sample (green dashed line).

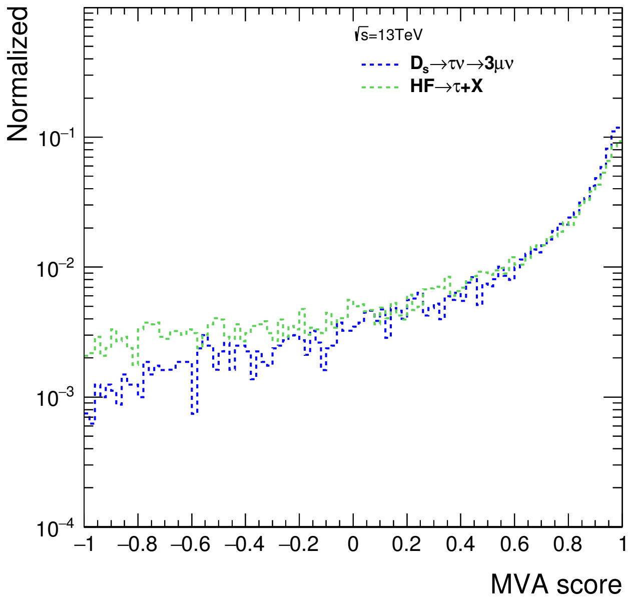


Figure 11.21: BDT response of the $D_s \to \tau\nu \to 3\mu\nu$ and the inclusive HF $\to \tau \to 3\mu$ signature. The difference can be understood as the effect of a^0_{xy}.

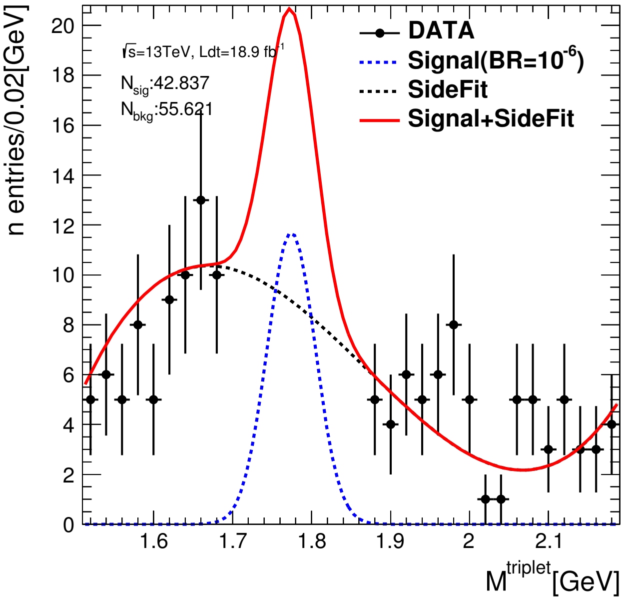


Figure 11.22: Mass distributions for inclusive HF $\rightarrow \tau \rightarrow 3\mu$ signature and the sideband data in the same manner as Figure 11.12. The yield of the $D_s \rightarrow \tau\nu \rightarrow 3\mu\nu$ limited case is an under-estimated value, because the τ leptons from decay of B mesons is not considered. Therefore, factorization for the inclusive search is necessary.

Table 11.1: $A \times \epsilon$, N_{Ds} ($\mathcal{BR}(D_s \to \tau\nu)$), $F_{Ds/HF}$, and N_{bkg} in the SR determined in the previous sections. Integrated luminosity of 100 fb^{-1}is assumed.

$A \times \epsilon$	$0.36\% \pm 0.065$ ($syst. = 18\%$)
N_{D_s}	$(3.67 \pm 0.40) \times 10^{10}$ ($syst. = 11\%$)
$F_{Ds/HF}$	$1/(1.69 \pm 0.09)$ ($syst. = 5.6\%$)
$N_{bkg}(SR)$	294.2 ± 200.1 ($syst. = 68\%$)

inclusive HF → $\tau \to 3\mu$ signature is estimated to be 0.36% by the simulation. Therefore, the uncertainty is determined to be $\{1 - (0.34/0.36)\} \times 100 = 5.6\%$.

11.4 Prospect of the HF $\to \tau \to 3\mu$ search

The expectations of the signal efficiency, the number of background, and the number of HF-originated τ leptons are described above. In this section, sensitivity to $\mathcal{BR}(\tau \to 3\mu)$ from HF $\to \tau \to 3\mu$ is estimated.

11.4.1 Numbers for the $\mathcal{BR}$ calculation

The $\mathcal{BR}(\tau \to 3\mu)$ is calculated by the following equation in this channel:

$$\mathcal{BR}(\tau \to 3\mu) = \frac{N_{obs}}{(A \times \epsilon) \times N_{HF \to \tau + X}} \tag{11.7}$$

$$= \frac{N_{obs} \times F_{Ds/HF}}{(A \times \epsilon) \times N_{Ds} \times \mathcal{BR}(D_s \to \tau\nu)} (\because \text{Equation 11.6}). \tag{11.8}$$

$A \times \epsilon$ is determined to be 0.36% by the selection efficiency for the simulation samples of $D_s \to \tau\nu \to 3\mu\nu$, and its systematic uncertainty is assumed to be 18% as the same as that of $W \to \tau\nu \to 3\mu\nu$ search. N_{Ds} is estimated to be 3.67×10^{10} by the production cross-section of $D_s \to \tau\nu \to 3\mu\nu$, which is validated by the measurement of $D_s \to \phi\pi \to \mu\mu\pi$ signature. The systematic uncertainty is set to be 11% by this validation measurement. $F_{Ds/HF}$ is determined by the fraction of the $D_s \to \tau\nu \to 3\mu\nu$ in the inclusive process of HF $\to \tau \to 3\mu$. The value is set to be 1.69 fully based on the simulation, and the systematic uncertainty is estimated to be 5.6% by the difference of the selection efficiency for the same procedure. In addition, N_{bkg} is also estimated by the fitting of the triplet invariant mass in the sideband region. This is estimated to be 294.2 for the integrated luminosity of 100 fb^{-1}, based on the result with 18.9 fb^{-1}. The systematic uncertainty is conservatively set to be 68%, which is the same as that in $W \to \tau\nu \to 3\mu\nu$ analysis. These measured and estimated numbers are summarized in Table 11.1.

11.4.2 Expected sensitivity in Run2

The sensitivity is estimated with the pseudo observed number of events (N_{obs}) as the same as N_{bkg}. The upper limit for the $\mathcal{BR}(\tau \to 3\mu)$ with $100\,\mathrm{fb}^{-1}$of data is expected as follows:

$$\mathcal{BR}(\tau \to 3\mu)_{N_{obs}=N_{bkg}} < 5.6 \times 10^{-7}. \tag{11.9}$$

This is worse than that of $W \to \tau\nu \to 3\mu\nu$ search, however it can be understood by the lower efficiency and worse separation from background events due to the lack of discriminating variables. Nevertheless, it is the importance of this prospect that another approach in the ATLAS experiment is shown. This analysis procedure is still preliminary for the inclusive search of HF-originated $\tau \to 3\mu$ search. The development of the variables enhancing the signal feature will improve the efficiency. In addition, the systematic uncertainty of the background estimation should be reduced by using large amount of the sideband data in the integrated luminosity of $100\,\mathrm{fb}^{-1}$. Therefore, it is expected to be improved in the actual analysis using the whole Run2 data.

Part IV

Conclusion

Chapter 12

Summary of the Analysis and Future Prospect

In this thesis, existence of BSM physics is investigated by searching for a cLFV signature, which is strictly suppressed in the SM framework. This search is a complementary approach to direct searches for the production of new heavy particles. Experiments with electron-positron colliders and muon beams have played main roles in such rare decay searches of leptons. On the other hand, the LHC provides a large number of τ leptons in the high energy proton-proton collisions. It enables us to search for rare decays of τ leptons in hadron collider experiments.

The target of search in the ATLAS experiment is $\tau \to 3\mu$ signature, which has several characteristic features. $W \to \tau\nu$ and HF$\to \tau + X$ are dominant processes of τ production in pp collisions. However, these signatures have different kinematic features in $E_{\mathrm{T}}^{\mathrm{miss}}$, collinear hadron jets, flight length, and their momenta. Therefore, it is necessary to optimize the analysis strategy separately for each channel. One of the most challenging issues of this search is the efficiency of the triggers used in data acquisition. The tight thresholds for the muons and inefficiency for highly collimated muons at the trigger requirements kill large amount of the events relevant to the signature. Hence, estimations and improvements of the trigger efficiencies are critical tasks.

Using the data taken in 2012, analysis of $W \to \tau\nu \to 3\mu\nu$ search was performed and the first result from ATLAS was presented. In the analysis in 2012, it was clarified that there is a sizable room for further improvements in this search in ATLAS including the addition of HF $\to \tau \to 3\mu$ channel. Hence, in addition to the result of 2012 analysis, prospects for Run2 with several improvements were studied in this thesis.

In this chapter, the summary and conclusion of the search are described.

12.1 Result and lessons from 2012

The target signature of the analysis in 2012 was $W \to \tau\nu \to 3\mu\nu$. This search has characteristic features in several distributions of physics objects; three muons, hadron jets, $E_{\mathrm{T}}^{\mathrm{miss}}$, and their combination. In order to maximize the sensitivity of the signature, MVA technique is employed in the final event selection instead of cut-based approach. The selection with MVA achieves to suppress number of background events below one event in the signal region, keeping the signal efficiency high.

On the other hand, the efficiency of the trigger and reconstruction needs to be estimated. The correlation between highly collimated muons has to be considered, therefore the efficiency for the target signature cannot be estimated by simply multiplying the efficiency for each muon. Evaluation of the correlation effect as a function of ΔR between muons enables to take account for this effect for efficiency calculation of multi-muon triggers. The estimated efficiency with the correlation factor measured in the data shows a good agreement with that in the simulation, and the difference is used as a systematic uncertainty.

The number of background events and $W \to \tau\nu$ events are estimated. The sideband, which is defined as the region around τ lepton mass excluding the signal region, is used for the background estimation. This number is also used for optimization of the final selection criteria for MVA score. Number of $W \to \tau\nu$ events is estimated by extrapolating the measurement in 7 TeV analysis. This is used for the calculation of $\mathcal{BR}(\tau \to 3\mu)$ from the observed number of events.

As a result of the search, no event is found in the signal region. Hence, the upper limit for the $\mathcal{BR}(\tau \to 3\mu)$ is set to be 3.76×10^{-7} at 90% CL. Although this is not the most stringent limit, it is the first result of the search with the specific analysis procedure at a hadron collider and shows the possibility of a search for $\tau \to 3\mu$ signature in the future of the ATLAS experiment. Also, this shows that there are rooms for improvements. Especially for the efficiency, it should be considered to avoid close-by muon inefficiency. This should be solved by the implementation of dedicated triggers and an alternative reconstruction algorithm. The addition of HF-originated analysis is also the important extension for the sensitivity improvements. This channel was not the target of this search in 2012 because of the low efficiency and worse separation from background events, although it has ten times larger amount of τ leptons. Hence, this also will be sensitive in future analyses with the improvement of efficiency and selection tuning.

12.2 Future of $\tau \to 3\mu$ search in ATLAS

In order to improve the sensitivity of the search using the data taken in Run2 period (2015-2018), several approaches are considered. One of them is the improvement of the efficiency of the trigger and reconstruction. Implementation of dedicated trigger chains and the alternative reconstruction algorithm is expected to increase the signal efficiency. In addition, the production rate of W boson also increases by the high center-of-mass energy of 13 TeV. Therefore, sensitivity of the $W \to \tau\nu \to 3\mu\nu$ search will be improved and reach the level of 4.1×10^{-8} if there would be no event in the signal region.

On the other hand, HF $\to \tau \to 3\mu$ search is suggested to be effective in addition to $W \to \tau\nu \to 3\mu\nu$ search because of the larger production cross-section of HF-originated τ lepton. The dominant process of HF$\to \tau + X$ is $D_s \to \tau\nu$. Therefore, this method is designed to focus on $D_s \to \tau\nu \to 3\mu\nu$ events. The D_s production in the simulation is validated using $D_s \to \phi\pi \to \mu\mu\pi$ channel, which can be measured in the data. Not only the number of $D_s \to \tau\nu$ events, but the fraction of the direct and indirect process is also validated with this channel. Finally, this is extended to inclusive HF $\to \tau \to 3\mu$ search because other HF-originated processes cannot be rejected. The number of τ leptons are scaled by the fraction of the $D_s \to \tau\nu$ in all HF-originated processes, assuming similarity of the event feature. The sensitivity with the integrated luminosity of $100\,\text{fb}^{-1}$ and presence of SM background in the signal region is expected to be 5.6×10^{-7}. This sensitivity is not more stringent than W-originated analysis in Run2, although it can reach the same level as that of result in 2012. In addition, this analysis procedure is still the baseline, which has several possibilities to improve by the discriminant variables, the selection optimization, and the uncertainty reduction. Therefore, this approach is also worth considering in the future analysis.

12.3 Conclusion

A $\tau \to 3\mu$ search at a hadron collider is a challenging search due to the low efficiency of the trigger and reconstruction, and large number of background from QCD events. Nevertheless, it should be considered to use the large number of τ leptons from the decay of W bosons and heavy flavored hadrons as effective as possible.

In the analysis of the ATLAS experiment in 2012, the limit for $\mathcal{BR}(\tau \to 3\mu)$ is set for the first time [88]. Although this limit does not improve that by the Belle experiment, this is important for showing the possibility of the search in hadron collider experiments. In addition, a room for improvement is realized by the experience of the analysis development.

For the W-originated signature search, its sensitivity should be improved by the increase of the efficiency and the integrated luminosity from 2016 to 2018. Furthermore, HF-originated analysis can also be developed thanks to the improvement of the trigger efficiency. This is a new approach in ATLAS, and the production cross-section should be larger than W-originated signature. In fact, there are rooms for optimization in event selections, estimation of the number of background, and the normalization. Furthermore, the combination of the two approaches is not considered in this thesis, but it should be considered in the future. The ATLAS experiment also plans high luminosity runs after 2022, which will record huge amount of data of 3000 fb^{-1}. It is hoped that the $\tau \to 3\mu$ signature will be observed with more sensitive searches with these and further improvements.

This analysis focusses on the characteristic feature of the highly collimated muons. Therefore, the analysis procedure, the technique of the efficiency measurement and uncertainty verification can also be applied for other analyses, not only $\tau \to 3\mu$ search. For example, in the event that Higgs boson decays into new light neutral particles decaying to muon pairs ($h \to X + X \to \mu\mu + \mu\mu$), the muon pairs are highly collimated, called ”muon-jets” [94]. it is necessary for the search for the ”muon-jets” to apply a similar strategy to this $\tau \to 3\mu$ search. Also, these techniques can be applied for rare decay searches of B and D mesons. $B_s \to 2\mu$ (or 4μ) are the probes for the verification of the SM and signature of the BSM [95, 96]. Since these signatures are composed of the muons from the decay of light mass of a few GeV, the analysis for the collimated muons is necessary. Thus, the analysis technique which is developed in the $\tau \to 3\mu$ search is expected to be applied for the other searches in ATLAS.

Finally, in order to explore for the BSM evidence, it is necessary to continue many kinds of approaches from various angles. The ATLAS experiment has several direct and indirect approaches for cLFV searches. Hence, this experiment should play an important role to find out the existence of the BSM by several approaches.

Glossary

η Pseudo rapidity defined as Equation 2.4. 26, 28, 30–36, 60–62, 64, 68, 69, 71, 90, 93–96, 127, 128, 137

ΔR Angular distance determined by η and ϕ of the particles as Equation 2.7. 44, 78, 91, 93, 96–98, 102, 105, 118, 119, 125, 151, 163

BDT Boosted Decision Tree. 76–83, 85–88, 107–109, 134, 135, 143, 145

BSM Beyond the Standard Model. 8, 14, 15, 18–20, 54, 150, 153

cLFV Lepton-Flavor-Violation in the charged sector. 14–21, 45, 150, 153

CSC Cathode Strip Chamber. 34, 35, 39, 61, 63, 65, 67

HF Heavy Flavor. 43, 44, 47–49, 60, 62–65, 67, 70–72, 74, 75, 111, 112, 114–116, 124, 125, 127, 132, 134, 135, 143, 147, 148, 150–152

ID Inner Detector. 26, 27, 30, 34, 39, 41, 48, 59–61, 63, 64, 66, 69, 70, 90, 102, 103, 112, 118, 126, 136, 137

IP Interaction Point. 24, 26, 33, 36, 60

LFV Lepton-Flavor-Violation. 14, 15, 18

LHC Large Hadron Collider. 19, 21–24, 49, 52, 111, 114, 150

MDT Monitored Drift Chamber. 34–36, 39, 61, 63, 65, 67

MS Muon Spectrometer. 27, 34, 35, 39, 41, 48, 61, 63, 64, 70, 102, 126

MVA Multi-Variate-Analysis. 50, 57, 59, 73, 76, 79, 84, 127, 132, 134, 151

PDF Parton Distribution Function. 53, 54, 88, 114, 115

PV Primary Vertex. 45, 47, 48, 61, 71, 72, 78

RPC Resistive Plate Chamber. 34–36, 39, 63–68

SM Standard Model. 8–10, 13–20, 22, 41, 43, 54, 123, 150, 152

SV Secondary Vertex. 45, 47, 61, 71, 72, 78, 107, 112, 127, 132–134, 137, 138, 143, 145

TGC Thin Gap Chamber. 34, 36, 39, 63–68

TPa One of the types of the reconstructed muons in ATLAS, which is called Muon-Associated-Track-Particle. 61, 64, 66, 67, 70, 125

References

[1] D. H. Perkins, *Introduction to High Energy Physics.* Cambridge University Press, 2000.

[2] F. Halzen, A. D. Martin, *Quarks & Leptons.* John Wiley & Sons, Inc., 1984.

[3] C. N. Yang and R. L. Mills, "Conservation of Isotopic Spin and Isotopic Gauge Invariance," *Phys. Rev.*, vol. 96, 101, 1954.

[4] M. Gell-Mann, "A schematic model of baryons and mesons," *Phys. Lett.*, vol. 8, pp. 214–215, 1964.

[5] M. Y. Han, Y. Nambu, "Three Triplet Model with Double SU(3) Symmetry," *Phys. Rev. B*, vol. 139, pp. 1006–1010, 1965.

[6] A. Ceccucci, Z. Ligeti, and Y. Sakai., "THE CKM QUARK-MIXING MATRIX," 2016. http://www-pdg.lbl.gov/2016/reviews/rpp2016-rev-ckm-matrix.pdf.

[7] S. L. Glashow, "Partial-symmetries of weak interactions," *Nucl. Phys.*, vol. 22, pp. 579–588, 1961.

[8] S. Weinberg, "A Model of Leptons," *Phys. Rev. Lett.*, vol. 19, 1264, 1967.

[9] A. Salam and J. C. Ward, "Electromagnetic and weak interactions," *Phys. Lett.*, vol. 13, pp. 168–171, 1964.

[10] P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons," *Phys. Rev. Lett.*, vol. 13, pp. 508–509, 1964.

[11] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons," *Phys. Rev. Lett.*, vol. 13, pp. 321–323, 1964.

[12] ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC," *Phys. Lett. B*, vol. 716, pp. 1–29, 2012.

[13] The CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," *Phys. Lett. B*, vol. 716, pp. 30–61, 2012.

[14] V. Trimble, "Existence and nature of dark matter in the universe," *Annual review of astronomy and astrophysics*, vol. 25, pp. 425–472, 1987.

[15] Z. Xing, S. Zhou, *Neutrinos in Particle Physics, Astronomy and Cosmology.* Advanced Topics in Science and Technology in China, 2011.

[16] The Class for Physics of the Royal Swedish Academy of Sciences, "NEUTRINO OSCILLATIONS," 2015. Scientific Background on the Nobel Prize in Physics.

[17] Z. Maki. M. Nakagawa. S. Sakata., "Remarks on the Unified Model of Elementary Particles," *Prog. Theo. Phys.*, vol. 28:870, 1962.

[18] "Neutrino Oscillations." http://www.lppp.lancs.ac.uk/neutrinos/theory.html.

[19] R. H. Bernstein, P. S. Cooper, "Charged Lepton Flavor Violation: An Experimenter's Guide," *Phys. Rep.*, vol. 532, 2, pp. 27–64, 2013.

[20] E. Arganda and M. J. Herrero, "Testing supersymmetry with lepton flavor violating τ and μ decays," *Phys. Rev. D*, vol. 73, 055003, 2006.

[21] P. Langacker, "The Physics of Heavy Z' Gauge Bosons," *Rev. Mod. Phys.*, vol. 81, pp. 1199–1228, 2009.

[22] J. Barry and W. Rodejohann, "Lepton number and flavour violation in TeV-scale left-right symmetric theories with large left-right mixing," *JHEP*, vol. 153, 2013.

[23] W. Chang, J. N. Ng, "Lepton Flavor Violation in Extra Dimension Models," *Phys. Rev. D*, vol. 71, 053003, 2005.

[24] H.Miyazawa, "Baryon Number Changing Currents," *Prog. Theo. Phys.*, vol. 36(6), pp. 1266–1276, 1966.

[25] J. L. Gervais and B. Sakita, "Field theory interpretation of supergauges in dual models," *Nucl. Phys. B*, vol. 34, pp. 632–639, 1971.

[26] D. V. Volkov, V. P. Akulov, "Is the neutrino a goldstone particle?," *Phys. Lett. B*, vol. 46, pp. 109–110, 1973.

[27] J. Wess and B. Zumino, "Supergauge transformations in four dimensions," *Phys. Lett. B*, vol. 70, pp. 39–50, 1974.

[28] A. Brignole and A. Rossi, "Anatomy and Phenomenology of mu-tau Lepton Flavour Violation in the MSSM," *Nucl. Phys. B*, vol. 701, pp. 3–53, 2004.

[29] P. Paradisi, "Higgs-Mediated $\tau \to \mu$ and $\tau \to e$ transitions in II Higgs doublet Model and Supersymmetry," *JHEP*, vol. 50, 02, 2006.

[30] R. Barbier, et al., "R-parity violating supersymmetry," *Phys. Rept.*, vol. 420, pp. 1–202, 2004.

[31] A. de Gouvea, S. Lola, K. Tobe, "Lepton-flavor violation in supersymmetric models with trilinear R-parity violation," *Phys. Rev. D*, vol. 63, 035004, 2001.

[32] J. N. Esteves, et al., "LHC and lepton flavour violation phenomenology of a left-right extension of the MSSM," *JHEP*, vol. 77, 12, 2010.

[33] MEG Collaboration, "New Limit on the Lepton-Flavor-Violating Decay $\mu^+ \to e + \gamma$," *Phys. Rev. Lett.*, vol. 107, 171801, 2011.

[34] MEG Collaboration, "Search for the lepton flavour violating decay $\mu^+ \to e+\gamma$ with the full dataset of the MEG experiment," *Eur. Phys. J. C*, vol. 76, 434, 2016.

[35] *Results of the SINDRUM-II experiment*, vol. C980420, 1998, 534-546 of *Lepton and baryon number violation in particle physics,astrophysics and cosmology. Proceedings, 1st International Symposium, Lepton-baryon'98, Trento, Italy, April 20-25,1998*, Conf. Proc.

[36] M. Aoki, et al., "Proposal of an Experimental Search for $\mu - e$ Conversion in Nuclear Field at Sensitivity of 10^{-14} with Pulsed Proton Beam from RCS," *KEK_J-PARC-PAC*, vol. 13, 2010.

[37] The COMET Collaboration, "Conceptual Design Report for Experimental Search for Lepton Flavor Violating $\mu^- - e^-$ Conversion at Sensitivity of 10^{-16} with a Slow-Extracted Bunched Proton Beam," *KEK*, vol. 10, 2009.

[38] The Mu2e Project, Collaboration, "Mu2e Conceptual Design Report," *Fermilab-TM*, vol. 2545, 2012.

[39] A. Blondel, et al., "Research Proposal for an Experiment to Search for the Decay $\mu \to eee$." arXiv:1301.6113 [physics.ins-det].

[40] SINDRUM Collaboration, "Search for the decay $\mu^+ \to e^+e^+e^-$," *Nucl. Phys. B*, vol. 299, pp. 1–6, 1988.

[41] T. Mori and W. Ootani, "Flavour violating muon decays," *Prog. Part. Nucl. Phys.*, vol. 79, pp. 57–94, 2014.

[42] *CLEO Contributions to Tau Physics*, vol. 123, 2003.

[43] G.D.Lafferty, "Rare tau decays from BABAR," *CPC(HEP & NP)*, vol. 34, 6, pp. 1–5, 2010.

[44] Belle Collaboration, "Search for lepton-flavor-violating τ decays into three leptons with 719 million produced $\tau^+\tau^-$ pairs," *Phys. Lett. B*, vol. 687,2-3, pp. 139–143, 2010.

[45] Belle Collaboration, "Search for lepton-flavor-violating τ decays into a lepton and a vector meson," *Phys. Lett. B*, vol. 699,4, pp. 251–257, 2011.

[46] LHCb Collaboration, "Searches for violation of lepton flavour and baryon number in tau lepton decays at LHCb," *Phys. Lett. B*, vol. 724,1-3, pp. 36–45, 2013.

[47] LHCb Collaboration, "Search for the lepton flavour violating decay $\tau^- \to \mu^-\mu^+\mu^-$," *JHEP*, pp. 121, 02, 02, 2015.

[48] HFAG-Tau Summer 2014 Report. https://www.slac.stanford.edu/xorg/hfag/tau/summer-2014/lfv-limits-plot.html.

[49] ATLAS Collaboration, "Search for lepton-flavour-violating $H \to \mu\tau$ decays of the Higgs boson with the ATLAS detector," *JHEP*, p. 211, 11, 2015.

[50] CMS Collaboration, "Search for lepton-flavour-violating decays of the Higgs boson," *Phys. Lett. B*, vol. 749, pp. 337–362, 2015.

[51] ATLAS Collaboration, "Search for the lepton flavor violating decay $Z \to e\mu$ in pp collisions at $\sqrt{s} = 8\,\text{TeV}$ with the ATLAS detector," *Phys. Rev. D*, vol. 90, 072010, 2014.

[52] CMS Collaboration, "Search for Lepton Flavor Violation in Z decays in pp collisions at $\sqrt{(s)} = 8\,\text{TeV}$," *CMS-PAS*, vol. EXO-13-005, 2015.

[53] ATLAS Collaboration, "Search for a Heavy Neutral Particle Decaying to $e\mu$, $e\tau$, or $\mu\tau$ in pp Collisions at $\sqrt{s} = 8\,\text{TeV}$ with the ATLAS Detector," *Phys. Rev. Lett.*, vol. 115, 031801, 2015.

[54] CMS Collaboration, "Search for Lepton Flavour Violating Decays of Heavy Resonances and Quantum Black Holes to electron/muon Pairs in pp Collisions at a centre of mass energy of 8 TeV," *CMS-PAS*, vol. EXO-13-002, 2015.

[55] O.Steinkamp, N.Serra, B.Storaci, "Lepton Flavor Violating searches." http://www.physik.uzh.ch/lectures/flavour/14/lectures/LFV.pdf.

[56] T. Abe, et al., "Belle II Technical Design Report," *KEK Report*, vol. 1, 2010.

[57] http://atlas.kek.jp/sub/photos/Accelerator/PhotoAccelerator.html.

[58] https://lhc-machine-outreach.web.cern.ch/lhc-machine-outreach/lhc_in_pictures.htm.

[59] https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults.

[60] "Computer generated image of the whole ATLAS detector." https://cdsweb.cern.ch/ record/1095924.

[61] "Computer generated image of the ATLAS inner detector." https://cdsweb.cern.ch/ record/1095926.

[62] ATLAS Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider," *JINST*, vol. 3, S08003, 2008.

[63] ATLAS Collaboration, "ATLAS Insertable B-Layer Technical Design Report," *ATLAS TDR*, vol. 019, 2010.

[64] Patrick Czodrowski, "The ATLAS Trigger System: Ready for Run II," in *European Physical Society Conference on High Energy Physics*, 2015.

[65] ATLAS Collaboration. http://www.atlas.ch.

[66] S. van der Meer, "Calibration of the efective beam height in the ISR," *CERN-ISR-PO*, vol. 68-31, 1968.

[67] ATLAS Collaboration, "Luminosity Determination in pp Collisions at $\sqrt{s} = 8\,\mathrm{TeV}$ using the ATLAS Detector at the LHC," *CERN-PH-EP*, vol. 117, 2016.

[68] https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults.

[69] S. Agostinelli et al., "GEANT4: A Simulation toolkit," *Nucl. Instrum. Meth.*, vol. A506, pp. 250–303, 2003.

[70] ATLAS Collaboration, "The ATLAS Simulation Infrastructure," *Eur.Phys.J.*, vol. C70, pp. 823–874, 2010. arXiv:1005.4568 [physics.ins-det].

[71] T. Sjostrand, et al., "An Introduction to PYTHIA 8.2," *Comput. Phys. Commun.*, vol. 191, pp. 159–177, 2015. arXiv:1410.3012 [hep-ph].

[72] ATLAS Collaboration, "Summary of ATLAS Pythia 8 tunes," 2012. https://cds.cern.ch/record/1474107.

[73] A. Martin et al., "Parton distributions for the LHC," *Eur. Phys. J.*, vol. C63, pp. 189–285, 2009. arXiv:0901.0002 [hep-ph].

[74] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, "ALPGEN, a generator for hard multiparton processes in hadronic collisions," *JHEP*, vol. 07, p. 1, 2003.

[75] J. M. Butterworth and M. H. Seymour., "JIMMY4: Multiparton Interactions in Herwig for the LHC," https://jimmy.hepforge.org/.

[76] R. Fruhwirth, "Application of Kalman filtering to track and vertex fitting," *Nucl.Instrum.Meth. A*, vol. 262, pp. 444–450, 1987.

[77] ATLAS Collaboration, "Calorimeter Clustering Algorithms: Description and Performance." ATL-LARG-PUB-2008-002.

[78] ATLAS Collaboration, "Local Hadronic Calibration." ATL-LARG-PUB-2009-001.

[79] M. Cacciari, G. P. Salam and G. Soyez, "The anti-kt jet clustering algorithm," *JHEP04*, vol. 063, 2008. arXiv:0802.1189 [hep-ph].

[80] G. Corcella et al. arXiv:hep-ph/0210213.

[81] B.P.Roe, et al., "Boosted decision trees as an alternative to artificial neural networks for particle identification," *NIM*, vol. A543, pp. 577–584, 2005. physics/0408124.

[82] R. Schilling, R. Song, Z, Vondracek, "Bernstein functions. Theory and Applications," *de Gruyter Studies in Mathematics*, vol. 37, 2010.

[83] ARGUS Collaboration, "Measurement of the polarization in the decay $B \to J/\psi K$," *Phys. Lett. B*, vol. 340, 3, 8, pp. 217–220, 1994.

[84] ATLAS Collaboration, "Measurement of the inclusive $W^{\pm}$ and Z/γ^{*} cross sections in the e and μ decay channels in pp collisions at $\sqrt{s} = 7\,\text{TeV}$ with the ATLAS detector," *Phys. Rev. D*, vol. 85, 072004, 2012.

[85] ATLAS Collaboration, "Performance of the ATLAS muon trigger in pp collisions at $\sqrt{s}$=8 TeV," *Eur. Phys. J. C*, vol. 75:120, 2015.

[86] ATLAS Collaboration, "Determination of the jet energy scale and resolution at ATLAS using Z/γ*–jet* events in data at $\sqrt{s} = 8\,\text{TeV}$," *ATLAS-CONF*, vol. 57, 2015.

[87] A L Read, "Presentation of search results: the CLs technique," *J.Phys. G*, vol. 28, pp. 2693–2704, 2002.

[88] ATLAS Collaboration, "Probing lepton flavour violation via neutrinoless $\tau \to 3\mu$ decays with the ATLAS detector," *Eur. Phys. J. C*, vol. 76:232, 2016.

[89] R. Anders, et al., "EvtGen: A Monte Carlo Generator for B-Physics," *EVTGEN-V00-11-07*, 2005. http://evtgen.warwick.ac.uk/static/docs/EvtGenGuide.pdf.

[90] "EvtGen Project." http://evtgen.warwick.ac.uk/.

[91] https://nnpdf.hepforge.org/html/PDFsets.html.

[92] The ATLAS collaboration, "Measurement of W and Z Boson Production Cross Sections in pp Collisions at $\sqrt{s} = 13\,\text{TeV}$ in the ATLAS Detector," *ATLAS-CONF-2015-039*, 2015.

[93] "Particle Data Group." http://www-pdg.lbl.gov/.

[94] A. Falkowski, J. T. Ruderman, T. Volansky, et al., "Hidden Higgs decaying to lepton jets," *JHEP*, vol. 77, 2010. arXiv:1002.2952 [hep-ph].

[95] CMS Collaboration & LHCb Collaboration, "Observation of the rare $B^0_s \to \mu^+\mu^-$ decay from the combined analysis of CMS and LHCb data," *Nature*, vol. 522, pp. 68–72,, 2015.

[96] S. V. Demidov and D. S.Gorbunov, "Flavor violating processes with sgoldstino pair production," *Phys. Rev. D*, vol. 85, 077701, 2012.

Appendix A

Trigger Efficiency for Collimated Three Muons

Although trigger efficiency measurement for the collimated multi-muons are described in Section 7.2, the calculation for the three muons are omitted due to the complicated formation. This is designed as an extension of that of two muon case (Equation 7.5) using properties of three muons and ΔR between them. But the terms of correlation need to be considered for the three pairs, as well as the overlap of the each term. First, details of design philosophy of the two muon case are described, including the case with a single muon trigger requirement. The extension to three muon case is based on the same philosophy.

A.1 Efficiency calculation for two reconstructed muons

Using two reconstructed two muons for the analysis procedure, the data which is acquired by a single-muon trigger and di-muon trigger can be used. The most simple case is the calculation of the di-muon trigger efficiency shown as Equation 7.5. This equation consists of the efficiency for each muon and correlation factor (CF) between these triggered muons. In other words, these factors are understood as the probability of triggered and separated.

The single-muon trigger efficiency for an event which consists of two reconstructed muons are also described as:

$$\begin{aligned}\epsilon_{\texttt{muX},\mu1,\mu2} =& \epsilon_{\texttt{muX},\mu1} \times \epsilon_{\texttt{muX},\mu2} \\ &+ \epsilon_{\texttt{muX},\mu1} \times (1-\epsilon_{\texttt{muX},\mu2}) \times CF(\Delta R_{\mu1,\mu2}) \\ &+ \epsilon_{\texttt{muX},\mu2} \times (1-\epsilon_{\texttt{muX},\mu1}) \times CF(\Delta R_{\mu1,\mu2}).\end{aligned} \tag{A.1}$$

In the first term of the equation, CF does not a matter because both muons are assumed to pass the trigger requirement. On the other hand, the second (and third) term considers the $\mu2$ ($mu1$) does not pass the trigger requirement.

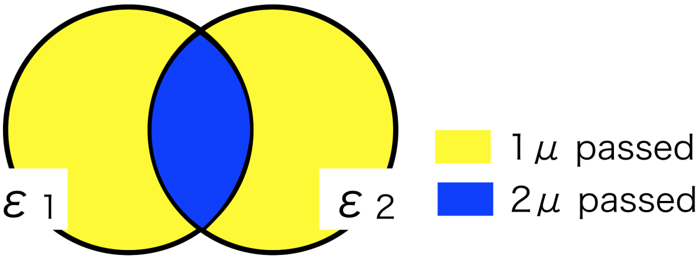


Figure A.1: Venn diagram of the efficiency for two muon events. The blue and yellow regions should be considered separately in the calculation because the correlation effects are different.

Therefore, these muons are necessary to be recognized independently, i.e. these terms should be factored by CF. Figure A.1 describes this equation as a Venn diagram. This is the essence of the multi-muon trigger efficiency calculation considering the correlation between muons.

A.2 Efficiency calculation for three reconstructed muons

Venn diagram for the three-muon case is described in Figure A.2. For the event with three muons, three-muon trigger is also used for the data acquisition. This is the most simplest because all muons are required to be triggered and separated. $CF(\Delta R_{\mu\mathrm{i},\mu\mathrm{j}})$ is simplified as CF_{ij} because each term is long. Therefore, the efficiency is calculated as:

$$\epsilon_{\texttt{3muX},\mu1,\mu2,\mu3} = \epsilon_{\texttt{muX},\mu1}\epsilon_{\texttt{muX},\mu2}\epsilon_{\texttt{muX},\mu2} \times CF^2_{12}CF^2_{13}CF^2_{23}, \tag{A.2}$$

where the CF for each pair is multiplied twice because each muon should be separated from all the other muons, i.e. it is obtained by the tranceformation from:

$$\epsilon_{\texttt{3muX},\mu1,\mu2,\mu3} = \prod_{i=1}^{3} \epsilon_{\texttt{muX},\mu i} \times CF_{ij}CF_{ik}, \tag{A.3}$$

where i, j, k are assumed to be 1, 2, 3, and $i \neq j \neq k$.

The di-muon trigger efficiency is also calculated by the similar equation. In this case, at least two in the three muons are required to be triggered. Therefore, number of the combination patterns increased from three-muon case. The

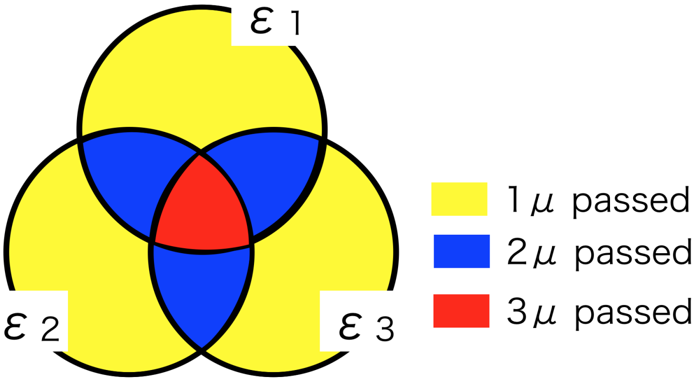


Figure A.2: Venn diagram of the efficiency for three muon events. This is the simple extension of Figure A.1 with an additional muon. The red, blue, and yellow regions should be considered separately in the calculation because the correlation effects are different.

equation is defined as:

$$\begin{aligned}\epsilon_{\mathtt{2muX},\mu1,\mu2,\mu3} =& \epsilon_{\mathtt{muX},\mu1}\epsilon_{\mathtt{muX},\mu2}\epsilon_{\mathtt{muX},\mu3} \times (1-(1-CF_{12})(1-CF_{13})(1-CF_{23})) \\ &+ \epsilon_{\mathtt{muX},\mu1}\epsilon_{\mathtt{muX},\mu2}(1-\epsilon_{\mathtt{muX},\mu3}) \times CF_{12}^2 CF_{13} CF_{23} \\ &+ \epsilon_{\mathtt{muX},\mu1}\epsilon_{\mathtt{muX},\mu3}(1-\epsilon_{\mathtt{muX},\mu2}) \times CF_{13}^2 CF_{12} CF_{23} \\ &+ \epsilon_{\mathtt{muX},\mu2}\epsilon_{\mathtt{muX},\mu3}(1-\epsilon_{\mathtt{muX},\mu1}) \times CF_{23}^2 CF_{12} CF_{13}, \end{aligned} \tag{A.4}$$

where the first term represents of the case that all muons pass the single-trigger requirement, while the other terms are for the case that one of the muon does not pass. In the former case, at least one of the pairs has to be separated. On the other hand, the trigger-failed muon does not have to separated from the other (trigger-passed) muons in the latter case. Therefore, only the CF for the trigger-passed muons is multiplied twice in each term.

Efficiency calculation of a single-muon trigger for the three muons is the most complicated case because three cases should be considered: one of the three muons is triggered; two of these are triggered; and all of these muons are triggered by the requirement. The equation of this case can be defined as:

$$\begin{aligned}\epsilon_{\mathtt{muX},\mu1,\mu2,\mu3} =& \epsilon_{\mathtt{muX},\mu1}\epsilon_{\mathtt{muX},\mu2}\epsilon_{\mathtt{muX},\mu3} \\ &+ \epsilon_{\mathtt{muX},\mu1}\epsilon_{\mathtt{muX},\mu2}(1-\epsilon_{\mathtt{muX},\mu3}) \times (1-(1-CF_{13})(1-CF_{23})) \\ &+ \epsilon_{\mathtt{muX},\mu1}\epsilon_{\mathtt{muX},\mu3}(1-\epsilon_{\mathtt{muX},\mu2}) \times (1-(1-CF_{12})(1-CF_{23})) \\ &+ \epsilon_{\mathtt{muX},\mu2}\epsilon_{\mathtt{muX},\mu3}(1-\epsilon_{\mathtt{muX},\mu1}) \times (1-(1-CF_{12})(1-CF_{13})) \\ &+ \epsilon_{\mathtt{muX},\mu1}(1-\epsilon_{\mathtt{muX},\mu2})(1-\epsilon_{\mathtt{muX},\mu3}) \times CF_{12}CF_{13}CF_{23} \\ &+ \epsilon_{\mathtt{muX},\mu2}(1-\epsilon_{\mathtt{muX},\mu1})(1-\epsilon_{\mathtt{muX},\mu3}) \times CF_{12}CF_{13}CF_{23} \\ &+ \epsilon_{\mathtt{muX},\mu3}(1-\epsilon_{\mathtt{muX},\mu1})(1-\epsilon_{\mathtt{muX},\mu2}) \times CF_{12}CF_{13}CF_{23}. \end{aligned} \tag{A.5}$$

If the all muons are passed the single muon requirement, CF does not matter. However, if one of the muons is failed to pass the single trigger requirement, at least one of the other muons has to be separated from the trigger-failed muon. In addition, only one of the muons passes the single-muon trigger requirement, all muons have to be separated.

Although the trigger efficiency can be extended to the three muon case for the normal chains, the Full-Scan (FS) triggers also should be considered. As explained in Section 2.2.4, this trigger chain requires an additional muon (or muons) is required at the last step of the HLT. Therefore, the efficiency can be assumed as single-muon trigger which is multiplied by the additional efficiency of the FS muon requirement. The efficiency of the trigger chain composed by a single normal muon ($> X$ GeV) and an FS muon ($> Y$ GeV) requirements is defined as:

$$\begin{aligned}\epsilon_{\mathtt{muX\_muY\_FS},\mu1,\mu2,\mu3} =& \epsilon_{\mathtt{muX\_muY\_FS},\mu1}\epsilon_{\mathtt{muX\_muY\_FS},\mu2}\epsilon_{\mathtt{muX\_muY\_FS},\mu3} \\ &+ \epsilon_{\mathtt{muX\_muY\_FS},\mu1}\epsilon_{\mathtt{muX\_muY\_FS},\mu2}(1-\epsilon_{\mathtt{muX\_muY\_FS},\mu3}) \times (1-(1-CF_{13})(1-CF_{23})) \\ &+ \epsilon_{\mathtt{muX\_muY\_FS},\mu1}\epsilon_{\mathtt{muX\_muY\_FS},\mu3}(1-\epsilon_{\mathtt{muX\_muY\_FS},\mu2}) \times (1-(1-CF_{12})(1-CF_{23})) \\ &+ \epsilon_{\mathtt{muX\_muY\_FS},\mu2}\epsilon_{\mathtt{muX\_muY\_FS},\mu3}(1-\epsilon_{\mathtt{muX\_muY\_FS},\mu1}) \times (1-(1-CF_{12})(1-CF_{13})) \\ &+ \epsilon_{\mathtt{muX\_muY\_FS},\mu1}(1-\epsilon_{\mathtt{muX\_muY\_FS},\mu2})(1-\epsilon_{\mathtt{muX\_muY\_FS},\mu3}) \times CF_{12}CF_{13}CF_{23} \\ &+ \epsilon_{\mathtt{muX\_muY\_FS},\mu2}(1-\epsilon_{\mathtt{muX\_muY\_FS},\mu1})(1-\epsilon_{\mathtt{muX\_muY\_FS},\mu3}) \times CF_{12}CF_{13}CF_{23} \\ &+ \epsilon_{\mathtt{muX\_muY\_FS},\mu3}(1-\epsilon_{\mathtt{muX\_muY\_FS},\mu1})(1-\epsilon_{\mathtt{muX\_muY\_FS},\mu2}) \times CF_{12}CF_{13}CF_{23},\end{aligned} \tag{A.6}$$

where $\epsilon_{\mathtt{muX\_muY\_FS},\mu i}$ represents the single-muon trigger with the FS requirement as described as:

$$\epsilon_{\mathtt{muX\_muY\_FS},\mu i} = \epsilon_{\mathtt{muX},\mu i}(1-(1-\epsilon_{\mathtt{muY\_FS},\mu j} \times CF'_{ij})(1-\epsilon_{\mathtt{muY\_FS},\mu k} \times CF'_{ik})), \tag{A.7}$$

where $\epsilon_{\mathtt{muY\_FS},\mu i}$ shows the efficiency of the FS requirement, and CF'_{ij} shows the correlation factor with a FS object. This equation form is similar to the single-muon trigger case, but $\epsilon_{\mathtt{muX},\mu i}$ is replaced by $\epsilon_{\mathtt{muX\_muY\_FS},\mu i}$. The efficiency for the two muon requirement at FS level, can be estimated by the same method while $\epsilon_{\mathtt{muX\_muY\_FS},\mu i}$ is replaced by $\epsilon_{\mathtt{muX\_2muY\_FS},\mu i}$ which is defined as:

$$\epsilon_{\mathtt{muX\_2muY\_FS},\mu i} = \epsilon_{\mathtt{muX},\mu i}(\epsilon_{\mathtt{muY\_FS},\mu j}\epsilon_{\mathtt{muY\_FS},\mu k} \times CF'_{ij}CF'_{ik}). \tag{A.8}$$